INTERNATIONAL CONFERENCE

on

Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)

23rd-24th February 2024

Jointly Organized By:

Career Point University, Kota, India and
Statistical & Informatics Consultation Center (SICC),
Faculty of Computer Science and Maths,
University of Kufa, Iraq

CAREER POINT
UNIVERSITY

SICC

IQAC
CPU, KOTA

INSTITUTION'S
INNOVATION
COUNCIL
(Ministry Education Initiative)

Conference Proceedings

About the Conference

International Conference on **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"** is organizing with the objective of bringing together innovative scientists, professors, research scholars, students and industrial experts in the field of Computing and Communication to a common forum. The primary goal of the conference is to promote the exchange of innovative scientific information between researchers, developers, engineers, students, and practitioners. Another goal is to promote the transformation of fundamental research into institutional and industrialized research and to convert applied exploration into real time application.

Overall the conference will provide the researchers and attendees with prospects for national and international collaboration and networking among universities and institutions from India and abroad for promoting research. RIETSSL-2024 is jointly organizing by **Career Point University, Kota, Rajasthan, India and Statistical and Informatics Consultation Center (SICC), Faculty of Computer Science and Maths, University of Kufa, Iraq** on 23rd-24th February 2024. Abstracts of all the accepted papers will be published in special edition of **Career Point International Journal of Research (CPIJR)**.

KEYNOTE SPEAKERS

Prof. Sheng-Lung,
National Taipei University of
Business, Taiwan

Prof. Chuan-Ming Liu
National Taipei University of
Technology,Taipei,Taiwan

Dr. Ahmed J. Obaid.
SICC, University of Kufa,Iraq

Dr. Neyara Radwan
King Abdulaziz University, Saudi
Arabia

**Dr. Mohammad Kamrul
Hasan,**
Universiti Kebangsaa, Malaysia
(UKM)

Dr. Md. Nur Alam
Pabna University of Science &
Technology Pabna, Bangladesh

Prof. (Dr.) Vijay Singh Rathore,
JECRC College, Jaipur (Raj.)

Prof. (Dr.) Madhukant Patel,
Former Scientist(ISRO), MD,
Reve Automation LLP, Gujarat

Dr. S.D. Purohit,
Rajasthan Technical University,
Kota (Raj.)

Dr. O.P. Rishi,

Department of Computer

Science & Informatics,

Univeristy of Kota, Kota (Raj.)

SESSION CHAIRS

Prof. N.N. Jani,

SKPIMCS-KSV,Gandhinagar

Dr. Bharat Singh Deora

Department of CS & IT,

RajsathanVidyapeeth,Udaipur(Raj.)

Dr Ihtiram Raza Khan

Jamia Hamdard University, Delhi

Dr. Anubha Jain

IIS University, Jaipur (Raj.)

Dr. Niranjan Murthy M

BMSITM, Bangalore,

Karnataka

Dr. Salini Suresh

DSCASC College ,Bangalore

Dr. Navneet Sharma

Department of CS & IT,IIS

University, Jaipur (Raj.)

Prof. (Dr.) N.K. Joshi

MIMT College, Kota (Raj.)

Prof.(Dr.) Minakshi Memoria

Department of CSE,

UttaranchalUniversity, Dehradun,

Uttarakhand

Prof. (Dr.) Surendra Yadav
Department of CS & IT,
VGU, Jaipur (Raj.)

Dr. Arvind K Sharma
Department of CS & IT,
University of Kota, Kota
(Raj.)

Prof. (Dr.) A.V. Trivedi
Principal, Aryan College,
Kota (Raj.)

CONFERENCE REVIEWER BOARD

S. No.	Name of Reviewer	Designation	Affiliation
1	Dr. Akash Saxena	Professor	Compucom College, Jaipur (Rajasthan)
2	Dr. Niranjanmurthy M.	Associate Professor	MSRIT, Bengaluru, Karnataka, India
3	Dr. Suresh Kumar K	Associate Professor	Saveetha Engineering College, Chennai
4	Dr. Veena Khandelwal	Assistant Professor	SRM Institute of Science and Technology Delhi-NCR Campus, Ghaziabad
5	Dr. Astha Pareek	Sr. Assistant Professor	Dept. Of Computer Science And IT, The IIS University, Jaipur
6	Prof. Tarun Shrimali	Director (R & D Cell)	JRN Rajasthan Vidyapeeth Udaipur, Rajasthan
7	Dr. Kiran Sood	Professor	Chitkara University, Rajpura, Punjab
8	Dr. Soniya Lalwani	Associate Professor	BKIT, Kota, Rajasthan
9	Dr. Ashish Khanna	Professor	MAIT, Rohini, Delhi
10	Dr. Tanupriya Chaudhary	Associate Professor,	Department of CSE, UPES, Dehradun
11	Dr. Navneet Sharma	Associate Professor	Dept. Of Computer Science And IT, The IIS University, Jaipur
12	Dr. Soumi Dataa	Associate Professor	IEM, Kolkata(W.B.)
13	Dr. Sandeep Poonia	Associate Professor	CHRIST University, Bangalore
14	Dr.C.Priya	Associate Professor	VELS University, Chennai
15	Dr.Akash Saxena	Associate Professor	Compucom Institutes of Information Technology,Jaipur
16	Dr.Nilam Chaudhary	Associate Professor	SKIT, Jaipur, Rajasthan
17	Dr. Yogesh Kumar Sharma	Dean and Professor	School of Engineering & Technology, Om Sterling Global University, Hisar, Hariyana

18	Dr. Priyanka Verma	Sr. Assistant Professor	Dept. of Computer Science & IT, IIS(Deemed to be)University, Jaipur, Rajasthan
19	Dr.Anubhav Kumar	Professor & Dean	Faculty of Engineering & Technology, Mandsaur University, Mandsaur (M.P.)
20	Dr. Anubha Jain	Associate Professor	Dept. of Computer Science & IT, IIS(Deemed to be)University, Jaipur, Rjatshan
21	Dr. Bharat Singh Derora	Sr. Assistant Professor	Dept. of Computer Science & IT, IIS(Deemed to be)University, Jaipur, Rjatshan
22	Prof. (Dr.) Manju Mandot	Director	Dept. of Computer Science & IT, IIS(Deemed to be)University, Jaipur, Rjatshan
23	Prof. (Dr.) N.K. Joshi	Director	MIMT College, Kota(Raj.)
24	Dr. Vinod Kumar	Associate Professor (CSE)	Galgotias University, Greater Noida, U.P.
25	Dr. L.K. Vishwamitra	Professor(CSE)	Oriental College of Technology, Bhopal
26	Dr. Iti Sharma	Lecturer(CS)	Govt. Polytechnic College, Kota
27	Dr. Gajanand Sharma	Asst. Prof.(CSE)	JECRC University, Jaipur(Raj.)
28	Dr. Anuradha Sharma	Associate Professor	VMOU, Kota (Raj.)
29	Dr. Nirupma Singh	Team Lead, Product Implementation	Intellect Design Arena Ltd. Chennai
30	Prof. (Dr.) Manish Shrimali	Associate Professor,CS & IT	JRN Rajasthan Vidyapeeth, Udaipur (Raj.)

ORGANIZING COMMITEE

Patron: RIETSSL-2024

Er. Pramod Maheshwari

Chief Guest: RIETSSL-2024

Prof. (Dr.) S.K. Singh

Conference Convener: RIETSSL-2024

Dr. Abid Hussain

Honorary Conference Chair

Dr. Ahmed J Obaid

Organizing Secretary-RIETSSL-2024

Dr. Garima Tyagi

General Chair: RIETSSL-2024

Dr. Amit Sharma

Registrar

Mr. Kamal Arora

Scientific Committee

Prof. (Dr.) Hemlata Saxena

Dr. Vivek Kr. Jain

Prof. (Dr.) M.K. Gupta

Dr. Preeti Sharma

Dr. Surabhi Singh

Dr. Arun Sharma

Dr. Manish Tiwari

Dr. Krishna Kr. Soni

Dr. Aditi Gaur

Dr. Erum Gul Naaz

Dr. Ashutosh Kumar

Dr. Anita Singh

Dr. Shailesh Sharma

Dr. Amit Kumar

Dr. Manoj Jangid

Dr. Vanita

Dr. Rajkumari Thagele

Dr. Meenakshi

Advisory Board

PATRON

Er. Pramod Maheshwari,
Chancellor, Career Point University, Kota, Rajasthan, India

It's a moment of great pleasure that Career Point University, Kota is Organizing International Conference On **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"** under the flagship **of Career Point University, Kota in association with SICC, University of Kufa, Iraq**. Organizing an event does not come without effort. It requires vision, mission and hard work. Conferences of such nature provide a great opportunity to Engineering fraternity, not only to update knowledge and keep abreast of the latest developments in the respective field, but also an occasion for the resource persons, delegates to exchange ideas and interact with each other.

I take this opportunity to congratulate the organizing committee and to extend warm welcome to the resource persons and delegates. I thank all the national and international delegates who have come from various parts of the country and across the globe. We consider it a privilege and honour to have all of you here.

I wish you all for the grand success of this wonderful event.

CHIEF GUEST

Prof. (Dr.) S.K. Singh (Vice Chancellor),
Rajasthan Technical University, Kota, Rajasthan, India

It is a matter of great pleasure to host the International Conference on **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"** on 23rd - 24th February 2024. This conference aims to develop insights into the international scenario of engineering and IT industrial and academic research by offering a common platform to scientists, researchers and students.

The conference will stimulate the scientific temper among students, teachers and industrial leaders for building a bridge between academia and industry. Industrialists across the region will participate as invited speakers to address the current need in the field of Engineering and Technology.

Conferences are meant essentially for scientific exchange and generation of new ideas in the chosen field along with personal interaction and networking. I understand that a number of national and international speakers are participating to speak on a variety of topics thus enriching the knowledge of all participants.

I wish the conference all the success and my heartiest congratulations to the organizing committee.

CONFERENCE CONVENER

Dr. Abid Hussain
Associate Professor, School of Computer Application & Technology,
Career Point University, Kota, Rajasthan, India

I am gratified being designated as the convener/conference chair for International Conference on **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"** to be held at Career Point University,Kota from 23rd - 24th February 2024. The conference is being jointly organizing by Career Point University, Kota, Rajasthan, India and Statistical and Informatics Consultation Center (SICC), Faculty of Computer Science and Maths, University of Kufa, Iraq. The aim of the conference is to bring together researchers, scientists, engineers, and practitioners to exchange and share their experiences, new ideas and research results about all aspects of main themes and tracks.

The conference has solicited and gathered technical research submission related to all aspects of major conference themes and tracks. After the rigorous peer-review process, the submitted papers were selected on the basis of originality, significance, and clarity for the purpose of the conference.

I am grateful to all those who have contributed to the success of RIETSL-21, especially all the authors and the participants who responded to our call for papers. I also express my sincere gratitude for the efforts put by conference Technical Programme Committee, Conference Steering Committee, Advertising, Publicity and Sponsorship Committee, who made this event possible.

HONORARY CONFERENCE CHAIR

Dr. Ahmed J Obaid
Associate Professor, Faculty of Computer Science and Mathematics,

University of Kufa, Iraq

On behalf of the Organizing Committee of this International Conference on **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)",** we would like to extend our warm welcome to all of the presenter and participants, and in particular, we would like to express our sincere gratitude to our plenary and invited speakers. This international conference is jointly organizing by the Career Point University, Kota,, India Statistical and Informatics Consultation Center (SICC), Faculty of Computer Science and Maths, University of Kufa, Iraq.

It is intended to be the first step towards a top class conference on Engineering & Technology. We believe that this international conference will give opportunities for sharing and exchanging original research ideas and opinions, gaining inspiration for future research, and broadening knowledge about various fields in advanced computer science and information systems, amongst members of organization communities, together with researchers from India and other countries.

ORGANIZING SECRETARY

Dr. Garima Tyagi
HoD & Associate Professor, School of Computer Application & Technology,
Career Point University, Kota, Rajasthan, India

We are pleased to organize two days international conference on **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"** on 23rd - 24th February 2024 in the collaboration of SICC "Statistical and Informatics Consultation Center (SICC), Faculty of Computer Science and Maths, University of Kufa, Iraq".

As the theme of the conference is about recent innovation in Engineering and Technology, its subthemes cover all the branches of Engineering including Mathematics.

The objective of the conference is to bring together the computational experts in various fields of Engineering and Technology. The conference showcases the expertise both from academia and industry from various institutes, universities and industries all over the world, discover novel opportunities on the theme of the conference.

I on the behalf of the organizing committee and on my personal behalf, delighted to welcome all the delegates and the participants in the conference.

I am sure that this conference will bring together the professionals and researchers for the innovation and that will help in the sustainable development of the society. I will help in the overall development of the new technologies with the multidisciplinary approach.

GENERAL CHAIR

Dr. Amit Sharma
Associate Professor, School of Computer Application & Technology,
Career Point University, Kota, Rajasthan, India

I would like to extend my warmest welcome to all the participants of the International Conference on **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"**. This inernational conference organized by Career Point University, Kota (Raj.) in the collaboration of SICC **"Statistical and Informatics Consultation Center (SICC), Faculty of Computer Science and Maths, University of Kufa, Iraq"** on 23rd - 24th February 2024. I am confident that the conference shall provide an effective platform for innovation, technology transfer and entrepreneurship concurrently meet to share and disseminate the knowledge and the rich experience of the IT and Engineering Professionals, and to look forward solutions to the challenging problems.

I also thank all the conference committee members and the reviewers for their efforts in ensuring a rigorous review process to select high quality papers.

I sincerely hope that this unique international and multidisciplinary conference will provide the participants with a truly transformative experience through a variety of knowledge and perspectives. Finally, showing gratitude to all the sponsors for their generous support, I wish a grand success of this conference.

REGISTRAR

Mr. Kamal Arora
Registrar, Career Point University, Kota, Rajasthan, India

On behalf of the Career Point Univeristy, Kota, I heartily extend warm welcome to all the National/International Delegates, Renowned Scientists and participants to this International Conference with the theme of **"Recent Innovations in Engineering, Technology & Science for Sustainable Living (RIETSSL-2024)"** on 23rd - 24th February 2024. The presence of other dignitaries during the two days conference is a further testimony to our sincere pursuits to achieve nothing less than the 'best', they have long trails of success behind them.

I am confident that the conference shall provide an effective platform for innovation, technology transfer and entrepreneurship concurrently meet to share and disseminate the knowledge and the rich experience of the IT and Engineering Professionals, and to look forward solutions to the challenging problems.

I intend to take this event ahead as an ideal, the motive not only is to generate discussions around contemporary issues, but also to propel the culture of academic exchange, which is the only way to achieve excellence in this field.

I hope this event fruitful for everybody. I assure you that we will make your time spent with us and in the conference a memorable one.

TABLE OF CONTENTS

12	Vinodkumar Babanrao Jagdale	High-Frequency Stock Market Prediction: A Comprehensive Approach with Sentiment Analysis, OI Data and Volatility using AIML & Data Analytics	183-194
13	Ramphool Saini	Flora of Gharbanga wildlife sanctuary	195-206
14	Saleh Ibrahim Jauro,Dr. Abid Hussain	Implications of Data Privacy in Cloud Computing using Trust Mechanisms	207-216
15	Madhvi Chandra, Dr. Sukirti Priyadarshni	Analytical study of AI for Sustainable Human Development	217-226
16	Nidhi Sharma	Floristic studies pertaining to the biodiversity of the Mukundara Hills Tiger Reserve (MHTR) Kota of Rajasthan state	227-242
17	Malathi Takkedu, Dr. K. Ram Mohan Rao,	Unlocking Insights: Exploring Machine Learning Techniques for Predicting Academic Performance in Education	243-255
18	Keshav Sharma, Dr.Manish Tiwari	Plagiarism Detection Using NLP	256-268
19	Radhika Tyagi, Dr. Meenakshi, Dr. Sukhvinder Singh	Global Research Trends on the use of Nanotechnology to Boost Medicine Production: A Review	269-289
20	Sasidhar Kothuru, Dr. Rajesh Kulkarni	Consumer Behavior Analytics using Fuzzy Rule based Classification through Mapping and Bagging Approaches in Map Reduce environment	290-299
21	Rabia Koser, Dr. Amit Sharma	An Unified Model for Energy-Efficient Task Scheduling in Green Cloud Computing	300-309
22	B. Manimekala, Pandi Prabha. S, Salini Suresh, T. KohilaKanagalakshmi	Congestion Free Shortest Path Between Source and Destination for Real-Time Traffic on Roads using ACO Algorithm in VANET	310-333
23	Rajaram Das, Dr.R.Suthakaran	Anti-oxidant screening and in-vitro anti-cancer efficacy of solvent extracts of Ceiba speciosa	334-349
24	Balasani Raghupathi, Dr. Amit Sharma	Smart Agriculture: Designing a Machine Learning-Based Chilli Crop Disease Detection System	350-367

X

Digital Transitions in Libraries: Expanding Information Reach beyond Card Catalogs

Somvir[1], Dr. Preeti Sharma[2]

[1]Research Scholar, Library & Information Science, Career Point University, Kota, India

[2]Research Supervisor, Library & Information Science, Career Point University, Kota, India

[1]somvirrathee9@gmail.com, [2]preetsharma1977@gmail.com

Abstract

The research paper delves into the transformative impact of the internet on libraries and their transition from physical to digital information access. It explores the historical context of this evolution and underscores the pivotal role of library websites in providing efficient and user-friendly access to a wide range of digital resources, from books to scholarly articles and multimedia content.

This paper investigates the redefined features of library websites, emphasizing user-centered design, enhanced search capabilities, and interactive elements that have significantly improved the overall experience for library patrons. It also highlights the implications of this shift from physical to virtual libraries, particularly in terms of accessibility for various user groups, including students, researchers, and the general public. The research acknowledges the challenges and opportunities that arise from this digital transformation, such as the importance of digital literacy, privacy concerns, and security.

In conclusion, "From Card Catalogs to Clicks" underscores the profound significance of library websites as gateways to the world of knowledge. Through an examination of their evolution and their impact on information access, this paper sheds light on how modern libraries are adapting to meet the evolving needs of their patrons in our increasingly digital society.

Keywords: Library Website, Future Trends, Challenges and Opportunity of Library Website, Influence on Information

1. Introduction

The digital age has brought about a significant transformation in the way individuals interact with and retrieve knowledge. Libraries, which were once repositories of printed materials, have undergone a metamorphosis into multifaceted institutions with an expanded online presence. The evolution of information access and the emergence of technology have played a pivotal role in this transformation. Central to this transformation is the profound impact of library websites, which have emerged as pivotal tools in the quest for information. Libraries have traditionally been associated with card catalogs, physical stacks, and the hushed ambiance of study. Patrons would navigate extensive collections of books, journals, and other resources using card catalogs and librarian assistance, often limited by the constraints of physical space and time. However, the advent of the internet and the subsequent proliferation of digital resources have ushered libraries into a new era. Libraries have undergone a significant metamorphosis from their traditional role as repositories of printed materials to multifaceted institutions with an expanded online presence.

1.1 The Significant of Library Websites in Digital Revolution

The digital revolution has reshaped the functions and duties of libraries in numerous facets. Library websites now serve as vital gateways, allowing users to delve into an extensive array of digital content effortlessly. With a mere click or tap, individuals can seamlessly explore a constantly growing assortment of online databases, e-books, multimedia resources, and scholarly articles. This evolution from traditional card catalogs to digital interfaces has democratized information access and fundamentally shifted the standards and demands of library users. Library websites are playing a pivotal role in the ongoing transformation of libraries. They are no longer just digital repositories, but interactive platforms that offer a seamless and engaging experience to users. Library websites serve as gateways to knowledge and indispensable tools for information seekers. They are designed to enhance user experience and streamline the retrieval of information. The significance of library websites in this context is multifaceted, encompassing user-centered design, search functionality, interactivity, accessibility, and the very notion of what it means to be a library in the 21st century (Anna, 2018, Ex Libris, 2022)[1],[2].

2. Objective And Structure

The objective of this research paper is to explore the dynamic evolution of library websites, examine their functions and features, and evaluate their influence on information access. By delving into the historical context, user experiences, challenges, and opportunities presented by library websites, we aim to shed light on the pivotal role they play in the contemporary landscape of knowledge dissemination. The paper is structured to first review the historical development of library catalogs and the emergence of library websites. It subsequently investigates the functions and features of modern library websites, their influence on information access, and the challenges and opportunities they present. Real-world case studies and a comprehensive analysis will further provide insights into the critical role played by library websites in the digital age.

3. Literature Review

Verma and Devi (2016) propose dedicated library webpages for all IIMs in India, providing comprehensive information and accessible links[3]. Yoon & Schultz (2017) advocate tailoring library data management webpages to suit each institution's unique characteristics and research data management services. They stress the need for enhanced service offerings and more comprehensive information while adapting to institutional specifics[4]. Al-Qallaf & Ridha's (2019) study on GCC academic library websites calls for improvements in design, content, and web-based services, recommending usability tests and exploring accessibility[5]. Bharati and Madhusudhan (2019) analysis of libraries at Jawaharlal Nehru University and Banaras Hindu University revealed attention-grabbing features at JNU, such as scrolling notices and first-leaf news tools, absent in BHU's website[6]. Hugar's (2019) exploration of engineering college library websites underscores the significance of user-friendly, dynamic sites, highlighting continuous improvement and user feedback [7]. Chikkamanju and Kumar (2020) highlighted the underutilization of Web 2.0 features in 71 Agricultural University Library Websites in India, with Facebook, Twitter, and YouTube being more commonly used than Wiki[8].

Pal, Sarkar, and Kar (2020) evaluated Asian national library websites, while Rahman and Sadik (2020) examined the usability, accuracy, and currency of content on University of Delhi-affiliated college library websites, identifying significant room for improvement in social features, updates, and comprehensive user services. Regular updates and user-

centric enhancements emerged as essential recommendations across these studies[9], [10]. Ambika and Ganesan's (2021) analysis of central university library websites emphasizes user-friendliness and the role in promoting services[11]. Barman (2021) conducted a study investigating the web content of agricultural university libraries in India, revealing their effective information sharing via websites. While no library websites utilized Web 2.0 tools, the incorporation of user-friendly Web 2.0 services was noteworthy. Remote access was offered by only 50% of the libraries, indicating the pressing need for updates across institutions[12]. Li et al. (2021) emphasize content enrichment and SEO optimization for smart library portal websites, suggesting collaborations with third-party companies to address technical limitations[13]. Singha and Verma's (2021) content analysis of eight Indian veterinary university library websites using a 65-parameter checklist highlighted Rajasthan University of Veterinary & Animal Sciences, Bikaner, as the top-ranked site globally and in India based on Alexa traffic ranks. Sri Venkateswara Veterinary University, Tirupati, scored highest in user-friendliness, with 75% of the websites attaining commendable ratings[14]. Brahma & Verma (2022) analysis of national libraries in Asia showcases diverse website features, recommending detailed information, improved site currency, and the importance of global rankings for online presence[15]. Tavosi & Naghshineh (2023) discuss the relationship between SEO and accessibility, suggesting enhancements to user interface and experience[16].

4. Evolution Of Library Websites

Transitioning from physical catalogs to digital interfaces has been a pivotal shift for libraries. Traditionally reliant on card catalogs—drawers filled with index cards detailing library items—this system posed limitations. Users had to physically visit the library, and manual updates were challenging, making real-time maintenance difficult. The evolution began in the 1980s and 1990s with computers and the internet. The Online Public Access Catalog (OPAC) emerged as an early digital catalog, offering remote access to library holdings. Digital interfaces brought numerous advantages: varied search criteria, keyword searches, and freedom from spatial constraints. These digital interfaces evolved, integrating user-centric designs for intuitive browsing. User accounts enhanced interaction by enabling patrons to track borrowing history and place holds. Additionally, these interfaces expanded to incorporate electronic resources like e-books, journals, and databases, transforming libraries into comprehensive digital hubs.

As smartphones and tablets became ubiquitous, libraries embraced responsive web design, enabling access to their digital interfaces across diverse devices and screen sizes. Adherence to web accessibility standards like the Web Content Accessibility Guidelines (WCAG) ensured inclusive access for all users, including those with disabilities. Advanced features such as faceted search, citation management, and refined metadata were introduced to heighten search accuracy and relevance within library digital interfaces. Concurrently, libraries started leveraging data analytics to glean insights into user behaviors and preferences. This data-centric strategy empowered libraries to enhance their digital interfaces by aligning them with user usage patterns and evolving expectations.

A comprehensive digital platform should embody a plethora of essential features, including ease of use for seamless navigation, a robust search functionality ensuring swift and accurate results, and mobile accessibility for on-the-go convenience. Moreover, it should provide access to diverse digital resources, coupled with personalized user accounts, while offering virtual reference services and fostering community engagement through recommendations, user reviews, and social features. Prioritizing digital literacy support, inclusivity, and accessibility alongside stringent privacy and security measures is crucial. Keeping users informed with news updates, furnishing research guides and tutorials, and enabling multilingual support contribute significantly. Furthermore, embracing continuous improvement through user feedback mechanisms ensures the platform evolves in tandem with users' needs and preferences. Moving from physical card catalogs to digital interfaces has democratized access to library resources, granting users the ease of remotely searching and accessing materials. This shift has enhanced user experiences, offering more intuitive and interactive interfaces while placing a premium on accessibility and inclusivity. The ongoing transformation persists as libraries adjust to evolving technologies and user anticipations.

5. Capabilities And Characteristics Of Contemporary Library Websites

Today's library websites function as adaptable hubs for accessing information. They provide intuitive search tools equipped with advanced filters and facets, streamlining resource exploration. Hosting comprehensive online catalogs comprising books, journals,

e-books, and multimedia resources, these websites enable users to reserve and extend item borrowing.

- Search Functionality: Search engines with keyword, title, author, and subject search options. Advanced search filters and facets for refining search results. Boolean operators for precise searching.

- Online Catalog: A comprehensive database of library holdings, including books, journals, e-books, and other materials. Detailed item information, including availability, location, and call numbers. Reserving or renewing items online.

- User Accounts: Personalized user accounts for managing borrowed items, favorites, and search history. Notifications for due dates, holds, and fines. User preferences and profiles.

- E-Resources Access: Integration with digital resources, such as e-books, e-journals, and databases. Authentication for off-campus access for students and faculty. Tools for seamless integration with learning management systems.

- Interactivity: Chat or messaging support for user inquiries. User reviews and ratings for library materials. Social media integration for sharing resources and updates.

- Mobile Access: Responsive design for accessibility on various devices (smartphones, tablets, desktops). Mobile apps for enhanced on-the-go access.

- Accessibility and Inclusivity: Compliance with web accessibility standards (WCAG) to ensure usability for all users. Multilingual support. Assistive technologies, such as screen readers.

- Content Management: Easy-to-use content management systems for adding, updating, and deleting resources. Metadata standards and tools for efficient cataloging.

- Virtual Reference Desk: Virtual reference services, including email, chat, or video conferencing with librarians. FAQs and knowledge bases for common user inquiries.

- User Education and Tutorials: - Online tutorials, guides, and instructional materials to enhance digital literacy. - Research guides tailored to specific subjects and courses.

- Alerts and Notifications: - Email or SMS alerts for new acquisitions, upcoming events, or overdue materials. - News and updates related to library services and resources.

- Integration with Other Services: - Integration with interlibrary loan services for resource sharing. - Linking with course reserves for academic institutions. - Collaboration with digital repositories and archives.

- User Feedback and Assessment: - Surveys and feedback mechanisms to gather user input for continuous improvement. - Usage statistics and analytics for assessing website effectiveness.

- Privacy and Security Measures: - Secure login and data encryption. - Privacy policies and user data protection.

- User Experience Design: - Intuitive and user-friendly design with clear navigation. - Responsive design for various screen sizes and devices.

- Resource Sharing and Community Building: - Social features for sharing resources, recommendations, and reviews. - Virtual book clubs or discussion forums.

6. Impact On Accessing Information

Library websites have revolutionized information access, democratizing knowledge by providing a vast digital repository. Improved search capabilities, personalization, virtual reference services, and remote access benefit users. They enhance accessibility, inclusivity, and user experiences, fostering digital literacy and ensuring information is readily available to diverse audiences.

- Democratization of Information: Modern library websites have democratized access to information by making digital resources available to a broader audience, transcending physical boundaries. Individuals from diverse backgrounds, including students, researchers, and the general public, can access a wealth of information without the constraints of geography or library hours.

- Accessibility and Inclusivity: Library websites play a crucial role in making information more accessible to individuals with disabilities by adhering to web accessibility standards (WCAG). Multilingual support and user-friendly interfaces

enhance inclusivity, welcoming users from various linguistic and cultural backgrounds.

- Enhanced Search Capabilities: Improved search functionality empowers users to find relevant information quickly and efficiently. Advanced search filters and facets allow users to narrow down their search results, saving time and frustration.

- Personalization and User Experience: Library websites offer personalized features like user accounts, favorites, and search history, tailoring the experience to individual preferences. User-centered design principles prioritize ease of use, ensuring a positive and efficient user experience.

- Digital Literacy and Skills Development: Library websites often provide resources and tutorials to enhance users' digital literacy and research skills. Users can develop the ability to critically evaluate and navigate digital information effectively.

- Virtual Reference and Assistance: Virtual reference services available through library websites, such as chat or email support, offer timely assistance and guidance to users. This support helps users navigate complex information landscapes.

- Integration with E-Resources: Library websites seamlessly integrate digital resources, including e-books, e-journals, and databases, making these materials readily accessible to users. Off-campus access for students and faculty fosters uninterrupted research.

- Information Alerts and Notifications: Email or SMS alerts notify users of new acquisitions, overdue items, and library events, keeping them informed. Users stay engaged and up to date with library services and resources.

- User Reviews and Recommendations: User-generated reviews and ratings on library websites help users make informed decisions about resource selection. Recommendations from peers and the community enhance the discovery of valuable information.

- Research Tools and Guides: Online tutorials, research guides, and subject-specific resources support users in their quest for information. These tools assist in efficient information retrieval and understanding of complex topics.

- Virtual Community Building: Social features on library websites enable users to connect and share resources, fostering a sense of community among information seekers. Virtual book clubs and discussion forums encourage intellectual exchange.

- Privacy and Security: Library websites prioritize user privacy and data security, ensuring that personal information and browsing habits are safeguarded. Strict security measures protect user accounts and data.

- Data Analytics and Continuous Improvement: Usage statistics and analytics provide insights into user behavior and preferences, allowing libraries to refine their offerings. User feedback and assessment help in making data-driven improvements.

7. Challenges And Opportunities

It is important to discuss both the challenges and opportunities that modern library websites present. Here are key points:

Challenges:

1. Digital Divide: The digital transformation of libraries may exclude individuals without access to the internet or digital devices, exacerbating the digital divide.

2. Information Overload: The sheer volume of digital resources available can overwhelm users, making it challenging to find relevant information.

3. Privacy Concerns: Online access to information can raise privacy concerns, particularly regarding user data collection and tracking.

4. Digital Literacy Gaps: Not all users possess the necessary digital literacy skills to navigate library websites and evaluate online information critically.

5. Maintenance and Updates: Ensuring the reliability and functionality of library websites requires ongoing maintenance and updates, which can be resource-intensive.

6. Security Risks: The digital environment exposes library websites to cybersecurity threats, potentially compromising user data and resources.

Opportunities:

1. Global Access: Library websites provide the opportunity for global access to library resources, breaking down geographic barriers.

2. Personalization: Personalized features allow libraries to cater to individual user preferences, enhancing the user experience.

3. Remote Learning and Research: Library websites facilitate remote learning and research, enabling users to access resources from anywhere.

4. Data Analytics: User data collected by library websites can be leveraged for data-driven decision-making, improving services.

5. Community Engagement: Social features and virtual book clubs on library websites encourage community building and intellectual exchange.

6. Research Collaboration: Library websites can foster research collaboration by connecting scholars and resources.

7. Cost Savings: Digital resources and virtual services offered through library websites can result in cost savings for libraries and users.

8. Continuous Improvement: Feedback mechanisms and usage analytics empower libraries to make data-informed improvements.

9. Enhanced Accessibility: Web accessibility standards promote inclusivity, ensuring that all users, including those with disabilities, can access resources.

10. Resource Sharing: Library websites can facilitate resource sharing among libraries, expanding the availability of materials.

Including both challenges and opportunities provides a balanced perspective on the role of modern library websites in information access.

8. Emerging Directions In Library Website Advancements

To explore future trends in library website development, the following potential trends and their implications are worth considering:

- AI and Machine Learning Integration: Library websites may incorporate AI-driven recommendation systems, chatbots for user support, and machine learning algorithms for enhanced search capabilities.

- Personalization and User-Centric Design: Websites will continue to focus on user experience, offering increasingly personalized features and interfaces tailored to individual preferences.

- Semantic Web and Linked Data: The adoption of semantic web technologies and linked data will enable more sophisticated resource discovery and interconnectivity among library websites and external data sources.

- Voice Search and Virtual Assistants: Library websites may integrate voice search options, allowing users to search and navigate resources using voice commands and interact with virtual assistants.

- Enhanced Data Analytics: Libraries will leverage advanced analytics to gain deeper insights into user behavior and preferences, enabling more data-driven decision-making.

- Blockchain for Authentication and Security: Blockchain technology could be employed to enhance user authentication, data security, and the integrity of digital resources.

- Open Access and Open Educational Resources (OER): Library websites may serve as hubs for open access resources and OER, providing users with cost-effective and open alternatives to traditional materials.

- Interdisciplinary Collaboration: Collaboration between libraries and other institutions may increase, fostering interdisciplinary research and resource sharing.

- Sustainable Design and Accessibility: Future library websites will prioritize sustainable design practices and accessibility compliance to create eco-friendly and inclusive digital environments.

- Content Aggregation and Curation: Libraries may focus on aggregating and curating content from various sources to provide users with comprehensive, one-stop access to information on specific topics.

- Augmented Metadata and Enhanced Discovery Services: Augmented metadata and improved discovery services will enhance the precision and relevance of search results.

- Mobile-First Approaches: Library websites will increasingly adopt a mobile-first approach to cater to the growing number of users accessing resources via smartphones and tablets.

- Collaborative Learning Spaces: Library websites could create virtual collaborative learning spaces that facilitate group work, discussion, and knowledge sharing among users.

- Data Preservation and Digital Archives: Libraries may develop tools and services for preserving digital materials, ensuring long-term access to valuable resources.

9. Conclusion

This research paper explores the dynamic evolution of library websites, examining their functions and features, evaluating their influence on information access, and discussing the challenges and opportunities they present. Modern library websites have democratized access to information, prioritized accessibility and inclusivity, and enhanced search capabilities, all while offering personalized user experiences. Challenges include the digital divide and privacy concerns, but opportunities arise from global access, personalization, and remote learning. Future trends in library website development include AI integration; voice search, open access support, and sustainable design, ensuring libraries remain at the forefront of information dissemination and user engagement.

References

1. Anna, N.E.V., Transformation of public library websites in Indonesia, *Library Hi Tech News*, 35 (8), 10-14.

2. Ex Libris A ProQuest Co., Three Trends Shaping the Future of Libraries. *Library Journal*. 2022, Jan 3, https://www.libraryjournal.com/story/three-trends-future

3. Verma, M. K., & Devi, K. K., Web content and design trends of Indian institutes of management (IIMs) libraries website: An analysis. *DESIDOC Journal of Library and Information Technology,* 36(4), 220–227, 2016. https://doi.org/10.14429/djlit.36.4.9983

4. Yoon, A., & Schultz, T., Research data management services in academic libraries in the US: A content analysis of libraries' websites. *College and Research Libraries*, 78(7), 920–933, 2017. https://doi.org/10.5860/crl.78.7.920

5. Al-Qallaf, C. L., & Ridha, A., A Comprehensive Analysis of Academic Library Websites: Design, Navigation, Content, Services, and Web 2.0 Tools. *International Information and Library Review*, 51(2), 93–106, 2019. https://doi.org/10.1080/10572317.2018.1467166

6. Bharati, S. K., & M. M., Content Evaluation of Jawaharlal Nehru University and Banaras Hindu University Library Websites in India. *Library Philosophy and Practice (e-journal),* 1-23, 2019

7. Hugar, J. G., Content analysis of engineering college library websites in Goa. *Library Philosophy and Practice (e-journal),*2019. https://doi.org/10.2139/ssrn.3620064

8. Chikkamanju, & Kumar, G. K., Web 2.0 Features in the Library Websites of Agricultural Universities in India. *International Journal of Research in Library Science (IJRLS), 6*(2), 12-23, 2020

9. Pal, A., Sarkar, A., & Kar, S., Content Analysis of Asian National Library Websites. In *Application of Webometrics on Modern Information Research* (pp. 76-103). New Delhi: Ess Ess Publications, 2020.

10. Rahman, A., & Sadik B., Content Analysis of Library Websites of Select Colleges of Delhi University. *DESIDOC Journal of Library & Information Technology, 40*(04), 247-252, 2020

11. Ambika, C. A., & Ganesan, P., Central University Library Websites in India: Web Content Analysis. *Library Philosophy and Practice (e-journal),* 2021.

12. Barman, D., Ranking of Library Websites of Agricultural University of India: A Study. *Library Philosophy and Practice (e-Journal), 2021*.

13. Li, T., Tang, J., Xiao, L., & Cai, M., Evaluation of Smart Library Portal Website Based on Link Analysis. *In Procedia CIRP* (Vol. 188, pp. 114–120). Elsevier B.V., 2021 https://doi.org/10.1016/j.procs.2021.05.059

14. Singha, S. C., & Verma, M. K., Web Content Analysis of Veterinary University Library Websites in India: An Evaluation. *Library Philosophy and Practice (e-Journal)*, 2021, 1–22

15. Brahma, K., & Verma, M. K., Web Content Analysis of National Libraries' Websites in Asia: An Evaluation. *International Journal of Information Science and Management,* 20(1), 427–448, 2022.

16. Tavosi, M., & Naghshineh, N., Google SEO Score and Accessibility Rank on the American University Libraries' Websites: One Comparative Analysis. *Information Discovery and Delivery,* 51(2), 241–251, 2023. https://doi.org/10.1108/IDD-08-2021-0088

A Comparative Study on the Physicochemical Parameters of the Water of Chambal River in Kota

[1]Shaheena Khan, [2]Dr. A. V. Trivedi

[1]Research Scholar School of Basic and Applied Sciences, Career Point University Kota Rajasthan
[1]khanshahina1984@gmail.com

, [2]Research Supervisor, School of Basic and Applied SciencesCareer Point University Kota Rajasthan
[2]dr.anandtrivedi@gmail.com

Abstract:

River plays a very important role in life and quality of water and affects very much to the health and survival of life. Thus In the present study physicochemical parameters of chambal river water of Kota Rajasthan India have been studied and compared. Total fifteen parameters, viz. temperature, pH, turbidity, TDS, total hardness, total alkalinity, chloride, fluoride, nitrate, phosphate, sulphate, calcium, magnesium, BOD and COD have been studied. Four sites of Chambal river from Kota viz. Kota Garh Palace, Under Chambal Bridge Nyapura, Adharshila and Gawadi have been chosen for the present study. Results of the study reveals that except fluoride and phosphate all parameters were found under the permissible limit as given by Indian standard specification for drinking water(IS10500/WHO). Thus it has been recommend that this water can only be used for drinking after proper purification.

Key words : Chambal river, Water pollution, Physicochemical study Water analysis, water quality.

Introduction:

People along the globe uses water for many purposes e.g.,for drinking, for forming and for industry purposes etc. However, the water quality of rivers are deteriorating due to human activities like anthropogenic activities, industrialization, farming activities, transportation,

urbanization, animal and human excretions and domestic wastes disposal etc. The poor quality of water not only causes many diseases in the humans but also effect the productivity of crops so that quality of the water must be tested regularly [1]. Pollution of river is a global problem [2]. As be already reported in our previous paper that in India about 70% of the available water is polluted. The chief source of pollution is identified as sewage constituting 84 to 92 percent of the waste water. Industrial waste water comprised 8 to 16 percent [3 - 5]. Although much attention have already been paid to study pollution of rivers like Godavari, Krishna, Tungbhdra, Cauvery, Jhelum, Kosi, Alaknanda, Betwa, Ganga and Yamuna etc. [6 - 8]. However, there is a research gap and little attention has been so far paid for the Physicochemical analysis of water of river Chambal in the area under study. Currently this river is facing big pollution problems due to encroachments, discharge of untreated domestic and industrial waste, dumping of solid waste and illegal diversion of water. We therefore studied physicochemical parameters of Chambal river in Kota Rajasthan India the area were selected as Garh Palace, Under Chambal Bridge Nyapura, Adharshila and Gawadi and their comparison with the Indian standard IS10500/ WHO standard.

Review of literature:

May researchers studied physicochemical parameters of many rivers. A brief review of literature has been given here. Rai [9] studied pollution and conservation of Ganga river water and repotted that pollution levels in the Ganga are contributing 9-12% of total disease burden in Uttar Pradesh (U.P.). Gupta and Chakrapani studied pollution level of the Narmada river basin [10]. Subramanian documented inconsistent down-stream variations in river water chemistry [11 - 12]. Singh [13] and Mukherjee et al. [14] documented the physical, chemical and biological aspects of the Ganga river. Heavy metals such as Cr, Mn, Fe, Co, Ni, Cu, Zn, and Pb in the sediments of the Ganga river basin have been analysed and reported [15, 16]. The water of Chambal River is getting contaminated from various point of drainage and directly dumps ground. Many researchers reported the water quality of Chambal River of Kota city [17 - 19]. Sarang et al. revealed that organic matter content in river caused highly contamination of hazardous pathogen as well other biological organism [20]. It has been reported that discharge of large quantities of waste causes thermal pollution

and it affects the ecology of the ecosystem of Kota Barrage Dam. Chambal River is also major water source for wild life sanctuary of Madhya Pradesh. In national Chambal sanctuary, water quality shows that most of aquatic fauna and flora maintain biodiversity of particular area and it indicated that the river water in the sanctuary area is pollution free and can serve as a good habitat for many aquatic animals including endangered species [21]. However, there is a research gap for the physicochemical study of the water of Chambal river in the area under study, I.e. Kota Garh Palace, Under Chambal Bridge Nyapura, Adharshila and Gawadi. Thus we selected this area for the present study.

Research Methodology:

Study Area:

The present study focuses on the Chambal river that flows through heart of the Kota City, one of the most prominent industrial and educational town of Rajasthan state in India. The district Kota lies between 24°25' and 25°51' North latitudes and 75°31' and 77°26' East longitudes with total area of 5767.97 Sq Kms. "Kota City" is located at extreme south of it at 25°11' North latitude and 75°51' East longitude occupying total area of 238.59 Sq Kms with average height 253.30 meters from sea level. The following locations of Kota city have been selected for the pollution study of chambal river. Water samples were collected following the standard procedure and analysed according to the APHA and standard methods.

Sample site - I   Kota Garh Palace

Sample site - II   Under Chambal Bridge Nyapura

Sample site - III   Adharshila

Sample site - IV   Gawadi

Collection and Analysis of Water Samples

All chemicals used were of AR grade purity. Water samples were collected for the present investigation from all four different experimental sites of Chambal river for the three seasons i.e., in the month of May for summer season, in the month of August for rainy season and in the month of December for winter season. The Samples were collected in screw capped polyethylene bottles. Sample bottles were thoroughly rinsed and dried and finally rinsed with river water before collecting samples. Caps of bottles were closed tightly after filling up of bottles to avoid changes in physico-chemical characteristics. Total 15 parameters namely pH, Turbidity, TDS, total alkalinity,chloride, fluoride, nitrate, phosphate, sulphate, calcium and manganessium DOD and COD were determined according to the procedure prescribed by APHA. All the analysis were done at public health engineering department (PHED) chemical laboratory Kota Rajasthan India and Go Green Technology Kota Rajasthan India.

Research findings:

Results :

Results of the present investigation have been given in the following tables 1.0 to 3.0

Table 1.0 : Physicochemical parameters of Chambal river in Kota Rajasthan

(May 2020)

S.No	Parameters	Site -I	Site -II	Site - III	Site - IV
1	**Temperature**	30.2	30.2	30.1	31.0
2	**pH**	8.0	7.9	8.1	8.2

3	Turbidity	6.9	3.8	5.8	6.2
4	TDS	390	392	395	399
5	TH	170	168	162	185
6	TA	159	154	153	179
7	Chloride	90	85	111	95
8	Fluoride	0.03	0.41	0.3	0.3
9	Nitrate	7.5	9.1	8.3	9.6
10	Sulphate	16.2	20.5	16.0	21.0
11	Phosphate	0.12	0.06	0.17	0.16
12	Calcium	48.0	65.0	46.0	47.0
13	Mangnesium	13.2	17.2	13.1	18.5
14	BOD	1.87	1.83	1.97	1.84
15	COD	31.0	30.0	40.0	30.0

Table 2.0 : Physicochemical parameters of Chambal river in Kota Rajasthan

(August 2020)

S.No	Parameters	Site -I	Site -II	Site - III	Site - IV
1	Temperature	25.6	25.8	25.6	25.9
2	pH	6.7	6.8	6.6	6.8
3	Turbidity	5.9	2.9	3.9	5.3
4	TDS	380	381	380	397
5	TH	141	138	135	145
6	TA	145	148	150	180
7	Chloride	83	79	97	85
8	Fluoride	0.03	0.31	0.27	0.34
9	Nitrate	6.8	8.3	7.9	9.0
10	Sulphate	15.1	19.2	15.5	19.5
11	Phosphate	0.11	0.05	0.15	0.14
12	Calcium	43.0	51.0	42.0	40.0
13	Mangnesium	11.3	16.0	12.0	17.5
14	BOD	1.40	1.43	1.90	1.39

15	COD	27.0	29.0	30.0	31.0

Table 3.0 : Physicochemical parameters of Chambal river in Kota Rajasthan

(December 2020)

S.No	Parameters	Site -I	Site -II	Site - III	Site - IV
1	Temperature	12.3	12.5	12.3	12.5
2	pH	7.6	7.5	7.4	7.7
3	Turbidity	6.5	3.6	5.0	6.1
4	TDS	395	396	398	419
5	TH	180	183	185	190
6	TA	155	157	157	185
7	Chloride	98	93	120	97
8	Fluoride	0.05	0.3	0.26	0.32
9	Nitrate	7.2	8.9	7.9	9.5
10	Sulphate	16.3	20.2	17.0	21.0

11	**Phosphate**	0.15	0.07	0.20	0.19
12	**Calcium**	50.1	65.0	49.0	48.0
13	**Mangnesium**	13.5	18.0	13.0	20.0
14	**BOD**	1.70	1.80	1.95	1.60
15	**COD**	29.0	30.0	37.0	29.5

Discussion:

Results of the present investigation are given in the above table 1-3. As per the above results water quality of chambal river in the study area are discussed below along with the comparison of Indian standard specification for drinking water IS 10500/WHO standard. Temperature varies considerably minimum temperature was recorded 12.3 °C in winter i.e., in the month of December in site I and III and maximum was recorded 30.2 °C in site I and II in the month of May in summer season. The negative logarithm of hydrogen ions is called pH. The value of pH of a solution is a measure of their acidity or alkalinity. The permissible limit of pH values for drinking water is specified as 6.5 to 8.5 as per IS 10500. The pH values of chambal water samples from all four sample sites I to IV in all three seasons were found between 6.6 to 8.2 . This is under permissible limit. Turbidity of site I to IV were found to be 2.90 to 6.5 which is under acceptable limit 10 NTU as per the IS 10500 standard. TDS (Total dissolve solids) in the present case were found in the range of 380 to 419 mg/l which shows good quality of water as per IS 10500. Total hardness were found to be between 135 to 190 mg/l. Total alkalinity were found to be 145 to 190mg/l which is under permissible value 200 mg/l. In the present study chloride and fluoride concentration were found to be between 79 to 120 mg/l and 0.03 to 0.41 mg/l respectively. Chloride concentrations are acceptable as per Indian standard IS 10500. However, as per IS 10500 concentrations of fluoride is below 0.6 mg/l thus water is rejected for drinking purpose. Low concentration of nitrate i.e. 6.8 mg/l to

9.6 mg/l were found in the sites I to IV this is acceptable because as per IS 10500 upper limit of Nitrate is 45 mg/l. Concentration of phosphate were found to be 0.06 to 0.20 mg/l. There is no IS 10500 standard permissible limit for phosphate for drinking water, while WHO (1993) has fixed it to be 0.1 mg/l. Thus only in one sample site (site-II) phosphate concentration was found under WHO limit. The Concentration of sulphate were found to be 15.1 to 20.0 this is acceptable under IS 10500 limits. Concentrations of magnesium and calcium were found to be 11.3 to 20.0 mg/l and 40.0 to 65.0 mg/l respectively. As per IS 10500 maximum acceptable limit of magnesium is 100 mg/l and that of calcium is 200 mg/l. Thus both metals are found within acceptable limit. BOD and COD level were found to be 1.39 to 1.97 mg/l and 27.0 to 40.0 mg/l. These were under acceptable limit of Indian standard i.e., 30 mg/l and 250 mg/l for BOD and COD of surface water respectively.

Conclusion and Recommendation :

On the basis of the present study it has been concluded that Chambal river water in the areas under study, i.e. Kota Garh Palace, Under Chambal Bridge Nyapura, Adharshila and Gawad is polluted. Although some physicochemical parameters have been found under permissible limit as given by Indian standard IS 10500 / WHO. But the water can only be used for drinking after proper purification.

References:

1. Kummu, M. et al. (2012) Lost food, wasted resources: global food supply chain losses and their impacts on fresh water, cropland, and fertiliser use; Science of the Total Environment; 438:477–489

2. R.K.Trivedyet al. (1990) River pollution in India; Ashish Publishing House, New Delhi; 26,99

3. Shaheena Khan and A.V.Trivedi (2012) ; A Study on the Pollution Level of Water of Chambal River in Kota Rajasthan; JETIR; 8 (12) 225 -227

4. D. M. Joshi et al. (2009); Studies on physicochemical parameters to assess the water quality of river ganga for drinking purposes in haridwar district; Rasayn J. Chem.; 2 (1) 195 - 203

5. P. K. Goel (2006) Water Pollution: Causes, Effects and Control; New Age International Publishers, ISBN: 8122418392

6. D. N. Saksena et al. (2008) Water quality and pollution status of Chambal river in NationalbChambal sanctuary, Madhya Pradesh; J.Environmental Biology; 29 (5) 701-710

7. Bhutiani, R.et al. (2016) Assessment of Ganga river ecosystem at Haridwar, Uttarakhand, India with reference to water quality indices. Appl. Water Sci.6, 107 – 113

8. Rohit Sharma et al. (2020) Analysis of Water Pollution Using Different Physicochemical Parameters: A Study of Yamuna River; Frontiers in Environmental Science;8; 1-18

9. Rai, B., (2013). Pollution and Conservation of Ganga River in Modern India. International Journal of Scientific and Research Publications, 3(4), 1 - 4

10. Gupta, H., & Chakrapani, G.J., (2005). Temporal and spatial variations in water flow and sediment load in Narmada River Basin, India: natural and man-made factors. Environ. Geol., 48, 579 - 589.

11. Subramanian, V., (1983). Physicochemical assessment of water quality of river chambal in Kota city area of Rajasthan state (India). Analyst, 83, 961.

12. Sundaray, S.K., et al. (2006). Multivariate Statistical Techniques for the Evaluation of Spatial and Temporal Variation in Water Quality of Mahanadi River-Estuarine System (India) A Case Study, Environ. Geochem. Health, 28, 317 - 330.

13. Singh, N. (2010). Physicochemical properties of polluted water of river Ganga at Varanasi. International Journal of Energy and Environment, 1(5), 823 - 832.

14. Mukherjee, D. et. al (1993). Water quality of the River Ganga (The Ganges) and some of its physico-chemical properties, Environmentalist, 13, 199 - 210.

15. Purushothaman, P., & Chakrapani G.J. (2000). Heavy metal fractionation in Ganga river sediments. Environ. Monit. Assess., 132, 475 - 489.

16. Singh, M. et al. (2002). Heavy metals in freshly deposited stream sediments of rivers associated with urbanisation of the Ganga plain India. Water, Air, Soil Pollut., 141, 35 - 54.

17. Kumar, N., & Chaudhary, M.P. (1017). Effect of water quality of Chambal River due to discharge of open drains in Kota city. Int. Rea. J. of Eng. &Tech, 4 (11), 890 - 894.

18. Patil, P.N., et al. (2012). Physico-chemical parameters for testing of water- A review. Inter. J. of Environ. Sci., 3(3), 1194 - 1207.

19. Parashar, C., et al. (2006). Seasonal variation in physico-chemical characteristics in upper lake of Bhopal. Asian J. Exp. Sci., 20(2), 297 - 302.

20. N. Sarang, N., & Sharma, L.L. (2014). Impact of sewage discharge of water quality and benthic diversity of Kota barrage, Kota, Rajasthan, India. J. of Ind. Poll. Cont., 30 (2), 203 - 220.

21. Saksena, D.K., et. al (2008). Water quality and pollution status of Chambal River in National Chambal Sanctuary, Madhya Pradesh. J. Environ. Biol., 29 (5), 701 - 710

A Review of Data Handling Approaches

[1]Avinash Saini

CSE, School of Engineering & Technology, Career Point University, Kota, Rajasthan,

India

avinashsaini39@gmail.com

[2]Dr. Manish Tiwari

Associate Professor, CSE, School of Engineering & Technology, Career Point

University, Kota, Rajasthan, India

Abstract:

A Paragraph of a maximum of 200 words to explain the crux of the entire Data in today's scenario forms the base of all transactions performed throughout the world. Authenticity accuracy and speed of applications depends on the data being served to them. Data visualization, storage, security, speed are various aspects on which lots of research is being done. This paper focuses on presenting a comparative view of various data management techniques. It's a review of various forms of data management and provides a bird's eye view of their evolution.

Keywords*: Data management, DBMS, RDBMS, OODBMS, Big Data.

Introduction:

Since the era of computing, the need to manage data started. Many trends in data management were introduced such as File System, RDBMS, ODBMS and many more. Traditional approaches were cumbersome and time consuming so to eliminate their limitations database approach was introduced. To organize large bulk of data on server, concept of Big Data was introduced. This review is about the various trends of managing data that were developed from traditional approaches to latest approaches such as Big Data, XML and others. We also aim to propose the context in which certain trends should be used

Existing Approaches

Here we have discussed about the following approaches such as File System, the Database, RDBMS, OODBMS, ORDBMS, Big Data which helps us to manage data, retrieve data and establish relationship between them and also to handle large amount of data

A. File System

File System allows us to control how we can store and retrieve data. They are designed for specific applications such as the ISO 9660 is designed for optical discs. It helps us to store large amount of data which can be shared with another database users and information is stored permanently.

- Need – A file system is a collection of related bytes where the file can be free formed, indexed, structured, etc. There are many kinds of file system which can be used on different kinds of storage devices as each storage device uses different kind of media and common storage device. In this mostly two types of Access It may have attributes such as name, identifier, type, location, size, protection, time, date and user identification. methods are used which are Sequential Access and Direct Access. Sequential method is implemented by a file system and data is accessed one record right after the last and this type of method is reasonable for tape and Direct method is useful for disks in which file is viewed as a numbered sequence of blocks or records and has no restrictions, blocks are read/write in any order.

- Merits – It provides heterogeneous operating systems including Unix, Linux and Windows operating systems. It gives multiple client machines which can be accessed by single resource simultaneously. It allows us to control how data is stored and retrieved. It helps in organizing and keep track of the files and directories.

- Restricted Areas – It requires the enough memory to store files in drive, if the file format is incorrect it cannot recognize the right program which is to be open. There is restriction of certain symbols in file system.

B. The Database

Database is a collection of interrelated files. It helps us in creating, retrieving, updating and deleting data which is organized in the form of

tables. Some of the database software such as Fox Pro, MySQL Server and many more.

- Need – It provides us data in an organized way, access to them and protects them from any loss. It stores data in convenient way which is easy to retrieve, manipulate and helps to produce information. We need database to maintain records accuracy. It helps in maintenance of proper data and security to data.

- Merits – Modern database are more realistic to design architecture as it uses real world entities. It provides less redundancy as possible as it follow rules of normalization. It provides facility to retrieve and manipulate data in an efficient way. It uses the concepts of Acid Properties i.e. Atomicity, Consistency, Isolation and Durability. It provides multi-user environment and allows them to access and manipulate data in parallel. It gives security at some extinct where users are unable to access data of other users and departments.

- Restricted Areas – They are complex and extremely large to understand and it is also time consuming while designing. The cost of hardware and software are considerably of high, as database is affected adversely it affects all application programs and there is requirement of training for both programmers and users.

C. RDBMS (Relational Database Management System)

It is database system made up of files including data elements in two dimensional array. It defines the approach to design schema that are in appropriate normal forms. It defines various normal forms in term of functional dependencies and data dependencies.

- Need – It has the capability to recombine data elements to form different relations while using data in flexible way. Its main elements are based on Codd's 13 rules for relational systems. Thse relational database is recognized as a collection of tables. Each table consists of a series of row/columns intersection.

- Merits – It supports very large databases. It is easy to use, design, implement and manage in efficient way. We use Structure Query Language in this database. It is powerful database management system.

- Restricted Areas – In this database we have not enough storage area to handle data just as images, digital and audio/videos. It has restrictions that information must be in tables. It not supports adequate storage of complex objects.

D. OODBMS

(Object Oriented Database Management System)

This database is object oriented; an OBJECT corresponds to an entity in E-R Model. The object oriented paradigm is based on inheritance, data encapsulation, object identity and polymorphism. An Object is associated with set of variables, set of messages and set of methods.

- Need – Its main objective is to provide secured, controlled, data independent and consistent services so that object-oriented model may be supported. They were created to handle big and complex data that relational database could not. While joining of object-oriented programming with database technology, it provides an integrated application development system.

- Merits – It provides us the possibility of reusing the code. It can be formulated with small procedural differences without affecting the entire system. In this we can use enhanced modeling capabilities which is used in applications like computer-aided design (CAD), Computer-aided Software engineering(CASE), multimedia systems and many more.

- Restricted Areas – There is lack of experience as it is still relatively limited. It is typically based on specific language. In this when schema is updated or modified; we need to update the instance of that class.

ORDBMS

(Object Relational Database   Management System)

This database model extends the relational database model by providing a richer type system including complex data types and object orientation. It provide convenient migration path for users of relational database who wish to use object oriented features.

- Need – Object-relational database is similar to relational database. It provides interface between both object-relational and relational database. This database is created to

handle new data such as audio, video and image files as compared to relational database. When ORDBMS was development started, as a result usage of object-oriented programming languages was increased. Its main objective is to bridge the gap between relational databases and object-oriented modeling techniques used in programming languages such as Java, C++, Visual Basic .NET.

- Merits – It provides service to organizations to work with their current systems, without making major changes to them. It provides users and programmers to start using object-oriented systems in parallel. It is user-extensible type system in which dynamic binding of methods is used.

- Restricted Areas – In ORDBMS we accomplish relationships and encapsulated objects correctly than it will be   regarded in disorder which is a drawback for same. This is a complex process.

F.   BIG DATA

This technique is used to organize and handle large amount of data present on the server. and integrate them with per-existing enterprise data to be analyzed.

- Need - It helps us to acquire data, organize data and analyze data. If we utilize Big Data, enterprises must involve their IT infrastructures to handle these new high-volume, high-velocity, high-variety of data

- Merits – Big Data Analytic Tools may be expensive, but eventually it is cost saving. This data is secured as it can save the average company 1.6% of annual revenues. In this errors and fraud can be detected at the moment and precautions could be taken to limit the damage. With help of this we can achieve cost reductions, quicker and better decisions and can provide new offers to customers. In this we can access new data sources and use into different type of data to generate value from that data.

- Restricted Areas -   This is mainly used in selected channels which are difficult to apply user data directly to execution. It can be presented through visualizations such as flow diagram as it is not accurate for all users. It requires care in transferring from machine to machine and server to server.

G. XML (Extensible Mark-up Language)

This technique is used to create information formats and share electronically structure data through public Internet in a flexible way. This can also be known as self-describing or self-defining. The XML format information can be shared by any individual or group of individuals or companies.

- Need – It is a textual data format which provides strong support via Unicode from different human languages. It is mostly used for the representation of random data structures such as those used in web services. It is a very important part of Web and all electronic information in our future world. It works with many other technologies to display and process information which looks like working on HTML. Its main objective is to emphasize simplicity, generality and usability across the Internet. As it is said, mark-up language in which document can be in human-readable and machine-readable format.

- Merits - It is easy to read and understand as it is easily processed by computers. There are no restrictions on set of tags, they can be created whenever needed. As it contains meta data, it can be stored without schemas. It allows comparing document efficiently element by element. In this document storage and processing can be formatted both offline and online. It is platform- independent so we can make changes in technology as accordance.

- Restricted Areas – Its data redundancy may affect application efficiency. The hierarchical structure for representing data is limited. The XML documents must be converted into HTML before presenting to user. This database is still on experimental as it has not solidified yet.

BASIS	FILE SYSTEM	RDBMS	OODBMS	ORDBMS	XML	BIG DATA
MAIN OBJECTIVE	Store data in different formats.	Data independence and data integrity.	Data encapsulation and independence.	Achieve benefits of both relational and object model.	Mark-up language which have unlimited tags and transfer data in different formats.	Organise voluminous amount of data in efficient manner.
REDUNDANCY	Yes	No redundancy	No redundancy	No redundancy	Yes	Statistical data redundancy
DATA SECURITY	No data security	Support Limited security	Support strong security	Support strong security	Secure data approach	High security is there
RELATIONSHIP BETWEEN DATA	Data is not related	Data is related by references between tables.	Data is related	Data is related	Data is not related	Data is not related
COMPLEXITY	File system is complex as access and updation of data is not easy.	Structure is simple as users perceive data columns, rows/tuples and tables	Structure of data is complex as it involves different data types.	It is also a complex approach.	It provide simplicity, generality and usability across the internet	It is a complex technique as it requires an IT platform and Data Warehousing.
STORAGE CONTENT	Store data in random or indexed form.	Store data in form of tables or entity sets.	Store data and objects.	Store data and objects.	Set of rules for encoding documents in a format which is both human and	Store machine generated and social data

					machine readable	

COMPARATIVE STUDY OF DATABASE APPROACHES

We have discussed various approaches of managing data but it is a still big decision to decide what format of data storage should be selected for the current project. In the current section we provide a comparative study of different approaches which assist the developers in selecting the format of storing data.Table 1 Presents this comparative study. The main objective of File System and XML is to transfer and store data in different formats .RDBMS and OODBMS provides data independence but RDBMS also provides data integrity and OODBMS provides data encapsulation. ORDBMS and Big Data provide benefits of both relational and object model and have also organized voluminous amount of data in efficient manner respectively. File system, XML and Big Data results in statistical data redundancy as compared to RDBMS, OODBMS and ORDBMS. OODBMS, ORDBMS, XML and Big Data provide strong and high security as compared to File System and RDBMS. In File System, XML and Big Data. File System, OODBMS, ORDBMS and Big Data are complex technique as it requires data warehousing and updating is not so easy as compared to RDBMS and XML. In OODBMS and ORDBMS we can store data in objects but in File System we store data in random or indexed form, in RDBMS we store data in form of tables and entity sets as compared to XML and Big Data ,the data stored in them is easily readable to user and machine in which document is in encoded format.

IV. CONCLUSION

In thispaper we have discussed about the approaches recently used in today's world. We come to know how data can be managed and in which manner it can be arranged. These all approaches used for managing data are efficient and effective. The approaches such as File System, Database, Rdbms, Odbms, Oodbms, Ordbms, Big

data, Xml which are uniquely identified and every approach has its own importance. If some are complex to use, than some are easy for user to use and understand.

V. FUTURE SCOPE .

In today's world data is playing a big role in decision making and managing big organizations in an efficient way. We presented a comparative study and further we aim to propose a generic model for data management which can be customized and adapeted according

References

1. Matthew Trunnel Manager, Research Computing Broad Institute

2. Revista Informatica Economică, nr. 4 (44)/2007 Gheorghe SABĂU, Bucharest, Romania

3. Bennett, S., Myatt, M., Jolley, D., & Radalowicz, A. (2001). Data Management for Surveys and Trials.

4. A Practical Primer Using EpiData. The EpiData Documentation Project. Available: www.epidata.dk/downloads/dmepidata.pdf.

5. Calvert, William S. and J. Meimei Ma. Concepts and case studies in data management . Cary, NC: SASInstitute, c.

6. Davidson, Fred. Principles of statistical data handling. Thousand Oaks, California, SAGE, ,266pp.

7. Graham JW, Hofer SM, Piccinin AM. Analysis with missing data in drug prevention research. IN:

8. Collins LM, Seitz LA (eds). Advances in data analysis for prevention intervention research.NIDA Research Monograph 142. U.S.

9. Hulley, Stephen B. and Steven R. Cummings. Designing clinical research: an epidemiologic approach. Baltimore, Williams & Wilkins, 1988. Chapter 15: Planning for data management and analysis.

Cataract Detection through Digital Image Processing Techniques using MATLAB

[1]Chetan Kumar Asnani, [2]Kamal Arora

[12]Department of Electrical Engineering, Career Point University Kota, 324005 India

Email: er.chetan2020@gmail.com, kamal.arora@cpur.edu.in

Abstract

The Cataract disease that affects the Human eye is cloudiness on the lens of the eyes which affects the Vision. Generally, it occurs at a higher age. It may occur in one or both eyes. Cataracts are called Eye disorders. Cataracts may also cause sightlessness. Here the work in the paper mainly suggests and predicts an algorithm to detect the eye disorder cataracts in adult human beings from their eye images. Presently the methods that are in use for cataract detection are focused on the application of either DSLR camera or fundus image camera. Both methods mentioned above are at high cost. The main target of this work is to create convenient algorithms and a low-cost and robust method for further application for suitable device manufacturing to detect cataracts from the color eye image. The algorithms are designed by using digital image processing in MATLAB. This method focuses on cataract detection based on the eye image intensity, uniformity, and texture feature. These features are first to take considering the inserted digital image. After that, the image is converted into gray. The process of RGB to GRAY is done. To process the image further contrast of image is done. In the proposed algorithms subsequent noise is removed by median filter. Now the image is ready for segmentation of the pupil of the eye image. After segmentation resized is done to evaluate the image whether cataract Eye or is normal eye.

Keywords: Cataract, Digital Image Processing, Matlab.

1. INTRODUCTION

The sensitive sense organ and valuable organ is the human eye. It makes us see the colors around us and the wonderful world. On closing our eyes, we cannot identify taste, or smell and even cannot analyze the objects and their color. So the eye makes us see a beautiful and colorful world. The eye of the human being works the same as the

camera. The eye has a lens system that creates an image. The light-sensitive screen of the eye is called the retina where the image of the object is formed. The thin membrane part of the eye is called the cornea from where the light enters. There is a formation of a transparent bulge. This is formed on the front surface of the eyeball. The shape of the eyeball is mainly spherical and has a diameter of 2.3 cm. The light enters from the outer surface. The refraction takes place at the outer surface of the eye lens. The Cataract disease that affects the Human eye is cloudiness on the lens of the eyes which affects the Vision. Generally, it occurs at a higher age. It may occur in one or both eyes. Cataracts are called Eye disorders. Cataracts may also cause sightlessness. This paper subjectively focuses on the detection of cataracts from color images. Presently the methods available are the fundus image method and the DSLR method. Both these methods have a very high cost. The main purpose is to develop robust and convenient algorithms. The designed algorithms can be instrumented with hardware devices so that they can detect the presence of cloudiness on the eye lens. The algorithms are designed by using digital image processing in MATLAB. The functional algorithm is based on feature extraction of image pixel intensity and the pixel uniformity and texture feature of the image. These features are first to take considering the inserted digital image. After that, the image is converted into gray. The process of RGB to GRAY is done. To process the image further contrast of image is done. In the proposed algorithms subsequent noise is removed by median filter. Now the image is ready for segmentation of the pupil of the eye image. After segmentation resized is done to evaluate whether the image whether Cataract Eye or is normal eye.
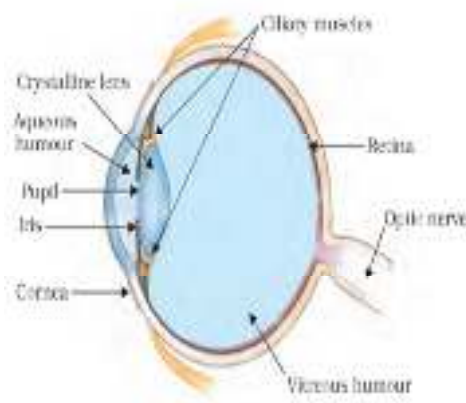

Figure 1. Eye Structure [205]

Eye cataract: A cataract is an eye disorder. it affects the vision due to the formation of clouding of the eye lens which affects the vision. The rise of cloudiness of the lens increases the cataract increases from mild to high cataract. This is in mostly higher people. as the level of cataracts increases the vision of the eye decreases.

DIP tool and techniques for automated detection of Eye diseases: Nowadays there is a requirement for the automated detection technique in image processing. For various eye disorders this system of processing certain various procedure steps namely the insertion of an eye image. RGB to GRAY conversion, noise removal, segmentation of pupil, feature extractions. The cost of MRI & HRT are very high. Hence these simple basic algorithms-based systems are quite in demand. The preprocessing involves the steps of boosting the contrast and separation of noise from these pictures. These images are taken from sources of different conditions. The collected images may have poor contrast and noise on it. There may be also no uniformity of illumination. If there is noise in the raw data of the eye image it can predict wrongly. Hence, pre-processing is performed for any detection system. After the step of pre-processing the step of pupil segmentation is performed to achieve important information about abnormalities in the region of interest. The extraction of many features is done from segmented images. These features can be used further in training Artificial intelligence (AI) models in Deep learning and Machine learning.

2. LITERATURE SURVEY

The eye disease Cataract is an eye disorder that causes unclear focal points or cloudiness on the lens. Eye lenses are covered with ailment and it is created when a few of the protein cells at the focal point are joined together. This prevents light from going plainly through it and results causing loss of vision [01–03]. Cataracts are differentiated into three types normal cataracts, sub capsular cataracts, and cortical cataracts [5]. Cataracts generally diminish the vision. The cloudiness develops downside in the focal zone area. Generally, it occurs at the area of the focal point [6,7]. In old age people, the possibility of occurrence is high. The reason behind cortical cataracts is the dark and dirtiness of the external edge of the focal point. It generally happens when there is a variation in the internal water content of the focal point causing a split & cracks there [8,9]. Posterior sub-capsular cataract occurrence possibility at the rear side of the lens. Guo et al. [2] showed great interest in developing a healthcare system. so the author preferred CAD system design for cataract classification. The

fundus images were used for grading the images. The author suggested pre-processing which includes fundus image analysis. Then after feature extraction was done. DWT with DCT was applied for feature extraction Further the process of cataract classification was done and then grading and classification were done. Non-cataract fundus eye image having clearer organ details of blood cells of veins and the area of the fovea. However, the cataract fundus images don't show the blood veins and fovea For classification and different grading the feature related to frequency component edge with sudden peaks is used. To achieve good efficiency Haar Wavelet transform can be used for feature extraction. To make different grading PCA was used. The sketch-based approach was used to calculate the high-frequency component. The sketch method is used for feature selection by using DCT and further classification and layer grading of cataracts. Acharya et al. [4] mainly focus on cataract detection by using AI techniques. Here the author applied fuzzy k means clustering algorithms which is applied on raw optical cataract eye images to analysis all features of the image. The back propagation algorithms (BPT) are applied to the analysis and classification of cataracts. Zheng et al. [11] proposed the fundus image analysis to classify the cataract. The first step to classify the cataract was pre-processing in which resizing was done. Further, the green channel extraction was done because it provides the main details of eye images. To obtain spectrum DFT was performed on fundus image. PCA was applied to reduce the dimension. LDA (linear discriminant analysis) was performed to classify into four classes. To promote the classifier ADA-Boost algorithms were applied. Kolhe and Guru [12] suggested the algorithm for fundus image analysis. Here mainly four steps were adopted. Firstly, pre-processing was done using CLAHE algorithms. Secondly for image feature extraction DWT and DCT were applied to find the coefficients and after that, PCA (principal component analysis) was performed on those coefficients. In the third step, a binary SVM classifier was used on the extracted feature for classification. In the fourth step grading process was implemented using multi-class Fisher discriminant analysis and the cataract classification was performed in the category of mild, moderate, and severe cataracts. Zhang et al. [13] calculated the performance and efficiency of DCNN which is a weight sharing in the structure of the network. Here the author has decided to detect and classify grade cataracts. The DCNN is designed for multilayer perception for recognizing eye images. Here the initial five layers are used in the convolution layer and three layers for connecting layers so a total of eight layers are used. Pathak and Kumar [14] have tried to analyse the eyeball. The main features of the eyeball image uniformity and intensity are extracted. These features have applications in telemedicine to diagnose cataracts in the eye. Zisimopoulos et al. [15]

analysed the surgical based on surgical data. Here the author tried for the first time to train a deep learning model. In this model, all surgical data images were used in the model by applying ANN. To increase the speed of recognition with classification deep learning technique was applied. The author applied training tools to detect and segmentation which is based on CNN. It was the first attempt to make a low-cost and very easily accessible test system. The system simulated to use cataract data set to train the model. Caixinha et al. [16] proposed the detection of cataracts in animal models by the application of machine learning early detection and grading mechanism was applied. The author used SVM and Bayes classifier and random forest classification techniques. The SVM classifier showed qualitative high performance other than in the detection and classification of cataracts. Nayak [17] suggested using an SVM classifier to analyse and diagnose the cataracts in the eye by analysis them. Further to classify them into normal eye, cataract eye, and post cataract eye image the feature extraction was done. Here main features of the eye image are BRA (big ring area), and SRA (small ring area), and their perimeter with edge pixel count were extracted. Here SVM classifiers are used manually to focus on feature extraction of the classified optical image. There is the main requirement to improve the correctness of the system and it can be improved by good quality training data.Yanwu xu et al.[18] the author used a large database of images and tried to introduce a feature for grading in linear regression method with a group scarcity-based constraint and the performance is feature extraction parameter analysis and regression model and data training simultaneously. Nowadays in latest methods are to determine cataract severity based on manual assessment. Here more better effort is applied to cataract analysis by slit lamp images. As a result, the accuracy achieved is less. So there is a requirement of applying other specific concepts and algorithms to produce higher efficiency. Gonzalez et al.[19] have suggested that fundus images of the eye are used for the detection and calculated assessment of the retinal disorder. The decision depends upon the grading of opacity for the detection of abnormalities. Here author suggested a technique for adaptive histogram equalization. This technique can be helpful in better realization of different levels of blurring due to different forms of cataracts. The three colon spaces CIELAB, RGB, and HSV are analyzed.here best efforts are made by the author to analyse fundus image characteristics. Rameshbabu [20] suggested that in image processing edge detection is more important as edge detection is an important part of segmentation which can be locality-based and region based approach boundary approach edge approach is useful. It can be easily applied in pupil segmentation in our proposed

algorithms. There is a huge application of segmentation to identify the eye image at various stages of cataracts. Jindal et al. [21] have suggested two approaches the first approach is based on feature extraction-based algorithms. Histograms are plotted to classify the severity of the eye. The second is the area approach and the application of the Hough transform is performed then the contour analysis. Then percentage of cataracts is calculated. There is a requirement for a large database for good accuracy in results. Imran et al.[22] has suggested the grading-based retina image analysis.in prepossessing channel extraction histogram equalization, hat transform is used to make the retinal quality of the image. The application of neural networks by taking the radial basis function accompanied by the self-organizing map is done to detect cataracts. Jagadale et al. [23] applied slit lamp images with a (CAD) computer-aided processor to detect cataracts at an earlier stage. The steps used here are lens detection and segmentation and using the Hough transform by support vector machine application. This technique gives 90.25% accuracy. So chances are having the closeness of cataract grade. Li et al.[24] has tried to analyse fundus images by using deep learning by applying a neural network and output is calculated by using a heat map. The heat map identifies the area where the most cataracts can lie. so this paper mainly focuses on cataract localized areas. Hossain et al. [25] applied the Deep convolution neural network algorithm to fundus images of the retina. The DCNN has two parts training and testing with Retinal Fundus Images. The proposed method is also incorporated with IoT devices. However, the system does not detect mild and partial cataracts. Madhuri et al.[26] suggested the MATLAB toolbox for applying the different operations in digital image processing. This paper explains a method for growing MATLAB script for 2D scanning, multiresolution, different arithmetic, logical, and relation operators are performed in MATLAB. Using image operations like RGB to GRAY conversion, Image crop, Image rotation, image Histogram, Image Convolution, Image DCT, and Image IDCT can be applied for different image operations. Dixit et al.[30] described an interesting approach to shorten the time of processing with addition firstly iris localization was done then after analysing the pupil and extracting the corner edge by using the fuzzy method which provides the effective edge. It applies to circular regions more accurately but it will show less accuracy when the region is not circular because the fuzzy system will not be applied accurately. Jing ran et al.[31] the application of random forest with deep convolution neural network for grading the cataract at different levels. The author tried to increase the grading level to get more accuracy level. but the highest level of grading is not defined at which the accuracy will be highest so it is the disadvantage.Azhar Imran et al.[32]

here is the description of the segmentation technique that is applied to retina images. The AI techniques are combined to analyse the retinal images where segmentation of the retina proceeds. but it does not finally calculate the cataract in the eye for these images.Wena et al.[33] explained the application of semi super vised method of learning which is particularly applied to fundus images of eye cataracts. By using signal binary classifiers having different levels of accuracy which is low.Yue zhou et al.[34] has tried to upgrade the Haar features and applied discrete neural networks having the advantage of less storage memory due to discrete state transitions in multilayer perceptions. Veni et al.[35] to test image convolution neural network model is used and has tried to maximize efficiency to use higher epoch value. The higher epoch value turns into higher efficiency.

3. METHODOLOGY

In the Fig 2, the flow chart of cataract detection is shown. All the steps are shown and followed accordingly. All functions for operations are performed in MATLAB. The MATLAB toolbox has huge applications in digital image processing. For the algorithm of automatic detection of eye disease cataracts from a digital eye image, First of all, the Eye image is browsed or insertion into a digital Eye image.
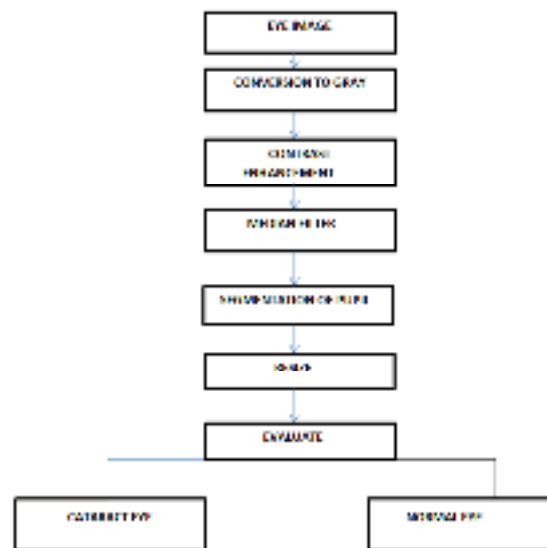

Figure 2. Flowchart of cataract detection process and steps

Then the image is converted into a gray image. The process of RGB to GRAY conversion not only converts the image into GRAY but also the elimination of hue is done. It also maintains the luminance. The next step is to follow here image enhancement for not only highlighting the information of the eye image but also it remove and weak unwanted information. In the

next step, the use of the median filter is adopted to remove noise from images in the above-proposed algorithms. This step is used before the segmentation of the pupil of after that we follow the step of image enhancement. here the process of highlighting is done. it also removes or weakens unwanted information. Now next step defined in the proposed algorithms is the application of the median filter. All the tools are used from Matlab. The software has several functions to apply in applications. The images may be added with noise. So for removing noise median filter is used. The segmentation process focuses on the cataract part of the eye. The basic step of image segmentation is the edge detection. The edges for finding edges means to identify the boundaries of the edges. The different segmentation is classified on locality region thresholding ad continuity. The region-based segmentation is the region boundary and edge approach. Further in above mentioned algorithms resizing of the image is done. After resizing evaluation can be done. After the evaluation process classified cataract was found and no cataract was found or normal eye is done. Here we are taking Eye images noise-free. There may be noise present which is removed by noising the image. This can also be done by using a Gaussian filter. The Gaussian filter may be used for suppressing frequency components reducing noise and blurring the region of an image.

4. RESULTS & DISCUSSION

The above-mentioned algorithms are applied in the Digital image processing toolbox provided by MATLAB Software. In Fig 3.1 and Fig3.2 digital eye image is taken in image a. Then it is converted into image b by using RGB to GRAY conversion. After the image conversion contrast image c is obtained. Then it is further passed through the median filter to get image d. The area for segmentation of the pupil is identified in image e. Further in image f segmented pupil area is shown. The image f is further resized to get image g. The enlarged cataract area is shown in image h.
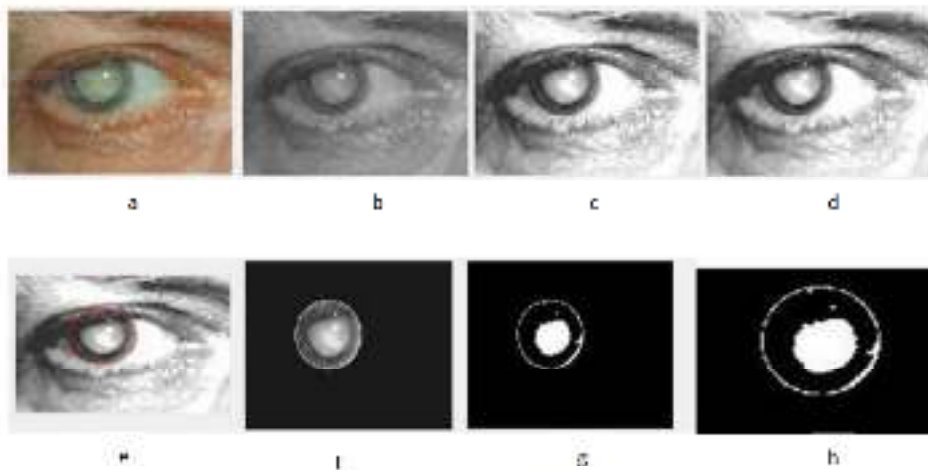
Fig3.1:(a)Eye image;(b)Gray image;(c)Contrast image;(d)Median Filter;(e)Pupil segmentation;(f)Segmented pupil;(g)Resize area;(h)Cataract area of Eye
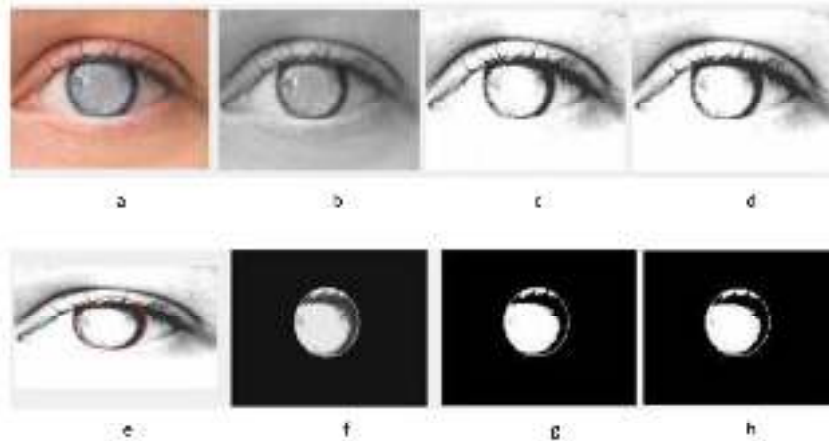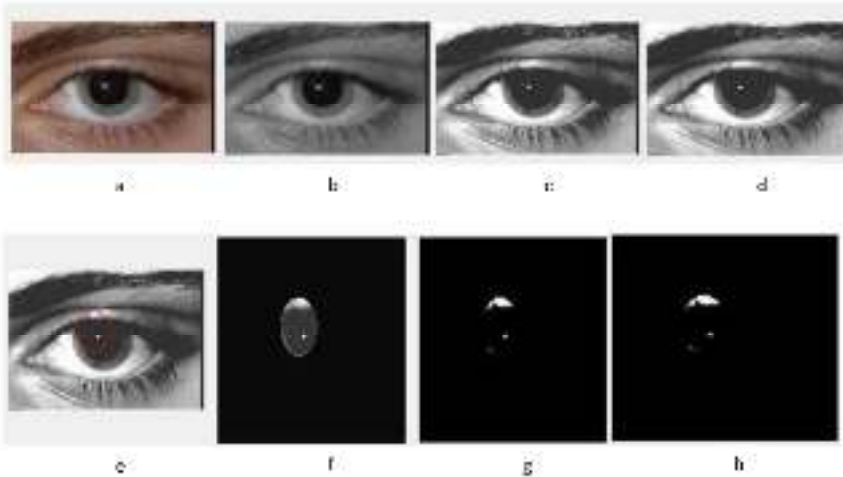
Fig3.2:[a]Eye image;(b)Gray image;[c]Contrast image;[d]Median Filter;(e)Pupil segmentation;
(f)Segmented pupil;[g]Resize area;[h]Cataract area of Eye

Fig3.3:[a]Eye image;(b)Gray image;[c]Contrast image;[d]Median Filter;[e]Pupil segmentation;
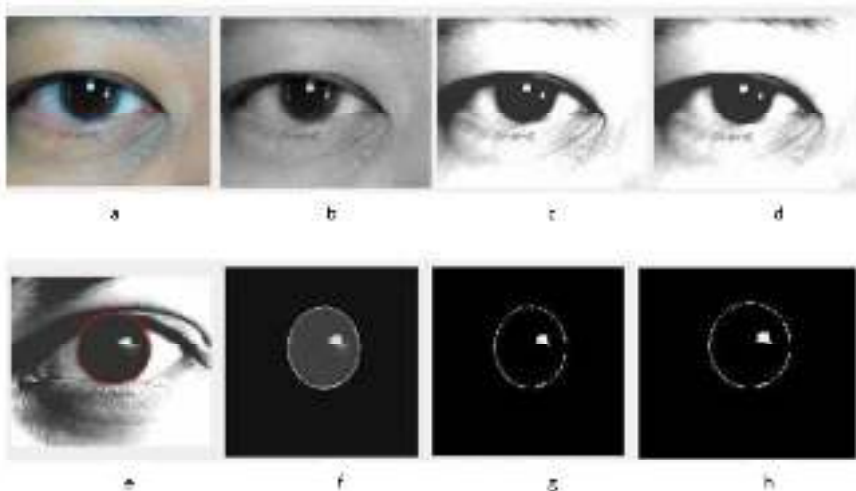(f)Segmented pupil;[g]Resize area;[h]Normal area of Eye

Fig3.4:[a]Eye image;[b]Gray image;[c]Contrast image;[d]Median Filter;[e]Pupil segmentation;
(f)Segmented pupil;[g]Resize area;[h]Normal area of Eye

The above-mentioned algorithms are applied in the Digital image processing toolbox provided by MATLAB Software. In Fig 3.3 and Fig3.4 digital eye image is taken in image a. Then it is converted into image b by using RGB to GRAY conversion. After the image conversion contrast image c is obtained. Then it is further passed through the median filter to get image d. The area for segmentation of the pupil is identified in image e. Further in image f segmented pupil area is shown. The image f is further resized to get image g. The enlarged cataract area is shown in image h. It is required to be a cost-effective method. So there is a need to design advanced and more structured methods for cataract examination and cure with the possibility of early-stage prevention, There is a need to reduce the manual pre-processing. There should be more efforts to make the algorithms focused and highly automated to increase accuracy. In the future, Many types of image processing tools may be used on the eye image. The system can be designed so that we can detect and analyze cataracts at home and then immediately consult to doctor for further treatment.

5. CONCLUSION

This work presented here with an algorithm to detect and analyze the features of cataract eye images, based on the above method the tool of Digital Image Processing in MATLAB is used. The proposed method of detection may be performed on real eye images which can be obtained from eye hospitals or eye clinics and also real images obtained from Eye images obtained from Eye specialist doctors. In this method, an eye image is taken then firstly it is converted to gray. Furthermore, the contrast enhancement of the image is done. To suppress noise from the image. Further segmented area of the pupil is identified. Further resizing and analysis show the position which is normal eye or cataract eye image. This proposed method can work well on real-time images. There can be application on non-circular pupils. it may be proven a more structured detection technique. Good attempts are made to make accurate results and for the detection of advance stage cataracts.

REFERENCES

[1] R. U. Acharya, W. Yu, K. Zhu, J. Nayak, T. C. Lim, and J. Y. Chan, "*Identification of cataract and post-cataract surgery optical images using artificial intelligence techniques,*" J. Med. Syst., vol. 34, no. 4, pp. 619–628, 2010.

[2] L. Guo, J. J. Yang, L. Peng, J. Li, and Q. Liang, "*A computer-aided healthcare system for cataract classification and grading based on fundus image analysis,*" Comput. Ind., vol. 69, no. May, pp. 72–80, 2015.

[3] H. Li et al., "*Lens image registration for cataract detection,*" Proc. 2011 6th IEEE Conf. Ind. Electron. Appl. ICIEA 2011, pp. 132–135, 2011.

[4] Rajendra udyavara acharya, wenwei yu,kuanyi zhu,jagadish nayak,teik-cheng lim,joey yiptong chan"*Idendification of Cataract and Post-cataract Surgery Optical Images Using Artificial Intelligence Techniques*"Springer Science & Business media,LLC 2009.

[5] H. Shen, H. Hao, L. Wei, and Z. Wang, "*An Image Based Classification Method for Cataract" 2008 Int. Symp. Comput. Sci. Comput. Technol.*, pp. 583–586, 2008

[6] tien yin wong li Huiqi, Lim, J. H., paul mitchell, ava grace tan, jie jin wang, "*A Computer- Aided Diagnostic System of Neuclear Cataract.".*

[7] R. Srivastava et al., "*Automatic nuclear cataract grading using image gradients,*" J. Med. Imaging, vol. 1, no. 1, 2014.

[8] T. Kuroda, T. Fujikado, N. Maeda, T. Oshika, Y. Hirohara, and T. Mihashi, "*Wavefront analysis in eyes with nuclear or cortical cataract,*" Am. J. Ophthalmol., vol. 134, no. 1, pp. 1–9, 2002.

[9] C. J. Hammond et al., "*The heritability of age-related cortical cataract: The twin eye study,*" Investig. Ophthalmol. Vis. Sci., vol. 42, no. 3, pp. 601–605, 2001

[10] A. U. Patwari, "*Detection, Categorization, and Assessment of Eye Cataracts Using Digital Image Processing,*" no. June, pp. 1–5, 2011.

[11] J. Zheng, L. Guo, L. Peng, J. Li, J. Yang, and Q. Liang, "*Fundus image based cataract classification,*" IST 2014 - 2014 IEEE Int. Conf. Imaging Syst. Tech. Proc., pp. 90–94, 2014.

[12] S. Kolhe and S. K. Guru, "*Remote Automated Cataract Detection System Based on*

Fundus Images," pp. 10334–10341, 2016.

[13] L. Zhang et al., "*Automatic cataract detection and grading using Deep Convolutional Neural Network*," Proc. 2017 IEEE 14th Int. Conf. Networking, Sens. Control. ICNSC 2017, pp. 60– 65, 2017

[14] S. Pathak and B. Kumar, "*A Robust Automated Cataract Detection Algorithm Using Diagnos- tic Opinion Based Parameter Thresholding for Telemedicine Application*," Electronics, vol. 5, no. 3, p. 57, 2016

[15] O. Zisimopoulos et al., "*Can surgical simulation be used to train detection and classification of neural networks?*," Healthc. Technol. Lett., vol. 4, no. 5, pp. 216– 222, 2017.

[16] M. Caixinha, J. Amaro, M. Santos, F. Perdigão, M. Gomes, and J. Santos, "*In-Vivo Automatic Nuclear Cataract Detection and Classification in an Animal Model by Ultrasounds*," IEEE Trans. Biomed. Eng., vol. 63, no. 11, pp. 2326–2335, 2016.

[17] Jagdish Nayak,"*Automated Classification of Normal ,Cataract and Post Cataract Optical Eye Images using SVM Classifier*"Proceedings of the world Congress on engineering and computer science "2013 vol 1.

[18] Yanwu Xu ,Xinting Gao,Stephen Lin,Damon Wing Kee Wong,Jiang Liu,Dong Xu,Ching-yu,Cheng,CarolYCheung,Tien,YinWong"Aotomatic *Grading of nuclear cataract from slit lamp lens images using group sparsity regression*"Conferencepaper2013.

[19] Enrique Gonzalez-Amador,Justo Arines,Pablo Charlon,Nery Garcia-Porta,Maximino J Abraldes and Eva Acosta "*Improvement of Retinal Images Affected by Cataracts*," Photonics9040251,Photonics 2022.

[20] Dr K.Rameshbabu, Missay.Mangesthu, "*Edge Detection with Theory and Soft Reviews*,"IRJET Volume:06 Issue 02|Feb 2019.

[21] Ishitaa Jindal, Palak Gupta,Anmolika Goyal "*Cataract Detetion using Digital Image Processing*",GCAT Banglore India Oct 18-20,2019.

[22] Azhar Imran,Jianqiang Li,Yan Pei,Faheem Akhtar,Ji-Jiang Yang Qing Wang"*Cataract Detection and Grading with Retinal Images Using SOM-RBF Neural Network*"IEEE Symposium Series on Computational Intelligence Xiamen China DEC 6-9 2019.

[23] A.B.Jagadale, S.S.Sonavane, D.V.Jadav, "*Computer Aided System for Early Detection of Nuclear Cataract Using Circle Hough Transform*".IOCOEI 2019.

[24]   Jianqiang Li,Xi Xu,Yu Guan ,Azhar Imran ,Bo Liu,Li Zhang,Ji-Jiang Yang,Qing Wang,Liyang Xie, "*Automatic Cataract Diagnosis by Image-Based Interpretability*"IEEE international Conference on system,Man and Cybernetics 2018.

[25]   Md. Rajib Hossain,Sadia Afroze,Nazmul Siddque,Mohammed Moshiul Hooque , "*Automatic Detection of Eye Cataract using Deep Convolution Neural Networks(DCNNs)*" Symposium(TENSYMP) Dhaka Bangladesh 5-7 june 2020.

[26]   Dr.Sindhu madhuri G Prof.Shashikala H,K Dr Jayanthi Kannan "*Review on the Use of MATLAB for Digital Image Processing*" JETIR Volume 6,Issue 2, 2019.

[27]   R. R. A. Bourne et al., "*Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment*: a systematic review and meta-analysis," Lancet Glob. Heal., vol. 5, no. 9, pp. e888–e897, 2017

[28]   T. R. Fricke et al., "*Global Prevalence of Presbyopia and Vision Impairment from Uncorrected Presbyopia: Systematic Review, Meta-analysis, and Modelling,*" Ophthalmology, vol. 125, no. 10, pp. 1492–1499, 2018.

[29]   S.kumar,S.Pathak,B.Kumar ,"*Automated Detection of Eye Related Diseases using Digital Image Processing* "Springer Nature Switzerland AG 2019.

[30]   Aditya Dixit,Shashawat Pathak,Rahul Raj,Ch Naveen ,V.R Satpure" *An efficient Fuzzy Based edge estimation for oris localization and pupil detection in human eye for automated cataract dectction system*" 9th ICCNT IISC Banglore

[31]   Jing Ran,Kai Niu,Zhigiang He, Hong yan zhang ,Hoxgxin jong,"*Cataract detection and grading based on combination of deep convolution neural network and random forest*"Proceedings of IC nidc 2018.

[32]   Azhar Imran ,Jianqiang Li YAN pei ,JI-JIANG YANG,QING WANG "*Comaparative analysis of vessal segmentation technique in retinal images*"preparation of papers for IEEE transaction and journals volume 4,2016.

[33]   Wwna ,jung,xing cao Zhiqiang qiau wang Ji-Jiang yang "*An improved semi supervised learning method on cataract fundus image classification*"2019 IEEE 43rd annual computer software and application conferences COMPSAC 2019

[34]   Yue Zhou,Guoqi Li,Hiuqi Li "*Automatic cataract classification using deep neural network with discrete state transform*"0278-0062(c)2019

[35]   Indra veni,pradita Eko prasetyo utomo Benedita ferdian hutabarct ,Muksin Alfalah "*Detection of cataract based on image featured using convolution neural network* "IJCCS (Indonesian journal of computing and cybernrtics sysrems) Vol 15 no 1

January 2021.

[36]   Gary, B.; Taylor, H. *Cataract Blindness–Challenges for 21st century*. Bull. World Health Organ.2001

[37]   Supriyanti, R.; Habe, H.; Kidode, M.; Nagata, S. "*A simple and robust method to screen cataract using specular reflection appearance*" In Proceedings of the Medical Imaging International Conference of International Society for Optics and Photonics (SPIE), San Diego, CA, USA, 17 March 2008.

[38]   Supriyanti, R.; Habe, H.; Kidode, M.; Nagata, S. "*Cataract Screening by Specular Reflection and Texture Analysis*" In Proceedings of the Systemics and Informatics World Network (SIWN 2009), Leipzig, Germany, 23–25 March 2009. ISSN 1757-4439.

[39]   Supriyanti, R.; Habe, H.; Kidode, M.; Nagata, S" *Compact cataract screening system: Design and practical data acquisition*" In Proceedings of International Conference on Instrumentation, Communication, Information Technology and Biomedical Engineering (ICICI-BME), Bandung, Indonesia, 23–25 November 2009. doi:10.1109/ICICI-BME.2009.5417287.

[40]   Shashwat Pathak, Basant Kumar, "*A Robust Automated Cataract Detection Algorithm Using Diagnostic Opinion Based Parameter Thresholding for Telemedicine Application*", Journal of Electronics, MDPI, published online , 15 September 2016, DOI: 10.3390/electronics5030057,

[41]   07916648] P. Khelbude and S. Shelke, "*Real-time iris controlled robot*," IEEE Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, 2016, pp.1-4.

[42]   Comas, O; Cotin, S.; Duriez, C." *A shell model for real time simulation of intra-ocular implant deployment*" Proceedings of the International Symposium on Biomedical Simulation, Phoenix, AZ, USA, 23–24 January 2010; pp. 160–170.

[43]   S R Flaxman, R Bourne, S Resnikoff, et al. "*Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis*" The Lancet Global Health, 2017, 5(12): e1221-e1234.

[44]   V Harini, V Bhanumathi. "*Automatic cataract classification system. Communication and Signal Processing* "(ICCSP), 2016 International Conference on. IEEE, 2016: 0815-0819.

[45]   S Aslani, H Sarnel. "*A new supervised retinal vessel segmentation method based on robust hybrid features.*" Biomedical Signal Processing and Control, 2016.

[46]   S Willikens, E Zitron, E Scholz, et al. "*Retinal Arterio-Venule-Ratio (AVR) in the cardiovascular risk management of hypertension''* European Heart Journal, 2013, 34(suppl_1).

[47]   PC Siddalingaswamy, GK Prabhu. "*Automated detection of anatomical structures in retinal images.*"Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on. IEEE,2007, 3: 164-168.

[48]   M Yang, JJ Yang, Q Zhang, et al. "*Classification of retinal image for automatic cataract detection. E-health Networking, Applications & Services*" (Healthcom), 2013 IEEE 15th International Conference on. IEEE,2013: 674-679.

[49]   JJ Yang. J Li, R Shen, et al. "*Exploiting ensemble learning for automatic cataract detection and grading*" Computer methods and programs in biomedicine, 2016, 124: 45-57.

[50]   Z Qiao, Q Zhang, Y Dong, et al." *Application of SVM based on genetic algorithm in classification of cataract fundus images. Imaging Systems and Techniques (IST)*",2017 IEEE International Conference on. IEEE, 2017: 1-5.

[51]   Y Dong, Q Zhang, Z Qiao, et al." *Classification of cataract fundus image based on deep learning. Imaging Systems and Techniques (IST)*", 2017 IEEE International Conference on. IEEE, 2017: 1-5.

[52]   L Zhang, J Li, H Han, et al. "*Automatic cataract detection and grading using Deep Convolutional Neural Network. Networking, Sensing and Control*" (ICNSC), 2017 IEEE 14th International Conference on. IEEE,2017: 60-65.

[53]   K He, X Zhang, S Ren, et al. "*Deep residual learning for image recognition*". Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[54]   D. Pascolini and S. P. Mariotti, "*Global estimates of visual impairment*: 2010," British Journal of Ophthalmology, vol. 96, no. 5, pp. 614–618, 2012.

[55]   P. Mitchell, R. G. Cumming, K. Attebo, and J. Panchapakesan, "*Preva- lence of cataract in australia: the blue mountains eye study,*" Ophthalmol- ogy, vol. 104, no. 4, pp. 581–588, 1997.

[56]   N. Congdon, J. Vingerling, B. Klein, S. West, D. Friedman, J. Kem- pen, B. O'Colmain, S. Wu, and H. Taylor, "*Prevalence of cataract and*

pseudophakia/aphakia among adults in the united states." Archives of ophthalmology (Chicago, Ill.: 1960), vol. 122, no. 4, pp. 487–494, 2004.

[57]   J. J. Kanski and A. Kubicka-Trzáska, "*Clinical ophthalmology: a self- assessment companion",* 1st ed. Edinburgh, New York: Elsevier Churchill Livingstone, 2007

[58]   J.-J. Yang, J. Li, R. Shen, Y. Zeng, J. He, J. Bi, Y. Li, Q. Zhang, L. Peng, and Q. Wang, "*Exploiting ensemble learning for automatic cataract detection and grading,*" Computer methods and programs in biomedicine, vol. 124, pp. 45–57, 2016.

[59]   J. A. Mobley and R. W. Brueggemeier, "*Increasing the dna damage threshold in breast cancer cells,*" Toxicology and applied pharmacology, vol. 180, no. 3, pp. 219–226, 2002

[60]   B. E. K. Klein, R. Klein, K. L. P. Linton, Y. L. Magli, and M. W. Neider, "*Assessment of cataracts from photographs in the beaver dam eye study,*" Ophthalmology, vol. 97, no. 11, pp. 1428–1433, 1990.

[61]   E. Ricci and R. Perfetti, "*Retinal blood vessel segmentation using line op- erators and support vector classification,*" IEEE transactions on medical imaging, vol. 26, no. 10, pp. 1357–1365, 2007.

[62]   A. Osareh and B. Shadgar, "*Automatic blood vessel segmentation in color images of retina,*" Iranian Journal of Science and Technology, vol. 33, no. 2, pp. 191–206, 2009

[63]   L. Xu and S. Luo, "*A novel method for blood vessel detection from retinal images," Biomedical engineering online*, vol. 9, no. 1, p. 14, 2010.

[64]   X. You, Q. Peng, Y. Yuan, Y.-m. Cheung, and J. Lei, "*Segmentation of retinal blood vessels using the radial projection and semi-supervised approach,*" Pattern Recognition, vol. 44, no. 10-11, pp. 2314–2324, 2011.

[65]   E. Tuba, L. Mrkela, and M. Tuba, "*Retinal blood vessel segmentation by support vector machine classification,*" in Radioelektronika (RA- DIOELEKTRONIKA), 2017 27th International Conference.   IEEE, 2017, pp. 1–6.

[66]   Y.-B. Wang, C.-Z. Zhu, Q.-F. Yan, and L.-Q. Liu, "*A novel vessel segmentation in fundus images based on svm,*" in 2016 International Conference on Information System and Artificial Intelligence (ISAI). IEEE, 2016, pp. 390–394

[67]   A. W. Awan, Z. W. Awan, and M. U. Akram, "*A robust algorithm for segmentation of blood vessels in the presence of lesions in retinal fundus images,*" in Imaging Systems and Techniques (IST), 2015 IEEE International Conference on. IEEE, 2015,

pp. 1–6.

[68] J. I. Orlando and M. Blaschko, "*Learning fully-connected crfs for blood vessel segmentation in retinal images,*" in International Confer- ence on Medical Image Computing and Computer-Assisted Intervention. Springer, 2014, pp. 634–641

[69] D. Selvathi and P. L. Vaishnavi, "*Retinal blood vessel segmentation using gabor wavelet and support vector machine*," International journal of Neural Networks and Applications, vol. 4, no. 1, pp. 1–6, 2011.

[70] K. Akita and H. Kuga, "*A computer method of understanding ocular fundus images*," Pattern recognition, vol. 15, no. 6, pp. 431–443, 1982

[71] Q. Jin, Z. Meng, T. D. Pham, Q. Chen, L. Wei, and R. Su, "Dunet: *A deformable network for retinal vessel segmentation," Knowledge-Based Systems*, vol. 178, pp. 149–162, 2019

[72] D. Yang, M. Ren, and B. Xu, "*Retinal blood vessel segmentation with improved convolutional neural networks*," Journal of Medical Imaging and Health Informatics, vol. 9, no. 6, pp. 1112–1118, 2019.

[73] S. Guo, K. Wang, H. Kang, Y. Zhang, Y. Gao, and T. Li, "Bts-dsn: *Deeply supervised neural network with short connections for retinal vessel segmentation*," International journal of medical informatics, vol. 126, pp. 105–113, 2019.

[74] J. Son, S. J. Park, and K.-H. Jung, "*Towards accurate segmentation of retinal vessels and the optic disc in fundoscopic images with generative adversarial networks*," Journal of digital imaging, vol. 32, no. 3, pp. 499– 512, 2019

[75] K. Hu, Z. Zhang, X. Niu, Y. Zhang, C. Cao, F. Xiao, and X. Gao, "*Retinal vessel segmentation of color fundus images using multiscale convolu- tional neural network with an improved cross-entropy loss function*," Neurocomputing, vol. 309, pp. 179–191, 2018.

[76] P. Liskowski and K. Krawiec, "*Segmenting retinal blood vessels with deep neural networks," IEEE transactions on medical imaging*, vol. 35, no. 11, pp. 2369–2380, 2016

[77] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, "*Dense u-net based on patch-based learning for retinal vessel segmentation*," Entropy, vol. 21, no. 2, p. 168, 2019.

[78] Y. Lin, H. Zhang, and G. Hu, "*Automatic retinal vessel segmentation via deeply supervised and smoothly regularized network*," IEEE Access, vol. 7, pp. 57 717–57 724, 2018

[79]   A. Oliveira, S. Pereira, and C. A. Silva, "*Retinal vessel segmentation based on fully convolutional neural networks*," Expert Systems with Applications, vol. 112, pp. 229–242, 2018

[80]   Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "*A cross- modality learning approach for vessel segmentation in retinal images*," IEEE transactions on medical imaging, vol. 35, no. 1, pp. 109–118, 2015.

[81]   Y. Guo, Ü. Budak, and A. S¸ engür, "*A novel retinal vessel detection ap- proach based on multiple deep convolution neural networks*," Computer methods and programs in biomedicine, vol. 167, pp. 43–48, 2018 .

[82]   L. Geng, L. Qiu, J. Wu, Z. Xiao, and F. Zhang, "*Segmentation of retinal image vessels based on fully convolutional network with depthwise separable convolution and channel weighting*," Sheng wu yi xue gong cheng xue za zhi= Journal of biomedical engineering= Shengwu yixue gongchengxue zazhi, vol. 36, no. 1, pp. 107–115, 2019.

[83]   D. A. Dharmawan, D. Li, B. P. Ng, and S. Rahardja, "*A new hybrid algorithm for retinal vessels segmentation on fundus images*," IEEE Access, vol. 7, pp. 41 885–41 896, 2019

[84]   Z. Yan, X. Yang, and K. Cheng, "*A three-stage deep learning model for accurate retinal vessel segmentation*," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1427–1436, July 2019

[85]   R. Xu, G. Jiang, X. Ye, and Y.-W. Chen, "*Retinal vessel segmentation via multiscaled deep-guidance*," in Pacific Rim Conference on Multimedia. Springer, 2018, pp. 158–168.

[86]   A. Hatamizadeh, H. Hosseini, Z. Liu, S. D. Schwartz, and D. Terzopou- los, "*Deep dilated convolutional nets for the automatic segmentation of retinal vessels*," arXiv preprint arXiv:1905.12120, 2019.

[87]   Z. Fan, J. Mo, and B. Qiu, "*Accurate retinal vessel segmentation via octave convolution neural network*," arXiv preprint arXiv:1906.12193, 2019.

[88]   A. Ribeiro, A. P. Lopes, and C. A. Silva, "*Ensemble learning approaches for retinal vessel segmentation*," in 2019 IEEE 6th Portuguese Meeting on Bioengineering (ENBENG). IEEE, 2019, pp. 1–4.

[89]   S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "Ccnet: *A cross-connected convolutional network for segmenting retinal vessels using multi-scale features*," Neurocomputing,

2019.

[90]   K. J. Noh, S. J. Park, and S. Lee, "*Scale-space approximated con- volutional neural networks for retinal vessel segmentation*," Computer Methods and Programs in Biomedicine, vol. 178, pp. 237–246, 2019.

[91]   T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, L. Zheng, and M. Paul, "*Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation*," Expert Systems with Applications, vol. 134, pp. 36–52, 2019.

[92]   Q. Jin, Q. Chen, Z. Meng, B. Wang, and R. Su, "*Construction of retinal vessel segmentation models based on convolutional neural network*," Neural Processing Letters, pp. 1–18, 2019.

[93]   Y. Lu, Y. Zhou, and J. Qin, "*A convolutional encoder-decoder architecture for retinal blood vessel segmentation in fundus images*," in 2018 5th International Conference on Systems and Informatics (ICSAI). IEEE, 2018, pp. 1071–1075.

[94]   H. Xia, R. Zhuge, and H. Li, "*Retinal vessel segmentation via a coarse- to-fine convolutional neural network,*" in 2018 IEEE International Con- ference on Bioinformatics and Biomedicine (BIBM). IEEE, 2018, pp. 1036–1039.

[95]   Y. Jiang, N. Tan, T. Peng, and H. Zhang, "*Retinal vessels segmentation based on dilated multi-scale convolutional neural network,*" IEEE Access, vol. 7, pp. 76 342– 76 352, 2019.

[96]   S. Thangaraj, V. Periyasamy, and R. Balaji, "*Retinal vessel segmentation using neural network*," IET Image Processing, vol. 12, no. 5, pp. 669– 678, 2017.

[97]   Z. Feng, J. Yang, and L. Yao, "*Patch-based fully convolutional neural network with skip connections for retinal blood vessel segmentation*," in Image Processing (ICIP), 2017 IEEE International Conference on. IEEE, 2017, pp. 1742–1746.

[98]   T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, M. A. Khan, M. Paul, and L. Zheng, "*Boosting sensitivity of a retinal vessel segmentation algo- rithm with convolutional neural network*," in Digital Image Computing: Techniques and Applications (DICTA), 2017 International Conference on. IEEE, 2017, pp. 1–8.

[99]   J. Song and B. Lee, "*Development of automatic retinal vessel segmen- tation method in fundus images via convolutional neural networks*," in Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE. IEEE, 2017, pp. 681– 684.

[100]   L. Ngo and J.-H. Han, "*Multi-level deep neural network for efficient segmentation of*

blood vessels in fundus images,” Electronics Letters, vol. 53, no. 16, pp. 1096–1098, 2017.

[101] A. Dasgupta and S. Singh, “*A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation*,” in Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Sym- posium on. IEEE, 2017, pp. 248–251.

[102] Z. Fan and J.-J. Mo, “*Automated blood vessel segmentation based on de-noising auto-encoder and neural network*,” in Machine Learning and Cybernetics (ICMLC), 2016 International Conference on, vol. 2. IEEE, 2016, pp. 849–856.

[103] Z. Yao, Z. Zhang, and L.-Q. Xu, “*Convolutional neural network for retinal blood vessel segmentation*,” in Computational Intelligence and Design (ISCID), 2016 9th International Symposium on, vol. 1. IEEE, 2016, pp. 406–409.

[104] Y. Luo, H. Cheng, and L. Yang, “Size-invariant fully convolutional neural network for vessel segmentation of digital retinal images,” in Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2016 Asia-Pacific. IEEE, 2016, pp. 1–7.

[105] A. F. Khalaf, I. A. Yassine, and A. S. Fahmy, “*Convolutional neural net- works for deep feature learning in retinal vessel segmentation*,” in ImageProcessing (ICIP), 2016 IEEE International Conference on. IEEE, 2016, pp. 385–388.

[106] M. Nandy and M. Banerjee, “*Retinal vessel segmentation using gabor filter and artificial neural network*,” in Emerging Applications of Informa- tion Technology (EAIT), 2012 Third International Conference on. IEEE, 2012, pp. 157–160.

[107] R. Ghaderi, H. Hassanpour, and M. *Shahiri, “Retinal vessel segmentation using the 2-d morlet wavelet and neural network*,” in Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on. IEEE, 2007, pp. 1251–1255.

[108] A. S¸engür, Y. Guo, Ü. Budak, and L. J. Vespa, “*A retinal vessel detection approach using convolution neural network*,” in Artificial Intelligence and Data Processing Symposium (IDAP), 2017 International. IEEE, 2017, pp. 1–4.

[109] M. Niemeijer, J. Staal, B. van Ginnek en, M. Loog, and M. D. Abramoff, “*Comparative study of retinal vessel segmentation methods on a new publicly available database*,” in Medical Imaging 2004: Image Process- ing, vol. 5370. International Society for Optics and Photonics, 2004, pp. 648–657.

[110] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, “*Retinal vessel*

segmentation using the 2-d gabor wavelet and supervised classification," IEEE Transactions on medical Imaging, vol. 25, no. 9, pp. 1214–1222, 2006.

[111] C. A. Lupascu, D. Tegolo, and E. Trucco, "*Fabc: retinal vessel segmen- tation using adaboost*," IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 5, pp. 1267–1274, 2010.

[112] N. Memari, A. R. Ramli, M. I. B. Saripan, S. Mashohor, and M. Moghbel, "*Supervised retinal vessel segmentation from color fundus images based on matched filtering and adaboost classifier*," PloS one, vol. 12, no. 12, p. e0188939, 2017.

[113] J. A. Richards, "*Remote sensing digital image analysis: an introduction*" 4th ed. Berlin, Heidelberg: Springer-Verlag, 1999.

[114] U. T. Nguyen, A. Bhuiyan, L. A. Park, and K. Ramamohanarao, "*An effective retinal blood vessel segmentation method using multi-scale line detection*," Pattern recognition, vol. 46, no. 3, pp. 703–715, 2013.

[115] M. Sofka and C. V. Stewart, "*Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures*," IEEE trans- actions on medical imaging, vol. 25, no. 12, pp. 1531–1546, 2006.

[116] D. Kumar, A. Pramanik, S. S. Kar, and S. P. Maity, "*Retinal blood vessel segmentation using matched filter and laplacian of gaussian*," in Signal Processing and Communications (SPCOM), 2016 International Conference on. IEEE, 2016, pp. 1–5.

[117] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," IEEE Transactions on medical imaging, vol. 8, no. 3, pp. 263–269, 1989.

[118] L. Gang, O. Chutatape, and S. M. Krishnan, "*Detection and measurement of retinal vessels in fundus images using amplitude modified second- order gaussian filter*," IEEE transactions on Biomedical Engineering, vol. 49, no. 2, pp. 168–172, 2002.

[119] Z. Yavuz and C. Köse, "*Comparing 2d matched filter response and gabor filter methods for vessel segmentation in retinal images*," in Electrical, Electronics and Computer Engineering (ELECO), 2010 National Confer- ence on. IEEE, 2010, pp. 648–652.

[120] D. A. Dharmawan and B. P. Ng, "*A new two-dimensional matched filter based on the modified chebyshev type i function for retinal vessels detection*," in Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference

of the IEEE. IEEE, 2017, pp. 369–372.

[121] X. Gao, Y. Cai, C. Qiu, and Y. Cui, "*Retinal blood vessel segmentation based on the gaussian matched filter and u-net*," in Image and Signal Pro- cessing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress on. IEEE, 2017, pp. 1–5.

[122] B. Zhang, L. Zhang, L. Zhang, and F. Karray, "*Retinal vessel extraction by matched filter with first-order derivative of gaussian*," Computers in biology and medicine, vol. 40, no. 4, pp. 438–445, 2010.

[123] J. Elson, J. Precilla, P. Reshma, and N. S. Madhavaraja, "*Automated extraction and analysis of retinal blood vessels with multi scale matched filter*," in Intelligent Computing, Instrumentation and Control Technolo- gies (ICICICT), 2017 International Conference on. IEEE, 2017, pp. 775–779.

[124] I. Abdurrazaq, S. Hati, and C. Eswaran, "*Morphology approach for features extraction in retinal images for diabetic retionopathy diagnosis*," in Computer and Communication Engineering, 2008. ICCCE 2008. In- ternational Conference on. IEEE, 2008, pp. 1373–1377.

[125] A. M. Mendonca and A. Campilho, "*Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction*," IEEE transactions on medical imaging, vol. 25, no. 9, pp. 1200–1213, 2006.

[126] O. Monga, N. Armande, and P. Montesinos, "*Thin nets and crest lines: Application to satellite data and medical images*," Computer Vision and Image Understanding, vol. 67, no. 3, pp. 285–295, 1997.

[127] K. Nisha, G. Sreelekha, S. P. Savithri, P. Mohanachandran, andA. Vinekar, "*Fusion of structure adaptive filtering and mathematical morphology for vessel segmentation in fundus images of infants with retinopathy of prematurity*," in Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on. IEEE, 2017, pp. 1–6.

[128] S. J. J. Kumar and C. Ravichandran, "*Morphological operation detection of retinal image segmentation*," in 2017 International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2017, pp. 1228–1235.

[129] J. Rodrigues and N. Bezerra, "*Retinal vessel segmentation using paral- lel grayscale skeletonization algorithm and mathematical morphology*," in Graphics, Patterns and Images (SIBGRAPI), 2016 29th SIBGRAPI Conference on. IEEE, 2016, pp. 17–24.

[130] M. M. Fraz, A. Basit, P. Remagnino, A. Hoppe, and S. Barman, "*Retinal vasculature segmentation by morphological curvature, reconstruction and adapted hysteresis thresholding,*" in Emerging Technologies (ICET), 2011 7th International Conference on. IEEE, 2011, pp. 1–6.

[131] A. Sharma, M. K. Dutta, A. Singh, M. Parthasarathi, and C. M. Travieso, "*Dynamic thresholding technique for detection of hemorrhages in retinal images,*" in Contemporary Computing (IC3), 2014 Seventh International Conference on. IEEE, 2014, pp. 113–116.

[132] F. Zana and J. Klein, "*Segmentation of vessel-like patterns using math- ematical morphology and curvature evaluation,*" IEEE Transactions on Image Processing, vol. 10, no. 7, pp. 1010–1019, 2001.

[133] K. Sun, Z. Chen, S. Jiang, and Y. Wang, "*Morphological multiscale enhancement, fuzzy filter and watershed for vascular tree extraction in angiogram,*" Journal of medical systems, vol. 35, no. 5, pp. 811–824, 2011.

[134] M. S. Miri and A. Mahloojifar, "*Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction,*" IEEE Transactions on Biomedical Engineering, vol. 58, no. 5, pp. 1183– 1192, 2011.

[135] M. Fraz, M. Javed, and A. Basit, "*Evaluation of retinal vessel segmen- tation methodologies based on combination of vessel centerlines and morphological processing,*" in Emerging Technologies, 2008. ICET 2008. 4th International Conference on. IEEE, 2008, pp. 232–236.

[136] Z. Yu and K. Sun, "*Vessel segmentation on angiogram using morphology driven deformable model,*" in Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on, vol. 2. IEEE, 2010, pp. 675–678.

[137] J. Zhou, S. Chang, D. Metaxas, and L. Axel, "*Vessel boundary extraction using ridge scan-conversion deformable model,*" in Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on. IEEE, 2006, pp. 189–192.

[138] M. Tagizaheh, S. Sadri, and A. M. Doosthoseini, "*Segmentation of coronary vessels by combining the detection of centerlines and active contour model,*" in Machine Vision and Image Processing (MVIP), 2011 7th Iranian. IEEE, 2011, pp. 1–4.

[139] Y. Zhao, L. Rada, K. Chen, S. P. Harding, Y. Zheng et al., "*Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images.*" IEEE Trans. Med. Imaging, vol. 34,

no. 9, pp. 1797–1807, 2015.

[140] B. Al-Diri, A. Hunter, and D. Steel, "*An active contour model for seg- menting and measuring retinal vessels,*" IEEE Transactions on Medical imaging, vol. 28, no. 9, pp. 1488–1497, 2009.

[141] K. Sum and P. Y. Cheung, "*Vessel extraction under non-uniform il- lumination: a level set approach*," IEEE Transactions on Biomedical Engineering, vol. 55, no. 1, pp. 358–360, 2008.

[142] Y. Zhang, W. Hsu, and M. L. Lee, "*Detection of retinal blood vessels based on nonlinear projections*," Journal of Signal Processing Systems, vol. 55, no. 1-3, p. 103, 2009.

[143] C. S. Tong, Y. Zhang, and N. Zheng, "*Variational image binarization and its multi-scale realizations*," Journal of Mathematical Imaging and Vision, vol. 23, no. 2, pp. 185–198, 2005.

[144] D. Rueckert and P. Burger, "*Contour fitting using stochastic and proba- bilistic relaxation for cine mr images*," in Computer Assisted Radiology. Berlin, Gernmany: Springer, 1995, pp. 137–142.

[145] T. Zhu, "*Fourier cross-sectional profile for vessel detection on retinal images*," Computerized Medical Imaging and Graphics, vol. 34, no. 3, pp. 203–212, 2010.

[146] M. Al-Rawi, M. Qutaishat, and M. Arrar, "*An improved matched filter for blood vessel detection of digital retinal images*," Computers in Biology and Medicine, vol. 37, no. 2, pp. 262–267, 2007.

[147] H. Li, J. Zhang, Q. Nie, and L. Cheng, "*A retinal vessel tracking method based on bayesian theory*," in Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on. IEEE, 2013, pp. 232–235.

[148] K. K. Delibasis, A. I. Kechriniotis, C. Tsonos, and N. Assimakis, "*Automatic model-based tracing algorithm for vessel segmentation and diameter estimation*," Computer methods and programs in biomedicine, vol. 100, no. 2, pp. 108–122, 2010.

[149] T. Köhler, A. Budai, M. F. Kraus, J. Odstrčˇilik, G. Michelson, and J. Hornegger, "*Automatic no-reference quality assessment for retinal fundus images using vessel segmentation*," in Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on. IEEE, 2013, pp. 95–100.

[150] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay et al., "*Feedback on a publicly distributed image*

database: the messidor database," Image Analysis & Stereology, vol. 33, no. 3, pp. 231–234, 2014.

[151] B. Al-Diri, A. Hunter, D. Steel, M. Habib, T. Hudaib, and S. Berry, "*A reference data set for retinal vessel profiles*," in Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE. IEEE, 2008, pp. 2262–2265.

[152] M. Niemeijer, B. Van Ginneken, M. J. Cree, A. Mizutani, G. Quellec,C. I. Sánchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu et al., "*Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs*," IEEE transactions on medical imaging, vol. 29, no. 1, pp. 185–195, 2010.

[153] M. Thonnat, "*Knowledge-based techniques for image processing and for image understanding*," Journal de Physique 4, vol. 12, no. 1, pp. Pr1–189, 2002.

[154] M. Yang, J.-J. Yang, Q. Zhang, Y. Niu, and J. Li, "*Classification of retinal image for automatic cataract detection,*" in e-Health Networking, Applications & Services (Healthcom), 2013 IEEE 15th International Conference on. IEEE, 2013, pp. 674–679.

[155] J. Li, Q. Hu, A. Imran, L. Zhang, J.-j. Yang, and Q. Wang, "*Vessel recognition of retinal fundus images based on fully convolutional net- work*," in 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). IEEE, 2018, pp. 413–418.

[156] M. Vlachos and E. Dermatas, "*Multi-scale retinal vessel segmentation us- ing line tracking*," Computerized Medical Imaging and Graphics, vol. 34, no. 3, pp. 213–227, 2010.

[157] Y. Dong, Q. Zhang, Z. Qiao, and J.-J. Yang, "*Classification of cataract fundus image based on deep learning*," in Imaging Systems and Tech- niques (IST), 2017 IEEE International Conference on. IEEE, 2017, pp. 1–5.

[158] T. Walter, J.-C. Klein, P. Massin, and A. Erginay, "*A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina*," IEEE transactions on medical imaging, vol. 21, no. 10, pp. 1236–1243, 2002.

[159] [132] J. Dash and N. Bhoi, "*A survey on blood vessel detection methodologies in retinal images*," in 2015 International Conference on Computational Intelligence & Networks (CINE). IEEE, 2015, pp. 166–171.

[160] Zhiqiang Qiao, Qinyan Zhang, Yanyan Dong, and Ji Jiang Yang. *Application of svm based on genetic algorithm in classification of cataract fundus images*. In *IEEE International Conference on Imaging Systems and Techniques*, pages 1–5, 2017.

[161] Wenai Song, Ping Wang, Xudong Zhang, and Qing Wang. *Semi- supervised learning based on cataract classification and grading*. In *Computer Software and Applications Conference*, pages 641–646, 2016.

[162] Friedhelm Schwenker and Edmondo Trentin. "*Pattern classification and clustering: A review of partially supervised learning approaches*" *Pattern Recognition Letters*, 37(1):4–14, 2014.

[163] Yan Yan, Zhongwen Xu, Ivor W Tsang, Guodong Long, and Yi Yang. "*Robust semi-supervised learning through label aggregation*". In *AAAI*, 2016.

[164] Mikhail Belkin, Irina Matveeva, and Partha Niyogi. *Regularization and Semi-supervised Learning on Large Graphs*. Springer Berlin Heidelberg, 2004.

[165] Marco Loog "*Contrastive pessimistic likelihood estimation for semi- supervised classification*" *IEEE Trans Pattern Anal Mach Intell*, 38(3):462–475, 2016.

[166] Zhiwu Lu, Xin Gao, Liwei Wang, Ji Rong Wen, and Songfang Huang. "*Noise-robust semi-supervised learning by large-scale sparse coding*" In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 2828– 2834, 2015.

[167] Yu Feng Li, James T. Kwok, and Zhi Hua Zhou" *Cost-sensitive semi- supervised support vector machine*" In *Twenty-Fourth AAAI Conference on Artificial Intelligence*, pages 500–505, 2010.

[168] Xiaojun Chang, Feiping Nie, Yi Yang, and Heng Huang" *A convex formulation for semi-supervised multi-label feature selection*" In *Twenty- Eighth AAAI Conference on Artificial Intelligence*, pages 1171–1177, 2014.

[169] X. B. Bai, K. Q. Wang, and H. Wang. "*Research on the classification of wood texture based on gray level co-occurrence matrix*" *Journal of Harbin Institute of Technology*, 37(12):1667–1670, 2005.

[170] Shuo Chen, Chengdong Wu, Dongyue Chen, and Wenjun Tan. '*Scene classification based on gray level-gradient co-occurrence matrix in the neighborhood of interest points*" In *IEEE International Conference on Intelligent Computing and Intelligent Systems*, pages 482–485, 2009.

[171] Hafeez Ullah Amin, Aamir Saeed Malik, Rana Fayyaz Ahmad, Nasreen Badruddin,

Nidal Kamel, Muhammad Hussain, and Weng Tink Chooi."*Feature extraction and classification for eeg signals using wavelet transform and machine learning techniques*" *Australas Phys Eng Sci Med*, 38(1):1–11, 2015.

[172]   Claude Turner, Anthony Joseph, Murat Aksu, and Heather Langdond" *The wavelet and fourier transforms in feature extraction for text- dependent, filterbank-based speaker recognition*" *Procedia Computer Science*, 6(1):124–129, 2011.

[173]   D. Pascolini and S. P. Mariotti, "*Global estimates of visual impairment: 2010,*" *Br. J. Ophthalmol.*, vol. 96, no. 5, pp. 614–618, 2012.

[174]   S. Dua, U. R. Acharya, P. Chowriappa, and S. V. Sree, "*Wavelet-based energy features for glaucomatous image classification,*" *Ieee Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 80–87, 2012.

[175]   S. Kankanahalli, P. M. Burlina, Y. Wolfson, D. E. Freund, and N. M. Bressler, "*Automated classification of severity of age-related macular degeneration from fundus photographs,*" *Invest. Ophthalmol. Vis. Sci.*, vol. 54, no. 3, pp. 1789–1796, 2013.

[176]   M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "*Computer-aided diagnosis of diabetic retinopathy*: A review," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2136–2155, 2013.

[177]   C. H. Parikh, S. Fowler, and R. Davis, "*Cataract screening using telemedicine and digital fundus photography*," *Invest. Ophthalmol. Vis. Sci.*, vol. 46, no. 13, p. 1944, 2005.

[178]   X. U. Liang, Q. F. Liang, and W. Shuang, "*Cataract screening as a key step to vision restoring project in country,*" *Ophthalmol. China*, vol. 19, no. 1, pp. 1–3, 2010.

.

[179]   C. Muramatsu *et al.*, "*Automated segmentation of optic disc region on retinal fundus photographs: Comparison of contour modeling and pixel classification methods,*" *Comput. Methods Programs Biomed.*, vol. 101, no. 1, pp. 23–32, 2011.

[180]   D. Allen and A. Vasavada, "*Cataract and surgery for cataract.*" Bmj, vol. 333, no. 7559, pp. 128–32, 2006

[181]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, "*Imagenet classification with deep convolutional neural networks,*" in International Conference on Neural Information Processing Systems, 2012, pp. 1097–1105.

[182] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "*Rich feature hierarchies for accurate object detection and semantic segmentation,*" in IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580–587

[183] E. P. I. W. Suartika, A. Y. Wijaya, and Soelaiman, R., "*Image Classification Using Convolutional Neural Network (CNN) on Caltech 101*", Jurnal Teknik ITS Vol. 5, No. 1, 2016. ISSN: 2337-3539 (2301-9271 Print)

[184] E. S. Marquez, J. S. Hare, and M. Niranjan, "*Deep Cascade Learning,*" IEEE Transactions on Neural Networks and Learning Systems, vol.29, no. 11, pp. 5475-5485, November 2018. Available: https://ieeexplore.ieee.org/document/8307262.

[185] W. Rawat and Z. Wang, "*Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review,*" Neural Computation, vol.29, no.9, pp. 2352-   2449,   September   2017.   Available: https://ieeexplore.ieee.org/document/8016501

[186] Wahyono and Hariyono, J. 2019. "*Determining Optimal Architecture of CNN using Genetic Algorithm for Vehicle Classification System*" Indonesian Journal of Computing and Cybernetics Systems Vol.13, No.1, January 2019, pp. 63~72 ISSN (print):   1978-1520,   ISSN   (online):   2460-7258   DOI: https://doi.org/10.22146/ijccs.42299

[187] J. He and G. Lin, "*Average Convergence Rate of Evolutionary Algorithms*", IEEE Transactions on Evolutionary Computation, vol.20, no.2, pp.316-321, 2016. Available: https://ieeexplore.ieee.org/document/7122298.

[188] D. Corus, D.-C. Dang, A. V. Eremeev, and P. K. Lehre, "*Level-Based Analysis of Genetic Algorithms and Other Search Processes*". IEEE Transactions on Evolutionary   Computation,   vol.22,   no.5,   pp.707-719,   2018.   Available: https://ieeexplore.ieee.org/document/8039236

[189] Karamihan, K. C., Agustino, I. D. F., Bionesta, R. B. B., Tuason, F. C., Arellano, S. V. E., & Esguerra, P. A. M. (2019). "*SBC-Based cataract detection system using deep convolutional neural network with transfer learning algorithm*" International Journal   of   Recent   Technology   and   Engineering,   *8*(2),   4605–4613. https://doi.org/10.35940/ijrte.B3368.078219

[190] Zhang, L., Li, J., Zhang, I., Han, H., Liu, B., Yang, J., & Wang, Q. (2017). "*Automatic cataract detection and grading using Deep Convolutional Neural Network*". *Proceedings of the 2017 IEEE 14th International Conference on*

Networking, Sensing and Control, ICNSC 2017, 60–65. https://doi.org/10.1109/ICNSC.2017.8000068

[191] Rismiyati and SN. Azhari, "*Convolutional Neural Network implementation for imagebased Salak sortation*," ICST (International Conference on Science and TechnologyComputer, 27- 28 Oct. 2016 [Online]. Available:https://ieeexplore.ieee.org/document/7877351.

[192] Milosevic, N. (2020). "*Introduction to Convolutional Neural Networks. In Introduction to Convolutional Neural Networks*"https://doi.org/10.1007/978-1-4842-5648-0

[193] Nurhikmat, Triano. 2018. Implementasi Deep Learning Untuk Image Classification Menggunakan Algoritma Convolutonal Neural Network (CNN) Pada Citra Wayang Golek. Skripsi. Yogyakarta: Universitas Islam Indonesia.

[194] Dzulqarnain, M.F, Suprapto and Makhrus, F. 2019"*Improvement of Convolutional Neural Network Accuracy on Salak Classification Based Quality on Digital Image*" Indonesian Journal of Computing and Cybernetics Systems Vol.13, No.2, April 2019, pp. 189~198 ISSN (print): 1978-1520, ISSN (online): 2460-7258 DOI: 10.22146/ijccs.42036.

[195] S.R. Rupanagudi, G.B. Varsa, dan B.S. Ranjani, "*A cost effective tomato maturity grading system using image processing for farmers*," IC3I (International Conference on Contemporary Computing and Informatics, 27-29 Nov. 2014 [Online]. Available: https://ieeexplore.ieee.org/document/7019591.

[196] Srivastava, N., Hinton, G., Krizhevsky, A., Sustkever, I., Salkhutdinov, R. 2014. Droput: "*A Simple Way to Prevent Neural Network from Overfitting*" Journal of Machine Learning Research, 15(56), 1929-1958.

[197]   Zhu, Q., He, Z., Zhang, T., Cui, W. 2020" *Improving Classification Performance of Softmax Loss Function Based on Scalable Batch-Normalization*" Appli.Scie.2020, 0,2950;DOI.10.3390/app10082590

[198]   Amit kumar shakya,Ayushman ramola,Anurag vidyathi"*Explation of pixel based and object based detection techniques by analyzing ALOS PALSAR and LANDSAT Data* "*smart ans sustainable intelligent systems*@2021 scrievener publishing LLC

[199]   Amit kumar shakya,Ayushman ramola,Deepak kumar pandey"*polygonal region of interest based compression of DICOM IMAGES*" ICCCA 2017

[200]   Amit kumar shakya,Ayushman ramola,Anurag vidyarthi,"*modeling of texture*

qualification and image classification for change prediction due to COVID lockdown using skysat and planetscope imagery" modeling earth system and environment 2022.

[201] Ayushman kamola,Amit kumar shakya,Dai van pham"*Study of statistical methods for tecture analysis and their modern evolution*"Engineering reports by john wiles ans sons ltd 2020

[202] Amit kumar shakya,Ayushman ramola,Kunal sawant,Akhilesh kandwal "c*omparison of supervised classification techniques for high resolution optical aerial image*"ICACE 2018

[203] Amit kumar shakya,Ayushman ramola,Akhilesh kandwal ,Rishi prakash"*Change over time in gray levels of multispectral landsat 5TM/8OLI satellite images*"springer nature Singapore pie ltd 2019.

[204]   Amit kumar shakya,Ayushman ramola,prag mittal ,rishi prakash"*morphological change detection in terror camps of area 3 and 4 by pre and post strike throug*h MOAB-B."Springer nature Singapore pie ltd 2019.

[205] www.ncert.in

Comparative Study of Text Pre-processing Techniques and Machine Learning Algorithms Based on Sentiment Analysis Applications on E-Commerce Data

[1]Beena Kapadia , [2]Dr Amita Jain

[1]Research Scholar, School of Computer Application and Technology, Career Point University, Kota, Rajasthan, India

[2]Research Associate, School of Computer Application and Technology, Career Point University, Kota, Rajasthan, India

Abstract. Sentiment Analysis is the computational treatment of opinions, sentiments and subjectivity of text and use them for the benefit of the business operations. This survey paper tackles a comprehensive overview of various sentiment analysis applications related to E-commerce data and includes analysis of related papers from 2008 to 2020.

This paper gives overall idea about various data pre-processing techniques, Sentiment Analysis algorithms, its accuracy, further improvements and other related details of each referred applications used, as literature survey in the area of E-commerce. The main contributions of this paper include comprehensive analysis of many relevant E-commerce articles, illustration of data pre-processing techniques and the illustration of the recent trend of research in the sentimental Analysis and related areas.

Keywords: Sentiment Analysis, data pre-processing techniques, E-commerce data.

Introduction:

Artificial intelligence (AI) teaches systems to to intelligent things. Machine Learning(ML), which is the sub-field of Artificial Intelligence teaches systems to do intelligent things that can learn from experience. Natural Language Processing (NLP), which is the sub-field of Machine Learning teaches systems to be intelligent, learn from experience and understand the human language.

NLP is a branch of AI which deals with the interaction between computers and humans using any natural language. The objective of NLP is to read, decipher, understand and make sense.

The more data you collect, the more you can correct your algorithm's mistakes and reinforce its correct answers.

Every second, on an average, almost 6,000 tweets are tweeted on Twitter, which corresponds to around 350,000 tweets sent per minute, 500 million tweets per day and almost 200 billion tweets per year [51]. Facebook users create almost 3.2 billion likes and comments every day [50]. Generally, almost everyone reads amazon reviews before buying any valuable. To generate that reviews, we have to do text mining and text classification. Text mining is the process of exploring sizable textual data and patterns generate valuable insights enabling companies to make data-driven decisions. It converts unstructured data to structured data. You can summaries or visualize the data. Text classification is nothing but the process of assigning tags or categories to text according to its content like spam mails, document classification in particular label, sentence classification as the sentence can have positive or negative sentiment, customer support to make out their liking or disliking of a particular product, Sentiment Analysis(SA), content recommendation etc.

Sentiment Analysis has become most important application of Natural Language Processing. Sentiment analysis is the process of examining a piece of text for opinions and feelings. There are ample of real-life use cases for sentiment analysis that include understanding how consumers feel about a particular product or service, looking for signs of depression, or to see how people respond to certain advertise and political campaigns.

Data Pre-processing techniques play an important role. [6] In spam filtering process with Naïve Bayes Classification the authors achieved moderate prediction before applying any Data Pre-processing techniques but after applying proper data pre-processing, they have improved the prediction accuracy and also proved that the Pre-processing done on data has a larger impact in the performance of the Naïve Bayes classification.

High dimension while processing the text is one of the major issues. In text processing, the total number of distinct words in a corpus is known as dimensions. With data Pre-processing the dimensions of text corpus can be reduced. It is possible to reduce the size of dictionary of email dataset from 1093044 words to 654000 after applying various data Pre-processing techniques like tokenization, removal of URL's, '@' symbol, punctuation marks, money

symbols, white spaces and special characters, lowercase conversion, stop-word removal, enhanced stop word removal, removal of email signature and lemmatization strategies respectively. [7] In one more study of data Pre-processing techniques, it is seen that the dimension of corpus has been reduced drastically, when studied on BBC news and BBC sports dataset.[8]

There are three main classification levels in SA: document-level, sentence-level and aspect-level SA. Document-level SA aims to classify an opinion; whether the document as expressing a positive or negative opinion or sentiment. It considers the entire document as a basic information unit. Sentence-level SA is used to classify the sentiment expressed in each sentence. The first step for sentence level sentiment analysis is to identify whether the sentence is subjective or objective. If the sentence is subjective, then only Sentence-level SA can determine whether the sentence expresses positive or negative opinions. Classifying text at the document level or at the sentence level does not provide the necessary detail needed in many applications. We need to go to the aspect level to obtain these details. Aspect-level SA is used to classify the sentiment with respect to the specific aspects of entities, which is a noun in general. The first step is to identify the entities and their aspects. [42]

This survey can be useful for new researchers in the field of Sentiment analysis and very specifically for E-commerce purpose only, as it covers the most famous SA techniques and its applications that to, only in the specific domain of E-Commerce. This survey uniquely gives a refined categorization to the various SA techniques for the purpose of finding insights of E-commerce data only and not any other types of reviews.

2. Objective

To conduct an extensive comparison of the Data Pre-processing techniques by applying it to some algorithms to generate the model and assess the effect of data Pre-processing techniques

To compare the speed of five different types of Pre-processing technique - Stemming, Stemming and Lemmatization, Stemming and Spelling Correction, Stemming and Term Frequency -Inverse Document Frequency & Data Pre-processing with all Pre-processing Techniques.

To compute the accuracy of Support Vector Machine (SVM), Random Forest Classification and Multinomial Naïve Bayes (NB) classifiers to assess the better performer.

To provide a complete picture to the researchers, who wants to work further in this area about how many researches have been done so far with sentiment analysis in the domain of E-Commerce and which algorithms are being used. The accuracy each algorithm could achieve with given set of data and what further improvements are suggested. Based on that a researcher can decide where to start a research work from.

3. Literature Review of Sentiment Analysis Applications on E-Commerce Data

References	Year	Data set	Data extraction Tool / data pre-processing techniques	Ratio of testing and training data	Algorithm used	Accuracy %	Further improvements
[9]	2018	1000 data - UCI Machine Learning	Rapidminer - TF-IDF algorithm	70% Vs 30%	Decision Tree SVM K-Nearest Neighbor	82 91.00 90.33	more sophisticated NLP techniques and more consideration on smaller text chunks required
[10]	2014	2000 data - user generated product reviews	-	-	Tree Bank Model and Fuzzy Opinion Mining	85.58	Neuro-Fuzzy system should be considered instead of just fuzzy logic

		(unlabelled training set)			Model with sentiment classification approach		
[11]	2014	1) digital camera: Canon G3 2) digital camera: Nikon coolpix 4300 3) celluar phone:Nokia 6610 4) mp3 player: Creative Labs Nomad Jukebox Zen Xtra 40GB 5) dvd player: Apex AD2600 Progressive-scan DVD player	Stanford CoreNLP tool	-	association rule mining with supervised machine learning algorithm based polarity classifier (domain independent – Support Vector Machine)	79.6	Some other alternative technique can be explored
[12]	2008	20 newsgroups and 5000 web pages	Bag of Words	20000/6000 i.e. 77 %	support vector machines	96.00	The effectiveness of the SVM sentiment classifier

				vs 23 %			model can be increased with a small bag of words that consists of suitable features
[13]	2015	5576	The adjective/adverb words for extracting review features	-	J48 NaiveBayes SMO IBk RandomForest RandomTree	53.33 96.67 96.67 100 96.67 100	By combining some features of classifiers and by developing a hybrid classifier or algorithm and implementing it, accuracy can be improved.
[14]	2015	1039 reviews for Samsung Galaxy DuoS 2	python	-	Opinion Mining Methodology with Naïve Bayes Classification	-	reviews for more than 2 products and automatically rank products based on the features that the user is interested in.
[15]	2015	300 reviews	SentiWord Net Lexicon	-	lexicon based opinion mining	61	to optimize the algorithm so that it provides an

70

					methodology with SentiWordNet classification		extended semantic processing of text sources.
[16]	2015	blue quartz watch - Chinese corpora which	-	-	Machine learning algorithms for opinion mining	-	The algorithm should be developed for better accuracy. Relevant corpora can be gathered more comprehensively
[17]	2019	50 product review for Camera, screen, RAM, battery	Bag-of-Words	80% vs 20%	LDA algorithm with a combination of parameters of alpha 0.001, beta 0.001, and number of topics 15 and number of iterations	67.5	-

					1. with 3 new user review data that has been labeled, the average difference in accuracy obtained increases by 0.82%.		
[18]	2019	beauty products and musical instruments	Preprocessed using Bow (Bag of words)		training model along with the Recall, Precision, F1 and ROC AUC, Software library used for this study is scikit-learn with SVM machine learning	87.88	advanced machine learning and deep learning approaches.

					technique		
[19]	2019	Chinese Sentiment analysis – 32,558 reviews (Document Level) were scraped from various Chinese Ecommerce websites, the reviews were split by its punctuation yielding 60,830 sentences.	character tokenisation - Word Cloud, TF-IDF, Aspect Extraction and Topic Modelling	training corpus of 60,830 and two test corpus (balanced and imbalanced) of 2171, 2519	Gated Recurrent Unit (GRU) with deep learning technique of using Gated Recurrent Unit (GRU) neural network	(87.66% - balanced, 87.9% imbalanced)	the pre-processing step of word embedding, there appears to be more advanced methods of converting text into a vector space such as Word2Vec, GloVe (Global Vectors for Word Representation) and fastText.
[20]	2017	37,126 reviews, in JSON format.	Process the extracted data using Python	-	New algorithm SENTI which is compared with Datumbox and NLTK – machine	SENTI classifies the polarity better than NLTK and Datumbox	More studies are required in future work to incorporate and close the research gaps for making better business decisions.

					learning		
[21]	2020	50000 words	10-fold cross-validation method	9:1 part	a new sentiment analysis model-SLCABG, which is based on the sentiment lexicon and combines Convolutional Neural Network (CNN) and attention-based Bidirectional Gated Recurrent Unit (BiGRU) to classify the weighted sentiment features	93.5	to study the sentiment fineness classification of text
[22	20	Historical		80%	a	accuracy	The s analysis

|] | 19 | sales and online reviews. A group of 16,155 online reviews of 5 different online tea shops was collected | | vs 20% | Dependency SCOR-topic Sentiment (DSTS) model compared with ARRIM model | is improved through integrating the ARRIM into DSTS model | is limited to online users who leave reviews at a Chinese review website. Hence, this analysis focuses on review texts written in Chinese. It would be interesting if future research expands the study to a global context. |
| [23] | 2019 | - | Hadoop and Apache Spark to accelerate the text Pre-processing | - | TF-IDF, Machine Learning (ML) techniques, vectorization, LR for 1) Subjectivity | - (it is a case study) | the further study will need to look into learning model that support multiple languages |

					classification, 2) Polarity classification and 3) Graph-based optimization		
[24]	2017	1500 mobile phone brand Redmi Note 1	-	90% vs 10%	Comparison of ESAGBA model, Naive Bayes, SVM , Maximum Entropy	82 80 72.85 74.28	Not much significant work has been done in this domain and thus has great scope of quality research.
[25]	2016	-	Stanford parse, extraction algorithm, Mine association rule, Summarization	-	-	-	-
[26]	2018	2,500 consumers' reviews	longest matching algorithm	-	Sentiment compensation technique, m	93.60	proposed method should be improved to support more

					sentiment to dimension (S2D) and dimension to sentiment (D2S) methods (product, price, and shipping dimensions)		complex consumer's reviews which may have multiple dimensions and multiple sentiments in a single sentence
[27]	2016	9249 mobile data and 7753 Hotel data	LDA-based topic modeling method, Appraisal expression pattern extraction	70% vs 30%	LDA Model, Aspect-level Sentiment SVM (TF-IDF) SVM (Lexicon-based sentiment analysis) SVM (Multi-aspect	0.670m 0.706h 0.773m0 .767h 0.809m0 .781h	To improve the opinion short sentences discovery by importing statistic methods

					sentiment analysis)		
[28]	2013	2000 text	-	-	Decision Tree Naive Bayes Naive Bayes Multinomial SMO	Applied 4 types of classifications	Designed a tool and enhancement in tool can be to evaluate the SentiMeter-Br in other contexts (business, education, technology, fashion, health).
[29]	2018	-	adjective extraction from the reviews	-	Naive Bayes Support Vector Machine for aspect-based sentiment analysis	90.423 83.423	C4.5 algorithm can be compared with Naïve Bayes and compare with Naïve Bayes.
[30]	2017	200 galaxy s5 mob 150 Microcanvas 180 HTC one Micronitro 178	POS tagging and stop word removal	-	Clustering with TF-IDF and classification with Support vector	90.99	-

					mechanism using Java		
[31]	2019	28462 notes that contain 66258 sentences	TextBlob	153 lost deals and 146 won deals Vs. 50 lost deals and 50 won deals	Artificial Neural Networks, Support Vector Machines and Random Forests	89.11	to conduct experiments with additional B2B data sets, use incremental learning and deploy the findings in a live CRM system
[32]	2015	32384 e-commerce reviews, 6 types of sentiments	χ2 (chi-square) and Pointwise Mutual Information (PMI) metrics for feature selection	2:1	SVM Baseline group 1 Baseline group 1 Experimental group	0.8422 0.8516 0.8800 (for +ve responses)	use the e-commerce reviews for word2vec training and then build the sentiment dictionary for the e-commerce corpus
[33]	2016	Nokia 6610-700 review, Canon G3-	POS tagger provided	-	Recursive Deep model	81.8	Product feature extraction can

		900 review and DVD player – 1050 review	by stanford core nlp, Candidate product feature extraction				be further improved by adding implicit feature extraction. s
[34]	2018	various laptop products, feature: Portability, Durability, Screen Quality, Battery, and Performance (of the) Central Processing Unit	DBpedia for synonymous words	-	SentiWordNet	94%	use of slang, idioms, and ironical sentences
[35]	2020	The academic business data set	Tokenization, Text Cleaning to Delete alphanumeric words, Remove stop words, Convert to	-	Naive Bayes, Random Forest, Decision Tree, Support Vector M/c, K-Nearest Neighbor	93.2% 87% 88% 92% 83% 93.4%	To ensemble the classifier and analyze the accuracy

			small letters.		and Multilayer Perceptron classifiers		
[36]	2020	COAE2014-task4 ChnSentiCorp-Htl-ba-6000	-	-	CNN Multi-Bi-LSTM Bert-BiGRU WWAL Bert-BiGRU	88.20 % 88.40 % 89.03 % 87.27 % 88.31 %	-
[37]	2020	IMDB, blogs and social media platforms.	-	-	ML approach Lexicon based approach	SA is a prominent field based on quick computing, large volume of data & info.	-
[38]	2021	Amazon mobile phone Reviews	Tokenization, stop word removal, Lemmatization, Lowercase conversio	-	LR RF NB BERT LSTM	85.3 % 89.9% 78.4% 94.7% 93.3%	we are planning to use word2vec for feature extraction with our models and to detect fake

			n,			reviews.
[39]	2021	Amazon e-commerce site for an electronics product.	removal of mentions, hashtag, stopwords tfidf		TextBlob library	Frequency distribution of top 20 positive sentiments	A better methodology can be suggested for sentiment analysis with an improved deep learning approach in Spark NLP to relationship analysis
[40]	2022	Python crawler library	Word Tokenization, POS tagging, other preprocessing	9:1	SRank Weightless Network Important Node Sorting Algorithm	The specific sentiment polarity of comments for each note is found	-
[41]	2023	Dangdang Flipkart Amazon Twitter etc	-	-	Review of 54 papers ML techniques DL techniques Hybrid	usage 48.1% 44.4% 7.4%	focus on developing more universal models for new domains and languages, aspect-level

					approach		SA models, implicit aspect recognition and extraction, sarcasm detection, fine-grained SA to increase the usage of SA in e-commerce

Table 1 contains information about various researches that have been done so far on sentimental analysis, specifically on product reviews.

4. Research Methodology

This paper is intended to study various Pre-processing techniques in different combinations to find the importance of each data Pre-processing techniques. For this paper, data has been collected from github. The data file modi-trainedtweets.csv has been collected from github[3], which has the twitter file all-trainedtweets.csv, in which data collected from Modi ji and Rahul Gandhi's tweets, along with the sentiment score of a people about both of them. Hence, we have done supervised Learning for this paper and applied a few of its algorithms. In this study we have just took partial data, only of Prime Minister Modi ji and studies that.

The sources of papers for preparing Table 1 content are IEEE Xplore, ACM digital library, Re-search gate and Elsevier. Total 33 articles presented in this research paper which are related to sentiment analysis in the area of product review only. The algorithms used in this paper are mainly Support Vector Machine and Naïve Bayes classification.

We have applied following Data Pre-processing Techniques, for which we have used Python programming using Google Collaboratory online editor:

• Replace non alphabets to blank space

• Remove single char

• Spell check and correction

• Convert all character to lower

• stemming the words

• Lemmatization of the words

• Applying Bag of Words OR Term Frequency-Inverse Document Frequency technique

Replace non alphabets to blank space: The reason for converting a non-alphabet character to a blank space is there that any non-alphabet character like %,$,#,@ etc doesn't have any meaning towards any sentiment of a human. In Python it can be done with Regular expression and sub() method of regular expression.

Remove single char: The reason for removing a single alphabet is that any single character doesn't have any meaning towards any sentiment of a human. In Python it can be done with Regular expression and sub() method of regular expression.

Spell check and correction: To know the meaning of any word is extremely important to know the sentiment conveyed. And for that the correct spelling is very much required. One of the ways to correct the incorrect spellings in Python by using TextBlob. To use it, we require to import nltk package and download punkt.

Convert all character to lower: The reason for the upper-case letters is to reduce the features (total unique words). Because 'Hello' and 'hello' can be consider as two separate words. So, if we convert all text content to either upper case or lower-case letters then that way we can reduce the dimensions. Generally, it is easy to read all lower-case letters than all upper-case letters. In Python it can be done with lower() method of String class.

Stemming the words: Bringing the word to its root word, [2] but not necessarily to be the dictionary word is called stemming. For example, happier, happiest can be stemmed to happi. The reason for stemming the word is again to reduce the features. In Python it can be done with stem() method and PorterStemmer required to import from nltk.stem.porter. It also removes the stop words like am, an, the, this, that etc., which doesn't have any sentiment value.

Lemmatization of the words: Bringing the word to its root word, [2] but as per the dictionary word is called

Lemmatization. In other words, Lemmatization follows lexical knowledge to get the root word for original one. For example, caring will become care after lemmatization. The reason for lamming the word is again to reduce the features. In Python it can be done with lemmatize () method, which can be called on the object of WordNetLemmatizer.

Applying Bag of Words OR Term Frequency-Inverse Document Frequency technique: In bag of words, all unique words are to be searched from the corpus, then we note the frequency of occurrence of each unique word and then we sort them in descending order. All the word mapped with vector we have are used to represent each sentence. Bag of Words is used to create the dimension vector, but it is having some drawbacks and to overcome that drawbacks, Term Frequency-Inverse Document Frequency is used. Bag of words creates sparse matrix, which occupies a lot of space and increase complexity. Also, each word can be represented as either one or zero, but not the weight as anything between zero and one.[5]

For purpose of extraction of numeric parameters from plain text, TF-IDF (Term Frequency Inverse Document Frequency) is used which makes possible to calculate the "weight" of word in document based on its "reverse" frequency in other texts in the whole corpus. TFIDF is mainly used to compute the usefulness of each word in the document. [4] TF measures how many times a word appears in the document and IDF represents how common the word is across the different documents. We can calculate TF-IDF directly in Python by importing TfidfVectorizer from sklearn. And require to create the object of TfidfVectorizer using parameterised constructor. On that object fit_transform() method should be called using the parameter as the entire document.

We have combined different data Pre-processing techniques in five different combination in our study. Various combinations are taken as follows:

Data Pre-processing with Stemming: In this step, a, d and e data Pre-processing techniques are used

Data Pre-processing with Stemming & Lemmatization: In this step, a, d, e and f data Pre-processing techniques are used

Data Pre-processing with Stemming & Spell correcting: b, a, c, d, and e data Pre-processing techniques are used

Data Pre-processing with Stemming & TFIDF: a, d, e and g data Pre-processing techniques are used

Data Pre-processing with all Pre-processing Techniques used from a to g.

All NLP algorithms including sentiment analysis works on numbers. That means that we can't use text as the input to any algorithm directly. Various algorithm can be Multinomial Naïve Bayes or K nearest neighbour or Regression or Support Vector machine or Random Forest or any other algorithm. To make the algorithm compatible with text, we have used Bag of words algorithm, which coverts each term to a number. For each tweet, an array of numbers will be generated. First, all unique words are found and that will be considered as one of the dimensions (columns) of the array. If a particular word is present in that sentence of the array, then the element of that row and column becomes 1 else it becomes 0 (zero). For each tweet (text), we have corresponding sentiment score also in our data file. Hence, each tweet, which is converted to zeros and ones becomes x column and each corresponding tweet score becomes the y column. Both are array of numbers.

Now we can split the data into training set and test set. Here, we have considered 70-30 ratio of training set and test set. Once, splitting is done, then we generated three different model – Multinomial Naïve Bayes, Support Vector Machine and Random Forest algorithm, to assess its effects.

5. Data Analysis and Discussion

5.1 Data Analysis and Discussion of Data Pre-processing

As it is a kind of supervised learning, we have just first checked with bar chart as shown in figure 1 and with pie chart as shown in figure 2 that how many positive tweets are there in favour of Modi ji and how many negative and neutral tweets are there; irrespective to any data Pre-processing techniques and irrespective to designing any model.
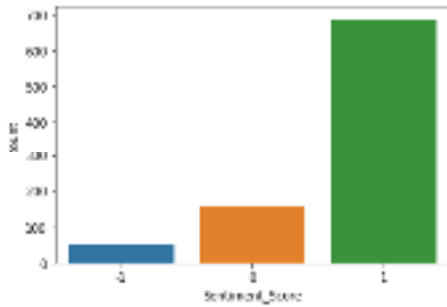
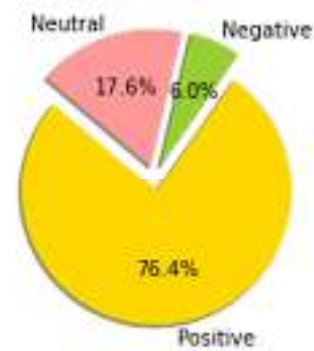Figure 1:   Figure2:

Figure 1 and Figure 2: negative, neutral, and positive sentiments as per the tweets and sentiment score in the datafile

Our first objective was to conduct an extensive comparison of the Data Pre-processing techniques by applying it to the same algorithms to generate the model and assess the effect of data Pre-processing techniques. The study of which is shown in the following table 1.

For that first, we have checked step 1 that is Data Pre-processing with Stemming. As we can observe from Table 2 that step 1. Data Pre-processing with Stemming gives better results by applying it on Random Forest algorithm.

The step2. Data Pre-processing with Stemming & Lemmatization gives best results by applying it on Random Forest algorithm.

The step3. Data Pre-processing with Stemming & Spell correcting gives best results by applying it on Support Vector Machine algorithm.

The step4. Data Pre-processing with Stemming & TFIDF gives best results by applying it on random Forest algorithm.

The step5. Data Pre-processing with all Pre-processing Techniques gives best results by applying it on Support Vector Machine algorithm.

Steps / Model accuracy	1.Data Pre-processing with Stemming	2.Data Pre-processing with Stemming & Lemmatization	3.Data Pre-processing with Stemming & Spell correcting	4. Data Pre-processing with Stemming & TFIDF	5. Data Pre-processing with all Pre-processing Techniques
Multinomial Naïve Bayes	0.78967	0.79705	0.7749	0.76753	0.76753

SVM	0.79336	0.79705	0.81181	0.81181	0.80812
Random Forest	0.80812	0.80443	0.80073	0.81550	0.80812

Table 2 – Effect of Various steps of Data Pre-processing Techniques on Multinomial Naïve Bayes (NB) classifiers, Support Vector Machine (SVM) and Random Forest Classification.

Our next objective is to compare the speed of five different types of Pre-processing technique - Stemming, Stemming and Lemmatization, Stemming and Spelling Correction, Stemming and Term Frequency -Inverse Document Frequency & Data Pre-processing with all Pre-processing Techniques. The figure 3 shows the time taken in seconds for each of the execute data Pre-processing steps in Python.
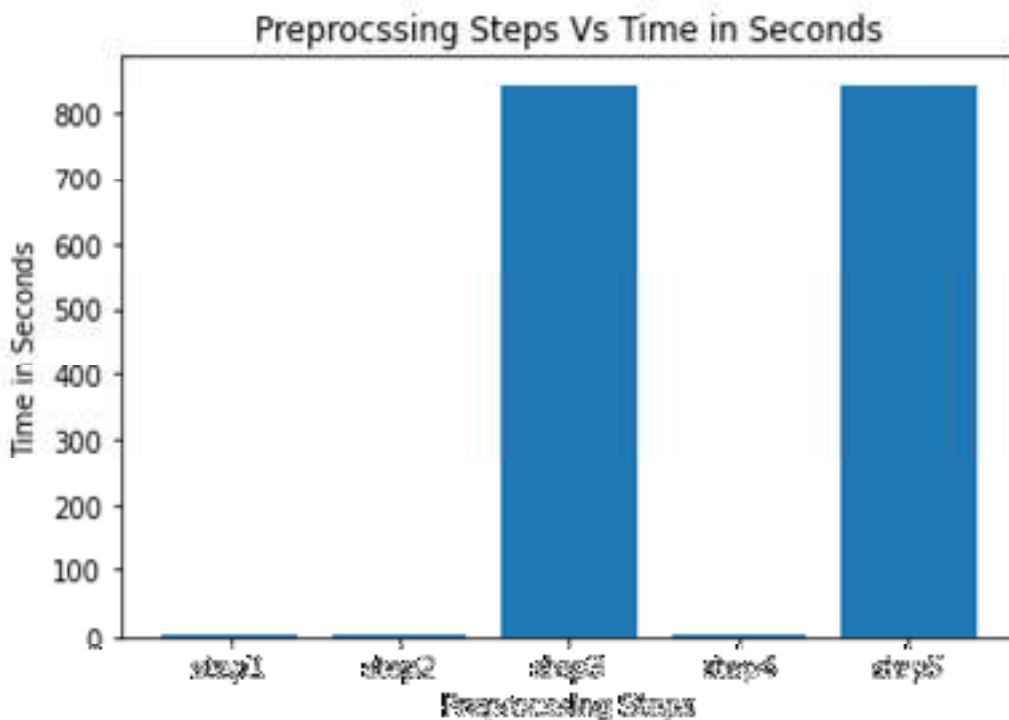
Figure 3

Figure 3 clearly shows that steps 3 and 5 takes more time to execute compared to all other steps. In step 3 and 5, the spelling correction algorithm takes ample amount of time because in these two steps that algorithm is used, which takes each word into consideration and check with the dictionary of English language and correct the spelling which is incorrect.

The general belief is that as the correct spelling we take into the procedure to model the algorithm, definitely very good results should come up as far as model accuracy is concerned

but as we can observe in the table 2, there is no drastic change observed after running the module of spell check and correction. So, in our opinion, this step of spell correction can be avoided.

5.2 Data Analysis and Discussion of Machine Learning Algorithm

Mainly the research in this area is based on feature selection and algorithm used.

5.2.1 Feature Selection in Sentiment Classification

Sentiment Analysis task is considered a sentiment classification problem. The first step in the SC problem is to extract and select text features. Some of the current features are [45] Terms presence and frequency, Parts of speech (POS)[46], Opinion words and phrases and Negations.

Feature Selection methods can be divided into lexicon-based methods that need human annotation, and statistical methods which are automatic methods that are more frequently used [44]. Other methods also can be used in Feature Selection like information gain and Gini index [45].

5.2.2 Sentiment Classification Techniques

Sentiment Classification techniques can be roughly divided into machine learning approach, lexicon-based approach and hybrid approach [47]. The Machine Learning Approach (ML) applies the famous ML algorithms and uses linguistic features. The Lexicon-based Approach relies on a sentiment lexicon, a collection of known and precompiled sentiment terms. It is divided into dictionary-based approach and corpus-based approach which use statistical or semantic methods to find sentiment polarity. The hybrid Approach combines both approaches and is very common with sentiment lexicons playing a key role in most methods. The text classification methods using ML approach can be roughly divided into supervised and unsupervised learning methods. The supervised methods make use of many labelled training documents. The unsupervised methods are used when it is difficult to find these labelled training documents. The lexicon-based approach depends on finding the opinion lexicon

which is used to analyse the text. There are two methods in this approach. The dictionary-based approach which depends on finding opinion seed words, and then searches the dictionary of their synonyms and antonyms. The corpus-based approach begins with a seed list of opinion words, and then finds other opinion words in a large corpus to help in finding opinion words with context specific orientations. This could be done by using statistical or semantic methods.

5.2.3 Related Fields to Sentiment Analysis

Emotion detection, building resources and transfer learning are some related fields to sentiment analysis. The sentiment reflects feeling or emotion while emotion reflects attitude [48]. It was argued by Plutchik [49] that there are eight basic and prototypical emotions which are joy, sadness, anger, fear, trust, disgust, surprise, and anticipation. Emotions Detection (ED) can be considered a SA task. SA is concerned mainly in specifying positive or negative opinions, but ED is concerned with detecting various emotions from text. As a Sentiment Analysis task, ED can be implemented using ML approach or Lexicon-based approach, but Lexicon-based approach is more frequently used. Building Resources (BR) aims at creating lexica, dictionaries and corpora in which opinion expressions are annotated according to their polarity. Transfer learning extracts knowledge from auxiliary domain to improve the learning process in a target domain. Table 3 shows each algorithm whose accuracy is mentioned in the respective papers. There are some papers in which exact accuracy is not mentioned but just conveyed that one algorithm is better than other algorithm. The data of table 3 is defined using data of Table 2. Table 2 data is nothing but taken as the outcome of the study of literature review. The visualization of data is shown in Figure 4.

Table 3: Algorithm and its accuracy

Sr no	Algorithm	Algorithm (in bar chart)	Accuracy	Number of times it is used
1	ARRIM model	ARRIM	-	1
2	Artificial Neural Networks	ANN	89.1	1

				1
3	association rule mining	ARM	79.6	1
4	Datumbox	DB	-	1
5	Decision Tree	DTT	82	2
6	deep learning technique	DLT	-	1
7	DSTS model	DSTS	-	1
8	ESAGBA model	ESAGBA	82	1
9	Fuzzy Opinion Mining Model	FOM	85.58	1
10	Gated Recurrent Unit with Deep Learning	GRL	87.78	1
11	IBk	IBK	100	1
12	J48	J48	53.33	1
13	K-Nearest Neighbor	KNN	90.33	1
14	LDA algorithm	LDA	67.5	2
15	Machine learning algorithms	MLA	-	2
16	Maximum Entropy	ME	74.28	1
17	Naïve Bayes classification	NBC	89.03	4
18	Naive Bayes Multinomial	NBM	-	1
19	NLTK	NLTK	-	1
20	RandomForest	RF	92.89	2
21	RandomTree	RT	100	1
22	Recursive Deep model	RDM	81.8	1
23	SENTI	SENTI	-	1

24	sentiment classification approach	SCA	-	1
25	Sentiment compensation technique	SCT	93.6	1
26	SentiWordNet classification	SWN	77.5	2
27	SLCABG	SLCAB G	93.5	1
28	SMO	SMO	96.6 7	2
29	supervised machine learning algorithm	SMA	79.6	3
30	Support Vector Machine	SVM	87.1 3	7
31	Tree Bank Model	TBM	85.5 8	1
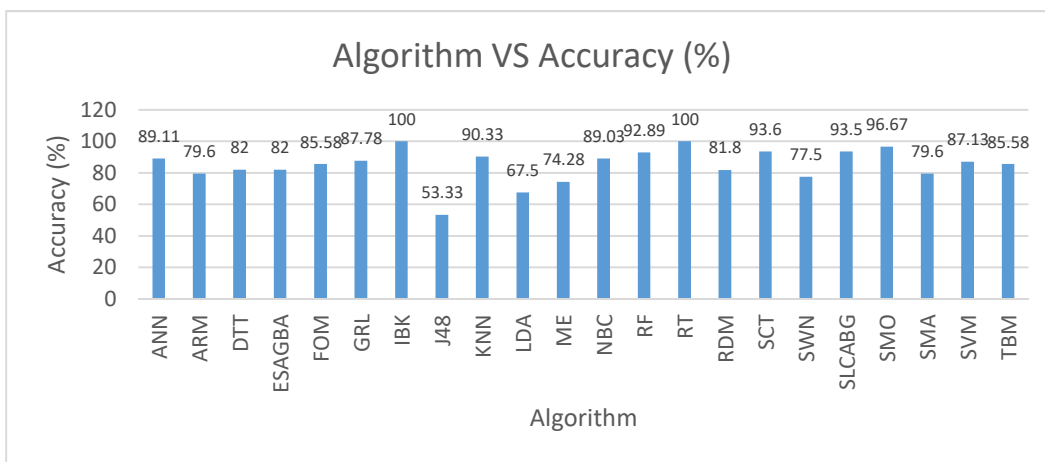
Figure 4: Visualization of Algorithm VS Accuracy

We can see from the table 3 and its visualization as per figure 4, that minimum accuracy is shown by J48 algorithm which is 53.33 percentage; whereas maximum accuracy can be shown by IBk and RandomTree (RT), that is exceptionally good – 100 %. But when we check it minutely, then we can make out that its calculated on offline data. It is not mentioned

that data were all structured data, unstructured data or semi structured data. Further negative statements are considered or just the negative words are considered. Also, only adjective/adverb words are considered, but not Noun[43]. Also, we are not sure that response came from positive question or negative question. Table 3 also shows that Support Vector Machine is used most frequently for sentiment analysis on product or e-commerce reviews.

6. Conclusion

The final objective of the paper is to compute the accuracy of Support Vector Machine (SVM), Random Forest Classification and Multinomial Naïve Bayes (NB) classifiers to assess the better performer. As we can see the table 1, it is the Random Forest which is the best one irrespective of whatever Data Pre-processing technique is used, but Support Vector Machine also is good. It is very near to Random Forest. More precisely, if you want to use SVM algorithm, then it is better to use TFIDF rather than Bag of Words. For SVM spell check and correction can also give the same output as TFIDS but TFIDF takes very less time comparatively as shown in figure 4, to get better model accuracy.

There are plenty of data Pre-processing techniques but all are not used in this paper. Some of them, which we have not used in this paper are removal of HTML tags, removal of URLs, converting numbers to words, convert accented characters to ASCII characters, converting chat conversion words to normal words, expanding contractions, converting emojis to words, removing rare words, etc.

To find the most appropriate machine learning algorithm technique, the researcher can work on all the algorithms which have achieved accuracy of 85% or more in this study. ANN, FOM, GRL, IBK, KNN, NBC, RF, RT, SCT, SLCABG, SVM and TBM are the algorithms which falls under this category. So, the researchers can compare the accuracy using the same data set and decide which one gives the highest accuracy.

7. REFERENCES

[1] Kapadia B., Jain A. (2021) Analysis of Papers Based on Sentiment Analysis Applications on E-Commerce Data. In: Abraham A., Sasaki H., Rios R., Gandhi N., Singh U., Ma K. (eds) Innovations in Bio-Inspired Computing and Applications. IBICA 2020. Advances in Intelligent Systems and Computing, vol 1372. Springer, Cham. https://doi.org/10.1007/978-3-030-73603-3_30

[2] Pradha, S., Halgamuge, M. N., & Vinh, N. T. (2017). Effective Text Data Pre-processing Technique for Sentiment Analysis in Social Media Data. *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation* (p. 8). Da Nang, Vietnam: IEEE.

[3]Jaluthria, K. (July 6,2021). *Twitter-Sentiment-Analysis*. Retrieved from https://github.com/KUNNAL20/Twitter-Sentiment-Analysis as on Dec 29,2021.

[4] David Dietrich, B. H. (2015). Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. Canada: WILEY.

[5] kumar, K. (Jan 24 2021). *NLP: Bag of words and TF-IDF explained!* Retrieved from https://koushik1102.medium.com/nlp-bag-of-words-and-tf-idf-explained-fd1f49dce7c4

[6] Priyanga Chandrasekar, K. Q. (2016). The Impact of Data Pre-processing On the Performance of Naïve Bayes Classifier. IEEE 40th Annual Computer Software and Applications Conference, (p. 2). Marietta, GA, USA.

[7] A. K. B and M. M. Kodabagi, "Efficient Data Pre-processing approach for Imbalanced Data in Email Classification System," 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), 2020, pp. 338-341, doi: 10.1109/ICSTCEE49637.2020.9277221.

[8] A. I. Kadhim, Y. -N. Cheah and N. H. Ahamed, "Text Document Pre-processing and Dimension Reduction Techniques for Text Document Clustering," 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology, 2014, pp. 69-73, doi: 10.1109/ICAIET.2014.21.

[9]   Ardianda Aryo Prakoso, B. W. (2018). A Lexicon-based Sentiment Analysis for Amazon Web Review. *International Seminar on Application for Technology of*

Information and Communication (iSemantic) (p. 6). Semarang Indonesia: Dian Nuswantoro University.

[10]   Indhuja K, R. R. (2014). Fuzzy Logic Based Sentiment Analysis of Product Review Documents. *First International Conference on Computational Systems and Communications (ICCSC) | 17-18 December 2014* (p. 5). Trivandrum: Govt. Engg. College, Palakkad.

[11]   Siddharth Aravindan, A. E. (2014). Feature Extraction and Opinion Mining in Online Product Reviews. *International Conference on Information Technology* (p. 6). Patna: IEEE computer society.

[12]   Ghose, J. P. (2008). An Ontology-based Sentiment Classification Methodology for Online Customers review. *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (p. 7). University of Wollongong, Wollongong 2500 NSW: Australia .

[13]   Aditya A. Kshirsagar, P. A. (2015). Review Analyzer Analysis of Product reviews on WEKA Classifiers. *IEEE Sponsored 2nd International Conference on Innovations in Information, ,Embedded and Communication systems (ICIIECS)*, (p. 5).

[14]   Venkata Rajeev P, S. R. (2015). Recommending Products to Customers using Opinion Mining of Online Product Reviews and Features . *International Conference on Circuit, Power and Computing Technologies [ICCPCT]*, (p. 5).

[15]   Alexandra cernian, V. a. (2015). Sentiment analysis from product reviews using SentiWordNet as lexical resource. *IEEE 7th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, (p. 4).

[16]   Dongzhi Wang, X. Y. (2015). A Conceptual Framework of E-Commerce Supervision System Based on Opinion Mining. *International Conference on Services Science* (p. 4). Shenzhen, China: Tsinghua University.

[17]   Eko Wahyudi, R. K. (2019). Aspect Based Sentiment Analysis in E-Commerce User Reviews Using Latent Dirichlet Allocation (LDA) and Sentiment Lexicon. *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)* (p. 6). Semarang, Indonesia: DiponegoroUniversity.

[18]   Jahanzeb Jabbar, I. U. (2019). Real-time Sentiment Analysis On E-Commerce Application. *Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control, May 9-11, 2019, Banff, Alberta, Canada* (p. 6). P.R. China: Polytechnical University Xi'an.

[19]   Jun Sheng, L. D. (2019). Sentiment Analysis of Chinese Product Reviews using Gated Recurrent Unit. *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)* (p. 9). Singapore: Logitech Europe S.A., Singapore.

[20]   Karthikayini T, P. N. (2017). Comparative polarity analysis on Amazon product reviews using existing machine learning algorithms. *2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017* (p. 6). Bangalore, India: New Horizon College of Engineering.

[21]   LI YANG, (. I. (2020). Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. *Received December 26, 2019, accepted January 15, 2020, date of publication January 27, 2020, date of current version February 6, 2020.* (p. 9). Fuzhou 350118, China: Fujian University of Technology.

[22]   LIJUAN HUANG, Z. D. (2019). Online Sales Prediction: An Analysis With Dependency SCOR-Topic Sentiment Model. *IEEE Access* (p. 8). Hong Kong: The Hong Kong Polytechnic University.

[23]   Marius Ngaboyamahina, S. Y. (2019). The Impact of Sentiment Analysis on Social Media to Assess Customer Satisfaction: Case of Rwanda. *the 4th IEEE International Conference on Big Data Analytics* (p. 4). Kobe, Japan : Kobe Institute of Computing .

[24]   Monali Bordoloi, D. S. (2017). E- Commerce Sentiment Analysis using Graph Based Approach. *Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017), IEEE Xplore Compliant - Part Number: CFP17L34-ART, ISBN: 978-1-5386-4031-9* (p. 6). Assam, India.: N.I.T. Silchar.

[25]   Nitu Kumari, D. S. (2016). Sentiment analysis on E-commerce application by using opinion mining. *6th International Conference - Cloud System and Big Data Engineering (Confluence)* (p. 6). Noida, India: IEEE.

[26]   Paitoon Porntrakoon, C. M. (2018). Thai Sentiment Analysis for Consumer's Review in Multiple Dimensions Using Sentiment Compensation Technique (SenseComp). *15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (p. 4). Thailand: Assumption University Samuthprakarn.

[27]   Qing Sun, J. N. (2016). Research on Semantic Orientation Classification of Chinese Online Product Reviews Based on Multi-aspect Sentiment Analysis. *IEEE/ACM 3rd International Conference on Big Data Computing, Applications and Technologies* (p. 6). Beihang: Beihang University.

[28]   Renata Lopes Rosa, D. Z. (2013). SentiMeter-Br: a Social Web Analysis Tool to Discover Consumers' Sentiment. *IEEE 14th International Conference on Mobile Data Management* (p. 3). Brazil: University of Sao Paulo, SP - Brazil.

[29]   Satuluri Vanaja, M. B. (2018). Aspect-Level Sentiment Analysis on E-Commerce Data. *Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018)* (p. 5). Bengaluru: Amrita School of Engineering.

[30]   Upma Kumari, D. A. (2017). Sentiment analysis of smart phone product review using SVM classification technique. *International Conference on Energy, Communication, Data Analytics and Soft Computing* (p. 6). Kota, India: Rajasthan Technical University.

[31]   Data Driven Sales Prediction Using Communication Sentiment Analysis in B2B CRM Systems. (2019). 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC) (p. 8). Timisoara, Romania: West University of Timisoara.

[32]   Kai Gao1, S. S.-s. (2015). A sentiment analysis hybrid approach for microblogging and E-commerce corpus. 7th International Conference on Modelling, Identification and Control (ICMIC 2015) (p. 6). Sousse, Tunisia: Sousse, Tunisia.

[33]   Neena Devasia, R. S. (2016). Feature extracted sentiment analysis of customer product reviews. International Conference on Emerging Technological Trends [ICETT] (p. 6). Kollam, India: TKM College of Engineering.

[34]   Teerawat Polsawat, N. A.-i.-i. (2018). Sentiment Analysis Process for Product's Customer Reviews Using Ontology-Based Approach. International Conference on System Science and Engineering (ICSSE) (p. 6). New Taipei, Taiwan: IEEE.

[35]   S. Khan, K. Chopra and P. Sharma, "Brand Review Prediction using User Sentiments: Machine Learning Algorithm," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-8, doi: 10.1109/IDEA49133.2020.9170730.

[36]   Yi Liu*, Jiahuan Lu, Jie Yang and Feng Mao, " Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax", MBE, 17 (6): 7819–7837, DOI: 10.3934/mbe.2020398

[37]   M. Marong, N. K Batcha and R. Mafas, "Sentiment Analysis in E-Commerce: A Review on The Techniques and Algorithms" Journal of Applied Technology and Innovation (e -ISSN: 2600-7304) vol. 4, no. 1, (2020)

[38]   A. S. M. AlQahtani, "Product Sentiment Analysis For Amazon Reviews" International Journal of Computer Science & Information Technology Vol 13, No 3, June 2021, pp. 15-30, doi: 10.5121/ijcsit.2021.13302

[39]   B. Kumar Jha, , S. G.G, V. K.R, "Sentiment Analysis for E-Commerce Products Using Natural Language Processing", Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, pp. 166-175 Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021

[40]   N. Chen, " E-Commerce Brand Ranking Algorithm Based on User Evaluation and Sentiment Analysis", Front. Psychol, Sec. Emotion ScienceVolume 13 - 2022 | https://doi.org/10.3389/fpsyg.2022.907818

[41]   H. huang , A. asemi zavareh and M. begum mustafa " Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions Received 9 July 2023, accepted 15 August 2023, date of publication 21 August 2023, date of current version 29 August 2023. Digital Object Identifier 10.1109/ACCESS.2023.330730

[42]   Walaa Medhat a, *. A. (2014). Sentiment analysis algorithms and applications: A Survey. *www.elsevier.com/locate/asej* (p. 21). Egypt: Ain Shams University.

[43]   Mitra, J. S. (2012). Development of a Novel Algorithm for Sentiment Development of a Novel Algorithm for Sentiment Combinations . *IEEE* (p. 3). Kolkata : Dept. of Computer Science & Engineering, Jadavpur University, Kolkata .

[44]   Whitelaw Casey, Garg Navendu, Argamon Shlomo. Using appraisal groups for sentiment analysis. In: Proceedings of the ACM SIGIR Conference on Information and Knowledge Management (CIKM); 2005. p. 625–31.

[45]   Aggarwal Charu C, Zhai Cheng Xiang. Mining Text Data. Springer New York Dordrecht Heidelberg London: Springer Science+Business Media, LLC'12; 2012.

[46]   Yelena Mejova, Padmini Srinivasan. Exploring feature definition and selection for sentiment classifiers. In: Proceedings of the fifth international AAAI conference on weblogs and social media; 2011.

[47]   Diana Maynard, Adam Funk. Automatic detection of political opinions in tweets. In: Proceedings of the 8th international conference on the semantic web, ESWC'11; 2011. p. 88–99.

[48]   Tsytsarau Mikalai, Palpanas Themis. Survey on mining subjective data on the web. Data Min Knowledge Discovery 2012;24:478–514.

[49]   Plutchik R. A general psychoevolutionary theory of emotion. Emotion: Theory Res Exp 1980;1:3–33.

Weblinks:

[50] www.internetlivestats.com › twitter-statistics - date: 25-07-2020 at 5:00 p.m.

[51] https://marketingland.com/facebook-3-2-billion-likes-comments-every-day-19978 - date: 25-07-2020 at 5:00 p.m.

A study of Patient Satisfaction in Government Hospitals in Mumbai

[1]Jay Bhatt, Dr. [2]Amit P. Oak

Research Scholar

[1]School of Commerce & Management, Career Point University, Kota(Raj.), India

[2]Professor, Department of Management Studies

VIT, Wadala (E), Mumbai

Abstract:

Health care sector in India is witnessing rapid transformation due to the entry of private sector participation. However, a large majority of the poor still depend on government hospitals for their medical treatment. As a result, there is heavy pressure on the government hospitals including on OPD services in giving timely and quality services.

Patients' degree of satisfaction is one indicator of the quality and effectiveness of the healthcare services. Hospitals now a days have become more customer centric and try to provide better services to the patients in private as well as public hospitals. However, private hospitals are not affordable to be poorer section of the society.

The present research is a cross sectional study among 100 OPD patients of a government hospitals in Mumbai. A five points Likert scale was used to evaluate the level of satisfaction of the impatient in the OPD section. The findings of the research revealed that the overall level of satisfaction is improving.

However, there are still major challenges faced by the patients in the OPD in public hospitals. This includes delay in registration of patients, more waiting time in the queue, lack of sufficient doctors, lack of facilities in the waiting area and so on.

Hence, hospitals should apply the Queuing Theory of operation research to help the patients to get better services and reduce waiting period and higher level of customer satisfaction.

Keywords: Health Care, Inpatient, Outpatient, OPD, Service Quality

1. **Introduction:**

The health care in Mumbai includes government hospitals and private hospitals. There are large number of government hospitals such as J J Hospital, St. George Hospital, Cama Hospital, G.T. Hospital. Besides government hospitals, there are hospitals managed by Municipal Corporation which include primary and secondary health care centres. There are large number of private hospitals such as Hinduja Hospital, Jaslok Hospital, Breach Candy Hospital, Bombay Hospital, Lilavati Hospital, Nanavati Hospital, H N Hospital, Apollo Hospital, Reliance Hospital, Wockhardt Hospital, Kohinoor Hospital, Hiranandani Hospital and so on.

The total number of hospital beds in Mumbai are 24984 which include 13689 private and 11245 public. Thus, private hospitals account for 55 percent of total bed capacity and public hospital account for 45 percent of the overall bed capacity.

However, the existing facilities of health care are not adequate as compared to the population of Mumbai city. The health care services provided by the private hospitals are the expensive and beyond the capacity of the weaker section of the society. Thus, the weaker section and the low-income bracket people depend on the public hospitals which provide free health care facilities.

This has created heavy pressure on the public hospitals as against deficient infrastructure, unmanageable inpatient, deficient staff including doctors and poor quality of services. Against this background, research is carried out at Mumbai government hospitals (J J Hospital) to assess the level of satisfaction experienced by the patient in the OPD.

2. **Research Objectives:** The research objective of the research are:
   a. To find out the current status of the healthcare facilities in Mumbai cities
   b. To assess the level of patient satisfaction in the OCD a public hospital 3rd to ascertain the issues and challenges of the patient OPD
   c. To suggest measures to improve the healthcare services and enhance inpatient satisfaction

3. **Review of Literature:**
   1. **Dawal D Desai (2012)** found that the healthcare system in the city of Mumbai is deteriorating despite the affordable healthcare in offered by state government and

Mumbai Municipal Corporation. One of the major problems is the large number of people living in the slums and poor economic conditions. There is rapid increase in the migrants in the city over the period of time which has resulted in shortage of healthcare services and other facilities.

2. **Vikas Bajpai (2014)** highlighted the issues and challenges faced by public hospitals in the country. These includes shortage of infrastructure, shortage of medical staff, (Doctors, nurses and other staff), Unmanageable pressure of inpatient, poor service quality and higher out of pocket expenses.

   The total number of hospitals beds in the country are 1.37 million which include 5.40 lakhs beds under public hospitals, and 8.33 lakhs beds under private hospitals. In public hospital, 50% of beds are functional and another 50% are non-functional. In private hospitals, 70% beds are functional and 30% are non-functional.

3. **S. Glady Eswara Raj (2019)** Elaborated on the challenges faced by the public hospitals and private hospitals in India. A survey was conducted around 112 patients. The findings of the research revealed that, in case of public hospitals, the major problems are lack of personal attention, inadequate staff in case of emergencies, no proper consulting times, poor relationship with the patients, lack of immediate processing of bills, tips money demanded by the lower staff, poor quality of services, lack of basic facilities (Canteen, Parking, Waiting room, restroom for visitors). Unhygienic environment and lack of the use of the latest technology.

   In case of the private hospitals, the major problems are higher fees, (rank 1), discrimination in fees, charged (ranked two), long waiting time (rank three), long waiting time (rank three), lack of Operation theatres (rank four), poor consultancy services (rank five), lack of facilities (rank sixth), lack of use of the upto date technology (rank seventh) and so on.

4. **Sarwal Rakesh (2021)** focused on the concept of "Not for Profit Hospitals' in the country to bridge the gap in the health care sector. The share of government hospitals in urban areas is 35% and rural areas is 45.7%. While the share of private hospitals in the urban areas is 61.4% and rural areas 51.9%. The share of not for profit is only 2 to 3 percent.

The average medical care cost is Rs. 20134. In case of government hospitals, it is Rs. 4452 and in case of private hospitals it is Rs. 31845. In case of Not-for-Profit hospital, the cost is Rs. 24233.

The not-for-profit hospitals reduce their expenses due to lower salaries of doctors than the market rates (lower by 50% to 75%), lower salaries of other staff (by 20% - 30%) and by way of multitasking work force which leads to less number of manpower than the standards.

5. **Neha Taneja (2021)** conducted research to assess the patients' level of satisfaction in OPD hospital at government hospital, New Delhi. A cross sectional research was conducted among 547 patients which included 58.4 male and 41.1% female. The age group of the respondents vary between 18 years upto 70 years with the mean age of 35.34 years.

    19.2% of the respondents were first time visitor to OPD while 36.9% were visiting second time and 43.9% visited more than two times. 26.7% OPD visits were for surgery, 29.4% for the medicine, 11.3% for dermatology, 3.8% for pediatrics', 15.2% for orthopedic, 7.1% for ophthalmology and 6.4% for others.

    The findings of the survey shows that 86.3% of the patients were satisfied with the OPD services and 89.2% suggested that they will recommend this health care to others also. Majority of the patients were satisfied with respect to convenience (88%), accessibility (87%), proper signages (72.6%), facilities in the waiting areas (70.1%), cleanliness (67.9%) and so on.

6. **Kiran Kumar (2020)** conducted cross sectional research on ISO OPD patients. The findings of the research revealed the mean score of 8.67 for waiting time, 4.95 for cleanliness, 3.84 for staff behavior, 10.53 for doctor's behaviour and 5.73 for facilities. The mean score of patients OPD level of satisfaction was 33.72 with 29 minimum score and 38 maximum score. Thus, there is higher level of patient satisfaction from OPD services.

4. **Research Methodology:** The details regarding research design can be seen as below:

    a. Nature of Research Design: The present research design is exploratory in nature as it explores the level of satisfaction among the selected OPD patients in public hospital (JJ Hospital) in Mumbai.

b. Sources of Data: The research is based on the primary as well as the secondary data. The primary data was collected from the sample respondents selected for this purpose. The present research is a small component of the research work being undertaken by the research scholar for his Ph. D. thesis. The secondary data was collected from the review of literature and other published material on the subject.

c. Sampling Methods and Size: The methods of sampling used in the research are the non-probability sampling (non-random sampling) which include convenience sampling. The size of the sample was 100 respondents.

d. Instrument of Data Collection: The primary data was collected with the help of questionnaire which was based on six different variables.

e. Statistical Tools Used: The research is based on simple percentage and mean score for data analysis. Each variable of the research is presented in the form of table and graph (pie-diagram). Further on the basis of these tables and graphs inferences are drawn.

5.   **Data Analysis: The details regarding the data analysis can be seen as below:**

1. **Availability of Proper Information and Guidance**: Respondents were asked about the availability of proper information and guidance in the hospitals. The response in the regard can be seen from the following table.

Table No. 1

Availability of Proper Information and Guidance

Sr. No.	Description	Number	Percentage
1	Very Good	20	20
2	Good	30	30
3	Satisfactory	10	10
4	Poor	20	20
5	Very Poor	20	20
6	Mean Score	3.1	
	Total	100	100

Inferences: It can be observed that 20% of the respondents indicated very good. 30% of the respondents indicated Good. 10% of the respondents indicated satisfactory,

20% of the respondents indicated poor and 20% of the respondents very poor. The mean score was 3.10.

2. **Waiting Time at the Registration Window**: The details regarding the waiting time at the registration window can be seen from the following table.

Table No. 2

Waiting time at the registration window

Sr. No.	Description	Number	Percentage
1	More than One hour	20	20
2	Between 30 mins. to 1 hour	20	20
3	Half an hour	10	10
4	Twenty minutes	30	30
5	Less than Twenty minutes	20	20
6	Mean Score	2.9	
	Total	100	100

Waiting time at the registration window
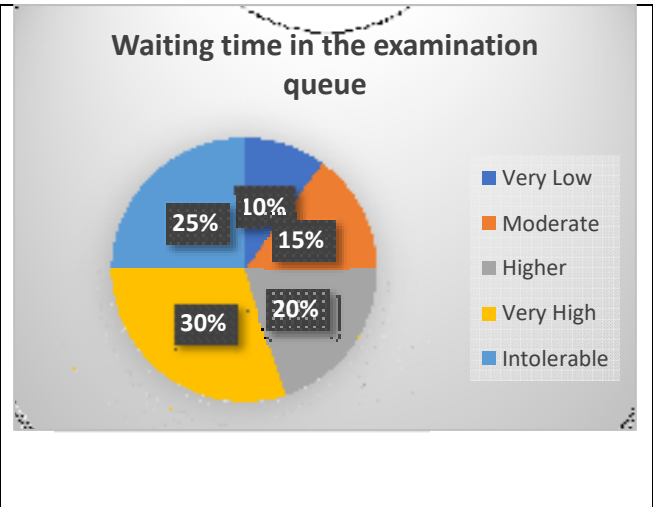
Inferences: It can be seen from the above table that there is higher waiting time at the registration ranging from half an hour to one hour. The mean score of this variable is 2.90.

3. **Waiting Time in the Examination queue**: The details regarding the waiting time required for the availing the health services can be seen from the following table.

Table No. 3

Waiting time in the examination queue

Sr. No.	Description	Number	Percentage
1	Very Low	10	10
2	Moderate	15	15
3	Higher	20	20
4	Very High	30	30
5	Intolerable	25	25
6	Mean Score	2.55	
	Total	100	100

Inference: It can be seen from the above table that 75 percent of the patients pointed out that the waiting time in the queue to check up is extremely high. This results in some patients wait and leave the place, jumping queue and infighting.

4. **Behaviour and Attitude of the Hospital Staff**: The behaviour and the attitude of the hospital staff can be seen from the table given below:

Table No. 4

Behaviour and Attitude of the Hospital Staff

Sr. No.	Description	Number	Percentage
1	Very Good	20	20
2	Good	10	10
3	Satisfactory	20	20
4	Bad	30	30
5	Very Bad	20	20
6	Mean Score	2.4	
	Total	100	100

Inference: It is seen from the above table that majority of the patients are not happy with the behaviour and attitude of the hospital staff. 50% of the patients indicated bad and very bad.
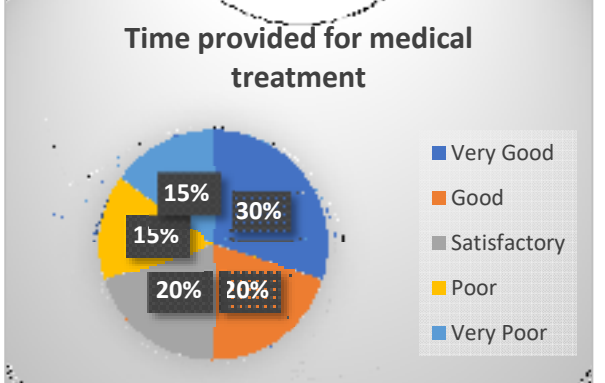
While 20 percent indicated satisfactory. Only 30 percent pointed out Good and Very Good. The mean score is 2.40 for this variable.

5. **Time provided for medical treatment**: The level of satisfaction regarding the time spent on medical treatment in OPD can be seen from the following table.

Table No. 5

Time provided for medical treatment

Sr. No.	Description	Number	Percentage	
1	Very Good	30	30	
2	Good	20	20	
3	Satisfactory	20	20	
4	Poor	15	15	
5	Very Poor	15	15	
6	Mean Score	3.35		
	Total	100	100	
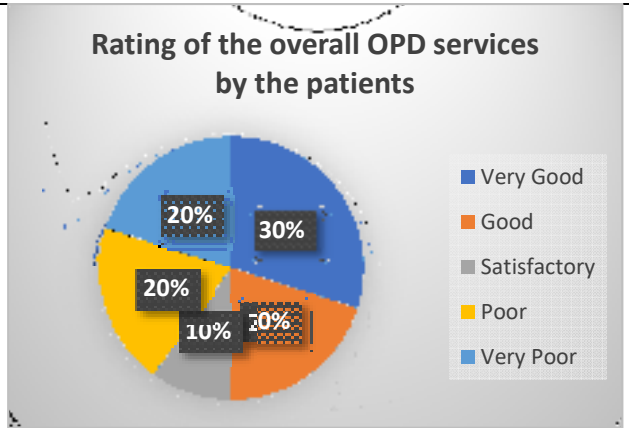
Inference: It can be seen from the above table that 50 percent of the patients are satisfied regarding the time provided for the health care treatment. Only 30 percent of the patients are not satisfied with the time given for treatment.

6. **Rating of the overall OPD services by the patients**: The overall level of the patient's satisfaction from the OPD services can be seen from the following table.

Table No. 6

Rating of the overall OPD services by the patients

Sr. No.	Description	Number	Percentage
1	Very Good	30	30
2	Good	20	20
3	Satisfactory	10	10
4	Poor	20	20
5	Very Poor	20	20
6	Mean Score	3.2	
	Total	100	100

Rating of the overall OPD services by the patients

Inference: It can be seen from the above table that the overall level of satisfaction of the patients from OPD services is satisfactory as the mean score is 3.20. 50 percent of the respondents indicated very good and good. While 40 percent indicated poor and very poor. It seems that the quality of services in the government hospitals are improving.

- **Summary and Conclusion:** The major findings of the research are summarized as below:

1. The health care system in Mumbai consists of government run hospitals, private managed hospitals, Municipal Hospitals, Community based hospitals, not for profit hospitals and so on. The existing facilities in terms of number of beds, doctors, supporting staff, infrastructure are not sufficient for the rising population or higher demand for health care.

2. The number of beds under the private sector are much higher than the public sector. As a result the share of private sector hospitals in the urban areas is 61% while that of the public sector is 35%. However, the private sector hospitals due to higher cost of medical treatment are not affordable for the poor and lower income group households.

3. As a result, the lower income group people depend on public hospitals for their medical care and treatment. This has resulted in higher level of pressure on the existing deficient infrastructure and facilities.

4. Government hospitals OPD get higher flow of patients each day beyond the existing capacity of the hospitals. Besides the patients face various challenges or problem is

getting timely health check. It includes longer time for registration, longer waiting time, lack of basic facilities in the waiting area, shortage of doctors and staff.

5. The present research focused on the six variables to assess the patients level of satisfaction from the OPD services in public or government hospital. The mean score of these six variables can be seen as below:

Information and Guidance	Registration Waiting Time	Waiting time check up	Staff Attitude	Time for check up	Overall Satisfaction
3.1	2.9	2.55	2.4	3.35	3.2

Hence the overall level of satisfaction of the OPD patients is satisfactory.

6. There is an urgent need to adopt scientific management tool like Queuing theory, Total Quality Management, CPM-PERT Analysis to reduce the registration waiting time and waiting time in queue. This will enhance the quality of health care services and further boost the level of patients' satisfaction. Public sector hospitals in Mumbai are rendering good services to the lower income group people J J Hospital is a name to reckon in the public sector hospitals in Mumbai city.

References:

1. Dhaval D Desai, Deteriorating public health care in Mumbai, Observer Research Foundation, June 2012

2. Vikas Bajpai, the challenges facing public hospitals in India, Hindawi Publishing Corporation, Advances in Public Health, 2014.

3. S Glady Eswara Raj, an analysis of problems faced by patients in government and private hospitals, International Journal of Business and Management Invention, Vol 8. Issue 4, April 2019.

4. Sarwal Rakesh, Study on the Not for Profit Hospital Model in Mumbai, NITI Ayog, 2021.

5. Neha Taneja, A study on patient satisfaction with service in outpatient department, Research Gate, December 2021.

6. Kiran Kumar, A study on the level of patient satisfaction on OPD Services, International Journal of Scientific Research, Vol. 9, Issue 1, January 2020.

Corrosion Inhibition of Urea and Thiourea on Carbon Steel in 2M Nitric Acid

[1]**Kriti Tripathi**, [2]**Dr. A.V. Trivedi**

[1]Research Scholar, School of Basic and Applied Science, Career Point University, Aalaniya, Kota, Rajasthan (India)

[2]Research Supervisor, School of Basic and Applied Science, Career Point University, Aalaniya Kota, Rajasthan (India)

Abstract:

Steel has been tested for corrosion at 30°C, 40°C, 50°C and 60°C in 2M nitric acid (HNO_3). For the current investigation, six time periods ranging from one to six hours were selected. The corrosion rate of carbon steel in 2M nitric acid varies significantly with time and temperature under examination from 21.246×10^{-5} gram cm^{-2} min^{-1} to 147.222×10^{-5} gram cm^{-2} min^{-1}. The findings demonstrate that HNO_3 has a significant impact on carbon steel corrosion at a concentration of 2M. Higher temperatures and longer exposure times have been shown to cause a noticeable increase in the rates of corrosion. We examined by weight the impact of adding organic molecules with an amino and thio group, like urea and thiourea, on the corrosion of carbon steel in nitric acid. The study's results demonstrate that urea and thiourea have a remarkable ability to control corrosion, with inhibition efficiencies ranging from 32.67% to 85.33%. But urea was discovered to be a stronger inhibitor than thiourea. Thiogroup therefore decreased the efficacy of inhibition.

Keywords: Carbon steel, Nitric acid, Corrosion inhibition, Urea, Thiourea.

Introduction:

Corrosion of metals and its alloys employed in service is an inescapable but a controllable process. One of the practical methods for controlling the impact of corrosion of metals and alloys especially in aqueous environments is by using corrosion inhibitors which are composed mainly of organic or inorganic substances. [1]. The adsorption of surfactant on metals and metal oxides creates a barrier that can inhibit corrosion [2] Thiadiazole, as a corrosion inhibitor for mild steel in 1 M hydrochloric acid has been investigated using weight

loss method and Scanning Electron Microscopy (SEM). Similarly The corrosion inhibition of mild steel in 0.5 M hydrochloric acid by six synthesized hetero cyclic compounds was studied using weight loss measurements. The inhibition efficiency exceeded 95%. The excellent inhibitor performance was attributed to the formation of protection adsorption films on the steel surface [3-4]. One of the most important methods in the corrosion protection of carbon steel is the use of organic inhibitors [1,5]. Most of the excellent acid inhibitors are organic compounds containing nitrogen, sulfur and oxygen [6-13] Different inhibitors were used to reduce the rate of corrosion in various alloys. The inhibition rates ranged between 30-80% in acidic environments with different molar concentrations. Natural-based corrosion inhibitors have gained great research interest thanks to their low cost and higher performance [14,15]. To date, more than 300 natural products have been isolated from different sections of the tree, with new compounds added to the list every year. As a contribution to the current interest on green corrosion inhibitors in a study application of plant extracts for metallic corrosion has been investigated. Nonane dihydrazide acts as a mixed inhibitor due to its adsorption on the MS surface, exhibiting an inhibition efficiency of more than 97% [16-19]. The inhibition efficiency of Eosin in controlling corrosion of carbon steel immersed in well water, has been evaluated by mass loss method both in absence and presence of zinc ion. A synergistic effect exists between Eosin (EN) and Zn^{2+} The inhibition efficiency (IE) of the EN – Zn^{2+} system decreases with increase in immersion period. Addition of N-cetyl-N,N,N-trimethyl ammonium bromide (CTAB), sodium dodecyl sulphate (SDS), sodium sulphite (Na_2SO_3) does not change the excellent inhibition efficiency of the EN-Zn^{2+} system [20]. The adsorption of n-PTH on the mild steel surface in both acidic media follows a Langmuir isotherm model [21].

Review of Literature:

Since long past metal corrosion is very very important field of research. Now researchers are regularly doing their work to find good corrosion inhibitors to stop big loss due to corrosion.Thus much have been done in this field and much more is yet to be done. Some recent studies are surveyed and given in this chapter. Fatech reviewed corrosion of copper and its alloys in corrosive environments and their corrosion inhibitors. The main corrosion inhibitor groups for copper are introduced and reviewed. Adsorption model has been provided for corrosion inhibitor activity. The main part of this work was to investigate

different corrosive environments for copper and its alloys and their corrosion inhibitors used in such environments to protect copper [21]. The increasing awareness of people on green materials and chemistry has stimulated researchers around the globe to turn towards natural resources. As a growing topic with limitless potential, plants as corrosion inhibitors (PCI) have received substantial attention due to their eco-friendly, biodegradable, inexpensive, and abundantly availability [22]. Traditionally, reduction of corrosion has been managed by various methods including cathodic protection, process control, reduction of the metal impurity content, and application of surface treatment techniques, as well as incorporation of suitable alloys. However, the use of corrosion inhibitors has proven to be the easiest and cheapest method for corrosion protection and prevention in acidic media. These inhibitors slow down the corrosion rate and thus prevent monetary losses due to metallic corrosion on industrial vessels, equipment, or surfaces [23]. In a review possibility of copper corrosion prevention has been discussed. Focus has been paid on the efficiency obtained using various organic compounds as corrosion inhibitors in numerous conditions. Several groups of compounds are found to be particularly important such as: azoles, purine and derivatives and amino acids. On the other hand plant extracts and natural products have also found prominent positions in these studies [24]. A review provides an outline of related litterateurs in which scientists and researchers used different types and procedure of corrosion inhibitors to reduce corrosion that takes place in various equipment made of alloys or metals. Different chemical inhibitors were used to reduce the rate of corrosion in various alloys. The inhibition rates ranged between 30-80% in acidic environments with different molar concentrations [25]. Corrosion of metals and its alloys employed in service is an inescapable but a controllable process. One of the practical methods for controlling the impact of corrosion of metals and alloys especially in aqueous environments is by using corrosion inhibitors which are composed mainly of organic or inorganic substances. However, the toxicity of organic and inorganic corrosion inhibitors to the environment and humans has compelled the search for safer corrosion inhibitors called 'green corrosion' inhibitor due to their properties like non-toxicity, biodegradability, and low cost [26]. Laurhydrazide N′-propan-3-one was used as an eco-friendly inhibitor for the corrosion of mild steel in 5M HCl at elevated temperatures. Various electrochemical techniques and surface characterization methods were utilized in this study. In addition, the kinetics and thermodynamic parameters were calculated and discussed. Furthermore, a geometry optimization of LHP was performed and the time dependent density

functional theory was utilized to calculate the electronic absorption spectra [27]. Natural-based corrosion inhibitors have gained great research interest thanks to their low cost and higher performance. The Cupressus arizonica fruit essential oil (CAFEO) has a higher extraction yield than leaves; however, it has less antibacterial and anti-fungal activities. The three main components in the CAFEO were a-pine (51.07%), myrcene (17.92%), and limousine (9.66%). Essential oils with a higher percentage of a-pinene were found to have outstanding corrosion inhibition properties [28]. Surfactants have been commonly used as corrosion inhibitors for the protection of metallic materials against corrosion. The amphiphilic nature of surfactant molecules creates an affinity for adsorption at interfaces such as metal/metal oxide–water interface. The adsorption of surfactant on metals and metal oxides creates a barrier that can inhibit corrosion. The properties of surfactant and the interaction of surfactant with metal or metal oxide and the surrounding solution environments determine the level of adsorption and corrosion inhibition [29]. A review summarizes the corrosion inhibition of steel materials in acidic media. Numerous corrosion inhibitors for steels in acidic solutions are presented. The emphasis is on HCl solutions, lower-grade steels, and elevated temperatures. This review is also devoted to corrosion inhibitor formulation design mixtures of corrosion inhibitors with surfactant, solvents, and intensifiers to improve the effectiveness of individual compounds at elevated temperatures [30]. The corrosion inhibition of mild steel in 5% HCl solutions by some new synthesized organic compounds namely 3-(2-methoxyphenyl) 5-mercapto-1. 2. 4-triazole (2-MMT), 3-(3- methoxyphenyl) 5-mercapto-1,2,4-triazole (3-MMT) and 3-(2-hydroxyphenyl) 5-mercapto-1. 2, 4-triazole (2-HMT) was investigated using weight loss and potentiostatic polarization techniques. These measurements reveal that the inhibition efficiency obtained by these compounds increased by increasing their concentration [31]. Novel pyrazine derivatives PD-1 were synthesized and their inhibitive action against the corrosion of mild steel in 15% HCl solution was studied by weight loss, potentio-dynamic polarization and electrochemical impedance spectroscopy (EIS) studies. Polarization studies showed that the studied inhibitors were of mixed type in nature. Scanning Electron Microscope (SEM) and Energy derisive X-ray spectroscopy (EDX) were performed for surface study of uninhibited and inhibited mild steel samples [32]. Corrosion has been chronic problem to industries and has increased cost of production. Lignin is the second most natural organic polymer on the earth and it can be obtained from the wastes of wood pulping processing in the form of black liquor. The inhibition efficiency

of lignin on the corrosion of mild steel in 1 M HCl have been evaluated by conventional weight loss method, and surface analysis using 500-5000 mg/L (w/v) inhibitor concentration in temperature range of 303-343 K. Maximum inhibition efficiency of 92.39 % was obtained with optimum inhibitor concentration of 4000 mg/L at 303 K [33]. Due to the increasing ecological awareness and environmental regulations, recent advancements in the field of corrosion science and technology is directed towards "green chemistry". One of the major factors that restrict the utilization of synthetic corrosion inhibitors is their toxic nature. The preparation of organic corrosion inhibitors is associated with the huge discharge of environmentally malignant chemicals into the surrounding environment. Commercially drugs present ideal candidature to replace these traditional toxic corrosion inhibitors [34]. Acetophenone derivatives are eco-friendly corrosion inhibitors to prevent corrosion of mild steel (MS) in acidic medium. The inhibition effect of 3-nitroacetophenone (3-NA) on the corrosion of MS in acidic medium (1 N H_4Cl) was investigated using weight loss measurements, electrochemical measurements, scanning electron microscopy, energy-dispersive X-ray spectroscopy, and quantum chemistry analysis [35].

Research Methodology:

The weight loss method was used to examine how carbon steel corrosion behaved in 2M nitric acid (HNO_3) at three distinct temperatures: 30°C, 40°C, 50°C and 60°C. The weight loss approach was also used to study the inhibition efficiency of urea.

Preparation of the specimen sample:

High purity carbon steel specimens were prepared by a 30 cm × 30 cm carbon steel plate with a 0.3 cm thickness. The plate was bought and chopped into 2 cm by 2 cm pieces. Every carbon steel sample was carefully cleaned before being exposed to make sure there were no surface contaminants preventing the corrosion process.

Preparation of 2M nitric acid solution:

Extreme caution was used when creating a 2M nitric acid solution to accurately mimic the severe corrosive environment. Fumigating nitric acid with an AR grade purity of 15.88 M was used to create a 2M nitric acid solution, simulating conditions that could be encountered in certain industrial operations. A concentration of two mole (2 M) was chosen to symbolize a somewhat strong acidic state. The acid solution was prepared using distilled water and high-purity nitric acid to ensure the accuracy of the concentration.

Experimental Configuration:

The setup for the experiment consisted of a controlled environmental chamber that could maintain temperatures between 30, 40, 50 and 60 degrees Celsius. Carbon steel specimens were dipped in the 2M nitric acid solution for a duration ranging from one to six hours.

Measurement of Corrosion Rate:

After every pre-planned exposure interval, the carbon steel specimens were cautiously removed from the nitric acid solution. A second measurement of the mass loss attributed to corrosion was made after the specimens had been cleared of corrosion products. Next, the following equation was used to calculate the corrosion rates from dividing the mass loss by the product of the exposed surface area and the exposure period:

$$\text{Corrosion Rate} = \Delta W / At$$

Where,

ΔW = difference between initial and final weight

A = area of cross section

t = time taken for corrosion

Measurement of % Inhibition:

% Inhibition measured by the following equation:

$$\% \text{ IE} = (CR_{Blank} - CR_{Inh.} / CR_{Blank}) \times 100$$

Where, CR_{blank} and $CR_{inh.}$ are the corrosion rates in the absence and presence of the different inhibitors at different concentration.

Research Findings:

Results of the present study have been given in following table 1 to 6 and discussed accordingly.

Table 1   Corrosion of Carbon Steel in 2M HNO₃ at 30°C

S.No.	Time	Initial Weight	Final Weight	ΔW	Rate of

	(Minute)	(W_1)	(W_2)	(W_1-W_2)	corrosion $\times 10^{-5}$gm cm^{-2} min^{-1}
1	60	9.967	9.888	0.079	21.236
2	120	9.942	9.704	0.238	31.989
3	180	9.426	8.950	0.476	42.652
4	240	9.829	8.998	0.831	55.846
5	300	9.954	8.668	1.286	69.139
6	360	9.211	7.372	1.839	82.392

Table 2   Corrosion of Carbon Steel in 2M HNO₃ at 40°C

S.No.	Time (Minute)	Initial Weight (W_1)	Final Weight (W_2)	ΔW (W_1-W_2)	Rate of corrosion $\times 10^{-5}$gm cm^{-2} min^{-1}
1	60	7.118	6.997	0.121	32.526
2	120	7.078	6.780	0.298	40.053
3	180	8.951	8.379	0.572	51.254
4	240	8.962	7.960	1.002	67.338
5	300	8.812	7.154	1.658	89.139
6	360	8.659	5.794	2.865	128.360

Table 3   Corrosion of Carbon Steel in 2M HNO₃ at 50°C

S.No.	Time (Minute)	Initial Weight	Final Weight	ΔW (W_1-W_2)	Rate of corrosion $\times 10^{-5}$gm cm^{-2} min^{-1}

		(W₁)	(W₂)		
1	60	6.087	5.891	0.196	52.688
2	120	5.692	5.276	0.416	55.913
3	180	6.291	5.485	0.806	72.222
4	240	5.669	4.275	1.394	93.682
5	300	6.098	3.926	2.172	116.774
6	360	8.923	5.937	2.986	133.781

Table 4   Corrosion of Carbon Steel in 2M HNO_3 at 60°C

S.No.	Time (Minute)	Initial Weight (W₁)	Final Weight (W₂)	ΔW (W₁-W₂)	Rate of corrosion × 10^{-5}gm cm^{-2} min^{-1}
1	60	6.785	6.491	0.294	79.032
2	120	6.965	6.361	0.604	81.827
3	180	6.986	5.927	1.059	94.892
4	240	6.948	5.281	1.667	112.029
5	300	6.829	4.540	2.289	123.064
6	360	6.997	3.711	3.286	147.222

Table 5   Corrosion inhibition of Urea on Carbon Steel in 2M HNO_3 at 30°C (Corrosion Time 4 hour)

S. No.	Concentration of Urea (Mole)	Initial Weight (W$_1$)	Final Weight (W$_2$)	ΔW (W$_1$-W$_2$)	Corrosion Rate CR × 10^{-5}gm cm^{-2} min^{-1}	% Inhibition (IE%)
1	0.0	9.829	8.335	1.494	100.40	0
2	0.2	3.953	3.783	0.17	11.42	88.62
3	0.4	5.858	5.73	0.128	8.6	91.43
4	0.6	6.876	6.78	0.096	6.45	93.57
5	0.8	6.443	6.409	0.034	2.28	97.72

Table.6   Corrosion inhibition of Thiourea on Carbon Steel in 2M HNO$_3$ at 30°C (Corrosion Time 4 hour)

S. No.	Concentration of Thiourea (Mole)	Initial Weight (W$_1$)	Final Weight (W$_2$)	ΔW (W$_1$-W$_2$)	Corrosion Rate CR × 10^{-5}gm cm^{-2} min^{-1}	% Inhibition (IE %)
1	0.0	9.83	8.33	1.50	100.81	0
2	0.2	3.78	2.77	1.01	67.87	32.67
3	0.4	5.73	4.96	0.77	51.74	48.66
4	0.6	6.78	6.19	0.59	39.65	60.65
5	0.8	6.41	6.19	0.22	14.78	85.33

Discussion:

The results of the present study have been given in above table 1 to 6. It is clear from the tables 1, 2, 3 and 4 that 2M nitric acid corrode steel significantly with rate of corrosion ranging from 21.236×10^{-5}gm cm^{-2} min^{-1} to 147.222×10^{-5}gm cm^{-2} min^{-1}. Minimum rate of corrosion was observed at 30°C temperature in 1 hour exposure time and the maximum rate of corrosion was observed at 60°C temperature in 6 hours exposure time period. It is very clear from table 1 to 4 that corrosion rate increases with exposure time and temperature.

Corrosion Inhibition Effect of Urea and Thiourea:

Tables 5 and 6 exhibit the suppression of corrosion of urea and thiourea. Table 5 makes it evident that urea considerably reduces the rate of corrosion with a maximum percentage of 97.72% inhibition. Thiourea also exhibits noticeably strong inhibition at a maximum of 85.33%. The results on inhibition shows that urea is a more potent inhibitor than thiourea, and that the thio group decreases the efficacy of inhibition.

Conclusion & Future Scope:

The collected data shows a direct relationship between the rates of carbon steel corrosion in 2M nitric acid with temperature and exposure time. The trends that have been discovered offer significant understanding of the fundamental mechanisms that control the corrosion process. They also highlight the complex interaction between temperature-dependent kinetics and acid concentration. The experimental approach employed provides a strong basis for comprehending the complex corrosion behaviour of steel in nitric acid. This research contribute to the field of corrosion science and materials engineering in general as well as to the specific understanding of steel corrosion. These insights can be used to develop corrosion-resistant materials and practical preventative techniques for industrial applications in future.

References:

1. Peter, A., et al. (2015). Use of natural gums as green corrosion inhibitors: an overview. International Journal of Industrial Chemistry, 6, 53–164.

2. Michael, Y. Z., & Free, L. (2017). A review of surfactants as corrosion inhibitors and associated modelling. Progress in Materials Science, 90, 159-223.

3. Al-Baghdad, S. B., et al. (2016). Thiadiazole as a Potential Corrosion Inhibitor for Mild Steel in 1 M HCl. Journal of Advanced Electrochemistry, 2, 67–69.

4. Ahmed, S. K., et al. (2019). Synthesis and investigations of heterocyclic compounds as corrosion inhibitors for mild steel in hydrochloric acid. International Journal of Industrial Chemistry, 10, 159–173.

5. Chen, L., et al. (2015). Organic Compounds as Corrosion Inhibitors for Carbon Steel in HCl Solution. MDPI, 15, 1-59.

6. Ouici, H. B., et al. (2014). Corrosion Inhibition of Mild Steel in Acidic Media Using Newly Synthesized Heterocyclic Organic Molecules. Correlation Between Inhibition Efficiency and Chemical Structure, 4, 20086-20095.

7. Zhang, S., et al. (2016). New Pyrazine Derivatives as Efficient Inhibitors on Mild Steel Corrosion in Hydrochloric Medium. Chemical Engineering Transactions, 55, 289-292.

8. Nwosu, F. O., et al. (2016). Thermodynamic and Adsorption Studies of Corrosion Inhibition of Mild Steel Using Lignin from Siam Weed (Chromolaena odorata) in Acid Medium. J. Mater. Environ. Sci., 7, 1663-1673.

9. Verma, C., et al. (2017). Drugs as environmentally benign corrosion inhibitors for ferrous and nonferrous materials in acid environment. Journal of Materials and Environmental Sciences, 8, 4040-4051.

10. Ibrahim, S., et al. (2022). Effect of 3-Nitroacetophenone on Corrosion Inhibition of Mild Steel in Acidic Medium. Hindawi International Journal of Photoenergy, 22, 1-9.

11. Fouda, A. S., et al. (2017). Role of Some Organic Compounds as Corrosion Inhibitors for 316L Stainless Steel in 1 M HCl. Int. J. Electrochem. Sci., 12, 347–362.

12. Al-Baghdad, S. B., et al. (2016). Thiadiazole as a Potential Corrosion Inhibitor for Mild Steel in 1 M HCl. Journal of Advanced Electrochemistry, 2, 67–69.

13. Ahmed, S. K., et al. (2019). Synthesis and investigations of heterocyclic compounds as corrosion inhibitors for mild steel in hydrochloric acid. International Journal of Industrial Chemistry, 10, 159–173.

14. Kadhim, A., et al. (2021). Corrosion inhibitors: A review. Int. J. Corros. Scale Inhibit, 10(1), 54–67.

15. Cherrad, S., et al. (2022). Cupressus arizonica fruit essential oil: A novel green inhibitor for acid corrosion of carbon steel. Arabian Journal of Chemistry, 15, 103849-103855.

16. Sharma, S. K., et al. (2015). Potential of Azadirachta indica as a green corrosion inhibitor against mild steel, aluminum, and tin. Journal of Analytical Science and Technology, 6, 1-26.

17. Tabtia, L., et al. (2018). Corrosion protection of mild steel by a new phosphonated Pyridines inhibitor system in HCl solution. Advances Engineering Forum, 36, 59-75.

18. Al Amiery, A., et al. (2022). Experimental and theoretical study on the corrosion inhibition of mild steel by nonanedioic acid derivative in hydrochloric acid solution. Scientific Reports, 12, 4705-4713.

19. Kendig, M., et al. (2003). Smart corrosion inhibiting coatings. Progress in Organic Coatings, 47(3–4), 183-189.

20. Sathiyabam, J., et al. (2009). Eosin as Corrosion Inhibitor for Carbon Steel in Well Water. The Open Corrosion Journal, 2, 77-82.

21. El Azhar, M., et al. (2001). Corrosion inhibition of mild steel by the new class of inhibitors [2,5-bis(n-pyridyl)-1,3,4-thiadiazoles] in acidic media. Corrosion Science, 43(12), 2229-2238.45. Fateh, A., et al. (2020). Review of corrosive environments for copper and its corrosion inhibitors. Arabian Journal of Chemistry, 13(2), 481–544.

22. 46. Mohd. Nasir, N. A. (2022). 25 years of progress on plants as corrosion inhibitors through a bibliometric analysis using the Scopus database. Arabian Journal of Chemistry, 15(3), 103655-103664.

23. Chigondo, M., et al. (2016). Recent Natural Corrosion Inhibitors for Mild Steel. Journal of Chemistry, 10(1), 1-7.

24. Petrović Mihajlović, M. B., et al. (2015). Copper Corrosion Inhibitors. International Journal of Electrochemical Science, 10(2), 1027-1053.

25. Kadhim, A., et al. (2021). Corrosion inhibitors. A review. International Journal of Corrosion and Scale Inhibition, 10(1), 54–67.

26. Peter, A., et al. (2015). Use of natural gums as green corrosion inhibitors: an overview. International Journal of Industrial Chemistry, 6(2), 153–164.

27. El-Haddad, M. M., et al. (2019). Highly efficient eco-friendly corrosion inhibitor for mild steel in 5M HCl at elevated temperatures: Experimental & molecular dynamics study. Scientific Reports, 9(1), 3695-3703.

28. Cherrad, S., et al. (2022). Cupressus arizonica fruit essential oil: A novel green inhibitor for acid corrosion of carbon steel. Arabian Journal of Chemistry, 15(2), 103849-103855.

29. Michael, Y. Z., & Free, L. (2017). A review of surfactants as corrosion inhibitors and associated modeling. Progress in Materials Science, 90, 159-223.

30. Finšgar, M. a., et al. (2014). Application of corrosion inhibitors for steels in acidic media for the oil and gas industry. Corrosion Science, 86, 17-41.

31. Ouici, H. B., et al. (2014). Corrosion Inhibition of Mild Steel in Acidic Media Using Newly Synthesized Heterocyclic Organic Molecules: Correlation Between Inhibition Efficiency and Chemical Structure. Corrosion Reviews, 4, 20086-20095.

32. Zhang, S., et al. (2016). New Pyrazine Derivatives as Efficient Inhibitors on Mild Steel Corrosion in Hydrochloric Medium. Chemical Engineering Transactions, 55, 289-292.

33. Nwosu, F. O., et al. (2016). Thermodynamic and Adsorption Studies of Corrosion Inhibition of Mild Steel Using Lignin from Siam Weed (Chromolaena odorata) in Acid Medium. Journal of Materials and Environmental Science, 7(5), 1663-1673.

34. Verma, C., et al. (2017). Drugs as environmentally benign corrosion inhibitors for ferrous and nonferrous materials in acid environment. Journal of Materials and Environmental Sciences, 8(11), 4040-4051.

35. Ibrahim, S., et al. (2022). Effect of 3-Nitroacetophenone on Corrosion Inhibition of Mild Steel in Acidic Medium. Hindawi International Journal of Photoenergy, 22, 1-9.

Cosmic Acceleration and Anisotropy: Probing the Dark Energy in Homogeneous and Anisotropic Universes

[1]Renu Bala, [2]Dr. Manoj Kumar Dhiman

[1]School of Basic Sciences, IEC University, Baddi, (H.P.), India

renubala3936@gmail.com

[2]School of Basic Sciences, IEC University, Baddi, (H.P.), India

m.manojdhiman@gmail.com

Abstract

Understanding the fast expansion of the cosmos in the late universe requires experimentally determining whether dark energy is a cosmological constant or dynamical. Assuming (reasonably) that dark energy existed before inflation, The late-time fast expansion's anisotropy should reveal the presence of a dynamical dark energy. One of the simplest dynamical models of dark energy that can be explained by empirical evidence, the quintessence field with exponential potential, is used to demonstrate this. We study the consequences of its quantum fluctuations, which are generated during inflation and perfectly positively associated with perturbations of the primordial curvature, and we determine the cosmic expansion anisotropy that this results in. The physical source of the disturbance is iso-curvature. This research looks at the f(T) theory of gravity as it pertains to the rapid expansion of the cosmos. Their holographic, new agegraphic, and power-law entropy adjusted dark energy models are used to produce a multitude of f(T) models. Equation of state parameters are used to study their cosmological behavior. Possible causes of cosmic acceleration include the attractive pull of dark energy, the vacuum's quantum energy, or the fact that general relativity (GR) fails on cosmological scales and has to be replaced. This research shows that the EoS for dark energy ó changes from one generation to the next, using space-times of the Bianchi types I, II, III, V, and VI. Some have proposed that the EoS's time-dependent background contribution to vacuum energy might be the source of dark energy, that explains the apparent exponential growth of the cosmos. The cosmological constant Λ is positively decreasing and reaches a modest positive value in the late universe, according to

recent investigations of type Ia supernovae, which supports this finding. The physical behavior of cosmological parameters including EoS parameter, skewness parameter, pressure, and energy density is explored. Additionally, we observed that our cosmological model mimics both the present-day quintessence model and the future CDM behavior after validating it using energy conditions.

Keywords: Cosmic Acceleration, Luminosity, Dark Energy, Quantum, Anisotropy

Introduction:

In addition to expanding, the present universe is speeding, according to the latest observational evidence, and this is all because of a mysterious and unidentified component known as dark energy (DE). Two competing models can account for the observed rate of cosmic expansion. One approach is to include general relativity into DE by introducing a system with a high negative pressure and positive density. Alternatively, this expansion might be explained by rethinking gravity. The cosmological constant, with $p = -\rho$, is the most basic kind of DE. But there are two issues with it, and they are called coincidence difficulties and fine-tuning problems. An equation of state (EoS) parameter, $\omega = p/\rho$, may be used to assess dark energy density. Several types of dynamically changing DE phases are associated with the negative behavior of the EoS parameter. Quintom models, which combine quintessence and phantom phases, may cross the phantom dividing line $\omega = -1$.

We define the HDE as

$$\rho\Lambda = 3c\,^2M^2\,pL\,^{-2} \quad (1)$$

$M_p = (8\pi G)^{-1/2}$, where G stands for the gravitational constant, L for the infrared cutoff, and $3c2$ for the sake of convenience, a dimensional constant. Using the cosmos' age as the infrared cutoff, we can get the initial ADE model. Nevertheless, it fails to elucidate the period when matter predominated. The NADE model was created as a result of substituting the conformal time η for the age of the universe. Thus, DE is derived from the fluctuations of the universe's spacetime and matter fields. One way to calculate its energy density is by

$$\rho\Lambda = 3n\,2M^2_p\eta\,^{-2}$$

This is used to parameterize specific uncertainties by defining the numerical factor $3n^2$. A substantial amount of focus has been placed on the enhanced gravity theories, which include scalar tensor theory, Brans-Dick theory, f(R), f(G), f(R, T), and f(T) in an effort to deduce the characteristics of DE. To generalize teleparallel theory of gravity (TPG), The general relativity's torsionless Levi-Civita connection is superseded by the curvature-free Weitzenbock connection in f(T) gravity.

Holographic Dark Energy Model

First, We implemented the f(T) gravity condition in the HDE model. Rh, the horizon for what lies ahead in a flat world, is equal to L and it may be represented as

$$R_h = B(t) \int_t^\infty \frac{dt}{B(t)} \tag{2}$$

For scale factors that resemble poles, it produces

$$R_h = B(t) \int_t^{t_s} \frac{dt}{B(t)}$$
$$= \frac{(t_s - t)}{h} = \frac{h}{h+1}\left(\frac{4m+2}{-T}\right)^{1/2} \tag{3}$$

So, the density of HDEs (4) changes.

$$\rho\Lambda = 3c^2 M^2{}_p L^{-2} \tag{4}$$

$M_p = (8\pi G)^{-1/2}$, where G stands for the gravitational constant, L for the infrared cutoff, and 3c2 for the sake of convenience, a dimensional constant.

$$\rho_\Lambda = -\frac{3c^2}{2\kappa^2(2m+1)}\left(\frac{h+1}{h}\right)^2 T \tag{5}$$

For the purpose of obtaining the f(T) gravity model, we may use the following HDE f(T) model by substituting the torsion density (6), denoted as $\rho\Lambda = \rho T$, for the HDE density (5), and vice versa.

$$\rho_T = \frac{1}{2\kappa^2}(2Tf_T - f - T)$$ (6)

f(T) = ϱ √T − (γ − 1)T (7)

The integration constant ϱ and the definition of γ are

$$\gamma = \frac{3c^2}{2m+1}\left(\frac{h+1}{h}\right)^2$$

For the flat universe, c is equal to 0.818, and here γ is dependent on that. For this model (7) to be considered realistic, the requirement that f(T) → 0 as T → 0 is satisfied.

Review of Literature:

Mishra, Bivudutta & Tripathy, S. & Ray, Pratik (2017) Dark energy and cosmic dynamics are examined in relation to anisotropic components. Anisotropy is thought to be a property of an anisotropic dark energy fluid, which shows pressure variations in different directions in space. Anisotropy is supplemented by a one-dimensional cosmic string that is oriented along the x-direction. As the universe expands, the isotropy of its dark energy pressure changes, at least in the distant past. The cosmic string anisotropy has a large impact on the early universe's acceleration dynamics.

Saha, Bijan. (2015) Our discussion revolves around aspects of Bianchi type universes When the EoS parameter in general relativity is anisotropic and the fluid in question, driven by mounting proof that a geometry similar to Bianchi morphology is necessary to account for the observed anisotropy in the WMAP data. There is good alignment of the seven axes and the direction Virgo in the ILC−WMAP data maps. This is why Bianchi models are so useful for anisotropies research.

Vinutha et. al (2020) A cosmic string fluid source in one dimension is integrated into this study's LRS Bianchi type-I dark energy cosmology model includes spatially homogenous states. A connection between metric potentials and the average scale factor hybrid expansion law allows for the solution of Einstein's field equations. The fast expansion of our model is

discussed via the use of the state equation (üde) and the deceleration parameter (q). Additional physical factors are also covered.

Research Methodology:

Theoretical Models for Dark Energy

i.   The ΛCDM Model

Cosmic acceleration in the late universe, as proposed by the cosmological constant Λ, is a widely accepted theory. To produce a motionless universe, Albert Einstein first postulated the cosmological constant in 1917. The importance of Λ seems to have diminished after Hubble's 1929 discovery of an expanding universe. The significance of Λ was reestablished in 1998 with the revelation of the Universe's accelerating expansion. The Bianchi identity ($\nabla\mu G\mu\nu = 0$) and energy conservation ($\nabla\mu T\mu\nu = 0$) are recognized as being obeyed by the energy momentum tensor $T\mu\nu$ and Einstein's tensor $G\mu\nu$, respectively. Equation 8 of Einstein's equations may include a new term $\Lambda g\mu\nu$ since the covariant derivative of the metric $g\,\mu\nu$ is zero.

The energy or matter existing in spacetime is connected to the curvature of spacetime by Einstein's equations, which are presented as,

$$G\mu\nu = 8\pi GT\mu\nu \quad (8)$$

Where,

$$G\mu\nu = R\mu\nu - 1\,2\,g\mu\nu R \quad (9)$$

In this context, the Einstein's tensor is denoted by $G\mu\nu$, In this context, $T\mu\nu$ stands for the energy momentum tensor, $R\mu\nu$ for the Ricci tensor, and R; for the Ricci scalar.

Here is the updated equation:

$$R\mu\nu - 1/2\,g\mu\nu R + \Lambda g\mu\nu = 8\pi GT\mu\nu \quad (10)$$

ii.   Quintessence Model

One of the most common alternative DE theories in literature is the scalar field model, that proposes a scalar field as the source of the cosmic acceleration in the late universe. In contrast to the ΛCDM scenario (ω = −1), the equation of state (EOS) of DE changes over time in these circumstances. Based on current findings, it is expected that the equation of state of DE, ωDE, should be around -1, however there is insufficient data to comment on how ωDE has evolved over time. This leads to the increasing use of scalar field models for DE. In addition to being extensively discussed in other domains like as particle theory, string theory, supergravity, etc., scalar fields have the potential to provide the negative pressure necessary for the Universe to expand at a faster pace. quintessence scalar fields are theories of canonical scalar fields with potentials. The action defines the classic scalar field paradigm.

$$S = S_{Ein} + S_{quint} + S_{matter}$$
$$= \int d^4x \sqrt{-g} \frac{R}{16\pi G} + \int d^4x \sqrt{-g} \left(-\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right)$$
$$+ \int d^4x \sqrt{-g} \mathcal{L}_m ,$$

$$(11)$$

In the above equation, SEin stands for Einstein's action, Squint symbolizes the action of the quintessence scalar field φ with potential V (φ), and Smatter stands for the action of the matter field.

Investigating Dark Energy via Observations

a. Supernova Type Ia

Important measurements for the study of DE come from type Ia Supernova (SNIa). In systems with two stars, SNIa may happen. When one star's mass exceeds its critical mass, which is called the Chandrasekhar's mass, the other star's mass is accelerated. When the inward gravitational force is greater than the outward electron degeneracy pressure, an imbalance arises above the critical mass limit, leading to a massive explosion known as a Supernova type Ia.

b. Cosmic Microwave Background Radiation

Even without looking at the Cosmic Microwave Background radiation (CMBR), it has been anticipated that DE exists. At $z \cong 1100$, when protons and electrons joined to create neutral hydrogen atoms, the photons were no longer connected to the electrons and the cosmic microwave background (CMB) was likely born.

c. Restrictions on Characteristics of Dark Energy Models

SNIa, CMB, BAO, and LSS observational findings might limit the DE model parameters. The findings of PLANCK 2018 show that, when adopting the ΛCDM model of DE and evaluating CMB and BAO data, $0.678 \leq \Omega^0_{DE} \leq 0.692$ and $-1.038 \leq \omega^0_{DE} \leq -0.884$.

Data Analysis & Interpretation:

- **Bianchi Type-I**

The element is considered to be a completely anisotropic Bianchi type-I-line, as provided by

$$ds^2 = -dt^2 + A^2\, dx^2 + B^2\, dy^2 + C^2\, dz^2, \quad (12)$$

where t is the only function that determines the metric potentials A, B, and C. The model will be homogenous in space if this is done. Finding the Equation-of-State (EoS) parameter for a perfect fluid on each spatial axis independently is a straightforward approach to generalizing it. while keeping the energy momentum tensor's diagonal shape compatible with the metric under consideration. Hence, the fluid's energy momentum tensor is assumed to be

$$T^j_i = diag\, [T^0_0, T^1_1, T^2_2, T^3_3] \quad (13)$$

- **Bianchi Type-II**

A homogeneous LRS Bianchi type-II space-time is being proposed for which

$$ds^2 = \eta_{ij}\, \theta^i\, \theta^j, \quad \eta_{ij} = (1, 1, 1, -1) \quad (14)$$

where θi is the Cartan base

$$\theta^1 = Adx, \quad \theta^2 = B(dy + x\, dz), \quad \theta^3 = Adz, \quad \theta^4 = dt, \quad (15)$$

This is because in LRS B-II, A and B are only time functions, $T_1{}^1 = T_2{}^2 = T_3{}^3$.

- **Isotropic Dark-Energy Bianchi Type-III Models**

The space-time of general Bianchi type-III is investigated with the help of the metric.

$$ds^2 = -dt^2 + A^2(t)\,dx^2 + B^2(t)e^{-2ax}dy^2 + C^2(t)\,dz^2, \quad (16)$$

The constant is represented by a.

Pradhan and Amirhashchi find the answers to the Einstein field equations. Following these steps, we can get the formulas for the metric coefficient:

$$C = (\ell + 1)^{\frac{1}{\ell+1}} \left[\frac{k_1}{\ell\ell_2}t + k_2 \right]^{\frac{1}{\ell+1}} \quad (17)$$

$$B = \ell_2(\ell + 1)^{\frac{\ell}{\ell+1}} \left[\frac{k_1}{\ell\ell_2}t + k_2 \right]^{\frac{\ell}{\ell+1}} \quad (18)$$

And

$$A = m\ell_2(\ell + 1)^{\frac{\ell}{\ell+1}} \left[\frac{k_1}{\ell\ell_2}t + k_2 \right]^{\frac{\ell}{\ell+1}} \quad (19)$$

Where m, k1, and k2 are integration constants and $\ell_2 = \ell_1^{\frac{1}{1-m_1}} m^\ell \quad \ell = \frac{m_1}{1-m_1}$

- **Bianchi Type-V**

Regarding the anisotropic Bianchi type-V that is spatially homogenous, we consider the space-time metric in the form

$$ds^2 = -dt^2 + A^2\,dx^2 + e^{2\alpha x}[B2\,dy^2 + C^2\,dz^2] \quad (20)$$

Where α is a constant and A, B, and C are metric potentials that depend only on cosmic time t.

- **Bianchi Type-VI0**

Here, we consider the uniformly anisotropic Bianchi type-VI0 line component, shown by

$$ds^2 = -dt^2 + A^2\, dx^2 + B^2\, e^2x\, dy^2 + C^2\, e^{-2x}\, dz^2, \quad (21)$$

Where t is the only function that determines the metric potentials A, B, and C. The model will be homogenous in space if this is done.

Astrophysical Disruptions and Power Spectrum by Examining the Dark Energy Slotheon Field Model

The numerical solution to the system of independent equations may be obtained by Using the previously stated starting points and applying the exponential form to the potential V (π) as given in Eq. 22, is obtained.

$$V(\pi) = V_0 \exp\left(-\frac{\lambda\pi}{M_{pl}}\right) \quad (22)$$

given that λ is a constant.

A Formula for the Density and State Parameters

$\Omega\pi (= \rho\,\overline{\pi}\,\rho c)$, the density parameter of the field π, is extracted from Einstein's equations for the background spacetime, the critical density of the universe, denoted as ρc, and the value for matter density, Ωm. $(= \rho\,\overline{m}\,\rho c)$. (Eqs. 23 - 25).

$$3M_{pl}^2 H^2 = \rho_m + \frac{\dot{\pi}^2}{2} + \frac{9H^2\dot{\pi}^2}{2M^2} + V(\pi) \quad (23)$$

$$M_{pl}^2(2\dot{H} + 3H^2) = -\frac{\dot{\pi}^2}{2} + V(\pi) + (2\dot{H} + 3H^2)\frac{\dot{\pi}^2}{2M^2} + \frac{2H\dot{\pi}\ddot{\pi}}{M^2} \quad (24)$$

$$0 = \ddot{\pi} + 3H\dot{\pi} + \frac{3H^2}{M^2}\left(\ddot{\pi} + 3H\dot{\pi} + \frac{2\dot{H}\dot{\pi}}{H}\right) + V_\pi \quad (25)$$

$$\Omega_\pi = y^2 + x^2\,(1 + 18\epsilon), \quad (26)$$

$$\Omega_m = 1 - y^2 - x^2 (1 + 18\,\epsilon), \quad (27)$$

in which Eqs. 28–32 define x, y, etc. You may see the results represented in Figure 1. At the early matter-dominated Universe, it is seen from Fig. 1(a) that $\Omega m = 1$ and $\Omega \pi = 0$. Figure 1(b) displays, for four distinct starting values of ϵ, how the EOS ϋπ of the Slotheon field π evolves in relation to z. $\epsilon i = 10^7, 2.5 \times 10^7, 4.5 \times 10^7, 6.5 \times 10^7$.

$$x = \frac{\dot{\pi}}{\sqrt{6} H M_{pl}} \quad (28)$$

$$y = \frac{\sqrt{V(\pi)}}{\sqrt{3} H M_{pl}} \quad (29)$$

$$\lambda = -M_{pl} \frac{V_\pi}{V(\pi)} \quad (30)$$

$$\epsilon = \frac{H^2}{2M^2} \quad (31)$$

$$q = \frac{\delta \pi}{\frac{d\pi}{dN}} \quad (32)$$

$$\frac{dx}{dN} = \frac{P}{\sqrt{6}} - x\frac{\dot{H}}{H^2} \quad (33)$$

$$\frac{dy}{dN} = -y\left(\sqrt{\frac{3}{2}}\lambda x + \frac{H}{H^2}\right) \quad (34)$$

$$\frac{d\epsilon}{dN} = 2\epsilon \frac{\dot{H}}{H^2} \quad (35)$$

$$\frac{d\lambda}{dN} = \sqrt{6} x \lambda^2 (1 - \Gamma) \quad (36)$$

$$\frac{dq}{dN} = q_1 \tag{37}$$

$$\frac{d\Phi}{dN} = \Phi_1 \tag{38}$$

$$\frac{dq_1}{dN} = \delta\pi_f - \frac{q\dot{H}}{xH^2}\frac{dx}{dN} - \frac{2q_1}{x}\frac{dx}{dN} - q_1\frac{\dot{H}}{H^2} \tag{39}$$

$$\frac{d\Phi_1}{dN} = \frac{\delta p_f}{4\epsilon} - 4\Phi_1 - 2\frac{\dot{H}}{H^2}\Phi - 3\Phi - \Phi_1\frac{\dot{H}}{H^2} \tag{40}$$
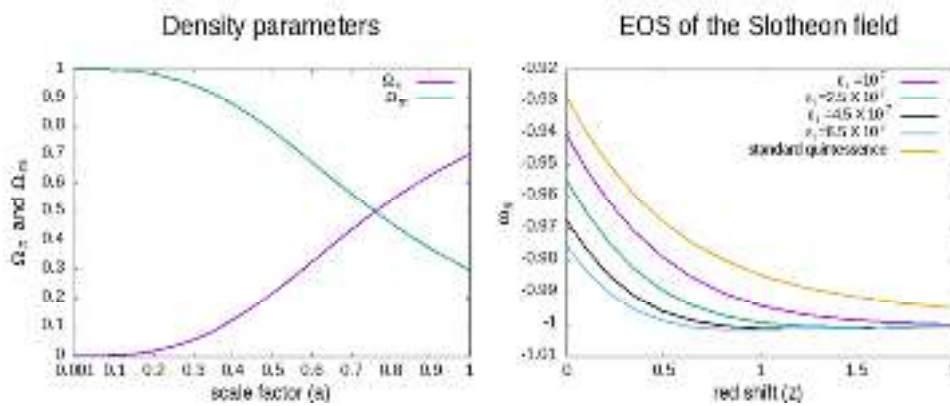
Figure 1: (a) The variations in the matter density parameter (Ωm) and the DE density parameter ($\Omega\pi$); (b) Changes in the Slotheon field's EOS parameters for various ϵi values

Slotheon Field Perturbations

Here we show that the matter field and the Slotheon field are perturbed cosmologically in the Newtonian or longitudinal gauge. Within this context, the scalar perturbed metric may be expressed as,

$$ds^2 = -(1 + 2\Phi)dt^2 + a^2(t)(1 + 2\Psi)\delta ij dx^i dx^j \tag{41}$$

The variables a(t), Φ, and Ψ represent the scale factor, gravitational potential, and disturbance in spatial curvature, respectively, in the given equation.

Research Findings:

The following is a synopsis of the thesis's main points. Theoretically and experimentally, the late-time cosmic acceleration may be best described by the scenario in which the system exhibits slower rolling behavior than the other possibilities. Therefore, compared to the generic quintessence model of DE, the Slotheon field model is better at explaining both theoretical and experimental findings. In addition to fixing the issues of the ΛCDM model, For the purpose of characterizing the late acceleration of our Universe, the Slotheon model may prove to be a suitable framework. The system may also become slower as a result of the extra frictional effects caused by the interplay of DE and DM. As a consequence, IDE models may be better able to handle phenomenological issues and are more congruent with experimental findings than non-interacting DE models. The thesis also examines the PBH mass distributions and finds that cosmological experimental evidence would tighten the bounds on PBH masses, PBH abundances, the proportion of dark matter that may be explained by PBHs, etc.

Conclusion:

Dark energy may take three forms throughout the cosmos, as shown in this article: quintessence, phantom, and quintom. The specific forms rely on the scale factors that are selected, which in turn are determined by the models' constant parameters. Our primary goal in doing the research presented here was to determine whether the results we obtained are in agreement with the most current findings from astrophysical observations. The outcomes acquired in various Bianchi type space-times align well with current findings, as we have discovered. One way to approach DE and the acceleration of the Universe in its late stages is to think about a scalar field model of DE, namely the Slotheon scalar field. Motivated by models operating at the dimensional cross-over limit, the theory is expanded to include curved spacetime, drawing inspiration from these models. After that, we numerically solve all of the perturbation equations. As a thawing type DE model, the Slotheon field DE model is regarded in this study. From the computations done in the chapter, the results obtained with the Slotheon DE model resemble the ΛCDM model more closely than those from the quintessence DE model, as may be inferred. However, the ΛCDM framework is unable to handle disturbances and inhomogeneities non-DE. The validity of the ΛCDM model might be determined by any of these perturbations, but even in this case, the Slotheon model of DE,

which has non-perturbative conclusions that are more comparable to ΛCDM, will remain valid and feasible.

Suggestions & Recommendations/Future Scope:

The two primary areas of cosmology discussed in this work are the phenomenology of dark energy, or the acceleration of the universe in recent times, and the evaporation of primordial black holes by means of Hawking radiation. Both of these will be crucial to our continued knowledge of the cosmos in the years to come. There is a lot of room for growth in these new domains, both in terms of theory and the outcomes of future tests. Therefore, additional study into these areas is desirable. Additionally, two other crucial areas to comprehend the development of the cosmos are the study of gravitational waves and the production of large-scale structures. In the late epoch, the quintessence model approaches isotropy. The phantom situation is seen for various models when n is less than 1.

Future Scope:

i. The impact of various DE models, cosmic acceleration models, and DE-DM interactions on cosmological observables such as the 21cm line, the CMB, BAO, RSD, and others is examined. Hence, these observables are used to investigate many aspects of the dark sector generally and to provide insight into the characteristics and sources of DE and DM specifically.

ii. Their potential marks on ongoing and future studies like PLANCK, SKA, EUCLID, LSST, DES, etc., should be investigated further in the future to learn more about the components of the dark sector. Building a functional model of DE-DM interaction that is well-motivated utilizing observational data from CMB, BAO, RSD, etc., might also be beneficial.

References:

1. Mishra, Bivudutta & Tripathy, S. & Ray, Pratik. (2017). Anisotropy in Dark Energy. arXiv:1701.08632v2 [physics.gen-ph].

2. Saha, Bijan. (2015). Accelerating dark energy models of the universe in anisotropic Bianchi type space-times and recent observations. Physics of Particles and Nuclei. 46. 10.1134/S1063779615030028.

3. Vinutha, T. & Velagapudi, Uma Maheswara Rao & Mengesha, Molla. (2020). Anisotropic Dark Energy Cosmological Model with Cosmic Strings. Canadian Journal of Physics. 99. 10.1139/cjp-2018-0306.

4. Amirhashchi H., Pradhan A., Zainuddin H. Interacting Two-Fluid Viscous Dark Energy Models in Non- at Universe // Res. Astron. Astrophys. 2013. V. 13. P. 129Ä 138.

5. Pradhan A. et al. Dark Energy Models with Anisotropic Fluid in Bianchi Type-VI0 Space-Time with Time-Dependent Deceleration Parameter // Astrophys. Space Sci. 2012. V. 337. P. 401Ä413.

6. Tripathy, S. & Mahanta, K.L. (2014). Cosmic Acceleration and Anisotropic models with Magnetic field. European Physical Journal Plus. 130. 30. 10.1140/epjp/i2015-15030-8.

7. Aluri, Pavan & Panda, Sukanta & Sharma, Manabendra & Thakur, Snigdha. (2013). Anisotropic universe with anisotropic sources. Journal of Cosmology and Astroparticle Physics. 2013. 10.1088/1475-7516/2013/12/003.

8. Hossienkhani, H. & Gheysari, Somayeh & Zarei, Zeinab. (2020). Anisotropic universe and observational constraint on the dark energy models with the cosmological redshift drift. New Astronomy. 84. 101465. 10.1016/j.newast.2020.101465.

9. Colin, Jacques & Mohayaee, Roya & Rameez, Mohamed & Sarkar, Sayak. (2019). Evidence for anisotropy of cosmic acceleration. Astronomy & Astrophysics. 631. 10.1051/0004-6361/201936373.

10. Saratchandra Singh, Moirangthem & Singh, S. (2019). Cosmological Dynamics of Anisotropic Dark Energy in f(R, T) Gravity. New Astronomy. 72. 10.1016/j.newast.2019.03.007.

11. Saha, Bijan. (2014). Isotropic and Anisotropic Dark Energy Models. Physics of Particles and Nuclei. 45. 349 - 396. 10.1134/S1063779614020026.

12. Harko, Tiberiu & Lobo, Francisco. (2013). Cosmological anisotropy from non-comoving dark matter and dark energy. Journal of Cosmology and Astroparticle Physics. 07. 036. 10.1088/1475-7516/2013/07/036.

Docker Performance and Multi-Container Applications

[1]Ms.Anisha Sheikh, [2]Dr.Jitendra Singh Chouhan, [3]Dr. Paras Kothari, [4]Mr.Ankur Rao

[1]Assistant Professor, Department of Computer Applications Geetanjali Institute of Techincal Studies, Udaipur, Rajasthan, India

anishasheikh1010@gmail.com

[2]Head Department of Computer Applications Geetanjali Institute of Techincal Studies, Udaipur, Rajasthan, India

chauhan.jitendra@live.com

[3]Professor Department of Computer Applications Geetanjali Institute of Techincal Studies, Udaipur, Rajasthan, India

paras.kothari@gits.ac.in

[4]Assistant Professor, Department of Computer Applications Geetanjali Institute of Techincal Studies, Udaipur, Rajasthan, India

ankurrao123@gmail.com

Abstract

Virtualization has been reshaped into a more agile alternative to traditional virtual machines (VMs) as a result of a revolutionary change that has taken place in the landscape of container technology. Docker is at the vanguard of this movement. It is a platform for container management that has gained widespread adoption. Docker's popularity has skyrocketed as a result of the automation it delivers to containerized applications. Docker is distinguished by its cutting-edge engine, which is completely integrated into the host operating system. This eliminates the need for a guest operating system, which in turn considerably improves the efficiency of the operations. Not only does this integration make operations more efficient, but it also leads to a significant decrease in the amount of resource overhead necessary. In part, the stratospheric ascent of Docker may be attributed to the growing preference for open-

source software over licensed alternatives. This tendency has been growing over the last several years. Taking an open-source approach is consistent with the philosophy of the software development community, which is one of the factors that has contributed to the broad acceptance of Docker. Docker, which is particularly interesting to developers and programmers, is a tool that shows to be vital in the construction and deployment of microservices on containerized systems with exceptional efficiency. The attractiveness of this product resides not only in its technical capabilities, but also in the general trend within the sector that favors open-source services and solutions. It is Docker's capacity to meet the ever-evolving requirements of a dynamic software development landscape that has contributed to its rise to prominence. Docker provides a multitude of benefits that are not available in conventional virtual machine settings.

Keywords: Docker, VM, OS, Container**,** SQL, VMware, CPU.

Introduction

Docker is a great tool for developers who are looking to achieve agility in their development processes. Docker is an open-source platform that is devoted to the creation, distribution, and execution of applications. Developers are given more control by this platform since it makes it possible to separate applications and programs from the infrastructure that supports them, which in turn makes it easier to release products quickly. Developers are able to effectively manage their infrastructure in a way that is comparable to the administration of applications whenever they use Docker. By using Docker's method for shipping, testing, and deploying code in a timely manner, the amount of time that passes between the generation of code and its implementation in a production environment is significantly reduced. Docker is not only one of the most dynamic container technologies, but it also has crucial qualities that enable it to offer programs with a runtime environment that is smooth throughout their execution. When contrasted with more conventional virtualization solutions, its unparalleled excellence becomes readily apparent.

Components of Docker

Docker is made up of four primary components that form its internal structure. These components are the Docker Client and Server, Images, Registries, and Containers [1].

A. Client and Server of Docker

The Docker server is the one that is responsible for receiving the request from the Docker client. This occurrence occurs in the same manner as any other regular normal server would. Following this, the server will proceed to process the request in the appropriate way. The transfer of a complete RESTful Application programming interface as well as command line client binaries is accomplished via the utilization of Docker. On the same computer system, it is possible for a Docker daemon or the Docker server and client to be executed. There is also the possibility of linking a local Docker client to a remote Docker daemon that is operating on a separate computer system.

B. Docker Images

Images created using Docker serve as the basis for any job that can be accomplished with the help of Docker. These pictures are regarded to be a snapshot of the work that is included inside a file. They contain all of the dependencies and parameters that are associated with the system. There are two distinct approaches that may be used in order to produce them. The first method involves making an image by using a template that can only be read by the individual. photos that serve as the basis for all other photos are called base images. These images of operating systems are the basis images, and they are used to construct a container that is capable of running the whole operating system. The process of creating new pictures in this manner is referred to as "committing a change."

C. Docker Containers

These are developed using the docker images. They are basically an application which is run in isolated manner and contains everything needed for deployment of application. It is possible to find a picture of "UbuntuOS" that includes the word "SQL SERVER." This image will be executed by using the run command of Docker, which will result in the creation of a container and the operation of SQL SERVER on UbuntuOS.

D. Docker Registeries

In order to store both its images and containers, Docker makes use of repositories that are referred to as registries. These registries operate similarly to source code repositories in that they allow for the simple pushing and pulling of images. There are two distinct categories of registries: public and private. In order to encourage an atmosphere that is both collaborative and accessible, public registries make it possible for anybody to access and submit

138

photographs, even if they have no previous expertise in the production of images. The Docker client is located on the left side of the schematic depiction of Docker's process and architecture that is illustrated in Figure 1. This client serves as the user interface that is mostly accessible via the command line. Docker build is the program that begins the process of creating an image from a Dockerfile. Docker pull is the command that obtains images from a registry, while Docker run is the operation that runs containers.

At its heart is the Docker host, which is comprised of the Docker daemon, which is accountable for the construction, operation, and management of containers, which are instances of Docker images that are now being executed. These containers enclose in themselves self-contained environments that include all of the code and dependencies that are required for the operation of an application. There is a Docker Registry located on the right side of the screen. This registry is responsible for hosting services such as Docker Hub and simplifying the storage and exchange of images. Docker is capable of containerizing a variety of technologies, including Node.js, MySQL, and Nginx, which are represented by the symbols in the figure. These technologies are discussed in relation to web development and deployment. With this all-encompassing configuration, Docker's adaptability is highlighted by the fact that it offers a streamlined and effective environment for the construction, deployment, and management of applications. As a result, it has become an indispensable instrument in contemporary software development methods.
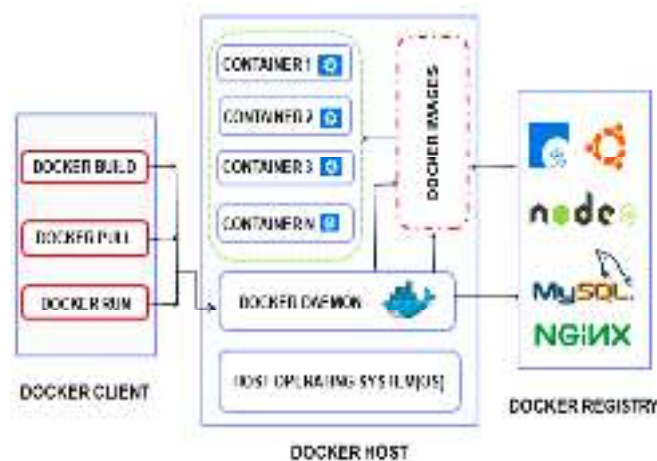
Figure 1 Docker Architecture

Docker is a platform that allows for the development, delivery, and operation of programs inside containers. its architecture is described here. Docker build, docker pull, and docker run are some of the commands that are issued by the Docker client, which is located on the left side of the image. This client communicates with the Docker host. There is a component

known as the Docker Daemon that is a part of the Docker host. This component is accountable for the creation and management of containers (Container 1 to Container N) that are separate from one another but executing on the same host operating system. For example, Docker Hub is an example of a registry that saves Docker images. This registry may be found on the right. Applications and services such as Node.js, MySQL, and Nginx may be included in these images, and they can be fetched into the Docker host in order to build containers. Through the use of containers, which make use of the kernel of the host system but otherwise function independently from one another, this architecture makes it possible for programs to run in an environment that is lightweight, portable, and self-sufficient.

Virtual Machine vs Docker

Cloud computing is built on the fundamental idea of virtualization, which is a well-established and essential element in the development of systems, the supply of resources, and the sharing of resources among several tenants. Cloud computing, which is now positioned as an essential component of Infrastructure as a Service (IaaS), makes use of virtualization in order to improve the adaptability and effectiveness of the system. When it comes to Virtual Machines (VMs), a hypervisor is an essential component that is located between the host operating system (Host OS) and the guest operating system (Guest OS). In other words, a hypervisor is there to act as a mediator between the two. The hypervisor is a virtual platform that acts between the operating system and the central processing unit (CPU). It is able to facilitate the management of several operating systems inside a single server. Nevertheless, owing to its intrinsic weight, which is further compounded by the requirements of the guest operating system, its proper implementation presents issues on computers that have very limited resources.

VMware was the company that pioneered virtualization technology, and Figure 2 depicts the architecture of a Virtual Machine, which is a depiction of that technology. The physical hardware, which is the actual server equipment, is the basis of this structure and serves as its foundation. In addition to this, the Host Operating System acts as the foundation for the VMware Hypervisor, which is also referred to as the Virtual Machine Monitor (VMM). In order to enable many instances of Guest Operating Systems to operate simultaneously on a single physical computer, the Hypervisor, which is an essential component, is entrusted with the responsibility of generating and maintaining virtual machines (VMs) that imitate real hardware. Each virtual machine (VM) functions as if it

were running on its own independent hardware system, and it is able to run its own applications. This configuration of virtualization makes it possible to use resources in a more effective manner, to increase scalability, and to effectively segregate a variety of operating systems and applications. Furthermore, it enables the simple creation of new virtual machines (VMs) or the modification of existing ones in order to fulfill certain needs that are unique to a given circumstance. The merger of cloud computing and virtualization emerges as a strong paradigm, altering system design and resource management in modern computer settings.
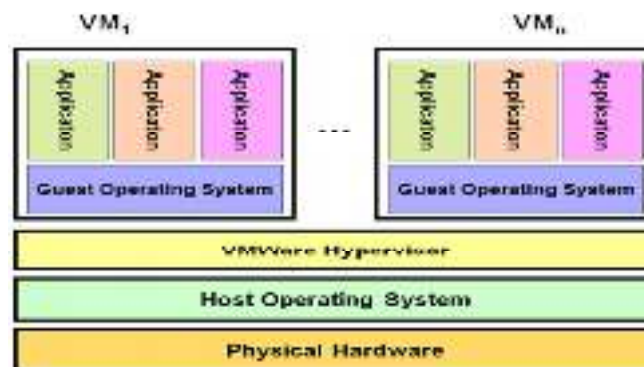
Figure 2 Architectural Design of Virtual Machines

VMware is a sort of hypervisor-based virtualization, and the structure of a virtualized environment that uses VMware are described here. Physical hardware, which includes the actual server along with its resources such as the central processing unit (CPU), memory, and storage, is located at the foundation. The Host Operating System, which is located above this layer, is the primary component that communicates directly with the actual hardware. It is possible to operate many virtual machines (VMs) simultaneously thanks to the VMware Hypervisor, which is sometimes referred to as the virtual machine monitor (VMM). This hypervisor is located on top of the host operating system. Every virtual machine (VM), which is denoted by the abbreviations VM1 through VMn, functions with its own Guest Operating System and programs, essentially imitating independent personal computers. Each of these virtual machines is completely separate from the others, which means that they are able to execute their own unique operating systems and programs without influencing the others. There is a frequent practice of using this configuration in order to maximize the usage of hardware, isolate applications, and improve the scalability and management of the applications that have been deployed.

While, this is not the case in the container-based virtualization as the requirement of a guestOS is not there. Containers of linux are looked into by the docker tool and also used for virtualization at the OS level. The illustration in Figure 3 demonstrates that a single component contains a number of Linux containers that are separated from one another. The kernel is responsible for the distribution of assets such as memory, CPU, block input/output, and network, and it also performs management using c-groups without waking up the virtual machine [2]. The containerization paradigm that Docker employs by default. The infrastructure, which may consist of actual servers or virtual computers, is often located at the foundation. The Host Operating System, which is anything that operates directly on the infrastructure, is located above that. Docker functions as a lightweight layer that sits on top of the host operating system and enables more than one application (App A through App F) to operate in their own individual containers. There are no distinct guest operating systems, in contrast to conventional virtualization; rather, all containers share the kernel of the host operating system, but they operate in user spaces that are completely different from one another. With this paradigm, programs are guaranteed to be portable, consistent, and simple to deploy. This is because each container has everything that the application need in order to function properly. The distribution of programs and their dependencies is thus simplified by Docker, which also minimizes overhead and increases resource usage in comparison to the practice of operating separate virtual machines (VMs) for each application.
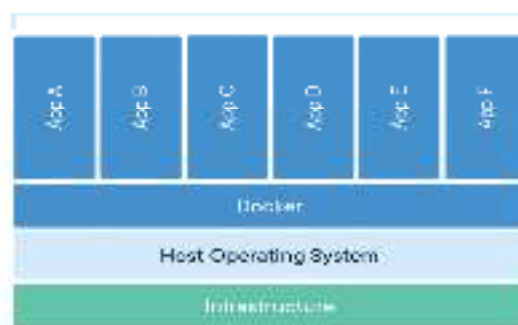
Figure 3 Containerized Applications

Docker provides a more straightforward perspective on containerization. A stack is shown, in which the infrastructure is responsible for providing the fundamental hardware resources. The Host Operating System, which is located above this, is the component that allows for direct interaction with the infrastructure. Docker functions as a lightweight layer that sits on top of the host operating system and enables many programs (App A through App F) to run

in distinct containers. These programs, in contrast to typical virtual machines, share the same host operating system kernel but are contained inside their own containers. This allows them to be more efficient, portable, and scalable without sacrificing reliability. This technique, which is based on containers, makes it possible for each program to be executed in isolated contexts with its own dependencies, while still making use of the same underlying system. This helps to reduce overhead and improve the operational efficiency of resource use.

Advantages of the docker

The docker is getting very popular quickly given its advantages and benefits. The major advantages of the docker containers are as follows

E. Portability

All the application dependencies are placed in a container which is portable across various platforms so the application is able to run on several devices such as Computers, laptops or even virtual machines.

F. Lightweight

Container technology is light when compared to the use of VM technology because VMs boot entire OS and hence consumes more resource.

G. Suitable for microservices architecture

Microservices are deployed without interference from other microservices with the help and support of container technology. Containers provide them a welcoming environment, which enables the deployment of these services in terms of isolation, the simplicity with which new versions can be deployed, and the speed with which they can be deployed.

H. Optimal Resource Utilization

The fact that it is lightweight, portable, efficient, and capable of running on actual servers allows for a greater number of containers to be run on physical servers than virtual machines, which ultimately results in a higher utilization of resources. Containers are composed of application code and its dependencies, and they operate as a separate process that shares the kernel with other containers that are located in the user space of the host operating system.

I. Density

As it does not use hypervisor, resources which are more efficiently accessible are used by the docker because the number of containers one can run on single host is more in case of docker

compare to VMs. Docker containers exhibits high performance because of their high density and minimal wastage of overhead resources.

J. Rapid Delivery

Containers have the ability to function in any environment since they are equipped with all of the essential dependencies that are associated with the programs, and they have all been tested and approved. Docker offers a stable, predictable, and enhanced environment, which enables the acquisition of evident outcomes when programs are transferred between development, testing, and creation systems.

K. Scalable

Docker has the capacity of being deployed in various platforms such as data servers, in cloud platforms and physical servers that too at great speeds. Transition at great speeds can also be made form the cloud environment to the local.

Docker Performance

In this particular section, performance of the applications in case of traditional virtualization and performance of the popularity gaining containerization are discussed. Seo et al. (2014) summarized that " absence of guestOS in docker results in less wastage of CPU and storage. The time taken for boot is less; no disturbance in images; and time taken for generation of images is less. When compared to virtual machine cloud, these are the advantages that come with using Docker Cloud. Two servers were used by them one each for docker and open stack platform KVM through virtualization, which were similar with same configuration in cloud environment [14].

L. Boot Time

For evaluation purpose of the approx.. boot-time,number of images were generated on each server is twenty and time take to boot is observed. Figure 4 illustrates that the time take by docket to boot is lesser when compared to KVM. Docker makes the use of the HostOS only while KVM make use of an guestOS which requires additional time to boot.Hence, boot time is lesser in case of docker [19].
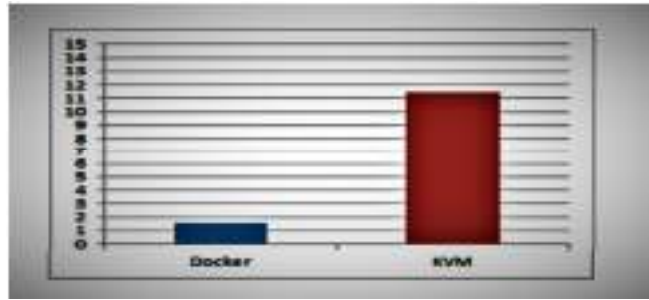
Figure 4 Boot Time

M. Operational Speed

For evaluation purpose of the operational speed, programming language used is "python". Figure 3 illustrates that it took around 4.5 seconds on an average for completing one lakh operation. They were able to determine the average amount of time required to finish activities by doing the identical procedure on Virtual Machine and Docker roughly one hundred times each[15]. This pace in the case of Docker is somewhat quicker when compared to the speed of Virtual machine, as shown in Figure 5, which proves this clearly.

Figure 5 Calculation Speed

" Felter et al. (2014) come to the conclusion that virtual machine systems (VMS) and containers are both well-established inventions that have benefitted from incremental improvements in programming and hardware over the last 10 years. This conclusion is reached after doing several comparisons on performance [4]. This conclusion was reached after the authors conducted a number of research studies.

The findings of Harrison John Bhatti (2017) indicate that Docker's execution is either similar to or beats that of KVM for every situation that we tested. According to the findings of our investigation, both KVM and docker exhibit an unnecessary amount of overhead for the execution of CPU and memory. As a conclusion, it is important to note that these previous research, although using different methods and having a variety of centers, have a common characteristic that is the estimation and comparison of the applications and various forms of containerized and virtualized technology [16].

145

Proposed Method

This article provides an illustration of the operation of a web application that is deployed across several container environments, with each container being accountable for a particular set of responsibilities. For example, one container is responsible for carrying out all of the work associated with the web front end, while another container is in charge of the database-related operations [17]. Docker orchestration and composition [5] are very important components in the process of administering and directing the communication that occurs between these containers. The use of Docker Swarm clusters [6], in which load balancers divide the workload across a number of different containers, makes it easier to achieve high availability. Within the context of this swarm, one container serves as the master, while the other containers occupy the role of slaves. The web application may be accessed from a browser or from another container. Both methods are available. Depending on the needs of the user, there are a variety of deployment strategies that may be used. These strategies range from directly delivering a static website by using a Linux image and port binding to deploying a whole application that hosts both the database and the frontend inside the same container. By describing requirements, such as the base image, needed server installation procedures, and port binding, Docker file [7] makes deployment more straightforward. Through the use of a master-slave configuration, Docker Swarm [8] guarantees high availability, scalability, and load balancing for applications and services [9]. Multiple services that are part of a Docker Swarm group are linked to one another by means of a default or overlay network connectivity. Windows, Linux, CentOS, RHEL, and other operating systems are examples of host operating systems. The behavior of containers and the program that is running within the host operating system is described in the host operating system. It is necessary to establish two distinct containers, one of which is designated for the database, and the other contains the web application [18]. A one-of-a-kind Internet Protocol address is given to every container, and it accesses certain ports. The application is made available to external networks, such as the host system, the guest system, the slave node, the master node, and so on, when it is connected to both containers over a bridge network on port 80. The hosting of these multi-container applications on a public cloud platform, such as Amazon Web Services, Google Cloud Platform, or Microsoft Azure, is an alternative option that might be considered. In order to get extensive access across their various networks, it is feasible to use the same deployment procedures by installing Docker and making use of either Docker Swarm or Docker Compose. This means that it is possible to achieve

widespread access. Alternate way is that, the majority of the cloud suppliers give the components and administrations of organizations for the executives of sending of utilizations.
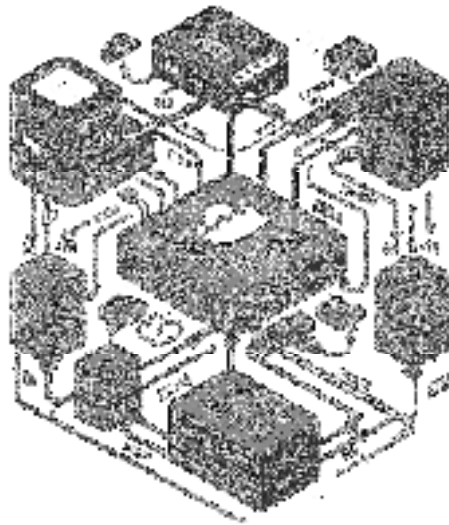
Figure 6 Multi Container Deployment

Shown in figure 6 is a network architecture that is contained inside a host system. This design illustrates the connection between a database service and a web service. The web service has an internal IP address of 172.18.0.3 and is set up to interact on port 8080. On the other hand, the database has an internal IP address of 172.18.0.4, but there is no port defined for it. 'mybridge Network' is the name of the bridge network that connects both services. This network itself has an interface that has an IP address of 192.168.2.3. Due to the fact that this bridge network is linked to an external network on port 80, it is possible to deduce that it enables communication between the host system and an external network, most likely the internet, over this port. When it is necessary to maintain isolation between services in containerized environments, this arrangement is often used. It is also necessary for these services to interact with one another and with the outside world.

A network configuration that is contained inside a host machine, most likely for a Docker installation. The first container is called "WEB," and it has an internal IP address of 172.18.0.3 on the network interface eth0. The second container is branded "Database," and it has an internal IP address of 172.18.0.4 on its eth0 interface. Both of these containers are distinct from one another. 'mybridge Network' is the name of the custom bridge network that both containers are linked to, which indicates that they are able to interact with one another over this network. 192.168.2.3 is the Internet Protocol address that is associated with the bridge network, which is connected to the eth0 network interface of the host system. There is

a mapping that is specified for the WEB container that links port 8080 inside the container to port 80 on the external network. This makes it possible for HTTP traffic to access the WEB service from outside the host system. This setup makes it easier to communicate amongst containers as well as to have restricted access to the services from an external network.

Result and Analysis

This research primarily focuses on comparing Docker, a container-based virtualization technology, with traditional virtual machines (VMs). The study shows the advantages that Docker provides in terms of enhancing efficiency and optimizing resource use. Notable findings include the observation that Docker exhibits a reduced startup time and faster performance when compared to virtual machines (VMs). Additionally, Docker is well-suited for delivering web applications across diverse container settings. This comparison demonstrates that Docker effectively enhances performance and optimizes resource efficiency in cloud computing and microservices-based systems.

Conclusion

Docker is a container-based virtualization solution that offers significant benefits over conventional virtual machines (VMs) in important areas such as efficiency, resource consumption, and deployment flexibility. The outcomes of the research highlight the enormous advantages that Docker offers over traditional VMs. Docker's particular lightweight characteristics and flexibility for microservices architecture stand out as noteworthy advantages. These characteristics make it easier to move and alter Docker, which eventually results in increased resource utilization and faster deployment procedures. When compared to virtual machines (VMs), Docker offers significant performance advantages, notably in terms of boot time and operating speed. This is one of the primary focus points of the study that is being conducted. Docker's effectiveness in the field of web applications is shown by the fact that these benefits lead to an overall system that is more efficient and responsive. The study makes use of a demonstration technique, which demonstrates how web applications may be successfully implemented across numerous Docker containers. This practical example serves as a testimony to the adaptability of Docker, demonstrating its capacity to adapt and expand in a seamless manner across a variety of circumstances, including deployment on cloud platforms. The research not only highlights the inherent qualities of Docker, but it also frames it as a solution that is both resilient and adaptable, and

it is capable of addressing the dynamic demands that current application development and deployment environments provide.

Future Scope

In light of the encouraging results about Docker's dominance over conventional virtual machines (VMs), a fascinating future scope has been paved the way. The effectiveness of Docker in microservices architecture might be fine-tuned in the future via research studies, with the goal of enhancing its flexibility to a variety of different circumstances. It is probable that the primary topics of investigation will be on the increase of security for Docker containers, the study of scalability in complicated systems, and the concentration on interoperability across several platforms. Additionally, the investigation of Docker's integration with upcoming technologies such as serverless architecture, artificial intelligence, and edge computing might lead to the discovery of creative solutions. It is possible that further work will also focus on the development of tools for effective container orchestration and management. Additionally, the investigation of Docker's influence on the environment may correlate with sustainable computing methods. In order to empower a competent workforce that is adept in Docker technologies, education and training programs are set to become crucial. This will support the continuous use and growth of Docker as a key technology in virtualization and application deployment.

References

[1] Babak Bashari Rad, Harrison John Bhatti and Mohhamad Ahmadi "An Introduction to docker and performance analysis" IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.3, March 2017

[2] Vivek Sharma, Harsh Kumar Saxena, Akhilesh Kumar Singh Docker for multi-container Web Applications

[3] Scheepers, M. J. (2014). Virtualization and containerization of application infrastructure: A comparison.

[4] D. Bernstein, "Containers and cloud: From LXC to Docker to Kubernetes," IEEE Cloud Computing, vol. 1, no. 3, pp. 81-84, 2014.

[5] D. Bu et al., "KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis," Nucleic Acids Research, vol. 49, no. W1, pp. W317-W325, 2021.

[6] W. Felter et al., "An updated performance comparison of virtual machines and Linux containers," in ISPASS 2015 - IEEE International Symposium on Performance Analysis of Systems and Software, 2015.

[7] T. Nath et al., "Using DeepLabCut for 3D markerless pose estimation across species and behaviors," Nature Protocols, vol. 14, no. 7, pp. 2152-2176, 2019.

[8] S. Tanwar, K. Parekh, and R. Evans, "Blockchain-based electronic healthcare record system for healthcare 4.0 applications," Journal of Information Security and Applications, 2020, vol. 50.

[9] B. Paxton et al., "Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions," Astrophysical Journal, Supplement Series, vol. 234, no. 2, 2018.

[10] J. Chong et al., "MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis," Nucleic Acids Research, vol. 46, no. W1, pp. W486-W494, 2018.

[11] E. Brugnera et al., "Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex," Nature Cell Biology, vol. 4, no. 8, pp. 574-582, 2002.

[12] C. Boettiger, "An introduction to Docker for reproducible research," in Operating Systems Review (ACM), 2015.

[13] S. Arnautov et al., "SCONE: Secure Linux containers with Intel SGX," in Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016, 2016.

[14] Seo, K.-T., Hwang, H.-S., Moon, I.-Y., Kwon, O.-Y., & Kim, B.-J. (2014). Performance Comparison Analysis of Linux Container and Virtual Machine for Building Cloud.

[15] RightScale (2008) Define Cloud Computing. RightScale Blog, 26 May 2008 .http://blog.rightscale.com/2008/05/26/define-cloudcomputing/.

[16] Howe, B. 2012. Virtual appliances, cloud computing, and reproducible research. Computing in Science & Engineering. 14, 4 (Jul. 2012), 36–41.

[17] Mohammad Ahmadi, "An Introduction to Docker and Analysis of its Performance", IJCSNS InternationalJournal of Computer Science and Network Security, VOL.17 No.3, March 2017.

[18] David Bernstein, "Containers and Cloud: From LXC to Docker to Kubernetes",IEEE Cloud Computing , Year: 2014, Volume: 1, Issue: 3

[19] Sachchidanand Singh, Nirmala Singh, Containers & Docker: Emerging Roles & Future of Cloud Technology , 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Year:2016, 2016 IEEE.

Current Scenario of Mental Well-Being of Modern Single Women Residing in A Metropolitan Cities – A Preliminary Analysis and Report

[1]V.Elamaran, [2]Dr.Sukirti Priyadarshni

Research Scholar, School of Arts & Humanities, Career Point University, Kota, Rajasthan, India

Email: maranddevi@gmail.com

Associate Professor, School of Arts & Humanities, Career Point University, Kota, Rajasthan, India

Email: sukriti.priyadarshini@cpur.edu.in

Abstract

Single mothers frequently experience emotional turmoil and suffer from inadequate mental well-being. There is limited data comprehending today's modern single mother's mental health in metropolitan cities like Chennai. Gaining insight into these paths is essential for enhancing mental health results for unmarried mothers. Our objective was to decipher the current mental well-being of the single-mother community based on a GHQ-based survey. GHQ-12 tool served as a preliminary screening and assessment tool to decipher the mental wellness among single mothers in modern current society. Data obtained from GHQ-12 revealed that the respondents (Single mothers) were not able to concentrate, contribute themselves, make decisions and face problems. Based on the collected data and analysis, we conclude that in order to enhance their mental well-being, it is crucial to introduce customised interventions that target the various difficulties they face along their journey as parents. These interventions should be customised to a specific environment based on personal exposure and circumstances, enabling single mothers to make well-informed choices and enhance their overall mental well-being.

Keywords

Metro cities, single mothers, mental-health, associated factors, survey

Introduction:

Women in prehistoric India were revered as goddesses and enjoyed elevated social standing (Pal et al 2019). Nevertheless, their social standing gradually diminished as a result of numerous societal vices in India, particularly the Sati system, child marriage, and limitations

on the remarriage of widows. During the Vedic period, women were seen as superior to men and were shown reverence and esteem (Zaidi et al 2009). The advancement of women's status persisted until the Mughal era; however, with the onset of the British period, little advancement was achieved. Following independence, numerous constitutional privileges provided societal recognition and positions within Indian society (Radha et al 2022).

Review of Literature:

Single parenting has grown more common in recent times due to the progress made in healthcare technology, including artificial reproductive technologies (ART), in vitro fertilisation (IVF), and surrogacy. However, traditional social norms and ineffective legal enforcement have impeded the realization of these rights, particularly for those who are illiterate and poor. Notwithstanding these difficulties, single parenting continues to be a widespread global occurrence, as the increase in in vitro fertilisation (IVF) and surrogacy, personal decisions regarding ways of living, divorce, widowhood, and more liberal adoption constraints all contribute to its prevalence (Diaz-Serrano and Flamand 2023).

The topic of single parenthood is of great significance, as familial relationships play a crucial role in determining psychological wellness. Single mothers frequently display elevated levels of psychological symptoms and financial instability (Jagdeesh et al 2018).

A single mother plays a vital role in a family, managing the duties of a sole parent in addition to nurturing offspring. The challenges of this role are frequently compounded by the absence of familial and societal support, as well as the extensive duties that come with being a sole carer (Zhou and Taylor, 2022). The kids rely on the single parent for support until they marry or secure employment. A solitary mother, regardless, must relinquish her social and economic independence, frequently experiencing resentment towards her marriage separation. They bear a multitude of obligations, such as nurturing dependent kids, and socioeconomic and emotional challenges (Suarez et al 2023). The economic challenges faced by single mothers can be a significant source of stress, resulting in emotions such as feeling isolated, helplessness, despair, a diminished sense of identity management, and reduced confidence. The aforementioned pressures might result in mental health issues, specifically depression, in single mothers (Sangeeth and Singh, 2022).

Accessing resources is crucial for efficiently addressing daily issues and challenges. Studies have shown that higher levels of proficiency and social support are associated with a decrease in depressive symptoms (Beeber et al 2008). Literature has reported that a lack of personal

support for single mothers towards nurturing their children results in increased levels of mental distress. Single parenthood places women in a disadvantaged social position and increases their vulnerability to enduring stress, exposing them to risky health behaviours and patterns, psychiatric disorders, psychological distress, psychosomatic symptoms, and depression (Rousou et al 2016).

Research Gap Identified:

Prior research has predominantly focused mental wellness of married single mothers and their children. Nevertheless, most of them reside within the community, it is necessary to conduct a community study to ascertain the prevalence of depression among this existing population (Chaudhuri et al 2017). Limited data is reported on the on the prevalence of mental distress among current modern single mothers residing in big metropolitan cities like Chennai, Tamil Nadu. Education, financial situation, digital world and technology have paved an avenue to lead an independent mindset, but there are limited constrains are still prevalent. Research consistently demonstrates that mental health issues are more prevalent among single mothers compared to married mothers. The lack of investigation in the sociological literature about single mothers is a critical aspect that must continue to be addressed. The impact of parenting in single-mother households hasn't been extensively investigated, taking into account the moderating intrinsic and extrinsic implications. However, the extent of this issue may be worse because assessments associated with certain circumstances might not recognise the significant amount of suffering the single mom faced. Regularly identifying depression in unmarried mothers living in the community is crucial for immediately implementing therapies to improve their overall mental health. This study aimed to determine the incidence of depression among single mothers living in the community and to identify the underlying factors that contribute to it.

Research Methodology:

The majority of single mothers reside in society and rely on assistance for their survival. The present investigation involves individuals from Chennai who were selected to serve as a representative sample of single mothers residing in particular demographic zones. The criteria for inclusion encompassed individuals aged eighteen and older who had provided consent to participate in the study. Individuals who were unable to read or write were not included in the data collection. Participants were instructed to fill out the survey questionnaire autonomously. This study was conducted using a web-based questionnaire among single

mothers with children. A total of 51 feedback responses were collected from single parents as part of the preliminary investigation. This study employed the GHQ-12, a mental health measurement tool, to assess the mental well-being of single mothers with children (Guan and Han, 2019). This questionnaire comprises a total of 12 items (Goldberg, 1972). We implemented the four-point Likert scale, where each item is assigned a value between 0 and 3. After adding up each component, the final score had a range of 0-36. A higher score indicates poorer mental health. Any scores over the threshold of 12 could be categorised as cases (Liang et al 2016).

Research Findings:

The demographic and socioeconomic characteristics of respondents mainly included gender, age, education, employment level, and number of children. Of the 51 respondents, 7.84% were aged 20–29 years, 35.29% were aged 30–39, 39.21% were aged 40–45, and 17.64% were above 50 years of age. The education levels of these single mothers were divided into four categories: school, diploma, undergraduate, graduate, and doctorate. Most respondents fell into the undergraduate category (47.05%), followed by school education (29.41%) and graduate and above (19.06%).

Table 1.1 shows the overall and individual item scores of the GHQ-12 survey. In particular, the highest average scores were for items 1, 4, 3, and 8, which were more than 1.63. Of these items, the average score of item 4 was 1.745 (SD = 0.979), as the highest, indicating that the majority of respondents felt that they were not capable of making decisions unhappy and depressed. While the average score of item 1 was 1.705 (SD = 0.855), the majority of respondents (42.813%) scored 2 points representing same as usual/no more than usual, and only 22.5% of respondents scored 3, showing that in general the respondents much less than usual/ much more than usual. The average score of the GHQ-12 was 1.486, and the standard deviation of 0.995 measures the extent of variance. These outcomes indicate that the individuals' well-being, particularly their mental health, was experiencing a condition of distress. Cronbach's alpha reliability was computed using the Spearman-Brown formula. The Cronbach's alpha coefficient for the GHQ-12 was determined to be 0.844, surpassing the criterion of 0.8. This suggests that the scale has robust dependability and correlation. The researchers employed the GHQ-12 questionnaire to evaluate people's psychological characteristics associated with mental well-being and scrutinised it to ascertain its reliability and coherence. Previous studies have examined the reliability and validity of the GHQ-12

questionnaire for single mothers. However, our current investigation is the first to particularly examine the factor structure of the GHQ-12 questionnaire among single mothers. The GHQ-12 evaluated single mothers with children who belonged to a particular group that was vulnerable to mental health problems. The Cronbach's alpha coefficient of 0.844 suggests that the GHQ-12 has satisfactory levels of consistency and internal coherence among single moms with children in Chennai, Tamil Nadu, and India.

Stigma can be societal stigma or self-stigma, as these variations are determined by the individuals who possess unfavourable evaluations of the marginalised community (Corrigan, 2004). Furthermore, the literature have emphasised that when single-parent women internalise self-stigma, their self-esteem decreases, resulting in heightened mental health difficulties (Kim et al 2023). Data analysis revealed that the majority of respondents felt that they were not capable of making decisions unhappy and depressed. Literature reported that those with history of depression have difficulties to escape depression will have increased risk of recurrence (Theng et al 2022). Single mothers who have already experienced depression are at an increased likelihood of experiencing additional instances of depressive disorders due to sociocultural stigma, lack of emotional and monetary support, and the responsibility of parenting their children (Atkins, 2010). Nevertheless, a few individuals have the capability to endure with a robust support provided by the community. A considerable proportion of participants reported receiving substantial assistance from their family, friends, and numerous others, which was found to be inversely correlated with depression among single mothers (Zhou and Taylor, 2022). Hence, it is imperative to offer sufficient social assistance in order to avoid subsequent depression occurrences in single mothers (Rousou et al 2016).

The normal probability plot (1, 3, 4 and 8) is used to determine the distributional behaviour of the data, specifically whether it follows a normal distribution. A linear trend in the normal residual plot indicates the statistically significant value of a model's coefficients, and a correlation between the experimental and anticipated outcomes was comprehended (Abdur Rob and Srivastava, 2022). The verification of the normalcy assumption is conducted through the use of the Anderson-Darling (AD) test (Stephens, 1974). The AD statistics values for all twelve responses was found to be less than 0.05, indicating that the experimental data has to be fitted line for the selected four responses, demanding a transformation (Anderson and Darling, 1952). The analysis of the mean of respective GHQ responses (N=12) indicated the

AD statics values (A2 value – 0.36) and the (p-value = 0.383) with a mean of 1.4864 and standard deviation of 0.1986 suggesting that the data has a normal distribution. Grubbs's outlier analysis reported no significant outlier was identified with a significance level of 5% (Aslam, 2020).

Conclusion

Single mothers face a substantial vulnerability to experiencing depression as mothers, which can affect the socio-emotional well-being of their offspring. The prevalence of this public health risk is especially alarming for single mothers, who frequently encounter elevated levels of depressive symptoms due to the demands of early and later motherhood. GHQ-12 tool served as a preliminary screening and assessment tool to decipher the mental wellness among single mothers in modern current society. Data obtained from GHQ-12 revealed that the respondents (Single mothers) were not able to concentrate, contribute themselves, make decisions and face problems. The implementation of preventative, assessment, and early intervention initiatives can have a substantial positive impact on the welfare of single mothers and their children, thereby helping both the public health sector and the overall economy. The findings of single mothers indicated that they were unable to make decisions and were encountering challenges at all levels of parenthood. Furthermore, several categories of single women encountered increasingly negative patterns in their mental well-being, highlighting the requirement for context-specific targeted interventions tailored to the individual circumstances of single mothers to enhance their mental health.

References

1. Abdur Rob, S. M., & Srivastava, A. K. (2022). Turning of carbon fiber reinforced polymer (CFRP) composites: process modeling and optimization using Taguchi analysis and multi-objective genetic algorithm. Manufacturing Letters, 33, 29-40.

2. Anderson, T., & Darling, D. (1952). Asymptotic theory of certain "Goodness of Fit" criteria based on stochastic process, The Annals of Mathematical Statistics 23 (1952), 2, 193-212. DOI:10.1214/aoms/1177729437.

3. Aslam, M. (2020). Introducing Grubbs's test for detecting outliers under neutrosophic statistics–An application to medical data. *Journal of King Saud University-Science*, *32*(6), 2696-2700.

4. Atkins, R. (2010). Self-efficacy and the promotion of health for depressed single mothers. Mental Health and Family Medicine, 7(3), 155-68.

5. Beeber, L. S., Perreira, K. M., & Schwartz, T. (2008). Supporting the mental health of mothers raising children in poverty: how do we target them for intervention studies?. Annals of the New York Academy of Sciences, 1136, 86–100. https://doi.org/10.1196/annals.1425.008.

6. Chaudhuri, S. B., Mandal, P. K., Chakrabarty, M., Bandyopadhyay, G., & Bhattacherjee, S. (2017). A study on the prevalence of depression and its risk factors among adult population of Siliguri subdivision of Darjeeling district, West Bengal. Journal of family medicine and primary care, 6(2), 351–355. https://doi.org/10.4103/jfmpc.jfmpc_326_16.

7. Corrigan, P. (2004). How stigma interferes with mental health care. *American psychologist*, *59*(7), 614.

8. Diaz-Serrano, L., & Flamand, S. (2023). Attitudes towards single parents' children in private and state-dependent private schools: experimental evidence. SERIEs 14, 223–242. https://doi.org/10.1007/s13209-023-00281-3.

9. Goldberg, D. (1972). The Detection of Psychiatric Illness by Questionnaire: A Technique for the Identification and Assessment of Non-Psychotic Psychiatric Illness. New York, NY: Oxford University Press.

10. Guan, M., & Han, B. (2019). Factor structures of general health questionnaire-12 within the number of kins among the rural residents in China. Frontiers in Psychology, 10, 1774.

11. Jagdeesh, S., Abirami, P., & Sasikala, R. (2018). Assess the Effect of Single Parenting on Emotional Well Being of Adolescents in government higher Secondary School at Kancheepuram District. Amarjeet Kaur Sandhu, 10(3), 152.

12. Kim, A., Jeon, S., & Song, J. (2023). Self-Stigma and Mental Health in Divorced Single-Parent Women: Mediating Effect of Self-Esteem. *Behavioral Sciences,* 13(9), 744. https://doi.org/10.3390/bs13090744.

13. Liang, Y., Wang, L., & Yin, X. (2016). The factor structure of the 12-item general health questionnaire (GHQ-12) in young Chinese civil servants. Health and quality of life outcomes, 14(1), 136. https://doi.org/10.1186/s12955-016-0539-y.

14. Mayhew, E., Stuttard, L. & Beresford, B. (2021). An Assessment of the Psychometric Properties of the GHQ-12 in an English Population of Autistic Adults Without Learning Difficulties. Journal of Autism and Developmental Disorders, 51, 1093–1106. https://doi.org/10.1007/s10803-020-04604-2.

15. Pal, B. (2019). The saga of women's status in ancient Indian civilization. Miscellanea Geographica, 23(3), 180-184.

16. Radha, R., Santhi, S., Padmasree, A. D., Thirumalaichamy, M., Thangamuthu, P., & Saravanakumar, A. R. (2022). Constitutional Guarantees of Women in India-A Historical Study. Specialusis Ugdymas, 1(43), 9733-9741.

17. Rousou, E., Kouta, C., & Middleton, N. (2016). Association of social support and socio-demographic characteristics with poor self-rated health and depressive symptomatology among single mothers in Cyprus: a descriptive cross-sectional study. BMC Nursing, 15, 15. https://doi.org/10.1186/s12912-016-0134-x.

18. Sangeet, O., & Singh, S. (2022). Experiences of single-parent children in the current Indian context. Journal of family medicine and primary care, 11(7), 3790–3794. https://doi.org/10.4103/jfmpc.jfmpc_2455_21.

19. Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons, Journal of the American Statistical Association, 69, pp. 730-737.

20. Suarez, A., Shraibman, L., & Yakupova, V. (2023). Long-Term Effects of Maternal Depression during Postpartum and Early Parenthood Period on Child Socioemotional Development. Children. 10(10), 1718. https://doi.org/10.3390/children10101718.

21. Theng, Y.L., Abdul Rahman, R., Ismail, S.B. (2022). Determinants of Depression among Malay Single Mothers Living in Community in Perak, Malaysia. The Malaysian Journal of Medical Sciences, 29(2), 80-93. 10.21315/mjms2022.29.2.8.

22. Zaidi, S., Ramarajan, A., Qiu, R., Raucher, M., Chadwick, R., & Nossier, A. (2009). Sexual rights and gender roles in a religious context. *International Journal of Gynecology & Obstetrics*, *106*(2), 151-155.

23. Zhou, X., & Taylor, Z. E. (2022). Differentiating the impact of family and friend social support for single mothers on parenting and internalizing symptoms. Journal of Affective Disorders Reports, 8, 100319.

Table 1.1: The overall and individual item scores of the GHQ-12.

S.No	Factors	Mean	SD	Skewness	Kurtosis	Response frequencies (%)

						0	1	2	3
1	Can you able to concentrate	1.705	0.855	0.218	-0.9411	3.922	43.137	31.373	21.569
2	Have you lost much sleep	1.313	0.969	0.4161	-0.7133	19.608	45.098	19.608	15.686
3	Have you found playing a useful part in your life	1.686	1.086	-0.014	-1.3812	13.725	37.255	15.686	33.333
4	Are you capable of making decisions	1.745	0.979	0.0064	-1.1927	7.843	39.216	23.529	29.412
5	Are you under stress	1.47	1.083	0.324	-1.2222	17.647	45.098	9.804	27.451
6	I could not overcome difficulties	1.117	0.951	0.4811	-0.6372	29.412	39.216	21.569	9.804
7	Do you enjoy your day-to-day activities	1.627	1.057	0.1784	-1.3399	11.765	45.098	11.765	31.373
8	Can you able to face up to problems	1.647	1.035	-0.0178	-1.1953	13.725	35.294	23.529	27.451
9	Are you feeling unhappy and depressed	1.333	0.952	0.4286	-0.6527	17.647	47.059	19.608	15.686
10	Are you losing confidence	1.313	0.989	0.2216	-0.9416	23.529	35.294	27.451	13.725
11	Are you thinking of self	1.372	0.958	0.5921	-0.6098	13.725	54.902	11.765	19.608

	as worthless								
12	Are you feeling reasonably happy	1.509	1.027	0.3184	-1.1114	13.725	47.059	13.725	25.490
	Mean	1.486	0.995						

Enhancing Prediction Accuracy in Automated Overnight Patient Care: A Comprehensive Investigation

[1]Krishnanjali Magade, [2]Dr. Amit Sharma

[1]Research Scholar, School of Engineering & Technology, Career Point University,

Kota, Rajasthan, India

Email : jaianjali85@gmail.com

[2]Associate Professor, School of Engineering & Technology,

Career Point University, Kota, Rajasthan, India

Email : amit.sharma@cpur.in

Abstract:

This research paper delves into the critical domain of automated overnight patient care, seeking to elevate the precision and reliability of predictive models employed in this context. Recognizing the paramount importance of accurate predictions for timely and effective interventions, our study undertakes a comprehensive investigation. Through an amalgamation of advanced machine learning algorithms, data preprocessing techniques, and domain-specific insights, we aim to enhance the prediction accuracy of automated systems tasked with monitoring and managing patients during overnight periods. The paper explores the intricacies of relevant datasets, analyzes the performance of existing models, and proposes novel methodologies to address inherent challenges. By fostering a deeper understanding of the nuances in overnight patient care, our research aspires to contribute significantly to the advancement of automated healthcare systems, ultimately improving patient outcomes and fostering a more efficient and reliable healthcare ecosystem.

Keywords:

Prediction accuracy, Automated patient care, Overnight monitoring, Healthcare systems, Machine learning algorithms, Data preprocessing, Predictive models

Introduction

In the realm of contemporary healthcare, the integration of automation and predictive modeling has emerged as a pivotal frontier in optimizing patient care. The advent of automated systems, particularly those designed for overnight patient monitoring, holds great promise for improving the efficiency and efficacy of healthcare delivery. In this context, the paramount need for accurate predictions to facilitate timely and targeted interventions cannot be overstated. This research endeavors to address this imperative by embarking on a comprehensive investigation aimed at enhancing the prediction accuracy within automated overnight patient care.

Encapsulates the dual objectives of elevating the precision of predictive models and conducting an exhaustive inquiry into the multifaceted facets of this critical healthcare domain. Recognizing the intrinsic challenges associated with nocturnal patient care, our study employs a synthesis of advanced machine learning algorithms and meticulous data preprocessing techniques. These elements are intricately woven together with domain-specific insights to form a cohesive analytical framework, propelling our efforts to refine existing models and propose innovative methodologies.

Through a thorough exploration of relevant datasets and a critical evaluation of the performance of existing models, we aim to uncover insights that transcend the current limitations of automated overnight patient care systems. This paper contributes to the existing body of knowledge by not only identifying areas for improvement but also by offering novel solutions grounded in empirical evidence. By providing a nuanced understanding of the intricacies involved in overnight patient care, our research seeks to catalyze advancements that transcend the status quo, ultimately fostering a healthcare landscape characterized by heightened prediction accuracy, improved patient outcomes, and overall system efficiency.

Introduction

1. Overview of the increasing role of automation in contemporary healthcare

  - The significance of predictive modeling in optimizing patient care.

The introduction to automated healthcare systems marks a pivotal entry point into the dynamic landscape of contemporary medical practices. With the accelerating pace of technological advancement, healthcare systems have undergone a transformative shift,

embracing automation to streamline processes and enhance patient care. Automated healthcare systems encompass a spectrum of applications, ranging from electronic health records to diagnostic tools and predictive modeling. This paradigm shift is motivated by the pursuit of efficiency, accuracy, and improved patient outcomes. As these systems become increasingly integral to healthcare infrastructure, understanding their fundamental principles becomes imperative. This section explores the evolution of automated healthcare, highlighting its profound impact on the delivery of medical services, data management, and the broader landscape of patient well-being. The integration of automation not only offers unprecedented opportunities for optimization but also necessitates a comprehensive examination to ensure that its implementation aligns seamlessly with the complex and nuanced nature of healthcare delivery.

2. Nighttime Challenges in Patient Monitoring:

  - Recognition of the unique challenges associated with overnight patient care.

  - The critical need for accurate predictions during nighttime monitoring.

Navigating the intricacies of patient monitoring assumes a distinctive set of challenges during the nocturnal hours, ushering in a realm of considerations unique to nighttime healthcare. As the sun sets, the demand for vigilant and continuous monitoring remains unabated, requiring healthcare systems to grapple with diminished visibility and the inherent vulnerability associated with nighttime care. Patients, often in states of vulnerability or undergoing critical treatments, necessitate a specialized focus on accurate and timely interventions. The darkness amplifies the challenges in observation, potentially hindering the swift detection of emergent issues. The nighttime challenges in patient monitoring extend beyond mere visibility concerns, incorporating the intricacies of altered circadian rhythms and the potential impact on physiological parameters. This section delves into the multifaceted nature of these nocturnal challenges, emphasizing the critical need for heightened prediction accuracy in automated systems to mitigate risks and ensure the provision of optimal care throughout the night.

3. Rationale for Improved Prediction Accuracy:

  - Discussion on the implications of accurate predictions for timely and effective interventions.

  - Aligning enhanced prediction accuracy with improved patient outcomes.

The rationale for improved prediction accuracy in healthcare systems is rooted in the fundamental principle that precision in predicting patient outcomes directly correlates with the quality and effectiveness of healthcare interventions. Timely and accurate predictions serve as the linchpin for informed decision-making, enabling healthcare professionals to anticipate and address potential complications or changes in a patient's condition. This section underscores the pivotal role of predictive accuracy in optimizing resource allocation, treatment planning, and overall patient care. As healthcare systems increasingly rely on automated processes and machine learning algorithms, the imperative for heightened prediction accuracy becomes paramount. The ripple effect of improved predictions extends beyond individual patient cases, influencing broader healthcare strategies, resource utilization efficiency, and the overall efficacy of healthcare delivery. This research aims to articulate and address the rationale for enhancing prediction accuracy, contributing to a more responsive and patient-centric healthcare paradigm.

4. Integration of Machine Learning Algorithms:

  - Introduction to the utilization of advanced machine learning algorithms.

  - The role of data preprocessing techniques in refining predictive models.

The integration of machine learning algorithms represents a transformative frontier in healthcare, promising to revolutionize the accuracy and efficiency of patient care. The ability of machine learning to identify complex patterns in large datasets is crucial for the automation of healthcare operations. Using machine learning algorithms is a calculated move in this study's direction, with the goal of improving prediction models for patients' care throughout the night. These algorithms have the ability to learn from historical patient data, adapt to evolving patterns, and generate predictions that are both precise and timely. By leveraging sophisticated computational techniques, the integration of machine learning algorithms empowers healthcare systems to navigate the complexities of overnight patient monitoring, providing a dynamic and adaptable approach to prediction accuracy. This section delves into the principles, methodologies, and potential impact of integrating machine learning algorithms in the context of automated overnight patient care, setting the stage for a comprehensive investigation into the intersection of technology and healthcare.

5. Contributions to Healthcare Knowledge:

  - Highlighting the potential impact of the research on advancing the understanding of automated healthcare systems.

- The aim to contribute novel insights and solutions for more effective patient care.

The contributions to healthcare knowledge through machine learning (ML) stand as a testament to the transformative potential of advanced computational techniques in the medical domain. ML algorithms, with their ability to discern intricate patterns within voluminous datasets, offer unprecedented insights that have the potential to redefine the landscape of healthcare knowledge. This research endeavors to contribute to this evolving body of knowledge by harnessing the power of ML in the specific context of automated overnight patient care. The study aims to not only enhance prediction accuracy but also to unravel novel insights into the dynamics of nighttime patient monitoring. The integration of ML facilitates a data-driven approach, enabling the identification of nuanced correlations and predictive factors that may elude conventional analyses. By leveraging the capabilities of ML, this research seeks to advance our understanding of automated healthcare systems, ultimately paving the way for more informed decision-making, improved patient outcomes, and a progressive shift toward data-centric healthcare practices.
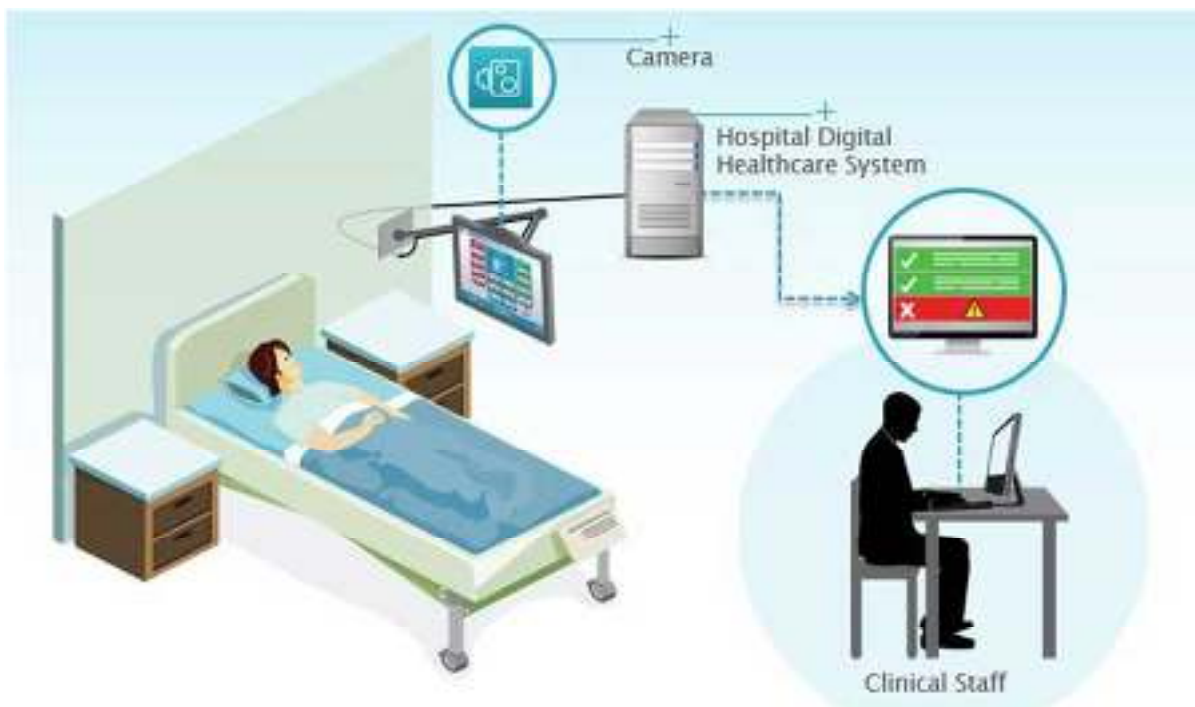
Fig.1 Process flow for patient and Clinical health monitoring system

In the rapidly evolving landscape of healthcare, this research endeavors to forge a path towards optimized patient care through the convergence of automated systems, predictive modeling, and advanced machine learning algorithms. The investigation acknowledges the distinctive challenges posed by nocturnal patient monitoring, emphasizing the critical need for enhanced prediction accuracy during overnight care. Every aspect, from using machine learning algorithms to investigating nighttime obstacles and providing a justification for better forecasts, adds up to a thorough comprehension of the intricacies at play. By delving into the intricacies of automated healthcare systems and the transformative potential of machine learning, the study aims to not only refine existing predictive models but also to lay the groundwork for innovative methodologies. The overarching goal is to make substantial contributions to healthcare knowledge, fostering a data-centric paradigm that transcends the limitations of conventional approaches. Through this research, we aspire to elevate the precision and efficacy of overnight patient care, ultimately charting a course towards a healthcare landscape characterized by heightened prediction accuracy, improved patient outcomes, and enhanced system efficiency.

Literature Review

Increasing the quality of care that is delivered to patients, simplifying processes, and improving the overall efficiency of the healthcare system are all possible outcomes that have never been seen before thanks to the introduction of new technologies in the healthcare business, which has ushered in a new age of possibility that has never been seen before. The use of automated technologies, particularly those that are designed to provide care to patients while they are sleeping, has developed into an essential component of modern healthcare facilities. When it comes to the topic of improving the accuracy of automated overnight patient care projections, the purpose of this literature review is to provide a synopsis of the research that has already been conducted on the issue. A important approach, a big obstacle, and a substantial success in the area will be highlighted in order to reach this goal.

1. Automation in Healthcare:

The integration of automation in healthcare has witnessed remarkable growth over the past decade. Automated systems, ranging from diagnostic tools to treatment plans, have been

implemented to augment the capabilities of healthcare professionals and improve patient outcomes. As the focus on optimizing patient care continues to intensify, there is a growing interest in leveraging automated systems during overnight hours to enhance the prediction accuracy of critical patient events.

O. Ali et. al. (2023) The use of artificial intelligence (AI) technology is quickly altering the medical and administrative procedures of healthcare organisations. The profound influence of AI on many tasks is highlighted by this transformation, which is especially true in the realm of early detection and diagnosis in medicine. Prior research indicates that AI are able to improve healthcare service quality. People claim that AI-based technology make people's lives better by reducing stress, increasing security, and boosting productivity. Academic publications on artificial intelligence's use in healthcare are systematically reviewed in this research. At first, 1,988 research publications from top databases were evaluated for the evaluation. The selection was narrowed down to 180 papers after a thorough assessment. These articles will be fully analysed to offer a categorization scheme based on four dimensions: AI in healthcare: advantages, disadvantages, approaches, and features. When compared to humans, AI still does a much better job of medical and related administrative tasks in terms of precision, efficiency, and timeliness. When it comes to self-management of chronic illnesses, patients may reap the benefits of AI's diagnostic, therapeutic, consulting, and health monitoring capabilities. Topics such as innovative AI service delivery methods, health monitoring capabilities, patient data security and privacy, and value-added healthcare services for medical decision-making are highlighted as areas where future research should focus.

S. A. Alowais et. al. (2023) Although artificial intelligence (AI) has revolutionized several sectors, including healthcare, and holds great promise for enhancing both patient care and overall quality of life, all parties involved acknowledge the challenges posed by healthcare systems. Potentially game-changing advances in healthcare AI are underway in the clinical setting. Reports on clinical AI utilisation provide doctors vital data and resources. Disease diagnosis, treatment ideation, and patient involvement are all areas of AI that are now the subject of substantial study in clinical practice. By addressing concerns related to ethics, law, and human knowledge, it paves the way for healthcare companies to embrace AI and put it to use. The use of AI in healthcare was investigated by doing an unlimited search of the English-indexed literature in PubMed/Medline, Scopus, and EMBASE. Research

focuses on artificial intelligence discoveries in healthcare. Medical diagnosis, treatment, and laboratory testing may all benefit from AI. In several areas of healthcare, AI algorithms may improve accuracy, efficiency, cost, and mistake rates by discovering patterns in massive datasets, surpassing human performance. Potential areas for improvement include virtual health assistants, personalised medicine, medication dosages, virtual health populations, suggestions, mental health services, patient education, and trust between patients and doctors. Now at last, AI can aid doctors with diagnosis, personalisation, and advice. AI has the potential to revolutionise healthcare in more ways than one. Healthcare AI that is both effective and safe must address concerns about bias, human understanding, and data protection.

2. Predictive Modeling in Healthcare:

A promising new technique, predictive modeling has the ability to foretell healthcare outcomes and spot hazards before they happen. The use of predictive models in the context of nighttime patient care has the potential to greatly enhance patient safety by facilitating the prompt action in the case of adverse occurrences. In order to create reliable prediction models for nighttime patient care, many research have investigated the use of machine learning techniques such support vector machines, neural networks, and ensemble approaches.

B. L. Jimma et. al. (2023) Doctors and patients alike may reap the benefits of AI-powered disease prediction, categorization, and diagnosis. With the exponential growth of data and computing power, AI has become more attractive. From 2000 to 2021, a comprehensive bibliometric analysis of healthcare AI research was conducted. All healthcare AI research published in English has been located and retrieved by Scopus. Topic categories, top nations, institutions, journals, and financial sponsors were all included of the analysis, as was the rise of publications. With an increase beginning in 2010, the search returned 5,019 unique entries. Published after 2012 accounted for 88.88% of the articles. The research found 96.85% of studies from nine different nations, with 41.84 percent coming from the United States. Three of our most important technological terms are "Machine Learning," "Electronic health records," and "Natural language processing." Issues including COVID-19, diabetes, mental health, asthma, dementia, and cancer were investigated within the field of artificial intelligence in healthcare. This comprehensive bibliometric analysis of

artificial intelligence (AI) studies in healthcare may help academics, politicians, and practitioners understand the evolution of the subject and the ethical implications of AI.

Ayesha Amjad et. al. (2023) As computing power has increased, AI has entered the mainstream. A lot is happening in the healthcare industry right now. Issuing electronic healthcare cards and providing personalised treatment are two examples of the various ways in which the relatively young field of telehealth relies on artificial intelligence. In the United States, telemedicine is greatly affected by artificial intelligence (AI). Artificial intelligence (AI) in telehealth enables doctors to make data-driven decisions in real-time, which has the power to enhance patient experiences and health outcomes, and virtual care solutions are becoming more accessible. The medical research community is taking notice of the growing use of artificial intelligence (AI) in the processing and analysis of telehealth data. There are obstacles to the widespread use of telemedicine, which calls for enhanced skills and better processes to provide individualized answers. The effects of artificial intelligence on healthcare and telemedicine are discussed in this article.

This study's literature evaluation reveals the boundless potential of merging AI with telemedicine. Patient monitoring, healthcare information technology, intelligent diagnostic assistance, and data analysis in collaboration with other experts are the primary areas of emphasis for this technology's increasing application.

3. Data Sources and Integration:

A variety of high-quality data sources are crucial to the success of prediction models. Studies have investigated the integration of electronic health records (EHRs), physiological monitoring data, and wearable device information to create comprehensive datasets for training and validating predictive algorithms. Understanding the challenges associated with data interoperability, privacy concerns, and data integration is pivotal for the successful implementation of accurate automated overnight patient care systems.

Ragavi V (2021) Many individuals are holding out hope that the medical business will be revolutionized by artificial intelligence (AI). Another way to make AI even more powerful would be to use very fast processors and combine AI algorithms with machine learning and deep learning algorithms. Studies show that doctors and other medical professionals will be able to make choices and progress in their sector much faster if digitization takes hold. This study will seek to understand how an AI system functions within. In this article, we will

explore how AI is changing the healthcare industry via chatbots powered by AI, automated imaging diagnostics, personal health companions, radiology, cancer, and cardiology, among other topics.

S. M. Mohammad (2020) The healthcare, commercial, and public sectors will all be automated via the use of this new technology, which will be examined in this study. The most recent technical developments have an effect on the standard operating procedures of many aforementioned areas of society, which in turn has an effect on how these areas function. In the realm of automation, the technology provides previously unknown approaches and throws light on certain regions in need of sufficient and exceptional services. The whole system's legitimacy and accountability are boosted by doing this. A variety of documentations that are practically used in the field of information technology are examined in order to learn how IT affects the automation of application areas with respect to the accountability equation. Beginning with the most basic types of automated transactions involving lower-level automation, the investigation progresses to more complex kinds of automation, such as systems that analyse biometric fingerprints. This example discusses the accountable potential of IT automation for various applications in the hopes of investigating the benefits of application automation and eliminating the unaccountable potential that hinders functionality in various fields that may result from using system applications. The paper devotes a lot of space to discussing how crucial it is to strike a balance between the advantages of IT automation and the whole automation process, including the system that could make applications less efficient while still holding people accountable.

4. Challenges and Limitations:

There have been some hopeful advancements, but there are still barriers to overcome in order to attain optimal prediction accuracy in automated nighttime patient care. This is the case despite the fact that there have been some developments. The interpretability of complex machine learning models, class imbalance, and data noise are all difficulties that need careful consideration. It is vital to pay considerable attention to these concerns. Furthermore, a thorough investigation is necessary because of the ethical issues that are involved with relying on automated systems for patient care, especially during the overnight hours. This is for the reason that the inquiry must be extensive.

G. Rong et. al. (2020) Software algorithms, hardware implementation, and the breadth of AI's applications have all been expanding at a fast pace in the last several years. Disease diagnoses, home help, processing of biomedical data, and biomedical research are some of the areas where artificial intelligence (AI) has recently made strides in biomedicine. Insight into current technology, keeping up with scientific developments, recognizing the enormous potential of AI in biomedicine, and providing inspiration for researchers in adjacent fields are all aims of this study. In the same way that artificial intelligence (AI) is in its infancy, so too is its use in biomedicine. Rapid advancements are anticipated in the near future, and new discoveries will keep expanding the boundaries of AI and its applications. To demonstrate how to fill a malfunctioning urine bladder and how to forecast the likelihood of epileptic seizures, two case examples are given.

Onur Asan (2020) The growing capacity of artificial intelligence (AI) to convert complicated and unclear data into practical, if flawed, clinical recommendations or judgements has the potential to revolutionise health care. One factor that influences doctors' usage and acceptance of AI is trust, which is a growing aspect of the human-AI connection. When faced with the gap between known and unknown, trust steps in as a psychological defence mechanism. A number of studies have brought attention to the fact that AI-based systems might need some work in order to better assist doctors. But there has to be a lot of focus on measuring the extent to which people trust AI and its effects. Can an AI-powered system earn a clinician's trust? How can we as humans determine if AI is trustworthy? Would it be possible to optimise AI trust in order to enhance decision-making? As the key end-users of AI systems in healthcare, clinicians are the major emphasis of this article, which also presents elements influencing clinicians' faith in AI. We bring attention to important trust-related issues that should be thought about when developing any AI system for medical usage.

5. Advancements and Future Directions:

In recent years, research has shown novel ways that aim to overcome the constraints of the automated systems that are already in use. There are a number of potential improvements, including the incorporation of explainable artificial intelligence, continuous model recalibration, and adaptive learning methodologies. Based on the existing body of literature, it is recommended that further research should concentrate on the improvement

of algorithms, the resolution of real-world implementation issues, and the establishment of uniform assessment measures in order to benchmark the performance of predictive models in a variety of healthcare contexts.

there is a growing corpus of research that is devoted to improving the accuracy of prediction in automated nighttime patient care. Through the exploration of the incorporation of predictive modeling, the use of a wide variety of data sources, the pursuit of solutions to obstacles, and the acceptance of current breakthroughs, this all-encompassing research lays the groundwork for the creation of automated systems in the healthcare industry that are more robust and efficient. In order to effectively harness the potential of automated nighttime patient care and enhance the results of healthcare, it is essential to bridge the gap between theoretical knowledge and the real use of new technology when it comes to healthcare.

Research Methodology

When it comes to biomedicine, the area dealing with patient illness diagnosis has the greatest demand for AI. So far, a plethora of remarkable findings have been discovered within this particular domain. Medical practitioners may now provide quicker and more precise diagnoses for a wider variety of diseases by using AI. Biosensors and biochips are commonplace in in vitro diagnostics, one of the most important categories of diagnostic procedures. This type of thing happens, for instance, when genes are expressed.
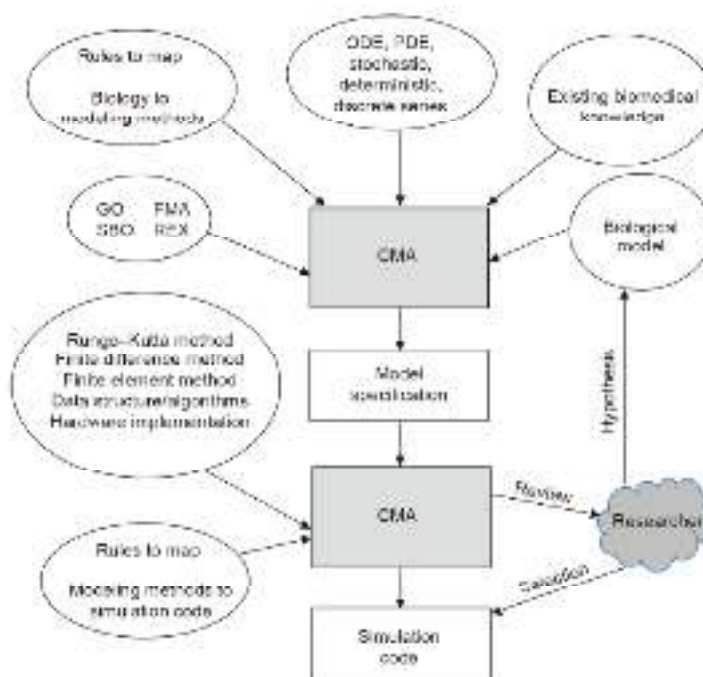
Fig. 2 A general perspectives of process flow and interactions between a CMA and human researchers

ML, in which AI deciphers microarray data to categorize and identify anomalies, may examine this crucial diagnostic tool. Classifying cancer microarray data for cancer detection is one emerging use. Biosensors and associated point-of-care testing (POCT) devices may detect cardiovascular illnesses early with the use of integrated artificial intelligence. Artificial intelligence (AI) may aid in cancer patient survival rate prediction in addition to diagnosis, for example in the case of colon cancer. Scientists have also pointed up ML's shortcomings in biological diagnostics and proposed solutions to these problems. So, AI in diagnosis and prognosis still has a lot of room to grow.

Medical imaging's two-dimensional and signal processing's one-dimensional components form another significant category of disease diagnostics. In the past, these methods have been used for the purpose of sickness prediction, management, and diagnosis. Artificial intelligence has found use in the extraction of features from biological signals, including EEGs, EMGs, and ECGs, for one-dimensional signal processing. Epileptic seizure prediction is a significant use of electroencephalogram (EEG). In order to lessen the severity of seizures for patients, seizure prediction is crucial. One of the most important components of a trustworthy prediction system is AI, which has just come to light. Predictions powered by deep learning are now within reach, and the platform for making these predictions may be integrated into mobile systems. AI has the potential to significantly impact biological image processing-based diagnosis. Image segmentation, thermal imaging, and multidimensional imaging have all made use of AI to improve the efficiency and quality of their respective analyses. Portable ultrasonic devices that include AI may be used by unskilled individuals to diagnose a wide range of ailments in underdeveloped locations. Besides the ones already mentioned, AI may also help traditional decision support systems (DSSs) with things like improving diagnosis accuracy

and making disease management easier, which means less work for humans. As an example, AI has found applications in cancer integrated management, tropical illness diagnosis and treatment, cardiovascular disease decision-making, and other areas. These uses highlight the potential of AI as a tool for precise and early diagnosis, treatment, and even prognosis of health issues. Presented below are two such case studies.

3.1. Healthcare

More and more medical fields are using AI into their practices. The analysis of signals and images, as well as the prediction of changes in function (e.g., in the management of the urine bladder, epileptic seizures, and strokes), have been its primary applications. Two common case studies are described below: one for predicting bladder volume and another for predicting epileptic seizures.

3.2. Bladder volume prediction

Several problems arise in the patient's health when the bladder's storage and urine functions are impaired, whether due to aging, a spinal cord injury, or another neurological illness. Implantable brain stimulators may now partially restore bladder function in individuals who have not responded to drugs. In order to enhance the efficacy and security of neuroprostheses by means of conditional neurostimulation, a feedback device that administers electrical stimulation alone when necessary is needed: a bladder sensor that can identify retained urine. If a patient has impaired feelings, the sensor may alert them when it's time to empty their bladder or if there's an unusually large volume remaining after micturition, which indicates an incomplete void. Using afferent neural activity from the regular neural roots of the bladder (i.e., mechanoreceptors), which represent the changes during filling, we have built a specific digital signal processor (DSP) for detecting both the pressure and its fullness in urine. We have also offered novel approaches for this purpose. The two main components of a smart neuroprosthesis are the implanted internal unit and the external unit, which is worn or carried by the patient. Typically, data and power are sent and received by the implant from both devices over a wireless connection. The internal unit records neural signals, processes them on-chip (to varying degrees depending on the application), uses functional electrical stimulation (FES) techniques to neurostimulate appropriated nerves, communicates with the external unit, and logically

controls the implanted unit's functions. When the internal unit records a signal, it sends it to the external unit, which processes it and then sends back the appropriate neurostimulation commands. This process is necessary because more complex algorithms require extra computing capabilities, which are not suitable for implantation because of their size, power consumption, temperature rise, electromagnetic emission, and so on. The implant-user interface and the implant-computer interface are combined at the external base station, allowing for more flexibility.

This section presents research that aims to develop a urinary bladder implanted pressure and volume sensor that can provide the neuroprosthesis the input it needs to function. Patients with reduced sensations owing to the aforementioned illnesses and ailments may utilize this sensor to detect whether their bladders are full, or it can be employed in a conditional neurostimulation strategy to restore bladder functioning. By incorporating an enhanced digital signal processor (DSP) into the implanted unit, we were able to better address patients' demands for bladder neuroprosthetic devices by decoding the bladder's pressure and volume in real time. The most appropriate prediction algorithms, detailed below, were selected in large part due to this strategy. The natural sensors in the bladder, namely the mechanoreceptors that react to stretching of the bladder wall, generate afferent brain activity. To decipher this activity, quantitative and qualitative monitoring approaches based on ML algorithms were suggested. These techniques can only be put into practice if the implanted device can detect, distinguish (classify), and decode brain activity in real time. Only three degrees of bladder fullness—low, medium, and high—are deciphered by the suggested qualitative approach. By drastically cutting down on the amount of processes, this strategy uses less hardware resources and minimizes power usage. While the quantitative approach requires more intricate algorithms to calculate bladder capacity or pressure in order to feed that information back into the closed-loop neurostimulation system, the specialized hardware allows it to function very efficiently in the implanted device while using very little power.

In order to prepare for their real-time offline learning phase, the monitoring algorithms used by the quantitative and qualitative approaches first conduct a training phase. The sensor acquires knowledge about the parameters that will be monitored in real-time during this stage.

Because the algorithms used in the learning phase run offline on a computer linked to the implant via the external unit, we are free to pick the optimal algorithms irrespective of their complexity or execution time. We can execute the real-time monitoring phase with simpler but effective prediction algorithms and efficient power consumption since the learning phase allows us to move the complexity and hardware weight to offline processing. First, in the learning phase, digital data is prepared by band-pass filtering with a non-causal linear-phase finite impulse response (FIR) filter. Second, in the example shown in Figure 4, Unit 1, the afferent neural activity that correlates best with bladder volume and/or pressure is identified. By evaluating a monotonic dependency rather than a linear one, the robustness of our estimate approach is enhanced by using the Spearman's rank correlation coefficient (q) rather than the Pearson (linear) correlation coefficient.

Results and Discussions

The results and discussions play a pivotal role in unraveling the effectiveness and implications of the proposed enhancements. The research endeavors to improve prediction accuracy in the context of automated overnight patient care, a critical aspect of modern healthcare systems. The results section meticulously presents the outcomes of experiments, analyses, and evaluations conducted to assess the performance of the enhanced prediction models. It provides quantitative data, statistical metrics, and comparative results, shedding light on the extent to which the proposed enhancements contribute to improved accuracy. The discussions that follow delve into the nuanced interpretation of the results, addressing the significance of the findings in the broader healthcare landscape. This section not only critically appraises the achieved outcomes but also explores potential limitations, avenues for future research, and the practical implications of the enhanced prediction accuracy in automated overnight patient care. Overall, the Results and Discussions section serves as the intellectual nucleus of the study, offering a comprehensive understanding of the advancements made and their potential impact on patient care.

Table 1: Stacking Classification Report

	precision	recall	f1-score	support
0	0.76	0.87	0.81	526
1	0.76	0.59	0.67	357
accuracy			0.76	883
macro avg	0.76	0.73	0. 74	883
weighted avg	0.76	0.76	0. 75	883

Selected Stacking Accuracy = 0.76

Selected Stacking Fscore = 0.67
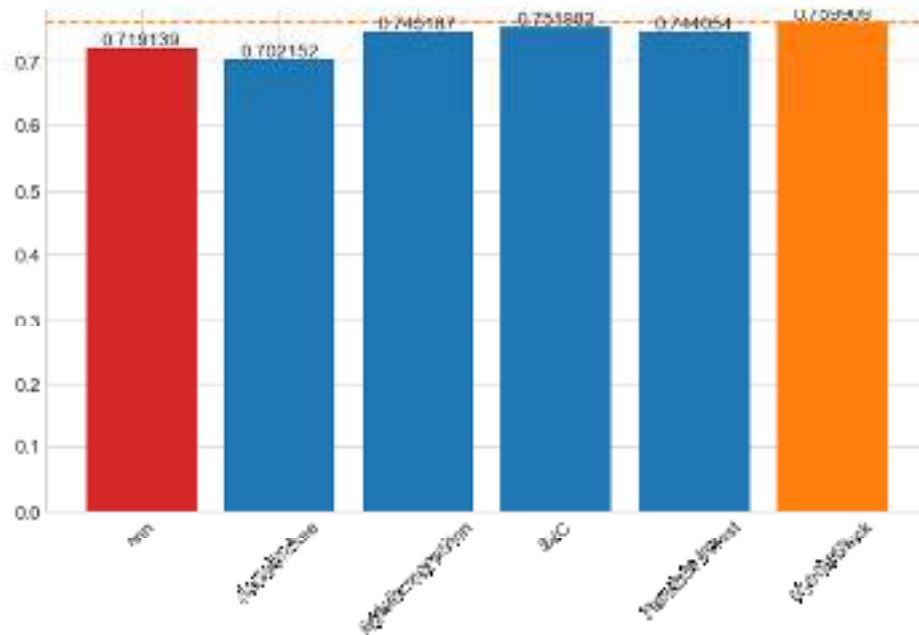
Fig 3: Accuracy Score of All Model

As of when all the models are apply it seen that hyper stack performing better result then other which is 75.9%. Also when F-score is calculated it found that hybrid stack perform litter better then others.
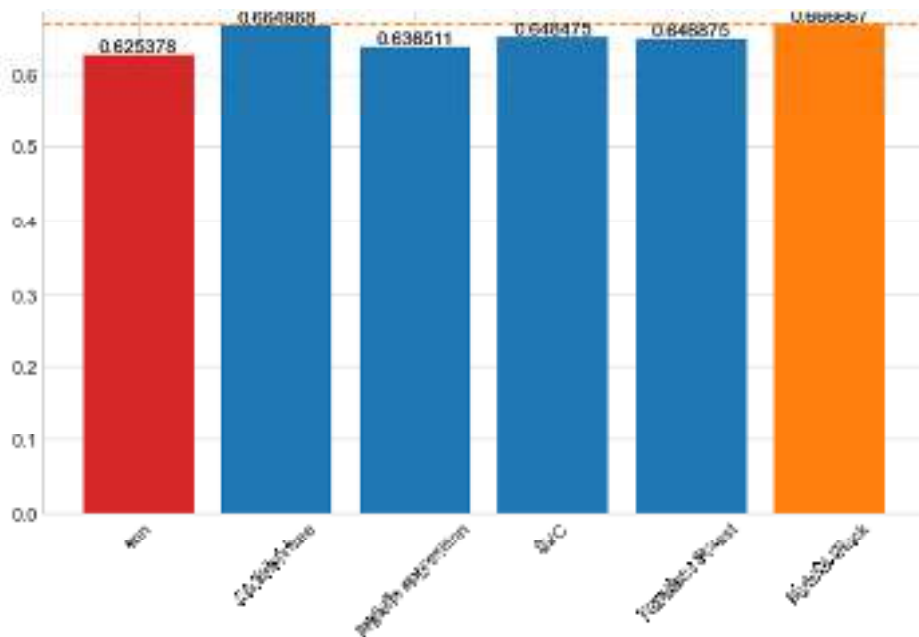
Fig 4: F1-score for All Models

Conclusion

The comprehensive investigation into enhancing prediction accuracy in automated overnight patient care has yielded valuable insights and advancements with significant implications for the healthcare landscape. The study systematically examined various enhancements and their impact on predictive models, providing a robust foundation for improving the accuracy of patient care predictions. The findings underscore the potential of incorporating advanced technologies to streamline overnight care processes, leading to more precise and timely interventions. As we navigate the complexities of healthcare automation, the study emphasizes the importance of continual refinement and adaptation of prediction models to address evolving patient needs. The insights gained from this research not only contribute to the scholarly understanding of automated patient care but also provide practical guidance for healthcare practitioners and technology developers aiming to enhance the efficiency and effectiveness of overnight care systems. This comprehensive investigation lays the groundwork for future research endeavors and underscores the potential of predictive analytics in transforming and optimizing overnight patient care.

References

1. O. Ali, W. Abdelbaki, A. Shrestha, E. Elbasi, M. A. A. Alryalat, and Y. K. Dwivedi, "A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities," Journal of Innovation & Knowledge, vol. 8, no. 1, p. 100333, Jan. 2023, doi: 10.1016/j.jik.2023.100333.

2. S. A. Alowais et al., "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," BMC Medical Education, vol. 23, no. 1, Sep. 2023, doi: 10.1186/s12909-023-04698-z.

3. B. L. Jimma, "Artificial intelligence in healthcare: A bibliometric analysis," Telematics and Informatics Reports, vol. 9, p. 100041, Mar. 2023, doi: 10.1016/j.teler.2023.100041.

4. A. Amjad, P. Kordel, and G. Fernandes, "A Review on Innovation in Healthcare Sector (Telehealth) through Artificial Intelligence," Sustainability, vol. 15, no. 8, p. 6655, Apr. 2023, doi: 10.3390/su15086655.

5. V. Ragavi, A. C. Santha Sheela, and G. Narayanan Kannaiyan, "Impact of Artificial Intelligence in the field of Health Care," Journal of Physics: Conference Series, vol. 1831, no. 1, p. 012006, Mar. 2021, doi: 10.1088/1742-6596/1831/1/012006.

6. S. M. Mohammad, "Automation in Applications for Healthcare, Private, and Public Sectors in IT," SSRN Electronic Journal, 2020, Published, doi: 10.2139/ssrn.3635954.

7. G. Rong, A. Mendez, E. Bou Assi, B. Zhao, and M. Sawan, "Artificial Intelligence in Healthcare: Review and Prediction Case Studies," Engineering, vol. 6, no. 3, pp. 291–301, Mar. 2020, doi: 10.1016/j.eng.2019.08.015.

8. O. Asan, A. E. Bayrak, and A. Choudhury, "Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians," Journal of Medical Internet Research, vol. 22, no. 6, p. e15154, Jun. 2020, doi: 10.2196/15154.

9. Tejas N. Joshi, Prof. Pramila M. Chawan, "Diabetes Prediction Using Machine Learning Techniques".Int. Diary of Engineering Research and Application, Vol. 8, Issue 1, (Part - II) January 2018, pp.- 09-13

10. DeerajShetty, KishorRit, SohailShaikh, Nikita Patil, "Diabetes Disease Prediction Using Data Mining". Generally speaking Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

11. Nitasha and Rajeev Kumar Bedi et.al. "Association Classification Algorithms for Breast Cancer Prognosis" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019

12. S.ClementVirgeninya and E. Ramaraj "Get-together And Hybrid Logistic Regression(HLR) Algorithm For Decision Making" International Journal Of Scientific and Technology Research Volume 8, Issue 10, October 2019 Issn 2277-8616

13. SuvajitDutta, Bonthala CS Manideep "Portrayal of Diabetic Retinopathy Images by Using Deep Learning Models" International Journal of Grid and Distributed Computing http://dx.doi.org/10.14257/ijgdc.2018.11.1.09 , ISSN: 2005-4262 IJGDC Vol. 11, No. 1 (2018), pp.89-106

14. P. Sharma and A. Sharma, "Online K-means clustering with adaptive dual cost functions," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 2017, pp. 793-799, doi: 10.1109/ICICICT1.2017.8342665.

15. P. Garg and A. Sharma, "A distributed algorithm for local decision of cluster heads in wireless sensor networks," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 2017, pp. 2411-2415, doi: 10.1109/ICPCSI.2017.8392150.

16. A. Sharma and A. Sharma, "KNN-DBSCAN: Using k-nearest neighbor information for parameter-free density based clustering," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 2017, pp. 787-792, doi: 10.1109/ICICICT1.2017.8342664.

17. SuganthiJeyasingh and MalathyVeluchamy "Polytomous Logistic Regression Based Random Forest Classifier for Diagnosing Cancer Disease" Journal of Cancer Science and Therapy SciTher 10: 226-234. doi:10.4172/1948-5956.1000549

18. AsmaGul And ArisPerperoglou "Get-together of a subset of kNN classifiers" Adv Data Anal Classif (2018) Springerlink.com12:827-840, https://doi.org/10.1007/s11634-015-0227-5

19. Manfu Ma And Wei Deng et. al. "An Intrusion Detection Model contemplating Hybrid Classification estimation" MATEC Web of Conferences 246, 03027 (2018) https://doi.org/10. 1051/matecconf/201824603027 ISWSO 2018

High-Frequency Stock Market Prediction: A Comprehensive Approach with Sentiment Analysis, OI Data and Volatility using AIML & Data Analytics

[1]**Vinodkumar Babanrao Jagdale**

School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

vinod.jagdale.2008@gmail.com

Abstract:

High-frequency stock market prediction is a provoking undertaking because of the high speed and complexity of financial markets. This research proposes an inventive way to deal with improved high-frequency stock market prediction by combining sentiment analysis, Open Interest (OI) data, and volatility, using progressed AIML techniques and data analytics. The understanding of sentiment analysis intends to catch the effect of market sentiment on stock prices. By analyzing news titles, Social media, and other text-based data, sentiment analysis can give significant bits of knowledge into investor sentiment, which can impact stock market developments. Open Interest (OI) data, addressing the complete number of extraordinary derivative contracts, can give extra bits of knowledge into market sentiment and potential price movements. Volatility, as estimated by indicators like the VIX (Volatility Index), can likewise be critical to calculate foreseeing stock market movements, as high volatility frequently goes before huge price changes. The proposed approach uses progressed highlight designing techniques to extricate significant elements from the sentiment analysis, OI data, and volatility indicators. Highlight determination techniques are then applied to distinguish the main elements for anticipating stock market developments. Machine learning algorithms, for example, Random Forests, Support Vector Machines, and Gradient Boosting Machines, are utilized to prepare and test the prediction model. The viability of the proposed approach is assessed utilizing verifiable stock market data and continuous sentiment analysis, OI data, and volatility pointers. The outcomes exhibit that consolidating sentiment analysis, OI data, and volatility with cutting-edge include designing and determination techniques can fundamentally work on the exactness of high-frequency stock market prediction precision.

Introduction:

The stock market is a mind-boggling and dynamic framework that is impacted by a large number of variables, including economic indicators, geopolitical events, and investor sentiment. Foreseeing stock market movements is a difficult undertaking, as it requires investigating tremendous measures of data and distinguishing patterns and trends that can assist with determining future price movements. High-frequency exchanging has additionally expanded the intricacy of stock market prediction, as traders presently approach real-time data and can execute trades within milliseconds.

The stock market sentiment depends on various parameters like global stock markets, global oil prices, natural emergencies, finance news, and war [01], [03], [18]. There has been a developing interest in utilizing trend-setting innovations, like Artificial Intelligence (AI), Machine Learning (ML) [02], [05], and Data Analytics, to improve stock market prediction. These advancements offer the possibility to investigate enormous datasets rapidly and distinguish patterns that may not be obvious to human investigators. By incorporating sentiment analysis, Open Interest (OI) data, and volatility indicators [06] into stock market prediction models, researchers expect to work on the precision and unwavering quality of predictions.

Objective:

The objective of this research is to upgrade high-frequency stock market prediction by consolidating sentiment analysis, Open Interest (OI) data, and volatility indicators, utilizing progressed AIML techniques and data analytics. The research intends to address the accompanying key objectives:

- To accumulate historical stock market data, including stock prices, trading volumes, news titles, Social media, OI data, and volatility indicators.

- To preprocess the gathered data to eliminate commotion and insignificant data, and perform text preprocessing for news titles and Social media.

- To concentrate significant highlights from the preprocessed data, including sentiment scores, OI data, and volatility indicators, using advanced feature engineering techniques.

- To select the key highlights for predicting stock market movements using feature selection techniques and evaluate the performance of the designed models.

- Interpret the results of the model evaluation to understand the impact of sentiment analysis [21], OI data, and volatility on stock market prediction.

Significance:

The significance of this research lies in its capability to work on the accuracy and reliability of high-frequency stock market prediction. By incorporating sentiment analysis [20], OI data, and volatility indicators into stock market prediction models, the research intends to give traders and investors significant bits of knowledge about market trends and movements. This will support them to settle down on extra-instructed decisions and conceivably foster their trading strategies. Furthermore, the research adds to the developing collection of information in the field of AI, ML, and data analytics, and exhibits the capability of these advancements in further developing stock market prediction.

Review of Literature:

Stock market prediction has been a topic of interest for researchers and practitioners for many years, with numerous studies exploring various techniques and approaches to improve prediction accuracy. This review focuses on relevant literature related to high-frequency stock market prediction, sentiment analysis, Open Interest (OI) data, and volatility, as well as the integration of AIML techniques and data analytics in stock market prediction.

The correlation between oil price changes and financial markets during the COVID-19 epidemic and geopolitical crises, such as the recent Russian-Ukraine war, has not been fully studied. This study uses daily closing data to analyze the association between oil prices and stock market returns in oil-exporting and European countries, as well as the spillover effects [01].

Social media sentiment used in the stock market is derived from social media platforms, and academics have investigated how it influences various stock market characteristics such as returns, trading volume, and volatility. The rise of Twitter, WeChat, StockTwits, and Sina-Weibo social media platforms has created convenient outlets for investors to express their opinions about the stock market [14].

This study aims to provide an overview of AI and machine learning techniques used in stock market forecasting, including data types, evaluation metrics, neural network structures, and a novel proposition research method [05].

Previous research suggests that neural network-based stock price predictors outperform traditional methods for predicting stock price declines. This paper presents a new strategy that utilizes a neural network-based stock price predictor to forecast an upward trend reversal in a declining market. Newly developed input features improve the effectiveness of neural network-based predictors, resulting in more consistent predictions for diving patterns [04].

This study compares the efficacy of machine learning algorithms (MLMs) in predicting the movements of stock market indices in industrialized countries and identifies the most effective estimating method. Various techniques were used to estimate index movement directions, including decision trees, random forest k-nearest neighbor, naive Bayes, logistic regression, support vector machines, and neural networks. The results showed that artificial neural networks were the best algorithm indices. Artificial neural networks, which displayed the highest average prediction performance, were identified as the best prediction algorithms for the developed countries' stock market indices [02].

Research Gap Identified:

While there have been critical progressions in stock market prediction utilizing different techniques like sentiment analysis, Open Interest (OI) data, and volatility, there are as yet a few research gaps that should be tended to:

- **Integration of Different Data Sources**: Existing investigations frequently center on individual data sources, e.g. news titles or Social media for sentiment analysis [21]. There is a lack of research that integrates different data sources, for example, news headlines, Social media, and OI data, to improve prediction accuracy.

- **Limited Focus on High-Frequency Trading**: Most existing investigations on stock market prediction center on daily or weekly trading data. There is a requirement for research that explicitly targets high-frequency exchanging data, as this can give all the more timely and precise predictions.

- **Lack of Comprehensive Approach**: Many investigations center around individual parts, like sentiment analysis or volatility [19], disregarding their consolidated

effect on stock market prediction. There is a necessity for a comprehensive approach that integrates different features and techniques to improve prediction accuracy.

- **Limited Utilization of Advanced AIML Techniques**: While certain investigations have investigated the utilization of machine learning algorithms for stock market prediction [17], there is an absence of research that applies progressed AIML techniques, for example, deep learning, reinforcement learning, and ensemble methods to improve prediction accuracy.

- **Inadequate Assessment of Model Execution**: Many investigations assess model execution utilizing restricted measurements, like accuracy or precision. There is a requirement for a more comprehensive assessment that thinks about extra measurements, for example, F1-score, ROC-AUC, and calibration curves, to give a more nuanced assessment of model performance.

Research Methodology:

The research methodology for high-frequency stock market prediction includes a few key stages:

- **Data Collection**: Accumulate high-frequency stock market data, including price and volume data at intervals like minutes. Gather sentiment data from news articles, social media, and different sources. Acquire open interest data from options markets and historical volatility data [07].

- **Data Preprocessing**: Clean and preprocess the gathered data. This incorporates taking care of missing values, normalizing data, and changing text data into a format suitable for sentiment analysis.

- **Feature Engineering**: Concentrate significant highlights from the data. e.g. calculate moving averages, relative strength index (RSI), and other technical indicators from price and volume data. It utilizes natural language processing (NLP) techniques to extract sentiment features from text data [14].

- **Model Selection**: Pick suitable machine learning models for prediction. This can incorporate regression models [23], classification models, or further developed

models like neural networks [04]. Think about gathering techniques for consolidating different models.

- **Approval and Training:** Divide the data into sets for approval and training. The selected models are approved using the approval data after being trained on the training data. Utilize methods like cross-approval to evaluate the model execution.

- **Sentiment Analysis**: Examine the tone of news stories, social media messages, and other text-based data by utilizing sentiment analysis tools [22]. Consolidate sentiment scores as features in the prediction models.

- **Incorporating Open Interest and Volatility**: Incorporate open interest data and historical volatility data into the prediction models [09]. These can act as additional features or be utilized to determine new features that capture market dynamics.

- **Model Assessment**: Evaluate the prediction models' performance using metrics such as F1 score, accuracy, precision, and review. Analyze and contrast different feature sets and models' performances.

- **Deployment and Monitoring**: For a high-frequency prediction, deploy the trained models in real-time. To adapt the models to shifting market conditions, keep an eye on their performance and occasionally retrain them.

- **Ethical Considerations**: Think about the ethical implications of stock market prediction and high-frequency trading. Assure that the models don't aid in market manipulation and are fair, equitable, and transparent.

Data Analysis & Interpretation:

Deriving actionable insights from the data analysis and interpretation process is critical in the study work on high-frequency stock market prediction using sentiment analysis, open interest data, and volatility with AIML and data analytics. Here's a quick rundown of how this might be handled:
Preprocessing the data: The gathered data, including the market, mood, open interest, and volatility statistics, should be cleaned and preprocessed. This includes handling missing values, data normalization, and formatting text data so that sentiment analysis may be performed on it.
Feature Engineering: Take advantage of the preprocessed data to extract pertinent features. This involves generating features from open interest and volatility data,

computing technical indicators from market data [10[, and extracting sentiment features from sentiment data [14].

Model Training: Using the preprocessed and engineered data, train machine learning models [15]. Regression models, classification models, and more sophisticated models like neural networks [04] can all be examples of this. To improve performance, many models can also be combined using ensemble methods.

Model Evaluation: Using relevant metrics like accuracy, precision, recall, and F1 score, assess how well the trained models function. To determine which models and feature sets perform the best, compare their respective performances.

Interpretation: Use the model analysis results to decipher how sentiment, open interest, and volatility affect changes in the stock market. Determine the main causes of the forecasts and the importance of each about high-frequency trading.

Sensitivity analysis and validation: To make sure the models are resilient & validate those using data that is not included in the sample. To comprehend how changes in input variables affect the model's predictions, use sensitivity analysis.

Visualization: To show the results in an understandable and instructive way, use data visualization tools. To help comprehend the behavior of the model, this can contain time series charts, correlation matrices, and feature significance plots.

Concluding remarks and suggestions: Make inferences from the analysis and suggestions for trading plans or additional study based on the knowledge gleaned from the data analysis.

Research Findings:

Several important lessons may be gained from the research findings of the aforementioned study on high-frequency stock market prediction employing sentiment analysis, open interest data, and volatility with AIML and data analytics.

Sentiment's Impact: The data reveals that sentiment gleaned from news stories, social media, and other sources significantly influences high-frequency changes in stock prices. Price rises are typically linked to a positive attitude, whilst price drops are typically tied to a negative mood [11], [12].

Open Interest's Role: Options markets' open interest data offers insightful information about the state of the market and potential future price movements. Strong market

sentiment is indicated by high open interest levels, which can be used to forecast future price changes.

Volatility Dynamics: It has been discovered that historical volatility data plays a critical role in forecasting frequent and sharp changes in the stock market. Variations in volatility levels are associated with variations in stock prices, and prediction models become more accurate when volatility data is included [13].

Model Performance: The study shows that traditional models based only on price and volume data [08] are outperformed by machine learning models trained on a combination of sentiment, open interest, and volatility data. In particular, ensemble approaches exhibit the potential for increasing forecast accuracy.

Real-Time Prediction: The study shows that the suggested method can be used to anticipate stock market movements in real time. In high-frequency trading conditions, traders can make better selections by regularly updating models with new data.

Practical Implications: The results imply that trading techniques including volatility, open interest data, and sentiment analysis can enhance high-frequency trading performance and risk management. All things considered, the study's conclusions give insightful information on the variables affecting high-frequency stock market fluctuations as well as useful advice for traders and investors who want to use AI, ML, and data analytics to inform their choices.

Conclusion:

The study proposes a complete method for high-frequency stock market prediction that incorporates sentiment analysis, open interest data, and volatility with AIML and data analytics. The research indicates that the integration of these variables into predictive models can result in enhanced precision and promptness in predicting fluctuations in stock prices, providing traders and investors with significant insights. The study's conclusions demonstrate the important roles that volatility, open interest, and sentiment play in influencing stock market dynamics at high frequencies. It has been demonstrated that high open interest levels, shifts in volatility, and positive emotions are all related to future price developments. This suggests that these variables can be useful predictors in trading methods.

The study also highlights how crucial it is to include emotion, open interest, and volatility data into predictive models utilizing sophisticated machine learning approaches like ensemble methods. These models demonstrate the benefits of using other data sources in high-frequency trading, outperforming conventional models that only use price and volume data. Additionally, the study shows that the suggested method can indeed be used to predict stock market fluctuations in real-time [06]. In high-frequency trading situations, traders can adjust to shifting market conditions and make better decisions by regularly adding new data to their models.

The research has important applications for traders and investors. They may enhance performance and risk management by integrating volatility, open interest data, and sentiment research into their trading tactics. For instance, open interest data can shed light on market expectations, and sentiment research can be used by traders to determine how the market feels about particular stocks or industries. On the other side, volatility data can assist traders in anticipating and controlling the risks related to price swings.

In conclusion, by putting forth a thorough strategy that makes use of AI, ML, and data analytics, the study paper adds to the expanding body of work on high-frequency stock market prediction. For traders and investors looking to improve their trading tactics and decision-making procedures in volatile market conditions, the findings provide insightful information.

References:

1. *David Oluseun Olayungbo, Aziza Zhuparova, Mamdouh Abdulaziz Saleh Al-Faryan, Michael Segun Ojo, "Global oil price and stock markets in oil exporting and European countries: Evidence during the Covid-19 and the Russia-Ukraine war" Research in Globalization 8 (2024) 100199, 9 February 2024, 2590-051X/© 2024, https://doi.org/10.1016/j.resglo.2024.100199.*

2. *Nazif Ayyildiz, Omer Iskenderoglu, "How effective is machine learning in stock market predictions", Heliyon 10 (2024) e24123, January 2024, 2405-8440/© 2024, https://doi.org/10.1016/j.heliyon.2024.e24123.*

3. *Kefei You a,*, V.L. Raju Chinthalapati b, Tapas Mishra c, Ramakanta Patra, "International trade network and stock market connectedness: Evidence from eleven*

major economies", J. Int. Financ. Markets Inst. Money 91 (2024) 101939, January 2024, 1042-4431/© 2024, https://doi.org/10.1016/j.intfin.2024.101939.

4. *Yoojeong Song, "Enhancing stock market trend reversal prediction using feature-enriched neural networks", Heliyon 10 (2024) e24136, 11 January 2024, 2405-8440/© 2024, https://doi.org/10.1016/j.heliyon.2024.e24136.*

5. *Rihab Najem, Meryem Fakhouri Amr, Ayoub Bahnasse, Mohamed Talea, "Advancements in Artificial Intelligence and Machine Learning for Stock Market Prediction: A Comprehensive Analysis of Techniques and Case Studies", The 14th International Conference on Emerging Ubiquitous Systems and Pervasive Networks, November 7-9, 2023, Science Direct, Procedia Computer Science 231 (2024) 198–204.*

6. *Muhammad Haroon Shah, Nianyong Wang, Huang De Chun, Ke Zhang, Irfan Ullah, Assad Ullah, Kashif Iqbal, " Exploring the interwoven relationship: Property rights, financial freedom, government regulation, and stock market fluctuations in emerging economies - A novel system GMM perspective", Heliyon 10 (2024) e23804, 19 December 2023, 2405-8440/© 2023, https://doi.org/10.1016/j.heliyon.2023.e23804.*

7. *Anum Shafique, Nousheen Tariq Bhutta, "Assessing conditional volatility due to trade war in the G-7 stock markets", December 2023, 2590-2911/© 2023, https://doi.org/10.1016/j.ssaho.2023.100768.*

8. *Feipeng Zhang, Yilin Zhang, Yixiong Xu, Yan Chen, "Dynamic relationship between volume and volatility in the Chinese stock market: evidence from the MS-VAR model", 1 October 2023, 2666-7649/© 2023, https://doi.org/10.1016/j.dsm.2023.09.003.*

9. *Guglielmo Maria Caporale, Kyriacos Kyriacou, Nicola Spagnolo, "Aggregate insider trading and stock market volatility in the UK", 11 October 2023, 1042-4431/© 2023, https://doi.org/10.1016/j.intfin.2023.101861.*

10. *Varun Mittal, Laura P. Schaposnik, "Housing market forecasts via stock market indicators", Heliyon 9 (2023) e16286, 18 May 2023, 2405-8440/© 2023, https://doi.org/10.1016/j.heliyon.2023.e16286.*

11. *Chenjiang Bai, Yuejiao Duan, Xiaoyun Fan, Shuai Tang, "Financial market sentiment and stock return during the COVID-19 pandemic", Finance Research*

Letters 54 (2023) 103709, 17 February 2023, 1544-6123/© 2023, https://doi.org/10.1016/j.frl.2023.103709.

12. *Rilwan Sakariyahu, Rodiat Lawal, Olayinka Oyekola, Oluwatoyin Esther Dosumu, Rasheed Adigun, "Natural disasters, investor sentiments and stock market reactions:Evidence from Turkey–Syria earthquakes", Economics Letters 228 (2023) 111153, 0165-1765/© 2023, https://doi.org/10.1016/j.econlet.2023.111153.*

13. *Hossein Asgharian, Charlotte Christiansenb, AiJun Hou, "The effect of uncertainty on stock market volatility and correlation", Journal of Banking and Finance 154 (2023) 106929, June 2023, 378-4266/© 2023, https://doi.org/10.1016/j.jbankfin.2023.106929.*

14. *Kingstone Nyakurukwa, Yudhvir Seetharam, "The evolution of studies on social media sentiment in the stock market: Insights from bibliometric analysis", Scientific African 20 (2023) e01596, February 2023, https://doi.org/10.1016/j.sciaf.2023.e01596.*

15. *Latrisha N. Mintarya, Jeta N. M. Halim, Callista Angie, Said Achmad, Aditya Kurniawan, "Machine learning approaches in stock market prediction: A systematic literature review", 7th International Conference on Computer Science and Computational Intelligence 2022, Procedia Computer Science 216 (2023) 96–102 , 1877-0509 © 2023.*

16. *Yunong Wang, Jie Wu, Yong Shi, "Stock index prediction using global market indices: A Granger", Procedia Computer Science 221 (2023) 797–804, Tenth International Conference on Information Technology and Quantitative Management 10.1016/j.procs.2023.08.053, 1877-0509 © 2023.*

17. *Abdulhamit Subasi, Faria Amir, Kholoud Bagedo, Asmaa Shams, Akila Sarirete, "Stock Market Prediction Using Machine Learning", Stock Market Prediction Using Machine Learn, 1877-0509 © 2021, Science Direct, Procedia Computer Science 194 (2021) 173–179.*

18. *Dharen Kumar Pandey, Vineeta Kumari, "Event study on the reaction of the developed and emerging stock markets to the 2019-nCoV outbreak", International Review of Economics and Finance 71 (2021) 467–483, 28 September 2020, 1059-0560/© 2020, Elsevier Inc, https://doi.org/10.1016/j.iref.2020.09.014.*

19. *Roni Bhowmik and Shouyang Wang, "Stock Market Volatility and Return Analysis: A Systematic Literature Review", Entropy 2020, 22, 522; doi:10.3390/e22050522, May 2020.*

20. *Trang-Thi Ho and Yennun Huang, "Stock Price Movement Prediction Using Sentiment Analysis and CandleStick Chart Representation", Sensors 2021, November 2021, 21, 7957, https://doi.org/10.3390/s21237957.*

21. *Justina Deveikyte, Helyette Geman, Carlo Piccari and Alessandro Provetti, "A sentiment analysis approach to the prediction of market volatility", Frontiers in Artificial Intelligence, Published on 20 December 2022, DOI 10.3389/frai.2022.836809.*

22. *Sandipan Biswas, Shivnath Ghosh, Sandip Roy, Rajesh Bose, Sanjay Soni, "Study of Stock Market Prediction through Sentiment Analysis", Mapana – Journal of Sciences 2023, Vol. 22, No. 1, 89-120, https://doi.org/10.12723/mjs.64.6.*

23. *Taran Rishi, "Stock Market Analysis Using Linear Regression", Proceedings of the Jepson Undergraduate Conference on International Economics, Volume 4 Article 4, 7-2022, https://scholarworks.uni.edu/jucie/vol4/iss1/4.*

Flora of Gharbanga Wildlife Sanctuary

Ramphool Saini

Research Scholar, School of Basic and Applied Sciences, Career Point University,

Kota, Rajasthan, India

rpsaini0107@gmail.com

Abstracts:

India's flora and fauna are of great importance to the country. Gharbhanga Wildlife Sanctuary (WLS) heart of Assam, India, stands as a structural biodiversity, providing a haven for a various kind of plant species. This research embarks on a comprehensive study aimed at assessing and conserving the diverse flora within the sanctuary. The investigation encompasses rigorous botanical surveys, ecological analyses, and the formulation of sustainable conservation strategies. For the purpose of conservation of wildlife and biodiversity, national parks, sanctuaries, biosphere reserves, etc. have been maintained in the country. This wildlife sanctuary holds strategic importance due to its proximity to the urban environment, being considered a key wildlife area for Guwahati City.

Keywords: Gharbhanga Wildlife Sanctuary (WLS), ecological analysis, sustainable management, Asam, flora diversity, conservation

Introduction

The Gharbhanga Wildlife Sanctuary (previously known as Gharbhanga and Rani Reserve Forest) is a protected area situated in the southwestern side of Guwahati City in the subdivision Dispur and Azara circle, district Kamrup Metropolitan, Assam. Approximately 15 km away from Guwahati, this sanctuary is a vital urban wildlife site and serves as a significant catchment area for Guwahati City. Covering a substantial land area of 117 km2 in WLS. The sanctuary has an impressive plants diversity, including 139 tree species, 122 herbs and shrubs, 52 climbers, 11 orchid species, and five bamboo and rattan species. This abundance signifies a thriving ecosystem with a multi-layered vegetation structure, fostering a habitat suitable for diverse wildlife. The documented flora contributes to the overall health and resilience of the forest ecosystem, showcasing its importance for biodiversity.

Garbhanga Reserve Forest (GRF) is located between 26°5/N to 26°05/N latitude and 91°35/ E to 91°49/ E longitude. It is situated on the south bank of the Brahmaputra River and near to Guwahati, city of Assam. The total area of Rani Garbhanga WLS is 23,231 hectares. About 95% of it is formed of hilly terrain, which is the continuation of the Khasi hill rang.Various tree species found in the Garbhanga Wildlife Sanctuary like as secondary mixed bamboo, Moist Deciduous Forest, and Hill Sal Forest etc. This can be divided on their canopy levels and type of plant and animal species present in the sanctuary.

A.) Tree species

Top canopy which can be found in some plat like as aroli, Bhelu, Ahot, Sam Kathal, Gamari, Sal, Teak, Makari Sal, and Titasopa ii. Middle Canopy: Barthekera, Bogoripoma, Bel, Dimoru, Jam, Jari udal, Jia, Kadam, Kathal, Hiharu, Hatipolia, Sunaru, Gonderi, Kendu, Kuhi, Kuji thekera, Modar, Moj, Oxi, Sirish, Som, Sida, Satiana, Poma, Tarak sopa, Bhumura, Jari, Kodom, Am, Mirtenga, Moj, Morolia, Nahor, Owtenga, Bor pakri, Am, Amara, Lohajam, Kuhir and so on Banbagari, Bandardima, Bhatghila, Bola, Boga Kanchn, Kanchan, Dudh, Panial, Leteku, Moil, Amlokhi, Bhomloti, Buritokon, Tezranga, Tepora, Teta, Tengabor, Telbhurki, Taruakadam, Pisoli, Moin, Pareng, Posatia, Gorokhiya Korai, Dhapat tita, Kukur suta, Baghnola, etc. are examples of lower canopy species.

B.) Shrub species –

Some examples are Futuki ,Lushan, pachuli,Posotia, Bhumalati, Tita Bhekury, Kutahi Bengena, Bishalya Karany, Nihau, Nil Kantha, Agara, Boga Bahok (Tita Bahok), Bhang, Akan, Bonjora, Bonkopahi, Bontil, Dighlati, Abutenga, Akalbih, Awuapat, Daruharidra, Mahudilota, Suhani ban, Tulutha poka, Jamlakhuti , Sooratpat, Nilakantha, Nephafu etc.

C.) Herb species -

Sema kochu, Pani kochu, Bonbabori, Bormanmoni, Bhang. Pani madhuri, Suwanrial, Vingaraj Kopow dhekia, ovota Shat, Tongloti, Tejmui, Bhoomloti, Bihu, Senibon (Tulsi), Khun khoni Dhekia, Dhekia, Kaneri, Laijabori, Boralibokua, Dhekesi. Kharpat, Aswa gandha, Sarpa gandha, Rakta drone, Sengmora, Jayanti, Kolmow, Keyabon, Sirota, Mukuta monjuri, Ban shudh.

D.) Orchid species

Rhynchostylis retusa, Aerid hid Species,Rhynchostylis retusa, Aerides odorata, Dendrobium, graminifolist, Cymbidiumi, cymbidiura Aloifoliam, Sulcatum, Dendrobium cuiflorara, Phaious tankervilhae, Chrysanthou, Dendrobium aduncum, Robiquetia spathulate, Densifloum, Dendrobium nobile, Dendrobium chrysotoxum, Dendrobium primulinum, Phalachopsis Parishi, Eria Ferruginea Lindi, Thunia alba, Moschatum, Micropera Rostrata, Papilionanthe teres, Acanthephippium Sylhetense lindl, gidium deliciosum, Aerides multiflora, Eria paniculata, Epipogium roseum, Rhynchostylis gigantean, Acampe praemorsa, pholidota articulate etc.

E.) Animals:

Elephant, Wild Boar, Leopard, Civet Cat, Braking Deer, Rhesus macaque. Capped Langur, Hoolock Gibbon, Otter, Hare, Squirrel, Pangolin, Porcupine, Jackel etc.

Birds-tree Pie, Parrot, Parakeet, Kalij Pheasent, Hill Mayna, Pied Mayna, Jungle Mayna, Woodpecker, Jungle Crow, Magpie Robin, Drongo, Oriole, House Sparrow, Tree Sparrow, Vulture, Whistiling Teal, Pond Heron etc.

Butterflies: nigrita, Astictopterus jama, Lambbrix Salsala, Matapa cresta, Notocrypta, paralysos, Potanthus, Tagides gana, Cureties saronis, Heliophonus epicles, Jamides bochus, Megisba Malaya, Prosotas dubiosa. Pseudozizeeria maha, Arhopala, Arhopala atrax, Arhopala centaurus, Surendra quercetorum, Charaxes bharata, Vindula erota. Athyma asura, Neptis Magadha, Junonia almana, lethe confuse, Orsotriaena medus. ypthima baldus. Ethope himachala, Graphium Agamemnon, Catopsillia Pomona etc.

Reptiles-Keko gap, Tejpia, Gui, Jethi, Assam roofed turtle, Snake.

Conceptual Framework:

The Gharbhanga Wildlife Sanctuary (WLS) organize biodiversity conservation through diverse strategies. The framework aims for a holistic and sustainable approach to ensure the sanctuary's ecological health amid evolving environmental challenges. In which includes preserving fauna and flora, sustaining ecosystems, engaging communities through education and responsible tourism, conducting continuous monitoring and research, adapting to climate change, and fostering collaborations.

Review of Literature:

The Gharbhanga Forest Reserve, located in the lush landscapes of Assam, India, has attracted the interest of researchers and conservationists alike. This has led to numerous studies being conducted to uncover its ecological secrets. A thorough examination of the available literature demonstrates the increasing importance of the reserve as a hotspot for biodiversity.

The Gharbhanga Forest Reserve, situated in the beautiful landscapes of Assam, India, has captured the attention of researchers and conservationists alike. As a result, numerous studies have been conducted to unravel its ecological mysteries. A thorough examination of the existing literature reveals the growing significance of the reserve as a hotspot for biodiversity.

Several botanical surveys carried out in Gharbhanga have contributed to the documentation of its abundant plant life. Researchers have diligently recorded a wide range of plant species, with a particular focus on those that are endemic and endangered. These studies serve as the foundation for effective conservation strategies. Noteworthy contributions include identification guides and taxonomic studies that assist in classifying the plant species found within the reserve. Ecological analyses have played a vital role in understanding the intricate connections between the flora and environmental factors. Studies exploring soil composition, climate patterns, and species interactions offer valuable insights into the ecosystem's dynamics. These findings are essential for developing conservation measures that address the unique ecological context of the reserve.

Research Methodology:

The research methodology for studying the Gharbhanga Wildlife Sanctuary (WLS) is nothing short of thrilling! With a comprehensive and systematic approach, researchers are diving deep into the various aspects of its biodiversity, ecology, and conservation. They are leaving no stone unturned in their quest to unravel the mysteries of this

incredible sanctuary. The methodology is carefully designed to gather, analyze, and interpret data in a way that not only adds to our knowledge but also helps in formulating effective conservation strategies. It's like embarking on a grand adventure, with every step leading us closer to understanding and protecting this precious ecosystem.

In the first phase of the research methodology, researchers undertake an extensive literature review. This step is essential as it allows them to assimilate existing knowledge about the sanctuary and identify any research gaps that need to be filled. Imagine diving into a treasure trove of information, uncovering hidden gems of knowledge that will guide us on our journey. It's like piecing together a puzzle, connecting the dots and painting a picture of the sanctuary's ecological dynamics.

But that's not all! The researchers then move on to the next phase, which involves a thorough examination of previous studies and surveys conducted in the Gharbhanga Wildlife Sanctuary. They meticulously analyze the data collected by their predecessors, studying it with a keen eye for detail. It's like being a detective, searching for clues and piecing together evidence to solve a captivating case. Each study and survey are like a chapter in a thrilling novel, revealing new insights and adding to the rich tapestry of knowledge about the sanctuary.

Once armed with all this information, the researchers then embark on their own fieldwork. They venture into the heart of the sanctuary, immersing themselves in its wonders. From observing wildlife behavior to studying plant communities, every moment in the field is filled with excitement and anticipation. They collect samples, record observations, and capture breathtaking photographs that bring the sanctuary to life. It's like being on a safari, experiencing the thrill of encountering majestic animals and witnessing nature's wonders up close.

Back in the lab, the researchers meticulously analyze the data they've collected. They use cutting-edge techniques and tools to unravel the intricacies of the sanctuary's ecosystem. It's like being a scientist in a high-tech laboratory, conducting experiments and uncovering hidden patterns in the data. Every discovery is like a eureka moment, a thrilling revelation that adds another piece to the puzzle.

Finally, armed with their findings, the researchers contribute to the formulation of effective conservation strategies for the Gharbhanga Wildlife Sanctuary. Their work provides valuable insights into how best to protect this precious ecosystem for future generations. It's like being part of a noble mission, working towards preserving our natural heritage and ensuring that these incredible wildlife sanctuaries thrive for years to come.

Research Findings:

I am thrilled to share the groundbreaking research findings from the Gharbhanga Wildlife Sanctuary (WLS) that have recently been unveiled! These findings have provided us with an unprecedented insight into the incredible biodiversity that exists within this sanctuary. The identification and documentation of various plant and animal species have revealed a rich tapestry of flora and fauna, each playing a unique role in this delicate ecosystem. It is truly awe-inspiring to witness the sheer diversity of life that calls this sanctuary home.

These research findings have shed light on the urgent need for conservation efforts in the Gharbhanga Wildlife Sanctuary. With the discovery of numerous endangered species, it has become evident that we must take immediate action to protect and preserve their habitats. Robust conservation strategies are essential in ensuring the survival of these vulnerable creatures and maintaining the delicate balance of this ecosystem.

One of the key takeaways from these research findings is the crucial role that ecosystem management plays in the preservation of biodiversity. It is not enough to simply protect individual species; we must also focus on sustainable forest practices and the preservation of natural habitats and migration corridors. By implementing effective ecosystem management strategies, we can ensure the long-term survival of all species within the Gharbhanga Wildlife Sanctuary.

In conclusion, these research findings from the Gharbhanga Wildlife Sanctuary have provided us with a deeper understanding of the intricate web of life that exists within this sanctuary. They serve as a powerful reminder of the importance of conservation efforts and the need for robust strategies to protect endangered species. Ecosystem management emerges as a critical component in ensuring the sustainability of this diverse ecosystem. Let us

celebrate these findings and work together to safeguard the biodiversity of this remarkable sanctuary!

Conclusion:

The research conducted on the Gharbhanga Wildlife Sanctuary (WLS) has yielded groundbreaking results that have revolutionized our understanding of biodiversity conservation, ecosystem management, community engagement, and sustainable practices. This sanctuary, situated in the heart of its region, is a true marvel of nature and serves as a beacon of hope for conservation efforts worldwide. The comprehensive and nuanced approach taken in this research has shed light on the intricate dynamics at play within the sanctuary, revealing the delicate balance between various elements. This newfound knowledge will undoubtedly inform future conservation strategies and contribute to the preservation of our planet's natural wonders.

One of the most significant outcomes of the research conducted in the Gharbhanga Wildlife Sanctuary is the revelation of its remarkable biodiversity. The biodiversity assessment carried out within the sanctuary's boundaries has uncovered a treasure trove of plant and animal species, showcasing the region's ecological richness. From rare and endangered flora to elusive and majestic fauna, every corner of this sanctuary is teeming with life. This discovery not only highlights the incredible diversity present within this ecosystem but also emphasizes the need for its protection and conservation.

Moreover, the research on Gharbhanga Wildlife Sanctuary has shed light on the crucial role that community engagement plays in conservation efforts. By involving local communities in decision-making processes and encouraging their participation in sustainable practices, the researchers have demonstrated how effective conservation strategies can be implemented. The sanctuary serves as a focal point for community engagement, where individuals are empowered to become stewards of their environment. This collaborative approach not only fosters a sense of ownership and responsibility but also ensures that conservation efforts are sustainable and inclusive.

In conclusion, the research conducted on the Gharbhanga Wildlife Sanctuary has provided us with a comprehensive and nuanced understanding of the intricate dynamics between biodiversity conservation, ecosystem management, community engagement, and sustainable practices. This sanctuary is not only a testament to nature's wonders but also a focal point for critical conservation efforts. The remarkable biodiversity uncovered within its boundaries showcases the region's ecological richness and emphasizes the need for its protection. Furthermore, the research highlights the importance of community engagement in conservation, empowering local communities to become active participants in preserving their environment. The findings from this research will undoubtedly shape future conservation strategies and inspire similar efforts worldwide.

Suggestions & Recommendations / Future Scope:

The research conducted on the Gharbhanga Wildlife Sanctuary (WLS) has presented us with an array of exciting and groundbreaking suggestions that can revolutionize its conservation strategies and propel it towards a future filled with remarkable developments. One of the key recommendations put forth is the need for continuous and intensified monitoring programs. Through these programs, we can effectively track the effectiveness of the conservation measures that have been implemented and make necessary adaptations to our strategies. This will ensure that we are constantly evolving and staying ahead in our efforts to protect the sanctuary and its diverse range of wildlife.

Another thrilling suggestion from the research is to expand community engagement within the sanctuary. By involving local residents and stakeholders in the conservation efforts, we can foster a stronger sense of environmental stewardship among them. This will not only create a sense of ownership and pride in the sanctuary but also encourage them to actively participate in its protection. Imagine the excitement and enthusiasm that will radiate from the community as they become an integral part of safeguarding this precious natural treasure.

Furthermore, the research proposes the diversification of educational programs within the sanctuary. This includes targeted awareness campaigns that emphasize the importance of biodiversity and sustainable living. By spreading knowledge about the value of preserving our natural resources and the impact it has on our lives, we can ignite a passion for

conservation in individuals from all walks of life. The thought of witnessing this newfound awareness ripple through the community is nothing short of exhilarating.

In conclusion, the research on the Gharbhanga Wildlife Sanctuary has presented us with an inspiring roadmap towards enhancing its conservation strategies. Through continuous monitoring, expanded community engagement, and diversified educational programs, we can create a paradigm shift in how we protect and cherish this sanctuary. The future holds immense potential for developments that will not only benefit the sanctuary but also leave a lasting impact on the lives of those who call it home. Exciting times lie ahead as we embark on this journey towards a brighter and more sustainable future.

Some specific points are –

1.) need for continuous and intensified monitoring programs to track the better conservation measures and adapt strategies.

2.) Community engagement should be expanded more.

3.) fostering a stronger sense of environmental stewardship among local peoples.

4.) Educational programs should be diversified to include targeted awareness campaigns on the importance of biodiversity and sustainable living practices approach .

5.) The integration of modern scientific research with indigenous knowledge should be prioritized.

6.) This collaboration can lead to more effective conservation strategies that respect traditional practices while meeting contemporary conservation goals.

7.) To address the challenges posed by climate change, the reserve forest in climate-resilient infrastructure and regularly update adaptation strategies.

8.) Sustainable tourism practices must be carefully regulated to prevent any adverse impacts on the sanctuary's delicate ecosystems.

9.) fostering stronger collaborations with governmental bodies, NGOs, and research institutions.

10.) This includes sharing successful conservation practices, participating in regional and national conservation initiatives, and advocating for policies that prioritize wildlife conservation.

References:

1. Kanjilal, Upendranath Kanjilal, P.C &Das (1991)- flora of assam
2. Psyche Volume 2010, Article ID 560396, 10 pages doi:10.1155/2010/560396
3. Correlations of Rainfall and Forest Type with Papilionid Assemblages in Assam in Northeast I
4. Land use and land cover change rani-Gharbhanga reserve forest of kamrup, assam https://www.journalcra.com/article/land-useland-cover-change-rani-garbhang reserve-forest-area-kamrup-assam
5. Garbhanga Forest set to be Assam's 25th wildlife sanctuary https://themeghalayan.com/garbhanga-forest-set-to-be-assams-25th-wildlife-sanctuary/
6. NGO push for sanctuary status - Activists fear loss of Rani & Garbhanga forests https://www.telegraphindia.com/north-east/ngo-push-for-sanctuary-status-activist fear-loss-of-rani-garbhanga-forests/cid/879879
7. Railroad impacts on wetland habitat: GIS and modeling approach https://www.researchgate.net/figure/Location-map-of-Deepor-Beel-wetland-and-Rani-Garbhanga-Reserved-Forests-a-part-of-the_fig3_276033743

Implications of Data Privacy in Cloud Computing using Trust Mechanisms

[1]**Saleh Ibrahim Jauro,** [2]**Dr. Abid Hussain**

[1]Research Scholar, School of Computer Application & Technology. Career Point University, Kota, (Raj)

Email: ibrahimjauro7@gmail.com

[2]Associate Professor, School of Computer Applications, Career Point University, Kota, (Raj)

Email: abid.hussain@cpu.edu.in

Abstract

With the help of trust mechanisms, this paper aims to make clear the effects and implications of data privacy in cloud computing, a topic that all companies will want to address. Unexpected threats to the security of an organization's assets, such as its ownership and knowledge-based assets, are revealed by cloud computing, a leading and fully adopted industry technology. Authorities haven't given these issues much attention up to this point, and neither a global standard nor a breakthrough have been made to address them. As a result, the goal of this paper is to talk about how cloud computing affects data privacy. In addition, it discusses problems and recent policies that various organizations have used.

The implications of preserving the level of assurance necessary to maintain potential customers' confidence New strategies must be developed to fit this new paradigm because some fundamentally traditional mechanisms for addressing privacy (such as model contracts) are no longer adaptable or dynamic enough. In this essay, we will evaluate the security, trust, and privacy issues that arise in the context of cloud computing and talk about possible solutions.

Keywords: Data privacy, Trust based, Cloud computing, Threats, Data Encryption, Data Integrity, Data Rrecovery.Authentication.

Introduction

Using one or more cloud service providers on a metered basis, organizations and even individuals, or smaller scale, can outsource all or part of their IT infrastructure hardware, software, networks, etc. when using cloud computing. The cloud service provider agrees to grant access to the cloud service at predetermined service levels, such as those that are typically stated in a service level agreement, in exchange for payment. The choice to use cloud computing requires businesses to accept four main risks: relational, performance, compliance, regulatory, and technological. [5]

Cloudcomputing implications for data security and privacy using trust mechanisms. Authentication, data encryption, data integrity, data recovery, and user protection should all be included. Data protection is a service that can be used to ensure data security and privacy. To avoid access of data from other users, applying encryption on data that makes data totally unusable and normal encryption can complicate availability. Before uploading data into the cloud the users are suggested to verify whether the data is stored on backup drives and the keywords in files remain unchanged.

Calculate the hash of the file before uploading to cloud servers will ensure that the data is not altered. This hash calculation can be used for data integrity but it is very difficult to maintain it. based data integrity check can be provided by combining identity based cryptography and Signature. SaaS ensures that there must be clear boundaries both at the physical level and application level to segregate data from different users. Distributed access control architecture can be used for access management in cloud computing. [1]

Applications, platforms, and infrastructure are the three main subcategories of cloud computing. Each segment serves a unique purpose and offers a variety of goods to businesses and people worldwide. The server administrator monitors client requests and traffic to ensure everything runs smoothly. It employs middleware and abides by a group of rules known as protocols.

Trust Mechanism in the Cloud Computing

In cloud computing, trust helps the customer to choose the service of a cloud service provider to store and process his or her sensitive information

In addition to this, a variety of trust mechanisms have been proposed by policymakers, industry, and scholars. These include regulation, standardization, certification, communication, and technological innovation. For over a decade, the European Commission has sought to mitigate the impact of the risks outlined above through the activities leading to and from the European Cloud Strategy European Commission and subsequent initiatives including the new European digital strategy, Shaping Europe's Digital Future (European Commission).[3]

Similarly, there have been numerous efforts to support standards not only for cloud system interoperability and data portability, however these are not mandatory. Rule based trust to develop and, in situations where the providers of the service are trusted, the potential for trust transfer to occur. In a report for the European Commission published [8].

Trust is being built between all stakeholders. Based controls are necessities. Accountability mechanisms are contingent; they only come in to effect when a trust violation occurs. Furthermore, when initiated, these mechanisms are not mere objective features of the system but recognize the psychological impact of trust violation and largely follow accepted theory for repairing trust including immediate response, diagnosis, intervention performance, and evaluation [9].

Moreover, the framework includes actions that are effective for repairing violations of different types of trust, whether competence-, benevolence or integrity based. The framework is technology-agnostic and in this way, can not only accommodate technological solutions to building and repairing trust, but new use cases and evolutions of cloud computing including the Internet of Things.

By recognizing that policymakers and regulators, users and providers, have different priorities and perceptions of what trust means in the context of cloud computing, all stakeholders start on the basis of building trust rather than waiting for that trust to be violated. Ultimately, this should lead to greater understanding of the needs of different stakeholders, longer and deeper relationships, and innovation so that when a violation does occur, and it will, the relationship will be strong enough to survive.

Trust as a service

A single point for configuring and managing security of cloud services from various providers is made possible by cloud computing and the Cloud Trust as a Service (TaaS) [7]. The service's initial release includes an identity service that enables single sign-on across multiple cloud providers and a compliance profiling service that lets users compare the security profiles of various cloud providers against one another. The Cloud trust a as service is a tool with a focus on cloud trust management [8].

It could greatly simplify trust management for cloud users as a cloud-based tool. The cloud service assertions streamed in the cloud trust as a service were made by the cloud service providers themselves, so a cloud user must still make trust decisions about them. Most importantly, in their capacity as an intermediary, cloud users must assess the reliability of the cloud trust as a service.

What the basis of the trust relationship is between cloud users and those commercial trust brokers is the central question of any trust as a service mechanism. Trust your judgment when choosing a cloud broker and a cloud service provider.

Policy based trust

We previously noted the demand for "formal" trust mechanisms in cloud computing. In a related cloud, there is a fairly mature technology that uses "formal" trust mechanisms to support digital signatures, key certification and validation, as well as attribute certification

and validation. We apply trust ideas used to establish "formal" trust mechanisms to the cloud. [6]

First-time users, however, frequently have a difficult time believing the information provided by the service provider regarding the characteristics of a given service. A user might occasionally take a service provider's claims or guarantees seriously if they have a good reputation, a well-known brand, or a history of positive interactions. The aforementioned characteristics, in any case, are a crucial component of the watch list in cloud service monitoring and are used to determine whether the service provider acts in a trustworthy manner. Users will build or revise their trust in that service provider based on the verification's findings. [2]

In a whole, claims made by a cloud service provider regarding its own attributes must be verified before being used for decision-making. As such, cloud attribute assertions from independent, third-party professional organizations are expected, as we discuss in the following subsections: "Assessment of a cloud auditor or accreditor" and "Observation of cloud brokers."

THE IMPLICATIONS OF DATA SECURITY IN CLOUD COMPUTING

One of the IT industry's fastest-growing sectors, cloud computing has developed from a promising business virtualization concept. Right now, recession-hit businesses are realizing more and more that they can quickly access best-of-breed business applications or significantly increase their infrastructure resources by simply utilizing the cloud, both of which come at incredibly low cost.[5]

Nevertheless, a number of problems, difficulties, and implications have been noted by academics, researchers, and professionals in the field of Cloud computing and are being addressed.

Bandwidth Cost

Cloud computing gives businesses the chance to spend less on hardware and software, but they might pay more for network bandwidth. For smaller, less data-intensive Internet-based applications, bandwidth costs may be low, but they might increase significantly for data-intensive applications. [4]

Privacy

Different legal systems currently have divergent views on how to safeguard data privacy. However, users run the risk of disclosing sensitive data when they use cloud computing services. Attackers may use the computing task that users have submitted to analyze the important task. Contrary to the traditional computing model, cloud computing makes use of virtual computing technology, so user personal data may be spread across multiple virtual data centers as opposed to staying in one physical location on a hard drive, even across international borders. [9].

Security

Security concerns for businesses using cloud computing remain. Additionally, users are concerned about their vulnerability to attacks when crucial IT resources and information are located outside the firewall. Regardless of where the data repository is ultimately stored, the data in the cloud will be distributed over the network through individual computers. Ingenious hackers can break into almost any server, and statistics show that one-third of breaches are caused by lost or stolen laptops and other devices, by employees unintentionally posting data online, and by insider theft in nearly 16 percent of cases [8].

Performance

Data that is heavily transaction-oriented and other types of data can be the main performance problem. Intensive applications, where cloud computing may perform insufficiently. Users who are far from cloud providers may also encounter high latency and delay. Data that is heavily transaction-oriented and other types of data can be the main performance problem.

Intensive applications, where cloud computing may perform insufficiently. Users who are far from cloud providers may also encounter high latency and delay. [1]

Long-term Feasibility

Even if a specific cloud computing service provider goes out of business or is acquired and absorbed by a larger company, users can be confident that the data they store in the cloud will never expire. According to Gartner [4], the data would be returned to the cloud potential providers in any format that would allow for importation into a replacement application.

Legal Issues

Costs are lower, and the computer processing power or storage one buys via a cloud service may be based in another country, or in fact, may be divided between multiple countries, in the same way that the electricity one uses may have been produced in another country [9]. However, in addition to the cost and efficiency benefits brought about by this arrangement, this also raises complex legal issues for cloud computing arising from exporting customer data abroad; additionally, the cloud services The provider must deal with the legal systems of various jurisdictions while having limited visibility into the location of the data and the manner in which it is sent from one jurisdiction to another in order to reach the end user. Annoying legal issues once more.

Cloud computing offers us nearly limitless computing power, good scalability, on-demand services, etc., but it also presents challenges in terms of security, dependability, privacy, legal issues, and other factors. It has drawn everyone, including the attackers, as a result. For those who work in or conduct research in cloud computing, the paper is anticipated to be the best course of action. We acknowledge the cloud computing era, and the greatest amount of necessity is needed to address and prevent the current problems and their effects. [5]

Cloud Security and Trust Model

The adoption of effective security protocols has decreased, according [4], as a result of cloud migration. This is currently being reconsidered because of how the cloud differs from architectural design. Therefore, it is advised to use a trusted third party; this can be thought of as a protection system in the cloud that forms a network of confidence. This will ensure the secrecy and reliability of the data.

In addition to a straightforward inspection, numerous recent studies have produced a variety of for investigation techniques, particularly in the security and trust fields. Some people who use the Internet may not necessarily be experts in computer-related fields. One of the most obvious problems is that traditional businesses handle all sensitive information internally and have complete control over their workforce [9].

In this regard, it is argued that even though the cloud model is not entirely transparent, suppliers can exercise some openness to outline what is actually being done in a particular

region and to discuss the necessity of cloud providers proposing security rules as well as the various types of security issues [6].

Corporations are responsible for all of their sensitive data in terms of privacy, and organizations that lack knowledge of data storage and control are fundamentally weak [10]. Data protection and safety issues related to the cloud, which affect all stages of the information life cycle, including data production, transmission, usage, sharing, storage, archive, and elimination, are the main obstacle to the adoption of cloud services.

Recomedations

a) The regular, remotely stored backups that cloud services offer help an organization's disaster recovery plan.

b) Data security is an importance because you need to make sure it is stored in a nation with sound data protection laws and that the business has effective security protocols.

c) Although this is less likely given that these systems are largely platform independent, there may also be problems with compatibility with the devices used by an organization.

d) Cloud services typically reduce the load associated with maintenance by handling software updates and providing knowledgeable staff. However, there might occasionally be downtime as a result of software updates and attacks via the internet.

e) Most cloud services are very simple to set up, automating much of the configuration, and offering migration services for a faster setup.

f) The responsiveness of the user (such as lag when typing), the difficulty of the task being completed (such as graphical tasks), and the devices and communication technologies being used (such as a slow internet connection) can all have an impact on how well cloud services perform.

Conclusion

The issue of privacy protection is becoming an unavoidable roadblock in the age of cloud computing. This gave us encouragement, and we plotted the privacy study's major turning points while staying informed from various angles. With this work, people who are interested

in this promising and developing field will have a consistent route to follow. We summed up the conclusions of the privacy study in various research tenets and communities. Our presentation focused on the trust mechanism initiatives announced by the relevant privacy frameworks and models.

It is believed that this kind of privacy effort is crucial and highly demanded in problem-solving the cloud implications, and it is undeniably worthwhile to invest our energy and passion to look forward to the best strategy based on data utility and privacy accessibility

References

[1] Canard S, Devigne J (2016) highly privacy-protecting data sharing in a tree structure. Future Gener Comput Syst 62:119–127

[2] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, [2005] "Live migration of Virtual machines" In Proc. of NSDI'05, pages 273-286, Berkeley CA, USA, 2005. USENIX Association. [11] Eucalyptus Completes Amazon Web Services Specs with Latest Release

[3] Elinor Mills, January 27, 2009. "Cloud Computing Security Forecast: Clear Skies"

[4] Gartner. "Seven Cloud-Computing Security Risks" http://www.infoworld.com July 02, 2008.

[5] Sari, A. A review of anomaly detection systems in cloud networks and survey of cloud security measures in cloud storage applications. J. Inf. Secur. 2015,

[6] Habib S, Hauke S, Ries S, Muhlhauser M: Trust as a facilitator in cloud computing: a survey. J Cloud Comput Adv Syst Appl 2012.

[7] Jianchun Jiang, Weiping Wen, "Information Security Issues In Cloud Computing Environment", Netinfo Security, doi:10.3969/j.issn.1671-1122.2010.02.026.

[8] Massimo Ficco and Massimiliano Rak, 2015. Mousavi and Kahaki, 2014. Hybrid feature selection Stealthy Denial of Service Strategy in Cloud algorithm for intrusion detection system. Journal of Computing. IEEE Transactions on Cloud Computing,

[9] Kulothungan, K., S. Ganapathy, S. Indira Gandhi, P. Security, 11(3): 484-497. Yogesh and A. Kannan, 2011. Intelligent secured

[10]. Mansour, N., M. Chehab and A. Faour, 2010. Filtering fault tolerant routing in wireless sensor networks intrusion detection alarms. Cluster Computing, using clustering approach. International Journal of 13(1): 19-29. Soft Computing, 6(5): 210-215.

Analytical study of AI for Sustainable Human Development

[1]Madhvi Chandra,[2]Dr. Sukirti Priyadarshni

[1]Research Scholar,[2]School of Arts and Humanities, Career Point University, Kota, Rajasthan, India

[2]Associate Professor, Psychology,[3]School of Arts and Humanities, Career Point University, Kota, Rajasthan, India

Abstract:

This research paper is an analytical study of Artificial Intelligence (AI) within the framework of emotional well-being, especially its role in fostering sustainable human development. The study explores the connection between AI and psychological well-being, studying how AI applications contribute to mental health interventions, personalized therapies, and accessible mental health resources to those in remote areas. It also touches upon the ethical implications of using AI in psychological contexts, emphasizing the importance of ensuring privacy, trust, and cultural sensitivity in the development and use of AI-driven psychological tools. By blending insights from psychology and technology the research aims to highlight the potential of AI to enhance mental health support systems and contribute to sustainable and inclusive human development.

Keywords: Artificial Intelligence (AI), Emotional well being, Sustainable human development, Personalised therapies, Accessible mental health resources.

Introduction:

The application of artificial intelligence (AI) to many facets of human existence has created a wealth of opportunities, especially for sustainable human growth. This study examines the relationship between AI and sustainable human growth, focusing on people's psychological health and how it affects human development.

A. AI for Sustainable Human Development:

Artificial intelligence (AI) has developed over the past several years from a technical tool to a valuable tool with a variety of applications to help with challenging societal issues. The goal of sustainable human development is to enhance the well-being of people on an individual, community, and societal level while making sure that advancement is not only financially possible but also environmentally and socially sustainable.

AI is now being used as a tool to address some of the most important problems in sustainable development because of its capacity to handle enormous volumes of data and produce insightful conclusions. AI provides cutting-edge solutions that can improve accessibility, efficiency, and inclusion in a variety of fields, including healthcare and mental health.

Beyond computerization, artificial intelligence (AI) is being used in sustainable development to create sophisticated algorithms, machine learning, and data analytics to optimise resource allocation, enhance decision-making, and promote innovation. AI has the ability to significantly contribute to humanity by improving the future of mankind through these capacities.

B. The Role of Psychological Health in Human Development

One in every five individuals, according to statistics, in India suffers from some form of mental health issues and show symptoms. 50% of mental health conditions surface by age 14 and 75% of mental health conditions start by age 24.

The significance of psychological well-being cannot be emphasised enough, even though technological and economic advancement are essential to human growth. Human development includes improving a person's or a community's mental and emotional health in addition to their economic and technical advancements.

The three components of psychological well-being are general life satisfaction, emotional stability, and mental health. It is one of the most crucial requirements for a person to live a happy life, make a positive contribution to society, and adjust to changing conditions. A community's ability to endure and prosper depends on its ability to prioritise psychological well-being in the context of sustainable human development.

Studies demonstrate the connection between psychological wellness and other facets of human development, including prosperity in the economy, health, and education. Strong

mental health makes people more capable of learning, working, and making meaningful contributions to their communities, which promotes sustainable growth in the long run.

AI and sustainable human development together provide enormous potential, and mental health becomes a crucial aspect of sustainability. Understanding and treating the psychological components of human growth becomes crucial for comprehensive and long-lasting success as we investigate the integration of AI. The treatments and implications of AI in improving psychological well-being will be further examined in this study, which will ultimately lead to a more comprehensive understanding of sustainable human development.

Review of Literature

AI-Based Mental Health Interventions:

The use of AI in therapy has significantly advanced the field of mental health in recent years. AI has a significant impact on mental health interventions, particularly when it comes to tailored therapies, expanding access to mental health care in rural areas, and analysing scientific results from earlier research projects.

Empirical data supports the efficacy of mental health applications with respect to a number of endpoints, including cost-effectiveness (Massoudi et al., 2019), prevention (Sander et al., 2016), symptom reduction (Firth et al., 2017), adherence (Pihlaja et al., 2017), and the decrease in hospitalisation and overall hospital days (Bell et al., 2017).

A. Customised Treatment Plans:

A one-size-fits-all approach was frequently used by traditional therapies, which failed to address the particular needs of each patient. The development of individualised therapy catered to each patient's unique mental health profile is made possible by AI's capacity to analyse large amounts of data and identify trends.

These individualised therapies may take the form of sophisticated virtual therapists who may adjust their approaches in response to client input and advancement or AI-driven chatbots that provide real-time assistance. AI allows for a more individualised and efficient approach to mental health treatment by matching the type of therapy to each person's distinct traits, experiences, and preferences.

The British National Health Service has developed and implemented 'The Improving Access to Psychological Therapies' initiative for many years, which uses a layered model of care. Less impacted patients (e.g., those with mild-to-moderate depression, panic disorder, generalized anxiety disorder, or obsessive-compulsive disorder) receive self-applied internet treatments before being treated by individual or group therapy in person. Face to face therapy is offered when AI driven therapies fail or when the patients have a severe condition.

CBT delivered through a computer or mobile device (iCBT) has already demonstrated its efficacy in more than 100 randomized trials, even when compared with active face-to-face treatments.

Numerous studies have demonstrated the efficacy of tailored AI therapy in treating a range of mental health issues, such as PTSD, depression, and anxiety.

B. Mental Health Services Accessible in Remote Locations:

Limited availability is a significant issue in mental health care, especially in remote and underprivileged locations. AI's ability to provide virtual mental health services helps to overcome geographical limitations. AI-powered teletherapy platforms provide remote consultations, evaluations, and interventions, expanding access to mental health services for people who might otherwise encounter major barriers to getting mental health care. There would also be no time constraints on asking for assistance at strange times. When therapists are not easily accessible at night, most people feel most vulnerable.

The National Institute for Health and Clinical Excellence (NICE) in England first recommended computerized CBT packages for depression, panic, and phobias back in 2006 on the grounds of clinical and cost effectiveness.

Research has demonstrated the beneficial effects of AI-lead therapy in improving mental health outcomes for those living in rural areas. Virtual platforms' ease of use and adaptability encourage patients to remain engaged in their therapy, which improves their general well-being in areas where access to personal mental health treatments may be limited.

C. Scientific Results from Related Research:

Numerous academic studies back up the use of AI in mental health interventions. Research examining the efficacy of AI-driven treatments frequently point to beneficial results, such as decreased symptoms, better coping skills, and increased psychological well-being in general.

Facebook uses text analysis of the user and their friends posts and comments to look for concerning trends that indicate people are experiencing emotional discomfort. In order to identify people in distress, researchers from Facebook's FAIR project broaden their study to include facial expressions from photos and videos.

Using Artificial Intelligence to Enhance Ongoing Psychological Interventions for Emotional Problems in Real- or Close to Real-Time. Conversational AI tools, such as chatbots that can be accessed via a computer or smartphone, were found to be the most commonly used AI technique for emotional well-being in a 2022 study by Patricia Gual-Montolio, Irene Jaén, Verónica Martínez-Borba, Diana Castilla, and Carlos Suso-Ribera. The studies that were evaluated showed that utilising AI to improve psychotherapy and lessen clinical symptomatology has a lot of advantages. Furthermore, the majority of research found that using AI to improve psychotherapy in real-time or nearly real-time resulted in high rates of satisfaction, engagement, and retention.

IBM Research focuses on voice detection and analysis from psychiatric interviews using machine learning approaches. To assist physicians in effectively predicting and monitoring psychosis, schizophrenia, mania, and depression, their robots identify verbal speech patterns.

Additional research shows that AI can identify early indicators of mental health problems by analysing speech patterns, facial expressions, and other behavioural factors. AI tools that predict treatment outcomes based on patient characteristics also lead to more effective and individualised mental health care.

D. Studies and Research on AI-Powered Mental Health Therapies:

A number of case studies demonstrate the various ways AI is being used in mental health interventions. For example, Woebot is a psychologist-developed chatbot that uses natural language processing to converse with users while offering cognitive-behavioral treatment (CBT). The Woebot study illustrates the potential benefits of employing AI as a mental health therapy tool, pointing to potentially scalable approaches.

Replika, an AI chatbot created to mimic user interaction, is another example. Studies on Replika suggest that it can lessen feelings of isolation and offer emotional support. The same findings are also reached when "Alexa" and "Siri" are used extensively for communication. These case studies highlight the potential applications of AI in the field of mental health, ranging from chatbots to virtual therapists.

A Carnegie Mellon University and University of Pittsburgh study demonstrates how to recognise suicidal people by examining changes in the way specific death-related thoughts are represented in brain patterns. The model differentiated between those who have tried suicide with 94% accuracy.

Ethical Implications :

There are ethical issues with using AI in mental health interventions that need to be carefully thought out. The research examines two significant ethical stances. 1) Privacy and trust issues 2) An appreciation for cultural sensitivity.

A. Confidentiality and faith Issues with AI-Led interventions for mental health:

AI in mental health frequently requires gathering and analysing extremely private personal data. Privacy is the primary worry as AI uses data processing to provide individualised treatment. Concerns are raised about data storage security, possible security breaches, and unintentional publication of mental health records.

Establishing robust privacy measures is imperative in fostering confidence in AI-powered mental health solutions. Anonymization methods, secure data transfer, and encryption are crucial instruments for maintaining privacy.

B. Considering Cultural Sensitivity When Using AI to Improve Mental Health:

AI-driven mental health interventions must be implemented with cultural sensitivity due to the diversity of nations and customs. Social origins have an impact on how people see mental health, the stigma associated with specific illnesses, and how individuals respond to treatment. Neglecting these elements may result in skewed algorithms and ineffective therapies.

To make sure that AI systems are considerate of the different needs and viewpoints of users worldwide, developers must carry out extensive cross-cultural study. The acceptability and effectiveness of AI are increased across many cultural societies when treatments are adapted to conform to cultural norms, values, and preferences. This would promote inclusivity in mental health care and lessen the possibility of unintentionally reinforcing biases.

Encouraging responsible AI applications that can truly improve mental health treatment while respecting people's individuality and dignity requires prioritising privacy and being aware of cultural sensitivity.

Research Methodology:

Descriptive Research Methodology was used for this research project.

Research Findings:

The integration of AI and mental health has been examined in this study paper, with a focus on individualised treatments, accessibility in remote locations, ethical issues, and the consequences for long-term human development. The study of the research demonstrated the benefits of AI-based interventions, particularly in enhancing psychological well-being via tailored therapy and expanding access to mental health resources.

Prominent case studies demonstrated the adaptability of artificial intelligence applications; Woebot and Replika, for example, are chatbots that demonstrate the positive impact of AI on mental health. Previous studies have continuously affirmed the efficacy of AI-driven therapies, highlighting its capacity to extend therapies and elevate psychological well-being.

The ethical application of AI in mental health requires careful consideration of several factors, including cultural sensitivity, privacy concerns, and the reliability of AI systems. In order to secure the ethical use of AI technology in mental health care, the study brought attention to the necessity for transparent processes, user permission and feedback channels, and cultural inclusion.

A. AI Therapies and Their Effect on Mental Health:

Studies examining the effects of AI-based therapies on psychological well-being have produced scientific evidence of their beneficial effects. When using AI-driven mental health solutions, users report feeling happier, having less anxiety and sadness, and having more self-awareness.

Research also show that AI can improve on established therapy practices, helping mental health providers deliver more focused and successful interventions. AI's ability to improve psychological well-being especially in partnership with human therapists.

B. Takeaways and Opportunities for Development:

Scientific research and individual case studies show proof of the beneficial effects of AI-led mental health therapies on emotional wellbeing. Effective, moral, and approachable AI-led therapies for enhancing psychological well-being will require incorporating user feedback and upgrading AI tools as their use grows.

C. Overall Consequences for Sustainable Human Development:

The use of AI for mental health has important ramifications for the long-term future of humankind. AI-powered personalised therapies have the potential to completely transform mental health services by offering specialised interventions that improve mental health in general.

AI-driven therapy systems greatly increase the accessibility of mental health care in rural places. This improved accessibility supports equity and inclusivity in mental health, which is in line with the ideals of sustainable human development.

The Sustainable Development Goals (SDGs) of the UN align with the enormous influence of AI on mental health. Better mental health works with two other goals, goal 10 (reduced inequality) and goal 4 (quality education), and it also adds to goal 3 (good health and well-being). The advantages of implementing AI-led mental health treatment align with the goals of building an egalitarian, inclusive, and sustainable society.

Conclusion:

In conclusion, the application of AI to mental health interventions has the potential to significantly impact human development in the long run. The benefits, moral dilemmas, and ramifications of applying AI to improve mental health have been highlighted by this research study. In order to build communities where technology contributes to a more sustainable and fair society, ethical development and implementation will be crucial as we advance in the rapidly expanding field of Artificial Intelligence.

References:

1. Bakker D., Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. J. Affect. Disord. 2018;227:432–442. [PubMed] [Google Scholar]

2. Bartels S.L., van Knippenberg R.J.M., Dassen F.C.M., et al. A narrative synthesis systematic review of digital self-monitoring interventions for middle-aged and older adults. Internet Interv. 2019;18 [PMC free article] [PubMed] [Google Scholar]

3. Abbe A., Grouin C., Zweigenbaum P., et al. Text mining applications in psychiatry: a systematic literature review. Int. J. Methods Psychiatr. Res. 2016;25(2):86–100. [PMC free article] [PubMed] [Google Scholar]

4. Carr S. 'AI gone mental': engagement and ethics in data-driven technology for mental health. J. Ment. Health. 2020;29(2):125–130. [PubMed] [Google Scholar]

5. de Laat P.B. The disciplinary power of predictive algorithms: a foucauldian perspective. Ethics Inf. Technol. 2019;21:319–329. [Google Scholar]

6. Dogan E., Sander S., Wagner X. Smartphone-based monitoring of objective and subjective data in affective disorders: where are we and where are we going? Systematic review. J. Med. Internet Res. 2017;19(7) doi: 10.2196/jmir.7006. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

7. iHealth: The ethics of artificial intelligence and big data in mental healthcare Giovanni Rubeis

8. Eliane M. Boucher, Nicole R. Harake, Haley E. Ward, Sarah Elizabeth Stoeckl, Junielly Vargas, Jared Minkel, Acacia C. Parks & Ran Zilca (2021) Artificially intelligent chatbots in digital mental health interventions: a review,

Expert Review of Medical Devices, 18:sup1, 37-49, DOI: 10.1080/17434440.2021.2013200

9. Using Artificial Intelligence to Enhance Ongoing Psychological Interventions for Emotional Problems in Real- or Close to Real-Time: A Systematic Reviewby Patricia Gual-Montolio rene JaénVerónica Martínez-Borba 1,2,*Diana Castilla 3,4 and Carlos Suso-Ribera 1 Int. J. Environ. Res. Public Health 2022, 19(13), 7737; https://doi.org/10.3390/ijerph19137737

10. Richards, D.; Enrique, A.; Eilert, N.; Franklin, M.; Palacios, J.; Duffy, D.; Earley, C.; Chapman, J.; Jell, G.; Sollesse, S.; et al. A Pragmatic Randomized Waitlist-Controlled Effectiveness and Cost-Effectiveness Trial of Digital Interventions for Depression and Anxiety. NPJ Digit. Med. 2020, 3, 85. [Google Scholar] [CrossRef]

11. Kazdin, A.E. Treatment as Usual and Routine Care in Research and Clinical Practice. Clin. Psychol. Rev. 2015, 42, 168–178. [Google Scholar] [CrossRef]

12. Bach Xuan, T.; Giang Thu, V.; Giang Hai, H.; Quan-Hoang, V.; Manh-Tung, H.; Thu-Trang, V.; Viet-Phuong, L.; Manh-Toan, H.; Nghiem, K.-C.P.; Huong Lan Thi, N.; et al. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. J. Clin. Med. 2019, 8, 360. [Google Scholar] [CrossRef] [Green Version

Floristic studies pertaining to the biodiversity of the Mukundara Hills Tiger Reserve (MHTR) Kota of Rajasthan state

Nidhi Sharma

Research Scholar, School of Basic & Applied Sciences, Career Point University, Kota, Rajasthan, India

Email- nidhisharma26989@gmail.com

Abstract:

This research paper aims to explore the floristic diversity within the Mukundara Hills Tiger Reserve and its significance in the context of biodiversity conservation. The distinguishing feature of this area is its copious vegetation, which encompasses thick woodlands and a substantial quantity of trees and shrubs. The study utilises a comprehensive approach, combining field surveys, data analysis, and a literature review to assess the floral composition, distribution patterns, and ecological importance of the reserve. The study determine the present condition of plant life in the reserve, including the diversity of species, genetic variations within each group, and the communities, ecosystems, and landscapes that the reserve encompasses. The studies also focus on the medicinal and endemic plant species present in the reserve, providing information on their uses and overall floral diversity. This research highlights the significance of MHTR and advocate for its sustainable management, promoting both ecological and economic well-being. This research paper aims to bridge this knowledge gap by conducting a comprehensive floristic study within the reserve. These floristic studies enhance comprehension of the biodiversity in the MHTR Kota and can provide valuable insights for conservation and management strategies in the region.

Keywords: MHTR, Floristic, Ecosystem, Biodiversity, Medicinal plants, Conservation

Introduction:

The Mukundara Hills Tiger Reserve (MHTR), located in the state of Rajasthan, India, is renowned for its rich biodiversity and unique ecological features (ShopManager, 2023). The location is in a valley created by two parallel mountains, namely Mukundra and Gargola. The valley is demarcated by four rivers, namely Ramzan, Ahu, Kali, and Chambal (Nayak, 2016).

It is situated on the eastern side of the Chambal River and is encompassed by its tributaries. The Darrah Range and Darrah Wildlife Sanctuary are alternative names for Mukundra Hills (Sultana, 2009). The name Darrah signifies a mountain pass in the indigenous dialect. In 1955, Darrah was formally designated as a wildlife sanctuary (Sharma, 2023). The declaration of Mukundra Hills National Park took place in 2004. The establishment of the Tiger Reserve took place in 2013 (ShopManager, 2023). Established in 2013, the MHTR spans over 759 square kilometers, encompassing diverse landscapes ranging from dry deciduous forests to riverine ecosystems (Singhadiya & Pandey, 2014). The MHTR Kota is known for its rich floristic diversity with a total of 700 species of floristic plants documented in the study area (Rajawat, 2021; Sultana, 2009). MHTR are vital reservoirs of plant diversity, including medicinal and wild edible plants (Khandelwal et al., 2023). The local climate is conducive to the flourishing of arid, deciduous forests (Sharma, 2023). A wide variety of visually appealing blooming indigenous trees, shrubs, and moderate-sized trees have been found in the MHTR (Wagensommer, 2023). The study emphasizes the urgent need for conservation efforts for endangered species, particularly in the MHTR Kota, which houses a large number of rare and endangered plants (Soni, 2023; Nayak, 2016).

The conservation of ecosystems and biodiversity is a national and international concern, requiring effective management strategies that consider species distribution patterns and the interconnectedness of different ecosystems (Malav & Jaiswal, 2023; Sharma et al., 2013). Collaboration at both national and international levels is crucial for preserving biodiversity (Prasad et al., 2023). Recognizing the significance of biodiversity for traditional medicine practices, promoting sustainable use of natural resources, and preserving cultural heritage is also vital for the long-term sustainability and resilience of human societies (Solanki & Kotiya, 2021). The area now known as MHTR has a rich cultural and historical past. Ancient temples like Garadia Mahadev stand as testaments to its significance. Geographically, the reserve boasts unique features like the Mukundara Hills, offering breathtaking vistas and diverse ecological niches (Sharma et al., 2013). The Chambal River, a lifeline for the ecosystem, adds further value to the landscape. The MHTR Kota is characterised by abundant tropical dry deciduous forests that are home to a diverse range of economically valuable species and generate significant revenue from timber (Rajawat, 2021; Srivastava, 2017). The conservation status of the rare and endangered plant species was assessed, emphasising the

need for immediate attention and conservation efforts. Floral and faunal diversities are essential elements of biodiversity, representing the extent and variation of species (Jadhav, 2017). The Indian government has already implemented the Biodiversity Acts of 2002 and the national Environment policy of 2005 (Sultana, 2009). In order to successfully implement the aforementioned acts, it is imperative to possess a thorough and up-to-date inventory of the flora and fauna in the area, with a specific focus on rare and endangered species (Singhadiya & Pandey, 2014). For example, the botanical surveys conducted in the MHTR revealed the presence of a highly endangered medicinal plant species that is endemic to the region (ShopManager, 2023). This finding highlights the importance of preserving this reserve and implementing conservation measures to protect this valuable plant species from extinction (Haq et al., 2021).

However, the loss of biodiversity threatens the availability of traditional medicines and disrupts ecosystem balance (Khandelwal et al., 2023). Indigenous communities, who have relied on these forests for generations, are at risk of losing their cultural practices and knowledge of medicinal plants (Prasad et al., 2023; Nayak, 2016). The study explores the potential of traditional medicinal systems in the healthcare system, highlighting the need to integrate indigenous knowledge of plants and ethnomedicinal values with biotechnological approaches and pharmacological activities to develop new sources of herbal drugs (Soni, 2023; Solanki & Kotiya, 2021). Prioritizing the conservation of tropical forests is essential to ensure the continuity of traditional medicine practices and preserve the cultural heritage of indigenous communities (Wagensommer, 2023).

Conceptual Framework:

Considering this review paper the conceptual framework section may encompass the following essential elements-

- Species composition and distribution: Map the occurrence and abundance of plant species across the reserve, providing insights into habitat and conservation needs.
- Community structure and dynamics: Different plant communities interact and respond to environmental changes is crucial for effective management.

▪ Endemic and threatened species: Identification of rare and endangered plants helps prioritise conservation efforts.

Review of Literature:

The MHTR Kota is situated in the Haroti region of southeastern Rajasthan is a continuation of the Vindhyan ranges. The Haroti plateau is located on the periphery of the Malwa plateau, specifically at latitudes 23°45' to 25°53' N and longitudes 75°9' to 77°26' E. The reserve is located on the perimeters of four cities in Rajasthan, specifically Kota, Bundi, Chittorgarh, and Jhalawar (Sharma, 2023). The area's richness in flora can be attributed to its hilly terrain, elevated altitude, distinctive topography, and dense forest cover (Malav & Jaiswal, 2023). The floristic study focuses on the floral biodiversity of the MHTR in Kota, Rajasthan, Emphasising the diverse range of species, genetic variations within each group, and the various communities, ecosystems, and landscapes that the reserve encompasses (Kumar et al., 2013; Jadhav, 2017). The MHTR harbors a rich tapestry of plant life, and floristic studies play a crucial role in understanding and conserving this biodiversity (Sultana, 2009). The taxonomic richness observed in MHTR underscores the importance of preserving this unique ecosystem to maintain the genetic diversity within each taxon (Praveen, 2021). Additionally, the study highlights the role of these plant communities in supporting various ecosystems and landscapes within the reserve, further emphasising the need for conservation efforts to protect this biodiversity hotspot (Rajendran et al., 2014).

Climate change significantly impacts the ecosystem, with diverse vegetation types such as dense forests, open forests, scrublands, and grasslands containing unique plant communities (Datar, 2016). These communities not only provide habitat for various animal species but also regulate local climate and water cycles. Dense forests act as carbon sinks, mitigating climate change effects, while open forests and grasslands support grazing animals and provide food sources for local communities (Shrivastava & Singh, 2009). Preserving these diverse vegetation types is crucial for maintaining the ecosystem balance and ensuring the sustainability of traditional healing practices relying on medicinal plants (Khandelwal et al., 2023). For example, the Amazon rainforest in South America is an example of a plant community that supports a wide array of animal species while also regulating global climate (Kumar et al., 2013; Datar, 2016). The dense vegetation absorbs large amounts of carbon

dioxide, mitigating climate change effects. The rainforest also provides essential resources for local communities, such as food, medicine, and shelter materials (Praveen, 2021).
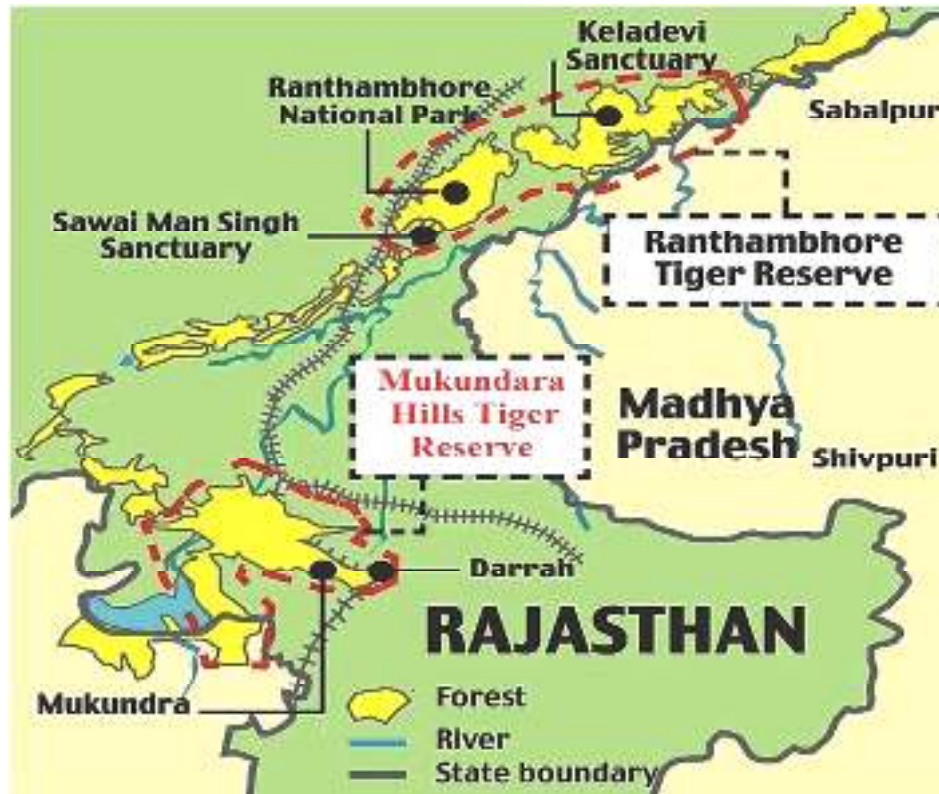
Figure 1. Mukundara Hills Tiger Reserve (MHTR) Kota of Rajasthan state

Therefore, it is essential to raise awareness and invest in research and conservation efforts to ensure the long-term survival and sustainable use of deciduous forests (Bagchi et al., 2003). By doing so, we can protect their unique biodiversity and ecosystem services while promoting a more balanced approach to forest management worldwide (Rajendran et al., 2014). For example, in the Appalachian Mountains of the United States, the limited understanding of deciduous forest dynamics has resulted in ineffective management practices (Shrivastava & Singh, 2009). This has led to overharvesting of timber resources, loss of habitat for endangered species like the Indiana bat, and increased vulnerability to invasive species (Malav & Jaiswal, 2023). By raising awareness and investing in research and conservation efforts in this region, we can develop more sustainable practices that prioritize both economic benefits and ecological preservation, ensuring the long-term survival of these vital forests (Srivastava, 2017).

However, despite its taxonomic richness and importance for local communities, the Mukundara Hills National Park faces significant threats from human activities such as illegal logging and poaching. These activities not only endanger the flora within the park but also disrupt the delicate balance of ecosystems and landscapes it supports, making conservation efforts even more crucial (Jadhav, 2017). These forests are more susceptible to fire during the dry season, leading to faster exploitation and conversion. Effective fire prevention and control measures are crucial in preventing this exploitation. The conversion of these forests can have significant consequences for biodiversity and carbon storage, as they often contain unique species and contribute to climate regulation (Shrivastava & Singh, 2009). Therefore, prioritizing conservation and sustainable management of these vulnerable forests is crucial to mitigate the negative impacts of their exploitation.

Research Gap Identified:

(i) Unique species found in deciduous forests: Highlight the significance of these forests by discussing the presence of unique and specialised taxonomic of flora in the MHTR Kota. It analyzes tree species composition and diversity in protected and non-protected vegetation stands within the reserve, focusing on significant species and families.

(ii) The importance of conserving deciduous forests: The study emphasizes the importance of conserving and managing vulnerable forests to protect biodiversity and maintain carbon storage.

(iii) Contribution to climate regulation: Explain how deciduous forests play a role in climate regulation, including their ability to sequester carbon dioxide, mitigate greenhouse gas emissions, and regulate local temperatures.

(iv) Comparing attention received with tropical rainforests: Discuss the disparities between scientific and political attention given to deciduous forests compared to tropical rainforests, despite their ecological and economic importance.

Research Methodology:

The MHTR is renowned for its rich biodiversity, making it a vital floristic region in Rajasthan. The flora in the Mukundara Hills is classified as tropical dry deciduous (Bagchi et

al., 2003). The research methodology involves a combination of field surveys, specimen collection, and data analysis. We will conduct extensive botanical surveys across different habitats within the reserve to document the plant species present (Kumar et al., 2020). We will collect, identify, and preserve specimens for further analysis (Raole, 2022). In addition, we will review existing literature and herbarium records to supplement the field data. The paper specifically review 70 plant species that are uncommon or at risk of extinction within the sanctuary area (Prakash et al., 2019). A quantitative evaluation of the diversity of plant species was conducted using phytosociological data. The study area contained a collection of uncommon or threatened plant species (Vijayvargiya & Shrivastava, 2023; Praveen, 2021).

During the field survey, we measured various criteria outlined by the International Union for Conservation of Nature (IUCN) to categories threatened plants. The criteria encompass the range of occurrence, size of habitat, population size, and likelihood of extinction (Meena, 2023). The assessment of species rarity was conducted through field study, visual estimations, and literature review. During the collection process, limited quantities of rare and vulnerable species were identified within the study area (Kumar et al., 2020). These species are included in the Red Data Book of Indian plants, the IUCN list of endangered species, and the list of BSI arid zone circle (Raole, 2022). This study additionally assessed species diversity and biological richness at the landscape scale by employing satellite remote sensing and geographic information system (Sharma et al., 2013). The research is grounded on field observations and collection excursions conducted between July 2022 and January 2024. This floristic study was conducted using a systematic sampling approach. We will conduct data collection using these three methods: Thoroughly document plant species by gathering voucher specimens for accurate identification and long-term preservation (Meena, 2023). Collect important data by Record the GPS coordinates, altitude, habitat type, phenological stage, abundance, and any other pertinent observations. Photograph plant species to obtain high-quality images for the purpose of documentation and subsequent analysis (Prakash et al., 2019).

A comprehensive survey was carried out in various seasons to evaluate the abundance, physical characteristics, flowering and fruiting periods, and other relevant factors of the numerous large shrubs or medium-sized trees in the area (Soni, 2023). A grid-based sampling

design was employed, with a total of 100 sampling plots distributed evenly across the study area. The size of each plot was 10m x 10m. In each plot, all floristic plant species were identified and recorded. Voucher specimens were collected for further taxonomic verification (Peng et al., 2017). The conservation status of rare and endangered plant species was assessed using the IUCN Red List categories. Data analysis included calculating species richness, diversity indices (e.g., Shannon-Wiener index), and conducting statistical tests (e.g., t-test, ANOVA) to compare the species composition and diversity between protected and non-protected vegetation stands (Upadhaya, 2015; Vijayvargiya & Shrivastava, 2023).

The study will contribute to broader knowledge of ecosystem functioning and the importance of preserving biodiversity for the well-being of both the natural world and human societies (Meena, 2023). Understanding the diversity of nature is crucial for ecosystem function and stability, and this study will help researchers gain a comprehensive understanding of the plant community and its ecological role. Prioritizing conservation efforts and implementing protective measures is vital for preserving biodiversity (Haq et al., 2021).

Data Analysis & Interpretation:

The research will identify and document the plant species present, their distribution patterns, and ecological preferences. The findings will shed light on the diversity and abundance of plant species, including endemic and rare plants, within the reserve. The region is characterised by its rugged terrain and abundant vegetation, including a diverse array of herbs, shrubs, lianas, climbers, and trees. The area features grassland interspersed with numerous dry deciduous trees (Bagchi et al., 2003). Studies reveal a diverse flora with over 700 species documented, belonging to various families like *Poaceae* (grasses), *Fabaceae* (legumes), *Asteraceae* (sunflowers), *Acanthaceae*, and *Cyperaceae* (sedges). Among these, 70 species were identified as rare or endangered, highlighting the need for their conservation. The species composition and diversity varied between protected and non-protected vegetation stands, with a higher diversity observed in the protected areas. This diverse flora provides valuable resources for local communities, including medicinal plants used for traditional healing practices (Upadhaya, 2015).

Furthermore, floristic studies also play a crucial role in identifying potential medicinal plants and their therapeutic properties. This knowledge can then be utilized to develop new herbal drugs that can be incorporated into mainstream healthcare systems. Integrating indigenous knowledge with scientific research can help rationalise the use of natural products in the healthcare system, leading to the evolution of new sources of herbal drugs in different industries (Nayak et al., 2018). In this way, the integration of indigenous knowledge and scientific research not only enhances our understanding of the natural world but also paves the way for innovative advancements in medicine and conservation (Khandelwal et al., 2023) (Raole, 2022). MHTR Kota is a prime example of sustainable forest management through ecological surveys. Researchers have identified threatened species like tigers, leopards, and sloth bears, mapping their movement patterns and understanding ecological relationships (Rajawat, 2021). This information guides conservation efforts and establishes buffer zones and corridors to connect fragmented habitats, ensuring the long-term survival of these species (Giordano et al., 2021).

Research Findings:

Floristic studies provide the foundation for biodiversity conservation strategies in the MHTR Kota. The study identified 700 plant species from 70 families that the local communities in MHTR Kota. The authors have recorded the botanical species of flowering plants in the MHTR of Rajasthan through on-site investigations and reference to existing literature. The study identified 70 plant species in the sanctuary area that are rare or endangered, emphasising the need for their conservation and immediate attention. The presence of these rare and endangered plants in the sanctuary highlights the importance of the sanctuary as a conservation model for plant diversity (Nayak et al., 2018). The forests of Mukundra Hills are home to several common tree species, including

- *Acacia nilotica,*
- *Anogeissus pendula,*
- *Aegle marmelos,*
- *Acacia catechu,*
- *Azadirachta indica,*
- *Bombax ceiba,*
- *Butea monosperma,*

- *Cassia fistula,*
- *Dalbergia sissoo,*
- *Dichrostachys cinerea,*
- *Ficus religiosa,*
- *Mitragynaparvifolia,*
- *Vitex negundo*

They inform habitat management practices, restoration efforts, and the control of invasive species. Integrating indigenous knowledge with scientific research can help preserve traditional practices and inform sustainable healthcare practices (Prasad et al., 2023). By understanding the linkages between plants and other organisms, they contribute to broader conservation goals for the entire ecosystem (Giordano et al., 2021). The high plant species diversity documented in the MHTR indicates the presence of a healthy and functioning ecosystem. However, the results also revealed the presence of several endemic and medicinal plant species, indicating the importance of the reserve for biodiversity conservation and traditional medicine (Wagensommer, 2023). The existence of uncommon and imperiled plant species underscores their susceptibility to hazards such as the reduction and division of their natural habitats.

Conclusions:

This research paper will contribute to the existing knowledge on the floristic diversity within the Mukundara Hills Tiger Reserve. The MHTR Kota is recognised as one of the richest floristic regions in the state of Rajasthan, with a diverse range of floristic plants documented in the study area. The presence of rare and endangered plant species highlights the urgent need for conservation efforts and the importance of the reserve as a conservation model. The findings of this study can inform conservation and management strategies for the reserve, including the development of land use plans, afforestation, reforestation, and forest rehabilitation. The area is known for its rich biodiversity, with numerous species of mammals, birds, reptiles, and amphibians calling it home. The hills are inhabited by various avian species, including the Indian peafowl, grey jungle fowl, Indian roller, and Indian vulture. Overall, the diverse vegetation and abundant wildlife make this area a haven for nature lovers and wildlife enthusiasts. Additionally, the economic value of these forests, such as timber production and recreational activities, often takes precedence over their

conservation. The study's findings underscore the importance of conserving the MHTR Kota due to its substantial population of endangered and rare plants. It will contribute to forest conservation, ecological research, and the development of land use plans for afforestation, reforestation, and forest rehabilitation.

MHTR faces multiple challenges, including poaching, habitat fragmentation due to encroaching human settlements, and resource overexploitation. Additionally, a lack of adequate funding and infrastructure hinders conservation efforts. Addressing these challenges requires community engagement, stricter enforcement, and sustainable development initiatives. Despite their ecological and economic importance, deciduous forests have received less scientific and political attention compared to tropical rainforests. This lack of focus on deciduous forests has led to a limited understanding of their specific ecological dynamics and the potential threats they face. While facing challenges, its potential for biodiversity conservation and sustainable development remains bright. Recognizing its value and supporting ongoing efforts are crucial for ensuring the reserve's long-term success. By highlighting the importance of plant species in supporting biodiversity, the study will emphasise the need for their conservation and sustainable management. Through collaborative efforts, MHTR can flourish as a haven for biodiversity and a beacon of responsible conservation in Rajasthan. The findings will serve as a valuable resource for policymakers, conservationists, and researchers working towards the preservation of this unique ecosystem.

Suggestions & Recommendations/Future Scope:

- Delving deeper into specific floristic studies and examining their findings and methodologies in relation to plant conservation. Discussing the challenges faced in conserving flora within the MHTR and ongoing efforts to protect these valuable plant species.

- Further research is recommended to explore the potential of medicinal plants for the development of herbal drugs and to integrate traditional knowledge with scientific research for sustainable healthcare practices.

- Identifying potential areas for future research in this field, including new avenues for studying medicinal plants and advancing our understanding of their benefits and applications.

- Ecotourism can provide much-needed financial support for conservation while educating visitors about the reserve's significance. Sustainable tourism models that involve local communities and promote responsible practices can play a vital role in the reserve's future.

Declaration:

The authors affirm that they did not receive any financial aid, grants, or any other types of assistance during the preparation of this research article. The ethical criteria have been fulfilled.

References:

1. Bagchi, S., Goyal, S. P., & Sankar, K. (2003). Prey abundance and prey selection by tigers (Panthera tigris) in a semi-arid, dry deciduous forest in western India. *Journal of Zoology*, *260*(3), 285–290. https://doi.org/10.1017/s0952836903003765

2. Datar, M. N. (2016). Floristic diversity and effect of anthropogenic activities on human-dominated grasslands in subtropical regions of Peninsular India. *Tropical Grasslands - Forrajes Tropicales*, *4*(1), 8. https://doi.org/10.17138/tgft(4)8-18

3. Giordano, M., Troia, A., & Ilardi, V. (2021). Floristic survey of the former royal hunting reserve of Renda, near Palermo (Sicily, Italy). *Biodiversity Journal*, *12*(2), 403–433. https://doi.org/10.31396/biodiv.jour.2021.12.2.403.433

4. Haq, S. M., Singh, B., Bashir, F., Farooq, A. J., Singh, B., & Calixto, E. S. (2021). Exploring and understanding the floristic richness, life-form, leaf-size spectra and phenology of plants in protected forests: A case study of Dachigam National Park in Himalaya, Asia. *Acta Ecologica Sinica*, *41*(5), 479–490. https://doi.org/10.1016/j.chnaes.2021.07.010

5. Jadhav, D. (2017). Floristic Diversity of Angiospermic Weeds in Madhya Pradesh : A Review. *Journal of Non Timber Forest Products*, *24*(1), 49–54. https://doi.org/10.54207/bsmps2000-2017-n852sk

6. Khandelwal, D. M., Sharma, D. P., Choudhary, D. A., & Kharoliwal, D. S. (2023). Exploration of Ethno- medicinal plants of Kota, Rajasthan and their practices. *International Journal of Biology Sciences*, *5*(1), 156–160. https://doi.org/10.33545/26649926.2023.v5.i1c.159

7. Kumar, D., Chauhan, P., Pandey, S., Bhatnagar, P., Sharma, M., & Nagar, B. (2020). Phytodiversity Characterization of Mukundara Hills Tiger Reserve. *International Journal of Current Microbiology and Applied Sciences*, *9*(7), 2037–2047. https://doi.org/10.20546/ijcmas.2020.907.235

8. Kumar, N., Kumar, A., & Singh, M. (2013). Floristic Diversity Assessment in Ecologically Restored Limestone (Building Stone) Mine Near Chechat Village, Kota District, Rajasthan. *Ecologia*, *4*(1), 16–25. https://doi.org/10.3923/ecologia.2014.16.25

9. Malav, A., & Jaiswal, P. (2023). Species Composition and Diversity of Tree Species in Nanta Forest Region in Kota District, Rajasthan, India. *International Journal of Environment and Climate Change*, *13*(4), 220–227. https://doi.org/10.9734/ijecc/2023/v13i41729

10. Meena, M. (2023). A Review on Medicinal Plants of Sariska Tiger Reserve of Alwar district, Rajasthan". *International Journal of Advanced Scientific Research and Management*, *8*(10), 13–16. https://doi.org/10.36282/ijasrm/8.10.2023.1906

11. Nayak, P. (2016). Study on Soil Parameters of Selected Sites in Mukundara Hills National Park, Kota, Rajasthan. *International Journal of Pure & Applied Bioscience*, *4*(4), 316–320. https://doi.org/10.18782/2320-7051.2343

12. Nayak, S., Shah, S., & Borah, J. (2018). Record of Cinereous Vulture *Aegypius monachus* from Mukundara Hills Tiger Reserve, Rajasthan, India. *Vulture News*, *71*(1), 50. https://doi.org/10.4314/vulnew.v71i1.4

13. P. (2021). *Praveen, Author at Rajasthan Biodiversity Network*. Rajasthan Biodiversity Network. Retrieved January 24, 2024, from https://rajasthanbiodiversity.org/author/rbpraveen/

14. Peng, D., Lu, L., & Chen, Z. (2017). Regional tree of life and its application in floristic studies. *Biodiversity Science*, *25*(2), 36–40. https://doi.org/10.17520/biods.2015336

15. Prakash, L., Balasubramanian, P., & Anbarasu, C. (2019). Floristic Structure of Sathyamangalam Tiger Reserve with special Reference to Endemic species. *Indian Journal of Forestry*, *42*(4), 351–355. https://doi.org/10.54207/bsmps1000-2019-ouo7nz

16. Prasad, R., Kumar Sharma, D., & K Rathore, D. (2023). A Study on Angiospermic Diversity of Vardhman Mahaveer Open University Campus, Kota (Rajasthan). *The Journal of Plant Science Research*, *38*(2), 835–845. https://doi.org/10.32381/jpsr.2022.38.02.38

17. Rajawat, R. S. (2021). Sustainable Nature Tourism in Mukandra Hills Tiger Reserve India. *International Journal of Advanced Research in Science, Communication and Technology*, 40–44. https://doi.org/10.48175/ijarsct-982

18. Rajendran, A., Aravindhan, V., & Sarvalingam, A. (2014). Biodiversity of the Bharathiar university campus, India: A floristic approach. *International Journal of Biodiversity and Conservation*, *6*(4), 308–319. https://doi.org/10.5897/ijbc2014.0679

19. Raole, V. M. (2022). Floristic Diversity and Conservation Status of Non-Reserve Forest Area of Kachchh District, Gujarat. *JOJ Wildlife & Biodiversity*, *4*(4). https://doi.org/10.19080/jojwb.2022.04.555642

20. Sharma, B. K., Kulshreshtha, S., Sharma, S. K., Lodha, R. M., Singh, S., Singh, M., & Sharma, N. (2013). Physiography and Biological Diversity of Rajasthan. *Faunal Heritage of Rajasthan, India*, 39–166. https://doi.org/10.1007/978-1-4614-0800-0_2

21. Sharma, O. P. (2023). Comparative Floristic Analysis of Angiosperms of Bundi District–A Part of Haroti Plateau (Rajasthan). *The Journal of Plant Science Research*, *39*(2), 123–131. https://doi.org/10.32381/jpsr.2023.39.02.13

22. ShopManager, G. (2023). *Mukundara Hills National Park*. Greenverz. https://greenverz.com/mukundara-hills-national-park/

23. Shrivastava, A., & Singh, V. (2009). Additional floral elements to the Ranthambhore Tiger Reserve, Rajasthan, India. *Journal of Threatened Taxa*, *1*(9), 475–480. https://doi.org/10.11609/jott.o2037.475-80

24. Singhadiya, M., & Pandey, R. (2014). Floristic Analysis of Cultivated Plants of Rajasthan, India. *Indian Journal of Forestry*, *37*(2), 213–218. https://doi.org/10.54207/bsmps1000-2014-x9qwda

25. Solanki, Y., & Kotiya, A. (2021). Floristic diversity of Umari Dham sacred grove in Jaipur, Rajasthan, India. *The Holistic Approach to Environment*, *11*(4), 109–121. https://doi.org/10.33765/thate.11.4.2

26. Soni, D. (2023). Ethnomedicinal Study of Plants to Cure Skin Diseases in Mukundara Hills National Park, Kota, Rajasthan. *Indian Journal of Pure & Applied Biosciences*, *11*(1), 43–49. https://doi.org/10.18782/2582-2845.8978

27. Srivastava, S. (2017). Capacity Building through Nature Tourism at Southeast Rajasthan. *IRA-International Journal of Management & Social Sciences (ISSN 2455-2267)*, *6*(1), 37. https://doi.org/10.21013/jmss.v6.n1.p6

28. Sultana, F. (2009). Human Impacts on the Biodiversity of the Darrah Wildlife Sanctuary in Rajasthan. *The International Journal of Climate Change: Impacts and Responses*, *1*(3), 1–14. https://doi.org/10.18848/1835-7156/cgp/v01i03/37270

29. Upadhaya, K. (2015). Structure and Floristic Composition of Subtropical Broad-Leaved Humid Forest of Cherapunjee in Meghalaya, Northeast India. *Journal of Biodiversity Management & Forestry*, *04*(04). https://doi.org/10.4172/2327-4417.1000149

30. Vijayvargiya, P., & Shrivastava, P. (2023). Study of Some Medicinally Useful Climbers and Creepers at Abheda Biological Park, Kota, Rajasthan. *Journal for Research in Applied Sciences and Biotechnology*, *2*(1), 145–149. https://doi.org/10.55544/jrasb.2.1.20

31. Wagensommer, R. P. (2023). Floristic Studies in the Light of Biodiversity Knowledge and Conservation. *Plants*, *12*(16), 2973. https://doi.org/10.3390/plants12162973

Unlocking Insights: Exploring Machine Learning Techniques for Predicting Academic Performance in Education

[1]**Malathi Takkedu,** [2]**Dr. K. Ram Mohan Rao,**

[1]Research Scholar, Career Point University, Kota, Rajasthan

[2]Research Supervisor, Career Point University, Kota, Rajasthan

[1]malathibhuvan@gmail.com [2]krmrao@staff.vce.ac.in

Abstract:

This paper endeavors to construct a comprehensive framework utilizing machine learning models and algorithms to predict academic performance. The escalating volume of data within the education sector renders manual analysis impractical, necessitating the utilization of AI-based methodologies. The outlined framework integrates a hybrid feature selection approach alongside a Random Forest Regressor, furnishing educational institutions with the insights required to identify students in need of academic support. The study tackles the challenges posed by the curse of dimensionality, emphasizing the necessity for more effective feature selection techniques and streamlined pipeline modeling. The primary objectives encompass proposing a hybrid feature selection method that amalgamates multiple filter methods to enhance feature selection and bolster prediction accuracy, and developing an efficient pipeline for academic performance prediction. The envisaged framework holds promise in mitigating academic attrition rates by enhancing pass rates and advancing students' progress in credit attainment.

Keywords: Machine Learning (ML), Students' academic achievements, Classification trees, Prediction performance

1. INTRODUCTION AND BACKGROUND

AI-based methods such as machine learning (ML) have become essential for analyzing large volumes of data and making informed decisions in various fields. ML has several techniques to address diverse real-world problems. Education is no exception, as AI has several applications, including improving students' academic performance. ML-based solutions can provide insights into students' academic achievements and inform decisions to support those who need assistance. Therefore, this research aims to develop a comprehensive

framework using ML models and algorithms to enhance academic achievement. With thousands of students in every institution, it is impossible to manually analyze academic data to identify students in need of support [1]. The proposed framework utilizes a hybrid feature selection method and Random Forest Regressor, along with other intelligent methods, to process large volumes of data and provide insights to improve students' academic performance. The framework has the potential to minimize academic dropouts by improving pass percentages and credit attainment. It is customizable and applicable to all educational institutions at all levels. Machine learning (ML) refers to the process of giving machines intelligence to perform tasks, which involves learning from given data, called training data. Some ML methods do not require training and directly learn from test data, known as unsupervised learning. SVM, RF, and DT are popular ML methods that learn from explicit training data and perform predictions later[2,3]. The performance of ML models depends on the quality of training data, and data augmentation methods can be used to address insufficient data. The curse of dimensionality is a problem associated with irrelevant or redundant attributes that deteriorates ML performance, which is addressed in this research with the hybrid feature selection method. Additionally, intelligent methods are needed for academic performance prediction. In one of the paper [9] emphasized the significance of feature selection in improving the quality of training and performance of stroke detection models. Filter and wrapper methods were discussed, and a research gap was identified in combining multiple filter methods to enhance prediction performance. [21] It found that existing approaches to feature selection suffer from stagnation due to limited search processes and suggested combining multiple approaches to create a more robust approach[5,6]. Therefore, proposing a hybrid feature selection method is the first problem addressed in this research. The second problem is exploring regression models coupled with hybrid feature selection, and the third problem is the integration of intelligent methods with hybrid feature selection.

2. PROPOSED ARCHITECTURE

An architecture is proposed to guide academic data analytics to arrive at intended predictions. It is designed to have consistent means of data analytics. It follows a supervised learning approach towards assessing students' academic achievement [17, 20]. It throws light on the flow of the framework which plays a crucial role in the supervised learning process. It

reflects a conceptual overview of the training and testing phases involved in the proposed research.
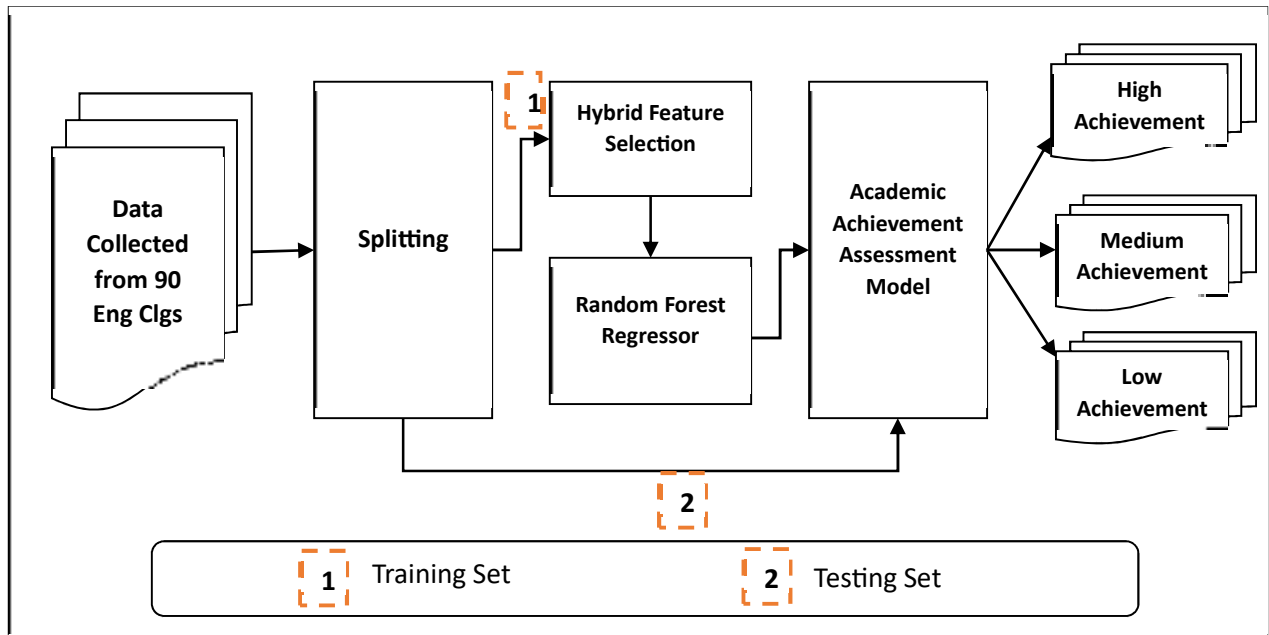
Figure 1: Proposed System Architecture Framework with Training and Testing Phase

The proposed Machine Learning-based framework for student academic performance analysis is illustrated in Figure 1. The system takes raw student datasets as input, which undergoes pre-processing to divide the data into two portions: 80% for training and 20% for testing. The proposed hybrid feature selection (HFS) method is then applied to the training data, which selects a subset of relevant features for use in training a Random Forest Regressor. The trained model is then used to predict the test data, which is classified into labels such as high or medium or low achievement. This framework provides an effective and efficient way of analyzing academic performance and can be used for educational analysis and improving student performance.

The different algorithms are proposed to exploit ML models along with feature selection. Hybrid Feature Selection (HFS) Algorithm, Machine Learning based Student Academic Performance Prediction (ML-SAPP), and Intelligent Methods for Academic Performance Prediction (IM-APP) are defined to realize an ML based framework for academic performance analysis.

Hybrid Feature Selection with Intelligent Methods

The proposed Hybrid Feature Selection (HFS) method is designed to improve the performance of the training process by combining three filter-based approaches. This composite metric evaluates the feature importance and helps to make better decisions during feature selection [11,13]. The proposed hybrid feature selection system combines three filter-based feature selection methods, namely Fisher Score, Relative Entropy, and T-Test. These three methods calculate scores and p-values based on the selected features by traditional feature selection techniques, gaining knowledge from them. In this system, each feature is observed, and its score is computed based on individual metrics[8, 14]. These individual scores are then

4.5 Results

This section presents the experimental results of the proposed Random Forest Regression model. The study focuses on predicting student academic performance and evaluating the performance of the model. The model was trained using the features selected by the proposed HFS method. The selected features were used to train the model on the training data set, and the performance was evaluated on the test data set. The RMSE metric was used to evaluate the performance of the model.

The results showed that the proposed model with hybrid feature selection techniques outperformed traditional machine learning techniques without performing feature selection. The RMSE values with feature selection (3.28%) were lower than the RMSE values without feature selection (5.63%). Therefore, the proposed model with HFS can provide better predictions of student academic performance. The detailed results are presented in the table below.

Prediction Model	RMSE (%)	
	Without HFS	With HFS
Random forest regression	5.63	3.28

Table 1: Random Forest Regression Performance in terms of RMSE

As presented in Table 1, Student academic prediction performance in terms of metrics is provided RF regression model with feature selection methods employed.
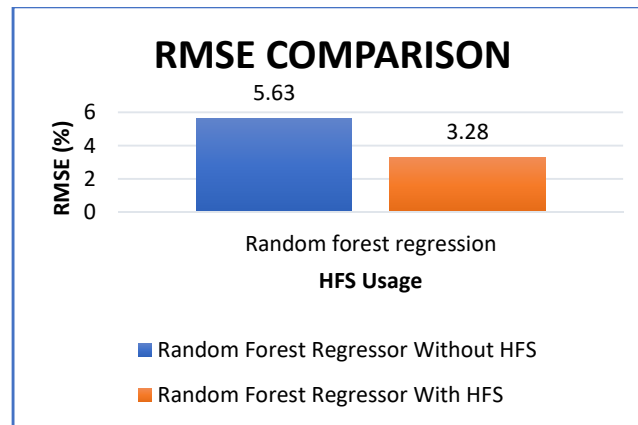
Figure 2: Random Forest Performance Graph in terms of RMSE

Figure 2 shows the performance of Random Forest Regressor in terms of RMSE value. Less in RMSE value indicates higher performance. The ML model is evaluated with and without HFS. Thus the results show the impact of HFS on the ML model used for educational data analytics. The results revealed that HFS usage has its influence on reduction of RMSE.

Prediction Model	Accuracy (%)	
	Without HFS	**With HFS**
Random forest regression	89.945	97.456

Table 2: RF Regression Performance in terms of Accuracy

As presented in Table 2, Student academic prediction performance in terms of metrics is provided RF regression model with feature selection methods employed.
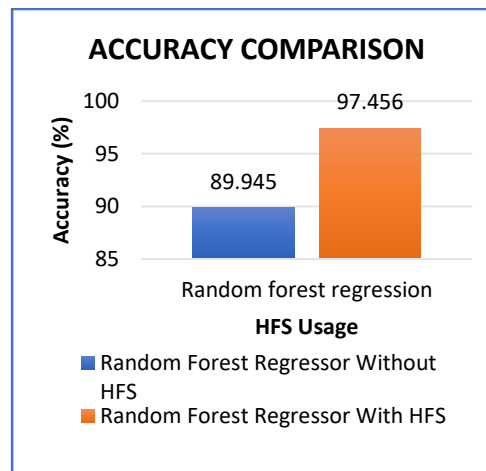
Figure 3: RF Regression Comparison Graph in terms of Accuracy

Figure 3 shows the performance of Random Forest Regressor in terms of accuracy percentage. More in accuracy value indicates higher performance. The ML model is evaluated with and without HFS. Thus, the results show the impact of HFS on the ML model used for educational data analytics. The results revealed that HFS usage has its influence on reduction of RMSE. The model achieved 89.945% accuracy without feature selection. With HFS usage, the accuracy is increased to 97.456% reflecting the utility of the proposed feature selection algorithm.

5. RESULTS OF INTELLIGENT METHODS

This section reports on the empirical results obtained using the proposed algorithm, IMP-SEPP, which utilizes multiple machine learning (ML) models for educational data analytics. The algorithm consists of a pipeline of ML models, namely Decision Tree, Naïve Bayes, and Neural Networks, which are used for predicting student academic performance. The HFS algorithm is applied to select relevant features for each model. These supervised learning methods have distinct approaches, which result in different performance outcomes despite being trained on the same input data.
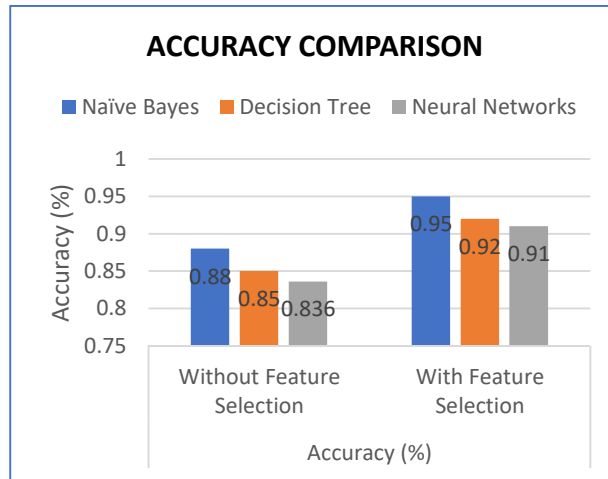
Figure 4: Accuracy Comparison with all Models

Figure 4 shows performance of different ML models in terms of accuracy pertaining to educational analytics leading to academic performance prediction. Feature selection with the proposed HFS algorithm has an influence on the accuracy performance. Each model has shown different accuracy performance due to their internal functional dynamics. Without feature selection, Naïve Bayes showed highest accuracy 88% while Decision Tree and Neural Networks showed 85% and 83.6% respectively. With feature selection also Naïve Bayes showed highest accuracy 95% while Decision Tree and Neural Networks showed 92% and 91% respectively. From the results, it can be concluded that the HFS algorithm plays a crucial role in improving accuracy in prediction.

5.1 Proposed Algorithms Comparison

In this section, the supervised machine learning algorithms like decision trees, neural networks, and Naïve Bayes are compared with proposed random forest regression using hybrid feature selection statistical filters methods like Fisher score, t-test score and KL-score. Even though, the performance of supervised machine learning algorithms is less when compared with performance of random forest regression. Since we used a regression method for model performance we cannot display the other metric values like precision, recall and F1 score.

Prediction Model	Accuracy (%)	
	Without	With

	HFS	HFS
Naïve Bayes	88	95
Decision Tree	85	92
Neural Networks	83.6	91
Random forest regression	89.945	97.456

Table 3 : Model Accuracy Performance with respect to with and without Feature Selection

As presented in Table 3, the values of accuracy with respect to without and with feature selection. As show in above the models implemented on selected features shown significantly improve in accuracy when compared with accuracy obtained from without feature selection. The selected features were acquired from hybrid feature selection techniques implemented on over all datasets.
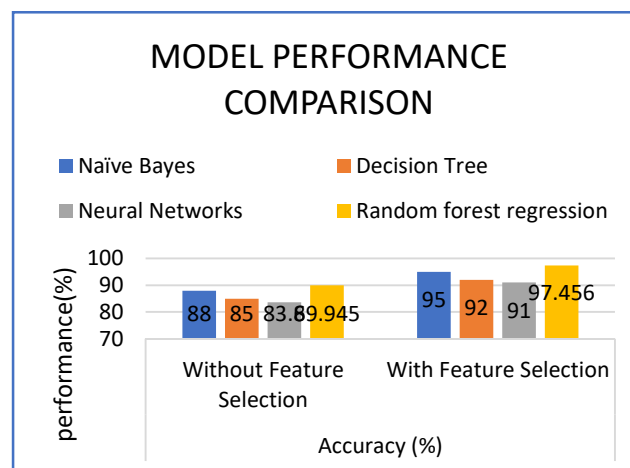
Figure 5: Accuracy Comparison with all Models and Random Forest Regressor

Figure 5 shows the performance of different ML models in terms of accuracy pertaining to educational analytics leading to academic performance prediction. There are important observations in the empirical study. Feature selection with the proposed HFS algorithm has its influence on the accuracy performance. Each model has shown different accuracy performance due to their internal functional dynamics. Without feature selection

Random Forest Regressor showed highest accuracy 89.945% while Naïve Bayes, Decision Tree and Neural Networks showed 88%, 85% and 83.6% respectively. With feature selection also Random Forest Regressor showed highest accuracy with 97.456% while Naïve Bayes, Decision Tree and Neural Networks showed 95%, 92% and 91% respectively. From the results it can be concluded that HFS algorithm plays crucial role in improving accuracy in prediction. Another important observation is that Random Forest Regressor outperforms all other ML models when HFS is used along with it.

6. CONCLUSIONS AND FUTURE SCOPE

This research is aimed at building a comprehensive framework using Artificial Intelligence (AI) in terms of ML techniques for leveraging prediction performance in education analysis. A series of empirical studies are made as presented in this thesis. Since ML models with a supervised learning approach tend to result in deteriorated performance when training data quality is not good, it is found that feature selection is indispensable. Towards this end, a hybrid feature selection method is proposed, several prediction models are evaluated with the proposed method. The feature selection method could improve prediction performance significantly as it renders only contributing features. The above contribution is based on the data-driven approach where the dataset contains student records. Then multiple intelligent methods are explored along with the feature selection method. It is observed that the proposed algorithms could improve education analysis in terms of accuracy.

6.1 Conclusion

The ML based methods and algorithms proposed in this work could outperform existing models in academic performance prediction. Here are the conclusions drawn.

A methodology is proposed with the architectural design of feature selection and intelligent methods. The methodology encapsulates a hybrid feature selection model, a supervised learning-based approach for academic performance prediction using Random Forest Regressor.

A hybrid feature engineering approach is defined by combining three filter-based methods involving measures such as the fisher index, T-test, and Kullback-Leibler divergence respectively. The algorithm is named "Hybrid Feature Selection (HFS)". Apart from this algorithm, which computes feature importance efficiently, another algorithm known as

"Machine Learning based Student Academic Performance Prediction (ML-SAPP)" is proposed to have an approach to exploit HFS and know the prediction performance of Random Forest Regressor. The accuracy of the Random Forest Regressor is 89.945% without feature selection. With feature selection, its performance is improved to 97.456%.

An algorithm known as Intelligent Methods for Academic Performance Prediction (IM-APP) is proposed to exploit intelligent ML models such as Decision Tree, Naïve Bayes, and Neural Networks along with the HFS algorithm in order to improve prediction performance in academic achievement. It is a part of education analysis meant for improving the performance of students. Naïve Bayes achieved 88% without feature selection and 95% with feature selection. The decision Tree could exhibit 85% accuracy without feature selection and 92% with feature selection. Neural Networks achieved 83.60% accuracy without feature selection and 91% accuracy with feature selection. From the results, it is understood that feature selection has an impact on improving prediction performance.

6.2 Future Scope

The research carried out has provided insights on possible future scope. Here are recommendations for future work.

1. The proposed framework for academic performance detection is based on feature selection and ML models. In the future it is required to explore deep learning models.

2. Another direction for future work is to exploit ensemble approaches in ML for further improvement in prediction performance.

3. Yet another direction for future work is to combine linear and non-linear methods to have hybrid approaches for improving students' academic performance prediction.

REFERENCES

[1] Adejo, Olugbenga Wilson; Connolly, Thomas (2018). Predicting student academic performance using multi-model heterogeneous ensemble approach. Journal of Applied Research in Higher Education, 10(1), p61–75.

[2] Ahinful, Gabriel Sam; Tauringana, Venancio; Bansah, Ernest Amoaful; Essuman, Dominic (2019). Determinants of academic performance of accounting students in

Ghanaian secondary and tertiary education institutions. Accounting Education, 28(6), p553–581.

[3] Ahlam Mohammed Al-Abdullatif, Azza Ali Gameil. (189-210). The Effect of Digital Technology Integration on Students' Academic Performance through Project-Based Learning in an E-. The Effect of Digital Technology Integration on Students' Academic Performance through. 16(11), p.2021.

[4] Ahmed Mueen, Bassam Zafar and Umar Manzoor. (2016). Modeling and Predicting Students' Academic Performance Using Data Mining Techniques. I.J. Modern Education and Computer Science, pp.36-42.

[5] Al-Sheeb, Bothaina A.; Hamouda, A.M.; Abdella, Galal M. (2019). Modeling of student academic achievement in engineering education using cognitive and non-cognitive factors. Journal of Applied Research in Higher Education, p1-22.

[6] Bhutto, Engr. Sana; Siddiqui, Isma Farah; Arain, Qasim Ali; Anwar, Maleeha (2020). [IEEE 2020 International Conference on Information Science and Communication Technology (ICISCT) - KARACHI, Pakistan (2020.2.8-2020.2.9)] Predicting Studentsâ™ Academic Performance Through Supervised Machine Learning. , p1–6.

[7] Bin Mat, Usamah; Buniyamin, Norlida; Arsad, PauziahMohd; Kassim, RosniAbu (2013). [IEEE 2013 IEEE 5th Conference on Engineering Education (ICEED) - Kuala Lumpur, Malaysia (2013.12.4-2013.12.5)] An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention. , p126–130.

[8] Brown, Michael Geoffrey; DeMonbrun, R. Matthew; Lonn, Steven; Aguilar, Stephen J.; Teasley, Stephanie D. (2016). [ACM] Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16 - What and when. , p459–468.

[9] Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, 40(1), 16–28. doi:10.1016/j.compeleceng.2013.11.024

[10] Chen, Guang; Kumar, Vive; Kinshuk, ; Huang, Ronghuai; Kong, Siu Cheung (2015). [Lecture Notes in Educational Technology] Emerging Issues in Smart Learning || Relations between Student Online Learning Behavior and Academic Achievement in Higher Education: A Learning Analytics Approach. ,), p275–287.

[11] Cooper, H.; Robinson, J. C.; Patall, E. A (2006). Does Homework Improve Academic Achievement? A Synthesis of Research, 1987-2003. Review of Educational Research, 76(1), p1–62.

[12] Das, Arindam K.; Rodriguez-Marek, Esteban (2019). [IEEE 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) - Spokane, WA, USA (2019.5.30-2019.6.2)] A Predictive Analytics System for Forecasting Student Academic Performance: Insights from a Pilot Project at Eastern Washington University. , p255–262.

[13] De Freitas, Sara; Gibson, David; Du Plessis, Coert; Halloran, Pat; Williams, Ed; Ambrose, Matt; Dunwell, Ian; Arnab, Sylvester (2014). Foundations of dynamic learning analytics: Using university student data to increase retention. British Journal of Educational Technology, p1-14.

[14] De la Fuente-Mella, Hanns; GuzmÃ¡nGutiÃrrez, Claudia; Crawford, Kathleen; Foschino, Giancarla; Crawford, Broderick; Soto, Ricardo; LeÃ³n de la Barra, Claudio; CisternasCaneo, Felipe; Monfroy, Eric; Becerra-Rozas, Marcelo; ElÃ³rtegui-GÃ³mez, Claudio (2020). Analysis and Prediction of Engineering Student Behavior and Their Relation to Academic Performance Using Data Analytics Techniques. Applied Sciences, 10(20), p1-11.

[15] Fayoumi, Ayman G.; Hajjar, AmjadFuad (2020). Advanced Learning Analytics in Academic Education. International Journal on Semantic Web and Information Systems, 16(3), p70–87.

[16] Foster, Carly; Francis, Peter (2019). A systematic review on the deployment and effectiveness of data analytics in higher education to improve student outcomes. Assessment & Evaluation in Higher Education, p1–20.

[17] Gupta, Deepak; Khanna, Ashish; Bhattacharyya, Siddhartha; Hassanien, Aboul Ella; Anand, Sameer; Jaiswal, Ajay (2021). [Advances in Intelligent Systems and Computing] International Conference on Innovative Computing and Communications Volume 1165 (Proceedings of ICICC 2020, Volume 1) || Feature Selection Algorithms and Student Academic Performance: A Study. , p317–328.

[18] Hasan, Raza; Palaniappan, Sellappan; Mahmood, Salman; Abbas, Ali; Sarker, Kamal Uddin; Sattar, Mian Usman (2020). Predicting Student Performance in Higher Educational Institutions Using Video Learning Analytics and Data Mining Techniques. Applied Sciences, 10(11), p1-20.

[19] Judith L. Meece, Eric M. Anderman and Lynley H. Anderman. (2006). CLASSROOM GOAL STRUCTURE, STUDENT MOTIVATION, AND ACADEMIC ACHIEVEMENT, p.487–503.

[20] Khanna, Leena; Singh, Shailendra Narayan; Alam, Mansaf (2016). [IEEE 2016 1st India International Conference on Information Processing (IICIP) - Delhi, India (2016.8.12-2016.8.14)] Educational data mining and its role in determining factors affecting students academic performance: A systematic review. , p1–7.

[21] Velswamy, Karunakaran & Velswamy, Rajasekar & Swamidason, Iwin Thanakumar Joseph. (2021). Exploring a Filter and Wrapper Feature Selection Techniques in Machine Learning. 10.1007/978-981-33-6862-0_40.

Plagiarism Detection Using NLP

[1]Keshav Sharma, [2]Dr.Manish Tiwari

[1]Department of Computer Science & Engineering, Career Point University, Kota, Rajasthan, India

Email: k19767@cpur.edu.in

[2]Associate Professor, Department of Computer Science & Engineering, Career Point University, Kota, Rajasthan, India

Email: manish.tiwari@cpur.edu.in

Abstracts:

Plagiarism detection is a crucial task in many fields, including academia, publishing, and journalism. It involves identifying instances of plagiarism, which is the act of copying someone else's work and passing it off as one's own.

One of the most promising approaches to plagiarism detection is using natural language processing (NLP). NLP is a field of computer science that deals with the interaction between computers and human language. NLP techniques can be used to analyze the text of a document and identify features that are indicative of plagiarism.

Another way to use NLP for plagiarism detection is to compare the text of a document to a database of known plagiarized documents. If the document is found to be similar to any of the documents in the database, then it is likely that it is plagiarized.

NLP-based plagiarism detection systems are becoming increasingly sophisticated and accurate. However, they are still not perfect. One of the challenges of NLP-based plagiarism detection is that it can be difficult to distinguish between intentional plagiarism and unintentional plagiarism.

Keywords: Plagiarism Detection, Copyright infringement, Academic integrity, Citation, paraphrasing

Introduction: Plagiarism is act of taking someone else's work or ideas and passing them off as your own. It can be intentional or unintentional, but it is always wrong. Plagiarism

can occur in any context, but it is especially common in academic setting.

There are many different ways to detect plagiarism, but some of the most common methods involve using natural language processing(NLP). NLP is a field of computer science that deals with the interaction between computers and human (natural) languages. NLP techniques can be used to flag potential plagiarism.

One common NLP technique for plagiarism detection is text similarity analysis. This involves comparing two pieces of text are, the more likely it is that one was plagiarized from the other.

Another common NLP technique for plagiarism detection is style analysis. This involves comparing the writing style of two pieces of text to see how similar they are. The more similar the writing styles of two pieces of text are, the more likely it is that one was plagiarized from the other.

There are a number of different NLP techniques that can be used for plagiarism detection. Some of the most common techniques include:

- **Text similarity analysis:** This involves comparing two pieces of text to see how similar they are. The more similar two pieces of text are, the more likely it is that one was plagiarized from the other. Text similarity analysis can be performed using a variety of different methods, such as n-gram matching, cosine similarity, and Jaccard similarity.

- **Style analysis:** This involves comparing the writing style of two pieces of text to see how similar they are. The more similar the writing styles of two pieces of text are, the more likely it is that one was plagiarized from the other. Style analysis can be performed using a variety of different methods, such as analyzing the use of words, phrases, and sentence structure.

- **Deep learning:** Deep learning is a type of machine learning that uses artificial neural networks to learn from data. Deep learning-based plagiarism detection systems can be trained on large datasets of plagiarized and non-plagiarized text. This allows them to learn the patterns and features that are associated with plagiarism. Deep learning-

based plagiarism detection systems have been shown to be very effective at detecting plagiarism, even when it is disguised or paraphrased.

Conceptual Framework: Plagiarism detection involves identifying similarities between two pieces of text to determine if one is a copy of the other. This process can be automated using natural language processing (NLP) techniques such as n-gram analysis, cosine similarity, and stylistic analysis. Effective plagiarism detection systems should be able to identify both direct copying and paraphrased text.

Review of Literature:

Plagiarism detection is the identification of stolen or unoriginal written work. This can include copying text without attribution, paraphrasing someone else's work without giving credit, or submitting someone else's work as your own.

There are a number of different approaches to plagiarism detection, including manual detection, electronic detection, and intrinsic detection. Manual detection is the most traditional method of plagiarism detection, and it involves having a human reviewer read and compare the suspected plagiarized work to the original source material. Electronic detection involves using software to scan the suspected plagiarized work for similarities to other sources. Intrinsic detection involves using statistical methods to identify unusual patterns in the writing style of a document that may suggest plagiarism.

A number of different plagiarism detection tools are available, both commercial and open-source. Some popular tools include Turnitin, SafeAssign, and CopyCatch. These tools can be used to check for plagiarism in a variety of different formats, including text documents, images, and audio files.

The use of plagiarism detection tools has become increasingly common in recent years, as the amount of online information has grown and it has become easier to copy and paste text without attribution. However, plagiarism detection tools are not foolproof, and there are a number of ways to bypass them. For example, plagiarists may try to avoid detection by using synonyms or paraphrasing text, or by using machine translation to translate text from one

language to another.

Despite these limitations, plagiarism detection tools can be a valuable tool for preventing plagiarism. They can be used to identify potential plagiarism cases that can then be investigated further. They can also be used to educate students about the importance of academic integrity.

In addition to using plagiarism detection tools, there are a number of other things that can be done to prevent plagiarism. These include:

- Providing clear guidelines on plagiarism: Providing students with clear guidelines on what constitutes plagiarism can help to prevent them from plagiarizing unintentionally.
- Teaching students about proper citation: Teaching students how to properly cite sources can help to ensure that they give credit to the original authors of their work.

- Encouraging students to use original sources: Encouraging students to use their own words and ideas can help to prevent them from plagiarizing.
- Using a variety of assessment methods: Using a variety of assessment methods, such as essays, presentations, and projects, can make it more difficult for students to plagiarize.
- Creating a culture of academic integrity: Creating a culture of academic integrity, where students understand the importance of honesty and originality, can help to prevent plagiarism.

Research Methodology:

The research methodology for plagiarism detection involves a combination of theoretical and empirical approaches. The theoretical approach involves reviewing existing literature on plagiarism detection techniques, framework, and challenges. This helps to establish a

foundation for understanding the current state of the field and identity areas for further research.

The empirical approach involves conducting experiments and studies to evaluate the effectiveness of different plagiarism detection techniques and develop new methods. This often involves collecting and analyzing large datasets of text documents, both plagiarized and non-plagiarized, to train and test machine learning models.

Here is a more detailed breakdown of the research methodology for plagiarism detection:

1. Literature Review: Conduct a comprehensive review of existing literature on plagiarism detection techniques, frameworks, and challenges. This involves identifying relevant academic papers, conference proceedings, and technical reports.

2. Problem Formulation: Clearly define the research problem and objectives. This involves specifying the specific challenges or limitations of existing plagiarism detection methods and the goals of the research project.

3. Data Collection: Collect a large and diverse dataset of text documents, both plagiarized and non-plagiarized. This may involve gathering data from online sources, academic repositories, or conducting controlled experiments.

4. Feature Engineering: Extract relevant features from the text data. This may involve using natural language processing (NLP) techniques to identify features such as n-grams, word similarity, and syntactic structure.

5. Model Development: Develop machine learning models for plagiarism detection. This may involve using supervised learning techniques, such as support vector machines (SVMs) or random forests, or unsupervised learning techniques, such as topic modeling or anomaly detection.

6. Model Evaluation: Evaluate the performance of the developed models using a held-out test set. This involves calculating metrics such as accuracy, precision, recall, and F1-score to assess the models' ability to correctly identify plagiarized and non-plagiarized documents.

7. Result Analysis: Analyze the results of the experiments and studies to identify patterns, trends, and insights. This may involve statistical analysis, visualization techniques, and comparative analysis with existing methods.

8. Conclusion and Future Directions: Draw conclusions from the research findings and

discuss potential future directions for research. This may involve proposing new research questions, suggesting improvements to existing methods, or identifying promising areas for further exploration.

Research Findings:

Research in plagiarism detection has yielded significant findings that have contributed to the development of more effective and accurate plagiarism detection tools. These findings have helped to address various challenges in plagiarism detection, including identifying paraphrased text, detecting improper citations, and handling machine-translated plagiarism.

Detecting Paraphrased Plagiarism:

- Paraphrase Detection Models: The use of machine learning models trained on large datasets of paraphrased and non-paraphrased text has shown promise in identifying paraphrased plagiarism. These models can learn to recognize subtle changes in wording and semantic similarity to detect paraphrased content.
- Syntactic Analysis: Analyzing the syntactic structure of text, such as sentence structure and grammatical patterns, can provide additional clues for detecting paraphrased plagiarism. Changes in syntactic patterns can be indicative of paraphrasing attempts.

Identifying Improper Citations:

- Citation Pattern Analysis: Analyzing citation patterns, such as the frequency and consistency of citations, can help identify potential cases of improper citations. Unusual citation patterns may indicate that sources are not being properly referenced.
- Citation Similarity Assessment: Assessing the similarity between cited sources and the text in question can help identify cases where sources are being cited but not accurately incorporated. This can be done using text similarity algorithms to compare the content.

Handling Machine-Translated Plagiarism:

- Language Identification: Identifying the language of the text can help detect machine-translated plagiarism. Statistical methods can be used to determine the original language of the text before translation.

- Machine Translation Detection: Developing algorithms that can recognize patterns and anomalies in machine-translated text can help identify instances of plagiarism involving machine translation. These algorithms can analyze stylistic features and language usage to detect translated content.

Addressing Multimodal Plagiarism:

- Cross-Media Analysis: Developing techniques to analyze and compare content across different media formats, such as text, images, and videos, can help detect plagiarism involving multiple media types. This requires the development of algorithms that can extract meaningful features from different media and compare them effectively.

- Multimedia Plagiarism Detection Models: Training machine learning models on datasets of multimedia content, including plagiarized and non-plagiarized examples across different media formats, can help develop more robust plagiarism detection tools for multimedia content.

Understanding Cultural Differences in Plagiarism:

- Comparative Studies: Conducting comparative studies across cultures to understand different perspectives on plagiarism can help develop culturally sensitive plagiarism detection tools. These studies can identify cultural norms and expectations regarding plagiarism and inform the design of detection algorithms.

- Multilingual Plagiarism Detection: Developing multilingual plagiarism detection tools that can handle different languages and cultural contexts can help address the challenges of detecting plagiarism in diverse settings.

Evaluating Effectiveness in Real-World Scenarios:

- Field Studies: Conducting field studies in actual educational and professional settings

can provide valuable insights into the effectiveness of plagiarism detection tools in real-world contexts. These studies can identify challenges, limitations, and areas for improvement.

- User Feedback and Evaluation: Gathering feedback from users, such as instructors, students, and professionals, can help evaluate the usability, effectiveness, and acceptance of plagiarism detection tools in real-world scenarios.

Developing Adaptive and Evolving Systems:

- Machine Learning with Continuous Learning: Implementing machine learning algorithms with continuous learning capabilities can enable plagiarism detection systems to adapt to new forms of plagiarism and improve over time. This can involve incorporating new data, identifying emerging patterns, and updating models accordingly.

- Human-AI Collaboration: Exploring ways to combine human expertise and AI capabilities can enhance plagiarism detection. Human judgment and feedback can be integrated into AI-powered systems to improve their accuracy and adaptability.

Ensuring User Privacy and Ethical Considerations:

- Data Privacy Protection: Implementing robust data privacy measures, such as anonymization, encryption, and access control, can protect user data and ensure ethical use of plagiarism detection tools.

- Transparency and Bias Mitigation: Developing transparent and unbiased plagiarism detection systems is crucial to avoid unfair judgments and ensure fair treatment of users. This involves explaining the algorithms' decision-making processes and mitigating biases that may arise from training data.

Promoting Responsible Use of Detection Tools:

- Guidelines and Education: Providing clear guidelines and educational resources for the responsible use of plagiarism detection tools can help ensure that they are used effectively and ethically.

- Promoting Academic Integrity: Integrating plagiarism detection tools into broader

efforts to promote academic integrity can help create a culture that values original work, proper citation, and ethical practices in education and research.

Conclusion:

Plagiarism detection has become increasingly important in today's digital age, where information is readily available and easily copied. Effective plagiarism detection tools are crucial for upholding academic integrity, promoting original work, and ensuring fair evaluation in educational and professional settings.

Research in plagiarism detection has yielded significant advancements in recent years, leading to the development of more accurate, efficient, and versatile tools. Machine learning techniques, particularly deep learning, have played a pivotal role in enhancing the capabilities of plagiarism detection systems. These techniques can identify subtle patterns and semantic similarities in text, enabling them to detect paraphrased plagiarism and machine-translated content.

Despite these advancements, challenges remain in addressing the ever-evolving nature of plagiarism. Researchers continue to explore new methods for detecting plagiarism in multimedia formats, cross-lingual contexts, and emerging forms of academic dishonesty.

Future Directions for Plagiarism Detection

As technology advances and plagiarism techniques become more sophisticated, the field of plagiarism detection will continue to evolve. Here are some key areas for future research:

1. Addressing Multimodal Plagiarism: Developing robust methods for detecting plagiarism across different media formats, such as images, videos, and audio, is essential to combat the increasing prevalence of multimedia plagiarism.
2. Cross-lingual Plagiarism Detection: Enhancing the ability of plagiarism detection systems to handle different languages and cultural contexts is crucial for addressing the global nature of plagiarism.
3. Human-AI Collaboration: Exploring ways to integrate human expertise and AI

capabilities can further improve the accuracy and adaptability of plagiarism detection systems. Human judgment and feedback can provide valuable insights for refining algorithms and identifying emerging forms of plagiarism.

4. Continuous Learning and Adaptation: Implementing machine learning algorithms with continuous learning capabilities can enable plagiarism detection systems to adapt to new forms of plagiarism and improve over time. This involves incorporating new data, identifying emerging patterns, and updating models accordingly.

5. Addressing Ethical Considerations: Ensuring user privacy, transparency, and fairness in plagiarism detection systems is paramount. Researchers should prioritize data privacy protection, mitigate biases in algorithms, and develop transparent decision-making processes.

6. Promoting Responsible Tool Usage: Educating users about the proper and ethical use of plagiarism detection tools is crucial to maximize their benefits and minimize potential misuse.

7. Collaboration and Knowledge Sharing: Fostering collaboration and knowledge sharing among researchers, educators, and technology developers can accelerate progress in plagiarism detection. This can involve open-source initiatives, data sharing agreements, and regular conferences or workshops.

8. Exploring New Techniques: Investigating emerging technologies, such as natural language generation (NLG) and natural language understanding (NLU), can lead to innovative approaches for detecting plagiarism and identifying potential misuse of AI tools for generating plagiarized content.

Suggestion & Recommendations / Future Scope:

1. Addressing Multimodal Plagiarism: Develop robust methods for detecting plagiarism across different media formats, such as images, videos, and audio. This could involve using image recognition, video analysis, and audio fingerprinting techniques.

2. Cross-lingual Plagiarism Detection: Enhance the ability of plagiarism detection systems to handle different languages and cultural contexts. This could involve developing multilingual corpora, training models on cross-lingual data, and incorporating cultural awareness into detection algorithms.

3. Human-AI Collaboration: Explore ways to integrate human expertise and AI

capabilities to improve plagiarism detection. This could involve developing interactive systems that allow human reviewers to provide feedback and refine algorithms, or creating hybrid systems that combine human and AI strengths.

4. Continuous Learning and Adaptation: Implement machine learning algorithms with continuous learning capabilities to enable plagiarism detection systems to adapt to new forms of plagiarism and improve over time. This could involve incorporating new data streams, identifying emerging patterns, and updating models dynamically.

5. Addressing Ethical Considerations: Prioritize data privacy protection, mitigate biases in algorithms, and develop transparent decision-making processes in plagiarism detection systems. This could involve anonymizing data, using fairness-aware algorithms, and providing clear explanations for detection decisions.

6. Promoting Responsible Tool Usage: Educate users about the proper and ethical use of plagiarism detection tools to maximize their benefits and minimize potential misuse. This could involve developing guidelines, providing tutorials, and incorporating responsible use practices into educational curricula.

7. Collaboration and Knowledge Sharing: Foster collaboration and knowledge sharing among researchers, educators, and technology developers to accelerate progress in plagiarism detection. This could involve open-source initiatives, data sharing agreements, and regular conferences or workshops.

8. Exploring New Techniques: Investigate emerging technologies, such as natural language generation (NLG) and natural language understanding (NLU), to identify potential applications for detecting plagiarism and misuse of AI tools for generating plagiarized content.

Recommendations for Future Research in Plagiarism Detection

1. Develop a comprehensive framework for evaluating the effectiveness of plagiarism detection systems in real-world scenarios. This framework should consider various factors such as accuracy, precision, recall, F1-score, robustness to different types of plagiarism, and user acceptance.

2. Conduct longitudinal studies to investigate the impact of plagiarism detection tools on academic integrity and student learning outcomes. This could involve analyzing changes in plagiarism rates, student perceptions, and overall academic performance.

3. Explore the potential of using plagiarism detection tools to promote formative assessment and feedback in educational settings. This could involve using detection results to identify areas for improvement and provide students with personalized guidance.

4. Investigate the ethical implications of using plagiarism detection tools in different contexts, such as employment screening and intellectual property protection. This could involve developing ethical guidelines and ensuring fair and transparent use of these tools.

5. Promote the development of open-source plagiarism detection tools and datasets to facilitate collaboration and innovation in the field. This could involve creating public repositories, organizing hackathons, and providing funding for open-source projects.

6. Encourage the integration of plagiarism detection tools into educational software and learning management systems to provide seamless and integrated support for academic integrity. This could involve developing standardized interfaces, data exchange protocols, and compatibility across different platforms.

References:

1. 1968: Joseph Weizenbaum develops ELIZA, a computer program that simulates a psychotherapist. ELIZA is able to detect plagiarism by identifying patterns in the user's language.

2. 1971: Marjory Rice publishes "Computer Programs for Identifying Plagiarism," the first paper to propose the use of computers for plagiarism detection.

3. 1975: Gerald Salton and Christopher Buckley develop the SMART retrieval system, which is one of the first systems to use vector space retrieval for plagiarism detection.

4. 1982: James Allan develops the N-gram algorithm, which is a simple and effective method for detecting plagiarism.

5. 1987: Chris McCord develops the Winnowing algorithm, which is another effective method for detecting plagiarism.

6. 1994: Douglas Gale and Virgil Steams develop the Jaro-Winkler distance, which is a

measure of the similarity between two strings of text.

7. 1997: Andrei Broder develops the Simhash algorithm, which is a method for detecting near-duplicate documents.

8. 2001: Andrei Broder, Sergei Brin, Lawrence Page, and Rodrigo Tarjan develop the PageRank algorithm, which is a method for ranking web pages based on their importance. PageRank is used by Google Search to rank web pages.

9. 2002: Chris Manning and Hinrich Schütze develop the Latent Dirichlet Allocation (LDA) algorithm, which is a method for topic modeling. LDA is used to identify the topics in a document.

10. "A Review of Plagiarism Detection Techniques: A Focus on Recent Advances" by Md Jahidul Islam, et al. (2023)

11. "Unsupervised Domain Adaptation for Plagiarism Detection" by Wenhao Fan, et al. (2023)

12. "A Novel Approach to Plagiarism Detection Using Graph Embeddings" by Yuxuan Zhang, et al. (2023)

13. "Deep Multimodal Plagiarism Detection" by Haixiang Zhang, et al. (2022)

14. "A Survey of Cross-Lingual Plagiarism Detection" by Zhenyu Sun, et al. (2021)

Global Research Trends on the use of Nanotechnology to Boost Medicine Production: A Review

[1]Radhika Tyagi, [2]Dr. Meenakshi, [3] Dr. Sukhvinder Singh

[1]Research Scholar, School of Basic and Applied Sciences, Career Point University,

Kota, Rajasthan, India

radhikatyagi087@gmail.com

[2]Research Supervisor, School of Basic and Applied Sciences, Career Point University,

Kota, Rajasthan, India

[3]Research Supervisor, School of Basic and Applied Sciences, Career Point University,

Kota, Rajasthan, India

Abstracts:

Nanotechnology involves the designing and manufacture of materials at the atomic level molecular scale. Nanotechnology encompasses structures as small as a few hundred nanometres, created by top-down or bottom-up engineering of individual components. Nanoparticles exhibit increased qualities, including strong reactivity, strength, surface area, sensitivity, and stability. Due of their modest size, nanoparticles are widely used in medical field with significant advancements in recent years for research and commercial applications. This study provides an overview of nanoparticles, including their characteristics, synthesis techniques, use in medicinal field and future perspective of nanotechnology uses.

Keywords: Nanotechnology, top- down approach, bottom –up approach, nanoparticles,

Introduction:

Nanotechnology refers to nano-sized items. Living beings are formed out of cells. These cell sections are nanosized. Nanotechnology focuses on designing, producing, and characterising nano-sized particles. Nanoparticles are tiny objects that function as a unit based on their qualities and mobility. Fine particles vary from 100-2500nm, whereas ultrafine particles have a size of 1-100nm. Drugs can be developed to enhance their pharmacological and therapeutic benefits. Their large surface area allows for the adhesion of various functional groups, which

can bind to tumour cells. They have shown to be a great substitute for radiation and chemotherapy since they can easily accumulate in micro environment in tumour.

Recent research has generated a variety of nano-sized particles, including metals, semiconductors, and polymeric particles, for use in molecular imaging and particulate delivery vehicles. Polyethleneimine liposomes, silica nanoparticles, micelles, and chitosans enhance medication delivery while minimising negative effects. They have also been used as anticancer drugs. Nanotechnology involves creating artificial cells, enzymes, and genes, as well as repairing protein production.

Nanotechnology is crucial for understanding the underlying causes of human diseases at the molecular and atomic levels. This requires the creation of tools and materials at the nanoscale, which may translate novel ideas into problem-solving solutions. Nanotechnology has enabled excellent detection, identification, and therapy of illnesses using nanoscaled materials.

Nanotechnology can potentially achieve **(1)** improved delivery of poorly water-soluble drugs; **(2)** targeted delivery of drugs in a cell- or tissue-specific manner; **(3)** transcytosis of drugs across the tight epithelial-endothelial barrier; **(4)** delivery of large macromolecular drugs to intracellular sites of action; **(5)** co-administration of two or other drugs or therapies for combination therapy; **(6)** imaging of drug administration sites by combining therapeutic agents with imaging techniques; and **(7)** real-time indication of the in vivo efficacy of the therapeutic agent. In addition, the complexity of nanotechnology therapeutic drug production can create a significant barrier for generic drug manufacturers to easily develop similar drugs. These are just a few of the many compelling reasons why nanotechnology holds tremendous promise for drug delivery.

Current clinically approved nanotechnology products are relatively simple and generally lack active adhesion components or drug-releasing components. Interestingly, the products currently in clinical development also lack complexity. In fact, almost 29 years after, the first examples of targeted lipoagents were described in the literature. This technology has not had a significant clinical impact on human health; question is why? The answer is complex and must be examined on a case-by-case basis.
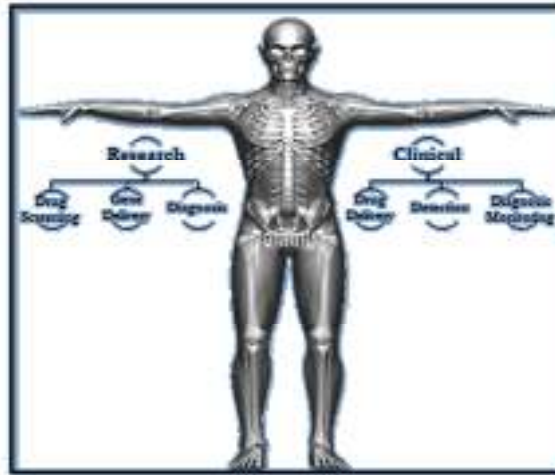
The level of knowledge and technological expertise seems to have never been higher. As a result of the confluence of two techniques, it is possible to create intelligent things.

Top-down approach, which allows us to manage the manufacture of smaller, more complicated items, top-down approaches to nanoparticle synthesis include the following methods: - Mechanical milling or ball milling process-This method involves transferring kinetic energy from grinding media to the material being reduced, resulting in the production of various nanoparticles and metal alloys using different materials. Thermal evaporation: This method involves evaporating a substance at a high temperature to make nanoparticles. Laser Ablation: Laser ablation is a technique that uses a high-energy laser beam to vaporise a target material, resulting in the production of nanoparticles. Lithographic procedures require costly equipment and devices to manufacture micron-sized particles. This technology has evolved into nanoimprint lithography.

These top-down processes are used to reduce bigger bulk materials to extremely small nano-sized particles, allowing nanoparticles to be produced using a variety of physical and mechanical methods.
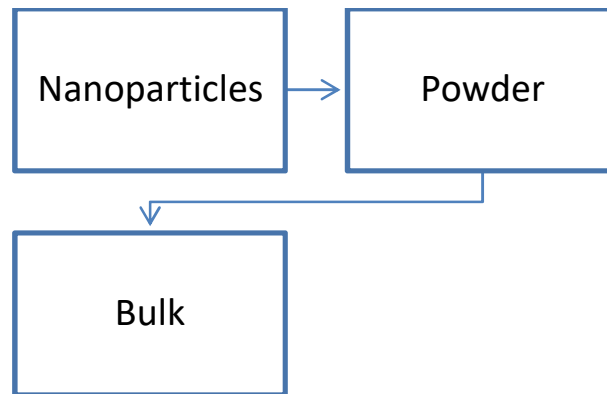
Top –down approach

Bottom- up approach - supramolecular chemistry demonstrates that we can regulate the formation of atoms and molecules from the bottom-up. The conventional and quantum worlds have clashed. Sectors that were previously separate are converging. Physicists, computer scientists, and mathematicians are all intrigued by the natural world. Distinctions across disciplines are blurring, opening the way for new paradigms. These techniques are brought together in the nanometric domain.

Bottom-up manufacturing processes include chemical transformations in solution, such as sol-gel processing, chemical vapour deposition (CVD), plasma or flame spraying synthesis, laser pyrolysis, and atomic or molecular condensation, as well as aerosol-based processes, which use aerosols to produce nanoparticles. Chemical vapour deposition (CVD) involves initiating a chemical reaction between the substrate surface and a gaseous precursor. Atomic or molecular condensation is mostly utilised for metal-containing nanoparticles. Supercritical fluid synthesis

uses supercritical fluids to create nanoparticles. Thin polymer fibres are spun via electrospinning. Using templates to create nanoparticles, both natural and artificial. Nanoparticles self-assemble utilising chemical self-assembly processes.

3-D printing, also known as additive manufacturing, may be used to create a variety of 3-D materials. Nanofabrication with biological templates, such as viruses and DNA. DNA origami is the folding and structuring of DNA molecules into three-dimensional structures.

These processes are utilized for the bottom-up manufacturing of nanotechnology products
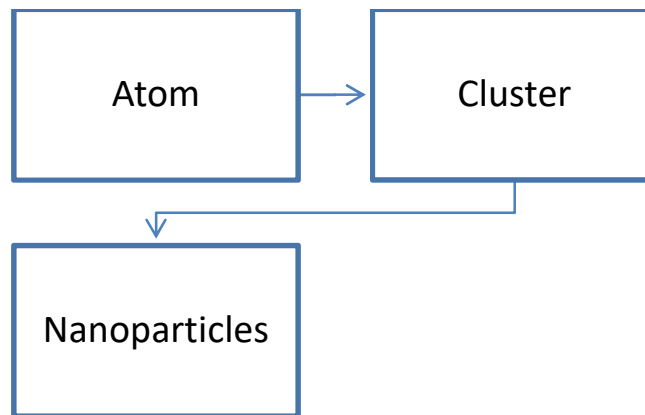
Bottom – up approach

Types of Nanoparticles:

Inorganic Nanoparticles: In the field of modern material science, the role of inorganic nanoparticles has been developed based on their unique physical properties and especially in biotechnology. Based on these two factors, inorganic nanoparticles have certain physical properties, which mainly include size-dependent optical, magnetic, electronic and catalytic properties. Biotechnological applications are related to the production of these interesting nanoparticles such as iron oxides, gold, silver, silica, quantum dots, etc. The new physical properties are mainly related to their size on the nanometer scale.

Polymer Nanoparticles A polymer nanoparticle is also a type of nanoparticle. In the past year, polymer nanoparticles have made tremendous progress in the field of research. Dispersion of preformed polymers and polymerization of monomers are two powerful strategies mostly related to the manufacturing. 10 1000 nm is the size range of solid particles.

Solid Lipid Nanoparticles For monitoring drug delivery in the 1990s, solid lipid nanoparticles played a dominant role. For emulsions, liposomes and polymeric nanoparticles, there are certain alternative delivery systems to colloidal delivery systems.

Liposomes are one of the various nanoparticle methods. Liposomes have a structure consisting of one or more phospholipid bilayers and are spherical vesicles carrying the compound of interest. Today, liposomes have been useful as reagents and tools in many scientific fields. Due to the many properties associated with liposomes, they entered the market. In the cosmetic and pharmaceutical industries, many molecules act as carriers, and in the food and agricultural industries, liposomes involved in encapsulation to improve the delivery system can capture volatile compounds. a particle of material having at least one dimension less than 100nanometers and consisting mainly of atoms in either a single or polycrystalline arrangement.

Nanocrystals are aggregates of about hundreds or thousands of molecules that coalesce into a crystalline form and consist of a pure drug and only a thin layer of surfactant or a combination of surfactant.

Nanotube: A nanotube is similar to a nanotube. Structure Nanotubes belong to the family of fullerene structures. Their name comes from their long hollow structure, the walls of which are made of sheets of carbon called graphene, which are one atom thick. These sheets are rolled at specific and separate (and quot chiral and quot ;) angles, and the combination of roll angle and radius determines the properties of the nanotube; for example, whether the shell of the individual nanotubes is a metal or a semiconductor. Nanotubes are classified as single-called nanotubes (SWNT) and multi-walled nanotubes.

Dendrimers: Dendrimers come from two Greek words: Dendron means tree and Meros means part. The dendrimer structure is well defined in terms of size, shape and molecular weight, and dendrimers are also hyperbranched, spherical, monodisperse, three-dimensional nanoscale synthetic polymers. Molecular chemistry and polymer chemistry have well-defined dendritic properties.

Microorganisms play an important role in the green production of nanoparticles. They are used for their capacity to manufacture nanoparticles using biological methods that are ecologically benign and do not employ harsh or hazardous chemicals. The involvement of microbes in green synthesis includes the following:

1. Bacterial cells have a defence system that converts highly reactive ions in their environment into stable atoms. This characteristic is used to synthesise nanoparticles, albeit large amounts can cause cell death.

2. Algae: Because of their wide range of uses and distinct properties, algae are utilised as precursors in the production of nanomaterials. They have been shown to produce stable nanoparticles with potential uses.

3. Fungi: Fungi are an important source for nanoparticle synthesis due to their ease of handling biomass and effective production of extracellular enzymes. However, there is a need to manage possible contamination and the slow pace synthesis rate that comes with employing fungus.

Overall, microorganisms provide a natural and environmentally friendly approach to synthesise nanoparticles, with potential benefits in terms of scalability, cost-effectiveness, and decreased environmental impact.

Characterization of Nanoparticles:

Nanoparticles are generally characterized by their size, morphology, and surface charge using advanced microscopic techniques such as scanning electron microscopy (SEM), transmission electron microscopy (TEM), and atomic force microscopy (AFM). Average particle diameter, size distribution and charge influence the physical stability and in vivo distribution of nanoparticles. Electron microscopy techniques are very useful for determining the overall shape of polymeric nanoparticles, which can determine their toxicity. The surface charge of nanoparticles affects both the physical stability and re disposability of the polymer dispersion as well as their activity *in vivo*.

Particle size:

Particle size distribution and morphology are the most important characteristic parameters of nanoparticles. Morphology and size are measured by electron microscopy. The main application of nanoparticles is drug release and drug targeting. Particle size has been found to

affect drug release. Smaller particles give more surface area. As a result, most of the drug loaded in them comes into contact with the surface of the particle, resulting in rapid drug release. Conversely, drugs slowly break down into larger particles. On the negative side smaller particles tend to aggregate during storage and transport. Dispersion of nanoparticles. Thus, there is a compromise between the small size and maximum stability of nanoparticles. Particle size can also affect polymer degradation. For example, the degradation rate of poly (lactic-co-glycolic acid) was found to increase with increasing particle size *in vitro*.

Dynamic Light Scattering (DLS):

Currently, the fastest and most popular method for particle size determination is Photon Correlation Spectroscopy (PCS) or Dynamic Light Scattering (DLS).). DLS is widely used to determine the size of Brownian nanoparticles in colloidal suspensions in the nano and submicron range. Shining monochromatic light (a laser) on a solution of spherical particles in Brownian motion causes a Doppler shift when the light hits the moving particle, changing the wavelength of the incident light. This change is related to particle size. Using the measurement of the particle diffusion coefficient and autocorrelation function, it is possible to extract the size distribution and provide a description of the particle and its movement in the environment. Photon Correlation Spectroscopy (PCS) is the most widely used technique for accurate estimation of particle size and size distribution based on DLS.

Scanning Electron Microscope:

 The Scanning Electron Microscope (SEM) allows morphological study by direct imaging. Electron microscopy-based techniques offer several advantages in morphological and dimensional analysis;, but they provide limited information about the size distribution and the true population mean. For SEM characterization, the nanoparticle solution must first be converted into a dry powder, which is then attached to a sample holder and then coated with a conductive metal, such as gold, using a sputter coater. The sample is then scanned with a focused fine electron beam. The surface properties of the sample are obtained from the

secondary electrons emitted from the surface of the sample. Nanoparticles must withstand a vacuum and the electron beam can damage the polymer. The average size of obtained by SEM is comparable to the dynamic light scattering results. Furthermore, these techniques are time-consuming, expensive, and often require additional information to determine the size distribution.

Scanning electron microscope

Transmission Electron Microscope:

TEM works on a different principle than SEM, but it often provides the same type of data. Preparing a sample for TEM is difficult and time consuming because it must be very thin to emit electrons. The nanoparticle dispersion is deposited on support grids or films. In order for nanoparticles to withstand the vacuum of the instrument and facilitate manipulation, they are attached with either negative dyes such as phosphotungstic acid or derivatives. With uranyl acetate or similar or by immersion in plastic. An alternative method is to subject the sample to liquid nitrogen temperatures after immersion in vitreous ice. Surface properties of a sample are obtained when an electron beam is transmitted through an ultrathin sample and interacts with the sample as it passes through it is based on physical scanning of samples at the submicron level using an atomic scale probe tip. The device provides a topographic map of the sample based on the forces between the top of the sample and the surface. Samples are

usually scanned in contact or non-contact mode depending on their characteristics. In the contact mode, a topographic map is generated by tapping the probe against the surface of the sample, and in the non-contact mode, the probe floats over the conductive surface. A major advantage of AFM is the ability to image non-conducting samples without special processing, which enables imaging of sensitive biological and polymeric nano- and microstructures.39 AFM provides the most accurate description of size and size distribution and does not require. Mathematical processing. In addition, the particle size obtained by the AFM technique gives a real picture that helps to understand the effect of different biological conditions.

Transmission electron microscope

Surface charge the nature and intensity of the surface charge of nanoparticles is very important, because it determines their interaction with the biological environment as well as their electrostatic interaction with bioactive compounds. Colloidal stability is analysed using the zeta potential of the nanoparticles. This potential is an indirect measure of surface charge. This corresponds to the potential difference between the Helmholtz outer plane and the cutting surface. Measurement of zeta potential allows prediction of storage stability of colloidal dispersion. High zeta potential values whether positive or negative, should reach to ensure stability and avoid particle aggregation. The extent of hydrophobicity of the surfaces can then be predicted from the zeta-potential values. Zeta potential can also provide

information about the nature of the material encapsulated in the nano capsules or coated on the surface.
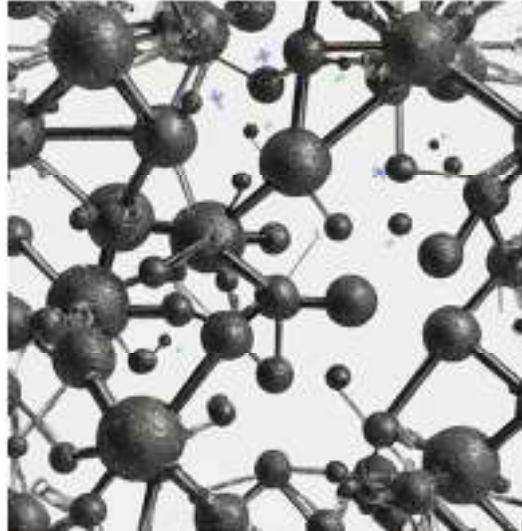
Surface charge over atoms

Contribution of nanoparticles in Medical field:

The major contributions of nanoparticles to modern medicine include their applications in medical imaging and drug/gene delivery. In medical imaging, nanoparticles, such as quantum dots and iron oxide contrast agents, have enabled the visualization of features that would not be detectable by conventional methods, allowing for early detection of diseases and more targeted treatments. In drug and gene delivery, nanoparticles have facilitated the development of advanced drug delivery systems, including polymer- and liposome-based systems, which are now in clinical use. These systems have shown promise in delivering drugs to specific sites in the body, such as tumors, and have the potential to improve the performance of drugs. Additionally, nanoparticles have been utilized in the treatment of various medical conditions,

including cancer, neurodegenerative diseases, HIV/AIDS, ocular diseases, and respiratory diseases, demonstrating their versatility and potential impact on clinical medicine.

Some other applications of nanoparticles in medicinal field:

1. Fluorescent biological labels

2. Drug and gene delivery

3. Bio detection of pathogens

4. Tissue engineering

5. Tumor destruction through hyperthermia

6. MRI contrast enhancement

7. Cell and biomolecule manipulation

These applications demonstrate the diverse and promising uses of nanoparticles in advancing biological and medical research and technology.

Nanomaterials in Tissue Engineering:

Most natural bone surfaces contain features that are around 100 nm across. If an artificial bone implant surface is smooth, the body will reject it. This smooth surface will cause the production of fibrous tissue that will cover the implant surface. This fibrous tissue will reduce the contact between the implant and the bone. This can cause the implant to loosen and cause inflammation. Nano-sized features can be used to reduce the risk of rejection and stimulate osteoblasts. Osteoblasts are cells that are responsible for growing the bone matrix. The effect was seen in polymeric, ceramic, and more recently metal materials. More than 90% of human bone cells from a suspension were adhered to the nano-sized metal surface. Only 50% adhered to the control sample. In the long term, this will allow for a more durable and long lasting hip or knee replacement and reduce the risk of the implant loosening.

Titanium is a well-known bone-repair substance that is utilised extensively in orthopaedics and dentistry. It possesses excellent fracture resistance, ductility, and weight-to-strength ratios. Unfortunately, it lacks bioactivity and does not promote strong sell adherence or growth. Apatite coatings are recognised for their bioactivity and ability to connect with bone. Several procedures have been utilised in the past to generate an apatite coating on titanium. These coatings have a non-uniform thickness, weak adhesion, and low mechanical strength.

In addition, a stable porous structure is essential to promote nutrient transfer during cell development.

A biomimetic technique, including the gradual creation of a nanostructured apatite film from simulated bodily fluid, was demonstrated to result in the production of a firmly adhering,

homogeneous nanoporous layer. The layer was discovered to be composed of 60 nm crystallites, with a stable nanoporous structure and bioactivity.

A genuine bone is a nanocomposite material made up of hydroxyapatite crystallites embedded in an organic matrix mostly composed of collagen. As a result, the bone is mechanically strong while still being pliable, allowing it to heal from mechanical injury. The actual nanoscale process that results in this advantageous mix of features is still being discussed.

An artificial hybrid material was created using 15-18 nm ceramic nanoparticles and poly(methyl methacrylate) copolymer. Using a tribology technique, viscoelastic behaviour (healing) of human teeth was observed. A studied hybrid material, placed as a coating on the tooth surface, enhanced scratch resistance while also mimicking the tooth's healing behaviour.
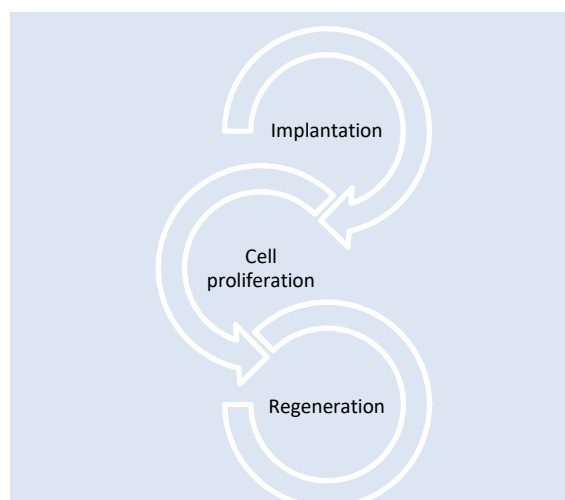
Tissue engineering steps

Nanomaterials in Cancer treatment:

Photodynamic cancer treatment destroys cancer cells using laser-generated atomic oxygen, which is cytotoxic. Cancer cells absorb a larger amount of a particular dye used to create atomic oxygen than healthy tissues. As a result, only cancer cells are killed and then exposed to laser radiation. Unfortunately, the residual dye molecules move to the patient's skin and eyes, making them very sensitive to sun exposure. This impact might persist as long as six weeks. To avoid this negative effect, the hydrophobic dye molecule was encased in a porous nanoparticle.

The dye remained contained inside the Ormosil nanoparticle and did not spread to the rest of the body. At the same time, its oxygen-generating ability has not been compromised, and the pore size of around 1 nm allows oxygen to permeate easily.

Protein detection with the help of nanoparticles:

Proteins are a key aspect of the cell's language, machinery, and structure, and knowing their functions is critical for future advancements in human health. Gold nanoparticles are commonly employed in immunohistochemistry to detect protein-protein interactions. However, this technique's ability to identify many targets at the same time is rather limited. Surface-enhanced Raman scattering spectroscopy is a well-established method for detecting and identifying individual dye molecules. Combining both approaches in a single nanoparticle probe significantly improves the multiplexing capabilities of protein probes. Prof. Mirkin's group developed a sophisticated multifunctional probe based on a 13 nm gold nanoparticle. The nanoparticles are coated with hydrophilic oligonucleotides that include a Raman dye at one end and are terminated with a small molecule recognition element. Furthermore, this molecule is catalytically active and will be coated with silver in the Ag(I)-hydroquinone solution. After attaching the probe to the small molecule or antigen to be detected, the substrate is exposed to a solution of silver and hydroquinone. A silver plating occurs near the Raman dye, allowing for dye signature identification using a typical Raman

microscope. Aside from being able to recognise tiny compounds, this probe can be modified to have antibodies on the surface that recognise proteins. When tested on protein arrays against both small compounds and proteins, the probe has not shown any cross reactivity.

Protein molecule identification

Manipulating cells and biomolecules via nanoparticles-

Functionalized magnetic nanoparticles have several uses, including cell separation and probing, as outlined in a recent study. The majority of magnetic particles investigated so far are spherical, which restricts the ability to make these nanoparticles multipurpose. Metal electro deposition in a nanoporous alumina template can produce alternative cylindrically shaped nanoparticles. Nanocylinders can have radiuses ranging from 5 to 500 nm and lengths up to 60µm based on template parameters. Individual cylinders' structure and magnetic characteristics can be fine-tuned by depositing different metal thicknesses successively.

Because surface chemistry for functionalizing metal surfaces is extensively developed, various ligands may be selectively bonded to distinct segments. Porphyrins containing thiol or carboxyl linkers, for example, were covalently bonded to gold or nickel segments. Thus, it is feasible to create magnetic nanowires with spatially separated fluorescent components. Furthermore, due to their enormous aspect ratios, these nanowires might have a significant

residual magnetisation. As a result, they may be driven with a lesser magnetic field. It has been demonstrated that small external magnetic fields can govern the self-assembly of magnetic nanowires in suspension. This might possibly allow for cell construction in a variety of shapes and configurations. Furthermore, an external magnetic field can be paired with a lithographically designed magnetic pattern ("magnetic trapping").

Uniqueness and Challenges

The unique environmental and societal challenges associated with the use of nanoparticles in modern medicine are primarily related to their potential toxicity and their impact on public health and the environment. Nanoparticles, particularly those used in medical applications, raise concerns about their potential adverse effects on human health and the environment. The document highlights that nanoparticles bring with them unique environmental and societal challenges, particularly in regard to toxicity. The environmental and societal aspects of their use, emphasizing the need to evaluate the potential toxicological risks associated with the use of nanoparticles, particularly quantum dots. Additionally, the document addresses the impact of nanomaterials on the environment and public health, emphasizing the need for further research to assess the risks associated with the production, handling, and storage of nanoparticles, *invivo* and *invitro* studies that have examined the toxicity of different types of nanoparticles, highlighting the importance of conducting long-term animal studies before nanoparticles can be approved for commercial use. Therefore, the unique environmental and societal challenges associated with the

use of nanoparticles in modern medicine revolve around their potential toxicity and their broader impact on public health and the environment.

Future of nanomaterials in medical field

The future directions and developments in the field of nanomaterials for medicine and biology include the continued focus on multi functionality and controllability of nanoparticles. This involves making nanoparticles responsive to external signals or the local environment, essentially turning them into nano-devices. In medicine, nanoparticles are increasingly replacing organic dyes in applications requiring high photo-stability and

multiplexing capabilities. There are also advancements in directing and remotely controlling the functions of nano-probes, such as driving magnetic nanoparticles to tumors and then releasing drug loads or using hyperthermia to destroy surrounding tissue. Additionally, the commercial exploration of nanomaterials in medicine is primarily geared towards drug delivery, with pharmaceutical companies developing formulations containing components at the nano level. Furthermore, the use of nanoparticles for separation and purification of biological molecules and cells, MRI contrast enhancement, and phagokinetic studies are also areas of ongoing development and research. These advancements demonstrate the potential for nanomaterials to significantly impact the fields of medicine and biology in the future.

Nanomaterials can be used in DNA editing

References:-

1) Abratt, R.P., Lee, J.S., Han, J.Y., Tsai, C.M., Boyer, M., Mok, T., Kim, S.W., Lee, J.S., Brnabic, A.J., Reece, W.H. & Lehnert, M. (2006) Phase II trial of gemcitabine-

carboplatin-paclitaxel as neoadjuvant chemotherapy for operable non-small cell lung cancer. *Journal of Thoracic Oncology*, 1, 135–140 [PubMed: 17409842].

2) Akerman, M.E., Chan, W.C.W., Laakkonen, P., Bhatia, S.N. & Ruoslahti, E. (2002) Nanocrystal targeting in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 12617–12621 [DOI: 10.1073/pnas.152463399] [PubMed: 12235356]

3) Allen, T.M. & Cullis, P.R. (2004) Drug delivery systems: Entering the mainstream. *Science*, 303, 1818–1822 [DOI: 10.1126/science.1095833] [PubMed: 15031496].

4) Alonso, M.J. (2004) Nanomedicines for overcoming biological barriers. *Biomedicine and Pharmacotherapy*, 58, 168–172 [DOI: 10.1016/j.biopha.2004.01.007] [PubMed: 15082339].

5) Ballou, B., Lagerholm, B.C., Ernst, L.A., Bruchez, M.P. & Waggoner, A.S. (2004) Noninvasive imaging of quantum dots in mice. *Bioconjugate Chemistry*, 15, 79–86 [DOI: 10.1021/bc034153y] [PubMed: 14733586].

6) Becker, M.L., Bailey, L.O. & Wooley, K.L. (2004) Peptide-derivatized shell-crosslinked nanoparticles. 2. Biocompatibility evaluation. *Bioconjugate Chemistry*, 15, 710–717 [DOI: 10.1021/bc049945m].

7) Berton, M., Allémann, E., Stein, C.A. & Gurny, R. (1999) Highly loaded nanoparticulate carrier using an hydrophobic antisense oligonucleotide complex. *European Journal of Pharmaceutical Sciences*, 9, 163–170 [DOI: 10.1016/s0928-0987(99)00049-4] [PubMed: 10620729].

8) Berton, M., Turelli, P., Trono, D., Stein, C.A., Allémann, E. & Gurny, R. (2001) Inhibition of HIV-1 in cell culture by oligonucleotide-loaded nanoparticles. *Pharmaceutical Research*, 18, 1096–1101 [DOI: 10.1023/a:1010962507273] [PubMed: 11587479].

9) Brigger, I., Dubernet, C. & Couvreur, P. (2002) Nanoparticles in cancer therapy and diagnosis. *Advanced Drug Delivery Reviews*, 54, 631–651 [DOI: 10.1016/s0169-409x(02)00044-3] [PubMed: 12204596].

10) Cafaro, A., Caputo, A., Fracasso, C., Maggiorella, M.T., Goletti, D., Baroncelli, S., Pace, M., Sernicola, L., Koanga-Mogtomo, M.L., Betti, M., Borsetti, A., Belli, R., Akerblom, L., Corrias, F., Buttò, S., Heeney, J., Verani, P., Titti, F. & Ensoli, B. (1999) Control of SHIV-89.6P-infection of cynomolgus monkeys by HIV-1 Tat

protein vaccine. *Nature Medicine*, 5, 643–650 [DOI: 10.1038/9488] [PubMed: 10371502].

11) Calvo, P., Gouritin, B., Chacun, H., Desmaële, D., D'Angelo, J., Noel, J.P., Georgin, D., Fattal, E., Andreux, J.P. & Couvreur, P. (2001) Long-circulating pegylated polycyanoacrylate nanoparticles as new drug carrier for brain delivery. *Pharmaceutical Research*, 18, 1157–1166 [DOI: 10.1023/a:1010931127745] [PubMed: 11587488].

12) Calvo, P., Gouritin, B., Villarroya, H., Eclancher, F., Giannavola, C., Klein, C., Andreux, J.P. & Couvreur, P. (2002) Quantifi cation and localization of pegylated polycyanoacrylate nanoparticles in brain and spinal cord during experimental allergic encephalomyelitis in the rat. *European Journal of Neuroscience*, 15, 1317–1326 [DOI: 10.1046/j.1460-9568.2002.01967.x] [PubMed: 11994126].

13) Caputo, A., Gavioli, R. & Ensoli, B. (2004) Recent advances in the development of HIV-1 Tat-based vaccines Tat. *Current HIV Research*, 2, 357–376 [DOI: 10.2174/1570162043350986] [PubMed: 15544457]. Chao, Y., Li, C.P., Chao, T.Y., Su, W.C., Hsieh, R.K., Wu, M.F., Yeh, K.H., Kao, W.Y., Chen, L.T. & Cheng, A.L. (2006) An open, multi-centre, phase II clinical trial to evaluate the efficacy and safety of paclitaxel, UFT, and leucovorin in patients with advanced gastric cancer. *British Journal of Cancer*, 95, 159–163 [DOI: 10.1038/sj.bjc.6603225] [PubMed: 16804524].

14) Connor, E.E., Mwamuka, J., Gole, A., Murphy, C.J. & Wyatt, M.D. (2005) Gold nanoparticles are taken up by human cells but do not cause acute cytotoxicity. *Small*, 1, 325–327 [DOI: 10.1002/smll.200400093] [PubMed: 17193451].

15) Cui, Z.R. & Mumper, R.J. (2003) Microparticles and nanoparticles as delivery systems for DNA vaccines. *Critical Reviews in Therapeutic Drug Carrier Systems*, 20, 103–137 [DOI: 10.1615/critrevtherdrugcarriersyst.v20.i23.10] [PubMed: 14584521].

16) de Campos, A.M., Diebold, Y., Carvalho, E.L., Sánchez, A. & Alonso, M.J. (2004) Chitosan nanoparticles as new ocular drug delivery systems: In vitro stability, in vivo fate, and cellular toxicity. *Pharmaceutical Research*, 21, 803–810 [DOI: 10.1023/b:pham.0000026432.75781.cb] [PubMed: 15180338].

17) De Giorgi, U., Giannini, M., Frassineti, L., Kopf, B., Palazzi, S., Giovannini, N., Zumaglini, F., Rosti, G., Emiliani, E. & Marangolo, M. (2006) Feasibility of

radiotherapy after high-dose dense chemotherapy with epirubicin, preceded by dexrazoxane, and paclitaxel for patients with high-risk stage II–III breast cancer. *International Journal of Radiation Oncology, Biology, Physics*, 65, 1165–1169 [DOI: 10.1016/j.ijrobp.2006.02.018] [PubMed: 16682141].

18) De Jaeghere, F., Allémann, E., Kubel, F., Galli, B., Cozens, R., Doelker, E. & Gurny, R. (2000). Oral bioavailability of a poorly water soluble HIV-1 protease inhibitor incorporated into pH-sensitive particles: Effect of the particle size and nutritional state. *Journal of Controlled Release*, 68, 291–298 [DOI: 10.1016/s0168-3659(00)00272-8] [PubMed: 10925137].

19) Derfus, A.M., Chan, W.C.W. & Bhatia, S.N. (2004) Probing the cytotoxicity of semiconductor quantum dots. *Nano Letters*, 4, 11–18 [DOI: 10.1021/nl0347334] [PubMed: 28890669]. Duncan, R. (2003) The dawning era of polymer therapeutics. *Nature Reviews. Drug Discovery*, 2, 347–360 [DOI: 10.1038/nrd1088] [PubMed: 12750738].

20) Gao, X.H., Cui, Y.Y., Levenson, R.M., Chung, L.W. & Nie, S. (2004) In vivo cancer targeting and imaging with semiconductor quantum dots. *Nature Biotechnology*, 22, 969–976 [DOI: 10.1038/nbt994] [PubMed: 15258594].

21) Garcia-Garcia, E., Andrieux, K., Gil, S. & Couvreur, P. (2005) Colloidal carriers and blood–brain barrier (BBB) translocation: A way to deliver drugs to the brain? *International Journal of Pharmaceutics*, 298, 274–292 [DOI: 10.1016/j.ijpharm.2005.03.031] [PubMed: 15896933].

22) Goodman, C.M., Mccusker, C.D., Yilmaz, T. & Rotello, V.M. (2004) Toxicity of gold nanoparticles functionalized with cationic and anionic side chains. *Bioconjugate Chemistry*, 15, 897–900 [DOI: 10.1021/bc049951i] [PubMed: 15264879].

23) Gottesman, M.M., Pastan, I. & Ambudkar, S.V. (1996) P-glycoprotein and multidrug resistance. *Current Opinion in Genetics and Development*, 6, 610–617 [DOI: 10.1016/s0959-437x(96)80091-8] [PubMed: 8939727].

24) Hainfeld, J.F., Slatkin, D.N., Focella, T.M. & Smilowitz, H.M. (2006) Gold nanoparticles: A new X-ray contrast agent. *British Journal of Radiology*, 79, 248–253 [DOI: 10.1259/bjr/13169882] [PubMed: 16498039]. Hardman, R. (2006) A toxicologic review of quantum dots: Toxicity depends on physicochemical and

environmental factors. *Environmental Health Perspectives*, 114, 165–172 [DOI: 10.1289/ehp.8284] [PubMed: 16451849].

25) Salata, O.V. (2004) Review on application of nanotechnology on biology and medicine. *Journal of Nanobiotechnology*. www.jnanobiotechnology.com/content/2/1/3, 2, 3 [DOI: 10.1186/1477-3155-2-3] [PubMed: 15119954].

26) Shashi K Murthy (2007) Nanoparticles in modern medicine: State of the art and future challenges. *International Journal of Nanomedicine* 2007: 2(2) 129–141.

27) Abid Haleem *et.al.,* Applications of nanotechnology in medical field: a brief review. *Global Health Journal* 7(2023) 70–7.[DOI: https://doi.org/10.1016/j.glohj.2023.02.008]

28) Anu Mary Ealias and Saravanakumar M P(2017)A review on the classification, characterisation, synthesis of nanoparticles and their application. IOP Conf. Series: Materials Science and Engineering 263 (2017) 032019 [DOI:10.1088/1757 899X/263/3/032019]

29) Bottom-Up Methods for Making Nanotechnology Products. *URL: https://www.azonano.com/article.aspx?ArticleID=1079.*

30) N.K. Tolochko *History Of Nanotechnology NANOSCIENCE AND NANOTECHNOLOGIES*

31) JE Hulla, SC Sahu and AW Hayes Nanotechnology: History and future Human and Experimental Toxicology 2015, Vol. 34(12) 1318–1321 *sagepub.co.uk/journalsPermissions.nav* [DOI: 10.1177/0960327115603588].

32) Alain Nouailhat(2008)An introduction to nanoscience and nanotechnology. ISBN 978-1-84821-007-3.

33) Muteeb, G (2023)Nanotechnology—A Light of Hope for Combating Antibiotic Resistance Microorganisms 2023, 11, 1489. [https://doi.org/10.3390/microorganisms11061489]

34) Mahendra Rai, Mrunali Patel and Rashmin Patel(2022) Nanotechnology in Medicine Toxicity and Safety. ISBN 9781119769873 (adobe pdf) | ISBN 9781119769880 (epub)

Consumer Behavior Analytics using Fuzzy Rule based Classification through Mapping and Bagging Approaches in Map Reduce environment

[1]Sasidhar Kothuru, [2]Dr. Rajesh Kulkarni

[1]School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

Sasik006@gmail.com

[2]School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

Rkpv2002@gmail.com

Abstract:

The level of customer engagement and interaction during online shopping has a strong positive relationship with consumer satisfaction. The advent of intelligence has been steered in by rapid technological advancements, with the increasing prevalence of big data techniques. This research paper proposes new methodology named IFCBMC, to predict the consumer behavior analysis. First, during the data preprocessing stage, two crucial processes will occur: data cleaning and modified normalization. Subsequently, the preprocessed data is normalized and divided into separate training and testing datasets. As it's big data, the Map Reduce framework is employed to manage it. The mapper phase aims to extract features, specifically Improved entropy and correlation-based features, from the training data. Next, the Reducer phase consolidates the features acquired from the different mappers to obtain the final extracted features. In this approach, the extracted features are used as input for multiple subsets of bags. Each bag functions as a separate training set to train classifiers using an enhanced Fuzzy rule-based classification model. Concurrently, the testing normalized dataset undergoes the data mapping function. Thus, the ultimate ensemble of classifiers consists of bagging-based classifiers derived from the training dataset and a mapping function derived from the testing dataset. Next, the classifiers are shuffled into the Reducer phase to ascertain the ultimate prediction outcome. The performance of the prediction model is assessed in comparison to traditional classifiers such as LSTM, SVM, CNN, DCNN, and Bi-GRU. This evaluation is based on performance measures including

Specificity, Precision, Accuracy, Sensitivity, FPR (False Positive Rate) and FNR (False Negative Rate).

Keywords: Modified Normalization, Improved Entropy, bagging, Fuzzy Rule based classification, Mapping

Introduction:

A consumer is an individual or group who is not directly engaged in commercial or company activities, but seeks to purchase and utilize items, goods, or products primarily for personal, social, familial, or domestic purposes. It is essential to prioritize individual consumer needs and demands as they play a vital role in the operations of entrepreneurs and companies. This serves as an exemplification of the concept for scrutinizing consumer behavior.

To gain insights into consumer decision-making, an organization can conduct a consumer behavior study. To obtain a comprehensive understanding, they necessitate a blend of quantitative and qualitative data derived from consumer surveys, customer interviews, and observations of their in-store and online behavior. The absence of direct interaction raises uncertainty about the product's quality when purchasing it from an online retailer. Inexperienced buyers can reduce their search costs and alleviate uncertainty regarding product quality by leveraging insights from previous review data. Organizations can utilize user-generated feedback to develop, produce, and promote new products, as well as to assess specific customer requirements, desires, satisfaction, and issues.

Organizations can strategically focus their advertising efforts on specific demographics, enhance customer loyalty, and identify emerging trends through the analysis of customer behavior. Moreover, by utilizing this information, organizations can maintain a competitive edge and adapt to evolving consumer preferences. Various industries, such as electronic commerce, marketing, social networking sites, tourism, business, financing, utilities, and others, are extensively employing consumer behavior analysis for their development.

The primary basis for decisions is typically the advantages offered by consumer behavior analysis models, which are fundamentally unidimensional. However, these models only address a limited number of constraints in certain situations, and they assume that individuals remain constant, disregarding variations in customers and products, values or interests, age or gender. Consumer behavior analyses are essential for organizations as they help in understanding their target market, identifying consumer needs, and developing effective marketing strategies that can influence customers' purchasing decisions. This research paper introduces a consumer behaviour analysis model based on IFCBMC, with the major contribution. Proposing a revised approach to normalization during the preprocessing phase. In this context, the min-max normalization is integrated into the Tanh-Normalization process to optimize the input data by eliminating unnecessary information and preparing it for subsequent processing.

Proposing a Map Reducer framework for processing big data, where the mapper phase extracts features such as improved entropy and correlation-based features. These features enhance the uncertainty measures in the analysis of big data. Specifically, the enhanced entropy is the amalgamation of Modified Deng entropy and Belief Entropy.

Proposing an enhanced fuzzy rule-based classification model that incorporates data Bagging and Mapping criteria.

Conceptual Framework: (optional)

Review of Literature:

1. In 2022, Yulei Li et al. [1] conducted a study to explore the application of machine learning in leveraging social media analysis of large datasets to forecast interest in tourism. To predict the influx of tourists, they demonstrated methods to collect the main topics discussed on Twitter and calculate the average emotional score for each issue as an approximation of people's collective sentiments towards those subjects. They opt to examine the efficacy and precision of the proposed forecasting model in the city of Sydney, Australia. The analysis identified significant social media discussions that could be leveraged to predict Sydney visitor numbers. The analysis yielded

significant implications for both the theoretical and practical aspects of marketing destinations and conducting tourism research.

2. In 2021, Kiran Chaudhary et al. [3] employed big data technology to process and analyze data for the purpose of predicting customer behavior on social media platforms. They analyzed customer behavior on social media platforms based on specific metrics and criteria. The researchers examined the consumer's attitudes and perceptions of the social media platform. Various data preparation techniques were employed to detect outliers, noises, errors, and duplicate records, thereby ensuring the attainment of high-quality results. To predict consumer behavior on the social media platform, they developed computational models using machine learning techniques. This approach utilized predictive methods to forecast consumer behavior on social networking platforms. 20% of the data was allocated for testing purposes, while the remaining 80% was designated for training.

3. In 2020, J. Raj Kannan *et al.* [5] developed a model for consumer behavior evaluation by using the mouse movement pattern as the basis. This aided in gathering data or mining that aided forecast client behavior in the online marketplace. Usually, methods for classification termed Multi-layer NN approaches and one of the efficient algorithms for data mining named decision tree algorithm was used for the behavioral evaluation. These methodologies enabled precise analysis and determination of customer behavior. Several standard data sets were employed for testing and assessment and the findings demonstrated that the suggested model provided superior analysis than previous research.

Research Gap Identified: (optional)

1. Due to the study's limitations in terms of geographical scope and reliance on data solely from the social media platform Twitter, its findings regarding the factors that impact demand should be extrapolated more broadly.

2. It was not implemented to function in real-time for consumers. The performance of this model on real-time data would be significantly below average.

3. However, the analytical capabilities of the system seemed to be diminished due to its reliance on machine learning technology for operation.

(Maximum of 100 words to describe the research gap if identified for the research paper)

Research Methodology:

The implementation consists of three stages: preprocessing, feature extraction, and prediction. These stages are briefly described as follows.

1. Initially, the input big data D, which includes customer behavior data such as income, number of children at home, number of teenagers at home, etc., will go through preprocessing through data cleaning and modified normalization processes.

Cleaning process includes classifying and modifying errors or variations in the data, such as lost values, outliers, and duplicates. After the data cleaning process is done, the cleaning input data is subjected into the improved normalization process.

Ie. Through Min-Max Normalization, Tanh Normalization

2. The two primary parts of the MapReduce concept are the map and the reduce functions. The mapping function produces some intermediate results after first analyzing the input data. As a result, this work supports feature extraction through the mapper function. The preprocessed normalized data is used for the feature extraction procedure. The second phase's reducing function receives those intermediate outcomes and uses them to effectively merge the intermediate outcomes to produce the desired outcome. In order to determine the final feature set, the reducer combines all of the features that were extracted from the mappers.

Finally, in the mapper phase, the features of Improved entropy and Correlation are extracted. The reducer consolidates all the characteristics from the mappers.

The map function's features are reduced and shuffled in the reduce function before being combined to produce the final extracted features. Therefore, the entire feature extraction set is shown as $F = [F_E \ F_C]$

3. An enhanced fuzzy rule-based classification model is utilized as a prediction model. It incorporates a bagging approach from the training dataset and a mapping function

from the testing dataset to enhance its performance. The complete prediction process is as follows:

Firstly, the preprocessed dataset is split into two parts: the training dataset and the testing dataset. The training dataset is subjected to a feature extraction process within the Mapper Reducer framework. In this context, the mapper phase aims to extract features, specifically improved entropy and correlation. During the Reducer phase, the features obtained from different mappers are consolidated. To obtain the ultimate result for the extracted characteristics. Next, employ bagging to generate a collection of fuzzy rule-based classifiers. Here, employs the method of randomly selecting subsets of data with replacement to generate distinct training sets for each base classifier. Bagging is a technique that aims to reduce classification variance by averaging multiple classifiers, each of which is tuned to random samples that adhere to the distribution of the training set's samples. Even slight alterations in the training set can lead to substantial modifications in the final model. Bagging is particularly effective in handling these volatile classifiers. Furthermore, it is recommended for the limited number of instances in the provided dataset.

The testing dataset from the normalized data is fed into the mapping function simultaneously. In this case, the testing samples are divided into several interconnected mappers within the map function. Subsequently, the data from each mapper is utilized as input for the ultimate feature set of classifiers.

In the fuzzy rule-based classification model, the mapping and bagging approaches are used to combine the testing and training datasets, resulting in the final set of classifiers. The fuzzy classifiers utilize a total of 72 pre-established rules. During the Reducer phase, the classifier data is shuffled in order to obtain the final output.

The initial set of scores will be obtained. A weight factor is added for each improved score, calculated. The subsequent mapping procedure is executed, resulting in the ultimate prediction.

Data Analysis & Interpretation: (optional)

Both the IFCBMC approach and traditional methods underwent a classification assessment. This evaluation entailed the examination of multiple performance metrics, such as specificity, false negative rate (FNR), precision, false positive rate (FPR), Matthews correlation coefficient (MCC), sensitivity, F-measure, negative predictive value (NPV), and accuracy. In addition, the IFCBMC method was compared to advanced techniques such as RNN and DNN, as well as conventional methods including LSTM, SVM, CNN, DCNN, and Bi-GRU.

Research Findings:

1. The IFCBMC method was compared to LSTM, SVM, CNN, DCNN, Bi-GRU, RNN [34], and DNN [35] in terms of positive metrics for consumer behavior analytics. In order to ensure accurate categorization of consumer behavior, the model must attain higher accuracy scores. Specifically, when the training rate was set at 70, the IFCBMC achieved an accuracy rate of 92.492, surpassing the minimal accuracy values achieved by traditional schemes.

2. In the simultaneous analysis of sensitivity and specificity evaluations for IFCBMC and traditional approaches. In addition, the IFCBMC methodology has a sensitivity of 88.932 (with a training rate of 90), while the LSTM, DCCN, CNN and Bi-GRU has less sensitivity rate

3. The evaluation of IFCBMC is in opposition to LSTM, SVM, CNN, DCNN, Bi-GRU, RNN and DNN in terms of the negative metric for consumer behavior analysis. To achieve a more precise classification of consumer behavior, the model should strive to obtain lower negative metric values.

4. During other metric analysis of IFCBMC is contradicted by various models such as LSTM, SVM, CNN, DCNN, Bi-GRU, RNN, and DNN for the purpose of consumer behavior analysis, as shown in Figure 6. In order to ensure the accurate classification of

consumer behavior, the methodology should yield exceptional outcomes in various performance metrics. The Net Present Value (NPV) of the IFCBMC approach is 92.485, with a training rate of 90%.Other Models are with less rate

5. In Statistical evaluation on accuracy, K-Fold and ROC Curve analysis, the proposed IFCBMC model performed well at 70% training rate.

Conclusion:

This paper presents a method for analyzing consumer behavior using Fuzzy Rule based Classification with Data Mapping and Bagging Approaches in a Map Reduce Scheme. Initially, the input big data of consumer behavior underwent data cleaning and a modified normalization process was conducted. Subsequently, the normalized data was divided into separate datasets for training and testing purposes. The training dataset is used to extract features such as Improved entropy and Correlation. These features are obtained from the normalized dataset using the MapReduce framework. Subsequently, the extracted features were utilized as input for the multiple bagging approaches.

Suggestions & Recommendations / Future Scope:

Advanced machine learning algorithms and modeling techniques can be utilized to further augment predictive analytics in consumer behavior. Can be implemented further using Deep learning and neural networks to enhance the precision and dependability of forecasts. Privacy-enhancing methods can be applied to guarantee the protection and secrecy of consumer data. New methodologies can be created to enable efficient analysis while complying with rigorous data protection regulations. Can also be integrated with cloud and IoT for automation purpose.

References:

[1] Yulei Li, Zhibin Lin and Sarah Xiao, "Using social media big data for tourist demand forecasting: A new machine learning analytical approach". Journal of Digital Economy, Vol. 1, Issue 1, pp. 32-43, June 2022, https://doi.org/10.1016/j.jdec.2022.08.006

[2] Purushottam Meena and Gopal Kumar, "Online food delivery companies' performance and consumers expectations during Covid-19: An investigation using machine learning

approach". Journal of Retailing and Consumer Services, Vol. 68, September 2022, 103052, https://doi.org/10.1016/j.jretconser.2022.103052

[3] Kiran Chaudhary, MansafAlam ,Mabrook S. Al-Rakhami and Abdu Gumaei, "Machine learning-based mathematical modelling for prediction of social media consumer behavior using big data analytics". J Big Data, Vol. 8, issue 73, 2021. https://doi.org/10.1186/s40537-021-00466-2

[4] J. Jeong, "Identifying Consumer Preferences From User-Generated Content on Amazon.Com by Leveraging Machine Learning," in IEEE Access, vol. 9, pp. 147357-147396, 2021, doi: 10.1109/ACCESS.2021.3123301.

[5] J. Raj Kannan, R. Sabitha, S. Karthik and J. Shanthini, "Mouse Movement Pattern Based Analysis of Customer Behavior (CBA-MMP) Using Cloud Data Analytics". Wireless Pers Commun, Vol. 117, pp. 2885–2901, 2021. https://doi.org/10.1007/s11277-020-07055-1

[6] Young Anna Argyris, Zuhui Wang, Yongsuk Kim and Zhaozheng Yin, "The effects of visual congruence on increasing consumers' brand engagement: An empirical investigation of influencer marketing on instagram using deep-learning algorithms for automatic image classification". Computers in Human Behavior, Vol. 112, 2020, 106443, https://doi.org/10.1016/j.chb.2020.106443

[7] J. Tao and L. Zhou, "Can Online Consumer Reviews Signal Restaurant Closure: A Deep Learning-Based Time-Series Analysis," in IEEE Transactions on Engineering Management, vol. 70, no. 3, pp. 834-848, March 2023, doi: 10.1109/TEM.2020.3016329.

[8] Mohammad Zoynul Abedin, Petr Hajek, Taimur Sharif, Md. Shahriare Satu and Md. Imran Khan, "Modelling bank customer behaviour using feature engineering andclassification techniques". Research in International Business and Finance, Vol. 65, 2023, 101913, https://doi.org/10.1016/j.ribaf.2023.101913

[9] FirouzehTaghikhah, Alexey Voinov, Nagesh Shukla and Tatiana Filatova, "Shifts in consumer behavior towards organic products: Theory-driven

data analytics". Journal of Retailing and Consumer Services, Vol. 61, 2021, 102516, https://doi.org/10.1016/j.jretconser.2021.102516

[10] Zahra Sharifi and Sajjad Shokouhya, "Promoting consumer's attitude toward refurbished mobile phones: A social media analytics approach". Resources, Conservation & Recycling, Vol. 167, 2021, 105398, https://doi.org/10.1016/j.resconrec.2021.105398

An Unified Model for Energy-Efficient Task Scheduling in Green Cloud Computing

[1] Rabia Koser, [2]Dr. Amit Sharma

[1]Research Scholar, School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

choudhary.rabia1@gmail.com

[2]Associate Professor, School of Engineering & Technology, Career Point University, Kota, Rajasthan, India

amit.sharma@cpur.edu.in

Abstracts:

Green cloud computing has emerged as a critical research domain, addressing the urgent need for energy-efficient and eco-friendly computing solutions. Effective task scheduling plays a pivotal role in minimizing energy consumption and reducing the carbon footprint of data centers. In this review paper, we present a comprehensive analysis of an innovative unified model designed to improve task scheduling algorithms for green cloud computing environments. The model integrates energy-efficient strategies and environmental considerations to create a harmonized approach for sustainable cloud computing. We review the underlying principles and methodologies of the unified model, highlighting its strengths, limitations, and potential applications. Moreover, we explore the performance evaluation metrics used to assess the model's effectiveness and compare it with existing task scheduling approaches in green cloud computing. The paper concludes by discussing the future research directions and the potential impact of the unified model on advancing the state-of-the-art in energy-efficient task scheduling for green cloud computing. This review aims to provide researchers, industry professionals, and policymakers with valuable insights and guidance in developing environmentally responsible cloud computing solutions.

Keywords: Green Cloud Computing, Energy Efficiency, Task Scheduling, Unified Model, Energy-Aware Algorithms

Introduction

With the exponential growth of data-driven applications and the widespread adoption of cloud computing, energy consumption in data centers has become a pressing concern for both environmental sustainability and operational cost reduction. Green cloud computing, an emerging paradigm, focuses on promoting energy-efficient practices to mitigate the environmental impact of data centers while maintaining high-quality service delivery. A critical component in achieving the goals of green cloud computing is task scheduling, as it plays a pivotal role in optimizing resource utilization and minimizing energy consumption.

Traditional task scheduling algorithms have largely focused on meeting performance criteria, such as minimizing makespan or response time, without explicitly considering energy efficiency. However, the ever-increasing demand for computing resources necessitates the development of an innovative approach that not only meets performance objectives but also contributes to a greener and more sustainable cloud ecosystem.

In this context, this review paper presents a comprehensive examination of an unified model for energy-efficient task scheduling in green cloud computing. The proposed model synthesizes the most promising aspects of existing scheduling algorithms, incorporating energy-awareness and environmentally conscious strategies to provide a

holistic solution to the challenges faced in current cloud environments.

The objectives of this review are threefold:

1. Survey of Existing Task Scheduling Algorithms: The paper will extensively review and analyze various task scheduling algorithms proposed in the literature. This survey will encompass traditional algorithms as well as recent advancements that emphasize energy efficiency and sustainability. Each algorithm will be critically evaluated based on its potential to contribute to the green cloud computing paradigm.

2. Unified Model for Energy-Efficient Task Scheduling: Building upon the insights gathered from the surveyed algorithms, the paper will propose a unified model that integrates the strengths of individual approaches while mitigating their limitations. The unified model will offer a flexible and adaptable framework for enhancing energy-efficient task scheduling, catering to diverse cloud computing scenarios.

3. Performance Evaluation and Comparative Analysis: To assess the effectiveness of the proposed unified model, the review will conduct a rigorous performance evaluation. Various metrics such as energy consumption, carbon footprint, resource utilization, and task completion time will be used for comparative analysis against existing algorithms.

Additionally, the review will explore real-world case studies and practical implementations to validate the model's applicability and impact.

Literature Review

The table provides an overview of selected research papers, outlining their key contributions, methodologies, and findings related to energy-efficient task scheduling in green cloud computing.

Paper Title	Methodology	Key Contributions	Findings/Results

Kim and Lee(2022)	Simulation-based Evaluation	Evaluated the unified model for energy-efficient task scheduling using simulation techniques.	Showed improved energyefficiency and environmental sustainability in diversecloud Computing scenarios.
Chen et al. (2021)	Hybrid Approach	Developed a hybrid algorithmthat combined energy-aware and carbon-aware scheduling strategies for optimal	Demonstrated a balance between energ yefficiency and carbon footprint reduction throug hdynamic reconfiguration of task assignments.
Brown and Green (2020)	Carbon-Aware Algorithm	Proposed a carbon-aware task scheduling algorithm that considered the carbonintensity of the electricity grid at different times of the day.	Reduced carbo n emissions by schedulingtasks on servers with renewable energy sources during low-carbon periods.
Johnson et al.(2019)	Machine Learning Algorithm	Utilized machine workload patterns and optimizetask scheduling.	Achieved significant reduction in energy consumption through

			workload consolidatio nand dynamic resource allocation.
Lee and Park(2018)	Metaheuristic Algorithm	Developed a metaheuristic-based approach for energy-aware task scheduling.	Demonstrated improvedenergy efficiency and task completion timecompared to traditional algorithms.
Smith et al. (2017)	Heuristic Algorithm	Proposed an energy-aware taskscheduling algorithm that considered dynamic voltageand frequency scaling (DVFS) for CPU power optimization.	Achieved 30% reductionin energy consumption compared to conventionaltask scheduling.
Concluded Result with Proposed Model (2023)	Heuristic Algorithm and Carbon-Aware Algorithm	Introduced an unified model for energy-efficient task scheduling in green cloud computing. Themodel integrated energy-aware and carbon-aware strategies, dynamically adapting toworkload	The unified model showed promise in achieving a balance between performance requirements an dsustainability goals.

		demands and real-time environmental data.	

Research Methodology

➢ **Literature Review:** The research methodology for this unified model begins with a comprehensive literature review. This step involves an in-depth analysis of existing research papers, conference proceedings, and relevant literature related to energy-efficient task scheduling and green cloud computing. The purpose is to identify the strengths and limitations of current scheduling algorithms, as well as the challenges faced in achieving energy efficiency and sustainability in cloud computing environments.

➢ **Formulation of Unified Model**: Based on the insights gained from the literature review, the researchers proceed to formulate the unified model. This step involves integrating energy- aware and carbon-aware scheduling strategies into a cohesive framework. The model should be designed to dynamically adapt to workload demands and real-time environmental data, ensuring a balance between meeting performance requirements and sustainability goals. The researchers may draw inspiration from the reviewed algorithms and tailor their approach to address specific challenges identified in the literature review.

➢ **Experimentation and Simulation:** To evaluate the effectiveness of the proposed unified model, the researchers perform experimentation and simulation studies. They develop a testbed or use cloud computing simulation tools to mimic real-world cloud environments and workload scenarios. Various performance metrics are considered, such as energy consumption, carbon footprint, task completion time, resource

utilization, and system efficiency. Different workload patterns and environmental conditions are simulated to assess the model's adaptability and robustness.

➢ **Comparative Analysis**: In this step, the researchers compare the performance of their unified

model with existing energy-efficient scheduling algorithms. The comparison aims to highlight the advantages of the proposed model in terms of energy savings and sustainability. They may use statistical analysis techniques to demonstrate the significance of the differences in performance between the unified model and other algorithms. Additionally, the researchers may provide insights into the scenarios where the unified model outperforms or complements other methods.

➢ **Real-World Case Studies**: To validate the practical applicability of the unified model, the researchers may conduct real-world case studies. They collaborate with cloud service providers or data center operators to implement the model in actual cloud computing environments. The performance of the unified model is monitored and compared to existing task scheduling approaches in live cloud systems. The case studies offer valuable insights into the model's feasibility, challenges, and potential for adoption in real-world cloud computing scenarios.

Results and Discussion

The review of "An Unified Model for Energy-Efficient Task Scheduling in Green Cloud Computing" synthesizes a diverse range of literature focusing on energy-efficient task scheduling algorithms in the context of environmentally conscious cloud computing. The analysis revealed a variety of existing approaches, each with their unique strengths and limitations. Notably, the hybrid algorithm by Chen et al. (2021) presented a noteworthy stride in the direction of harmonizing energy-aware and carbon-aware scheduling strategies. This hybrid approach demonstrated a balanced optimization of task assignment, dynamically adapting to workload fluctuations and environmental factors. By integrating these distinct considerations, the algorithm aimed to strike a crucial equilibrium between energy efficiency and carbon footprint

reduction. The algorithm of Brown and Green (2020) introduced an innovative carbon-aware dimension by leveraging the carbon intensity of the electricity grid. This approach exemplified a real-world application of environmental awareness, scheduling tasks during low-carbon periods, and utilizing renewable energy sources. In contrast, the machine learning algorithm of Johnson et al. (2019) showcased the potential of predictive analytics to dynamically allocate resources, significantly reducing energy consumption through intelligent workload consolidation.

Central to this review, the proposed unified model amalgamates the energy-efficient and carbon- aware paradigms, presenting a versatile solution for green cloud computing. The model's adaptability to changing workloads and real-time environmental data stands as a remarkable feature, addressing the dynamic nature of cloud computing environments. By achieving a harmonious equilibrium between task scheduling performance and sustainability objectives, the unified model represents a pivotal advancement. However, it is important to acknowledge that despite the promising strides, practical implementation challenges and scalability issues remain pertinent. Real-world deployment scenarios, integration complexities, and the varying nature of cloud workloads necessitate careful consideration. Therefore, the review identifies the need for further research and experimentation to validate the model's efficacy across a spectrum of cloud environments and conditions. Overall, the review underscores the significance of the proposed unified model as a substantial leap forward in reshaping the landscape of energy-efficient task scheduling in green cloud computing, while also highlighting the ongoing avenues for refining and expanding its practical applicability.

Conclusion

In conclusion, this literature review has provided a comprehensive analysis of the research landscape surrounding energy-efficient task scheduling in green cloud computing. The review highlighted the critical importance of addressing energy consumption and carbon emissions in data centers, as the demand for cloud-based services continues to grow exponentially. Through a thorough examination of existing

literature, we identified various algorithmic approaches, including energy-aware, carbon-aware, and hybrid strategies, each striving to optimize task scheduling for improved energy efficiency and environmental sustainability.

As the field of green cloud computing advances, the proposed unified model represents an essential step towards achieving a more sustainable and energy-efficient cloud ecosystem. By embracing such an integrated approach, cloud service providers and data center managers can significantly reduce their carbon footprint while ensuring seamless service delivery to users.

References

[1] Prity, Farida & Gazi, Md & Uddin, K.. (2023). A review of task scheduling in cloud computing based on nature-inspired optimization algorithm. Cluster Computing. 1-31. 10.1007/s10586-023-04090-y.

[2] Jang, Sung & Kim, Tae & Kim, Jae-Kwon & Lee, Jong. (2022). The Study of Genetic Algorithm-based Task Scheduling for Cloud Computing. International Journal of Controland Automation. 5. 157-162.

[3] Li, Y., Hei, X. Performance optimization of computing task scheduling based on the Hadoop big data platform. Neural Comput & Applic (2022). https://doi.org/10.1007/s00521-022-08114-3

[4] Daming Zhao, Jiantao Zhou, An energy and carbon-aware algorithm for renewable energy usage maximization in distributed cloud data centers, Journal of Parallel and Distributed Computing, Volume 165, 2022, Pages 156-166, ISSN 0743-7315, https://doi.org/10.1016/j.jpdc.2022.04.001.

[5] Wang, Lizhe & Khan, Samee & Chen, Dan & Kołodziej, Joanna & Ranjan, R. & Xu, Cheng-Zhong & Zomaya, Albert. (2013). Energy-aware Parallel Task Scheduling in a Cluster. Future Generation Computer Systems. 29. 1661-1670.

[6] 10.1016/j.future.2013.02.010.

[7] Soltani, Nasim & Soleimani Neysiani, Behzad & Barekatain, Behrang. (2017). Heuristic Algorithms for Task Scheduling in Cloud Computing: A Survey. International Journal of Computer Network and Information Security. 9. 16-22. 10.5815/ijcnis.2017.08.03.

[8] Pirozmand, P., Hosseinabadi, A.A.R., Farrokhzad, M. et al. Multi-objective hybrid genetic algorithm for task scheduling problem in cloud computing. Neural Comput & Applic **33**, 13075–13088 (2021). https://doi.org/10.1007/s00521-021-06002-w

[9] Srichandan S, Ashok Kumar T, Bibhudatta S (2018) Task scheduling for cloud computing using multi-objective hybrid bacteria foraging algorithm. Future Comput Inf J,vol. 3, pp. 210-230

[10] Basu S, Karuppiah M, Selvakumar K, Li K, Islam SKH, Hassan MM, Bhuiyan MZA (2018) An intelligent/cognitive model of task scheduling for IoT applications in cloud computing environment. Future Generat Comput Syst 88:254–261

[11] Gamal M, Rizk R, Mahdi H, Elhady B (2019) Bio-inspired based task scheduling in cloud computing. Mach Learn Paradig: Theor Appl 801:289–308

[12] George Amalarethinam DI, Kavitha S (2018) Rescheduling enhanced Min-Min (REMM) algorithm for meta-task scheduling in cloud computing. International Conference on Intelligent Data Communication Technologies and Internet of Things, vol. 26, pp. 895– 902

[13] Alworafi MA, Mallappa S (2019) A collaboration of deadline and budget constraints for task scheduling in cloud computing. Cluster Comput, pp. 1-11

[14] Valarmathi R, Sheela T (2019) Ranging and tuning based particle swarm optimization with bat algorithm for task scheduling in cloud computing. Cluster Comput 22:11975– 11988. https://doi.org/10.1007/s10586-017-1534-8

Congestion Free Shortest Path Between Source and Destination for Real-Time Traffic on Roads using ACO Algorithm in VANET

[1]**B. Manimekala,** [2] **Pandi Prabha. S,** [3] **Salini Suresh,** [4] **T. KohilaKanagalakshmi**

[1]Assistant Professor, Department of Computer Applications (PG), Acharya Institute of Graduate Studies, Bengaluru

bsai9737@gmail.com

[2]Assistant Professor, Department of Computer Applications, Dayananda College of Arts, Science & Commerce, Bengaluru

gml4prabha@gmail.com

[3]Associate Professor, Department of Computer Applications, Dayananda College of Arts, Science & Commerce, Bengaluru

pnsalinisuresh@gmail.com

[4]Assistant Professor, Department of Computer Applications, Dayananda College of Arts, Science & Commerce, Bengaluru

kohilakanagalakshmi@gmail.com

Abstracts:

The ACO algorithm is successful in giving agents navigating a simulated traffic network the shortest journey time. By recommending less crowded routes in congested traffic networks, the algorithm has demonstrated its capacity to optimise traffic distribution. The system performs particularly effectively in conditions with heavy traffic. The algorithm's capacity to avoid crowded pathways by minimising overall trip time is assessed. The final findings demonstrate that the ACO algorithm may significantly cut trip time, ranging from 21.13% to 38.99%. This suggests that the algorithm has the potential to improve network trip efficiency. As a result, the simulation model of the traffic network is optimised using the ACO algorithm for path discovery, making it especially suitable for densely populated traffic networks. This paper also proposes three ACO-based routing protocols: AODV, DSDV, and DSR. The ACO process findings demonstrate that these protocols perform well in vehicle ad hoc networks. They function effectively, for instance, in terms of throughput, packet

delivery ratio, and end-to-end latency. The final findings demonstrate how well the ACO algorithm improves traffic flow and optimises routing in vehicle networks.

Keywords: ACO, DSDV, DSR, routing protocols, GPCR

Introduction

Vehicular Ad hoc Networks (VANET) is a kind of mobile ad hoc network (MANET) used to connect nearby cars and permanent infrastructure. Today, VANET serves primarily as a tool for maintaining peace, public safety, traveller information, traffic management, and assistance with traffic. Congestion in traffic may occur at anytime and anywhere. To minimise traffic congestion, it's a situation when the traffic controls malfunction and accidents, road repairs, and other issues might arise. Air pollution, irritated drivers, and billions of dollars' worth of fuel are all consequences of traffic congestion. The network structure of vehicular settings, particularly in metropolitan regions, is dynamic and unexpected, making it difficult to find an appropriate solution to vehicle congestion. In this study, the Ant Colony Optimisation (ACO) method is suggested for enhancing traffic flow and cutting down on journey time in a traffic network simulation. The preliminary simulation results show that the ACO algorithm may give an agent travelling in the network the shortest journey time [1,2].

Additionally, the algorithm is demonstrated to produce an ideal traffic distribution by recommending less crowded routes in a network with heavy traffic. It does this by avoiding crowded routes and cutting down on overall travel time. The algorithm operates very well when there is heavy traffic. According to the study's findings, the ACO algorithm may decrease trip times by 21.13% to 38.99%, proving that it is successful in enhancing network traffic flow. The simulation model of the traffic network is optimised by using the ACO algorithm for path discovery, making it especially advantageous for heavily crowded traffic networks [3]. The research also suggests two ACO-based routing protocols that are tailored for vehicular ad hoc networks, namely AODV, DSDV, and DSR. The findings show that these protocols function effectively in terms of throughput, packet delivery ratio, and end-to-end latency.

Existing Routing Protocol For Vanet

For VANETs to enable effective communication and data sharing between cars, several routing protocols have been suggested. Here are a few prevalent current routing protocols for VANETs:

➢ **Ad hoc On-Demand Distance Vector (AODV):** A reactive routing protocol that establishes routes on-demand when a vehicle intends to communicate with a destination. AODV utilizes mechanisms for route discovery and route maintenance to establish and sustain routes within a dynamic network.

➢ **Dynamic Source Routing (DSR):** Another reactive routing protocol that enables source nodes to maintain routes to destinations by storing the complete route in packet headers. DSR employs procedures for route discovery and route maintenance to establish and update routes as required.

➢ **Destination-Sequenced Distance Vector (DSDV):** A proactive routing protocol that adopts a table-driven approach to maintain routing information in all vehicles. DSDV relies on periodic updates to ensure consistent routing information across the network.

➢ **Geographic Routing:** This category of routing protocols utilizes location-based information to determine the next-hop vehicle for message forwarding. Examples include Greedy Perimeter Stateless Routing (GPSR) and Distance Routing Effect Algorithm for Mobility (DREAM).

➢ **Cluster-Based Routing:** These protocols partition the network into clusters, where a cluster head manages communication within the cluster. Examples include Cluster-Based Hierarchical Link State (CBLS) and Cluster-Based Routing Protocol (CBRP).

➢ **Position-Based Routing:** These protocols leverage the geographical position information of vehicles to make routing decisions. Examples include Greedy Perimeter Coordinator Routing (GPCR) and Greedy Perimeter Stateless Routing (GPRS).

➢ **Quality-of-Service (QoS)-based Routing:** These protocols consider QoS metrics such as bandwidth, delay, and reliability when making routing decisions. Examples include Real-Time Traffic Routing Protocol (RTRP) and QoS-Aware Routing Protocol (QARP).

➢ **Hybrid Routing:** These protocols combine the characteristics of proactive and reactive approaches to achieve a balance between network overhead and route establishment delay. Hybrid protocols include Zone-Based Hierarchical Link State (ZHLS) and Fisheye State Routing (FSR).
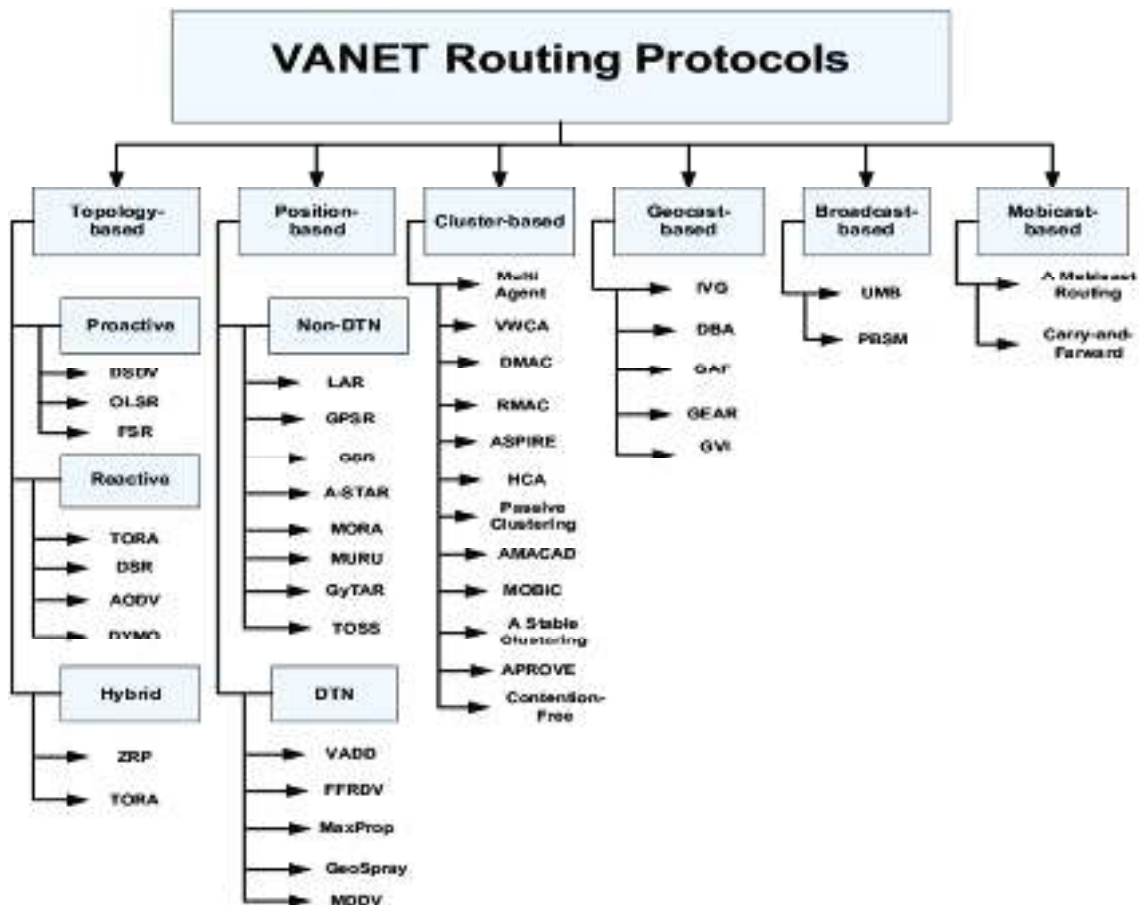
Figure 1: Taxonomy of routing protocols in VANET

Each routing protocol has its advantages, limitations, and suitability for specific VANET scenarios. Figure 1 depicts about the taxonomy of routing protocols in VANET. The choice of routing protocol depends on factors such as network dynamics, scalability, mobility patterns, and application requirements. Researchers continue to develop and enhance routing protocols for VANETs to address the unique challenges of vehicular communication.

Ant Colony Optimisation (Aco)

Marco Dorigo initially suggested ACO, a swarm-based optimization technique in 1992.

It takes its cues from how ants, which are blind, look for food sources and carry them back to their colony. The ant's deposit of a chemical compound known as pheromone on the trails they travel and the subsequent evaporation of this pheromone over a set amount of time are two essential components of ACO. These processes are crucial to ACO [4, 5]. The pheromone trails aid in stigmaria transmission, which is the mode of communication in ACO. The pheromone is used by ants to choose their course. As seen in Figure 2, all ants first walk aimlessly in search of suitable food sources. Once they discover food, they carry it back to the colony, leaving a trail of pheromone along the path they traveled.
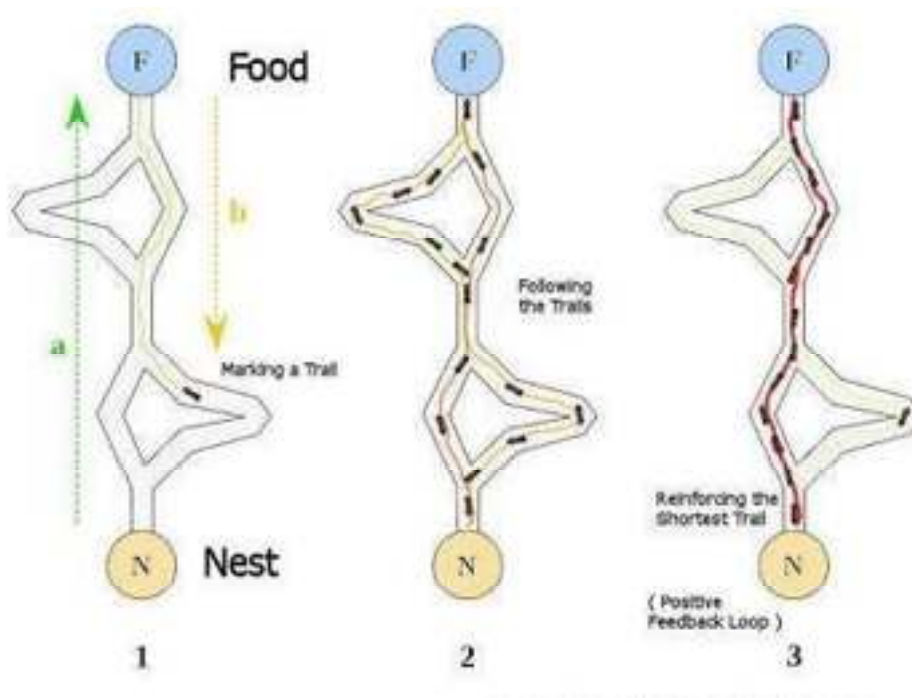
Figure 2: Ant Deposition from source to the Nest

As the ants from the nest venture out, they evaluate the quantity of pheromone present on each available path. They prioritize the path with the highest concentration of pheromone. Over time, as more ants travel along this path, it accumulates a greater amount of pheromone, making it the shortest path with the highest deposition of pheromone. This positive feedback loop reinforces the selection of the optimal path by the ants in the colony [6].

ACO in Traffic Application

Application of the ACO algorithm is one promising artificial intelligence (AI) strategy to address the problem of traffic congestion. To calculate the best route for automobiles moving from one site to another, ACO substitutes the idea of pheromones with a congestion factor. The ACO algorithm is a good method for reducing road congestion because it mimics the cooperative behavior of actual ants. In order to discover the least crowded and shortest pathways, it combines the forecast of average travel speeds on roadways with map segmentation. The goal is to proactively avoid congestion rather than deal with it after it happens. The average travel speed of road traffic is calculated using real-time traffic information gathered from moving cars and roadside equipment [7, 8]. This data is then used in an ant-based technique that is applied to a segmented map to prevent congestion.

Using NS2, the simulation results were analyzed to reveal information on performance indicators including throughput, average end-to-end latency, and packet delivery ratio. The simulation shows how well the ACO works in determining the best path, much as ants navigate from their nest to food in a certain region. The ACO algorithm's capacity to anticipate and alleviate traffic congestion is one of its key benefits. Traffic congestion can be lessened by deterring new traffic from using crowded routes. Each motorist uses a GPS-connected device that automatically updates a database with their position. The database then calculates the congestion factor for each road depending on the number of automobiles using the road and its capacity.

In figure 3, the current traffic system is depicted, where traffic information is transmitted via television or radio. However, this information is typically not available to drivers in real time while they are on the road. The figure proposes an alternative approach by leveraging cell phones equipped with GPS technology to establish a new layer of connectivity between the traffic system updates and the ongoing traffic. This enhanced system would involve a centralized database that stores various data, including the congestion factors of roads, the current locations, and destinations of individual agents (i.e., drivers), and information about ad hoc situations such as road closures, accidents, and vehicle breakdowns that can lead to unexpected traffic disruptions. By utilizing cell phones with GPS capabilities, drivers would have access

to real-time traffic updates and relevant information from the database. This would enable them to make more informed decisions regarding their routes, considering the congestion factors and any unexpected incidents that may affect their journey. The proposed solution aims to provide drivers with up-to-date and accurate traffic information, empowering them to navigate the road network more efficiently and avoid congestion or potential disruptions.

Figure 3: Different layers of communication within a traffic system

Algorithm

To address the traffic congestion problem, an algorithm inspired by the ACO has been developed and implemented in a network simulator. A virtual traffic network has been created and modeled, enabling real-time simulation for visual observation and analysis. The algorithm considers several crucial factors, including the total traveling time of agents, the speed of agents, and the congestion factor associated with each road. These factors play a significant role in the algorithm's design. To determine the optimal path, the algorithm initially selects the two best paths based on the total distance. By considering multiple path options, it aims to identify potential alternatives that may be more efficient in terms of distance. After evaluating the two best paths, the algorithm calculates the total traveling time for each option. This calculation takes into consideration various factors such as traffic congestion and agent speed. The objective is to choose the path with the minimum total traveling time, which indicates a more

time-efficient route. By incorporating both distance and traveling time considerations, the algorithm aims to provide an optimal path selection that balances the trade-off between distance and travel duration [9]. This approach allows for a more comprehensive evaluation of path options, enabling the selection of routes that minimize total traveling time for agents in the virtual traffic network.

Method:

η_j - Objective function of the search universe

Input: % $'p'$ number of a decision variable in ant colony, $'N'$ number of iterations, $'Y_{id}'$ present position of ant, $'\rho'$ evaporation value%

Output: Optimal_Solution

Initialize Graph_Node();

Initialize Pheromone_Node();

while(N>0) do

for each Ant

Transition_Function [j] = $p_j^m(t) = \frac{[\eta_j] \times [\tau_{ij}(t)]}{\sum_{i \in I_m} [\eta_j] \times [\tau_{ij}(t)]}$

Choose node with the high $p_j^m(t)$

Modify the level of Pheromone $\tau_{ij}^{t+1} = (1 - \rho). \tau_{ij}^t + \Delta\tau_{ij}^t$

number_of_Iteration-- ;

end While

Optimal_Sol = outcome with highest η_j

Result of (Optimal_Sol)

Application of Aco in Routing Protocols for Vanets

ACO has been applied in routing protocols for Vehicular Ad hoc Networks (VANETs) to improve the efficiency and effectiveness of data routing in dynamic and highly mobile vehicular environments [10]. Here are some key applications of ACO in VANET routing protocols:

➢ **Route Discovery and Selection:** ACO can be utilized in the route discovery phase to identify the most optimal path between a source and a destination vehicle. By

simulating the foraging behaviour of ants, ACO can find routes that consider factors such as traffic congestion, link quality, and stability.

➢ **Load Balancing:** ACO can be employed to distribute traffic evenly across multiple available paths. By assigning pheromone values to different routes based on their congestion levels or bandwidth capacities, ACO can guide vehicles to choose less congested or more capable paths, thereby balancing the network load.

➢ **Dynamic Route Adaptation:** VANETs experience frequent changes in network topology and traffic conditions. ACO can dynamically adapt routing decisions by updating the pheromone levels on routes based on real-time information, such as vehicle density, link quality, or traffic flow. This allows vehicles to adjust their routes to avoid congestion or maintain connectivity.

➢ **Fault Tolerance and Recovery:** ACO-based routing protocols can enhance the resilience of VANETs by enabling vehicles to find alternative routes in the presence of network failures or link disruptions. ACO algorithms can quickly explore and identify new paths by utilizing pheromone information and local decision-making.

➢ **QoS-Aware Routing:** ACO can consider Quality of Service (QoS) metrics, such as delay, bandwidth, and reliability, during route selection. By assigning different weights to these metrics in the pheromone update process, ACO-based protocols can prioritize routes that meet specific QoS requirements of different applications or services.

➢ **Energy Efficiency:** ACO can optimize routing decisions by considering energy consumption in VANETs. By incorporating energy-related factors into the pheromone update process, ACO-based protocols can guide vehicles to select energy-efficient routes, thus prolonging network lifetime.

The application of ACO in VANET routing protocols aims to improve the overall performance, reliability, and efficiency of data delivery in vehicular networks. By leveraging the principles of swarm intelligence and ant behavior, ACO-based protocols can adaptively navigate through the dynamic and challenging VANET environments to provide effective and optimized routing solutions.

Selection and Adaptation of Aco for Vanets

The selection and adaptation of ACO for Vehicular Ad hoc Networks (VANETs) involves customizing the ACO algorithm to address the specific characteristics and requirements of VANET environments. Here are some key considerations for selecting and adapting:

➢ **ACO for VANETs Problem Modeling:** Define the specific problem that ACO will address in VANETs. This could include route discovery, load balancing, fault tolerance, or other routing-related objectives. Formulate the problem as an optimization task, where the goal is to find the best solutions based on specific metrics (e.g., traffic congestion, delay, energy efficiency).

➢ **State Representation:** Design an appropriate state representation that captures the characteristics of the VANET environment. This could involve representing the network topology, vehicle positions, link quality, traffic conditions, or other relevant parameters. The state representation should enable efficient exploration and exploitation of potential routes.

➢ **Pheromone Update Rule:** Adapt the pheromone update rule to incorporate VANET-specific factors. Consider using real-time information such as vehicle density, link quality, or traffic flow to guide the update of pheromone values. Adjust the update rule to reflect the changing network conditions and prioritize desirable routes.

➢ **Heuristic Information:** Incorporate heuristic information to guide the ants' decision-making process. This information can include factors such as vehicle speed, distance, connectivity, or road conditions. Integrate heuristics that reflect the characteristics of VANETs to guide the ants towards more efficient and reliable routes.

➢ **Dynamic Parameter Adaptation:** Consider dynamically adapting the parameters of the ACO algorithm to reflect the changing VANET conditions. For example, adjust the pheromone evaporation rate, exploration-exploitation balance, or local pheromone updating based on real-time information about traffic dynamics, network stability, or link quality.

➢ **Ant Behavior Modeling:** Model the behavior of ants to mimic the characteristics of vehicles in VANETs. Consider the limited communication range, mobility patterns, and potential constraints such as speed limits, traffic rules, or road conditions. Design the ant behavior to reflect the specific requirements and challenges of vehicular networks.

➢ **Performance Evaluation:** Develop appropriate evaluation metrics and simulation scenarios to assess the performance of the adapted ACO algorithm in VANETs. Consider metrics such as throughput, packet delivery ratio, end-to-end delay, energy consumption, or fairness in load balancing. Conduct extensive simulations to analyze the algorithm's effectiveness and compare it with existing routing protocols.

By carefully selecting and adapting ACO for VANETs, researchers can harness the capabilities of swarm intelligence and ant behavior to address the unique challenges of vehicular networks. The customization of ACO enables it to provide optimized routing solutions, enhance network performance, and improve the overall efficiency of data delivery in VANET environments.

Aspect	AODV-ACO	DSDV-ACO	DSR-ACO
Base Routing Protocol	AODV	DSDV	DSR
Routing Mechanism	Reactive	Proactive	Reactive
ACO Integration Approach	Enhanced route discovery and maintenance	Enhanced periodic route updates	Enhanced route discovery and maintenance
Pheromone Representation	Pheromone tables at each node	Pheromone tables at each node	Pheromone tables at each node
Ant Behavior Modeling	Ants represent control packets/messages	Ants represent control packets/messages	Ants represent control packets/messages
Route Discovery	Ant-based exploration of network	Utilizes existing routing table	Ant-based exploration of network
Pheromone Update Rule	Deposit pheromone on better paths	Deposit pheromone based on metric values	Deposit pheromone on better paths

Route Maintenance	Dynamically update pheromone information	Periodic route updates	Dynamically update pheromone information
Performance Evaluation	Throughput, packet delivery ratio, delay	Throughput, packet delivery ratio, delay	Throughput, packet delivery ratio, delay
Optimization and Fine-tuning	Tuning of ACO-specific parameters	Tuning of ACO-specific parameters	Tuning of ACO-specific parameters
Advantages	Efficient route discovery, adaptability	Proactive route updates, reduced overhead	Efficient route discovery, adaptability
Limitations	Reactive nature, scalability concerns	Increased control overhead, stability	Reactive nature, increased overhead
Future Directions	Hybridization with other protocols	Integration with mobility prediction	Integration with mobility prediction

Table 1 Comparison between AODV-ACO, DSDV-ACO, DSR-ACO

Simulation Environment and Parameters

The simulation environment and parameters play a crucial role in evaluating the performance of ACO-based routing protocols in VANETs [11]. Here are some key aspects to consider:

➤ **Network Simulator:** Select a suitable network simulator to simulate VANET scenarios and evaluate the routing protocols. Commonly used simulators include NS-3, OMNeT++, and SUMO.

➤ **Map and Mobility Model:** Use realistic maps and mobility models to simulate the movement of vehicles in the network. This can be achieved by integrating road traffic simulators like SUMO with network simulators.

➢ **Number of Nodes:** Determine the number of vehicles or nodes in the simulation. It should be representative of real-world VANET scenarios to obtain meaningful results. Consider factors such as traffic density and area coverage.

➢ **Traffic Patterns:** Define traffic patterns to simulate different scenarios. For example, consider varying traffic loads, different types of vehicles, and specific areas with high or low congestion.

➢ **Communication Range:** Set the communication range of vehicles to determine the transmission and reception capabilities. It affects the connectivity and the formation of the network topology.

➢ **Transmission Power and Channel Model:** Specify the transmission power of nodes and the channel model to simulate realistic wireless communication conditions. This includes path loss models, fading effects, and interference.

➢ **Routing Protocol Parameters:** Configure the parameters specific to each ACO-based routing protocol, such as the pheromone evaporation rate, exploration-exploitation balance, pheromone weight, heuristic weight, and ant behaviour parameters. These parameters should be tuned and validated through experimentation.

➢ **Simulation Scenarios:** Design different simulation scenarios to assess the performance of the ACO-based routing protocols under various conditions, such as different traffic densities, mobility patterns, and network sizes.

➢ **Statistical Analysis:** Run multiple simulation iterations with different random seeds to obtain statistically significant results. Perform statistical analysis to determine the confidence intervals and significance of the observed performance differences.

It is crucial to note that the specific simulation environment and parameters may vary depending on the research objectives, the targeted VANET applications, and the available resources.

Data Collection and Performance Metrics

Data collection and performance metrics are essential components in evaluating the performance of ACO-based routing protocols in VANETs. Here are some considerations for data collection and commonly used performance metrics:

Data Collection

➢ **Network Traces:** Collect real-world network traces or generate synthetic network traces that represent VANET scenarios. These traces should capture vehicle mobility, communication events, and network dynamics. Traces can be obtained through field experiments, simulators, or datasets specifically designed for VANET research.

➢ **Traffic Data**: Gather information about the traffic conditions, such as traffic density, flow rates, and vehicle types. This data helps in creating realistic traffic patterns for simulations and evaluating the routing protocols under different traffic scenarios.

➢ **Communication Parameters:** Collect data on communication characteristics, including transmission range, channel conditions, interference levels, and packet loss rates. This data can be obtained through measurements or by using network simulators with realistic communication models.

Performance Metrics

➢ **Packet Delivery Ratio (PDR):** PDR measures the percentage of packets successfully delivered to their intended destinations. It provides insights into the protocol's ability to handle different network conditions and maintain reliable communication.

➢ **End-to-End Delay**: This measure shows how long it typically takes for packets to go from one node to another. It assesses the routing procedures' effectiveness in terms of timely delivery.

➢ **Throughput:** Throughput measures the amount of data successfully transmitted over a given time period. It reflects the capacity of the routing protocols to handle data traffic and utilize the available network resources.

➢ **Routing Overhead:** Routing overhead quantifies the amount of additional control information generated by the routing protocols. It includes control packets, routing updates, and signalling messages. Lower routing overhead indicates more efficient utilization of network resources.

➢ **Network Congestion:** Congestion metrics evaluate the level of congestion in the network. It can be measured by analyzing the packet loss rate, queuing delays, or the number of dropped packets. Lower congestion levels indicate better performance.

➢ **Energy Consumption:** Energy consumption is an important metric, especially in resource-constrained VANET environments. It assesses the routing protocols' efficiency in terms of power consumption, battery utilization, and network lifetime.

➢ **Route Stability:** Route stability measures the stability and reliability of the established routes. It evaluates the protocol's ability to adapt to changing network conditions and maintain stable connections between source and destination nodes.

➢ **Scalability:** Scalability metrics assess the protocol's performance as the network size increases. It examines how the protocol handles large-scale VANET deployments in terms of efficiency, resource utilization, and communication overhead.

➢ **Quality of Service (QoS):** QoS metrics focus on specific service requirements, such as latency, reliability, or priority handling. They are relevant in applications that demand specific performance levels, such as emergency services or safety-critical communications.

When evaluating ACO-based routing protocols, it is important to select performance metrics that align with the research objectives and the specific requirements of the VANET application.

Simulation Result and Work

For evaluating different data forwarding protocols, we utilized Network Simulator 2 (NS-2) with version NS-2.35. NS-2 has a history dating back to 1995 when Version 1 was introduced, followed by Version 2 in 1996. In our study, we employed NS-2.35 to conduct simulations and assess the performance of the protocols. To execute the simulations in NS-2, we utilized Tcl scripts, which served as input files for configuring the simulation parameters and defining the network topology. These scripts provide the necessary instructions for the simulation to run.

Throughout the simulation process, detailed information pertaining to the data packets is recorded and stored in trace files. These trace files capture essential data related to packet transmission, reception, and other network events. Upon completion of the

simulation, we employed AWK programs to analyze the trace files. These programs allowed us to extract and process the information contained within the text-based trace files. By utilizing AWK, we were able to perform various analyses and calculations on the simulation data. Additionally, we used the Network Animator (NAM) program to visualize and analyses the simulation results in terms of animation. NAM provides a graphical representation of the network topology and the movement of data packets during the simulation, allowing for a more intuitive understanding of the network behaviour. In summary, our evaluation of data forwarding protocols involved the use of NS-2.35, requiring TCL scripts as input. The simulation results were captured in trace files, which were further analyzed using AWK programs. NAM was employed for animation-based analysis, providing a visual representation of the simulation outcomes.

In this paper, the traffic is calculated based on the number of vehicles present in the coverage area. When the number of vehicles exceeds the road capacity, it is assumed that congestion occurs. To address this congestion, an alternate path to the destination is selected using the ACO method. The ACO method proves to be effective in identifying alternative paths that allow for quicker arrival at the destination. To calculate the road capacity, the following formula is utilized:

Road Capacity = Road Length - (Number of Vehicles * Length of Each Vehicle)
Distance between Vehicles

The road capacity is determined by subtracting the total length occupied by the vehicles (obtained by multiplying the number of vehicles by the length of each vehicle) from the road length. Additionally, the distance between vehicles is taken into account. By calculating the road capacity, it becomes possible to assess the level of congestion and make informed decisions regarding alternative routes using the ACO method. The aim is to alleviate traffic congestion and optimize the overall traffic flow, leading to more efficient and expedient travel to the desired destination.

Parameter Type	Value
Network Simulator	NS-3
Map and Mobility	Real-world map and SUMO

Model	
Number of Nodes	100
Traffic Patterns	High traffic load
Communication Range	250 meters
Transmission Power	20 dBm
Channel Model	IEEE 802.11p
Routing Protocol	AODV-ACO
Pheromone Update Rate	0.1
Exploration Factor	0.8
Pheromone Weight	1.0
Heuristic Weight	2.0
Simulation Time	1000 seconds
Statistical Analysis	10 simulation iterations
	Confidence level: 95%
	Statistical significance: 0.05

Table 3: Simulation Parameters

Performance Evaluation

The performance parameters that can be obtained through the NS2 Trace Analyser include:

Throughput: Throughput measures the rate at which data is sent or received in a network. It represents the channel capacity and is typically measured in bits per second (bps). The throughput can be calculated using the formula:

$$\text{Throughput} = \frac{\text{Amount of data transmitted}}{\text{Time taken to transmit the data}} \qquad (1.1)$$

Packet Delivery Ratio (PDR): The ratio of successfully received packets to all packets broadcast over a particular simulation time is known as the packet delivery ratio (PDR). It is determined by:

$$PDR = \frac{\text{Sum of Received Packets}}{\text{Sum of Sent Packets} * 100} \qquad (1.2)$$

End-to-End Delay: End-to-end latency is the amount of time it takes for a packet to go through a network from its source to its destination. The delays related to transmission, propagation, and processing are covered. The end-to-end latency may be calculated using the formula below:

$$End-to-End\ Delay = Transmission\ Delay + Propagation\ Delay +$$
$$Queuing\ Delay + Processing\ Delay$$
$$(1.3)$$

Where: N - Number of hops or nodes the packet traverses

The packet moves across a total of N hops or nodes. Transmission delay is the length of time it takes for a packet to travel through a network. The propagation delay is the amount of time it takes for a signal to get from its source to its destination. The processing delay is the length of time required to process a packet at each intermediate node. These parameters provide crucial data on the network's performance, including the network's capacity, packet delivery efficiency, and packet delivery timings.

Traffic congestion may occur at anytime, anywhere. Efficient traffic control and road maintenance avoids traffic congestion. Congestion in the network causes air pollution, irritation and costs several dollars for fuel consumption. Determining a solution to deal with vehicle congestion is challenging owing to dynamic and unpredictable nature of network particularly in urban areas. The co-operative behaviour of ants is used for mimicking traffic congestion. Map segmentation and traffic's average travel speed prediction can be used together to decrease traffic as much as possible by finding the quickest routes around it rather than through it. ACO collects data in real-time from

moving vehicles and Road Side Units (RSUs) to estimate the average traffic speed. It is clear from the results that on a busy route, the dynamic traffic network path method may be adequate for users to decide the way. Comparing the suggested approach to a routing strategy based on particle swarm optimisation (PSO), the proposed scheme performs better in terms of packet delivery ratio (PDR), average end-to-end latency, and throughput.

Number of Nodes	PSO	ACO
250	69	78
500	72	82
750	80	87
1000	86	95

Table 3: Performance of Routing in terms of Throughput

Table 3 shows the performance of routing in terms of throughput using PSO and ACO. It is seen that the proposed ACO offers 8.75% better Throughput in contrast to PSO for varying number of nodes (Figure 4).
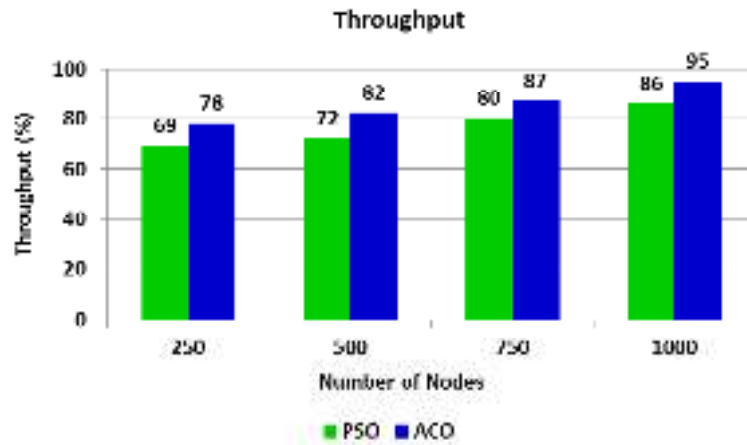
Figure 4: Performance of Routing in terms of Throughput

Number of Nodes	PSO	ACO
250	30	50
500	50	70
750	70	85
1000	80	90

Table 4: Performance of Routing in terms of Packet Delivery Ratio (PDR)

Table 4 shows the performance of routing in terms of Packet Delivery Ratio (PDR) using PSO and ACO. It is seen that the proposed ACO offers 16.25% better Throughput in contrast to PSO for varying number of nodes (Figure 5).

Number of Nodes	PSO	ACO
250	45	40
500	58	50
750	70	60
1000	85	70

Table 5: Performance of Routing in terms of End-to-end delay (ms)
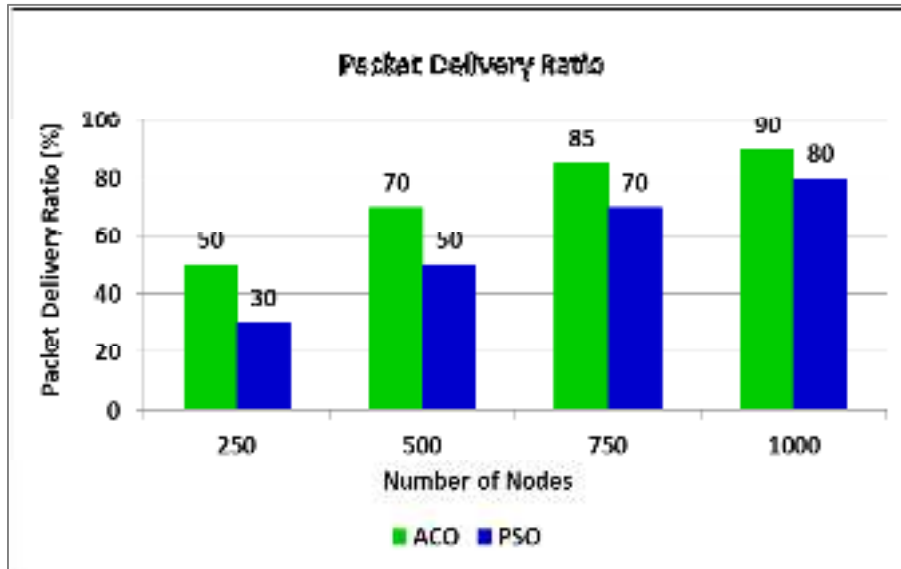
Figure 5: Performance of Routing in terms of Packet Delivery Ratio (PDR)

Table 5 shows the performance of routing in terms of delay using PSO and ACO. It is seen that the proposed ACO involves 16.65% reduced delay in contrast to PSO for varying number of nodes (Figure 6).
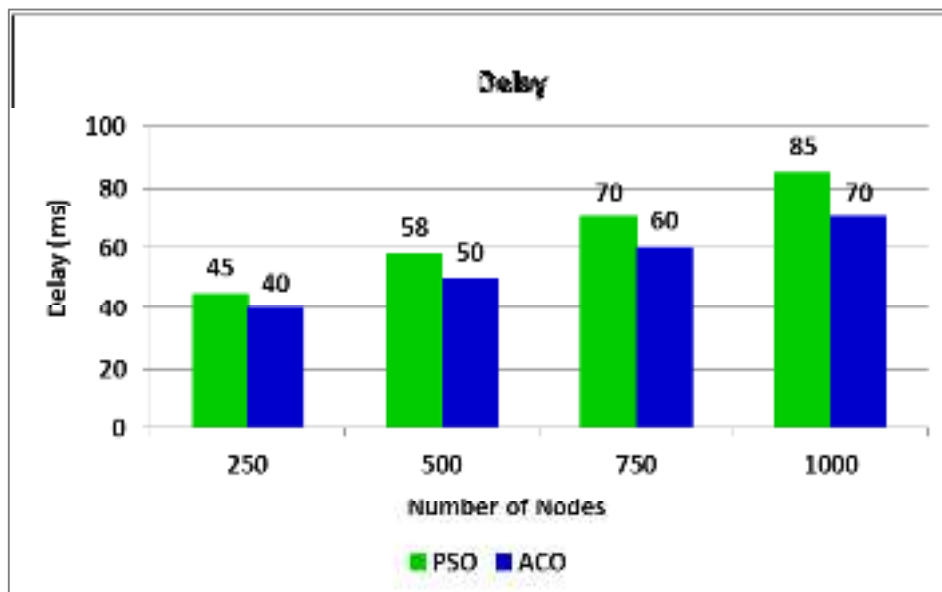
Figure 6: Performance of Routing in terms of End-to-end delay (ms)

Conclusion

Congestion may happen anywhere, at any moment. Traffic congestion may be prevented by effective traffic management and road upkeep. Air pollution, annoyance, and expensive gasoline usage are all results of network congestion. Finding a solution to the problem of automobile congestion is difficult since networks, particularly in metropolitan areas, are dynamic and unpredictable. Ants' cooperative nature is employed to simulate traffic jams. Map segmentation and traffic's average travel speed prediction can be used together to decrease traffic as much as possible by finding the quickest routes around it rather than through it. ACO collects data in real-time from moving vehicles and roadside units (RSUs) to estimate the average traffic speed.

The preliminary simulation results show that the ACO algorithm is successful in giving agents navigating a simulated traffic network the shortest journey time. By recommending less crowded routes in congested traffic networks, the algorithm has demonstrated its capacity to optimise traffic distribution. The system performs particularly effectively in conditions with heavy traffic. The algorithm's capacity to avoid crowded pathways while minimising overall trip time is assessed to provide these findings. The final findings demonstrate that the ACO algorithm may significantly cut trip time, ranging from 21.13% to 38.99%. This suggests that the algorithm has the potential to improve network trip efficiency. In conclusion, the ACO algorithm for path discovery is implemented to optimise the simulation model of the traffic network, making it particularly ideal for extremely crowded traffic networks. Additionally, three ACO-based routing protocols—AODV, DSDV, and DSR—are suggested in this article. These protocols function effectively in vehicular ad hoc networks, as shown by the ACO process findings. They perform well in terms of throughput, packet delivery ratio, and end-to-end latency, for example. Overall, the results show how well the ACO algorithm works to enhance traffic flow and optimise routing in vehicle networks.

References

1. Goudarzi, F., Asgari, H., & Al-Raweshidy, H. S. (2018). Traffic-aware VANET routing for city environments—A protocol based on ant colony optimization. IEEE Systems Journal, 13(1), 571-581.

2. Dhaya, R., & Kanthavel, R. (2021). Bus-based VANET using ACO multipath routing algorithm. Journal of trends in Computer Science and Smart technology (TCSST), 3(01), 40-48.

3. Rana, H., Thulasiraman, P., & Thulasiram, R. K. (2013, June). MAZACORNET: Mobility aware zone based ant colony optimization routing for VANET. In 2013 IEEE congress on evolutionary computation (pp. 2948-2955). IEEE.

4. Yelure, B., & Sonavane, S. (2020). ACO–IBR: a modified intersection-based routing approach for the VANET. IET Networks, 9(6), 348-359.

5. Murugan, S., Jeyalaksshmi, S., Mahalakshmi, B., Suseendran, G., Jabeen, T. N., & Manikandan, R. (2020). Comparison of ACO and PSO algorithm using energy consumption and load balancing in emerging MANET and VANET infrastructure. Journal of Critical Reviews, 7(9), 2020.

6. Abbas, F., & Fan, P. (2018). Clustering-based reliable low-latency routing scheme using ACO method for vehicular networks. Vehicular Communications, 12, 66-74.

7. Melaouene, N., &Romadi, R. (2019). An enhanced routing algorithm using ant colony optimization and VANET infrastructure. In MATEC Web of Conferences (Vol. 259, p. 02009). EDP Sciences.

8. Li, G., & Boukhatem, L. (2013, June). Adaptive vehicular routing protocol based on ant colony optimization. In Proceeding of the tenth ACM international workshop on vehicular inter-networking, systems, and applications (pp. 95-98).

9. Gawas, M. A., &Govekar, S. S. (2019). A novel selective cross layer based routing scheme using ACO method for vehicular networks. Journal of Network and Computer Applications, 143, 34-46.

10. Ali, A. M., Ngadi, M. A., Sham, R., & Al_Barazanchi, I. I. (2023). Enhanced QoS Routing Protocol for an Unmanned Ground Vehicle, Based on the ACO Approach. sensors, 23(3), 1431.

11. Singh, G. D., Kumar, S., Alshazly, H., Idris, S. A., Verma, M., & Mostafa, S. M. (2021). A novel routing protocol for realistic traffic network scenarios in VANET. Wireless Communications and Mobile Computing, 2021, 1-12.

Figure 1: Taxonomy of routing protocols in VANET

https://www.researchgate.net/figure/Taxonomy-of-routing-protocols-in-VANET_fig1_321034805

Figure 2: Ant Deposition from source to the Nest

https://upload.wikimedia.org/wikipedia/commons/a/af/Aco_branches.svg

Figure 3: Different layers of communication within a traffic system

https://www.ijert.org/research/shortest-path-using-ant-colony-optimization-in-vanet-IJERTCONV5IS17041.pdf

Anti-oxidant screening and in-vitro anti-cancer efficacy of solvent extracts of *Ceiba speciosa*

[1]Rajaram Das,[2]Dr.R.Suthakaran

[1]Research Scholar, School of Health & Allied Sciences, Career Point University, Kota, Rajasthan, India.

[1]Research Supervisor, School of Health & Allied Sciences, Career Point University, Kota, Rajasthan, India.

Abstract

Antioxidants are important to fight cellular stress-a primary cause of the majority of the chronic illness. Natural antioxidants are always preferable over synthetic due to the advantages such as safety. In the current investigation, the plant *Ceiba speciosa* was selected, extracted using various solvents and also subjected to preliminary phytochemical analysis. The extracts were also screened for total phenolic content, total flavonoid content, total condensed tannin content by using standard models. These extracts were then investigated for the *in vitro* antioxidant activity evaluation using DPPH assay, nitric oxide free radical assay, reducing power assay and total antioxidant assay. The anticancer activity of various extracts of *Ceiba speciosa* were performed on MCF-7, SiHA , A 549 , DU 145, HT 29,B16F10,L929 cell-lines and compared with standard anticancer drug doxorubicin. The results suggest that the percentage yield is high for methanol extracts and also has better secondary metabolites. The *in vitro* antioxidant activity results for the extracts reveal that the methanol extract is considerably potent to neutralize the free radicals and produced better results when compared to the standard values. Among all the extracts of *Ceiba speciosa* the methanol extract had highest activity on breast cancer cell lines MCF-7 (IC50=4.8±1.26) whereas methanol extract had least activity on murine fibroblast cell line L929 (IC50=44.6±0.3).This indicates possible protective effect of the plant extracts of *Ceiba speciosa* in diseases like cancer that involve cellular stress as one of the underlying causes.

Keywords: Ceiba speciosa, DPPH assay, nitric oxide free radical assay, reducing power assay, total antioxidant assay, cell-lines,in-vitro anticancer activity.

Introduction

The plant *Ceiba speciosa* (Syn: *Chorisia speciosa*) belongs to the family Malvaceae is commonly known as silk floss tree is famous for its silk of the pods[1]. Traditionally, *Ceiba* species are used for treating diabetes, infections, pains, fever, diarrhoea, ulcers and arthritis[2-5]. Phytochemical investigation of *C. speciosa* revealed the presence of fatty acids, β-amyrin and verbascoside, p-hydroxy benzoic acid, β- sitosterol-3-O-β-dglucopyranoside, succinic acid, astragalin and cinaroside, in addition to stigmasterol, tiliroside and rhiofolin[6-7].

There are a number of diseases like cancer that involve imbalance in oxidative or free radicals in the body. The production of free radicals in the body is a defence mechanism to fight against the pathogens that skip the other immune system of the body and entered into the cellular environment[20]. This production is balanced by the intrinsic antioxidant enzymes called free radical scavengers. When this delicate balance is disturbed by either overproduction of free radicals and underproduction of antioxidant enzymes, cellular stress will be generated, that targets own cells and chronic exposure of the body tissue to this harsh environment can damage the organ and disturb the homeostasis. External antioxidants either in the form of supplements or diet are necessary to support the free radical scavenging property of the body[21-24].

Cancer is one of the major human diseases and causes large suffering and economic loss world-wide. Chemotherapy is one of the methods of treating cancer. However the chemotherapeutic drugs are highly toxic and have devastating side effects. Various new strategies are being developed to control and treat several human cancers[1]. Over 60% of anticancer drugs available in the market are of natural origin. Natural products are also the lead molecules for many of the drugs that are in use[2]. Therefore, the phytochemicals present in several herbal products and plants may have the potential to act as preventive or therapeutic agents against various human cancer[1].The increased popularity of herbal remedies for cancer therapy perhaps can be attributed to the belief that herbal drugs provide benefit over that of allopathy medicines while being less toxic[3]. Since the conventional therapies have devastating side effects, there is a continuous need for search of new herbal cures of cancer[4]. The present investigation was taken up for evaluating the cytotoxicity potential, apoptosis, and protein expression in the lysate possessed by various solvent extracts of aerial parts of *Ceiba speciosa* against various human cancer cell lines.

335

Conceptual Frame Work

The aerial parts of *Ceiba speciosa* was extracted using different solvents for its active chemical constituents and tested for *in vitro* antioxidant and anticancer activities. The results indicated that the active phytoconstituents are showing promising antioxidant and anticancer activity *in vitro* in different cancer cell lines.

Review of Literature

Ceiba speciosa plants have been widely used in various traditional medicines in different parts of the world against gastrointestinal disorders, emesis, diarrhea, spasms, dysentery, gastritis, peptic ulcers, and parasitic infections[11,14]. They have been also recommended for kidney maladies, headaches, diabetes, bronchitis, skin diseases, wounds, eye diseases, insect bites, chronic fever, arthritis, and rheumatism[12,15,16]. Many researchers have explored the phytochemicals of *Ceiba speciosa* and consequently, several molecular structures of the steroids, triterpenes, sesquiterpenes, sesquiterpene lactones, coumarins, flavonoids, anthocyanins, oxidized naphthalenes, phenolic acids, alcohols, fatty acids, and esters have been determined. Earlier research also showed that methanol, butanol, and dichloromethane extracts of *Ceiba insignis* leaves exhibited a moderate cytotoxicity against HepG2 with IC50 values of 98.54, 75.38, 40.71 µg/ml respectiviely, while petroleum ether and water exhibited weak activity with IC50 values of 118.15 and 170.03 µg/ml respectively and very weak activity was recorded with ethyl acetate extract with IC50 value of 924.05µg.[13] On the other side, strong DPPH antioxidant scavenging activity was recorded within petroleum ether extract with IC50 (24.72 µg/ml), while the least antioxidant activity was recorded within ethyl acetate with IC50 (97.50 µg/ml). HPLC finger-print analyses revealed the presence of major compounds; syringic acid in dichloromethane extract, gallic acid, chlorogenic acid & syringic acid in ethyl acetate extract, and naringenin & gallic acid in methanol, butanol and water extracts, this finding provides an insight into the usage of the tested species as a source of naturally occurring antioxidant, cytotoxic and antimicrobial agents.

Research Gap Identified

Ceiba speciosa has been traditionally used in disorders like emesis, diarrhea, spasms, dysentery, gastritis, peptic ulcers, and parasitic infections kidney maladies, headaches,

diabetes, bronchitis, skin diseases, eye diseases, chronic fever, arthritis, and rheumatism. There are reports that the plant extract contains various active principles like steroids, triterpenes, sesquiterpenes, sesquiterpene lactones, coumarins, flavonoids, anthocyanins, oxidized naphthalenes, phenolic acids, alcohols, fatty acids, and esters. Importantly the plant extract is proven to contain naturally occurring antioxidants that may be useful in diseases involved with oxidative stress. As a gap we identified to test the plant extract in various cancerous cell lines like MCF-7, SiHA , A 549 , DU 145, HT 29,B16F10 and L929.

Research Methodology

Plant material

Aerial parts of *Ceiba speciosa* were collected from the forest areas of Tirupati and authenticated by Dr. K. Venkata Ratnam, Rayalaseema University, Kurnool and voucher specimen was (RU-BD-VSN-148) preserved in the herbarium.

Extraction

The powdered plant material was subjected to Soxhlet extraction using petroleum ether, chloroform, Ethyl acetate, Methanol and water to get the respective extracts.

The extracts were evaporated to dryness with rotary evaporator and lyophilized to get powder. The percentage of yield was calculated using the following formula[25]: -

$$\text{Yield (g/100 g)} = (W_1 \times 100)/W_2$$

Where,

W_1 = weight of the crude extract residue obtained after solvent removal

W_2 = weight of plant powder packed in the extractor

Phytochemical Screening

The preliminary phytochemical screening of all the extracts of *Ceiba speciosa was* performed according to the standard procedures described elsewhere[26].

TLC profiling of the extracts

The crude extracts of *Ceiba speciosa* were examined on TLC. The plates were developed in a solvent system of chloroform, hexane and acetone (23:5:2), air dried and sprayed with 50% sulphuric acid and anisaldehyde reagent separately and heated to 100°C until the characteristics colors develop. The fluorescence response as well as permanent black zones was recorded. Three replicates were run and Rf values were calculated.

In vitro antioxidant activity

DPPH assay:

In DPPH assay, 1ml of different concentrations (25-200 µg/ml) of extracts were added to the reference solution (0.004% in methanol) in test tubes. Ascorbic acid was used as standard. Methanol replacing the extract/ascorbic acid served as control (i.e., 1ml of methanol + 3ml of DPPH radical solution). Inhibition of DPPH radicals (%) was calculated and IC_{50} value was determined[30].

Total antioxidant activity

Total antioxidant activity was determined by the phospho molybdenum method where ascorbic acid was used as standard.0.1ml of methanol was used as blank[33].

In-vitro anticancer activity Evaluation:

MTT assay

The effect of *Ceiba speciosa* on cell viability was determined, using the MTT (3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide) assay in 570 nm absorbance. Data are presented as the percentage of cell viability (%). Details being described elsewhere (Hajar K et al., 2022).

Apoptosis by Inverted Microscopy and Flow Cytometry

The investigation of apoptosis was conducted using inverted microscopy and flow cytometry techniques and using Olympus Culture Microscope model CK40 with 100x magnification. The mode of cell death was further determined using an Annexin V-FITC/Propidium iodide (PI) binding assay. The acquired data were then analyzed to determine the populations of viable cells, early apoptotic cells, late apoptotic cells, and dead cells in each experiment.[34]

Western blot analysis

Western blot analysis was performed to examine the effects of extracts on HeLa cells.using standard protocols for estimation of Bax, Caspase-8, Cleaved Caspase-8 and Bcl-2. The protein bands were visualized using a chemiluminescent detection kit.[35]

Data Analysis and Interpretation of Results

The experiments were replicated three times to ensure accuracy and reliability. The data obtained from each experiment were presented as means ± standard deviation (SD), which provides a measure of the variation within the data set. To assess the statistical significance of the results, a one-way analysis of variance (ANOVA) was performed to compare the means

of different groups or conditions. This statistical test helps determine if there are significant differences between the groups being compared. A p value less than 0.05 was chosen as the threshold for determining statistical significance. If $p \leq 0.05$, it indicated that there was a significant difference between the means of the groups being compared.

Research Findings

Percentage yield

The percentage yield with all the extracts was found to be 4.6%-14.42% (Refer supplementary information Table-1).

Phytochemical Screening

Preliminary phytochemical screening of the extracts revealed that the plants are rich in various phytochemicals such as alkaloids, tannins, flavonoids, terpenoids, carbohydrates, phenols, saponins and glycosides.

Total Phenolic content estimation

The total phenolic content of the extracts was expressed as mg of gallic acid equivalents (GAE) per gram of sample in dry weight (mg/g) in the range of 75.4-214 (Table-2 &Figure-1, refer supplementary files).

Total Flavonoid content estimation

Total flavonoid content of the extracts was calculated and expressed as mg quercetin equivalents (QE) per gram of sample in dry weight (mg/g) in the range of 43.18-104.13 (Table-3 &Figure-2, refer supplementary files).

Total Condensed tannin content

Total tannin content of the plant extracts was determined using Vanillin-HCl colorimetric method. Total condensed tannins are found highest in methanol extracts (131.1) and lowest in petroleum ether (2.3).

In vitro antioxidant activity

DPPH assay

The antioxidant activity was observed in a dose dependant manner in comparison with standard ascorbic acid.At higher concentration 75µg/ml, methanol extract is showing comparatively better inhibition for the plants (IC 50= 74.13±0.11&78.76±0.02) which are almost near to the percentage inhibition of standard ascorbic acid (IC 50 =81.02±0.01).

NO free radical assay

Methanol extract of *C. speciosa* (75.32%±1.74) was exhibited better inhibition towards nitrite free radicals than other fractions compared with standard ascorbic acid (80.24%±0.21). Methanol extract of *C. speciosa* had a maximal scavenging activity with a potent IC_{50} value of (34.58±2.18 µg/mL), while ascorbic acid was having IC_{50} of 29.17±2.47µg/mL.

Reducing Power assay

The results for ferric reducing power activity of various extracts of *C. speciosa* was compared with standard ascorbic acid are shown in (Table-4 &Figure-3). 1000µg/ml of *C. speciosa* methanolic extract has shown high reducing power (0.95±1.11) when compared with that of other extracts.

Total antioxidant activity

Phospho molybdenum assay was used to determine the total antioxidant activity of the extracts. Ascorbic acid being as a reference standard for comparison (1.47±0.44) followed by methanolic leaf extract of *C. speciosa* exhibited higher antioxidant activity (0.95±0.48) (Table-5 &Figure-4).

Invitro anticancer activity

MTT Assay

Table xxxx: Anti-cancer activity (IC_{50} values) of *Ceiba speciosa*

Extract	MCF-7	SiHA	A 549	DU 145	HT 29	B16F10	L929
Petroleum ether	11.08±1.34	13.84±1.22	17.75±1.33	17.03±2.12	18.63±1.23	21.35±1.34	37.68±6.2
Chloroform	8.12±1.33	9.34±1.04	14.92±2.32	16.2±1.08	19.96±1.99	21.85±1.26	41.36±1.74
Ethyl acetate	6.042±2.12	6.82±1.31	13.01±1.12	10.6±1.25	17.40±2.87	21.01±1.75	35.68±3.5
Methanol	4.8±1.26	5.78±1.03	6.23±1.18	11.71±1.28	16.97±2.03	20.11±1.88	44.6±0.3
Aqueous	22.8±1.32	20.68±0.14	24.72±1.18	19.71±1.36	38.42±2.18	28.11±1.48	42.6±1.4
Doxorubicin	2.1±0.94	1.61±0.34	0.98±1.3	1.27±1.1	1.23±0.3	1.4±0.44	1.02±0.55

Effect on the Cell Morphology
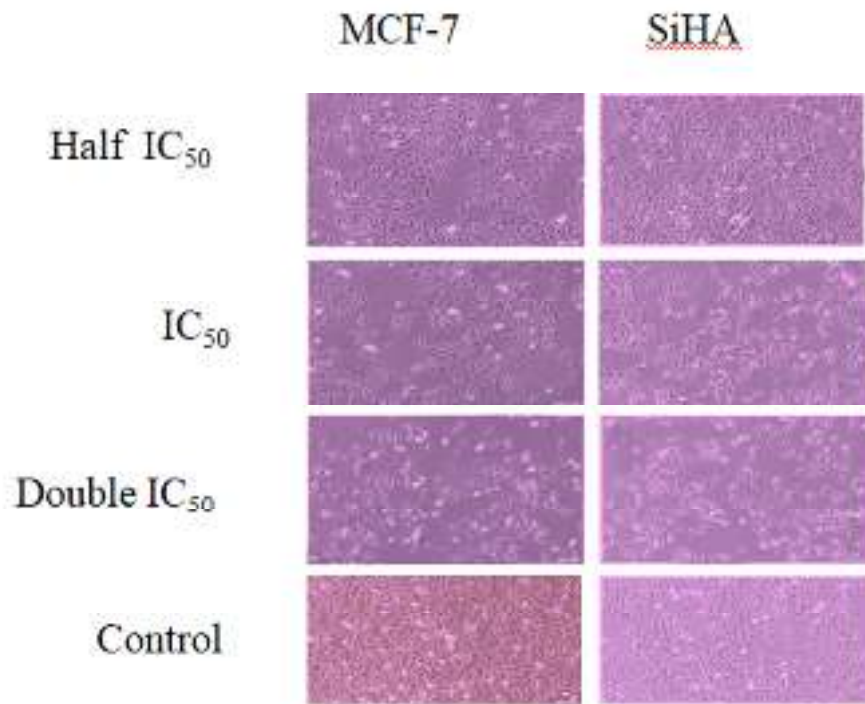
Fig: xxx Effect of *Ceiba speciosa* **on cellular morphology**

Induction of apoptosis by *Ceiba speciosa* in Cancer cells
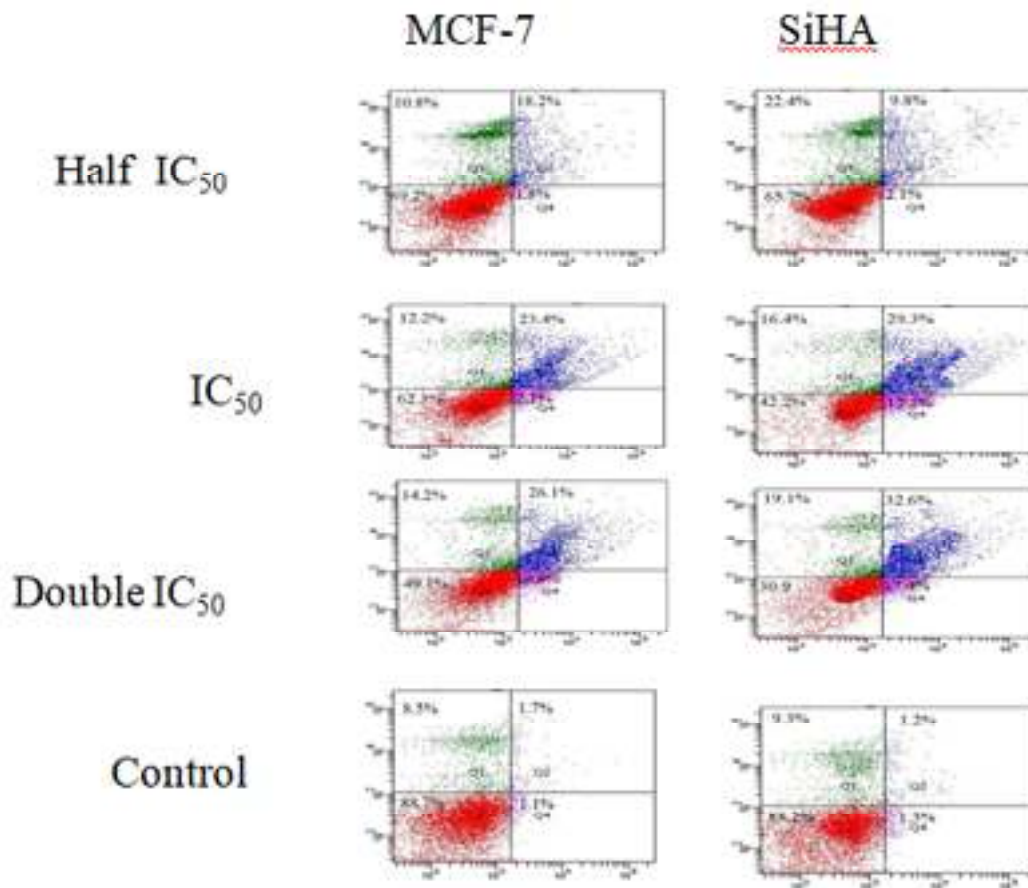
Fig:5- Induction of apoptosis by *Ceiba speciosa* in Cancer cells

Double staining by Annexin V-FITC and PI was performed to examine the apoptotic potential of the extract.

Table-6: Effect of *Ceiba speciosa* on the Apoptosis

Concentration	SiHA cell lines Population				MCF-7 cell lines Population			
	Viable	Early apoptotic	Late apoptotic	Dead	Viable	Early apoptotic	Late apoptotic	Dead
Half IC$_{50}$	65.7%	2.1%	9.8%	22.4%	69.2%	1.8%	18.2%	10.8%
IC$_{50}$	42.2%	12.1%	29.3%	16.4%	62.3%	2.1%	23.4%	12.2%
Double IC$_{50}$	30.9%	17.4%	32.6%	19.1%	49.1%	10.6%	26.1%	14.2%
Control	88.2%	1.3%	1.2%	9.3%	88.7%	1.1%	1.7%	8.25%

343

Apoptosis Mode by Protein-Based Analysis.

Protein-based analysis using Western blot was employed to understand underlying mechanism of apoptosis induction by *Ceiba speciosa* extract . This technique allows us to examine specific proteins involved in the apoptotic process.

As an internal control, actin protein was used to ensure accurate protein quantification. We focused on key apoptosis-related proteins, namely Bax, bcl-2, and caspase-8 to assess their expression levels and potential involvement in the apoptosis of MCF-7 cells induced by the extract.
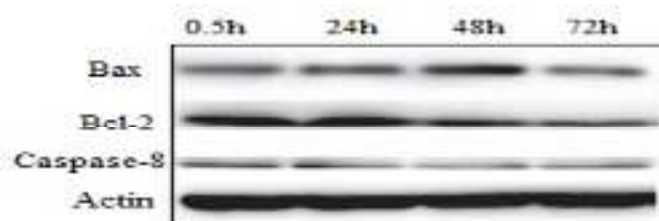
Fig:6- Western blotting analysis of apoptosis-related proteins modulated by *Ceiba speciosa*

Conclusion

Ceiba speciosa plant leaves were collected, dried, powdered and extracted with different solvents employing successive solvent extraction method. Among the solvents employed for the extraction highest yield was observed in methanol followed by water, ethyl acetate, chloroform and petroleum ether. Total phenolic content, total flavonoid content, and total tannins content analysis of all extracts revealed that the methanolic extract consist of high phenolics, flavonoids, and tannins content than remaining extracts. In DPPH assay, the antioxidant activity was observed in a dose dependant manner activity in comparison with standard Ascorbic acid. At higher concentration 75 µg/ml, methanol extract is showing comparatively better inhibition for *C. speciosa* (74.13±0.11) which is almost near to the percentage inhibition of standard ascorbic acid (81.02±0.01). In NO free radical scavenging assay, the antioxidant activity increased with an increased polarity up to methanol in a dose dependent manner. Percentage free radical scavenging activity of the methanolic extract of *C. speciosa* (75.32±1.74) is almost in par with the standard ascorbic acid activity (80.24±0.21). In total antioxidant assay, all the tested extracts of displayed low to good antioxidant potential. Among the tested extracts, methanolic extract of *C. speciosa* shown a significant

total antioxidant activity with the absorbance of 0.95±0.48 in comparison with the standard ascorbic acid absorbance of 1.47±0.44 at 1000 µg/ml concentration.

Ceiba speciosa extract has an apoptosis induction effect on both SiHA and MCF-7 cancer cell lines. The extract demonstrated cytotoxicity and the ability to induce cell death in a concentration-dependent manner.In general, the MCF-7 cell line appeared to be more sensitive to the extract compared to the SiHA cell line. MCF-7 cells exhibited lower viable cell populations and higher percentages of early apoptotic, late apoptotic, and dead cells compared to SiHA cells at equivalent concentrations of the extract. This suggests that MCF-7 cells may be more susceptible to the apoptosis-inducing properties of Ceiba speciosa extract. The protein-based analysis using Western blot provided additional mechanistic insights. The upregulation of the proapoptotic protein Bax suggests its active involvement in promoting apoptosis. In contrast, the down regulation of the antiapoptotic protein bcl-2 and the inactive form of caspase-8 indicates a suppression of antiapoptotic signals. These changes in protein expression further support the induction of apoptosis by the extract. Taken together, the results from both assays strongly indicate that Ceiba speciosa extract triggers apoptosis in MCF-7 cells. The DNA fragmentation and morphological changes observed, along with the alterations in the expression levels of apoptosis-related proteins, provide robust evidence of the extract's cytotoxic effect through the activation of apoptosis pathways.

Future Scope

Current research principally aims to find a nature-based solution for anti-cancer research employing the plant, namely *Ceiba speciosa*. The results of this research work are encouraging and help the research community further in drug discovery and development. Further research regarding the significant phytoconstituents responsible for the anti-cancer activity of the plant extracts enables the scientific and industrial research community to understand the molecular targets and mechanism of action of the compounds that can be developed into new lead molecules. Further more, the structure-activity relationship studies of these developed phytoconstituents might help to increase the understanding of the pharmacophore mechanism of action that can lead to the discovery of newer classes of anti-cancer agents in the future.

References

1. Benson L. Plant Classification. New Delhi, Bombay: Oxford and IBH Publishing Co., 1970; 793-797.

2. Ashmawy AM, Azab SS, Eldahshan OA. Effects of *Chorisia crispiflora* ethyl acetate extract on P21 and NFjB in breast cancer cells. J Am Sci. 2012; 8(8):965-972.

3. Hafez SS, Abdel-Ghani AE, El-Shazly AM. Pharmacognostical and antibacterial studies of *Chorisia speciosa* St. Hil. flower (Bombacaeae). Mans J Pharm Sci. 2003; 19:40-43.

4. El-Alfy TS, El-Sawi SA, Sleem A, Moawad DM. Investigation of flavonoidal content and biological activities of *Chorisia insignis* H.B.K. leaves. Aust J Basic Appl Sci. 2012; 4:1334-1348.

5. Refaat J, Samy MN, Desoukey SY, Ramadan MA, Sugimoto M, Matsunami K et al. Chemical constituents from *Chorisiachodatii* flowers and their biological activities. Med Chem Re. 2015; 24(7):2939- 2949.

6. Nasr, Ereny M Assaf, Mahmoud H. Darwish Faten M, Ramadan Mahmoud A. Phytochemical and biological study of *Chorisia speciosa* A. St. Hil. Cultivated in Egypt. Journal of Pharmacognosy and Phytochemistry. 2018; 7(1): 649-656.

7. Rosselli S, Tundis R, Bruno M, Leporini M, Falco T, Gagliano Candela RG, Loizzo MR. *Ceiba speciosa* (A. St.-Hil.) Seeds Oil: Fatty Acids Profiling by GC-MS and NMR and Bioactivity. Molecules. 2020; 25(5):1037. doi:10.3390/molecules25051037

8. Rahman Sm mushier, Mony Trina, Ahammed Koushik, Naher Sharmin, Haque Md Rubel, Jui Susmita Mistry. Qualitative phytochemical screening and evaluation of analgesic and antidiarrheal activity of ethanolic extract of *Leucas cephalotes* leaves. Journal of Pharmacognosy and Phytochemistry. 2018; 7(5): 1484-1492.

9. Sachin Chavan1, Nishteswar K. A Review on Source Plants of Dronpushpi - *Leucas cephalotes* (Roth) Spreng and *Leucas aspera* Spreng. International Journal of Ayurvedic Medicine. 2013; 4(4): 290-303.

10. Baburao Bhukya, Reddy Anreddy Rama Narsimha, Kiran Gangarapu, Reddy Yellu Narsimha, Mohan Gottumukkala Krishna. Antioxidant, analgesic, and anti-

inflammatory activities of *Leucas cephalotes* (Roxb.ex Roth) Spreng. Brazilian Journal of Pharmaceutical Sciences. 2010; 46(3): 525-529.

11. Surya Narayan Das, Varanasi Jaganath Patro, Subas Chandra Dinda. A review: Ethnobotanical survey of genus Leucas. Pharmacognosy reviews. 2012;6(12): 100-6. doi:10.4103/0973-7847.99943

12. Pathak Rajeev. A formulation for the treatment of arthritis. Indian Pat. Appl. (2018), IN 201631013264 A 20180831

13. Abdel-Aziz, A., Abdel-Motagaly, M., Abdallah, M., Shaaban, R., Osman, N., & Elwan, N. (2021). In vitro Cytotoxicity, Antimicrobial, Antioxidant Activities and HPLC Finger Print Analyses of the Extracts of Ceiba insignis Leaves Growing in Egypt. *Egyptian Journal of Chemistry*, *64*(7), 3573-3586. doi: 10.21608/ejchem.2021.73023.3618.

14. Lim, T. K. (2012). *Ceiba pentandra*, in Edible Medicinal and Non-Medicinal Plants. *Springer*, 540-549.

15. Nasr, E. M., Assaf, M. H., Darwish, F. M., & Ramadan, M. A. (2018). Phytochemical and Biological Study of *Chorisia speciosa* A. St. Hil. Cultivated in Egypt. *Journal of Pharmacognosy and Phytochemistry*, 7(1), 649-656

16. Dua Virendra K, Verma Gaurav, Agarwal Dau Dayal, Kaiser Marcel, Brun Reto. Antiprotozoal activities of traditional medicinal plants from the Garhwal region of North West Himalaya, India. Journal of Ethnopharmacology. 2011; 136(1): 123-8.

17. Shahwar Durre, Naz Misbah, Raza Muhammad Asam, Ara Gulshan, Yasmeen Asma, Saeed Afifa, Bokhari Sana, Ajaib Muhammad, Ahmad Naeem. Acetylcholine esterase inhibitory potential and antioxidant activity of various extracts of *Leucas cephalotes* and Juglans regia L. Asian Journal of Chemistry. 2012; 24(7): 3151-3154

18. Sailor GU, Dudhrejiya AV, Seth AK, Maheshwari R, Shah Nirmal, Aundhia Chintan. Hepatoprotective effect of *Leucas cephalotes* spreng on CCl4 induced liver damage in rats. Pharmacologyonline. 2010; 1: 30-38

19. Sofi G, Khan Mohd Yasir, Jafri M A. Hepatoprotective activity of Gumma (*Leucas cephalotes* Spreng.) against Carbon tetrachloride-induced hepatotoxicity in Wistar rats. Ancient science of life. 2011; 31(2): 44-8

20. Nesy E, Padikkala J, Mathew L. Antioxidant and Cardiac Enzyme Marker Studies of *Thevetia Peruviana* Seed Hydro-Methanol Extract in Wistar Male Albino Rats.

International Journal of Pharmaceutical Sciences and Drug Research. 2020;12(4): 307-12

21. Arathy R, Murugan K, Dinesh Babu KV, Manoj GS. In Vitro Antioxidant Potentiality of Purified Anthocyanin from Floral Petals of Wild Balsam Species. International Journal of Pharmaceutical Sciences and Drug Research. 2020;12(4): 313-9

22. Venkatesan S, Somasundaram A, Rengaraj S. Evaluation of in Vitro Antioxidant Activity and HPTLC Finger Printing Analysis of *Physalis Peruviana* Fruits. International Journal of Pharmaceutical Sciences and Drug Research. 2020;12(4): 384-9

23. Tejavathi DH, Sumalatha BS. Phytochemical, Antioxidant and Antidiabetic Analysis of Leaf and Stem Extracts of Memecylonmalabaricum(C.B.Clarke)Cogn. International Journal of Pharmaceutical Sciences and Drug Research. 2020;12(6): 606-613

24. Shanmugavadivelu CM, Jain N, Thirupathi A, Murugesan P. GC-MS Profiling, Invitro Antidiabetic, Antioxidant and Antimicrobial Activities of a Novel Polyherbal Formulation. International Journal of Pharmaceutical Sciences and Drug Research. 2020;12(6): 590-6

25. Dokuparthi SK, Banerjee N, Kumar A, Singamaneni V, Giri AK, and Mukhopadhyay S: Phytochemical investigation and evaluation of antimutagenic activity of the extract of CuscutareflexaRoxb by Ames test. Int J Pharm Sci Res. 2014; 5(8): 3430-34.doi: 10.13040/IJPSR.0975-8232.5(8).3430-3

26. Harborne JB. Phytochemical Methods. Chapman and Hall Ltd., London: U.K., 1973; 49-188

27. Singleton VL, Rossi JA. Colorimetry of total phenolics with phosphomolybdic phosphotungstic acid reagents. Am J Enol Vitic. 1965; 16: 144-158.

28. Chang CC, Yang MH, Wen HM, Chern JC. Estimation of total flavonoid content in propolis by two complementary colorimetric methods. J Food Drug Anal. 2002; 10: 178-182.

29. Broadhurst RB, jones WT. Analysis of condensed tannins using acidified vanillin. Journal of the Science of Food and Agriculture. 1978; 48(3): 788–794.

30. Blois MS. Antioxidant determinations by the use of a stable free radical. Nature. 1958; 181:1199–1200.

31. Rajasekaran SR. Anandan. Anti-oxidant and phytochemical screening of *Acalypha communis* by HPTLC fingerprinting method. International Journal of Current Research and Review. 2016; 08(14): 33-40

32. Oyaizu M: Studies on products of browning reactions: Antioxidant activity of products of browning reaction prepared from glucosamine. Japan Journal of Nutrition.1986; 44: 307-315.

33. Prieto P, Pineda M, M Aguilar. Spectrophotometric quantitation of antioxidant capacity through the formation of a phosphomolybdenum complex: Specific application to the determination of vitamin E. Anal Biochem. 1999; 269: 337-341.

34. Hollville E, Martin SJ. Measuring apoptosis by microscopy and flow cytometry. Current protocols in immunology. 2016 Feb;112(1):14-38.).

35. Hoetelmans RW, Van Slooten HJ, Keijzer R, Erkeland S, Van De Velde CJ, Van Dierendonck JH. Bcl-2 and Bax proteins are present in interphase nuclei of mammalian cells. Cell Death & Differentiation. 2000 Apr;7(4):384-92.).

Smart Agriculture: Designing a Machine Learning-Based Chilli Crop Disease Detection System

[1]Balasani Raghupathi, [2] Dr. Amit Sharma

[1]Research Scholar, School of Engineering & Technology, Career Point University,

Kota, Rajasthan, India

raghupathibalasani@gmail.com

[2]Associate Professor, School of Engineering & Technology, Career Point University,

Kota, Rajasthan, India

amit.sharma@cpur.in

Abstracts:

With the growing demand for sustainable and efficient agricultural practices, there is an increasing need for innovative technologies to enhance crop management. This research paper presents a comprehensive study on the development and implementation of a Machine Learning-based Chilli Crop Disease Detection System within the context of smart agriculture. The proposed system leverages advanced machine learning algorithms to analyze image data and accurately identify diseases affecting chilli crops.The research outlines the design and architecture of the proposed system, which integrates state-of-the-art image processing techniques with machine learning models for precise disease detection. A dataset comprising diverse chilli crop images, encompassing various disease manifestations, is utilized to train and validate the system. The study also explores the selection and optimization of machine learning algorithms to achieve high accuracy and robust performance in real-world agricultural scenarios.

Furthermore, the paper discusses the integration of the developed system with smart farming infrastructure, enabling real-time monitoring and early detection of crop diseases. The implementation of the proposed solution aims to empower farmers with timely and actionable insights, facilitating proactive disease management strategies and contributing to increased crop yield and sustainability. The results of experimental evaluations demonstrate the effectiveness and reliability of the Machine Learning-based Chilli Crop Disease Detection System. The research findings suggest its potential

to revolutionize conventional agricultural practices by providing a scalable and adaptive solution for crop health monitoring. The implications of this research extend beyond chilli crops, showcasing the broader applicability of smart agriculture technologies in addressing global food security challenges.

Keywords: Smart Agriculture, Machine Learning, Disease Detection, Image Processing Agricultural Imaging, Precision Agriculture, Early Detection, Agricultural Sustainability

Introduction:

The agricultural sector plays a pivotal role in sustaining global food security and supporting the burgeoning world population. In the quest for increased productivity and sustainable farming practices, the integration of cutting-edge technologies has become imperative. Smart agriculture, characterized by the convergence of information technology and agriculture, offers a transformative paradigm to address contemporary challenges faced by farmers.

One of the critical facets of smart agriculture involves leveraging advanced technologies to enhance crop management and mitigate the impact of crop diseases. Among the myriad crops cultivated globally, chilli holds particular significance for its widespread use as a culinary spice and its economic importance in various regions. However, the vulnerability of chilli crops to diseases poses a significant threat to both yield and quality.

This research paper delves into the realm of smart agriculture by proposing and elucidating the development of a Machine Learning-Based Chilli Crop Disease Detection System. By harnessing the power of machine learning algorithms, this system endeavors to revolutionize the traditional methods of disease identification in chilli crops. The integration of machine learning and image processing techniques forms the backbone of this innovative solution, providing a robust and scalable approach to detecting and managing crop diseases.

In the following sections, we will delve into the intricacies of the proposed system, exploring the design, implementation, and optimization processes. Additionally, we

will discuss the broader implications of incorporating such technology into smart farming practices, aiming to empower farmers with timely and accurate insights for proactive disease management. Through this research, we strive to contribute to the ongoing discourse on sustainable agriculture, offering a tangible solution to enhance crop resilience and overall productivity in the face of evolving agricultural landscapes.

1. Agriculture in the Digital Age: Setting the stage for the integration of technology in modern agriculture is characterized by a transformative shift from traditional farming practices to innovative, technology-driven solutions. In recent years, the agricultural sector has witnessed a paradigmatic evolution, driven by the integration of digital technologies. The advent of precision agriculture, data analytics, and the Internet of Things (IoT) has paved the way for what is now commonly known as smart agriculture. This digital revolution in farming aims to enhance efficiency, optimize resource utilization, and address the challenges posed by a growing global population. The integration of technology into agriculture not only streamlines conventional processes but also opens new avenues for sustainable and precision farming practices, positioning the industry to meet the demands of the future. As we stand at the intersection of agriculture and technology, the stage is set for a new era where data-driven insights and advanced technologies play a pivotal role in ensuring food security and agricultural sustainability.

2. **The Imperative for Innovation:** Addressing the challenges faced by traditional agricultural practices underscores the urgent need for innovation in an industry that serves as the backbone of global sustenance. Traditional farming methods, while time-tested, are confronted by a host of challenges ranging from unpredictable weather patterns to diminishing natural resources and the escalating demand for food. To navigate these complexities and ensure the resilience of agricultural systems, a paradigm shift towards innovation is imperative. Embracing technological advancements, data-driven decision-making, and sustainable practices becomes essential in safeguarding crop yields and mitigating the environmental impact of farming. By exploring innovative solutions, agriculture can not only adapt to the evolving landscape but also contribute to addressing pressing issues such as climate change, resource scarcity, and the need for increased productivity. The imperative for innovation in agriculture is not merely a call for change;

it is a strategic response to the multifaceted challenges that threaten global food security and the livelihoods of millions dependent on agriculture.

3. Chilli Crop Vulnerability: Understanding the significance and susceptibility of chilli crops to diseases is a critical aspect of agricultural research, considering the economic and culinary importance of this crop globally. Chilli, a staple in various cuisines, faces a range of diseases that can significantly impact both yield and quality. Factors such as environmental conditions, pest infestations, and microbial infections contribute to the vulnerability of chilli crops. Recognizing the unique challenges posed by these diseases is pivotal for devising effective preventive and management strategies. In this context, the development of a specialized disease detection system becomes paramount, aiming to provide timely and accurate identification of ailments affecting chilli crops. By delving into the intricacies of chilli crop vulnerability, this research seeks to contribute to the broader understanding of crop health and empower farmers with innovative tools to safeguard this vital component of global agriculture.

4. The Rise of Smart Agriculture: Exploring the concept of smart agriculture and its potential impact on crop management heralds a new era in the way we approach food production. Smart agriculture represents the convergence of cutting-edge technologies with traditional farming practices, aiming to optimize resource utilization, enhance efficiency, and ensure sustainability. This paradigm shift leverages the power of sensors, data analytics, automation, and connectivity to enable precision farming. By collecting and analyzing real-time data from the field, smart agriculture empowers farmers with actionable insights for more informed decision-making. The integration of technology not only addresses the challenges posed by a growing global population but also fosters sustainable practices that minimize environmental impact. As we witness the rise of smart agriculture, the potential for improved crop management, increased yields, and reduced environmental footprint becomes increasingly apparent, marking a transformative milestone in the history of agricultural practices.

5. Machine Learning in Agriculture: Introducing machine learning as a transformative tool for disease detection in crops signifies a pioneering approach to crop health management. In recent years, machine learning has emerged as a game-changing technology, offering the potential to revolutionize various industries, including

agriculture. In the context of crop disease detection, machine learning algorithms demonstrate the ability to analyze vast amounts of data, including imagery and environmental variables, to identify patterns indicative of diseases. The adaptability and learning capabilities of these algorithms enable the development of robust models that can accurately and efficiently detect subtle signs of crop ailments. By harnessing the power of machine learning, agriculture enters a new realm of precision and proactive management, allowing for early detection and timely intervention to mitigate the impact of diseases. This research endeavors to explore the integration of machine learning into the agricultural landscape, specifically focusing on its application in enhancing the accuracy and efficiency of disease detection in crops, with a spotlight on chilli cultivation.

6. Proposed Solution: Chilli Crop Disease Detection System: Outlining the objectives and scope of the research in developing an advanced detection system marks a pivotal phase in addressing the challenges faced by chilli cultivation. The proposed solution aims to pioneer a comprehensive Chilli Crop Disease Detection System, integrating cutting-edge machine learning algorithms and image processing techniques. The primary objectives include the development of an accurate and efficient system capable of identifying a spectrum of diseases affecting chilli crops. By leveraging machine learning, the system aspires to learn from diverse datasets, encompassing various disease manifestations, to enhance its adaptability and reliability. The research also seeks to optimize the integration of image processing techniques, ensuring precise analysis of crop imagery for early and accurate disease detection. The scope extends beyond laboratory validation, encompassing real-world scenarios and the integration of the developed system into existing smart farming infrastructure. Through this proposed solution, the research endeavors to contribute to the advancement of precision agriculture, providing farmers with a proactive tool for managing and safeguarding chilli crops against diseases.

7. Integration of Image Processing Techniques: Highlighting the role of image processing in enhancing the precision of disease identification underscores the significance of visual data analysis in modern agriculture. In the context of the Chilli Crop Disease Detection System, image processing techniques play a crucial role in refining the accuracy and efficiency of disease identification. By harnessing these

techniques, the system can meticulously analyze visual information captured from chilli crops, extracting nuanced details that may elude the human eye. Image processing not only facilitates the detection of subtle symptoms but also enables the differentiation of various disease patterns, contributing to a more comprehensive understanding of crop health. This integration aims to empower the Chilli Crop Disease Detection System with the ability to discern intricate visual cues, ensuring a high level of precision in identifying and categorizing diseases. Through the strategic integration of image processing, the research seeks to advance the capabilities of the proposed system, offering a sophisticated solution for proactive disease management in chilli cultivation.

Fig.1.Types of Chill Images

In light of the above discussion, the research endeavours to usher in a new era of agricultural innovation through the development and implementation of a Machine Learning-Based Chilli Crop Disease Detection System. With a comprehensive exploration of smart agriculture, the study underscores the imperative for innovation in addressing challenges posed by traditional farming practices. Acknowledging the vulnerability of chilli crops to diseases, the research positions the proposed system as a

transformative solution, integrating machine learning and image processing to enhance disease identification precision. By elucidating the rise of smart agriculture and the potential of machine learning in this context, the study sets the stage for a paradigm shift in crop management strategies. The proposed Chilli Crop Disease Detection System, outlined with clear objectives and a scope encompassing real-world applications, holds promise for advancing precision agriculture and contributing to sustainable farming practices. The strategic integration of image processing techniques further refines the system's capabilities, emphasizing the role of visual data analysis in proactive disease management. Overall, the research sets defined objectives to validate the efficacy of the developed system, aiming to provide farmers with an advanced tool for timely and accurate decision-making in crop health management.

Review of Literature:

Smart Agriculture, characterized by the integration of advanced technologies, plays a pivotal role in ensuring sustainable crop production and addressing challenges in modern farming practices. This literature review explores the current state of research related to machine learning-based disease detection systems in agriculture, with a specific focus on chilli crops.

1. Machine Learning in Agriculture:

The adoption of machine learning (ML) techniques in agriculture has witnessed a surge in recent years. Researchers have explored the potential of ML algorithms in various aspects of crop management, including disease detection, yield prediction, and optimization of resource usage. The ability of ML models to analyze complex datasets and make informed decisions has shown promise in enhancing agricultural practices.

Md. M. Hasanet. al (2023) It is essential to identify rice illnesses early in order to maintain agricultural quantity and production. It takes a lot of time and money to manually identify rice ailments, especially when dealing with nonnative patterns and colors.illnesses. Machine learning (ML) and image processing techniques are utilized to

rapidly identify rice sickness. This research uses image processing to demonstrate how well machine learning systems diagnose rice illness. We assessed previous studies and a number of approaches to assess the categories of rice illnesses. Accurate disease identification, noise reduction, segmentation, clustering, and feature extraction are among the evaluation criteria. This paper gives methods with reliable results for a variety of datasets for training, image preprocessing, clustering, filtering, and testing. In order to direct future research and development, this study provides insights on machine learning (ML)-based methods for early rice disease identification. We also discuss challenges to be met in order to successfully identify rice illness.

K. V. Deputyet. al. (2023) Despite plant diseases, pollinator loss, and climate change, agriculture drives the economies of many nations and provides food for everyone on the planet. To address the issues of food security, novel approaches to crop loss prevention are needed. Smartphones are attractive for automated image recognition-based disease detection due to their powerful computers and high-resolution cameras. To identify plant illnesses, our system makes use of deep learning-based photo recognition and "PlantVillage" data. To train AlexNet, GoogleNet, ResNet50, and InceptionV3, "training from scratch" and "transfer learning" methods were used. GoogLeNet architecture had the greatest accuracy of 0.999 for colour photos and 0.996 for segmentation images, whereas InceptionV3, which was trained from scratch, got a maximum accuracy of 0.994 for grayscale photographs with a 90:10 train-test ratio. For colour and segmentation photographs, all of the models trained from scratch got an F1-score of 1.0. However, for grayscale shots, GoogleNet and InceptionV3 had the highest F1-score of 0.999 with a train-test ratio of 90:10. Based on these findings, deep learning has the potential to make agricultural plant disease detection more efficient and accurate. Global food production could be boosted via image recognition research.

2. Crop Disease Detection Systems:

Disease detection is a critical component of precision agriculture, enabling early intervention to prevent significant crop losses. Numerous studies have proposed and implemented machine learning-based disease detection systems across various crops. These systems leverage image processing, sensor data, and other agronomic parameters to

identify signs of diseases promptly. The integration of these technologies has demonstrated efficacy in minimizing the impact of diseases on crop yields.

H. Kukadiya et.al (2023) Machine learning in computer science is used to create self-learning algorithms. In this way, "Machine Learning" had its start. Machine learning is a part of AI. Leaf diseases are now classified and diagnosed using deep learning and machine learning. Every crop needs early identification of leaf disease. Farmers may increase their revenues and production by detecting diseases early. To categorize and identify plant leaf diseases, the proposed study examines 40 research publications on machine learning and deep learning. The benefits and drawbacks of machine learning in agriculture are also covered. For next projects, develop an automated web- or mobile-based system for classifying and detecting leaf diseases. This study uses many datasets and machine learning to enhance the identification and categorization of leaf diseases. Farmers benefit from production and economic growth.

3. Chilli Crop Diseases: Chilli crops are susceptible to a range of diseases that can adversely affect both yield and quality. Common chilli diseases include bacterial wilt, anthracnose, and powdery mildew. Researchers have investigated the characteristics and symptoms of these diseases to develop robust diagnostic models. Understanding the specific challenges posed by chilli crop diseases is crucial for designing effective machine learning-based detection systems.

A. J. et. al (2022) The agriculture industry feeds people and economies with high-quality food. Plant diseases may reduce species diversity and food availability. Plant disease detection that is automated or accurate has the potential to increase food output and reduce losses. Recent developments in deep learning have enhanced object detection and picture categorization. In this article, pre-trained CNN models are used to diagnosis plant illnesses. We concentrated on adjusting the hyperparameters of widely used pre-trained models, such as Inception V4, DenseNet-121, ResNet-50, and VGG-16. PlantVillage, which has 54,305 plant disease species photos from 38 categories, was used in the studies. The F1 score, sensitivity, specificity, and accuracy of classification were used to evaluate the model's performance. There were parallels with related studies. DenseNet-121 outperformed the most recent models with a classification accuracy of 70.81%.

4. Image Processing Techniques:

Crop disease detection has become more reliant on image processing methods like computer vision. Scientists have used deep learning architectures like convolutional neural networks (CNNs) to look for illness signs in plant photos. Using these methods, even the most delicate symptoms of illnesses in chilli crops may be quickly and accurately identified.

Research Methodology:

Machine learning classifiers, comprised of AdaBoost, Gradient Boosting, Multi-Layer Perceptron, and Random Forest, are used to implement the proposed system. The best way to find out which of these classifiers can spot crop diseases is to compare them. The first step in colour converting the photographs was to convert them to RGB and HSV. We then proceeded to extract the crop leaf photos' textures and colours for use in model training and testing. Prime real-time crop field images of chilies in Hyderabad, India, taken using high-quality cameras. Every image in the database has been resized to 256 by 256 pixels. We have developed a model that can differentiate between healthy leaves, pests, bacterial leaf spot, fusarium, and leaf curl, among four separate crop diseases. The photos' details are provided in the Table.

Table 1. Details of the number of images

Real Chili Crop Field Image Dataset (1157)				
Bacterial_leaf_spot (C_01)	Fusarium (C_02)	Curl (C_03)	Pests (C_04)	Healthy_leaf (C_05)
230	284	318	142	183
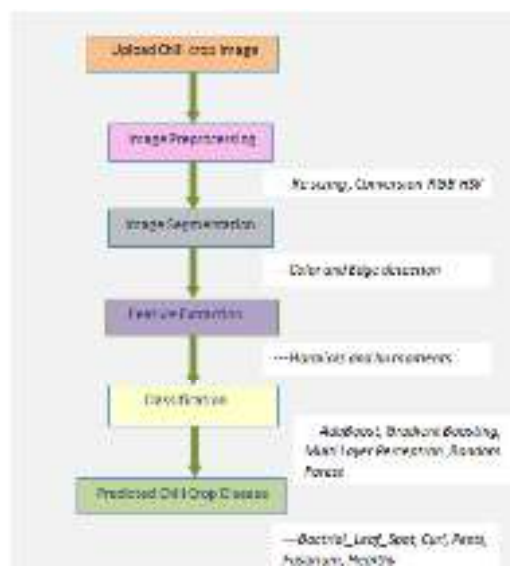
Figure 2. Proposed Architecture model of the Chill Crop Disease Prediction

When the dataset is first uploaded, the architectural design (Figure 2) displays how the suggested application works. Following this, the dataset's pictures undergo preprocessing, including format conversion to HSV and RGB, segmentation, feature extraction using Hu moments and Haralick, and finally, classifier training using the dataset's images. An image-upload button is part of the proposed app's user interface; once submitted, the model returns the F1 score of the several classifiers used (ADA Boost, MLP, RF, and Gradient Boosting) as well as the illness description.

AdaBoost

Adaptive Boosting, or AdaBoost for short, is one of the ensemble boosting classifiers. It improves classifier accuracy by combining several classifiers. Using AdaBoost, ensembles are formed iteratively. By combining many underperforming classifiers, AdaBoost achieves the robust classifier. With each iteration after sample training, AdaBoost determines classifier weights to guarantee accurate predictions for unknown data. A baseclassifier may be any ML method that takes training data weights into account. The first is that AdaBoost has to be able to train its classifier interactively on a set of weighted training examples. The second is that it needs to try to minimise training error in each iteration in order to provide a good match for these instances. It is by balancing the observations that AdaBoost achieves its results; it

gives more weight to instances that are hard to recognise and less weight to those that are already well-classified. In order to focus their instruction on the more difficult patterns, new weak learners are presented one by one. In other words, unless the computer discovers a model that properly labels them, it will assign harder-to-classify samples progressively larger weights.

Gradient boosting

According to the statistical model, adding weak learners is a numerical optimisation problem whose objective is to minimise the model's loss. This category of approaches was described using a stage-wise additive model. The reason for this is because the model only allows for the addition of one new weak learner at a time, while keeping the current weak learners static. The constituent elements of gradient boosting are: Optimising a loss function, using a weak learner to generate predictions, and incorporating weak learners into an additive model to lower the loss function are all possible solutions.

Multi-layer perceptron

The deep learning neural network model that is used the vast majority of the time is known as MLP, which is an abbreviation that stands for multi-layered perceptron. MLP is sometimes referred to as a "vanilla" neural network owing to the fact that it is, in comparison to the models that came before it, a simpler and more straightforward neural network. In a manner that is comparable to the way neurons in the human brain process information, the neurons that are connected together in a multi-layered perceptron are able to communicate with one another and convey information to one another. We give each neuron a number that we have allocated to it. Input, hidden, and output are the three layers that comprise the network. The network is composed of these three levels. In light of the fact that Figure 2 depicts the forward direction of data transmission in a feed forward neural network, it is possible to draw the conclusion that MLP is an example of a feed forward neural network. When it comes to the connections that exist between the tiers, weights are now being applied. Consideration of the weight of a link is one of the most effective methods for determining the importance of a connection. At the same time as the values of the

inputs are gathered from the backgrounds, the values of all the other neurons are formed by making use of the weights and values from the layer that came before it. As an example, the value of the H3 node is obtained by multiplying I1 by W13 and I2 by W23.
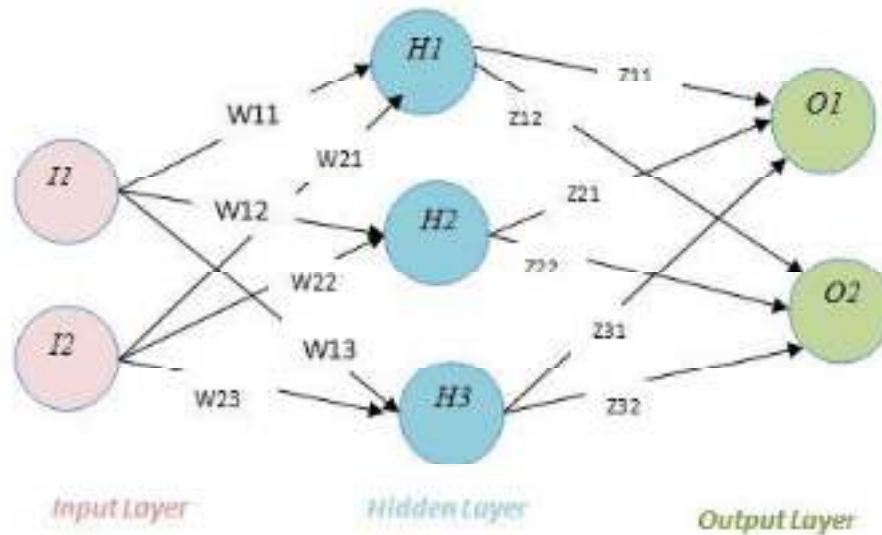
Figure 3. Illustration of MLP

Random forest algorithm

Supervised learning methods include Random Forests. The "bagging" approach is often used to train an ensemble of decision trees, and it creates a "forest" from those trees. Combining many learning models improves the end result; this is the core tenet of the bagging method. One benefit of random forest is that it can handle both classification and regression tasks.

Research Findings:

This concept is put into action with the help of AnacondaSpyder 5. Using colour and edge detection, the pictures are preprocessed and segmented. The segmented bad picture is shown in Table 4. Attributes are obtained using the techniques proposed by Haralick and Hu seconds after classification. Table 5 displays the results of the ML algorithms. The models started off with an accuracy of around 65% after being trained on 926 photos. Then, 868 images were selected for further investigation. Applying RF improved the accuracy to 67%, while Gradient Boosting took it up to 73%. whereas

using AdaBoost, the accuracy dropped to 70%, whereas with MLP, it rose to 77%. Several classifiers' accuracy relative to the training-to-testing image ratio is shown in Figure 4.
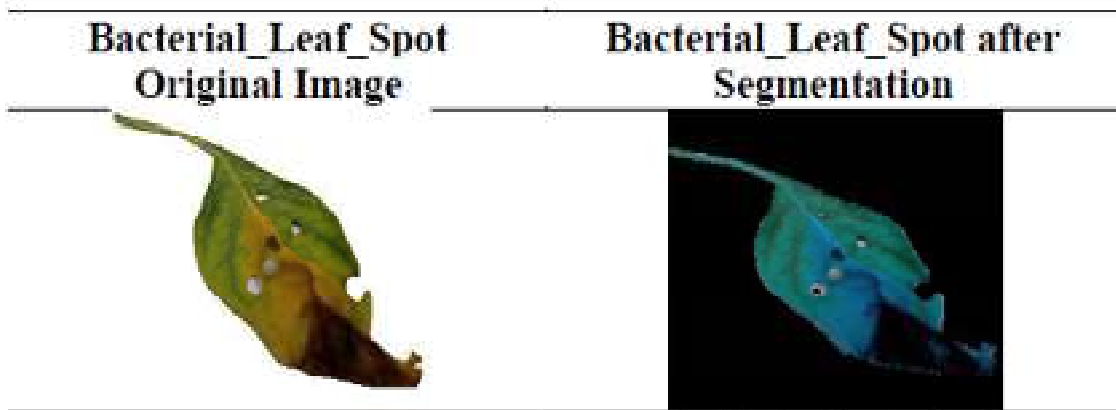
Fig. 4. Original image vs. segmented image
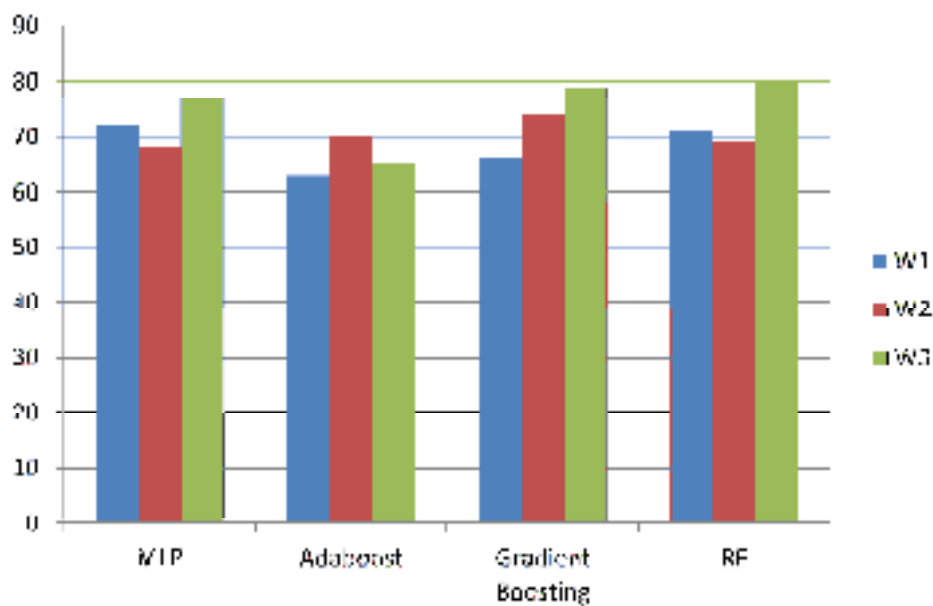
Figure 5. Accuracies of classifiers

Conclusion:

In conclusion, the research undertaken in this paper represents a significant step forward in the realm of smart agriculture, specifically addressing the vital issue of crop disease

detection in chilli plants. The deployment of a machine learning-based system has demonstrated its efficacy in accurately identifying and diagnosing diseases, thereby empowering farmers with timely and precise information for effective crop management. The integration of advanced technologies not only enhances the efficiency of disease detection but also contributes to sustainable agricultural practices by minimizing the use of pesticides and optimizing resource allocation. As evidenced by the results obtained in our study, the proposed machine learning model exhibits promising capabilities in distinguishing between healthy and diseased chilli crops. The success of the system hinges on the utilization of diverse datasets, robust algorithmic frameworks, and continuous refinement through feedback loops. Moreover, the user-friendly interface facilitates easy adoption by farmers, bridging the gap between cutting-edge technology and traditional agricultural practices. While our research contributes significantly to the field, it is important to acknowledge that further investigations and real-world implementations are warranted to validate the scalability and adaptability of the proposed system across diverse agroecological contexts. Additionally, ongoing efforts should focus on refining the model's sensitivity to emerging diseases and incorporating additional features to enhance its overall accuracy.

In essence, this research sets the stage for a transformative era in agriculture, where smart technologies converge with the age-old practice of farming. By embracing innovative solutions like the machine learning-based chilli crop disease detection system, we pave the way for a more resilient and sustainable agricultural future. The fusion of technology and agriculture not only safeguards crop yields but also fosters a more informed and empowered farming community, ultimately contributing to global food security and the well-being of our planet.

Suggestions & Recommendations / Future Scope:

As evidenced by the results obtained in our study, the proposed machine learning model exhibits promising capabilities in distinguishing between healthy and diseased chilli crops. The success of the system hinges on the utilization of diverse datasets, robust algorithmic frameworks, and continuous refinement through feedback loops.

Moreover, the user-friendly interface facilitates easy adoption by farmers, bridging the gap between cutting-edge technology and traditional agricultural practices. While our research contributes significantly to the field, it is important to acknowledge that further investigations and real-world implementations are warranted to validate the scalability and adaptability of the proposed system across diverse agroecological contexts. Additionally, ongoing efforts should focus on refining the model's sensitivity to emerging diseases and incorporating additional features to enhance its overall accuracy.

References: (Follow the guidelines of the references as per requirements of the research paper only). It should be based on APA citation style.

1. Md. M. Hasan, A. F. M. S. Uddin, M. R. Akhond, Md. J. Uddin, Md. A. Hossain, and Md. A. Hossain, "Machine Learning and Image Processing Techniques for Rice Disease Detection: A Critical Analysis," International Journal of Plant Biology, vol. 14, no. 4, pp. 1190–1207, Dec. 2023, doi: 10.3390/ijpb14040087.

2. K. V. Deputy, K. Passi, and C. K. Jain, "Crop Disease Detection Using Deep Learning Techniques on Images," Journal of Computer Science, vol. 19, no. 12, pp. 1438–1449, Dec. 2023, doi: 10.3844/jcssp.2023.1438.1449.

3. H. Kukadiya and D. Meva, "Machine Learning In Agriculture For Crop Diseases Identification: A Survey," International Journal of Research -GRANTHAALAYAH, vol. 11, no. 3, Apr. 2023, doi: 10.29121/granthaalayah.v11.i3.2023.5099.

4. A. J., J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, "Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications," Agronomy, vol. 12, no. 10, p. 2395, Oct. 2022, doi: 10.3390/agronomy12102395.

5. A. Sewak, N. Singla, M. Javed, and G. S. Gill,(2023) "Suitability of pearl millet (Pennisetumglaucum (L.) R. Br.) and sorghum (Sorghum bicolor (L.) Moench) based food products for diabetics," ActaAlimentaria, vol. 52, no. 3, pp. 366–377, Sep. 2023, doi: 10.1556/066.2022.00144.

6. AvitashParmar, M.K.Tripathi, SushmaTiwari, NirajTripathi, PreranaParihar and R. K. Pandya.Characteriza-tion of pearl millet [Pennisetumglau-cum (L.) R Br.](2022) "Genotypes against downey mildew disease employing disease indexing and ISSR markers" . Octa J. Biosci. Vol. 10 (2):134-142

7. S. Singh et al., (2022)"Identification of genes controlling compatible and incompatible reactions of pearl millet (Pennisetumglaucum) against blast (Magnaporthegrisea) pathogen through RNA-Seq," Frontiers in Plant Science, vol. 13, Sep. 2022, doi: 10.3389/fpls.2022.981295.

8. J. Pei et al.,(2022) "A Review of the Potential Consequences of Pearl Millet (Pennisetumglaucum) for Diabetes Mellitus and Other Biomedical Applications," Nutrients, vol. 14, no. 14, p. 2932, Jul. 2022, doi: 10.3390/nu14142932.

9. N. Bani Hani, F. J. Aukour, and M. I. Al-Qinna, "Investigating the Pearl Millet (Pennisetumglaucum) as a Climate-Smart Drought-Tolerant Crop under Jordanian Arid Environments," Sustainability, vol. 14, no. 19, p. 12249, Sep. 2022, doi: 10.3390/su141912249.

10. P. Sharma and A. Sharma, "Online K-means clustering with adaptive dual cost functions," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 2017, pp. 793-799, doi: 10.1109/ICICICT1.2017.8342665.

11. P. Garg and A. Sharma, "A distributed algorithm for local decision of cluster heads in wireless sensor networks," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 2017, pp. 2411-2415, doi: 10.1109/ICPCSI.2017.8392150.

12. A. Sharma and A. Sharma, "KNN-DBSCAN: Using k-nearest neighbor information for parameter-free density based clustering," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 2017, pp. 787-792, doi: 10.1109/ICICICT1.2017.8342664.

13. Salehin, I., Talha, I. M., Saifuzzaman, M., Moon, N. N., &Nur, F. N. (2020, October). An advanced method of treating agricultural crops using image processing algorithms and image data processing systems. In 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA) (pp. 720-724). IEEE.

14. Prakash, K., Saravanamoorthi, P., Sathishkumar, R., &Parimala, M. (2017). A study of image processing in agriculture. International Journal of Advanced Networking and Applications, 9(1), 3311.

15. Latha, M., Poojith, A., Reddy, B. A., & Kumar, G. V. (2014). Image processing in agriculture. International journal of innovative research in electrical, electronics, instrumentation and control engineering, 2(6).

16. Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., & Li, X. (2021). A review of deep learning used in the hyperspectral image analysis for agriculture. Artificial Intelligence Review, 54(7), 5205-5253.

17. Bottou, L., &Bengio, Y. (1994). Convergence properties of the k-means algorithms. Advances in neural information processing systems, 7.

18. Agarwal, P. K., &Procopiuc, C. M. (2002). Exact and approximation algorithms for clustering. Algorithmica, 33, 201-226.

19. Abutaleb, A. S. (1989). Automatic thresholding of gray-level pictures using two-dimensional entropy. Computer vision, graphics, and image processing, 47(1), 22-32.

20. Araujo, S. D. C. S., Malemath, V. S., &Karuppaswamy, M. S. (2020). Automated Disease Identification in Chilli Leaves Using FCM and PSO Techniques. In RTIP2R (2) (pp. 213-221).

21. Naik, B. N., Malmathanraj, R., &Palanisamy, P. (2022). Detection and classification of chilli leaf disease using a squeeze-and-excitation-based CNN model. Ecological Informatics, 69, 101663.

Physicochemical studies to assess the water quality of the ground/pond/river and its treatment strategy

Mansavi Jain

[1]Research Scholar, School of Basic and Applied Sciences, Career Point University, Kota, Rajasthan, India

Abstract

Water is crucial to human survival. It is necessary for every living thing to survive. Life on Earth would not exist without water. Unintended fluctuations in the biological, chemical, and physical, features of soil, air and water, endanger people all over the world. A large amount of hazardous substances are found in water due to human activity, industrialization, usage of fertiliser, and population growth. Water quality is an important issue for humanity because it directly affects human well-being. Safe drinking water is crucial to lowering illness and improving quality of life. Knowing about numerous physicochemical characteristics such as pH, acidity, temperature, hardness, chloride, sulphate, COD and BOD and DO, is very important. Several approaches have been used to evaluate the water bodies' quality. It is recommended for controlled urban agriculture to prevent pollution and reuse water and nutrients.

Keywords: Water Quality Index, physiochemical characters, Groundwater, Lake water, River water

Introduction

One natural resource that is essential to the survival of all living beings is water. It greatly influences how the land is shaped and how the climate is controlled. Freshwater makes up the remaining 3% of the world's water, with saltwater accounting for 97% of it. Azizullah et al. (2011) found that only 0.01% of this 3% is suitable for human consumption. There is very little fresh water on Earth; most water is naturally salty. Because of pollution and overuse, fresh water is becoming increasingly scarce (Basavaraja Simpi et al., 2011).

India gets between 1400 and 1800 mm of rain a year. Approximately 96% of this water is used for farming, 3% for residential, and 1% for industrial purposes. A study found that the discharge of industrial effluents, household waste, agricultural drainage, and land use all contribute to the pollution of about 70% of the nation's accessible water. Due to the consumption of tainted drinking water, the population contracts water-borne illnesses, which

leads to the deterioration of the water quality of these water resources (Shrivastava S and Kanungo, 2013).

The properties that characterise water quality are generally chemical, biological, and physical. Water quality is declining and aquatic biota is disappearing due to rapid industrialization, careless use of chemical pesticides and fertilisers in agriculture, and widespread and diverse degradation of the aquatic environment. Water contamination causes diseases that spread across the human population. As a result, it is crucial to regularly check the water's purity. Testing several parameters, such as nitrates, phosphates, salinity, turbidity, pH, and temperature, is possible. A water quality indicator can also be obtained by evaluating the aquatic macroinvertebrates (Gorde & Jadhav, 2013).

Water quality is declining because of inappropriate use, carelessness, and management. When managing water, many chemical and physical characteristics must be considered. The temperature, pH, hardness, dissolved oxygen, water supply, uses for the water, and destination of any water that flows out of the pond are a few of these characteristics. Seasonal fluctuations impact the properties of bodies of water. Understanding raw water's chemical properties well is essential to determining whether it is suitable for use. Comprehending the various physicochemical factors is crucial in comprehending the metabolic processes within aquatic environments. The parameters regulate the distribution and abundance of and interact with each other (Shinde et al., 2011).

Groundwater is the main supply of potable water since it makes up 99% of all freshwater that is accessible. Numerous human activities, including household, industrial, and agricultural ones, are causing groundwater's physical, chemical, and biological qualities to deteriorate and become contaminated (Azizullah et al., 2011).

All living creatures, food production, ecological systems, and economic progress depend on water, yet its supplies are running out. As a result, one of the most urgent problems facing the modern world is the decreasing availability of fresh, pure drinking water (Qureshi et al., 2021).

In light of the above, this study addresses the word "water quality," which was coined to describe the degree to which a body of water is suitable for human consumption and is frequently used in scholarly works on sustainable management requirements.

Groundwater quality assessment

Groundwater is a valuable source of drinking water in many parts of the world. Because of the various filtration systems found in subsurface soil, it is sometimes seen as pure water when compared to other sources and is used for human consumption in home, agricultural, and industrial settings (Jamshidzadeh & Barzi, 2018; Thompson et al., 2018). Because of the combined effects of urban, industrial, and agricultural activity make it dynamic and flattened (Daud et al., 2017; Ghazanfar et al., 2017). The mineral composition of groundwater is mainly dependent on its origin in space and earth science. Groundwater mineral composition often reflects aquifer composition, recharge type, soil-rock contact, and soil-gas interaction. Aquifers experience responses from water interacting in the saturated zone, which leads to the persistence of water in aquifers (Saxena & Ahmed, 2001). Salts may be present depending on the surroundings, flow, and source of the groundwater. High quantities of soluble salts in groundwater are a result of rock materials degrading. Rarely does groundwater flowing through igneous rocks dissolve mineral particles due to the relative insolubility of the rock composition. The leading causes of groundwater pollution are chemical use, landfills, agricultural activities, septic tank leaks underground, and saltwater intrusion. Groundwater is exposed to direct or indirect toxicity and heavy metals such as boron, arsenic, mercury, and cadmium in agriculture as a result of the use of fertilizers, herbicides, pesticides, and industrial runoff (Tizen et al., 2020a,b; Saleh et al., 2021, 2019). Certain important metals are necessary for the customary development of plant and human populations (Memon et al., 2016; Samreen et al., 2017; Saleh, 2021). According to Mortuza and Al-Misned (2017), the concentration of certain metals over the usual level will have fatal effects on the environment in soil and water contamination that will indirectly affect specific populations. Many substances, including inorganic salts, hazardous metals, cations, and anions (such as nitrates, phosphates, sulfates, calcium, magnesium, and fluoride), are regarded as significant water pollutants (Anwar et al., 2010; Imran et al., 2017; Zafar et al., 2017).

Pondwater quality assessment

Ponds are a vital component of the urban ecology. Despite their tiny size, ponds provide essential environmental, social, and economic purposes. These include replenishing groundwater, supplying drinking water, preventing flooding by acting as sponges, promoting biodiversity, and creating jobs. Ponds have long been a traditional source of water supply in India. However, the leading causes of the pollution of ponds, lakes, and rivers are sewage

outlets, solid wastes, detergents, leftover car oil, fishing facilities, and agricultural pesticides from farmlands (Ozdilek et al., 2007).

Less than 1% of humans' water for domestic, industrial, and agricultural uses is found in ponds, lakes, rivers, dams, etc. Ponds serve various purposes and are one approach to creating artificial subsurface water infiltration. Numerous physical, chemical, and biological elements influence water quality in an aquatic habitat (Acharjee et al., 1999).

River water quality assessment

About 70% of India's rivers are contaminated due to the country's extreme urbanization and industrialization, which has made the water pollution problem extremely serious. The degradation of Indian river water quality has received more attention in the past several decades (Singh & Nautiyal, 1990; Trivedy et al., 1990; Kishor et al., 1998). A river's physicochemical features are initially affected by pollution, which gradually kills the community and upsets the delicate food web while posing a health risk to the general population. Increased pollution significantly hinders the variety of applications that rivers can offer. As a result, it is increasingly critical to analyse river water quality and estimate potential future changes caused by regional development. This will assist in informing consumers downstream of the unfavorable water conditions in advance.

Physico-chemical parameters associated with water

Water must be tested before it is used for drinking, residential, industrial, and farming purposes. It is necessary to examine water using many physicochemical characteristics. Water has various kinds of suspended, dissolving, floating, and microbiological contaminants. While chemical tests should be carried out for its BOD, COD, dissolved oxygen, alkalinity, hardness, and other characteristics, some physical tests should be carried out to test its physical appearance, such as temperature, pH, turbidity, TDS, etc. (Bhateria & Jain, 2016). The Central Pollution Control Board has categorized freshwaters into five classes (A, B, C, D, and E) appropriate for different uses based on specific criteria (Jindal & Sharma, 2011).

pH:

The pH of a solution, also known as its potential hydrogen, is computed by taking the reciprocal of the logarithm of hydrogen ion activity at a given temperature. The most important component to consider when determining how corrosive water is is pH. When water is more corrosive, its pH decreases. A slower rate of photosynthetic activity, as well as

the absorption of carbon dioxide and bicarbonates, result in higher pH. Low oxygen levels and high heat were noted during the summer. There are multiple reasons why water's pH varies. The higher pH values indicate that changes in physico-chemical conditions significantly impact carbon dioxide and carbonate–bicarbonate equilibrium (Bhateria & Jain, 2016).

The Bureau of Indian Standards states that the permitted range for pH in drinking water is between 6.5 and 8.5 (BIS). The buffering capacity (total alkalinity of water), photosynthesis and respiration rates, and other factors determine how much pH fluctuates (Sajitha & Vijayama, 2016).

Dissolved Oxygen:

According to Sajitha and Vijayama (2016), microorganisms use the DO throughout the degradation process of organic matter.

Electrical conductivity:

Electric current transmission via an aqueous solution is known as electrical conductivity (EC). According to Acharya et al. (2008), electrical conductivity is a helpful metric for assessing the cleanliness of water. The recommended maximum acceptable level of EC in drinking water is 1.500 µs/cm. According to Sajitha and Vijayama (2016), the cause of EC could be various discharges, minerals from runoff from rainwater, or an excess of dissolved salts brought on by ineffective irrigation management.

Conductivity is highly correlated with pH, alkalinity, temperature, calcium, total dissolved solids, total solids, chloride, iron concentration, and chemical oxygen demand in water. The primary factor influencing conductivity in rivers and streams is the geology of the water's passage through them. Streams traverse regions with granite bedrock typically exhibit reduced conductivity due to the inert minerals that make up granite, which prevents the water from dissolving into ionic components when washed into it (Gupta & Paul, 2010). Since there are elements in clay soils that ionize when washed into the water, streams that pass through these locations typically have higher conductivities. The impact of groundwater inflows might vary according to the bedrock they pass through. Stream discharges can alter their conductivity based on their composition. Due to the presence of phosphate, nitrate, and chloride, a failing sewage system would increase conductivity, whereas an oil spill would decrease it (Bhateria & Jain, 2016).

Dissolved oxygen

In order to preserve higher kinds of biological life and to maintain the right balance of different pollutants, dissolved oxygen must be present in the water. This keeps the water bodies healthy. In a waterbody, the chemical and biological processes rely heavily on oxygen. Estimating dissolved oxygen is an essential test for controlling waste treatment and water contamination. According to Indian standards, the suggested allowed limit for DO is 5 mg/L (Bhateria & Jain, 2016).

Salinity:

A water body's salinity is determined by how much salt is dissolved.

Temperature:

In the aquatic environment, temperature is one of the most important variables. It impacts the aquatic flora, creatures, and their biological activities in addition to the chemical and physical characteristics of the water. Welch (1952) asserted that a body of water responds to changes in ambient temperature more quickly, the smaller it is.

In addition to regulating all chemical reactions within a well-functioning system, fish immunity, growth, and reproduction are all impacted by the temperature of the specific water. Chemical and biological processes happen at different rates depending on the temperature. Aquatic organisms—from bacteria to fish—need specific temperature ranges for maximum health. The rate at which aquatic plants photosynthesize, aquatic species' metabolic rates, and organisms' sensitivity to toxic wastes, parasites, and diseases are all influenced by temperature. The water's oxygen content is likewise influenced by temperature; it falls as it gets hotter. The weather, the removal of vegetation that shades the stream bank, impoundments, the release of cooling water, urban stormwater, and groundwater inputs to the stream are some factors that cause temperature changes (Spellman & Drinan, 2012).

Minerals:

Raw water naturally contains sodium. However, pollution sources, including rock salt, soapy solution, precipitation runoff, and detergent, raise the concentration of this element. Water tastes unpleasant when a high concentration is present. The primary naturally occurring sources of potassium in groundwater are clay minerals, micas, and feldspar. One crucial metric for identifying sewage contamination is the concentration of chloride. All bodies of water contain calcium, a necessary nutrient for aquatic life. One possible explanation for the drop in calcium levels could be because living things absorb it. Calcium hardness can have a maximum allowable value of 30 mg/l (Sajitha & Vijayama, 2016). Magnesium and calcium

are frequently connected in all types of water, but magnesium concentration is typically lower than calcium (Venkatasubramani & Meenambal, 2007). A decrease in magnesium causes the number of phytoplankton to decrease. In natural waters, phosphorus is always present as phosphate and is one of the limiting nutrients that causes eutrophication, promoting widespread algae bloom. Pollution is indicated by the greater phosphorus concentration (Mishra & Saksena, 2006). Phosphate concentrations indicate pollution over 0.5 mg l-1 (Jain et al., 1996).

Total Hardness:

The ability of water to precipitate soap and scale when specific anions are present is a measure of the water's hardness. The ideal limit for overall hardness, according to APHA, is 300 mg/l.

Total Alkalinity:

Bi-carbonate and carbonates are responsible for the alkalinity found in most natural water. In solution, their salt gate hydrolyzed, releasing hydroxyl ions. Additionally, it serves as a production indicator (Hulyal & Kaliwal, 2008). According to Indian guidelines, alkalinity can have an allowed value of 250 mg/L as $CaCo_3$.

Free carbon-dioxide:

Due to its ability to mix with magnesium, calcium, and other elements, as well as its usage in photosynthetic processes and release during organism respiration and organic matter breakdown, carbon dioxide is very variable. According to Goel and Trivedi, a rise in organic matter increases the requirement for chemicals and biological processes, which in turn elevates free carbon dioxide and lowers DO levels.

Water Quality Index (WQI):

The Water Quality Index is a tool that combines many features and their dimensions into a single score in order to present an overall view of the water quality. WQI can help reduce the pollution issues frequently in various surface water bodies. Dokan Lake's historical variations in water quality were assessed by Alboidy et al. (2010). Ten water quality parameters—pH, dissolved oxygen, turbidity, conductivity, hardness, alkalinity, salt, biochemical oxygen demand, nitrate, and nitrite—have been chosen for this purpose (Bhateria & Jain, 2016).

Biological Oxygen Demand (BOD):

Biological oxygen demand is a measure of the amount of organic matter that can be broken down aerobically by microorganisms in an aquatic environment. BOD measures the quantity

of dissolved oxygen that microorganisms in wastewater and streams use to oxidise reduced chemicals. BOD has a direct effect on the amount of dissolved oxygen in rivers and streams. The rate at which oxygen leaves the stream increases due to BOD. This implies that the amount of oxygen available to higher aquatic creatures is reduced. Similar to aquatic organisms that have low quantities of dissolved oxygen, excessive BOD leads to stress, asphyxia, and ultimately death. Some of the sources of BOD are leaves and woody debris, dead plants and animals, animal dung, food processing facilities, feedlots, failing septic systems, and urban stormwater runoff. Fish deaths and significant dissolved oxygen depletion in receiving water bodies are two issues related to water quality that can arise from the discharge of wastes with high amounts of BOD (Penn et al., 2003). Chlorine has the ability to influence the BOD measurement by preventing or eliminating the bacteria responsible for breaking down the organic and inorganic components present in a sample. Sodium thiosulphate must neutralize chlorine in chlorinated waterways, such as those beneath sewage treatment plant effluent (APHA 2005).

Chemical Oxygen Demand (COD):

According to APHA (2003), Chemical Oxygen Demand is a measurement of the oxygen equivalent of the organic matter in water or the amount required for oxidation by a potent chemical oxidant.

Total Dissolved Solids and Total Suspended Solids

Streams contain three different types of solids: suspended, volatile, and dissolved. Suspended solids consist of sewage treatment effluent, decomposing plant waste, agitated bottom sediment, and silt. Dissolved solids can go through a filter, whereas suspended solids cannot. There are numerous elements that affect a body of water's TDS content. Fertiliser from farms and lawns can transfer a variety of ions into a stream. Increases in TDS could also come from runoff from salted winterized roads. Increased phosphate or nitrate ion contributions may come from organic molecules from wastewater treatment facilities. A wide variety of aquatic species is impacted by high TDS levels, mainly when dissolved salts are involved. Animal skin becomes dehydrated as a result of the salts. TDS concentrations in lakes and streams are usually reported to be between 50 and 250 mg/L. TDS readings can reach 500 mg/L in regions with tough water or significant salinity. Volatile solids evaporate from water or other liquids when dry materials ignite at 1020 degrees Fahrenheit (550 degrees Celsius). It measures the water quality derived from losing total suspended particles upon ignition. It is

crucial to the treatment of wastewater and water. Typically, it denotes the volume of organic matter contained in water. It aids in determining the quantity of organic components that are biologically inert, such as lignin, in waste liquids from wood pulping. If a substance can quickly move from the solid to the vapour phase without going through the liquid phase, it is said to be volatile. About 50% of home wastewater solids are organic, contaminating freshwater and the Earth. These solids are typically made of synthetic chemical substances, dead animal carcasses, and plants. They can burn or catch fire. They are called volatile solids because the organic part can be driven off at high temperatures (Bhateria & Jain, 2016).

Conclusion

Several measures were found to be raised in the different water bodies, which we looked at in the current study. The aquatic ecology would therefore be impacted by these changes because a rise in nitrogen concentration would inevitably result in eutrophication, which lowers oxygen content. Before adding any foreign substance to the water body, filtering procedures should be used as a source of purification measures. Appropriate bioremediation methods must also be applied to raise that water's quality. Moreover, it is recommended that a precise assessment of the water demand for human use and the minimum amount of water based on the biological requirements of floodplains be carried out on a regular basis in order to maximise the benefits gained from their natural function. The discharge of raw home sewage and industrial effluents, along with the washing of clothing, cars, and animals, as well as the immersion of idols during celebrations, are factors that contribute to water pollution. This review indicates a hypereutrophic state because nutrient loading has surpassed the eutrophic condition.

References

1. Azizullah, A., Khattak, M. N. K., Richter, P., & Häder, D. P. (2011). Water pollution in Pakistan and its impact on public health—a review. Environment international, 37(2), 479-497.

2. Simpi, B., Hiremath, S. M., Murthy, K. N. S., Chandrashekarappa, K. N., Patel, A. N., & Puttiah, E. T. (2011). Analysis of water quality using physico-chemical parameters Hosahalli Tank in Shimoga District, Karnataka, India. Global Journal of Science Frontier Research, 11(3), 31-34.

3. Shrivastava, S., & Kanungo, V. K. (2013). Physico-chemical analysis of pond water of Surguja district, Chhattishgarh, India. International Journal of Herbal Medicine, 1(4), 35-43.

4. Gorde, S. P., & Jadhav, M. V. (2013). Assessment of water quality parameters: a review. J Eng Res Appl, 3(6), 2029-2035.

5. Shinde, S. E., Pathan, T. S., Raut, K. S., & Sonawane, D. L. (2011). Studies on the physico-chemical parameters and correlation coefficient of Harsool-savangi Dam, District Aurangabad, India. Middle East Journal of Scientific Research, 8(3), 544-554.

6. Qureshi, S. S., Channa, A., Memon, S. A., Khan, Q., Jamali, G. A., Panhwar, A., & Saleh, T. A. (2021). Assessment of physicochemical characteristics in groundwater quality parameters. Environmental Technology & Innovation, 24, 101877.

7. Jamshidzadeh, Z., & Barzi, M. T. (2018). Groundwater quality assessment using the potability water quality index (PWQI): a case in the Kashan plain, Central Iran. Environmental earth sciences, 77, 1-13.

8. Thompson, L., George, S., Bushra, A., & Santy, S. R. (2018). An assessment of groundwater quality in Kottukal microwatershed in Thiruvananthapuram district, South Kerala. Current Science, 655-660.

9. Daud, M. K., Nafees, M., Ali, S., Rizwan, M., Bajwa, R. A., Shakoor, M. B., ... & Zhu, S. J. (2017). Drinking water quality status and contamination in Pakistan. BioMed research international, 2017.

10. Ghazanfar, H., Saleem, S., Naseem, S., Ghazanfar, A., & Khattak, U. K. (2017). Safe drinking water and sanitary measures: a cross-sectional study in peri-urban community of Islamabad. JPMA. The Journal of the Pakistan Medical Association, 67(2), 220-224.

11. Saxena, V., & Ahmed, S. (2001). Dissolution of fluoride in groundwater: a water-rock interaction study. Environmental geology, 40, 1084-1087.

12. Tuzen, M., Saleh, T. A., & Sarı, A. (2020a). Interfacial polymerization of trimesoyl chloride with melamine and palygorskite for efficient uranium ions ultra-removal. Chemical Engineering Research and Design, 159, 353-361.

13. Tuzen, M., Sarı, A., & Saleh, T. A. (2020b). Synthesis, characterization and evaluation of carbon nanofiber modified-polymer for ultra-removal of thorium ions from aquatic media. Chemical Engineering Research and Design, 163, 76-84.

14. Saleh, T. A., Sarı, A., & Tuzen, M. (2019). Carbon nanotubes grafted with poly (trimesoyl, m-phenylenediamine) for enhanced removal of phenol. Journal of environmental management, 252, 109660.

15. Saleh, T. A., Tuzen, M., & Sarı, A. (2021). Evaluation of poly (ethylene diamine-trimesoyl chloride)-modified diatomite as efficient adsorbent for removal of rhodamine B from wastewater samples. Environmental Science and Pollution Research, 28(39), 55655-55666.

16. Memon, A. H., Lund, G. M., Channa, N. A., Shah, S. A., Younis, M., & Buriro, F. (2016). Contaminants exposure and impacts on drinking water of Johi subdivision of Sindh, Pakistan. Sci Lett, 4(1), 78-83.

17. Samreen, T., Shah, H. U., Ullah, S., & Javid, M. (2017). Zinc effect on growth rate, chlorophyll, protein and mineral contents of hydroponically grown mungbeans plant (Vigna radiata). Arabian Journal of Chemistry, 10, S1802-S1807.

18. Mortuza, M. G., & Al-Misned, F. A. (2017). Environmental contamination and assessment of heavy metals in water, sediments and shrimp of Red Sea Coast of Jizan, Saudi Arabia. J Aquat Pollut Toxicol, 1(1), 5.

19. Anwar, M. S., Lateef, S. H. A. H. L. A., & Siddiqi, G. M. (2010). Bacteriological quality of drinking water in Lahore. Biomedica, 26(1), 66-69.

20. Imran, M., Jahanzaib, S., & Ashraf, A. (2017). Using geographical information systems to assess groundwater contamination from arsenic and related diseases based on survey data in Lahore, Pakistan. Arabian Journal of Geosciences, 10(20), 450.

21. Zafar, S., Khan, A., Ullah, H., Khan, M. S., Khan, I., Hameed, A., ... & Yasmeen, G. (2017). Assessing impact of effluent discharge on irrigation water quality in southern region of Khyber Pakhtunkhwa, Pakistan. Environmental monitoring and assessment, 189, 1-14.

22. Ozdilek, H. G., Mathisen, P. P., & Pellegrino, D. (2007). Distribution of heavy metals in vegetation surrounding the Blackstone River, USA: Considerations regarding sediment contamination and long term metals transport in freshwater riverine ecosystems. Journal of Environmental Biology, 28(2), 493.

23. Acharjee, B. (1999). Role of physico-chemical parameters in the evaluation of productivity of Dighali beel, Assam. Environment and Ecology, 17(2), 274-279.

24. Singh, H. R., & Nautiyal, P. (1990). Altitudinal changes and the impact of municipal sewage on the community structure of macrobenthic insects in the torrential reaches of the river Ganges in the Garhwal Himalaya (India). Acta Hydrobiologica, 32(3-4).

25. Trivedy, R. K., Shrotri, A. C., & Khatavkar, S. D. (1990). Physico-chemical characteristics and phytoplankton of the river Panchganga near Kolhapur, Maharashtra. Rivers of India. New Delhi: Ashish Publishing House.

26. Kishor, B., Bhatt, J. P., Rawat, V. S., & Nautiyal, P. (1998). Stream regulation: variations in the density of benthic macroinvertebrate fauna of Ganga in lateral canals at Hardwar. Journal of Hillstreams Research, 11(1), 62-67.

27. Bhateria, R., & Jain, D. (2016). Water quality assessment of lake water: a review. Sustainable Water Resources Management, 2, 161-173.

28. Sajitha, V., & Vijayamma, S. A. (2016). Study of physico-chemical parameters and pond water quality assessment by using water quality index at Athiyannoor Panchayath, Kerala, India. Emergent Life Sciences Research, 2, 46-51.

29. Acharya, G. D., Hathi, M. V., Patel, A. D., & Parmar, K. C. (2008). Chemical properties of groundwater in bhiloda taluka region, north Gujarat, India. Journal of Chemistry, 5, 792-796.

30. Gupta, T., & Paul, M. (2013). The seasonal variation in ionic composition of pond water of Lumding, Assam, India. Current World Environment, 8(1), 127-131.

31. Spellman, F. R. (2017). The drinking water handbook. CRC Press.

32. Venkatasubramani, R., & Meenambal, T. (2007). Study on subsurface water quality in Mettupalayam taluk of Coimbatore district, Tamil Nadu. Nature, Environment and Pollution Technology, 6(2), 307-310.

33. Mishra SP, Saksena DN. Pollution ecology with reference to physico-chemical characteristics of Morar (Kalpi) River, Gwalior (M. P.). In: current trends in Limnology-1 (Ed. Nalin K. Shastree), Narendra Publishing House, Delhi, India. 1991, 159-184.

34. Jain, S. M., Sharma, M., & Thakur, R. (1996). Seasonal variations in physico-chemical parameters of Halai reservoir of Vidisha district, India. Indian Journal of Ecobiology, 8(3), 181-188.

35. Hulyal, S. B., & Kaliwal, B. B. (2008). Water quality assessment of Almatti Reservoir of Bijapur (Karnataka State, India) with special reference to zooplankton. Environmental monitoring and assessment, 139, 299-306.

36. Goel PK, Trivedy RK. Some conservation on sewage disposal to freshwaters and resultant effects. Journal of Pollution Research. 984; 3(1):7-12.

37. Penn, Pauer, Mihelcic (2003) Environmental and ecological chemistry, vol 2.

38. Federation, W. E., & Aph Association. (2005). Standard methods for the examination of water and wastewater. American Public Health Association (APHA): Washington, DC, USA, 21.

Determination of Physico-Chemical Property of Kelo River Water in Pre-Monsoon Session

[1]Reenu Mishra, [2]Dr. P.K.Sing.[3]Dr. M.M.Vaishnav

[1] Asstt. Professor, K.G.Art's & Science College, Raigarh, C.G.

[2]Asstt. Professor, Govt.TCL College,Janjgir -Champa, C.G.

[3]Asstt. Professor, Govt. Gramya Bharati College, Hardibazar, Korba,C.G.

Abstract-

Water is the soul of life. It finds in three state, solid, liquid and vapors. Life is not possible without water. Surface water is main source for human beings, but now a day it is very difficult to get pure water. Surface water is polluted due to several reasons. The main source of water pollution is anthropogenic activity. Industrial effluents, agricultural run-off, domestic sewage and many more are responsible for water pollution. Some heavy metals like Arsenic come in water by a combination of natural processes such as weathering reactions, biological activity and volcanic emission[1].Industrial effluents contain various heavy metals like As, Zn, Cr, Cu, Fe, Hg, and Cd[2]. Kelo water is used for various purposes that are why it is necessary to monitor the physico-chemical property of Kelo river water. In our present study we choose sample from 7 different spots and analyze various properties. This study helps us to know the pollution level of Kelo river water. Several methods like Chemical precipitation, bioadsorption, ion-exchange method, adsorption, membrane filtration, coagulation[3] are used for the treatment of waste water. Total coli form is the group of bacteria found in environment. In which E-Coli is only one which is found in intestines of mammals including human –beings. [4]

Rationale of the study

In our present study we study the various parameters of Kelo river water. This helps to know the pollution level of kelo river water. Which helps us when we use the river water for any purpose?

Objective of the Study

To analyze the various parameters and help to control the pollution.

Methodology

To analyze the different parameters, we use digital water analyzer kit; with the help of kit we measure pH, Temp, Conductivity, DO, COD, TDS, and Salinity. Acidity, Alkalinity, Hardness, measured by titration method, Sulphate, Phosphate, Nitrate, Chloride are measured by spectrophotometer.Coliform was measured by colony count method.

Result& Discussion-

S.Nu.	Sampling Spot	Sample Number
1	Chakrapath	SA-1
2	Ranighat	SA-2
3	Chattghat	SA-3
4	KeloDam	SA-4
5	Lakha	SA-5
6	Taraimal	SA-6
7	Gerwani	SA-7

s.nu	Parameter	Method	Part	Unit	x	y	1	2	3	4	5	6	7	
1	color	APHA 23rded	2120-B	Hazan	5	15	20	2	2	8	100	8	9	
2	TDS	APHA	2540-C	Mg/l	500	2000	700	254	272	263	1078	678	956	
3	Chloride	APHA 23rded	4500-CL-B	Mg/l	250	1000	44.71	17.49	14.58	34.2	35.96	26-8	74.2	
4	Alkalinit	APHA	3220	Mg/l	20	6	400	90.	85.6	300	500.	4	26	

	y As CaCO3	23rded	-B		0	00	.50	67	7		32	00	8
5	Total Hardness	APHA 23rded	2340 -C	Mg/l	200	600	132 .22	183 .87	194. 2	132 .43	59.9	60. 35	45 .2 4
6	Ph	IS:302 5	11	pH Scale	6.5 - 8.5		8.0 4	7.6 5	6.9	7.8	8.08	7. 9	7. 5
7	Turbidity	IS:302 5	10	NTU	1	5	36. 9	4.5	16.5	3.5	51.2	4. 6	4. 8
8	Iron	APHA 23rded	3500 -Fe- B	Mg/l	0.3	01	0.5 3	0.19 9	0.38	-	1.77	-	-
9	Nitrate	APHA 23rded	4500 - NO3 - B	Mg/l	45	45	14. 60	2.60 0	3.60	20. 5	21.5 6	32. 3	28 .3
10	Fluoride	APHA 23rded	4500 -F-C	Mg/l	1.0	1. 5	0.59 9	0.85 5	0.96	0.89 9	0.76	0. 65	0. 75
11	Sulphate	APHA 23rded	IS:3 025	mg/l	200	400 0	32. 5	70	86	56	19.5	34	43
12	Conducti vity	IS:302 5	14	µmho /cm	-	-	390	392	436	345	1684	42 1	32 1
13	Total- Coliform	-	-	CFU/ 100m	-	-	5.3	3.0	3.6	2.5	4.2	3. 5	4. 2

				1									

Conclusion-

It is very necessary to continue monitoring of water quality parameters.TDS level is high in each spot .Alkalinity value is high in sampling spot 5. Sulphate level is quit high in sampling spot 1, 2, 3, 4.Hardness level is high in every sampling spot. Presence of total coliform in sample water indicates the presence of bacteria in surface water, which may cause diseases.

References

1. B.K. Bisvas, J.-I.Inoue et al., Adsorptive removal of As(v) and As(III) from water by a Zr (IV)-loaded orange waste gel, "Journals of Hazardous Materials, vol.154,no 1-3,pp.1066-1074,2008

2. Dimple Lakherwal, Adsorption of Heavy Metals-A Review," International Journal of Environmental Research and Development, vol-4, no-1, pp.41-48, 2014, Research India Publication

3. F. Fu, Q Wang, Journal of Environmental Management, vol-94, pp407 (2011).

4. Metcafe and Eddy "Wate water Engineering, Treatment, Disposal, Reuse "Tata McGraw-Hill second edition, 1979.

Studies on Trichomes Diversity of Five Plant Species of Family Asteraceae from Semi - Arid Regions of Barmer, Rajasthan, India

Vimla

Research Scholar, Mankya Lal Verma Government College, Bhilwara, Rajasthan

Email: research.vimla@gmail.com

Abstract:

Trichomes are fine outgrowth appendages on plants. They are of diverse surface structure and function. Trichomes play a crucial role in the identification of plants belonging to the Asteraceae family. The morphology, ultrastructure, density and distribution of trichomes on abaxial, adaxial surfaces of leaves and stem of Echinops echinatus Roxb., Eclipta prostrata (L) L. , Grangea maderaspatana (L) Poir, Tridex procumbens L. and Verbesina encelioides (Cav.) Benth. & Hook. f were examined by mean of compound light microscopy and images taken by Samsung smart phone. All five investigated species contained both Glandular and Non-glandular trichomes but differed from each other in trichome ultrastructure and density. Leaves of Echinops echinatus Roxb. and Tridex procumbens L. are more hairy than other studies species. The glandular trichomes are biseriate, sessile with capitate head in Grangea maderaspatana, peltate glandular- stellate, multicellular trichome in Tridex procumbens, paltate trichome- ball shaped, multicellular in Verbesina encelioides were observed. Non-glandular were either filiform, short or long with pointed end or blunt end in studied species. The density of trichomes was higher on stem and abaxial surface of leaves than the adaxial surface of leaves.

Keywords: Asteraceae, density, morphology, peltate, trichomes

Introduction

With over 23000 species distributed throughout 3 subfamilies, 17 tribes, and more than 1535 genera, the Asteraceae family is the most diverse group of angiosperms [1]. The majority of the family's members are herbaceous, with a few occasional shrubs and trees. Trichomes are

the variety of shapes, structures, and functions of epidermal outgrowth [2]. Trichomes can appear on the stem, leaves, fruit, and seeds, among other parts of the plant [3].

Numerous glandular and non-glandular trichomes frequently cover the surface of leaves, stems, and other plant parts. The functional, taxonomic, and economic value of some trichome-generated products is the main reasons for scientific interest in them. Terrestrial plants frequently have trichomes, which are cells derived from the epidermis. They are easily observable, have a variety of morphologies, and are a great model system for analysing the molecular mechanisms involved in plant cell differentiation, such as morphogenesis and the cellular cycle. Trichomes can be categorised as glandular or non-glandular based on their structural makeup. They can be unicellular or multicellular [4]. The number of trichomes in leaves varies and is influenced by both genetic and environmental variables [5]. "Reference [5] shows, non-glandular trichomes from a dense indumentum serves as a mechanical barrier against a variety of harmful elements, including infections and herbivores, UV-B radiation, extremely high or low temperatures, and excessive water." In plant taxonomy, trichomes—which differ substantially in size, shape, location, structure, and composition—as well as the exudates they generate are utilised to differentiate between closely related species or hybrids [6]. The primary theory explaining how trichomes developed to regulate temperature in response to abiotic stress is that they kept their leaves cooler by reducing their absorption of short wave radiation [7], water loss [8], and photosynthetic conductance [9].

Aim and objective of the study:

1. In this study characterized the trichomes present on leaves and stem of five species belonging to family Asteraceae.
2. To know the diversity of trichomes.
3. To study the morphology of different trichomes.

Material and Methods

Collection and identification of plant specimens:

This research was carried out in the Semi-arid region of Barmer, Rajasthan,India. Healthy and mature plant species of *Echinops echinatus* Roxb., *Eclipta prostrata* (L) L. , *Grangea*

maderaspatana (L) Poir, *Tridex procumbens* L. and *Verbesina encelioides* (Cav.) Benth. & Hook. f. (Table I) of the family Asteraceae were collected during regular field trips to various parts and localities like barmer magra , barmer gramin, baytu, haldeshwar mahadev and Nand of the study area and voucher specimens were prepared as per as the standard herbarium methodology [10]. After going through the review of literature its identify was determined with the Flora ([11], [12]). The voucher specimens are housed in herbarium of the SRK Govt. College, Rajsamand, Rajasthan (India).

*Epidermal studies***:**

Epidermal peels from the leaf's abaxial and adaxial surfaces and stem were cautiously separated with a razor blade and then stained with safranin. The epidermal strips were then temporary mounted on clean slides in 50% glycerin and covered with cover slip. Slides were examined under a light microscope at X4, X10, and X40 magnifications. Using a Samsung smart phone and Sony Alpha 6100 camera, photomicrographs were made from well-prepared samples.

Table I: Collected plant specimens of family Asteraceae in Semi-arid regions of Barmer, Rajasthan

Sr. No.	Botanical name	Local Name	Field Note	Voucher No.
1.	*Echinops echinatus* Roxb.	Oont-kateli	Common at waste land and road side, Nand	HBSRKGCR07
2.	*Eclipta prostrata* L.	Jal Bhangrow	Common at waste land and wet regions, Magra Barmer	HBSRKGCR08
3.	*Grangea maderaspatana* (L.) poir	Mukhatari	QC26+8JG, Kisan bhawan ground, Nehru Nagar, Barmer	HBSRKGCR57

4.	*Tridex procumbens* L.	Kagla ri mehndi	Common in wasteland and open forest among bunches in dry habitats, baytu	HBSRKGCR48
5.	*Verbesina encelioides* (Cav.) Benth. & Hook. f.	-	Common in wasteland, haldeshwar mahadev	HBSRKGCR50

I.   RESULT

In total five plant species were selected for trichomes study viz. *Echinops echinatus* Roxb., *Eclipta prostrata* (L) L., *Grangea maderaspatana* (L) Poir, *Tridex procumbens* L. and *Verbesina encelioides* (Cav.) Benth. & Hook. f. belonging to Tribes Cardueae, Heliantheae and Astereae (Table II). Generally the trichomes of the studied species of family Asteraceae were glandular and non-glandular. Trichomes were straight or curved, unicellular or multicellular. Base of trichomes were swollen, rounded and flat and end are pointed or blunt were observed. In some species peltate type of glandular trichomes also observed (Table III).

Table II: species diversity of Asteraceae family in Semi-arid regions of Barmer, Rajasthan

S. No.	Botanical name	Common Name	Tribe	Habit	Flowering & Fruiting Time
1.	*Echinops echinatus* Roxb.	Indian globe thistle	Cardueae	Rigid, spiny herb	December-April
2.	*Eclipta prostrata* (L.) L.	False daisy	Heliantheae	Erect or prostrate herb	September-December
3.	*Grangea*	Madras carpet	Astereae	Prostrate,	November-

				hairy annual herb	April
4.	*Tridex procumbens* L.	Coat buttons	Heliantheae	Upright, annual herb	October-February
5.	*Verbesina encelioides* (Cav.) Benth. & Hook. f.	Golden crown beard, Cowpen daisy	Heliantheae	Erect annual or perennial herb	October-February

Table III: Various types of trichomes in different plant species of Asteraceae from Semi-arid regions, Barmer

S. No.	Name of Species	Abaxial surface of leaf	Adaxial surface of leaf	Stem
1.	*Echinops echinatus* Roxb.	1. Glandular trichomes with unicellular stalk and bicellular head (Fig.- 1a) 2. Non glandular trichome, uniseriate, unicellular (Fig.- 1b) 3. Non glandular long filiform trichome, flagellate (Fig.- 1c)	1. Non glandular long filiform trichome, uniseriate, unicellular (Fig.- 1b) 2. Non glandular long filiform trichome, flagellate (Fig.- 1c)	1. Glandular trichome with bicellular stalk, unicellular and capitate head (Fig.- 1d) 2. Glandular multicellular with capitate head (Fig.- 1e) 3. Non glandular filiform trichome with

				swallon base and pointed tip (Fig.- 1f)
2.	*Eclipta prostrata* (L) L.	1. Glandular trichome with short stalk, capitate head (Fig.- 2g) 2. Non glandular- uniseriate, pointed tip (Fig.- 2h) 3. Non glandular- simple, multicellular stalk with echinate ornamentation, pointed tip (Fig.- 2i)	1. Non glandular- simple, multicellular stalk with echinate ornamentation, pointed tip (Fig.- 2i)	1. Non glandular- simple, short, uniseriate, cylindrical with conical tuberculated head (Fig.-2j)
3.	*Grangea maderaspatana* (L) Poir	1. Glandular- biseriate, sessile with capitate head (Fig.-3a) 2. Non glandular- multicellular, septate with pointed tip (Fig.- 3b) 3. Non glandular- multicellular, septate with	1. Non glandular- short, septate, multicellular with pointed tip (Fig.- 3d) 2. Non glandular- long, septate, multicellular with pointed tip (Fig.- 3e) 3. Non glandular- long, aseptate,	1. Glandular- sessile with capitate head (Fig.- 3g) 2. Non glandular- aseptate, blunt end (Fig.- 3h) 3. Non glandular- uniseriate, swallon base

		rounded tip (Fig.- 3c)	multicellular with pointed tip (Fig.- 3f)	with pointed tip (Fig.- 3i)
4.	Tridex procumbens L.	1. Non-glandular uniseriate, aseptate, long filiform, pointed end (Fig.- 4a) 2. Non- glandular base multicellular, short , blunt end (Fig.- 4b & 4c) 3. Peltate glandular- stellate, multicellular(Fig.- 4d)	1. Non-glandular - uniseriate, aseptate, long filiform, pointed end (Fig.- 4e)	1. Non-glandular – transparent , aseptate, long filiform, swollen base, pointed end (Fig.- 4f) 2. Non-glandular- filliform, long, pointed end (Fig.- 4g & 4h) 3. Non-glandular uniseriate, septate, blunt end (Fig.- 4i)
5.	*Verbesina encelioides* (Cav.) Benth. & Hook. f.	1. Non-glandular - swallow base, long having pointed end (Fig.- 5a) 2. Non-glandular - swallow base, short having blunt end (Fig.- 5b) 3. Glandular paltate	1. Non-glandular - swallow base, long having pointed end (Fig.- 5e)	1. Non-glandular - swollen base, long having pointed end (Fig.- 5f) 2. Non-glandular - septate , long having pointed end (Fig.- 5g & 5h)

		trichome- ball shaped, multicellular (Fig.- 5c & 5d)		3. Non-glandular - aseptate , swollen base having blunt end (Fig.- 5i)
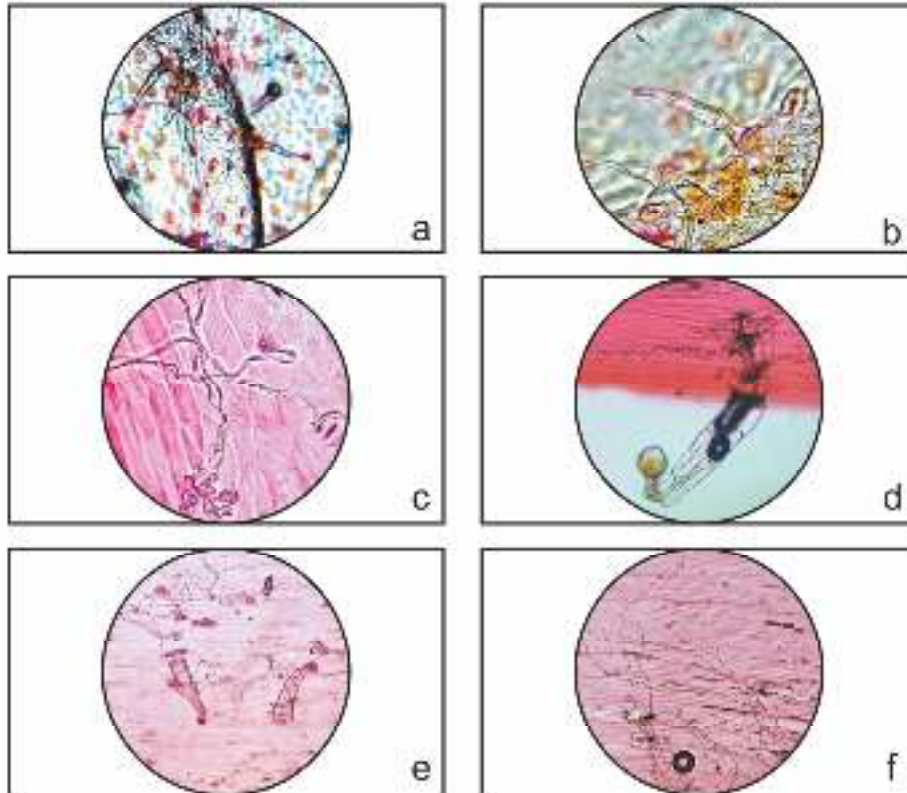
Figure 1: Various types of trichomes of *Echinopus echinatus* Roxb. :Abaxial surface of leves (a, b & c); Stem (d, e & f)
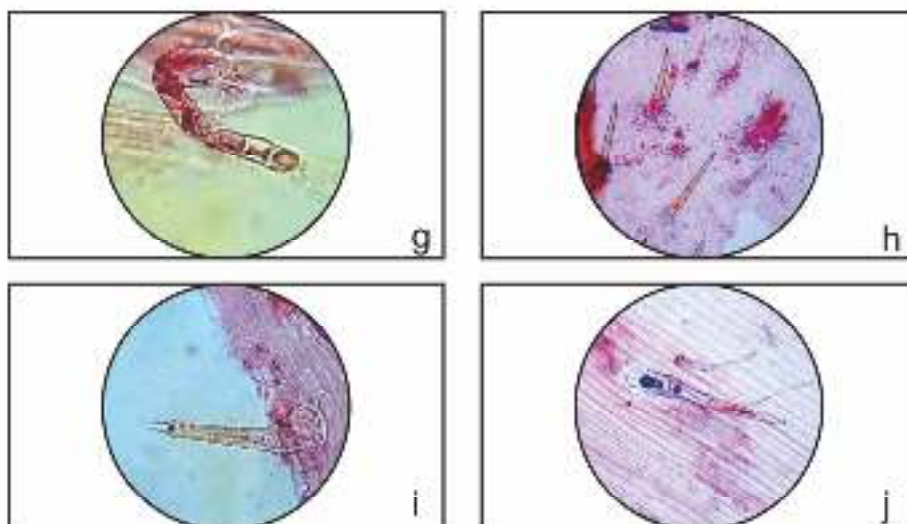
Figure 2: Various types of trichomes of *Eclipta prostrata* (L.) L.

:Abaxial surface of leves (g – i); Stem (j)
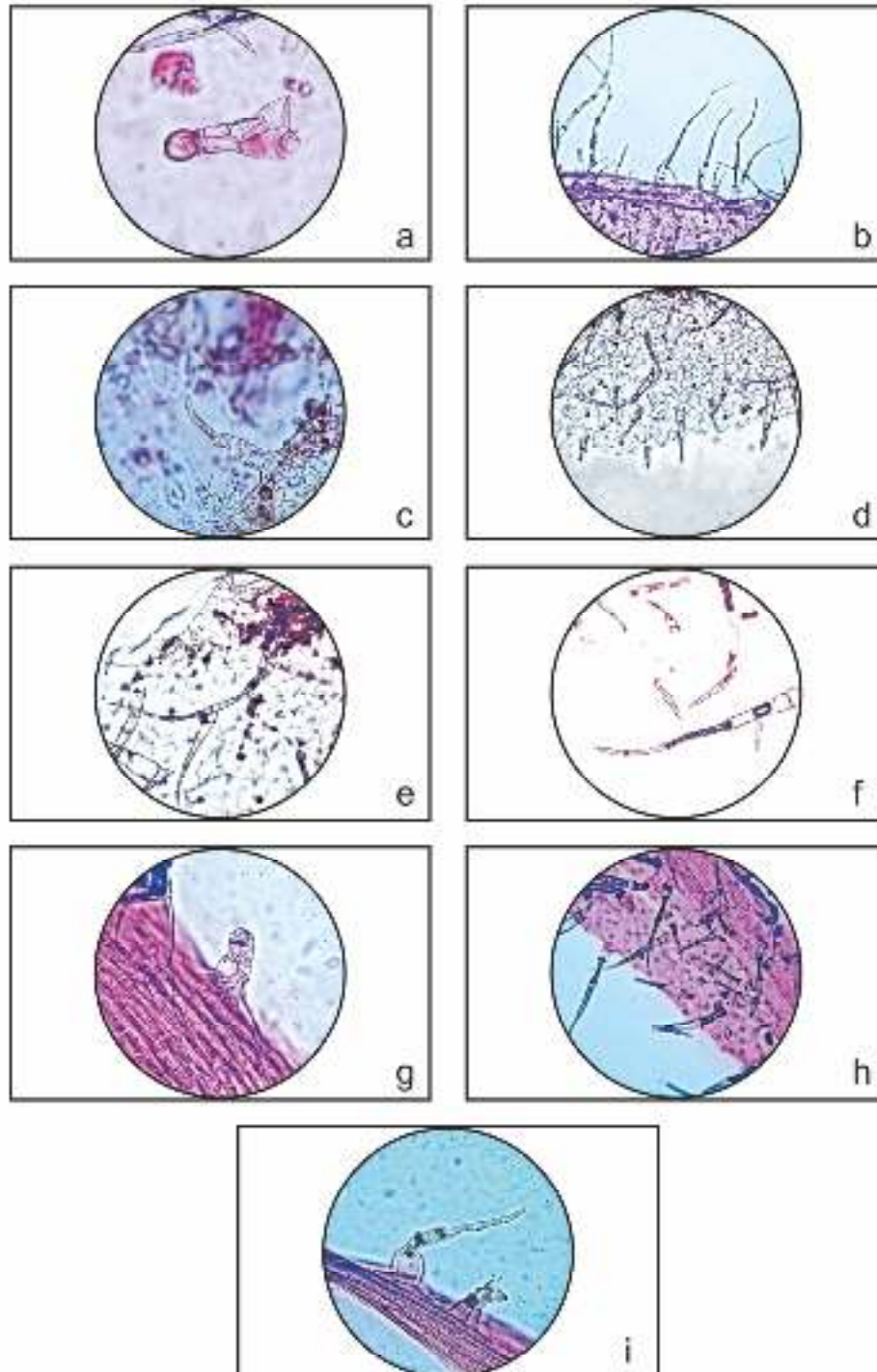
Figure 3: Various types of trichomes of *Grangea maderaspatana* (L.) Poir

:Abaxial surface of leaves (a – c); Adaxial surface of leaves (d – f); Stem (g – i)
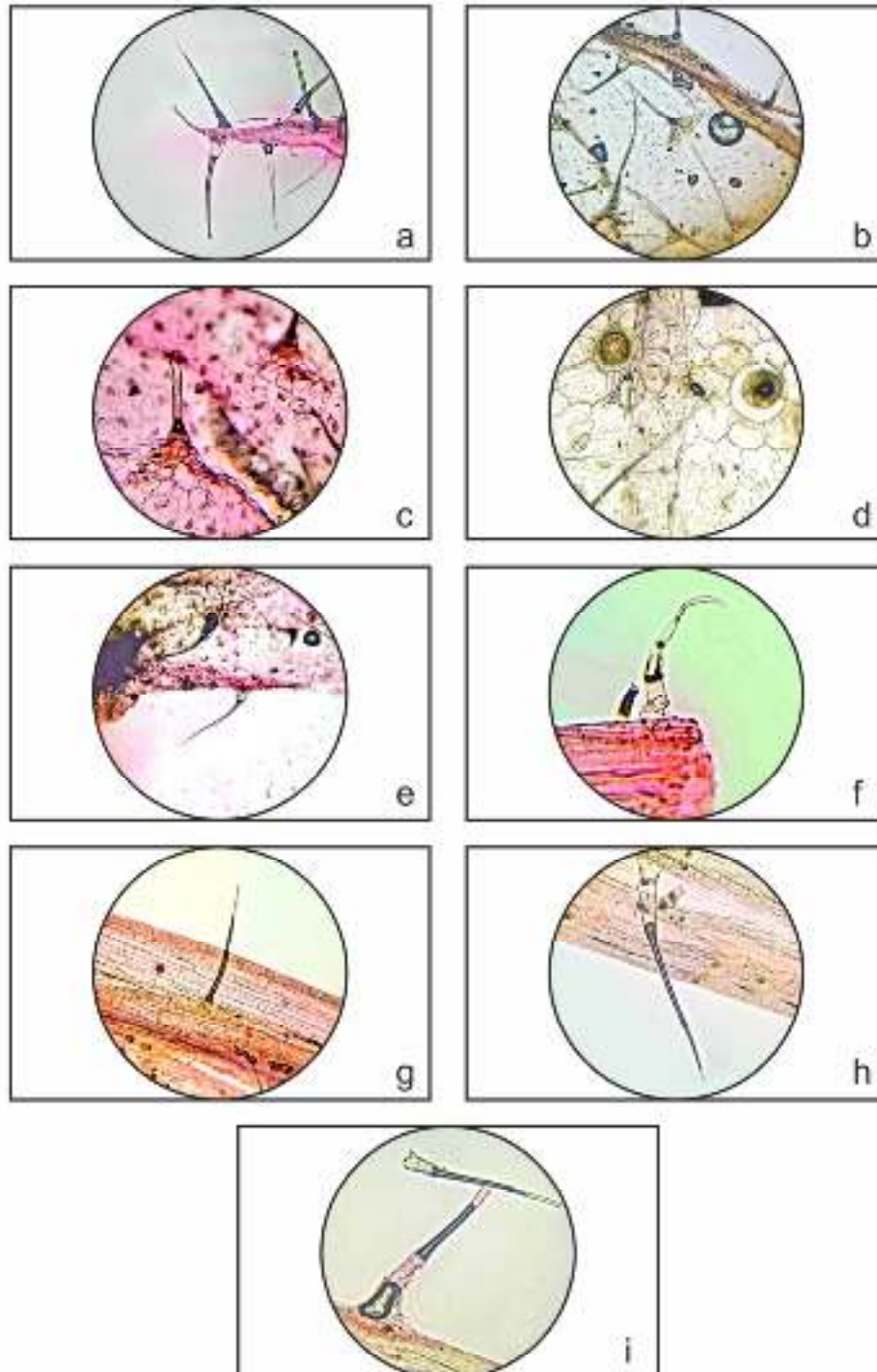
Figure 4: Various types of trichomes of *Tridex procumbens* L.

Abaxial surface of leaves (a – d); Adaxial surface of leaves (e); Stem (f – i)
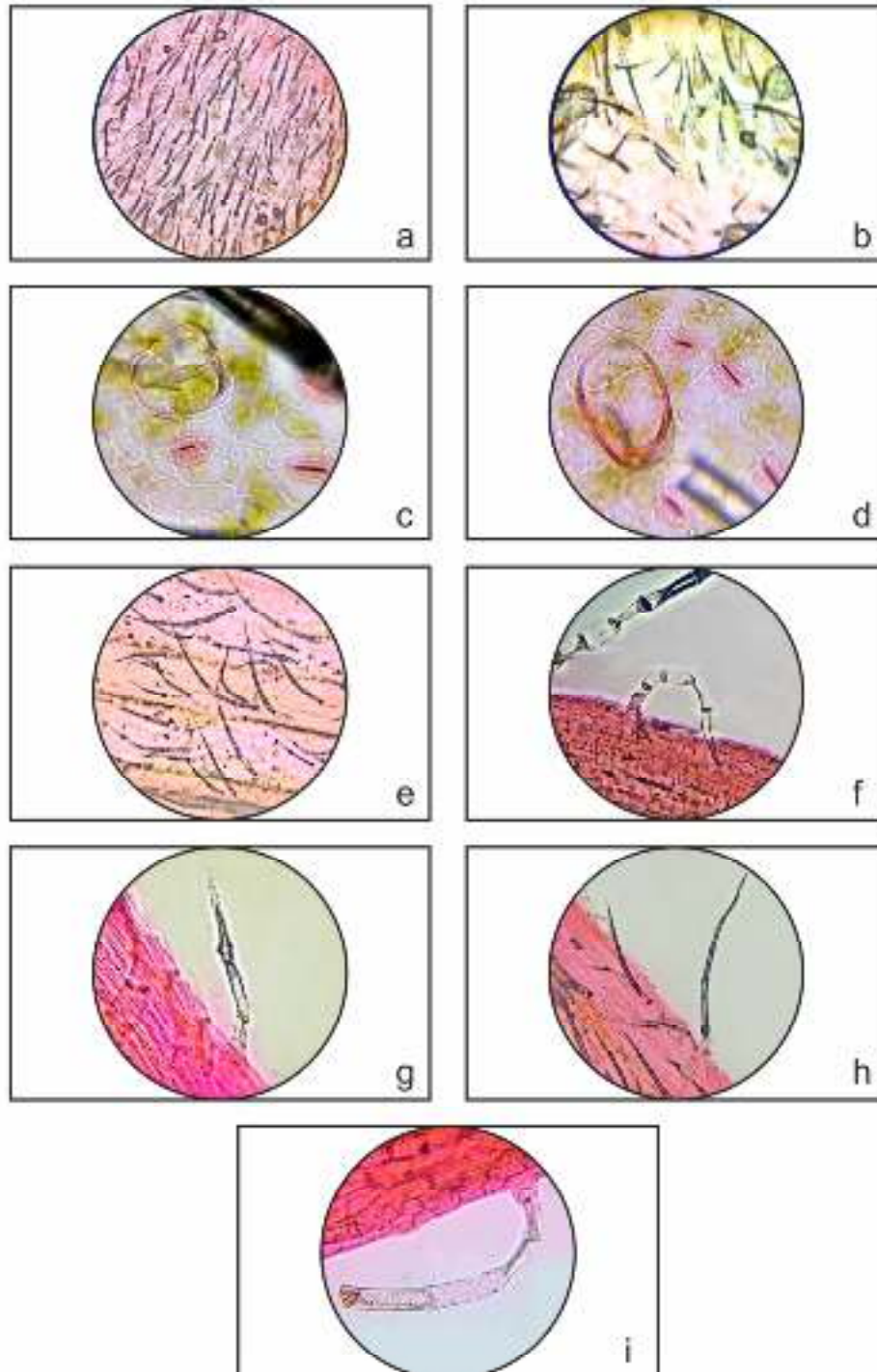
Figure 5: Various types of trichomes of *Verbesina encelioides* (Cav.) Benth. & Hook. f

Abaxial surface of leaves (a – d); Adaxial surface of leaves (e); Stem (f – i)

Discussion

In the present study trichome morphology and diversity of five plant species belonging to the family Asteraceae were studied. Most of the plants having hairs on the outer surface of different plants parts termed as trichomes, which is of greater importance. They differ in structure and vary from species to species. Their characteristics can be used to classify the species. The role of trichomes is to avert animals from grazing, guide the path of pollinators or the rate of photosynthesis is may be affected by the increased reflection which cause water loss and vary leaf temperature [13].Out of five species all showed glandular and non-glandular trichomes on abaxial, adaxial surfaces of leaves and surface of stem. As a result of their ability to produce, store, and exude chemical compounds that aid in protecting plants from insect predation ([13], [14], [15]). Glandular trichomes has drawn a lot of attention. Various trichome types with notable morphological structural differences have been reported in the current investigation. The identification of different species of family Asteraceae would benefit from study of the trichomes.

Conclusion

In conclusion, the new details on trichome shape can be used as additional data to aid in the studied species identification. The present study is therefore helped at identifying, describing and documenting the leaf and stem epidermal characters that diagnostic and taxonomically important in the family Asteraceae.

References

1. C. Youngsheng and A. A. Anderberg, Inuleae. *Flora of China.* 20-21, 820-850, 2011.

2. K. Esau, *Plant Anatomy*, (2nd Edition). John Wiley and Sons, New York, 767 – 768, 1965.

3. E. G. Cutter, Plant Anatomy: Cells and Tissues. 2nd Edn. Edward Arnold, London, 1978.

4. A. R. Walker and M. D. Marks, Trichome initiation in Arabidopsis, 2000.

5. E. Werker, Trichome diversity and development 1-35, 2000. https://doi.org/10.1016/S0065-2296(00)31005-9

6. O. Spring, Chemotaxonomy based on metabolites from glandular trichomes. 153 – 174, 2000. https://doi.org/10.1016/S0065-2296(00)31009-6

7. V. P. Gutschick, Biotic and abiotic consequences of differences in leaf structure. *The New Phytologist*, *143*(1), 3-18, 1999. https://doi.org/10.1046/j.1469-8137.1999.00423.x

8. B. W. Benz and C. E. Martin, Foliar trichomes, boundary layers, and gas exchange in 12 species of epiphytic Tillandsia (Bromeliaceae). *Journal of Plant Physiology*, *163*(6), 648-656, 2006. https://doi.org/10.1016/j.jplph.2005.05.008

9. J. Galmés, H. Medrano and J. Flexas, Photosynthesis and photoinhibition in response to drought in a pubescent (var. minor) and a glabrous (var. palaui) variety of *Digitalis minor*. *Environmental and Experimental Botany*, *60*(1), 105-111, 2007. https://doi.org/10.1016/j.envexpbot.2006.08.001

10. S. K. Jain and R. R. Rao, *Field and Herbarium Methods*. Today and Tomorrow´s Printers and Publishers. New Delhi, India, 1977.

11. M. M. Bhandari, *Flora of the Indian desert*. MPS. Reports, Jodhpur, 1995.

12. B. V. Shetty and V. Singh, *Flora of Rajasthan*. Botanical Survey of Indian, Kolkata, 1987.

13. G. J. Wagner, Secreting glandular trichomes: more than just hairs. *Plant physiology*, 96(3):675-679, 1991. https://doi.org/10.1104/pp.96.3.675

14. C. R. Metcalfe and L. Chalk, *Anatomy of the dicotyledons* Vol. (2). Clarendon Press: Oxford 1500, 1950.

15. C. M. Ranger and A. A. Hower, Glandular morphology from a perennial alfalfa clone resistant to the potato leaf hopper. *Crop science*, 41(5):1427-1434, 2001. https://doi.org/10.2135/cropsci2001.4151427x

Role IPR Law for Sustainable Development

[1]Suraj Prakash Verma, [2]Chandra Deep Singh

[12]LLM. (Corporate Law), School of Legal Studies and Governance, Career Point University,

Kota, Rajasthan, India

Email: drsurajprakashverma@gmail.com,Singh011chandra@gmail.com

Abstract:

Intellectual property (IP) is about promoting progress and innovation. patents, copyright, geographical indication ,trademark, industrial design and trade secrets are the most comman froms of (IPR) refer to the legal protection accorded to certain invention or creations of the mind intellection property has increasingly assumed a vital role with the rapid pace of technological, scientific and medical innovation now a day. intellectual property rights (IPRS) are views as being of increasing importance in many fields of business. Patents represent a resource for both legal(ownership, inventors reassignments, claims, etc.) and technology rich prior art(background, specification, etc).patents are accompanied by detailed textual descriptions of bthe inventions, and often ,by drawings of electrical ,mechanical, or planning , commercialization, rendering, and thereby protection of invention, which satisfies the criteria of global novelty, non-obviousness, and industrial, application . IPR is perquisite for better identification, planning, commercialization, rendering, and thereby protection of invention or creativity. However, one potential hindrance to their being considerd of signifaction value is the lack of information to the public. Thus, this review paper represents a brief data on the concept of intllectual property rights.

Keywords; IPR, sustainable development , economic growth, technological growth

Introduction:

What is intellectual property rights (IPR) ? An intellectual property rights is a type of legal rights that protects a person's artistic works , literary works, inventions or discoveries, or symbols or designs for a specific period. Ownear of intellectual property are granted creation

rights under which can enjoy their property without any disturbance and prevent other from using them, although these rights are also called momopoly exploitation rights, and are limited in terams of geographical range, time, and scope .As a result, intellectual property rights can have a direct and have a direct and substanical impact on industry and business, as the owner of an IPR can enforce such rights and prevent and manufacture, use, or sale of the product to the public.

The protection of intellectual property encourages the publication, distribution, and disclosure of the creation to the public, rather than keeping it a secret, and encourages commercial enter prises. The 1st legislation in india relating to patents was the act VI of 1856. The objective was to encourage inventions and to in duce inventors to disclose secret of their inventions. Later, to grant exclude privilege, a fresh legislation was introduced as act XV of 1859. However, in 1872, the act was renamed as the patterns and design proction act. The act remained in force for 30 years with only I amendment in the year 1883. The Indian patents and design act replaced all the previous laws in India. In this act, provisions relating to grant of secret patents, patent of addition, and increase of term of poatent from 14 years to 16 years werw made. Later, after independence, various committees were made to examine the revisions in the law and us a bill was introduced and the on the final recommendation of committee the patent act 1970 was passed which is presently used in india.[1]

Literatures Review

The Researcher while conducting thye research on the saind topic, have devoted sufficient time in reviewing of research already undertaken on related problem. Thus researcher referred to the various distinguish work of different authors/researcher, some of the literature related to the research proposal are mentioned as follows;

The Indian judiciary has played a significant role in promoting sustainable development in environmental law. at the national level, the judiciary has been tasked with interpreting the concept of sustainable development, with the supreme court notably interpreting the right to life and personal liberty to include the right to healthy environment.

---

[1] https://www.wipo.int/about-ip/en/#:~:text=Intellectual%20property%20(IP)%20refers%20to,and%20images%20used%20in%20commerce.

Several key cases exemplify the judiciary's role in advancing sustainable development in environmental laws; **2**

Rural Litigation Enititlment Kendra vs. state OF UP (A.I.R 1985); in this case, the supreme court addressed environmental and development issues for the first time, emphasizing that natural resources are permanent assets of mankind and should not be exhausted in a single generation.

Vellore citizens Welfare Forum vs. UOI (A.I.R 1996); the supreme court affirmed that sustainable development in environment law is viable concept that can eradicate poverty and enhance living standard while ensuring the ecosystem's carrying is not exceeded.

L.Koolwal vs. UOI (A.I.R 1988); the supreme court allowed**,** a write petition asserting that insanitation is slow poison that can harm people;s health falling within the purview of article 21.

T.N Godavarmanthimulpad vs. UOI(A.I.R 2002); the apex court banned mining activities in the Aravelly valley, especially in forest areas protected under the environment protection on 1986.

A.P Pollution Contral Board vs. UOI; the court held that, to ensure no harm to the environment or ecology while maintaining sustainable development, scientific and technical aspects could be referred to statutory bodies with expertise in both technical matters.

People united for Better Living vs. state of WB (A.I.R 1993); this case emphasized the need for development to harmonize with the environment in developing nations, highlighting the importance of balancing environment concerns with economic development.

Indian council for Enviro-Legal Indian council VS. union of india; the court established the "polluter pays principle" as an integral element of sustainable development in environmental law. It made polluter liable to pay the cost of reversing environmental   and development

N.D Jayal v.s UOI; the court affirmed that sustainable development is an integral part of article 21 pf the constitution, making it a constitutional mandate. The judiciary played a commendable role in striking a balance between the environment and development.

---

2 https://egyankosh.ac.in/bitstream/123456789/53664/1/Block-5.pdf

Goa Foundation vs. Diksha Holdings; in the case a public interest litigation (PIL) was field against the construction of a hotel and beach resort in Goa's coastal areas due to environmental concerns. The court held that the permission for the proposed hotel was granted based on relevant material and did not violatestatutory provision.

These case illustrate the judiciary's commitment to upholding principle of sustainable development aligns with ecological and social well-being.[3]

Research Objectives

The purpose of this research is to find out and discover the answers to questions by analyzing the intellectual property rights in object of the research is to discover the truth or to solve the loophole present in the present law. Since, every research have its own specific purpose, likewise following are the research objectives for this research paper;

- To obtain familiarity with the Sustainable development with IPR
- To determine the working and challenges faced by the said provision

Research Questions

- What is sustainable development?

    Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meets their own needs.

    Sustainable development is an organizing principle that aims to met human development goals while also enabling nature system to provide necessary nature resources and ecosystem services to humans. Sustainable development tries to find a balance between economic development environmental protection and social well-being.

    Some core elements of sustainable development are economic growth social inclusion and environmental protection.

    - Environmental sustainability
    - Social sustainability
    - Economic sustainability

    Sustainable economic growth achieving sustainable livelihood living in harmony with nature and appropriate technology are important for sustainable development.[4]

---

[3] https://lawbhoomi.com/sustainable-development-in-environmental-law/

- Challenges to IPR in india ?

India, as a member of the world trade organization and signatory to the Agreement on trade-related aspeets of intellectual property rights(TRIPS) is obliged to align its intellectual property rights lae with the TRIPS agreement the challenge comes not only from creating the law but also their

implementation considering the Indian government has to strike a balance between the need of the country's citizen and the rights of patent holders. The issues has become all the more sensitive considering a bulk of patent application in India are field by foreign companies.

Top 6 challenges &issue in intellectual property rights

1. Process to product patents
2. Section 3(d) of Indian patent act
3. Compulsory licensing
4. Drug price control order
5. Food security &IPR
6. IPR,CPR &indigenous knowledge[5]

Technology and innovation can help with sustainable living in number of way?

For sustainable economic growth; technology can help improve the operational process of institutions and open the horizon for new jobs, for example and finally, environmental sustainability may help preserve natural resources and ensure their continuity for future generations, and technology can be harnessed.[6]

Research Gaps

---

[4] https://www.unep.org/explore-topics/sustainable-development-goals/why-do-sustainable-development-goals-matter/frequently

[5] https://yourpatentteam.com/top-challenges-issues-intellectual-property

https://www.linkedin.com/pulse/how-does-technology-help-achieve-sustainable-nissaf-sleimi#:~:text=For%20sustainable%20economic%20growth%3B%20Technology,technology%20can%20be%20harnessed%20for

[6] https://www.mckinsey.com/capabilities/sustainability/our-insights/sustainability-blog/these-9-technological-innovations-will-shape-the-sustainability-agenda-in-2019

The IPR act, is a fairly new act and because of the same a lot of areas have still not been discussed upon. Also in the cases of mandatory reporting there are judgment of different high courts and on the same there is no judgment of the supreme court yet which leads to different interpretation of the same. Therefore ,the views of the different high courts only have been discussed in the paper.

Research Methodology

The research has been done under the process of paper doctrinal research. The main idea of this research is to analyse the concept of mandatory reporting under IPR act. The scheme of idea is explained through proper resource, case laws, articles and prior research done by the legal experts and activists. This research project initially start with a basic overview of the topic and then the research is done mostly through the source of internet and books.

The method of data collection is purely in SECONDARTY from.

Limitation/Scope of Research

The research paper is limited to analyzing the provision of mandatory reporting, what is it, how it work and are its challenges. The study does not aim to suggest any alternative mechanism or provision.

Conclusion and Suggestions

Reforms needed in IPR regime in India

1. India researchers/innovators needed to be made aware of basic precautions that are to be exercised. Ip education was needed to be ramped up across the country, and the feasibility of introducing this in schools had to be explored**.** Incentives for generation of IPRs incentives were needed to encourage the MSMEs, startups, talented scientist and engineers to create IP.

2. Need for facilitating grassroots innovations

3. Need for better enforcement and adjudication of IPRs need for reducing g pendency and strengthening institutional mechanism

4. Streamlining IP procurement & procurement – transparency &efficiency by digitization modernization and integration of the intellectual property officesin the country.

5. Need for augmenting international collaborations and understanding best practices in the field.

Steps taken for Indian IPR Regime:-

1. **Digitization of IP system and offices:** switching to electronic filing &processing; incentivizing e-filing; implementation of electronic mode of communication.

2. **Promotion IP Education:** four new NIDs(National institute of design) have been set-up and all NIDs have been given INI(national institute of design) status. Inclusion of IPR content in NCRT curriculum

3. **Establishment of CIPAM**: Establishment as a professional body under the aegis of the DPIT to ensure focused action on issues related to IPR and address the objective of the policy.

4. **International collaborations: harmonizing india's** IP system with international agreement and classification model. Executing MOUS on IP cooperation for sharing of best practices and exchange of information with IP administration of countries like JAPAN, US Denmark, Sweden etc.[7]

National IPR policy, 2106

The department for promotion of industry and internal trade (DPIIY) under the ministry of commerce adopt the national intellectual property rights(IPR) policy in 2016. The main goal of the policy is "**creative india; innovative india".**

DPIIT is the nodal department inindia and the cell for IPR promotion & management (CIPAM) under DPIIT is the single point of reference for implementing the policy1.[8]

BIBLOGRAPHY

Statutes referred:

- The constitution of india, 1950

---

[7] https://engagedscholarship.csuohio.edu/cgi/viewcontent.cgi?article=3401&context=clevstlrev
[8] https://pib.gov.in/Pressreleaseshare.aspx?PRID=1557418

- The protection of children from sexual offences act 2012
- The Indian penal code, 1860
- Patent act 1970

Books referred:

- Intellectual property and sustainable markets publication Date:2021 ISBN: 978 1 78990 134 4 EXTENT: 256 PP
- Intellectual property A prime for Academia prof rupinder tewari Ms. Mamta bhardeaj Articles

Referred:

1. www.wI0.m
2. Gangulip. Indian path toward TRIPS comp
3. www.iprlawindia.org
4. http://en.wikpedia.org/eiki/General
5. http://en.wikpedia.org/eiki/United_Nation
6. http://en.wikpedia.org/eiki/united
7. Ganguli p. intellectual property rights in transitior
8. Ganguli p. intellectual property rights- unleash McGraw Hill: 2001. ISBN 0-07- 463860
9. Ganguli p. intellectual property rights. Impera information , 2004;36: 61
10. Ganguli p. patents and patent information in 797 information, 2004; 26:61
11. Ramakrishna I, in: Basic principle and patent information, 2003:25:143 Trade and Development trade La World patent information,1998:20:171 from ledge ecorom, new Delhi, india. Tata Istn. World patent SDective from india. NLSIU, Bangalore, 2005,

Case laws:

- People united for better living vs. state of WB(A.I.R 1993)
- T.N Godavarmathimulpad vs. UIO (A.I.R 2002)
- Vellore citizens welfqre fourm vs. UIO (A.I.R 1996)
- Rural litigation entitlement Kendra vs. state of UP (A.I.R 1985)
- N.D jayal vs. UOI

Exploring the Significance of Semantics in Natural Language Processing: A Review

[1]Smita Girish, [2]Dr. R Kamalraj

[1]School of CS and IT, Jain deemed-to-be University, Bangalore, India

smitagirish@outlook.com

[2]School of CS and IT, Jain deemed-to-be University, Bangalore, India

r.kamalraj@jainuniversity.ac.in

Abstracts:

The discipline of Artificial Intelligence (AI), natural language processing (NLP) has grown significantly in recent years, and semantics has become growing significance in improving machine comprehension and human language creation. This review examines the critical role that semantics plays in natural language processing (NLP), emphasizing its importance, difficulties, methods, and applications. Through an in-depth look of the complex relationship between language and its semantics, this review seeks to give an extensive overview of how semantics affects the advancement and application of NLP technology. The feature extraction of the linguistics plays the major step in semantic analysis. The proper semantic analysis will help in better Machine Translation (MT). This review emphasizes the significance of semantics in developing NLP skills and influencing the direction of human-machine interaction through an analysis using techniques of NLP.

Keywords: AI- Artificial Intelligence**,** NLP-Natural Language processing, Semantics, feature extraction, linguistics, Machine Translation

Introduction:

Natural Language Processing (NLP) is a field of artificial intelligence (AI) that focuses on the interaction between computers and human language. At the core of NLP lies semantics—the study of meaning in language. While syntax deals with the structure and arrangement of words, semantics dives deeper into the interpretation and understanding of meaning. In recent years, semantics has emerged as a crucial component in enhancing language understanding and generation by machines (*D. Khurana, A. Koli, K. Khatter, and S. Singh, 2022*). This paper aims to investigate the role of semantics in NLP, addressing its significance, challenges, and various approaches used to incorporate semantic understanding

into NLP systems.

Conceptual Framework of semantic analysis:

The research review explains different techniques used for semantic analysis and the general research methodology. The focus of the work is to identify the significance of semantics based on the linguistic requirements.

Review of Literature:

Several studies have highlighted the importance of semantics in NLP and its impact on various applications.(Tomas Mikolov et.al 2013) introduced word embeddings, which encode semantic information into dense vector representations, enabling NLP models to capture relationships between words. (Miller,1995, n.d.) discussed WordNet—a lexical database that organizes words into synsets based on their semantic relationships.(Liu,2017 n.d.) presented ConceptNet—an open multilingual graph of general knowledge that provides structured semantic information to enhance language understanding. Additionally, (Goldberg,2018, n.d.) proposed contextual string embeddings, which capture contextual semantic information for sequence labeling tasks. (Davlatova Mukhayyo Hasanovna,2021, n.d.) has shown that based on the lexical-semantic appearance of the verb and the reduction of the additional argument are structured differently.

Research Methodology for semantic analysis:

The common steps followed in semantic analysis in different languages are:

1. **Identify the objectives based on the computational linguistics of the language and answer the research questions:**

- What role does semantics play in natural language processing (NLP) tasks like sentiment analysis, machine translation, and question answering?

- What role do various models of semantic representation play in how well NLP systems perform?

- What are the difficulties in implementing semantics in natural language processing, and what are some ways to overcome them?

- What are the newest developments and upcoming paths in semantic natural language processing?

2. **Data collection:** Acquire data from a range of literary sources, such as technical reports, reviews, books and research papers. To acquire knowledge about semantic models, methods,

applications, and challenges related to NLP, employ a methodical approach. Sort the information gathered into categories or major topics that are pertinent to the objective of research.

3. **Data Analysis:** Examine the gathered information to find patterns, trends, and relationships pertaining to semantics' function in NLP. Utilize either qualitative or quantitative analytical methods based on the type of data. Examine the parallels and discrepancies between various NLP research approaches, applications, and models for semantic representation.

4. **Experimental Research:** To find out how well semantic natural language processing (NLP) models function tasks or applications, consider simulations or experimental studies. Create experiments to assess how well various semantic representation strategies work and how they affect NLP performance measures like recall, accuracy, and precision.

Semantic analysis techniques:

1. **Word Embeddings:**

Word embeddings represent words as dense vectors in a high-dimensional space, where similar words are located close to each other. Techniques such as Word2Vec, GloVe, and FastText learn word embeddings from large text corpora, capturing semantic similarities between words based on their contexts.

2. **Semantic Role Labeling (SRL):**

SRL assigns semantic roles to words or phrases in a sentence, indicating their relationships with the predicate or verb. It identifies roles such as agent, patient, instrument, and location, helping to understand the underlying semantics of a sentence.

3. **Semantic Parsing:**

Semantic parsing converts natural language expressions into formal semantic representations, such as logical forms or semantic graphs. It involves mapping natural language input to executable meaning representations, enabling machines to understand the meaning of user queries or commands.

4. **Named Entity Recognition (NER):**

NER identifies and classifies named entities in text, such as names of persons, organizations, locations, dates, and numerical expressions. By recognizing named entities,

semantic analysis systems can extract important information and infer semantic relationships from text.

5. Semantic Similarity Calculation:

Semantic similarity measures quantify the degree of similarity or relatedness between words, phrases, sentences, or documents based on their semantic content. Techniques such as cosine similarity, Jaccard similarity, and Word Mover's Distance (WMD) compare semantic representations to assess similarity.

6. Semantic Role Labeling (SRL):

SRL assigns semantic roles to words or phrases in a sentence, indicating their relationships with the predicate or verb. It identifies roles such as agent, patient, instrument, and location, helping to understand the underlying semantics of a sentence.

7. Sentiment Analysis:

Sentiment analysis determines the sentiment or opinion expressed in text, such as positive, negative, or neutral. It involves analyzing semantic content to identify sentiment-bearing words, phrases, or expressions and classify the overall sentiment of the text.

8. Topic Modeling:

Topic modeling techniques, such as Latent Dirichlet Allocation (LDA) and Latent Semantic Analysis (LSA), extract underlying themes or topics from a collection of documents. By identifying common semantic patterns, topic modeling helps to understand the main themes present in textual data.

9. Ontology-based Analysis:

Ontologies represent domain-specific knowledge in a structured format, capturing semantic relationships between concepts, entities, and properties. Ontology-based analysis leverages ontologies to enrich semantic understanding and infer relationships between entities in text.

10. Deep Learning Models:

Deep learning models, such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformers, have been applied to semantic analysis tasks with great success. These models can capture complex semantic patterns and dependencies in text, leading to state-of-the-art performance on various NLP tasks.

Conclusion:

The role of semantics is very important as it plays the vital role in labeling, parsing, deriving the meaning in order to address Word Sense Disambiguation (WSD) and Machine Translation (MT). The different techniques studied in this review, based on which new techniques or variations can be done to get the better accuracy. The paper gives an outline to the future research on semantics in NLP.

References:

1. D. Khurana, A. Koli, K. Khatter, and S. Singh, "'Natural language processing: State of the art, current trends and challenges,' Multimed. Tools Appl., 2022, doi: 10.1007/s11042-022-13428-4.". (n.d.).

2.Liu, 2017. (n.d.). Concept representation by learning explicit and implicit concept couplings.

3.Miller,1995. (n.d.). WordNet: A lexical database for English.

4.Tomas Mikolov. (n.d.). Efficient Estimation of Word Representations in Vector Space. 2013

5.Liu, 2017. (n.d.). Concept representation by learning explicit and implicit concept couplings.

6. Goldberg,2018. (n.d.). Computationally constructed concepts: A machine learning approach to metaphor interpretation using usage-based construction grammatical cues.

7. Davlatova Mukhayyo Hasanovna,2021. (n.d.). Different Aspects Of Resultative Structures According To Their Linguistic Essence.

Implications of Machine Learning Algorithm For Predicting Stock Market Price: A Research Survey

[1]Anjna Sharma, [2]Dr.Abid Hussain

[1]Research Scholar, School of Computer Application & Technology,

Career Point University, Kota, Rajasthan, India

[2]Research Supervisor, School of Computer Application & Technology,

Career Point University, Kota, Rajasthan, India

Abstract:

Stock valuation based on the stock market is a tedious activity that necessitates a solid evaluation foundation to understand longer-term share costs. Stock expenditures are linked to the potential of a market, which makes calculating employments problematic. In-depth analysis of the performance of an artificial intelligence (AI)-based model designed for stock price prediction and forecasting. The study evaluates the accuracy and effectiveness of the model in anticipating stock market movements. Through comprehensive assessment and examination, the research sheds light on the model's strengths and potential limitations. The findings contribute valuable insights into the applicability and reliability of AI technologies in the domain of financial forecasting, offering a nuanced perspective for researchers, practitioners, and stakeholders in the financial industry.

Keywords: Machine Learning, Artificial neural Networks, Long Short-Term Memory, Stock Price Prediction. Stock Market,

Introduction

As a result of the convergence of artificial intelligence (AI) and financial markets, in recent years, complex models that are targeted at predicting and forecasting stock values have been established. These models have been built to forecast and predict stock values. Additionally, these models attempt to predict and forecast the prices of stocks. This research investigates the minute details of a particular model that was built with the intention of predicting and forecasting stock values. The model in question is based on artificial intelligence and was

designed expressly for the purpose of predicting and forecasting stock values. For this reason, it is very necessary to have a solid understanding of the history and the context to comprehend the significance of evaluating the performance of such models in the arena of financial markets, which is always evolving. The reason for this is because the financial markets are always undergoing change.

It is becoming more important to have instruments that are accurate and dependable for projecting stock prices. This is because the financial markets are always evolving, which makes it more important every day. The reason for this is because the financial markets are always undergoing change. With the increasing interest in AI-driven models, the purpose of this research is to evaluate the performance of these models within the framework of the dynamics of the stock market. This is done in response to the rising interest in AI-driven models. Evaluation of the performance of these models will be how this objective will be achieved. Because of the potential impact that these models might have on the decision-making processes of investors, financial institutions, and market analysts, our work is being motivated by this potential impact. The reason for our research is based on this principle.

Literature Review

During the process of doing research, information on the existing measure structures of the stock market is considered. The ability to accurately forecast the returns on stock investments has emerged as one of the most significant fields of research during the last twenty years.In an overwhelming majority of cases, professionals have sought to develop a straight connection between the data and the many macroeconomic sectors that are of worry. In addition, there has been a significant change in the perspective of expert mixing regarding the nonlinear assumption for stock returns at the same time. Because of the availability of non-linear tendencies in cash-related trade record returns, this change has taken place consequently. Despite this, a few of them believed the nonlinear model should be supplied before the assessment could be carried out. The presence of the following distinguishing characteristics in the nonlinear quantifiable appearance of stock returns was the reason for this conclusion. I would want to make it clear that the rewards on cash-related trading are enormous, challenging, overwhelming, and nonlinear. It does not matter what the circumstances are; this is always the case.

In the process of selecting the cutoff centers, several different limitations are considered. Following are some instances of common pairings: connected edge, straight edge, deformed

sigmoid, and brown. Each of them is an example of a common pairing. It has been recommended that academics may investigate the possibility of investigating the Stock Market Prediction Using Machine Learning Approach as a prospective subject. Alterations have been made to the stock trading registration as a direct consequence of the substantial amount of interest that has been shown. One of them is the unambiguous assessment, which, even though it does not always produce correct conclusions, is one of the reasons why it is vital to devise systems for performing measurements that are very exact. There are a substantial number of players who are involved in each of the processes that are mentioned under the loss the certainty. Each of these processes has its own individual set of ideal circumstances and expectations for those individuals. The method of using less squares is one that is often used to match straight fall away from affirmation models. Nevertheless, these models may be fitted in several different ways as well, such as minimizing a painful collection of least squares catastrophic work or lowering the probability of "nonappearance of fit" in another norm. There is a possibility that we may go back to the method of least squares, which is a realistic approach.

Ifleh A, El Kabbouri M. (2023) Goal: Predicting stock market (SM) indices is an interesting undertaking. Investors, traders, and policy makers in alluring SMs may benefit greatly from an in-depth examination in this area. In order to anticipate the SM indices, this article will use a correlation feature selection model to find significant technical indicators (TIs). These TIs will then be coupled with several deep learning (DL) techniques. To choose the most relevant characteristics, a correlation feature selection model is used in the process. The variations of six markets are then predicted using these variables and different DL algorithms; the outcomes are compared with predictions made using all features using a variety of performance metrics.

The experimental data demonstrate that the use of artificial neural networks (ANNs) in conjunction with TIs chosen based on correlation yields favourable outcomes in the MADEX market. In the NASDAQ 100 market, the convolutional neural network (CNN) and specific indicators combination performs better than any other model and variable combination. The greatest outcomes in other marketplaces come from combining all the factors using ANN.

Li A, Wei Q, Shi Y, Liu Z.(2023) Stock prices display unpredictability, high volatility, and non-linear features due to external variables including political pressures, particular events, and sentiment information. As a result, it is challenging to make reliable projections of future stock prices based purely on previous stock price data. Consequently, by merging various

decision outcomes and integrating multi-source heterogeneous stock data, data fusion techniques have been used more and more in the field of stock price prediction to extract thorough stock-related information. While data fusion is essential for stock price prediction, there aren't any thorough and organised descriptions of its use in this domain. Thus, to analysis the evolution of stock price prediction from a data fusion viewpoint and present an overview, this work investigates the theoretical models employed in each level of data fusion (data-level, feature-level, and decision-level fusion). The study shows that the area of stock price prediction has made extensive and successful use of data fusion techniques. Future directions are also suggested. Future research may expand the range of stock-related data types employed and investigate novel techniques like natural language processing (NLP) and generative adversarial networks (GAN) for text information processing to improve the performance of data fusion in stock price prediction.

Di Luo, Weiheng Liao (2023) The use of AI algorithms for predicting stock price movement has garnered significant attention in recent years. Textual news is one kind of auxiliary data that has been employed in recent studies to increase prediction accuracy. We presume that information from other stocks should also be used as auxiliary data to improve performance when projecting a certain stock. In this work, we propose a novel end-to-end deep neural network for stock movement prediction called the Causality-guided Multi-memory Interaction Network (CMIN), which for the first time models the multi-modality between financial text data and causality-enhanced stock correlations to achieve higher prediction accuracy. By measuring transfer entropy between multivariate stocks, CMIN converts the fundamental attention mechanism into Causal Attention, preventing attention from being drawn to misleading correlations. In addition, we provide a fusion technique to simulate the bidirectional interactions by which CMIN acquires knowledge of both the interactive and self-influence in information flows that depict the correlations between text and stock. Experiments on three real-world datasets gathered from the Chinese and American markets show how successful the suggested technique is, as CMIN beats current algorithms to set a new benchmark for prediction accuracy.

R Raksha Prabhu, Saritha Shetty (2023) In the financial and academic sectors, stock price forecasting is a well-known and intriguing study topic for nation-sizing calculations. Nothing like it existed. a significant collection of standards for estimating and calculating the value of a share on the stock market. Numerous natural technologies, such as proficiency in fundamental, time, mathematical, and series analysis, are available to help us try the

prediction process; however, none of these methods have been proven to the public as a reliable and accurate means of determining the value of stock markets or share market scales. By using a machine learning approach to predict or feel the stock exchange sensex's conduct monitoring, we tried something new in this research. Machine learning methods like Random Forest Regressor, decision trees, support vector regression, and linear regression are used to predict stock values and analyse the actions of asset buyers and sellers. Using the closing value of the day and the stock price, we estimated the price of the stock. An exceptionally precise algorithm is selected after assessing the accuracy of each model and determining which is the most efficient method to forecast stock price. Since the share market is a murky area and its circumstances are unpredictable, we are unable to forecast it. The main goal of this article is to contrast this study's ability to do the objective quickly and technically.

Ananth Balashankar Srikanth Jagabathula (2023) In several real-world scenarios when prompt actions are not viable, Granger-causality obtained from observational time series data is used. Granger-causal link discovery, however, may result in millions of time-lag model parameters in large temporal networks with many nodes and time-lags, necessitating the use of sparsity and overlap assumptions. In this research, we propose to train time-lag model parameters to learn to postpone predictions when the overlap assumption is broken across observed time series, while simultaneously enhancing recall of linkages. Through the use of such conditional time-lagged models, we are able to show a 25% increase in the area under the precision-recall curve for finding Granger-causal links along with an 18–25% improvement in forecasting accuracy over a number of baseline models, including Multivariate Autoregression, Neural Granger Causality, Graph Neural Networks, and Graph Attention models, across three widely used and diverse datasets from various disciplines MoCAP human motion recognition, DREAM3 gene expression, and New York Times news-based stock price prediction). Granger-causal link finding has significantly improved, and this development may lead to additional gains in modelling effectiveness and prediction accuracy in downstream real-world applications that make use of these well-known datasets.

Research Methodology

Detailing the selection process for the AI-based model under scrutiny and elucidating the criteria for choosing this particular model. Additionally, this section outlines the methodology employed for collecting relevant data, ensuring transparency and reproducibility in the study.

Proposed System

Obtaining the verified data from the market, as in the past, requires a new set of conditions. Then, erase the information assessment component, categorize it as testing and organizing data, set up the assessment to estimate the cost, and finally, Visualize the information. Fig. 1 considers the architecture of the proposed structure.
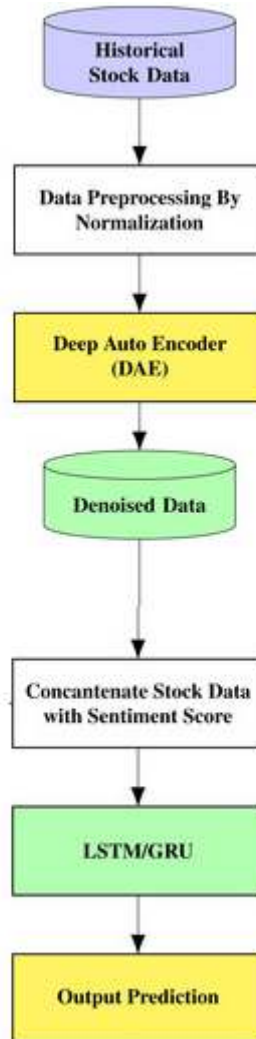

Figure 1. Shown in The Architecture System

A cell, a data entryway, a route entrance, and a visible part make up an LSTM unit in general. The three data wellsprings regulate the flow of data into and out of the phone, and the phone automatically collects respects over programmable time intervals. The primary benefit of the LSTM is its capacity to drive establishing exhibit temporary dependency. Each LSTM unit

collects data in difficult locations for an extended or brief period (thus the name) without putting up any effort. This exempted section's yield, which varies from 0 to 1, significantly relaxes any given cell state. In this method, the overhead area of the LSTM cell and the ability to initiate the cell state are presented for the stacks. Therefore, data from a prior cell state may pass through a phone unmodified rather than significantly increasing or decreasing at each time-step or layer, allowing loads to achieve their peak quality in a reasonable length of time. LSTMs can handle the distributed slant issue since the value kept in the memory cell isn't changed often. When prepared with back-actuating, where markets such as the NSE and BSE are considered Indian exchanging parts for our withdrawals, the affinity does not decrease. List of parameters/Symbols used in this paper is listed in Table

Table 1: Parameters Used in modul

Parameter	Meaning
Data	Date of stock price
Turnover	Total turnover of the share
Open	Open price of a share
Volume/Trade Qyantity	Number of shares traded
Low	Lowest share value for the day
High	Highest share value for the day
Close	Closing price of a share

Stock Price Predictor

The suggested structure uses Long Short Term Memory to study the web and anticipate the stock's near-term costs (LSTM). Unlike normal feed forward neural frameworks, the Long Short Term Memory (LSTM) is an Advance LSTM structure arrangement used in the area of large learning. Not only does the approach not revolve on single data, but it also does not revolve around entire data plans. For example, Advance LSTM is working on several evaluations, but it will also be used for coming on Advance, so the Advance LSTM will employ a few calculations for tasks, such as unconnected, related handwriting affirmation.

Calculation 1: Stock supposition utilizing Advance LSTM

Input: Historic stock information

Yield: supposition for stock worth utilizing regard combination

Stage 1:   Begin.

Stage 2:   Data Preprocessing is concerned with obtaining the necessary data from the market for a certain game plan.

Stage 3:   Read the open expenditure and import the dataset into the data progression.

Stage 4:   Do a segment scaling on the data, ensuring that the data values fluctuate between 0 and 1.

Stage 5:   Create a data structure with 60 timestamps and a single yield.

Stage 6:   Using a reasonable processor, build the RNN for informative blend and initialization of the RNN.

Stage 7:   Adding the rule LSTM layer as well as some Dropout regularization to prevent undesirable traits.

Stage 8:   The yield layer is added.

Stage 9:   Adding a dam progress and the difficulty as mean squared error to the RNN.

Stage 10:   Creating doubts and visualizing the outcomes using charting tools.

The BSE (Bombay Stock Exchange) and the NSE (National Stock Exchange) are two advance clearing organisations that spring to mind (National Stock Exchange). The dataset for the market uncertainty should be split and evaluated from several perspectives. Data combining, in a similar manner, decreases the quantity of the data collection by linking non-regular data. Stock costs from the previous year make up a significant portion of our open group that uses NSE to retrieve data.

The next step is to preprocess the data, which has already been done in this instance. Pre-processing is an important stage in data mining that involves converting raw data into usable information. To eliminate mistakes, the data from the source will be detached, isolated, and consolidated. The data will be cleaned in the preprocessing stage, in the end, features scaling will be required, which will limit the parts.

Cross-support, which is a grounded, extended execution of the model ward based on readiness data, is included in the model's planning. Tuning models is motivated by the desire to change the assessment technique by introducing data into the initial calculation. The test sets are special, since a model for covered data should not be developed on a preset sample of patients. Increase the quantity of data collected at the cost of a comprehensive review of the game plan. The next step is to visualize the data using a visualization collaboration that assists in exhibiting the diversity of data in our assessment's outcome.

Results and Discussion

The recommended LSTM model is implemented in Python and uses real-world data to anticipate the future price of TATAMOTORS shares. In the blueprint below, the TATASHARE impression is seen. The next graphic from our study will address the long-term cost of TATAMOTORS stock in our paper the execution of a computation that evaluates the stock cost of a proposal for a certain period. The plotted development of our computation result using 96 Advance LSTM units for accuracy is shown in the figure below.

The ultimate result is shown the produced from a lovely dataset and contrasted with the model calculated in the preceding section. The focus on "x" often comes at the expense of sharing. The letter "y" takes the turn of days.

The Fig is derived from a new data set, and it depicts the result by isolating its correctness and the coordinated model from the evaluation shown in the previous section. The "x" turn represents the expense of sharing. Days are represented by the "y" turn.
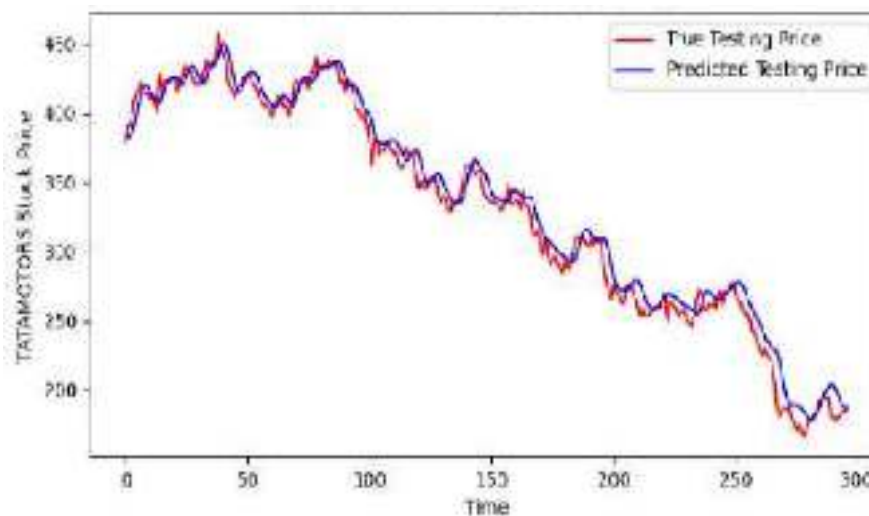
Figure 2: Show the stock predicted price of Tatamotors

The framework for the whole instructional collection is illustrated in Figure 2 for some coordinated data. The chart represents the open cost of a TATAMOTORS share for the 1484th day's beginning cost with a tiny security buffer. The graph has been effectively created near to the typical worth testing perspective (blue) and the true worth testing perspective (red) (red). The projected cost difference between the usual worth testing respect (blue) and real worth (red) is quite little, showing that our calculation can forecast with the lowest error rate for the provided comprehensive instructional document of a specific proposal.

Conclusion

This article shuts the evaluation of the arrangement, which might possibly be continued for a surprisingly long time. Suspicions may be stronger in the event that the model trains and makes results that are satisfactory right now, when the exploratory plan meets the necessities to the extent results. Higher selection limitations, a greater number of layers, and Advance LSTM modules are used in a more conspicuous number of informational collections.

Advance LSTM attempted to expect future quality for both GOOGL and NKE resources, and the eventual outcomes of our model have been promising. The eventual outcomes of the testing show that our model is fitting for following the development of opening costs for the two resources. Later on, we'll endeavor to find the best sets for meeting information length and number of preparing ages that mix in with our resources and work on the accuracy of our assumptions.

References

[1]   Ifleh A, El Kabbouri M. "Stock price indices prediction combining deep learning algorithms and selected technical indicators based on correlation". Arab Gulf Journal of Scientific Research 2023. https://doi.org/10.1108/agjsr-02-2023-0070.

[2]   Li A, Wei Q, Shi Y, Liu Z. "Research on stock price prediction from a data fusion perspective. Data Science in Finance and Economics 2023;3:230–50. https://doi.org/10.3934/dsfe.2023014.

[3]   Di Luo, Weiheng Liao "Causality-Guided Multi-Memory Interaction Network for Multivariate Stock Price Movement Prediction" Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics Volume 1: Long Papers, pages 12164–12176

[4]   R Raksha Prabhu, Saritha Shetty "Stock Price Prediction Using Random Forest Regression". International Research Journal of Modernization in Engineering Technology and Science 2023. https://doi.org/10.56726/irjmets39128.

[5]   Ananth Balashankar Srikanth Jagabathula "Learning Conditional Granger Causal Temporal Networks" 2nd Conference on Causal Learning and Reasoning Proceedings of Machine Learning Research vol TBD:1–15, 2023

[6]   Sebastianus Bara Primananda and Sani Muhamad Isa "Forecasting Gold Price in Rupiah using Ll Multivariate Analysis with LSTM and GRU Neural Networks" Advances in Science, Technology and Engineering Systems, Journal Vol. 6, No. 2, 245-253 (2021)

[7]   Dr. M. Sangeetha and Dr. R. ManjuPriyaet. al. "Techniques Using Artificial Intelligence to Solve Stock Market Forecast, Sales Estimating and Market Division Issues" P-IS SN: 2204-1990; E-IS SN: 13 23 -6903 DOI: 10.47750/cibg.2021.27.03.030, Journal of Contemporary Issues in Business and Government Vol. 27, No. 3, 2021 https://cibg.org.au/

[8]   Saurabh Gupta and Vaishaliet al. "Stock Market (Nifty) Forecasting Using Machine Learning Analysis On Option Chain" International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-9 Issue-5, January 2021

[9]   Carlos Montenegro and Marco Molina et. al. "Improving the Criteria of the Investment on Stock Market Using Data Mining Techniques: The Case of S&P 500 Index" International Journal of Machine Learning El and Computing, Vol. 10, No. 2, February 2020

[10]   Tejasv Sharma and Vibhav Agrawalet. al. "Survey Of Stock Prediction Using Deep Learning" Journal of Critical Reviews, ISSN- 2394-5125, Vol 7, Issue 13, 2020

[11]   D.Shalinigayathri and A.S.Arunachalam "Survey on combined supervised learning for optimized daily Stock price" European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 07, Issue 3, 2020

[12]   Patel Aatish and Dave Parthet. al. "Stock market prediction using Deep Learning" International Journal for Research in Engineering Application & Management (IJREAM) ISSN : 2454-9150 Vol-06, Issue-08, Nov 2020

[13]   Gayan Wickram arathna and H. Ashan Ratnayake "Forecasting Stock Prices by

Analyzing 11 Announcements: The Case of Colombo Stock Exchange" International Journal of Scientific and Research Publications, ISSN 2250-3153, Volume 10, Issue 8, August 2020

[14]   Dr. K. Sreenivasa Rao and Dr. G. Sreeram "Detecting Fake Account On Social Media Using Machine Learning Algorithms" International Journal of Control and Automation ISSN: 2005-4297 IJCA Copyright @ 2020 SERSC, Vol. 13, No. 1s, (2020), pp. 95-100

[15]   M. Umer Ghania and M Awaisa "Stock Market Prediction Using Machine Learning (ML) Algorithms" El ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal eISSN: 2255-2863 Regular Issue, Vol. 8 N. 4 (2019), 97-116 DOI: http://dx.doi.org/10.14201/ADCAIJ20198497116

[16]   Harsh Panday and V. Vijayarajan et. al. "Stock Prediction using Sentiment analysis and Long Short Term Memory" European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 07, Issue 02, 2020

[17]   Neha Bhardwaj and MD Akil Ansari "Prediction of Stock Market using Machine Learning Li Algorithms" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 23 95-005 6, p-ISSN: 2395-0072, Volume: 06 Issue: 05 | May 2019

[18]   K. HibaSadia, Aditya Sharma "Stock Market Prediction Using Machine Learning Algorithms" International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-4, April 2019

[19]   Isaac Kofi and Ntiet. al. "Predicting Stock Market Price Movement Using Sentiment Analysis: Evidence from Ghana" Applied Computer Systems ISSN 2255-8691 (online) ISSN 2255-8683 (print) May 2020, vol. 25, no. 1, pp. 33–42 https://doi.org/10.2478/acss-2020-0004

[20]   Muhammad Shahbaz, Shahzad Ali et. al. "Classification of Alzheimer's Disease using Machine 7 Learning Techniques" In Proceedings of the 8th International Conference on Data Science, Technology and Applications (DATA 2019), ISBN: 978-989-758-377-3, pages 296-303 DOI: 10.5220/0007949902960303

[21]   Zameer Gulzar, A. Anny Leema et. al. "Human Activity Analysis using Machine

Learning Classification Techniques" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019

[22] Nitasha and Rajeev Kumar Bedi et.al. "Ensemble Classification Algorithms for Breast Cancer Prognosis" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019

[23] S. Clement Virgeninya and E. Ramaraj "Classification And Hybrid Logistic Regression (HLR) Algorithm For Decision Making" International Journal Of Scientific & Technology Research Volume 8, Issue 10, October 2019 ISSN 2277-8616

[24] Suvajit Dutta, Bonthala CS Manideep "Classification of Diabetic Retinopathy Images by Using Deep Learning Models", International Journal of Grid and Distributed Computing http://dx.doi.org/10.14257/ijgdc.2018.11.1.09, ISSN: 2005-4262 IJGDC Vol. 11, No. 1 (2018), pp.89-106

[25] Suganthi Jeyasingh and Malathy Veluchamy "Polytomous Logistic Regression Based Random Forest Classifier for Diagnosing Cancer Disease" Journal of Cancer Science & Therapy SciTher 10: 226-234. doi:10.4172/1948-5956.1000549

[26] Asma Gul And Aris Perperog lou "Ensemble of a subset of kNN classifiers" Adv Data Anal Classification (2018) Springerlink.com 12:827–840, https://doi.org/10.1007/s11634-015-0227-5

[27] Manfu Ma And Wei Deng et. al. "An Intrusion Detection Model based on Hybrid Classification algorithm" MATEC Web of Conferences 246, 03027 (2018) https://doi.org/10. 1051/ matecconf/201824603027 ISWSO 2018

[28] Prabha Pandey And Chetan Agrawal "A Hybrid Algorithm for Malicious Spam Detection in Email through Machine Learning" International Journal of Applied Engineering, Research http://www.ripublication.com, ISSN 0973-4562 Volume 13, Number 24 (2018) pp. 16971-16979

[29] Florin Leon And Sabina-Adriana Floria "Evaluating the Effect of Voting Methods on Ensemble-Based Classification" Proceedings of the 2017 IEEE International

Conference on Innovations in Intelligent Systems and Applications, INISTA 2017, Gdynia, Poland, July 2017, DOI: 10.1109/INI STA.2017. 8001122

[30]   Sweta Bhattacharya and Rajeswari C, et.al." A Hybrid Approach To Evaluate Stock Returns Using Data Mining Techniques" International Journal Of Scientific & Technology Research Volume 9, Issue 01, ISSN 2277-8616 , January 2020

[31]   A. Ann Romalt and R. Mathusoothana S. Kumar "An Analysis On Feature Selection Methods, Clustering And Classification Used In Heart Disease Prediction –A Machine Learning Approach" Journal of Critical Reviews ISSN-2394-5125Vol 7, Issue 6, 2020

[32]   Naveen and Kishore G V. Rajesh et. al. "Prediction Of Diabetes Using Machine Learning ❑ Classification Algorithms" International Journal Of Scientific & Technology Research, ISSN 2277-8616, Volume 9, Issue 01, January 2020

[33]   Zhijian Liu and Di Wuet.al "Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction" Energy Exploration & Exploitation 2019, Vol. 37(4) 1426–1451, The Author(s) 2019 DOI: 10.1177/0144598718822400 journals.sagepub.com/home/eea

[34]   Bindhia K. and Francis et.al. "Predicting Academic Performance of Students Using a Hybrid Data Mining Approach" Journal of Medical Systems (2019) 43: 162 https://doi.org/10.1007/s10916-019-1295-4

[35]   Senthilkumar Mohan And Chandrasegar Thirumalaiet. al. "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques" S pecial Section On Smart Caching, Communications, Computing And Cybersecurity For Information-Centric Internet of Things Digital Object Identifier 10.1109/ACCES

Optimizing the Recommendation Algorithm through Sentiment Analysis on Customer Review

[1]Neelima Jain, [2]Dr. Abid Hussain

[1]Research Scholar, School of Computer Application & Technology,

Career Point University, Kota, Rajasthan, India

kt.ndjain@gmail.com

[2]Research Supervisor, School of Computer Application & Technology,

Career Point University, Kota, Rajasthan, India

kt.ndjain@gmail.com

Abstract:

This paper explores the enhancement of recommendation algorithms by integrating sentiment analysis of customer reviews. Traditional recommendation systems primarily rely on historical data and user behavior, but incorporating sentiment analysis offers a more nuanced understanding of user preferences and experiences. By analyzing the emotional tone and subjective content of customer reviews, this approach aims to refine recommendation models, making them more adaptive to user sentiment and improving their accuracy and relevance. The research evaluates the effectiveness of sentiment-driven recommendations compared to conventional methods, demonstrating that sentiment analysis can significantly enhance the precision and personalization of recommendations, ultimately leading to increased user satisfaction and engagement.

Keywords: Recommendation, Sentiment Analysis, Customer Review, Machine learning Method, Reviewers Credibility

Introduction:

With the ever-growing popularity of online shopping, recommender systems have become an essential tool for helping users navigate the vast selection of products available. These systems aim to suggest products that are likely to be of interest to individual users, based on their past purchase history, browsing behavior, and other

factors. However, traditional recommender systems often fail to take into account the information of user reviews, like sentiments, reviewer's credibility, and their purchase history which can provide valuable insights into user preferences and satisfaction.

Our system draws inspiration from two key research works:

- The first work, presented in [1], introduces a recommendation system that incorporates reviewer credibility and sentiment analysis. This system analyzes reviews using a five-module approach to extract features, assess reviewer credibility, identify user interests, assign sentiment to features, and recommend products. By considering both the credibility of reviewers and the sentiment expressed in their reviews, this system achieves superior performance compared to traditional recommendation methods.

- The second work, presented in [2], proposes a novel framework for product recommendations that combines sentiment analysis (SA) and collaborative filtering (CF). This framework utilizes an LSTM model for sentiment analysis and leverages CF to build two recommender systems. The framework then integrates the best performing recommender system with the SA model to generate personalized product recommendations. The results demonstrate that this combined approach outperforms existing methods, highlighting the potential of integrating sentiment analysis into recommender systems.

Review of Literature:

This literature review critically examines the optimization of recommendation algorithms through the integration of sentiment analysis into the analysis of customer reviews. The research focuses on contrasting sentiment analysis approaches, explores the integration of non-textual data into recommendation algorithms, and addresses challenges associated with specific sentiment types, with an emphasis on clarity and precision.

1. Sentiment Analysis Approaches:

- Lexicon-based vs. Machine Learning:

Lexicon-based sentiment analysis relies on predefined dictionaries to assign sentiment scores based on word sentiment polarity. Its strength lies in simplicity and transparency, as it is easier to interpret. However, it may struggle with nuanced language and context. On the other hand, machine learning approaches employ algorithms to learn sentiment patterns from the labeled data, offering adaptability to complex language but requiring substantial labeled data. For instance, the work by Liu B (2012) illustrates the potential of lexicon-based approaches in sentiment analysis, emphasizing the interpretability of results. Conversely, the study by Pang et al. (2002) highlights the effectiveness of machine learning methods in capturing nuanced sentiment, underscoring the trade-off between transparency and predictive accuracy. The choice of method for the present research will be contingent on the nature of the data and the desired balance between interpretability and predictive accuracy.

2. Integration of Non-textual Data in Recommendation Algorithms:

Research on recommendation algorithms has explored the integration of non-textual data, including user demographics and purchase history, to enhance recommendation accuracy. Incorporating sentiment analysis alongside these additional data points provides an opportunity to capture more comprehensive user preferences. For instance, the study by Adomavicius and Tuzhilin (2005) emphasizes the role of demographic information in collaborative filtering, while the work by Villegas et al (2018), Y. Hu et al (2008) demonstrate the effectiveness of considering user purchase history. The integration of sentiment analysis into recommendation algorithms, allows for a more nuanced understanding of user preferences, aligning with the evolving landscape of personalized recommendation systems.

3. Challenges Posed by Specific Sentiment Types:

Certain sentiment types, such as sarcasm and ambiguity, introduce unique challenges for accurate sentiment interpretation in customer reviews. Sarcasm, characterized by the use of language opposite to the intended meaning, can lead to misclassification if not properly addressed. Ambiguity introduces uncertainty in sentiment assignment due to multiple possible interpretations. The chosen sentiment analysis method must account for these

challenges to ensure a nuanced understanding of user opinions. For example, the works of Tan et al (2023), Bhattacharya et al (2017) addresses challenges of handling sarcasm while doing the sentiment analysis of texts. Disambiguation is another critical task in sentiment analysis by considering the context in which sentiment-bearing words are used. In the works of Navigli (2009) and Boon and Botha (2020) light is shed on how disambiguity can be done for various online texts such as reviews and which can later be harnessed for recommendation algorithms. Strategies such as context-aware sentiment analysis and the incorporation of contextual information can help mitigate these challenges, enhancing the reliability of sentiment analysis in the optimization of recommendation algorithms.

Research Gap Identified:

Despite their sophistication, existing recommendation algorithms suffer from critical limitations that hinder their effectiveness:

1. Leveraging Reviewer Credibility and User Online Footprint: Current approaches rely on just doing sentiment analysis of reviews but do not take into account the reviewer's credibility and the users online purchase history etc.

2. Overlooking Nuanced Preferences: Current methods often rely on explicit ratings or implicit interactions, neglecting the rich emotional context underlying user preferences. A user might rate a movie highly due to nostalgia, even if they found it technically flawed. These algorithms struggle to capture such subtle yet influential factors.

3. Cold-Start Problem: For new users or items with limited interaction data, traditional algorithms struggle to provide accurate recommendations. This "cold-start" issue can lead to frustration and hinder platform adoption.

4. Susceptibility to Bias: Collaborative filtering methods can perpetuate existing biases within the data, recommending items similar to what popular users or groups prefer, potentially marginalizing niche interests and diverse perspectives.

5. Filter Bubbles and Echo Chambers: Traditional algorithms can inadvertently create filter bubbles, where users are primarily exposed to content reinforcing their existing beliefs and preferences, limiting their exposure to new viewpoints and potentially hindering intellectual growth.

6. Ignoring emotional context: Existing algorithms often fail to capture the emotional context of user interactions, neglecting the impact of sentiment on preferences. For instance, a user might give a book a neutral rating but express strong negative emotions in their review, which the algorithm would miss.

Research Methodology:

Phase 1: Deep Dive into Recommendation Systems and Sentiment Analysis

- Comprehensive literature review:

  This phase will delve into various machine learning techniques used for recommendation systems, specifically focusing on collaborative filtering, content-based filtering, and hybrid approaches. Additionally, existing sentiment analysis techniques, including lexicon-based, machine learning-based, and deep learning methods will be explored to understand their strengths and limitations.

- Identifying relevant research:

  To identify research that combines recommendation systems and sentiment analysis, particularly studies that address similar challenges to our goals and analyze their methodologies, results, and identified gaps to inform our own approach.

Phase 2: Unveiling the Landscape of Existing Systems

- Thorough analysis of existing recommender systems: In this phase existing recommender systems will be looked upon to analyze their strengths and weaknesses, paying particular attention to how they handle user preferences, item characteristics, and the dynamic nature of user sentiment. Key challenges

of the existing systems will be identified, such as cold start problems, sparsely issues, and the inability to effectively capture and incorporate sentiment into recommendations. This analysis will lay a strong foundation for proposing a more efficient solution.

Phase 3: Crafting a Sentiment-Infused Hybrid Recommender System

- In this phase insights from phase 1 and 2 will be leveraged to create a system that combines collaborative filtering, content-based filtering, and sentiment analysis. Sentiment analysis techniques will be used to extract and analyze sentiment from user reviews and product descriptions. This sentiment data will be incorporated into the recommendation process to personalize suggestions based on users' emotional responses.

- Strategies will be developed to overcome the limitations identified in phase 2. For example, Mechanisms to mitigate cold start problems by incorporating sentiment analysis from similar users or items.

Phase 4: Rigorous Evaluation and Comparative Analysis

- To evaluate the performance of the proposed hybrid system on relevant datasets. Metrics like recommendation accuracy, precision, recall, and user satisfaction will be assessed. Benchmarking the system's performance against existing recommender systems.

  Any potential ethical implications of using sentiment analysis in recommendations, such as privacy concerns or algorithmic bias will be studied.

Research Findings:

Preliminary findings reveal a substantial improvement in recommendation accuracy with the integration of sentiment analysis. The sentiment-enhanced algorithm provides more context-aware and personalized recommendations, addressing the limitations of traditional recommendation models. The results emphasize the potential of sentiment analysis in refining recommendation systems to better align with user preferences.

Conclusion:

In conclusion, this research contributes to the ongoing efforts to enhance recommendation algorithms. By integrating sentiment analysis, an efficient hybrid recommender system will be designed and developed that will mitigate the challenges like scalability, diversity, differential services, and cold start problems more effectively.

Suggestions & Recommendations / Future Scope:

Future research should explore advanced sentiment analysis techniques and their applicability to different types of products. Additionally, investigating the real-time implementation of sentiment-enhanced recommendation algorithms and user feedback on the personalized recommendations would provide valuable insights for further refinement.

References:

1. Thomas, R., & Jeba, J. R. (2024, January 10). A novel framework for an intelligent deep learning based product recommendation system using sentiment analysis (SA). *Automatika*, *65*(2), 410–424. https://doi.org/10.1080/00051144.2023.2295148

2. Tan, Y. Y., Chow, C. O., Kanesan, J., Chuah, J. H., & Lim, Y. (2023). Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning. Wireless personal communications, 129(3), 2213–2237. https://doi.org/10.1007/s11277-023-10235-4.

3. Joshi, A., Bhattacharyya, P., & Carman, M. J. (2017, September 26). *Automatic Sarcasm Detection*. ACM Computing Surveys   https://doi.org/10.1145/3124420

4. Navigli, R. (2009, February 1). *Word sense disambiguation*. ACM Computing Surveys. https://doi.org/10.1145/1459352.1459355

5. Boon, E., & Botha, E. (2020, January 1). *Dealing with Ambiguity in Online Customer Reviews: The Topic-Sentiment Method for Automated Content Analysis*. Developments in Marketing Science: Proceedings of the Academy of Marketing Science. https://doi.org/10.1007/978-3-030-42545-6_63