3rd International Symposium

on

Smart Cities Challenges, Technologies and Trends

(SCCTT-2024)

29th November 2024

Organized by

Department of Computer Science & Engineering

Maharaja Agrasen Institute of Technology, New - Delhi

3rd International Symposium

on

Smart Cities Challenges, Technologies and Trends

(SCCTT-2024)

29th November 2024

Organized by

Department of Computer Science & Engineering

Maharaja Agrasen Institute of Technology

Agrasen Chowk, Sector-22, Rohini, New Delhi

Editorial Board

Inquiries to
The Editor
SCCTT-2024 3rd Symposium Souvenir
Maharaja Agrasen Institute of Technology,
Plot No 1 Rohini, Sector 22,
PSP Area, Delhi, 110086, India
Email: symposium.cse@mait.ac.in
Institute Website: www.mait.ac.in
Department Website: www.cse.mait.ac.in

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

Contents

Preface 5
Message from Respected Chairman Sir 6
Message from Respected Founder Chairman Sir 7
Message from Patron (Prof. (Dr.) Neelam Sharma) 8
Message from Dean Academics (Prof. (Dr). S. S. Deswal) 9
Message from General Chair (Prof. Namita Gupta) 10
Symposium Organizing Committee Members 11-16
Symposium Schedule 17

Accepted Papers

Support Vector Machine-Based Segmentation for Accurate Crowd Density Detection in Urban Spaces 18-34

Comparative Analysis of Data Redundancy Strategies for Wireless Sensor Networks in Smart Cities 35-54

Metaverse: Survey on future of internet and, its innovative applications 55-69

Optimizing Dehusked Arecanut Quality Segregation: CNN-Based Approach with Contrast Enhancement and Data Augmentation 70-81

Sentiment Analysis for Citizen Feedback in Smart Cities with XLNet-BiLSTM: Delhi Metro as a Case Study 82-94

Development of an interoperable Web Application to work in areas with limited connectivity in the Admissibility Processes of Activities, as a complement to Smart Cities 95-112

Enhancement of Accuracy in Soil Urea Estimation Using Machine Learning Tools 113-124

Crime Investigation Using Lip Reading 125-140

Smart Learning Ecosystems: The Influence of Online Skill-Based Education on Employability outcomes 141-158

Intelligent Web system for the management of pre-professional practices in a public university in Lima-Peru, using fuzzy logic in a Smart City environment 159-175

Smart Cities and AI: Leveraging User Data for Big Data Analytics Insights 176-185

About MAIT & CSE Department 186

Publication, Technical & Financial Partners 187

3rd International Symposium

on

Smart Cities Challenges, Technologies and Trends

(SCCTT-2024)

29th November 2024

Organized by

Department of Computer Science & Engineering

Maharaja Agrasen Institute of Technology

Agrasen Chowk, Sector-22, Rohini, New Delhi

Preface

We hereby are delighted to announce that Maharaja Agrasen Institute of Technology, Delhi, India, has hosted the eagerly awaited and much coveted "3rd International Symposium on Smart Cities Challenges Technologies and Trends" (SCCTT-2024) in Hybrid Mode. The third version of the symposium was able to attract a diverse range of engineering practitioners, academicians, scholars and industry delegates, with the reception of abstracts including more than 165 authors from different parts of the world. The committee of professionals dedicated towards the symposium is striving to achieve a high- quality technical program with track on Smart Cities, Smart Homes, Smart Hospitals, Smart Campuses, Green Computing, Smart Transportation System, Cyber Attacks, Smart light System, Smart Education, Smart urban pollution management, Smart urban waste management, IoT and Smart Applications. Therefore, a lot of research is happening in the above-mentioned track and its related subareas. More than 90 full-length papers have been received, among which the contributions are focused on theoretical, computer simulation-based research, and laboratory-scale experiments. Amongst these manuscripts, 11 papers have been included in the CEUR workshop proceedings after a thorough two- stage review and editing process. All the manuscripts submitted to the SCCTT-2024 were peer- reviewed by at least two independent reviewers, who were provided with a detailed review proforma. The comments from the reviewers were communicated to the authors, who incorporated the suggestions in their revised manuscripts. The recommendations from two reviewers were taken into consideration while selecting a manuscript for inclusion in the proceedings. The exhaustiveness of the review process is evident, given the large number of articles received addressing a wide range of research areas. The stringent review process ensured that each published manuscript met the rigorous academic and scientific standards. It is an exalting experience to finally see these elite contributions materialize into a book volume as SCCTT-2024 proceedings by CEUR workshop proceedings entitled "International Symposium on Smart Cities Challenges Technologies and Trends". All the contributing authors owe thanks from the organizers of SCCTT-2024 for their interest and exceptional articles. We would also like to thank the authors of the papers for adhering to the time schedule and for incorporating the review comments. We wish to extend my heartfelt acknowledgment to the authors, peer-reviewers, committee members and production staff whose diligent work put shape to the SCCTT-2024 proceedings. We especially want to thank our dedicated team of peer-reviewers who volunteered for the arduous and tedious step of quality checking and critique on the submitted manuscripts. The management, faculties, administrative and support staff of the college has always been extending their services whenever needed, for which we remain thankful to them. Lastly, we would like to thank CEUR workshop proceedings for accepting our proposal for publishing the SCCTT-2024 symposium proceedings

Prof. Namita Gupta, Dr. Deepak Gupta, Dr. Yogesh Sharma

Organizers, SCCTT-2024

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY
MAIT
उद्यमेन हि सिध्यन्ति कार्याणि न मनोरथै :

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY

(ISO 9001:2015 Certified)
APPROVED BY AICTE, MINISTRY OF HRD. GOVT OF INDIA
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)

Ref. No. MAIT/CSE/SCCTT-2024/2 Date 09/11/2024

Message from Founder Chairman & Chief Advisor's Desk

I am delighted to share my heartfelt appreciation for the upcoming 3rd International Symposium on Smart Cities Challenges, Technologies, and Trends (SCCTT-2024), scheduled to take place on 29th November, 2024, at the Maharaja Agrasen Institute of Technology in Delhi.

SCCTT-2024 promises to be a momentous event, uniting scholars, researchers, industry experts, and policymakers from across the globe. This symposium is designed to facilitate comprehensive discussions on the latest developments, challenges, and trends in the dynamic field of smart cities.

I extend my sincere gratitude to all the contributors, organizers, and partners who have devoted their time, expertise, and efforts to make this symposium a reality. I commend the Department of Computer Science and Engineering, particularly Prof. Namita Gupta, Head of the Department, for their dedicated and meticulous organization of this symposium.

As we anticipate the success of SCCTT-2024, I would like to express my best wishes to all involved, hoping for fruitful discussions, valuable networking opportunities, and the exchange of groundbreaking ideas.
Once again, congratulations to everyone involved, and I wish you all enduring success in your future endeavors.

Warm regards,

09. Nov. 2024

Dr. Nand Kishore Garg
Founder & Chief Advisor, MATES

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY
MAIT
उद्यमेन हि सिध्यन्ति कार्याणि न मनोरथै :

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY

(ISO 9001:2015 Certified)
APPROVED BY AICTE, MINISTRY OF HRD. GOVT OF INDIA
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)

Ref. No: MAIT/CSE/SCCTT2024/1

Date: 18/11/2024

Message from Chairman's Desk

It is with great pleasure that I acknowledge the remarkable efforts of the Department of Computer Science and Engineering at Maharaja Agrasen Institute of Technology in organizing the 3rd International Symposium on Smart Cities Challenges, Technologies, and Trends (SCCTT-2024), scheduled to take place on 29th November, 2024.

SCCTT-2024 will serve as a valuable platform for interdisciplinary exchange of ideas, fostering innovative solutions and contributing to the advancement of our understanding of the complex issues involved in transforming urban areas into smart, sustainable, and resilient cities.

I extend my heartfelt congratulations to Prof. Namita Gupta and her dedicated organizing team for their efforts in bringing together this momentous event. Additionally, I commend them on the successful creation of the souvenir magazine for SCCTT-2024.

May this symposium be a resounding success and may the endeavors of all those associated with SCCTT-2024 continue to prosper and contribute to the betterment of our society.

Wishing you all the success in your future endeavors.

Warm regards,

Vineet Kumar Gupta Lohia

Chairman, MATES

MAHARAJA AGRASEN CHOWK, SECTOR-22, ROHINI, DELHI- 110086
PHONES : 8448186931/35 TELEFAX:011-27582095 E-mail: mait@mait.ac.in www.mait.ac.in

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY
MAIT
उद्यमेन हि सिध्यन्ति कार्याणि न मनोरथै :

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY

(ISO 9001:2015 Certified)
APPROVED BY AICTE, MINISTRY OF HRD. GOVT OF INDIA
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)

Ref. No. MAIT/CSE/SCCTT-2024/3 Date 08/11/2024

Prof. (Dr.) Neelam Sharma
DIRECTOR, MAIT

MESSAGE

It gives me immense pleasure to know that a souvenir is being published by the Department of Computer Science and Engineering, MAIT for its 3rd International Symposium on Smart Cities Challenges Technologies and Trends organized on 29th November 2024.

Main objective of the symposium is to contribute significantly in area of research through high-quality research papers that present innovating ideas, novel approaches, developments, technologies, best practices, tools and techniques and provide future directions to young researchers and practitioners in the field of engineering.

Research material forming the contents of the souvenir will definitely be a developing tool to the readers.

I applaud the Symposium Chair Prof. Namita Gupta and her team to publish this issue. I wish them success.

Prof. (Dr.) Neelam Sharma
Director, MAIT

MAHARAJA AGRASEN CHOWK, SECTOR-22, ROHINI, DELHI- 110086
PHONES : 8448186931/35 TELEFAX:011-27582095 E-mail: mait@mait.ac.in www.mait.ac.in

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY

MAIT
उद्यमेन हि सिध्यन्ति
कार्याणि न मनोरथै :

(ISO 9001:2015 Certified)
APPROVED BY AICTE, MINISTRY OF HRD. GOVT OF INDIA
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)

Ref. No MAIT/CSE/SCCTT-2024/4 Date 08/11/2024

Prof. (Dr.) S. S. Deswal

Dean, MAIT

Message

It is gratifying to know that Department of Computer Science and Engineering, MAIT is organizing its 3rd International Symposium on Smart Cities Challenges Technologies and Trends (SCCTT-2024) organized on 29th November 2024.

Organizing such an event at this point of time reinforce our objective of developing an environment for the exchange of ideas towards technological development. I wish the Symposium would be able to deliberate on the current smart cities' challenges.

I extend my grateful wishes to Prof. Namita Gupta and her organizing team of SCCTT-2024 for making the Symposium a great success.

Prof. (Dr.) S. S. Deswal

Deam, MAIT

MAHARAJA AGRASEN CHOWK, SECTOR-22, ROHINI, DELHI- 110086
PHONES : 8448186931/35 TELEFAX:011-27582095 E-mail: mait@mait.ac.in www.mait.ac.in

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY

(ISO 9001:2015 Certified)
APPROVED BY AICTE, MINISTRY OF HRD. GOVT OF INDIA
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI)

MAIT
उद्यमेन हि सिध्यन्ति कार्याणि न मनोरथै :

Ref. No. MAIT/CSE/SCCTT-2024/5

Date 18/11/2024

Prof. Namita Gupta
Head of the Department, CSE
Message from Symposium Chair

The Department of Computer Science and Engineering takes great pleasure in organizing the 3rd International Smart Cities Challenges Technologies and Trends (SCCTT-2024) at Maharaja Agrasen Institute of Technology, Delhi. The Symposium is organized in association with various Academic, Industry, Technical, Finance and Publication Partners and is scheduled 29th November 2024. The Department of Computer Science and Engineering, established in the year 1999 has been the principal academic mover in the institute. The Department is known for its uniqueness in academic programs, state of art laboratories, research projects, industry collaboration, professional society activities, student and faculty development activities, and faculty awards. The symposium shall see contributions in varied tracks like Smart Cities, Energy-efficient Communications in Smart Cities, Green Computing, Big Data, and Analysis, Smart Transportation System, IoT and Smart Application. We are extremely happy to host distinguished personalities from academics as keynote speakers. These talks along with the presentations of selected papers are expected to be feast for the academics and research community.

On behalf of the organizing committee, I thank CSIR for sponsoring the symposium and being our Financial Partners, I thank CEUR Workshop Proceedings being our Partners. I thank the Management of Maharaja Agrasen Technical Education Society and the Director, MAIT for giving us the opportunity to organize this Symposium. Hope that the symposium leaves positive memories for you to cherish.

Prof. Namita Gupta
Symposium Chair and Head of Department
Computer Science and Engineering, MAIT

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

Symposium Organizing Committee Members

Chief Patron(s)

Dr. Nand Kishore Garg
Founder and Chief Advisor, MATES Chairman
Chancellor, MAU, Baddi, HP

Shri Vineet Kumar Lohia
MATES

Executive Committee

Shri S P Aggarwal
(Executive Chairman, IAS (Retd))

Shri Mohan Kumar Garg
(Joint General Secretary)

Sh. Gyanendra Srivastava
(Pro-Chancellor-MAU & CE-MATES, IAS (Retd.))

Shri Anand Gupta
(Treasurer)

Er. T. R. Garg
(General Secretary)

Shri. Rajnish Gupta
(Secretary)

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

Patron(s)

Prof. J.V. Desai
(Director General, MAIT)

Prof. (Dr.) Neelam Sharma
(Director, MAIT)

Prof. S.S. Deswal
(Dean, MAIT)

General Chair

Prof. Namita Gupta
(USICT, GGSIPU, Delhi)

Prof. Pravin Chandra
(MAIT, GGSIPU, Delhi)

Prof. Ahmed A. Elngar
(Beni-Suef University, Egypt)

Prof. Oscar Castillo
(Tijuana Institute Technology, Tijuana, Mexico)

Symposium Chair

Prof. Namita Gupta
(MAIT, GGSIPU, Delhi)

Prof. Amit Prakash Singh
(USICT, GGSIPU, Delhi)

Advisory Committee

Prof. Jafar a. Alzubi
(Al-Balqa Applied University, Salt-Jordan)

Prof. Vicente García Díaz
(University of Oviedo, Spain)

Prof. Oana Geman
(Universitatea Stefan cel Mare din Suceava, 720239 Romania)

Prof. Yu-Dong Zhangi
(University of Leicester, LE1 7RH, UK)

Prof. Sachin Gupta
(Dean (R&I), MAIT)

Prof. Utku Kose
(Suleyman Demirel University, Isparta, Turkey)

Prof. Sunil Kr. Mathur
(MAIT, Delhi)

Prof. V.N. Mathur
(MAIT, Delhi)

Prof. Amit Gupta
(MAIT, Delhi)

Prof. Rajveer Mittal
(MAIT, Delhi)

3rd International Symposium

on

Smart Cities Challenges, Technologies and Trends

(SCCTT-2024)

29th November 2024

Organized by

Department of Computer Science & Engineering

Maharaja Agrasen Institute of Technology

Agrasen Chowk, Sector-22, Rohini, New Delhi

Prof. M.L. Sharma
(MAIT, Delhi)

Prof. Puja Singh
(MAIT, Delhi)

Dr. Vaibhav Jain
(MAIT, Delhi)

Prof. Amita Goel
(MAIT, Delhi)

Dr. Pooja Gupta
(MAIT, Delhi)

Dr. Neeraj Garg
(MAIT, Delhi)

Dr. Vinay Kr. Saini
(MAIT, Delhi)

Dr. Monika Gupta
(MAIT, Delhi)

Organising Chair

Mr. Ashish Sharma
(MAIT Delhi)

Dr. Yogesh Sharma
(MAIT Delhi)

Dr. Deepak Gupta
(MAIT Delhi)

Dr. Sandeep Tayal
(MAIT Delhi)

3rd International Symposium

on

Smart Cities Challenges, Technologies and Trends

(SCCTT-2024)

29th November 2024

Organized by

Department of Computer Science & Engineering

Maharaja Agrasen Institute of Technology

Agrasen Chowk, Sector-22, Rohini, New Delhi

Convener(s)

Mr. Alok Sharma
(MAIT Delhi)

Ms. Savita Sharma
(MAIT Delhi)

Ms. Ruchi Goel
(MAIT Delhi)

Dr. Nisha Aggarwal
(MAIT Delhi)

Dr. Farzil Kidwai
(MAIT Delhi)

Co-Convener(s)

Dr. Ashish Khanna
(MAIT Delhi)

Ms. Garima Gupta
(MAIT Delhi)

Ms. Neetu Garg
(MAIT Delhi)

Ms. Mini Aggarwal
(MAIT Delhi)

Mr. Ajay Tiwari
(MAIT Delhi)

Ms. Karuna Middha
(MAIT Delhi)

Dr. Jyoti Kaushik
(MAIT Delhi)

3rd International Symposium

on

Smart Cities Challenges, Technologies and Trends

(SCCTT-2024)

29th November 2024

Organized by

Department of Computer Science & Engineering

Maharaja Agrasen Institute of Technology

Agrasen Chowk, Sector-22, Rohini, New Delhi

Technical Program Committee

- Prof. Oscar Castillo (Tijuana Institute Technology, Tijuana, Mexico)
- Prof. Yu-Dong Zhangi (University of Leicester, LE1 7RH, UK)
- Prof. Vicente García Díaz (University of Oviedo, Spain)
- Prof. Jafar A. Alzubi (Al-Balqa Applied University, Salt - Jordan)
- Prof. Utku Kose (Suleyman Demirel University, Isparta, Turkey)
- Prof. Oana Geman (Universitatea Stefan cel Mare din Suceava, 720239 Romania)
- Prof. Ahmed A. Elngar (Beni-Suef University, Egypt)
- Prof. Marcin Paprzycki (Polish Academy of Sciences, ul. Newelska,Warsaw, Poland)
- Prof. Lorenzo Donatiello (Univ. of Bologna, Italy)
- Prof. Jean-Luc Dugelay (Eurécom, France)
- Prof. Stefano Ferretti (University of Bologna, Italy)
- Prof. Manuel Perez Malumbres (Miguel Hernández University, Spain)
- Prof. Marco Roccetti (University of Bologna, Italy)
- Prof. Anil Kumar Ahlawat (KIET Group of Institutes, India)
- Prof. Balamurugan.B (ShivNadar University, India)
- Prof. K Sampath Kumar (AMET, India)
- Prof. Vishal Bhatnagar (NSUT, Delhi)
- Prof. Vikram Goyal (IIIT, Delhi)
- Dr. Inderjeet Kaur (Ajay Kumar Garg Engineering College, India)
- Parma Astya (Galgotias University, India)
- Rajesh Kumara (Galgotias University, India)
- Prashant Kumar (Mnnit Allahabad, India)
- Sandeep Sharma (Gautam Buddha University , India)
- Anita Mahajan (Vaishnav Engg. College, India)

3rd International Symposium on Smart Cities Challenges, Technologies and Trends (SCCTT – 2024)
Organised by
Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology, Delhi, India
(29th November 2024)

Symposium Schedule

Time	Program
9:30 am	**Registration** *Venue: outside room no. 114*
10:00 am – 11:20am	**Inauguration** **10:00am - 10:10am: Welcome and Saraswati Vandana** **10:10am - 10:20am: Introduction of SCCTT - 2024** **10:20am - 10:30am: Welcome Address by Prof. Namita Gupta, General Chair, SCCTT - 2024** **10:30am - 11:00am: Keynote Lecture on Smart Cities** **11:00am - 11:10am: Address by Prof. Neelam Sharma, Director MAIT, Patron, SCCTT - 2024** **11:10am – 11:20am: Vote of Thanks** *Venue: Room no. 114*
11:20am – 12:00pm	**Tea Break**
12:00 pm –1:00 pm	**Session (Offline)** **Session Chair: Prof. Rahul Kataria, DTU, Delhi** *Venue – Room No. 114*
1:00 pm –2:00 pm	**Lunch and Networking**
2:00 pm- 3:00 pm	**Session resume (Online)** **Session Chair: Prof. Rahul Kataria, DTU, Delhi**
3:00pm - 3:45pm	**Valedictory Session** **3:00 pm – 3:30 pm: Certificate Distribution ceremony** **3:30 pm – 3:45 pm Vote of Thanks** *Venue – Room No. 114*

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

Accepted Papers

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-1]: Support Vector Machine-Based Segmentation for Accurate Crowd Density Detection in Urban Spaces

Gourav Kalra[1], Rajeev Yadav[2], Satish Kumar Alaria[3]

[1]*M. Tech. Scholar, Department of CSE, Arya College of Engineering, Jaipur, Rajasthan*
[2]*Professor, Department of CSE, Arya College of Engineering, Jaipur, Rajasthan*
[3]*Computer Instructor, Education Department, Government of Rajasthan, Rajasthan*

Abstract

Estimating crowd density has become increasingly important in fields like public safety, event management, and urban planning. Accurate detection of crowd density helps in making informed decisions and ensuring safety in crowded areas. This study proposes a novel method for crowd density detection using segmentation and classification based on a Support Vector Machine (SVM). The method involves two key steps: crowd segmentation and density categorization. During segmentation, advanced image processing techniques like background removal and region-based segmentation extract crowd sections from input images or video frames. These segmented areas are then classified using an SVM model, known for handling complex data. The model is trained on a diverse dataset containing images with varying crowd densities. The approach captures crucial spatial and contextual information, and extensive testing on various datasets has demonstrated its accuracy and resilience in dynamic crowd scenarios. The proposed SVM-based method can be implemented in real-time, making it valuable for applications requiring quick decisions. This technique offers a reliable and efficient solution for crowd density detection, with significant implications for event management, public safety, and urban planning in congested environments.

Keywords

Crowd density detection, Support Vector Machine, crowd segmentation, image processing, real-time detection, region-based segmentation, urban planning, machine learning.

1. Introduction

The world has undergone rapid urbanization over the past two decades, leading to a significant increase in city populations. As cities become more crowded, the need for effective surveillance systems has grown, particularly to monitor people's movements and behaviors in public spaces, ensuring the safety and security of individuals and their possessions. Surveillance has become an integral part of maintaining public safety, with both public and private entities worldwide regularly employing video cameras for this purpose. However, traditional surveillance systems heavily rely on human operators, whose effectiveness can vary depending on their alertness and the available manpower. Given these limitations, modern surveillance is transitioning towards smart systems equipped with advanced technologies like intelligent video analysis, which enable automated decision-making without

* Corresponding author.
† These authors contributed equally.
gkalra144@gmail.com (Gourav Kalra); rajeevtpo@gmail.com (Dr. Rajeev Yadav); Satish.alaria@gmail.com (Satish Kumar Alaria)
0009-0008-4926-1929 (Gourav Kalra); 0000-0002-1976-4065 (Dr. Rajeev Yadav); 0000-0001-8298-1364 (Satish Kumar Alaria)

continuous human intervention.

Smart surveillance systems can be broadly categorized into two types: visual-based and multimodal. Visual-based systems utilize computer vision algorithms to process video data from cameras and drones in real time, offering solutions like facial recognition and license plate identification. On the other hand, multimodal systems integrate various data sources, including motion and audio sensors, alongside video data to provide comprehensive real-time insights. Companies like IBM and Intel have pioneered technologies that can detect traffic incidents, optimize routes, and even identify crime-related events using these advanced surveillance systems.

In today's world, smart surveillance plays a critical role, particularly in monitoring crowds. This becomes especially relevant during large public gatherings, where the potential for disasters, accidents, or criminal activity increases. Effective crowd control is vital in these scenarios, as seen in airports, concert venues, and religious gatherings. As crime, terrorism, and natural disasters rise, smart surveillance systems must rely on robust algorithms to manage and predict crowd behavior.

The analysis of crowd behavior is a key focus of this chapter. It begins by defining different types of crowds, highlighting their unique characteristics and behaviors in various contexts. A deeper exploration of collective crowd behavior from a psychological standpoint follows, offering insights into how crowds react in specific situations. From there, the discussion shifts to the challenges of analyzing crowd behavior through video footage, including the complexities involved in cognitive modeling for crowd behavior analysis. Ultimately, this chapter sets the stage for understanding the motivations behind this research and the primary contributions of the proposed approach.

Figure 1. Common applications of smart surveillance

A crowd is a large group of people gathered in one location, exhibiting a range of behaviors and attitudes. Based on movement patterns, crowds can generally be divided into two categories: dynamic and stationary. Dynamic crowds are constantly in motion and can be either organized or unstructured. In organized crowds, such as marathons or rallies, individuals move in the same direction, maintaining

consistent behavior over time. In unstructured crowds, such as those seen in airports or stadiums, individuals move in various directions with varying spatiotemporal characteristics. Stationary crowds, on the other hand, include audiences at rallies, concerts, or plays, where people remain in one place for a period of time.

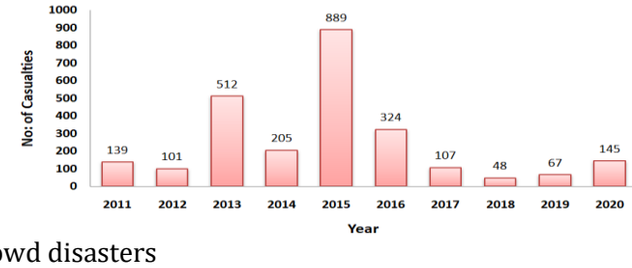


Figure 2: Statistics of crowd disasters

The characteristics of a crowd—such as its size, density, location, and time—are critical in understanding its behavior. Crowds can also be categorized into active and passive groups based on the behavior of their participants. While passive crowds primarily observe without engaging in activities, active crowds may exhibit behaviors ranging from aggression to panic or expressive actions, such as cheering at a concert or participating in religious events.

Analyzing crowd behavior is essential for smart surveillance systems, as it helps authorities understand crowd dynamics, develop control measures, and prevent crowd-related disasters. The behavior of a crowd is often influenced by the context in which it forms. For instance, in a shopping district, people might move peacefully alongside one another, while in a stadium, fans may express intense emotions in response to the game. These varying behaviors highlight the need for smart surveillance systems capable of monitoring and analyzing different crowd scenarios in real time. Crowd behavior is inherently complex, as it depends on the context and setting. Monitoring and understanding collective crowd behavior in both regular and emergency situations is challenging, particularly when individual identification is difficult in dense crowds. Over time, psychologists and sociologists have proposed numerous theories to explain crowd behavior. One of the earliest and most popular is Le Bon's Group Mind Theory, which suggests that crowd members lose their individual identity and are easily influenced by a leader. Freud's theories support the notion that individuals in a crowd open their unconscious minds, yet maintain control over their actions. McPhail's Pre-Disposition hypothesis posits that aggressive behavior in crowds is influenced by individual dispositions toward antisocial behavior. In contrast, the Emergent-Norm hypothesis suggests that crowds consist of people with common interests, leading to distinctive behavior patterns. These collective behaviors can often become impulsive, unpredictable, and volatile. Understanding these behaviors is crucial for developing smart surveillance systems that can anticipate and prevent crowd-related issues. Such systems must account for the social and psychological components of group behavior, including how crowd members concentrate their attention on a common cause, exchange ideas rapidly, and form homogenous groups based on shared beliefs and behaviors. Machine learning, particularly Support Vector Machines (SVM), is a key technology used in crowd behavior analysis. SVM models create distinct classes from input data features, enabling the classification of various crowd behaviors. Deep learning, especially Convolutional Neural Networks (CNN), is another powerful tool for crowd behavior research. CNN mimics the structure of neurons in the human visual cortex, allowing for the hierarchical processing of input data. Long Short-Term Memory (LSTM) networks, which resemble the brain's short-term memory, are also used to analyze and predict crowd behavior based on past events. These advanced AI models enable the system to learn from past examples, making it more effective in predicting crowd behaviors and

detecting anomalies. This research is motivated by the need to develop smart surveillance systems capable of detecting crowd anomalies, evaluating behaviors in real-time, and providing timely alerts. The current pandemic crisis has also highlighted the importance of monitoring crowd behavior to

ensure public safety, especially in terms of enforcing social distancing and detecting free-standing conversation groups. By combining video, audio, and other sensor data, this study aims to develop a comprehensive crowd behavior analysis system that can operate effectively in a variety of challenging scenarios.

In conclusion, the introduction of cognitive modeling and AI technologies into surveillance systems offers the potential to greatly improve crowd management, enhancing public safety and preventing disasters in crowded settings.

2. Related Works

Crowd behavior evaluation through computer vision techniques has been explored through various research studies, with each contributing to a broader understanding of how anomalies and movement patterns in large groups can be detected and analyzed. A review of these works highlights both advancements in this domain and the identification of gaps that future research must address. For instance, in 2007, Chen et al. proposed a framework for video event identification that proved essential for high-level video indexing and retrieval. This framework addressed challenges such as skewed data distribution and loose video structure, automating the determination of crucial thresholds that were typically manually set in conventional Association Rule Mining (ARM) techniques. The reduction in manual intervention in video analysis was a critical advancement towards fully autonomous video content analysis.

Following Chen et al., Su et al. (2013) introduced the Trajectory Segmentation and Multi-Instance Learning (TRASMIL) framework, which allowed for precise and adaptable local anomaly detection. This three-step method was found to outperform existing techniques in terms of identifying trajectories with local abnormalities. TRASMIL emphasized the importance of trajectory-based anomaly detection for accurately understanding crowd movement and behaviors. Similarly, Krishna et al. (2014) proposed a semantic video segmentation method that relied on One-Class Classification (OCC) techniques for identifying events through frame-by-frame processing. Their work highlighted the effectiveness of OCC in detecting unsupervised events, particularly through the use of Temporal Self-Similarity Maps (TSSMs), which were evaluated using a publicly available thermal video dataset. The use of OCC for unsupervised event detection opened new avenues for handling video data with minimal prior knowledge of the scene.

Building on these methods, Xia et al. (2015) introduced a dynamic time interval segmentation technique to improve item anomaly detection. Their segmentation approach dynamically validated the time interval length, grouping successive attack ratings. While effective, Xia et al. noted that the robustness of anomaly detection methods had received limited attention in terms of accuracy and consistency, pointing to a gap that future research must address. Meanwhile, Kaltsa et al. (2018) contributed by proposing an unsupervised method for scene analysis and anomaly detection in traffic video data recorded by stationary security cameras. By using local Hierarchical Dirichlet Process (HDP) models, Kaltsa et al. were able to achieve improved accuracy with lower computational costs, emphasizing the need for efficient solutions in processing large amounts of traffic video data.

Other researchers have approached the problem from a probabilistic standpoint. For example, Saligrama and Chen (2012) developed a probabilistic framework for identifying local spatiotemporal anomalies. This framework allowed for a more refined decision-making process by identifying ideal decision-making procedures based on score functions obtained from nearby neighbors' distances. Their work emphasized

the importance of spatiotemporal scales in accurately identifying anomalies. Thom et al. (2012) extended this concept to spatiotemporal anomaly detection using scalable aggregation and geolocated text visualization. They proposed a cluster analysis technique to automatically discover anomalies and presented these findings through a global map depiction. Their work demonstrated how scalable visualization could assist analysts in categorizing and evaluating event candidates on a global scale.
Chae et al. (2012) further refined the visualization of social media data with a visual analytics technique, which allowed users to extract significant subjects from a chosen collection of communications. By applying Latent Dirichlet Allocation (LDA) and visualizing topic time series, analysts could better understand abnormal events by identifying peaks and outliers in the data. Roshtkhari and Levine (2013) contributed with a probabilistic methodology that placed temporal and geographical constraints on video volumes, allowing for the identification of abnormal video configurations. Their approach, which avoided the need for motion estimation or background removal, proved particularly efficient for detecting rare events in video data.
In a related development, Cong et al. (2013) proposed an anomaly detection method that incorporated both spatial and temporal contexts. They introduced a region-based descriptor called Motion Context, which proved to be more reliable than statistical models when dealing with small training datasets. Their use of compact random projections sped up the search process, further enhancing the efficiency of the method. In 2013, Thida et al. developed a spatiotemporal Laplacian eigenmap technique to model crowd behavior and detect anomalies.

Their method, which identified both local and global anomalies, showcased the potential of regular crowd behavior modeling in accurately detecting abnormal crowd behaviors.
Yuan et al. (2014) took a different approach by developing a Structural Context Descriptor (SCD) to define crowd individuals, utilizing the potential energy function of particles from solid-state physics. Their SCD method used the 3-D Discrete Cosine Transform (DCT) to compute crowd SCD fluctuations and pinpoint issues through these variations. Xu et al. (2014) also focused on anomaly detection in complex crowd settings, using a hierarchical activity-pattern discovery framework. Their work factored in both local and global spatiotemporal contexts, creating an anomaly energy function that could quantify the abnormality of motion patterns. This method was particularly useful for detecting abnormal activity in densely packed crowds.
Continuing with anomaly detection in video monitoring, Li et al. (2015) proposed an unsupervised statistical learning framework for monitoring crowded environments. Their method, which relied on clustering and sparse coding to learn global and local activity patterns, utilized a multi-scale analysis approach to ensure precise anomaly localization. Wang and Xu (2016) advanced these techniques by developing a novel crowd video anomaly detection method based on spatiotemporal texture analysis. Their approach, designed for real-time applications, simplified machine learning procedures and demonstrated improved flexibility and efficiency compared to existing systems.
In 2017, Chong and Tay introduced a spatiotemporal architecture for anomaly detection, combining spatial feature representation with temporal changes in spatial features. This method proved to be effective for detecting anomalies in videos of crowded scenes. Omar et al. (2013) contributed by introducing an intrusion detection technique that detected normal behavior disturbances, signaling potential intentional or unintentional attacks. Their work explored both supervised and unsupervised methods for anomaly detection, emphasizing the importance of detecting disruptions in normal behavior patterns.
Du and Zhang (2014) developed an anomaly detection approach that utilized a reliable anomaly degree measure to increase the separability between anomaly pixels and background pixels. Their method divided pixels into potential anomaly sections and background sections, followed by discriminative information learning, highlighting the significance of feature extraction for accurate anomaly detection.
Gornitz et al. (2015) introduced a fresh approach to anomaly detection using a difference of convex

functions algorithm. Their method built a hidden Markov anomaly detector that extended the One-Class SVM and demonstrated improved performance across various datasets.

Wang et al. (2017) presented a sparse reconstruction-based method for detecting aberrant behavior, combining low-level visual features with causality analysis. By analyzing individual and group behaviors, they were able to detect abnormal interactions in multi-object settings. Plotnikov et al. (2018) focused on improving image classification performance through convolutional neural network (CNN) ensembles, showing how this approach could outperform both single CNN models and regular perceptrons in detecting abnormalities.

Sabokrou et al. (2018) introduced an unsupervised Fully Convolutional Network (FCN) for anomaly detection in videos. Their approach relied on temporal data and cascaded outlier detection, lowering computational complexity and improving both speed and accuracy. Anton et al. (2018) proposed a machine learning-based anomaly detection approach for detecting fraudulent traffic in Modbus and Transmission Control Protocol (TCP) connections. Their use of SVM, Random Forest, K-NN, and K-means clustering allowed for effective anomaly detection in an industrial scenario.

Wang and Xia (2019) applied deep learning to behavior detection, using a bag of vision words and the Agglomerative Information Bottleneck technique to compress vocabulary and minimize feature dimensions. Their sparse representation approach increased detection precision for deviant behavior. Garg et al. (2019) also leveraged deep learning in social multimedia to detect suspect flows, testing their method on a large-scale Carnegie Mellon University (CMU) dataset. Finally, Zhan Li and Jia Wei (2019) introduced the Inception-V3 neural network for feature extraction and classification, comparing its performance with traditional models like K-nearest Neighbor, random forest, and SVM, while Feizi (2020) developed a technique focused on maximizing the area under the ROC curve for hierarchical abnormal behavior detection, eliminating the need for manual labeling and offering a semi-supervised approach.

The literature on crowd behavior analysis demonstrates the continuous evolution of methods aimed at enhancing surveillance through anomaly detection. From trajectory-based techniques to deep learning and probabilistic models, researchers have developed increasingly sophisticated approaches to ensure real-time, accurate detection of abnormal behavior in crowds. These advancements have laid the groundwork for further research into the robustness and scalability of anomaly detection methods, while also identifying key areas for future exploration, such as improving computational efficiency and addressing issues like occlusion and multi-camera data integration.

3. Mathematical Modeling & Proposed Methodology

In the realm of image processing, feature extraction is pivotal for enhancing tasks like pattern recognition, face detection, and image classification. Features c an broadly be divided into two categories: general features such as color, texture, and shape, and domain-specific features like object detection or human face recognition. The efficiency of image annotation frameworks hinges on the ability to represent semantic concepts through low-level image features, which form the foundation of multimedia information retrieval, object recognition, and image annotation. In both Content-Based Image Retrieval (CBIR) and Automatic Image Annotation (AIA), key image features such as color, texture, and shape are employed to extract meaningful data. While CBIR primarily focuses on visual aspects of an image, AIA incorporates high-level concepts that better reflect the image content, addressing the challenge of locating images in large datasets. Hence, this research integrates both low-level features and high-level semantic concepts to improve image retrieval, focusing particularly on texture and shape as central features for efficient image annotation. Feature extraction is a dimensionality reduction process where the image is transformed into a feature set, representing its high-level characteristics. By condensing the image data into a feature vector, the system can quickly and accurately identify patterns within an image. For computational efficiency, a robust feature extraction system is required, and combining low-level and

high-level semantic concepts provides better retrieval accuracy. The proposed system uses fused feature extraction, employing texture and shape features to enhance the accuracy of image retrieval and reduce system complexity. This methodology combines multiple features to provide more accurate image information, avoiding the errors that might arise from relying on a single feature. In this study, the Haralick and Tamura texture features are fused with shape features, significantly improving image retrieval performance and reducing processing time. Image feature extraction forms the backbone of image retrieval systems, with features classified into two main categories: general features and domain-specific features. General features, including color, texture, and shape, describe the overall content of the image, while domain-specific features, such as face recognition or object detection, require specialized knowledge and fine-tuning. Low-level features like color and texture represent the visual aspects of an image, while high-level features correspond to semantic keywords or concepts.

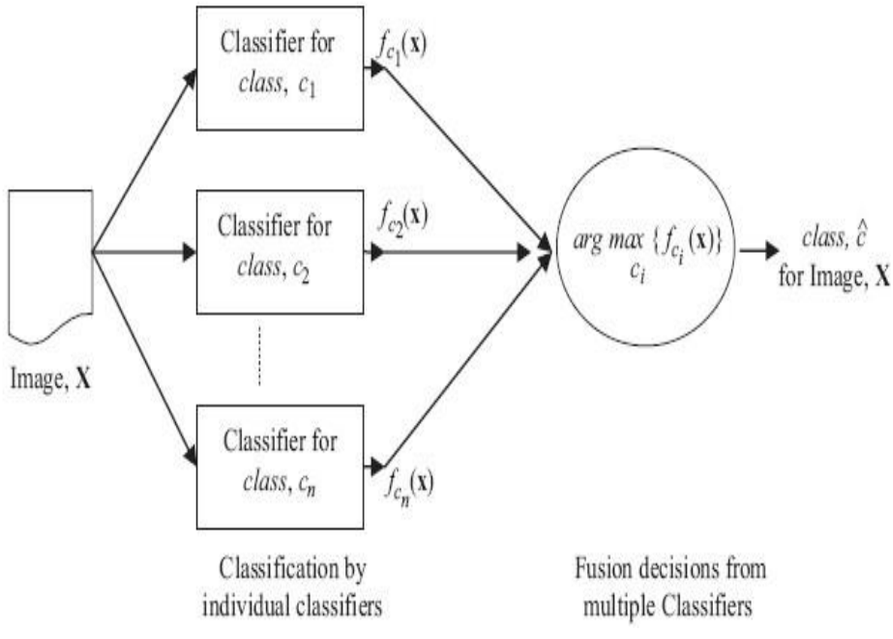


Figure 3: Multi-Class SVM classifier

In CBIR systems, visual similarity is calculated using distance measurements between the feature vectors of the query image and images in the database. The user feeds a query image, and the system ranks the database images based on similarity, often leading to incorrect results when only low-level features are considered. To overcome this issue, AIA systems incorporate semantic concepts based on visual content, enabling more accurate retrieval of relevant images. Pre-processing is crucial for pattern recognition and image classification, as it enhances the quality of input images by removing noise, resizing, and adjusting image features. In this research, the images are normalized through rescaling to (128x128) pixels, ensuring uniformity across datasets and improving computational efficiency, as shown. Additionally, color conversion to grayscale reduces the inherent complexity of the images, facilitating edge detection and pixel-based processing. In this research, edge-based segmentation is employed, relying on intensity differences and content. Edge detection using techniques such as Sobel, Prewitt, and Canny operators helps identify object boundaries by detecting intensity contrasts. Canny edge detection, in particular, is favored for its ability to produce sharp and fine edges, as demonstrated.The performance of various

segmentation techniques is evaluated using metrics such as Root Mean Square Error (RMSE), Signal-to-Noise Ratio (SNR), and Peak Signal-to-Noise Ratio (PSNR). RMSE measures the average difference between the original image and the segmented image, with a higher value indicating greater differences. SNR quantifies the noise present in an image, with higher values representing cleaner, noise-free images. PSNR is commonly used to measure the quality of edge detection between the original and segmented image, with higher values indicating better segmentation accuracy, where RRR is the maximum possible pixel value of the image. The performance evaluation results indicate that the Canny operator outperforms other edge detection techniques in terms of RMSE, SNR, and PSNR values. In thus section, we provide detailed mathematical expressions related to the proposed methodology, including image pre-processing, feature extraction, classification, and evaluation techniques. Each expression will be explained to illustrate its role in the overall image annotation and retrieval system. To normalize the size of images for consistent processing, we perform rescaling. If the original image has dimensions $W \times H$ (width W and height), and we want to resize it to a fixed siae $w_{10} \times h_0$. the rescaling factor S_x and S_y in the x and y directions can be expressed as:

$$S_z = \frac{w_n}{W},\ S_v = \frac{h_0}{H} \quad (1)$$

This ensures the image is resized uniformly for further processing.
To convert a color image to a gray-scale image, a weighted sum of the red, green, and blue (RG日) components is used:

$$I_{\text{gIn}} = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B \quad (2)$$

Where R, G, and B are the intensities of the red, green, and blue compconents of the image, respectively. This formula accounts for the different contributions of each color channel to perceived brightness. Thresholding is a simple segmentation technique used to separate objects from the background by converting an image into a binary format. Given a threshold value T, the binary image $I_{\text{biuny}}(x, y)$ is computed as

$$I_{\text{hinary}}(x, y) = \begin{cases} 1 & \text{if } I(x, y) > T \\ 0 & \text{if } I(x, y) \leq T \end{cases} \quad (3)$$

Where $I(x, y)$ represents the intensity of the pixel at location (x, y). Canny edge detection uses gradients to detect edges. The gradient magnitude G at each pixel is calculated using the partial derivatives in the x - and y-directions, G_x and G_y :

$$G = \sqrt{G G_x^2 + G_y^2} \quad (4)$$

The direction of the edge θ is caloulated as:

$$B = \tan^{-1}\left(\frac{G_p}{G_x}\right) \quad (5)$$

After calculating the gradient magnitude and direction, non-maximum suppression and double thresholding are applied to finalize the edge map.
The GLCM matrix is a statistical measure to describe texture features. For two picels separated by a distance d in a specific direction θ, the GLCM matrix element $p(i, j)$ is defined as:

$$p(i, j) = \sum_{x-1}^{N} \sum_{y-1}^{N} [1 \text{ if } I(x, y) = i \text{ and } I(x + d, y + d) = j] \quad (6)$$

Where $I(x, y)$ is the intensity of the pixel at (x, y), and i and j represent gray-level values. The contrast, a texture feature that describes the intensity contrast between a pixel and its neighbor over the whole image, is computed as

$$\text{Contrast} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2 \cdot p(i, j) \quad (7)$$

Where $p(i, j)$ is the element in the GLCM matrix coeresponding to the gray-level ca-occurrence between i and j. Entropy measures the randomness or complexity of the texture, and is given by:

$$\text{Entropy} = -\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p(i, j) \cdot \log\ p(i, j) \quad (8)$$

Entropy measures the randomness or complexity at the texture, and is grven by:

$$\text{Entropy} = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j)\cdot \log\, p(i,j) \quad (9)$$

This value indicates the level af disorder or unpredictability in the texture of the image. Coarseness measures the texture's roughness, where large differences in pixel intensities indicate coarser textures. The coarseness feature is calculated as

$$C = 2^k,\ k_{u\mu} = \arg\max_k\left(\sum_{x-1}^{N}\sum_{y-1}^{N}\left|A\left(x+2^k,y\right)-A(x,y)\right|\right) \quad (10)$$

Where $A(x,y)$ is the intensity at pixel (x,y) and k_{up} is the scale that maximizes the intensity difference. In Support Vector Machines (SVM), the goal is to find a hyperplane that separates data points of different classes. For a linear SVM, the decision boundary is given by:

$$w\cdot x+b=0 \quad (11)$$

Where w is the weight vector, x is the imput feature vector, and b is the bias term. The hyperplane is defined such that it maximizes the margin between the two classes. The margin M is the distance between the hyperplane and the closest data points, and is defined as

$$M=\frac{2}{\|w\|} \quad (12)$$

The objective is to maximize M, which is equivalent to minimizing $\| w \|^2$ - For non-linearly separable data, kernel functions transform the input space into a higher dimensional space. The polynomial kernel is given by:

$$K\left(x_i,x_j\right)=\left(x_i\cdot x_j+1\right)^2 \quad (13)$$

Where d is the degree of the polynomial, and x_i and x_j are input vectors. RMSE measures the difference between the original and predicted values, after used in evaluating edge detection. RMSE is computed as:

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i-E_i)^2} \quad (14)$$

Where O_i is the original image, E_i is the processed (e.g, edge-detected) image, and N is the total number of pixels.PSNR is used to measure the quality of an image after compression or transformation. It is defined as:

$$\text{PSNR}=10\log_{10}\left(\frac{R^2}{\text{MSE}}\right) \quad (15)$$

Where R is the maximum pixel value (e.g, 255 for 8 -bit imoges) and MSE is the Mean Squared Error between the original and processed image.

These mathematical expressions and their explanations provide a foundation for understanding the various components of the proposed image annotation and retrieval system, from feature extraction to classification and evaluation. Each formula plays a critical role in enhancing the accuracy and efficiency of the averall system. In the proposed methodology, the focus is on automatic image annotation using machine learning, specifically the Multi-Class Support Vector Machine (MCSVM) classifier. Automatic image annotation is a classification task where an image is automatically labeled with semantic keywords based on its visual content. Traditional binary SVM classifiers have limitations in handling multi-class problems, which are common in image annotation tasks. MCSVM extends the binary SVM approach to handle multiple classes by training classifiers for each class and combining their outputs to classify new images.

The proposed system incorporates the Semantic Keyword Transfer (SKT) algorithm to bridge the gap between low-level image features and high-level semantic concepts. Image classification involves training a model to recognize patterns in labeled images and applying this model to classify new images. Classification techniques such as Minimum Distance Classifier (MDC), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Decision Trees (DT) are commonly used in image processing.

The SVM classifier is particularly effective in high-dimensional data classification due to its ability to

create optimal class boundaries by maximizing the margin between classes. In the context of image annotation, MCSVM is used to classify images with multiple objects or regions.

The proposed methodology for automatic image annotation combines fused features (texture and shape) with the MCSVM classifier and SKT algorithm. This approach bridges the semantic gap between low-level image features and high-level semantic concepts, resulting in improved image retrieval accuracy. The integration of Haralick and Tamura texture features with shape features provides a comprehensive representation of image content, while the MCSVM classifier efficiently handles multi-class image annotation tasks. The evaluation results demonstrate that the proposed system outperforms existing methods in terms of retrieval accuracy, making it a promising solution for automatic image annotation and retrieval tasks.

4. Results and Analysis

This research proposes and examines a simple algorithm to perform this crowd behavior analysis. Given an aerial image of a crowd, the algorithm segments the image into crowd and non-crowd regions. On a large scale, we expect a crowd to contain some repetitive visual elements or textures that are significantly different from that of a non-crowd region. The proposed algorithm uses multiple Gabor filters to capture these different textures in an image and uses improved pre processing and support vector machines to segment the image into 2 groups corresponding to crowd and non-crowd regions. This research attempts to detect crowds of humans in still images. Given an image, the proposed algorithm segments out the regions that the crowd occupies. The data set consists of 1200 aerial images of crowds taken from the internet. Each images are tagged with a range 5 properties. By testing the algorithm on a range of images with varying properties, this research aims to choose a good set of parameters that can detect crowd well despite the diverse characteristics of crowds.

The ratio σ/λ determines the spatial frequency bandwidth and hence the number of parallel excitatory and inhibitory stripes in the Gabor filter. The half-response spatial frequency bandwidth b (in octave) related to the ratio σ/λ as follows:

$$b = \log_2 \frac{\frac{\sigma}{\lambda}\pi + \sqrt{\frac{\ln(2)}{2}}}{\frac{\sigma}{\lambda}\pi - \sqrt{\frac{\ln(2)}{2}}}, \quad \frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{\ln(2)}{2}}\frac{2^b+1}{2^b-1}. \tag{16}$$

In order to capture the repetitive texture of a crowd from many perspectives, we use 6 orientations with orientation separation angles of $d_\theta = 30°$:

$$\theta: 0°, 30°, 60°, 90°, 120°, 150° \tag{17}$$

We also use a range of wavelengths, evenly spaced in $\log_2$-space, ranging from some minimum wavelength to the radius of the image (or half its diagonal length). The choice of the minimum wavelength is adjusted when we apply the algorithm to some initial images. The general formula for the chosen wavelengths is

$$\lambda: \lambda_{\min} \times 2^k, k \in \mathbb{N} \tag{18}$$

For example, if we choose both $\lambda_{\min}$ and r_λ equal to 2 for a 288×512 image, there would be a total of 42 Gabor filters used from 6 orientations and 7 wavelengths.
In this work we set the value of the bandwidth b by default to 1 octave. In that case, the Equation gives

the approximation

$$\sigma = 0.5 \times \lambda \quad (19)$$

For each filtered image, we use a Gaussian smoothing function given by:

$$g(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (20)$$

where σ is the standard deviation that determines the windown size. The ratio σ/σ_g (where σ_g is the standard deviation parameter of Gabor filter) is estimated and adjusted when we apply the algorithm to some initial images. We first test them on minimum wavelength λmin = 3 and the gaussian vs gabor standard deviation ratio σ/σg = 3. The resulting segmentation is in Figure 5.

Figure 5: Test image of moderate crowd scenario

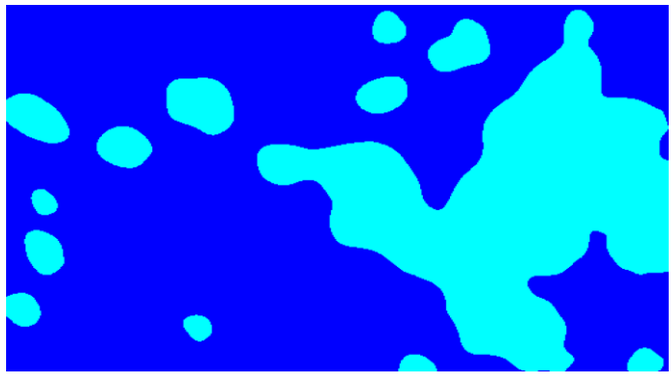


Figure 6: Labeling of images in moderate scenario

Figure 7: Segmentation of crowd scenario

Figure 8: Test image of high crowd scenario

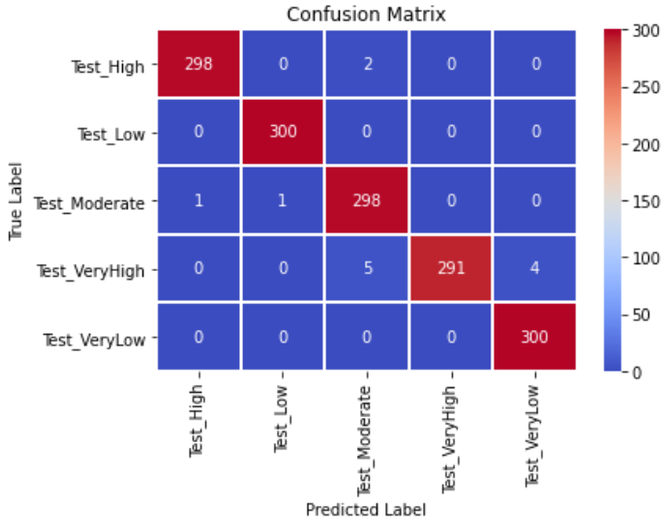


Figure 9: Plot of confusion matrix

Table 1: Analysis of performance parameters

Scenario	Precision	Recall	F1-Score
1	1	0.98	0.99
2	0.98	0.95	0.97
3	1	0.98	1
4	1	0.99	1
5	0.97	0.95	0.95

The algorithm does decently well with both of the picture. For both images, it pinpoints the correct regions where the crowds of people are. In the first image, it seems slightly over estimate the size of each crowd on the left and right. But the crosswalk stripes do not seem to confuse the algorithm. With the second image, the algorithm does a slightly worse job, as the shadow makes it overestimates the regions that the crowd occupies, and there are quite a few people who are not captured as belonging to the crowd.

Table 2: Comparative analysis of proposed methodology

Parameter	Previous Work	Proposed Work
Type of Detection	Segmentation	Segmentation and Classification
Type of Analysis	Single Level Scenario	Multiple Scenario
Performance Parameter	F Score	Precision, Recall and F Score
Implementation	Complex	Simple
Computational Time	Average	Faster

In order to lessen the algorithm's overestimation and be able to detect more people in a scattered crowd, we will reduce the value of both the minimum wavelength and the standard deviation ratio. The goal is that the algorithm can pick up smaller details in the picture and thus segment more precisely all the regions of the crowd.

In the second trial, we change the minimum wavelength to 2 and the standard deviation ratio to 1.6. The algorithm seems to improve for both images. For the first image, the algorithm seems to reduce the algorithm overestimation, although it seems to confuse a tiny part of the crosswalk stripes as parts of the crowd. For the second image, the algorithm seems to no longer include the majority of the shadow as parts of the crowd, and there are only 1-2 people who are no included as belonging to the crowd. As a result, we choose minimum wavelength equal 2 and standard deviation ratio equal 1.6 as the parameters for our algorithm, in addition to the other parameters

There are some defects inherent in Matlab average filters such as Gabor and Gaussian. In particular, they assume that pixels out of the image has intensity of 0, and thus it is possible the algorithm does not work well for pixels at the circumference of images. This problem did not arise with the 16 images in this data set, but it is a problem that may be needed to deal with when applying to more images in different circumstances. This program worked reasonably fast, needed from 20.839009 to 31.543316 seconds for each image of size 288 × 512. However, the time does add up when we want to process all the images multiple times when testing for different parameters. Crowd image segmentation and detection play a

significant role in various computer vision applications, including crowd monitoring, crowd behavior analysis, and public safety. This work presents a comprehensive study on the use of Gabor filters and Support Vector Machine (SVM) for crowd image segmentation and detection. The Gabor filter is employed to extract discriminative features from crowd images, and SVM is used as a classifier to distinguish between crowd and non-crowd regions. The results demonstrate the effectiveness of this approach in accurately segmenting and detecting crowds in complex visual scenes. This research concludes by discussing the potential applications of crowd image segmentation and detection using Gabor filters and SVM in real-world scenarios.

5. Conclusion

This research presents a novel approach to crowd behavior analysis using a combination of Gabor filters and Support Vector Machines (SVM) to detect and segment crowds in still images. The algorithm effectively segments an image into crowd and non-crowd regions by identifying repetitive textures that differentiate the crowd from the background. Through the use of multiple Gabor filters, the method captures various orientations and scales of these textures, enhancing the detection of crowd-specific characteristics. The SVM classifier is used to cluster the regions based on these features, ensuring that crowd regions are distinguished from non-crowd areas. The ability to detect crowds in public spaces is crucial for preventing congestion, ensuring safety, and enforcing social distancing measures. This research successfully demonstrates that crowd segmentation is a vital preprocessing step for more complex tasks such as crowd density estimation and behavior analysis. The algorithm's robustness is tested on a dataset of 1200 aerial images with varying properties, including crowd density, background variation, and lighting conditions, resulting in reliable crowd detection.Despite some limitations, such as overestimation in regions affected by shadows, the proposed methodology improves the precision and accuracy of crowd detection. By adjusting key parameters like the minimum wavelength and standard deviation ratio, the algorithm's performance was optimized, providing precise crowd segmentation. This research highlights the potential for further advancements in crowd detection, with applications in public safety, event management, and urban planning, offering a foundation for real-time crowd analysis systems in diverse environments.

REFERENCES

[1.] Aditya, CSK, Hani'ah, M, Bintana, RR & Suciati, N 2015, 'Batik classification using neural network with gray level co-occurence matrix and statistical color feature extraction', in 2015 International Conference on Information & Communication Technology and Systems (ICTS), pp. 163-8.

[2.] Ahmed, M, Mahmood, AN & Hu, J 2016, 'A survey of network anomaly detection techniques', Journal of Network and Computer Applications, vol. 60, pp. 19-31.

[3.] Anton, SD, Kanoor, S, Fraunholz, D & Schotten, HD 2018, 'Evaluation of machine learning-based anomaly detection algorithms on an industrial modbus/tcp data set', in Proceedings of the 13th international conference on availability, reliability and security, pp. 1-9.

[4.] Au, CE, Skaff, S & Clark, JJ 2006, 'Anomaly detection for video surveillance applications', in 18th International Conference on Pattern Recognition (ICPR'06), vol. 4, pp. 888-91.

[5.] Babenko, B, Yang, M-H & Belongie, S 2010, 'Robust object tracking with online multiple instance learning', IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 8, pp. 1619-32.

[6.] Belousov, A, Verzakov, S & Von Frese, J 2002, 'A flexible classification approach with optimal generalisation performance: support vector machines', Chemometrics and intelligent laboratory systems, vol. 64, no. 1, pp. 15-25.

[7.] Benabbas, Y, Ihaddadene, N & Djeraba, C 2011, 'Motion pattern extraction and event detection for automatic visual surveillance', EURASIP Journal on Image and Video Processing, vol. 2011, pp. 1-15.
[8.] Bertini, M, Del Bimbo, A & Seidenari, L 2012, 'Multi-scale and real- time non-parametric approach for anomaly detection and localization', Computer Vision and Image Understanding, vol. 116, no. 3, pp. 320-9.
[9.] Bezdek, JC, Ehrlich, R & Full, W 1984, 'FCM: The fuzzy c-means clustering algorithm', Computers & Geosciences, vol. 4, no. 10, pp. 191-203.
[10.] Brassil, J 2009, 'Technical challenges in location-aware video surveillance privacy', in Protecting Privacy in Video Surveillance, Springer, pp. 91-113.
[11.] Brutzer, S, Höferlin, B & Heidemann, G 2011, 'Evaluation of background subtraction techniques for video surveillance', in CVPR 2011, pp. 1937-44.
[12.] Castiglione, A, Cepparulo, M, De Santis, A & Palmieri, F 2010, 'Towards a lawfully secure and privacy preserving video surveillance system', in International Conference on Electronic Commerce and Web Technologies, pp. 73-84.
[13.] Chae, J, Thom, D, Bosch, H, Jang, Y, Maciejewski, R, Ebert, DS & Ertl, T 2012, 'Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition', in 2012 IEEE Conference on Visual Analytics Science and Technology (VAST), pp. 143-52.
[14.] Chandola, V, Banerjee, A & Kumar, V 2009, 'Anomaly detection: A survey', ACM computing surveys (CSUR), vol. 41, no. 3, pp. 1-58.
[15.] Chang, C-I & Chiang, S-S 2002, 'Anomaly detection and classification for hyperspectral imagery', IEEE transactions on geoscience and remote sensing, vol. 40, no. 6, pp. 1314-25.
[16.] Chapelle, O, Scholkopf, B & Zien, A 2009, 'Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]', IEEE Transactions on Neural Networks, vol. 20, no. 3, pp. 542.
[17.] Chen, M, Chen, S-C & Shyu, M-L 2007, 'Hierarchical temporal association mining for video event detection in video databases', in 2007 IEEE 23rd International Conference on Data Engineering Workshop, pp. 137-45.
[18.] Cheng, K-W, Chen, Y-T & Fang, W-H 2015, 'Video anomaly detection and localization using hierarchical feature representation and Gaussian process regression', in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2909
[19.] Cho, S-B & Park, H-J 2003, 'Efficient anomaly detection by modeling privilege flows using hidden Markov model', Computers & Security, vol. 22, no. 1, pp. 45-55.
[20.] Choi, Y-S 2009, 'Least squares one-class support vector machine', Pattern Recognition Letters, vol. 30, no. 13, pp. 1236-40.
[21.] Chong, YS & Tay, YH 2017, 'Abnormal event detection in videos using spatiotemporal autoencoder', in International symposium on neural networks, pp. 189-96.
[22.] Coello, CAC, Pulido, GT & Lechuga, MS 2004, 'Handling multiple objectives with particle swarm optimization', IEEE Transactions on evolutionary computation, vol. 8, no. 3, pp. 256-79.
[23.] Cong, Y, Yuan, J & Tang, Y 2013, 'Video anomaly search in crowded scenes via spatio-temporal motion context', IEEE transactions on information forensics and security, vol. 8, no. 10, pp. 1590-9.
[24.] Dasarathi, S 2015, 'Parametrization of Convolutional Neural Network for Image Classification', Dublin, National College of Ireland.
[25.] Davies, AC, Yin, JH & Velastin, SA 1995, 'Crowd monitoring using image processing', Electronics & Communication Engineering Journal, vol. 7, no. 1, pp. 37-47.

[26.] Davis, JW & Sharma, V 2005, 'Fusion-based background-subtraction using contour saliency', in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops, pp. 11.

[27.] Du, B & Zhang, L 2014, 'A discriminative metric learning based anomaly detection method', IEEE transactions on geoscience and remote sensing, vol. 52, no. 11, pp. 6844-57.

[28.] Du, B, Zhao, R, Zhang, L & Zhang, L 2016, 'A spectral-spatial based local summation anomaly detection method for hyperspectral images', Signal Processing, vol. 124, pp. 115-

[29.] Duan, L-Y, Xu, M, Tian, Q, Xu, C-S & Jin, JS 2005, 'A unified framework for semantic shot classification in sports video', IEEE Transactions on multimedia, vol. 7, no. 6, pp. 1066-83.

[30.] Feizi, A 2020, 'Hierarchical detection of abnormal behaviors in video surveillance through modeling normal behaviors based on AUC maximization', Soft Computing, vol. 24, no. 14, pp. 10401-13.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-14] COMPARATIVE ANALYSIS OF DATA REDUNDANCY STRATEGIES FOR WIRELESS SENSOR NETWORKS IN SMART CITIES

Mr. Ashish Sharma[1,†], Dr. Sandeep Tayal[1,†], Dr. Yogesh Sharma[1,†], Shivam Gangal[*,†], Jiya Verma[†], Vivek Kaushal[†]

[1]*Computer Science and Engineering Department, Maharaja Agrasen Institute of Technology, Affiliated to G.G.S.I.P. University.*

Abstract

Energy efficiency can be understood as getting desired outcome while consuming the least amount of energy possible. In context of wireless sensor networks (WSNs), tiny- battery-powered sensors work together to collect environmental data. These networks, often deployed in remote areas, rely on efficient energy use to function for extended periods. Since replacing batteries in these sensors can be difficult or impractical, maximizing their lifespan is critical. Therefore, designing WSNs with energy efficiency in mind is crucial. By minimizing energy consumption, WSNs can function for longer durations without intervention, leading to cost and effort reductions.

Keywords

Wireless Sensor Networks (WSNs), Energy Efficiency, Network's Lifetime, Challenges in Energy Efficiency, Sensor node components, Dynamic Power Management (DPM), Dynamic Voltage Scaling (DVS), Phases of WSN, Duplicate Data Elimination, Data Compression, Data Aggregation Techniques, Optimization Algorithms.

1. INTRODUCTION

Modern deployments of wireless technology encompass Wireless Sensor Networks (WSNs) [1], which offer a multitude of applications. These applications include surveillance, environmental monitoring, intrusion detection, healthcare, early warning systems for disasters, defence systems, target tracking, and security [1].

The rise of WSNs coincides with breakthroughs in low-bandwidth radio technologies, allowing for denser networks with faster data transfer [2]. Wireless networks are particularly advantageous in situations where traditional wired connections are impractical due to the environment being inaccessible. In these scenarios, collecting data directly is often difficult, making WSNs the ideal solution for sensing such areas [2].

However, a major challenge in WSNs is their limited energy supply. Because sensor nodes usually depend on batteries, ensuring extended sensor operation requires minimizing energy consumption.

SCCTT 2024: Proceedings of CEUR Workshop Proceedings,
*Corresponding author.
†These authors contributed equally.
shivamgangal7@gmail.com (Shivam Gangal); ashish@mait.ac.in (Mr. Ashish Sharma); sandeeptayal@mait.ac.in (Dr. Sandeep Tayal); yogeshsharma027@gmail.com (Dr. Yogesh Sharma); Jiyaverma3101@gmail.com (Jiya Verma); vivekkaushal501@gmail.com (Vivek Kaushal)

1.1. Challenges in Energy Efficiency

There are several challenges to consider when aiming to increase energy efficiency in WSNs:

Limited Energy Resources: Sensor nodes rely on compact batteries with limited capacity, necessitating efficient operation to stretch their lifespan. This highlights the importance of maximizing energy use from these finite resources.

Data Processing: Sensor nodes often perform data processing tasks before transmitting data. These computations consume additional energy. Optimizing data processing algorithms and techniques can help reduce energy consumption.

Network Topology: The arrangement of sensor nodes and their connectivity affect energy efficiency. Optimizing network topology, such as reducing the distance between nodes or employing clustering techniques, is vital to balance energy consumption across the network.

Routing and Data Aggregation: Efficient routing protocols and data aggregation techniques can significantly impact energy efficiency in WSNs. WSN mechanisms ensure data reaches the base station efficiently by minimizing redundant transmissions.

1.2. Early energy-efficient techniques in WSN

Pioneering research on energy conservation in WSNs explored two key methods: Dynamic Power Management (DPM) and Dynamic Voltage Scaling (DVS).

Dynamic Power Management (DPM): This approach advocates for temporarily turning off unused devices and reactivating them when needed. However, limitations exist. DPM relies on a combination of operating system integration and probabilistic modelling to anticipate upcoming device usage patterns.

Dynamic Voltage Scaling (DVS): This method adjusts power consumption based on the network's workload. By dynamically changing voltage and frequency, DVS effectively reduces overall power usage. The key lies in accurately predicting future workloads. Effective workload distribution hinges on considering both ongoing tasks and predicted future demands.

For embedded systems like Wireless Sensor Networks (WSNs) [3], conserving energy is critical. They also face challenges in setting up the network, data aggregation, monitoring specific locations/objects, and network safety. Despite these complexities, WSNs are a valuable tool for data acquisition in various applications.

Self-organizing WSNs equip sensor nodes with the ability to adapt through the use of adaptive algorithms. This approach complements dynamic power allocation techniques used in IP networks, which leverage power-saving modes, reliability, and prioritization techniques for reliable data delivery [3].

1.3. Energy Consumption in various phases of WSN

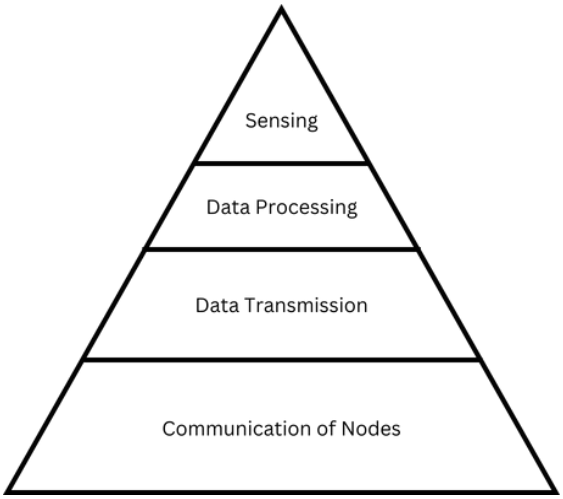


Fig 1 [4]**:** Energy Breakdown in WSN Operations

Data transmission significantly impacts energy use in Wireless Sensor Networks (WSNs), outweighing data processing. Transmitting a single data packet can consume roughly the same amount of energy as processing thousands of functions within a sensor node. While the sensor unit's energy consumption can fluctuate depending on the type of sensor, communication between nodes consistently represents the largest consumer of energy in WSNs. Sensor data acquisition itself consumes negligible energy compared to processing and communication [5].
Consequently, energy-efficient techniques for WSNs primarily target communication protocols and sensor operation. By combining various techniques, we can significantly extend the operational lifespan of WSN deployments [6].

1.4. Strategies for energy efficiency to enhance energy efficiency in WSN

Several strategies and techniques can be implemented:

Sleep Scheduling: Sleep scheduling involves adjusting the duty cycle of nodes to reduce power consumption. By letting nodes sleep during low-demand periods and waking them up only when necessary, significant energy savings can be achieved.
Data Compression: Data compression minimizes the information sent, reducing transmission demands, thereby lowering communication energy consumption. Compression algorithms are designed to minimize data size while retaining essential information.

Energy Harvesting: Sensor nodes can leverage energy harvesting technologies to extract power from their surroundings, like sunlight or vibrations, to supplement their battery power. By utilizing renewable energy sources, the nodes can prolong their operational lifetime.
Dynamic Power Management: Dynamically adjusting the power levels of sensor nodes according to the required operational level aids in optimizing energy consumption. Power management algorithms are designed to balance operational needs with energy usage.
Cross-Layer Design: Collaboration among various layers of the network protocol stack can result in energy-efficient designs. Cross-layer design facilitates improved coordination and optimization between layers, leading to decreased energy consumption.

2. ENERGY OPTIMIZATION ALGORITHMS FOR WIRELESS SENSOR NETWORKS

Addressing energy constraints is a significant challenge for Wireless Sensor Networks (WSNs) [6], especially as their use grows in areas like environmental monitoring, smart agriculture, and industrial automation. Prolonging the network's operational life requires optimizing energy usage by deploying effective algorithms [6].

2.1. Energy Efficiency in Wireless Sensor Networks

2.1.1. Understanding Energy Efficiency

Energy efficiency is a key principle in any system, aiming to achieve desired outcomes while minimizing energy consumption. In Wireless Sensor Networks (WSNs), where sensor nodes usually depend on limited battery power, energy efficiency is critical. By optimizing energy use, WSNs can extend their operational lifetime, minimizing disruptions caused by battery depletion or the need for frequent replacements.

2.1.2. Factors Affecting Energy Consumption in WSNs

Understanding the factors that contribute to energy consumption helps in identifying optimization opportunities. This section examines the primary factors influencing energy efficiency in Wireless Sensor Networks (WSNs):

Transmitting Data: Transmitting data requires a substantial amount of energy. This includes both radio transmission and data processing.
Receiving Data: Receiving data also requires energy, as the node needs to remain active and process the incoming data.
Sensing Environment: Sensing the environment using sensor nodes demands energy, particularly in cases where sensors need to sample and analyze data frequently.

Communication Range: Larger communication ranges necessitate higher transmission power, leading to increased energy consumption.
Data Aggregation: Combining sensor data before transmission minimizes the number of transmissions required, leading to significant energy conservation in WSNs.

2.2. Optimization Algorithms for Energy Efficiency

2.2.1. Adaptive Duty Cycling

Adaptive Duty Cycling (ADC) is a prominent optimization technique employed to reduce energy consumption in wireless sensor networks. It seeks to find a balance between energy conservation and the timely delivery of data.

2.2.2. Topology Control

WSN efficiency hinges on topology control algorithms, which optimize network structure to minimize energy use. By selectively activating certain nodes and adjusting transmission power levels, topology control algorithms minimize energy wastage.

2.2.3. Data Aggregation Techniques

Data aggregation techniques focus on reducing the amount of data transmitted by merging similar or redundant information into a single message. By aggregating data in a localized manner, energy consumption is significantly reduced since the number of transmissions is minimized.

2.2.4. Routing Protocols

Routing protocols play a vital role in energy efficiency as they determine the paths through which data is transmitted in the network. Examples of energy-efficient routing protocols include Low-Energy Adaptive Clustering Hierarchy (LEACH), Directed Diffusion, and Minimum Hop Routing (MHR).

2.2.5. Sleep Scheduling

Algorithms aim to strategically put sensor nodes into a deep sleep mode for extended periods to conserve energy. By coordinating sleep schedules across the network, energy consumption is reduced while ensuring connectivity and data delivery.

3. TECHNIQUES AND ALGORITHMS FOR DATA REDUNDANCY REDUCTION IN WIRELESS SENSOR NETWORKS (WSNs)

Data redundancy reduction techniques are essential for enhancing the efficiency and performance of Wireless Sensor Networks (WSNs) by decreasing the volume of redundant

information transmitted and stored [7]. Below are some common techniques employed for data redundancy reduction in WSNs:

3.1 Data Aggregation

Data aggregation is a cornerstone in Wireless Sensor Networks (WSNs), playing a pivotal role in optimizing network efficiency, conserving resources, and extending network lifespan. This section elucidates the essence of data aggregation, its significance, and diverse implementation methods, positioning it as a vital technique for maximizing WSN efficiency [7].

3.1.1 Essence of Data Aggregation

Data aggregation encompasses the in-network processing of raw sensor data, here intermediate nodes perform operations such as averaging, summation, or selection to generate aggregated data. This data is subsequently transmitted towards the sink node, hence reducing the overall volume of transmitted data and conserving network resources [8].
It is characterized by resource-constrained sensor nodes, and the direct transmission of raw data to the sink node poses significant challenges such as energy depletion and network congestion. Data aggregation faces the challenges of:

Reducing Transmission Overhead: By processing data closer to the source, data aggregation minimises the number of packets transmitted, hence conserving energy.
Mitigating Network Congestion: The reduced data volume reduces congestion on communication channels, thus enhancing overall network performance.
Extending Network Lifetime: Lower energy consumption due to fewer transmissions translates to a prolonged network lifespan, enhancing sustainability.

3.1.2 Implementation of Data Aggregation

The implementation of data aggregation in WSNs occurs at different levels within the network hierarchy:-

In-node Aggregation: Individual sensor nodes process the sensed data locally before transmission.
Cluster-based Aggregation: Sensor nodes are grouped into clusters, where cluster heads are tasked with aggregating data from member nodes before transmitting it to the sink.
Tree-based Aggregation: Nodes form a tree structure where data is progressively aggregated as it ascends towards the sink, offering flexibility in data routing.

3.1.3 Parameters for Effective Aggregation

The effectiveness of data aggregation techniques hinges on several parameters, including:

Aggregation Function: The choice of aggregation function (e.g., mean, median) that influences the level of information preservation at the time of aggregation.
Data Correlation: The degree of similarity between data from neighbouring nodes that affect the potential for efficient aggregation.
Network Topology: The spatial distribution of sensor nodes and the presence of cluster heads pushing data forwarding paths and aggregation opportunities.

3.1.4 Methods for Data Aggregation

Various methods have been proposed for implementing data aggregation in WSNs, each offering distinct advantages and limitations:

Min-Max Aggregation*:* Provides a concise overview of data trends by transmitting minimum and maximum values but may sacrifice detailed information [7].
Mean Aggregation: Calculates the average of sensed data, summarising statistically similar data but potentially overlooking outliers [7].
Median Aggregation: Offers robustness to outliers compared to the mean but may necessitate more complex calculations [7].
Histogram Aggregation: Constructs histograms locally to capture data distribution without transmitting raw data, suitable for applications requiring data distribution insights [9].
Fuzzy Aggregation: Utilizes fuzzy logic to handle uncertainty in sensor data, particularly beneficial for environmental monitoring applications [7].

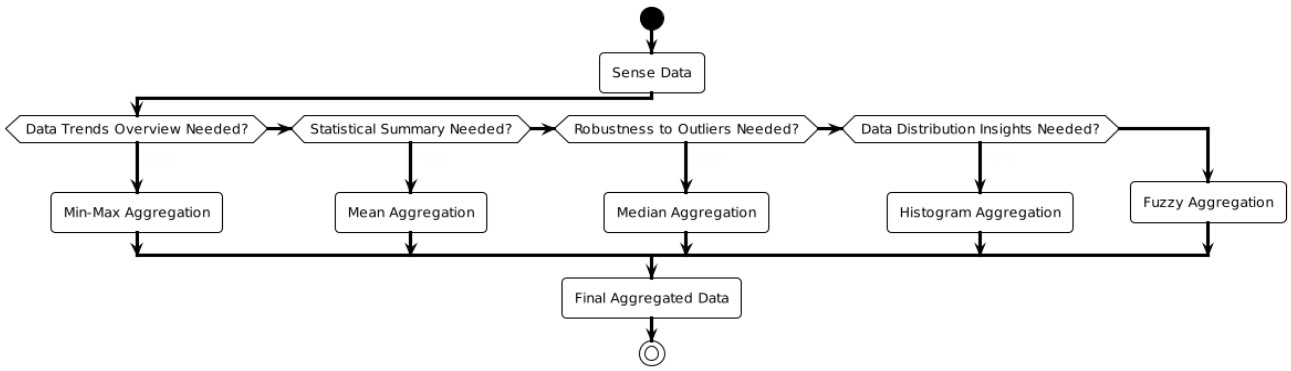


Figure 2: Various Aggregation techniques used according to the required metric

Innovating data aggregation techniques in WSNs enhances efficiency, reliability, and data fidelity, driving advancements across various applications and ensuring sustainable, efficient network operations.
Data aggregation guarantees a reduction in redundancy, ensuring that results are retained. This analysis reveals that the proposed algorithm exhibits improved network longevity and better energy consumption compared to other traditional algorithms.

3.2 Data compression

A data compression tool is a valuable tool for improving the efficiency and effectiveness of wireless sensor networks (WSNs). Network lifecycle provides an in-depth into the concept of data compression, its essence in reducing redundancy, various methods used in WSNs, and implementation of the same.

3.2.1 Essence of Data Compression

Data compression involves encoding data in a manner that minimizes the amount of storage required. WSNs are marked by resource-constrained sensor nodes, which have limited battery power and bandwidth. Data compression solves the redundancy problem by providing the following:

Reduced transmission Load: Compression reduces the number of transmissions by removing redundant data, leading to considerable energy savings. By eliminating redundant data, compression minimizes the number of bits transmitted, leading to significant energy savings. This can be calculated using the formula:

$$\textbf{Energy Saved (\%) = (1 - Compression Ratio) * 100} \quad \textbf{(1)} \text{ [10]}$$

Improved scalability: Reduces transfer rates can handle larger data, improving network scalability for dense sensor deployment.

Extending Network Lifetime: Reduced transmission translates to lower energy consumption, ultimately extending the operational lifespan of the network [11].

3.2.2 Implementation of Data Compression

Data Compression can occur at various levels due to its applicability being extremely diverse and effective in ensuring the reduction of repetition. Its functionality at different levels of WSN can be seen effectively as:

Intra-node compression: Intra-node compression: Each sensor node compresses the detected data before transmission, hence reducing node transmission overhead.

Network-wide compression: Data can be compressed at a specific network location (such as a card) before being sent to the recipient.

3.2.3 Compression quality parameters

The effectiveness of data compression technology in Wireless Sensor Networks (WSNs) relies on several parameters. Taking a broader look at these aspects, we can observe:

- **Compression ratio:** This parameter is determined by the compression algorithm. The smaller the size of the result files, the higher the ratio means more reductions. Formula:

Compression Ratio = Original Size / Compressed Size **(2)** [10]

- **Features:** Features of useful data (such as data type and classification) usually affect the suitability of various methods
- **Computational complexity:** The energy required for compression affects overall network efficiency and performance. For resource-constrained sensor nodes, fewer algorithms are preferred as complex algorithms consume excess energy.

3.2.4 Data compression methods

Many data compression methods have been examined for use in WSN, each method has advantages and limitations. Some would be:

- **Lossless Compression:** Huffman Coding and Lempel-Ziv (LZ) Coding allows the reconstruction of the original material after decompression. These methods ensure the applications where data accuracy is important, but they may not always achieve the highest compression ratio [12].
- **Lossy compression:** Techniques such as quantization and transfer coding permit data loss to be controlled in exchange for a higher compression ratio. This method is suitable for applications that result in some loss of quality data, such as environmental monitoring where small temperature changes may not make a bigger impact [12].
- **Dictionary-based compression:** This method exploits recurring patterns in data by creating a dictionary of frequently encountered words hence keeping a record. Characters or segments of data are encountered, stored and further used. This approach can achieve similar results for devices with core components but requires additional dictionary management and deployment overhead [13].

The choice of data compression technology in Wireless Sensor Networks (WSNs) depends on application requirements and network constraints, with various parameters influencing the decision Balancing compression ratio and data integrity is essential in many WSN applications.

3.3 Predictive Modelling

Predictive modelling in Wireless Sensor Networks (WSNs) forecasts future sensor readings from historical data patterns, reducing data redundancy through analysis of algorithms, implementation methods, and various techniques for improved efficiency and effectiveness. It entails creating mathematical models to forecast future outcomes based on historical data, which can be either recent or significantly older to enhance accuracy [14]. In WSNs, predictive models analyse past sensor readings to forecast future values, enabling proactive decisions.

Method	Description	Advantages	Disadvantage	Most Suitable for Data
Lossless Compression	Techniques like Huffman coding and Lempel-Ziv (LZ) coding achieve perfect reconstruction of the original data after decompression.	Guarantees data integrity	May not achieve the highest compression ratios	Sensor readings with high-fidelity requirements
Lossy Compression	Techniques like quantization and transform coding allow for controlled data loss in exchange for higher compression ratios.	Achieves higher compression ratios	Introduces data loss	Sensor readings where a certain level of accuracy is tolerable (e.g., temperature monitoring)
Dictionary-based Compression	These methods exploit repetitive patterns within the data by creating dictionaries of frequently occurring symbols or data segments, achieving high compression for data.	Highly effective for data with redundancy	Requires additional overhead for dictionary management	Sensor readings with recurring patterns (e.g., environmental monitoring)

Table 1: Various Data Compression Techniques and their comparison

3.3.1 The Essence of Predictive Modelling

Predictive modelling in Wireless Sensor Networks (WSNs) forecasts future sensor values using historical data, reducing data transmission by sending only differences or selective updates. This decreases energy consumption, extends operational lifespan, and enhances scalability, accommodating larger networks with minimal bandwidth limitations.
Overall, predictive modelling comes as a powerful technique for data redundancy reduction in WSNs, contributing to improved energy efficiency, prolonged network lifetime, and enhanced scalability.

3.3.2 Implementation of Predictive Modelling

Predictive modelling for redundancy reduction in WSNs can be implemented in various ways:

- **In-node Prediction:** Individual sensor nodes employ local prediction models to present their future values. This approach minimises communication overhead but requires sufficient processing power on each node.
- **Cluster-based Prediction:** Sensor nodes in a cluster further collaborate, sending only the prediction error or raw data exceeding a certain error threshold limitation to the cluster head for additional processing.
- **Centralised Prediction:** Sensor data is sent to a central node (sink) for comprehensive prediction using more sophisticated and complex models.

The optimal implementation strategy depends on the network architecture, resource constraints, and desired trade-off between prediction accuracy and communication efficiency [15].

3.3.3 Evaluation Parameters

The effectiveness of predictive modelling techniques in WSNs is evaluated using several parameters:

Prediction Accuracy: Measured using metrics such as Mean Squared Error (MSE) or Mean Absolute Error (MAE), lower values signify more accurate predictions.
Energy Consumption: The total energy spent on model training, prediction, and data transmission comprehends resource evaluation based on the provided network [16].
Computational Complexity: The processing resources required for training and executing predictive models to ensure error-free results.

3.3.4 Methods and Algorithms involved in Predictive Models

Several algorithms have been explored for predictive modelling in WSNs, each offering distinct advantages and limitations. Most are used on provided networks and their functionality keeping in mind the evaluation parameters. These common algorithms would be:
Auto-Regressive Integrated Moving Average (ARIMA): This widely used time series forecasting method leverages past observations and their lagged values to estimate future values.

Formula:

$$Yt = c + \phi 1 ydt - 1 + \phi pydt - p + \ldots + \theta 1 et - 1 + \theta qet - q + e \qquad \textbf{(3)}\ [17]$$

Simplified Formula:

$$Yt = c + \Sigma \varphi i * Y(t - i) + \Sigma \theta i * \varepsilon(t - i) \qquad \textbf{(4)}\ [17]$$

Where:-

Yt: Predicted value at time t; **c:** Constant term; **φ:** Autoregressive coefficient; **θ:** Moving average coefficients; **ε:** White noise error term at time t

Kalman Filter: This recursive estimation technique is well-suited for scenarios with dynamic sensor data and incorporates process noise for more accurate predictions [18].

Kalman Filter Equations (Simplified):

State prediction: **X_k = A * X_(k-1) + B * U_k** **(5)** [18]

Covariance prediction: **P_k = A * P_(k-1) * A^T + Q_k** **(6)** [10]

Kalman Gain: **K_k = P_k * H^T * (H * P_k * H^T + R_k)^(-1)** **(7)** [18]

State update: **X_k^est = X_k + K_k * (Z_k - H * X_k)** **(8)** [18]

Covariance update: **P_k^est = (I - K_k * H) * P_k** **(9)** [18]

Where:-

X_k: State vector at time k; A: State transition matrix; **B:** Control input matrix; **U_k:** Control input at time k [18]; **P_k:** Covariance matrix at time k; **Q_k:** Process noise covariance matrix [10]; **H:** Observation matrix; **R_k:** Measurement noise covariance matrix; **Z_k:** Measurement at time k; **X_k^est:** Estimated state at time k; **P_k^est:** Estimated covariance matrix at time k [18].

Artificial Neural Networks (ANNs): These data-driven models can learn complex relationships within sensor data and offer superior prediction accuracy, particularly for non-linear patterns. However, they often require significant training data and computational resources.

Linear Regression: Linear regression estimates the relationship between independent variables.

Formula:

$$y = mx + b \quad \textbf{(10)}\ [15]$$

Where:-

x and dependent variable; y by fitting a straight line to the data points.

Support Vector Machines (SVM): Support Vector Machines (SVM) create a hyperplane in a high-dimensional space to categorize data points and forecast future outcomes [15].

Predictive modelling offers a compelling approach for redundancy reduction in WSNs, but its efficiency and effectiveness depend on various factors [19]:

Data Characteristics: Data with strong temporal correlation (e.g., temperature readings) is more suitable for accurate predictions compared to rapidly changing data (e.g., seismic activity).
Computational Complexity: The training and execution of complex models (e.g., ANNs) can be computationally expensive for resource-constrained sensor nodes.
Communication Overhead: While predictive models aim to reduce overall data transmission, the communication cost associated with transmitting prediction errors or raw data exceeding thresholds needs to be balanced with the gains in reduced redundant data transmission.

3.4 Temporal Correlation

3.4.1 Energy Conservation in WSNs

The limited battery life of sensor nodes poses a significant challenge in Wireless Sensor Networks (WSNs) [20]. This paper explores temporal correlation exploitation, a powerful technique that leverages the inherent redundancy in sensor data collected over time to achieve this goal.

3.4.2 Temporal Correlation and its Exploitation

Temporal correlation refers to the tendency of sensor readings to exhibit similar values over short time intervals. Several algorithms have been developed for temporal correlation exploitation in WSNs. We discuss two common approaches:
Threshold-based Algorithms: These algorithms define a threshold value (δ). If the difference between the current sensor reading (S(t)) and the previously transmitted reading (S(t-1)) is below the threshold, the data is deemed redundant and will not be transmitted [1].

Formula: Transmit data only if-

$$|S(t) - S(t-1)| > \delta \quad \textbf{(11)}\ [1]$$

Predictive Algorithms: These algorithms predict future sensor readings based on past readings and statistical models. If the predicted value falls within a certain error margin (ε) of the actual reading, the data is deemed redundant.
Formula: Transmit data only if-

$$|S(t) - S'(t)| > \varepsilon \quad \textbf{(12)}\ [1]$$

Where:-

S'(t): Predicted value for time t.

The specific formulas and parameters used may vary depending on the chosen algorithm and application requirements.

Table 2: A brief comparison of various algorithms used in predictive models

Algorithm	Advantages	Disadvantages	Suitability for WSNs
ARIMA	Simple to implement Low computational complexity	Limited accuracy for non-linear patterns Requires pre-defined model order	Moderate Suitable for basic prediction tasks in WSNs with moderate resource constraints
Kalman Filter	Efficient for dynamic data with process noise	Increased complexity compared to ARIMA	Moderate Can handle dynamic data but may require more resources than ARIMA
ANNs	High prediction accuracy for complex relationships	High computational complexity Large training data requirements	Low Not ideal for resource-constrained WSNs due to high computational demands
Linear Regression	Easy to interpret Low computational cost	Limited to linear relationships Sensitive to outliers	Low Similar to ARIMA, suitable for basic linear prediction tasks but may not capture complex patterns
SVM	Effective for classification and non-linear data	Complex to tune hyperparameters May not be suitable for pure regression tasks	Low Primarily for classification tasks, not ideal for direct redundancy reduction in WSNs

3.4.3 Advantages of Temporal Correlation Exploitation

There are several compelling reasons to employ temporal correlation exploitation in WSNs:
Reduced Data Transmission: By eliminating redundant data transmissions, the technique significantly reduces energy consumption, leading to a prolonged network lifetime.
Improved Network Scalability: By minimizing data traffic on the network, temporal correlation exploitation can potentially handle a larger number of sensor nodes without compromising performance.
Extended Sensor Lifetime: Reduced communication translates to lower energy expenditure by individual sensor nodes, thereby extending their operational lifespan.

3.4.4 Applications in WSNs

Temporal correlation exploitation finds application in various WSN deployments, including:
Environmental Monitoring: Sensor readings for temperature, humidity, and pressure often exhibit slow temporal variations, making this technique highly effective.
Structural Health Monitoring: In monitoring bridges or buildings, sensor readings typically show gradual changes, allowing for efficient data reduction.
Target Tracking: While target location may change over time, the movement is likely to be gradual, enabling this technique to reduce redundant location updates.

3.4.5 Implementation Parameters

The effectiveness of temporal correlation exploitation hinges on several key parameters:
Sampling Rate: The frequency of data sampling significantly impacts the technique's performance. A higher sampling rate captures more detailed information but reduces redundancy reduction potential.
Threshold Value (δ) or Error Margin (ε): These parameters determine the sensitivity of the technique. A stricter threshold (lower δ or ε) transmits more data but reduces redundancy, while a looser threshold (higher δ or ε) transmits less data but risks missing important changes.
Data Compression Techniques: Integrating temporal correlation with data compression techniques can further improve efficiency by minimizing the size of the transmitted data packets.

3.4.6 Implementation Methods

There are two primary implementation methods for temporal correlation exploitation:
Local (in-node) Processing: In this approach, individual sensor nodes perform the necessary computations and comparisons (threshold-based) or predictions (predictive algorithms) to determine if data transmission is necessary.
In-network Processing: This method aggregates data from multiple sensor nodes and performs the correlation analysis at a central node or aggregator node.
The selection of an implementation method is influenced by factors such as network topology, the processing capabilities of sensor nodes, and the intended level of data aggregation.

3.4.7 Future Research Prospects for Temporal Correlation Exploitation in WSNs

Temporal correlation exploitation in WSNs reduces redundancy, with future research focusing on enhancing its efficiency and broader applicability. Here, we explore some promising directions:

- **Deep Learning for Adaptive Correlation Analysis:** Current algorithms use pre-defined thresholds for correlation analysis.
- **Hybrid Approaches with Compressed Sensing:** Integrating temporal correlation exploitation with compressed sensing could enhance sparse signal acquisition and reconstruction**.**
- **Exploiting Spatial and Temporal Correlations:** Future research could explore techniques that jointly exploit spatial and temporal correlations in dense WSN deployments to reduce data redundancy.
- **Security Considerations for Correlation Analysis Techniques:** Implementing temporal correlation exploitation algorithms may create security vulnerabilities in WSNs, allowing malicious actors to manipulate data.
- **Energy-Aware Algorithm Design:** Exploring algorithms computing techniques could minimize the energy consumption of temporal correlation exploitation in WSN, despite reduced data transmission.

By investigating these promising research avenues, we can improve the efficiency and applicability of temporal correlation exploitation in Wireless Sensor Networks (WSNs). This will ultimately result in the creation of more resilient, energy-efficient, and secure sensor networks capable of gathering and transmitting essential data over extended periods.

4. COMPARATIVE ANALYSIS OF TECHNIQUES FOR DATA REDUNDANCY REDUCTION IN WIRELESS SENSOR NETWORKS (WSNs)

In Wireless Sensor Networks (WSNs), managing data redundancy is crucial for optimizing network efficiency, energy consumption, and overall performance. Various techniques address these challenges, each with unique advantages and limitations.This enhances energy efficiency and scalability, especially in high-correlation scenarios.

Data compression encodes sensor data more efficiently, reducing transmission load while conserving resources. However, techniques vary in compression ratios and computational complexity, with lossy methods potentially compromising data fidelity. Predictive modeling uses historical data to forecast future values, allowing nodes to transmit only prediction errors, which effectively reduces redundancy but may struggle in dynamic environments and require significant computational resources.

Spatial and Temporal Correlation Exploitation identifies and eliminates redundant sensor data, enhancing energy efficiency and minimizing unnecessary transmissions. While effective in predictable environments, it can struggle with heterogeneous data distributions. Choosing the right redundancy reduction technique in Wireless Sensor Networks (WSNs) depends on data

characteristics, application needs, and computational constraints, requiring careful evaluation for optimal performance and resource utilization.

Table 3: Various algorithms and techniques for data redundancy reduction

Technique	Best Suited	Efficiency	Scalability	Advantage	Accuracy	Disadvantage	Suitability
Data Aggregation	High spatial and temporal correlation datasets	High	High	Simple, low-complexity	Depends on function, high for basic statistics	Information loss, limited for complex data	Basic redundancy reduction
Data Compression	Applications tolerating data fidelity loss, uniform data distribution	High	High	High compression ratio	High for moderate compression	Increased complexity	Various data types
Predictive Modelling	Strong temporal correlation, predictable data patterns	High	Moderate	Effective for temporal correlation, reduces overhead	Varies by model, high for stationary data	Training data, complex models	Data with strong temporal trends
Spatial and Temporal Correlation Exploitation	Predictable spatial and temporal data	Moderate	Moderate	Captures both spatial & temporal redundancy	Depends on correlation strength	Complex algorithms, processing power	Highly correlated data

5. CONCLUSION

Wireless Sensor Networks (WSNs) are crucial in a range of applications, including environmental monitoring and industrial automation. However, one major challenge in deploying WSNs is ensuring energy efficiency, as sensor nodes have limited battery life. Data transmission is a major factor for energy drain, so minimizing redundant data transmissions is crucial for extending network lifetime. This paper explores various energy-efficient strategies aimed at optimizing WSN performance.

Researchers aim to enhance performance and lifetime by employing data redundancy reduction techniques, optimization algorithms, and other energy-efficient strategies. Techniques like data aggregation and compression reduce transmitted data volume while improving accuracy and processing efficiency, significantly boosting overall network performance.
As WSN technology continues to evolve, advancements in hardware design, communication protocols, and data processing techniques will further contribute to achieving optimal energy efficiency in these versatile sensor networks.
In summary, by employing data redundancy techniques and utilizing optimization algorithms, we can significantly lower energy consumption and enhance the overall efficiency of Wireless Sensor Networks.

6. FUTURE SCOPES

While significant advancements have been made in energy-efficient techniques for WSNs, there's immense potential for further exploration and innovation. Here, we delve into some promising future research directions:

- Artificial Intelligence and Machine Learning for Dynamic Optimization
- Energy-Harvesting Advancements
- Security Considerations for Energy-Efficient Techniques

By actively pursuing these promising research areas, we can improve the energy efficiency, extend the operational lifespan, and strengthen the overall security of Wireless Sensor Networks, facilitating their broader use in various essential applications.

REFERENCES

[1] . I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “Wireless sensor networks: a survey.,” *Computer networks,* vol. 38, no. 1, pp. 393-422, 2002.

[2] S. Bandyopadhyay and E. J. Coyle, “An energy efficient hierarchical clustering algorithm for wireless sensor networks.,” *IEEE INFOCOM,* vol. vol. 3, pp. 1713-1723, 2003.

[3] S. Moh, C. Yu, B. Lee and H. Youn, “Energy efficient and robust multicast protocol for mobile ad hoc networks,” *Pacific Rim International Symposium on Dependable Computing,* pp. 145-152, 2002.

[4] A. Masood, “Survey on Energy-Efficient Techniques for Wireless Sensor Networks.,” *arXiv preprint arXiv:2105.10413,* no. 1, 2021.

[5] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Communications of the ACM,* vol. 5, pp. 51-58, 2002.

[6] C. Schurghers, V. Raghunathan, S. Park and M. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Processing Magazine,* vol. 2, pp. 40-50, 2002.

[7] S. Kokilavani, N. S. Kumar and A. S. Narmadha, "Study and Analysis of Energy Efficient Data Aggregation Techniques for Wireless Sensor Networks," *IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India,* pp. 1-5, 2022.

[8] S. S. G and D. S. M. Sundaram, "Data Aggregation Techniques Over Wireless- A Review," *International Journal of Engineering Research & Technology (IJERT),* 2017.

[9] G. Kollios, J. Byers, J. Considine, M. Hadjieleftheriou and F. Li, "Robust Aggregation in Sensor Networks".

[10] L. A. Fitriya, T. W. Purboyo and A. L. Prasasti, "A Review of Data Compression Techniques," *International Journal of Applied Engineering Research ISSN,* vol. 12, no. 1, pp. 8956-8963, 2017.

[11] A. Preethi and A. Ramakrishnan, "A Survey on Data Compression Techniques in Wireless Sensor Network," 2015.

[12] T. Sheltami, M. Musaddiq and E. Shakshuki, "Data compression techniques in Wireless Sensor Networks," *Future Generation Computer Systems,* vol. 64, pp. 151-162, 2016.

[13] K. L. Ketshabetswe, A. M. Zungeru, B. Mtengi, C. K. Lebekwe and S. Prabaharan, "Data Compression Algorithms for Wireless Sensor Networks: A Review and Comparison," *IEEE Access,* vol. 9, 2021.

[14] F. F. JURADO-LASSO, L. MARCHEGIANI, J. F. JURADO, A. M. ABU-MAHFOUZ and X. FAFOUTIS, "A Survey on Machine Learning Software-Defined," *IEE Access,* 2022.

[15] R. Rajagopalan and P. K. Varshney, "Data-aggregation techniques in sensor networks: A survey," *IEEE Communications Surveys & Tutorials,* vol. 8, pp. 48-63, 2006.

[16] S. Famila, A. Jawahar and S. Vimalraj, "Integrated Energy and Trust-Based Semi-Markov Prediction for Lifetime Maximization in Wireless Sensor Networks," 2021.

[17] J. Fattah, E. Latifa, Z. Aman and H. E. Mousaami, "Forecasting of demand using ARIMA model," *International Journal of Engineering Business Management,* vol. 1, p. 5, 2018.

[18] Z. Ziqiao, M. S. T., W. Wencen and Z. Fumin, "Distributed cooperative Kalman filter constrained by advection–diffusion equation for mobile sensor networks," *Frontiers in Robotics and AI* , vol. 10, 2023.

[19] A. Mehta, J. K. Sandhu and L. Sapra, "Machine Learning in Wireless Sensor Networks: A Retrospective," *Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC),* 2020.

[20] M. A. M. Vieira, C. N. Coelho, D. C. Da Silva and J. M. da Mata, "Survey on wireless sensor network devices.," *EFTA,* vol. 1, pp. 537-544, 2003.

[21] N. Kimura, V. Jolly and S. Latifi, "Energy restrained data dissemination in wireless sensor networks," *International Journal of Distributed Sensor Networks,* pp. 251-265, 2006.

[22] J. Corchado Rodríguez, B. P. J. M., D. I. Tapia Martínez and . A. P. Abraham, "Using Heterogeneous Wireless Sensor Networks in a Telemonitoring System for Healthcare," 2010.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-20] Metaverse: Survey on future of internet and its innovative applications

Ansh Sindhwani[1,†], Sagar Ruhela[1,†], Anshul Bhatia[1,†], Ruchika Lalit[1,†] and Rahul Johari[1,*]

[1]*SWINGER : Security, Wireless, IoT Network Group of Engineering and Research,*
University School of Automation and Robotics(USAR),
Guru Gobind Singh Indraprastha University, Delhi, India

Abstract

With key technologies like artificial intelligence, blockchain, internet of things and upcoming network capabilities of 6G mobile communication the current idea of the internet is bound to change. We expect a more immersive and interactive experience to enter our daily internet experience. Further creation of persistent, scalable and shareable virtual spaces will lead to the birth of Metaverse. The word Metaverse has been popularized in the common lexicon when the big tech giant 'Facebook' transformed itself as 'Meta'. Though the word itself traces its origin to Neal Stephenson's seminal 1992 novel, "Snow Crash", where it was introduced as a shared virtual reality space. Today the idea of metaverse remains similar where humans interact in a shared virtual environment that is beyond the current understanding of social media and internet browsing. The paper presents a survey on the future of the internet and considers its diverse applications, these applications include healthcare, education, digital tourism and digital twin which are pertinent to a modern idea of a more connected smart city. Through case studies like : "Gucci Visions" we aim to throw light on the current developments and future implications of this upcoming iteration in domain of Metaverse.

Keywords

Metaverse, Technology Integrations, Blockchain, Web3.0, City Digital Twin

* Corresponding author.
† These authors contributed equally.
rahul@ipu.ac.in(R. Johari); anshsindhwani007@gmail.com (A. Sindhwani); sagarruhela89@gmail.com (S. Ruhela); anshulbhatia.usar@ipu.ac.in(A.Bhatia);
0000-0002-7675-8550 (R. Johari); 0009-0007-9565-1277 (A. Sindhwani); 0009-0002-0615-5495 (S. Ruhela); 0000-0001-6243-2982(A.Bhatia); 0000-0002-9171-5014 (R.Lalit)

1. Introduction

Considered to be a post-reality universe, the 'Metaverse' is a compendium of perpetual and persistent multi-usability platform/ environment that imbibes a fusion of physical reality with the digital reality. Although, there is no consensual definition for metaverse in the published literature, we can view it as a new iteration of the internet that will utilize a 3D modeled virtual world that can be accessed through VR Headsets, Augmented Reality glasses or a browser, allowing for multimodal interactions with digital products/services/spaces. It will be enabled by technologies like blockchain, artificial intelligence and advanced network infrastructure.

As a future tech-enabling universe, the 'Metaverse' does not denote a specific technology; but it does indicate the evolving bandwidth that may encompass human-to-human, human-to-technology, human-to-products/services/spaces, so on and so forth pluriverse worlds. Metaverse is evolving through unique, innovative and tech-enabled methodologies to aid the future of communications and interactions in real-time. In the realm of the fast-evolving communications and aesthetics associated with the Open-Source Platforms, there are new challenges in the contemporary iterations of the Metaverse: featuring especially the social, immersive VR platforms that are compatible with massive multiplayer online video games, open game worlds and equivalent AR collaborative space(s) and Brain- Computer Interfaces (BCIs) / Brain-Machine Interfaces (BMIs)/ smart brain.

The metaverse traces its roots to Neal Stephenson's seminal 1992 novel, "Snow Crash," where it was introduced as a shared virtual reality. This initial conceptualization, though born of fiction, sparked the imagination of technologists and creators. The subsequent evolution of technology, particularly within the gaming sphere, began to transform this fiction into reality. Early instances of the metaverse were encapsulated within multiplayer online games, serving as precursors to the expansive interconnected digital landscapes we envision today. Games like Second Life[25] pioneered user-driven persistent virtual environments, where social interactions, creativity, and commerce flourished. These early endeavors laid the groundwork for the metaverse's development beyond entertainment.

Gaming, a cornerstone of the technological realm, played a pivotal role in the metaverse's evolution. Virtual worlds, exemplified by iconic titles such as World of Warcraft[26], Roblox[27] and Minecraft[28], emerged as foundational spaces that demonstrated the core attributes of the metaverse. These platforms transcended mere games, serving as interconnected realms where participants collaboratively shaped their experiences. The success of these gaming platforms highlighted the allure of immersive digital environments that facilitate real-time interactions, creativity, and exploration. Users found themselves drawn into these virtual landscapes, experiencing a sense of presence and connectivity that foreshadowed the metaverse's broader potential.

The concept of the metaverse has expanded beyond gaming and entertainment, infiltrating diverse sectors and applications. Social media platforms are increasingly incorporating metaverse-like features, enabling users to engage in virtual events, interactions, and commerce. Moreover, the advent of advanced technologies such as virtual reality (VR) and augmented reality (AR) has propelled the metaverse's progression. As these technologies continue to evolve, the metaverse's promise becomes even more apparent. Interconnected virtual worlds and digital identities hold the potential to revolutionize not only entertainment but also education, collaboration, commerce, societal interactions and even management of a city. The evolution from

science fiction to this burgeoning digital reality underscores the metaverse's significance as a transformative force that stands poised to reshape the way we connect, learn, and experience the digital world. It has implications in sectors like tourism, governance, traffic management, environmental sustainability and urban planning.

2. Review of Literature

The formal text explains that Metaverse comprises multiple components namely - hardware components, software components and content. The hardware consists of head mounted displays and input devices. Software cater to the need of motion rendering, sound and speech recognition, scene/object recognition and scene/object generation. The third component is the content. Content is the pillar of Metaverse that is crucial for providing the immersive experience promised by the idea.[1] The Metaverse relies on content as its foundational element, ensuring an engaging experience through well-crafted narratives and user-generated activities. Focusing on story authenticity, immersive engagement, and conceptual coherence, content creation can be approached through a paradigm shift or by repurposing existing material. Environmental design encompasses scene creation, color and lighting, audio elements, sampling and aliasing, navigational aspects, and real-life content. User movements, characters, and avatar personalities all contribute to behavioral modelling.

However, achieving the proposed metaverse idea is a multi-staged process. As elaborated by [2] achieving duality will require sequential stages, namely (I) digital twins, (II) digital natives, and eventually (III) co- existence of physical-virtual reality or namely the sureality. Digital twin refers to a replica of real-world objects and environments on digital platforms. They try to simulate the real characteristics of an object in the digital world. Once these digital copies are made the next stage is to create “digital natives.” This basically refers to creation of new content exclusive to the metaverse realm. In the third and last stage, the metaverse could become a self-sustaining and persistent virtual world that co-exists and interoperates with the physical world with a high level of independence. As such, the avatars, representing human users in the physical world, can experience heterogeneous activities in real-time characterized by unlimited numbers of concurrent users theoretically in multiple virtual worlds.

The role of blockchain technology is critical in realizing the shared virtual reality. In a shared virtual environment like metaverse, blockchain technology will provide a means of digital ownership. It enables trusted ways of proving authenticity as well as interoperability. A blockchain infrastructure can ensure that authentic data is loaded to the metaverse. Blockchain will also play a pivotal role in establishing the metaverse economy by providing a means of exchange in the digital world. [3]

Immutable Identifiers i.e. NFTs in Web3.0 open new challenges, but also dynamic experiences. Several industrial leaders have stated that the Metaverse’s full potential can only be attained if it is constructed on open standards. The Metaverse Standards Forum provides leading standards organizations and companies a platform to foster interoperability standards for an Open Metaverse. An NFT, for example can be utilized in a 2D Smartphone game or an 3D virtual headset apart from being shown in digital and physical art galleries.

Immutability feature of blockchains makes it a significant factor driving in new digital and hybrid economic activities. It allows for scalable and reliable methods to demonstrate authenticity and interconnectivity.

In [4] authors introduce Metaverse-as-a-service where they suggested a concept of metaverse platforms as a service to business owners and alike over the network. Facilitating scalable computing resources on demand over the internet is a common facility available to business users and individual users. Using products like google drive, office 365 is not uncommon. All these follow an "as-a-service" model under the aegis of cloud computing. similarly, we can offer services relevant to Metaverse where key technologies related to Metaverse can be provided. Services such as platforms, infrastructures, software, and artificial intelligence (AI), could be provided as through Metaverse-as-a-service (MaaS). As FAANG companies enter the Metaverse space, Like Microsoft, Samsung, NVIDIA, and others, it won't be long before many cloud-based services would offer Metaverse-as-a-Service (MaaS) option, allowing businesses to profit from the technology with lowered entry barriers.

3. Applications of Metaverse

Introduction of augmented reality and virtual reality in gaming and entertainment is the starting point of a more immersive internet and better human computer interfaces. Games like Pokémon Go, Fortnite and Minecraft provide a better social gaming environment allowing users to collaborate and compete in real time. Hence, entertainment and gaming are natural applications of metaverse. But the applications of metaverse are not limited to gaming and entertainment. Various studies regarding potential applications have been published. These applications include healthcare, advertising, education and e-commerce to name a few. Metaverse has a potential of changing our life, society and economy through its innovative and immersive applications.

I. EDUCATION

Education and experiential learning through metaverse are at the forefront of these applications. One can imagine being trained in virtual environments and through immersive simulations before practicing in the physical world. This will increase efficiency and proficiency of the apprentice. Learning in the metaverse will also lower the cost of training and lower the time of training as the apprentice will not be limited by the availability of teachers. Metaverse learning will expand the reach of education as it will not be limited to geographical bounds catering to a wider audience.

In [5] authors highlight the possibilities and limitations of Metaverse education. The authors also introduce the idea of 4 types of metaverse that can be used for educational purposes. These are (1) augmented reality, (2) virtual reality, (3) mirror world and (4) lifelogging. These modes of metaverse can be successfully implemented for creating virtual labs, training modules and overlaying on the physical world for an enhanced pedagogy. Students can learn visually through 3-dimensional digital twin, overlay information on physical objects, interact with artificial intelligence enabled characters of historic figures, create virtual tours of historic places and make science models in virtual reality to gain hands-on experience and learning.

Authors of [8] created a virtual campus of Nanyang Technical University, Singapore through the Nvidia omniverse platform called NTUniverse. NTUniverse allows students to engage in virtual

hands-on activities, online discussions, and collaborative team projects. It serves as a platform for interactive and immersive learning experiences in STEM education. The platform also supports the learning of abstract mathematics concepts through interactive manipulation of virtual objects. This approach helps students grasp complex ideas like Vector Cross Product and Skew line Distance by interacting with virtual models.

II. HEALTHCARE

Healthcare in metaverse points towards to the virtual environment consisting of various technologies like augmented reality (AR), virtual reality (VR), and extended reality (XR) to give the healthcare its functionality and medical process. In the healthcare virtual simulation or healthcare metaverse people can conduct lots of thing from their home like conducting more immersive and real consulting with doctors, patients can meet with doctors in 3d world. Use of digital twins for surgery planning and real time analysis and monitoring. These technologies offer innovative solutions for telemedicine, medical education, mental health support, and clinical care which is used for solving common person problem way better than before. On top of that healthcare in metaverse enables lots of immersive training for medical students through which the training is fast, accurate and precise. Inside the Virtual world students can practice surgeries and other medical procedures. The author(s) of [24] discuss a partnership with 8chili [30], a California-based software business and Global Healthcare Academy (GHA) aims to establish an EdTech platform that will enable medical education and healthcare to be accessible across the metaverse. More than 200 hours (about one week and a half's worth) of virtual reality (VR) content covering a wide range of subspecialties, such as paramedics, cancer, dentistry, orthopedic, neurology, spine, and ENT, have already been released by the alliance. The author(s) of [24]also describes an EdTech platform called Studyum[29] that offers spaces where group projects for student training or education would be organized. It encourages cooperation throughout society by using a reward system. Users receive tokens for each class they attend, video they watch, and homework they turn in. As a result, it offers various learning materials and training programs according to performance and training. An example of a trainer may be a "celebrity surgeon," who would be paid for his education and provide awards to pupils in honor of their accomplishments.

III. SMART CITIES AND DIGITAL TWIN OF CITIES

Smart Cities are going to use digital technology and data driven approach to make our lives and work easy. We can access lots of services from our home with the help of smart cities. Some of those services includes meeting people online, intercountry meeting, business opportunities, smart governance, more enhance online shopping and entertainment. With the help of metaverse and others technologies like artificial intelligence, digital twins, augmented and virtual reality, blockchain, and cloud computing, the concept of smart cities is possible which can also be used to create an interconnected and immersive virtual environment in which users or common public

easily interact with the help of avatars. Now the concept of smart cities not just concept but governments around the world are making cities smart by investing huge sum of money example Seoul (The capital city of South Korea) plans to invest $3.3 billion to become metaverse city, even UAE also announced its plans towards making a virtual or smart city. There are lots of benefits from smart cities like improving teamwork with the help of real time collaboration with the help of Extended reality (XR), new business opportunities with the help of Virtual Reality(VR).

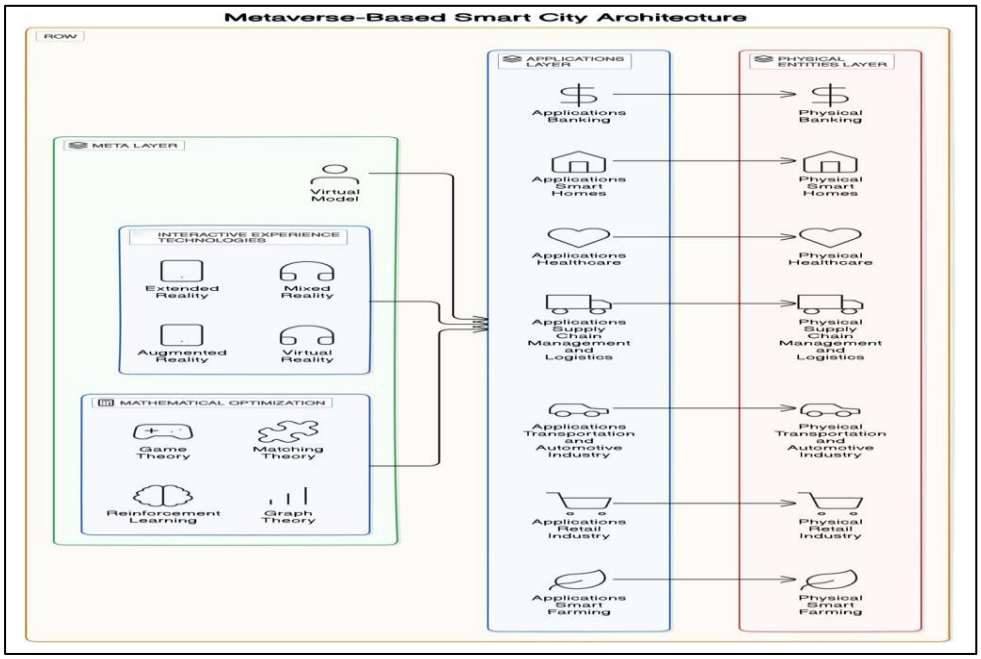


Figure 1: Metaverse Based Smart City Idea

I. DIGITAL TOURISM AND RECREATION

Digital tourism in metaverse refers to use of virtual reality to mimic the travel experience for user from their home. This will allow user to explore and interact with digital models of real-world monuments and exotic destinations. This concept leverages various immersive and scalable technologies. These include the use of artificial intelligence, AR/VR and 3D modelling. The concept goes beyond making a digital twin of real-world monument and opens doors for making new virtual experiences. People can travel virtually with friends or other users in the metaverse, joining group tours or exploring destinations together, replicating the social aspects of physical travel. The metaverse can also allow tourists to travel through time to witness historical events. Digital tourism brings accessibility to tourism for people with physical disabilities and financial constraints.

II. DIGITAL RETAIL AND E-COMMERCE

Digital retail and ecommerce through Metaverse bring a new stage of online shopping, with virtual environment providing tailored shopping experience. In metaverse customers can browse virtual stores and try products virtually. Metaverse allows user to purchase both NFTs and real-world goods that would be

delivered at customer's location after purchase. The retail stores can be equipped with AI powered sales avatars for a more realistic and immersive experience.

Figure 2: Evolution of commerce with metaverse

4. Key Technologies related to Metaverse Development

Metaverse is an amalgamation of multiple existing technologies which serve as fundamental pillars to metaverse and its immersive nature. The key technologies that are driving the democratization of metaverse include (i)Artificial Intelligence, (ii)virtual reality and augmented reality, (iii)cloud computing, (iv)5G internet, (v)blockchain and (vi) internet of things. These technologies provide scalability for virtual environment, immersion, help in content creation and provide interconnections

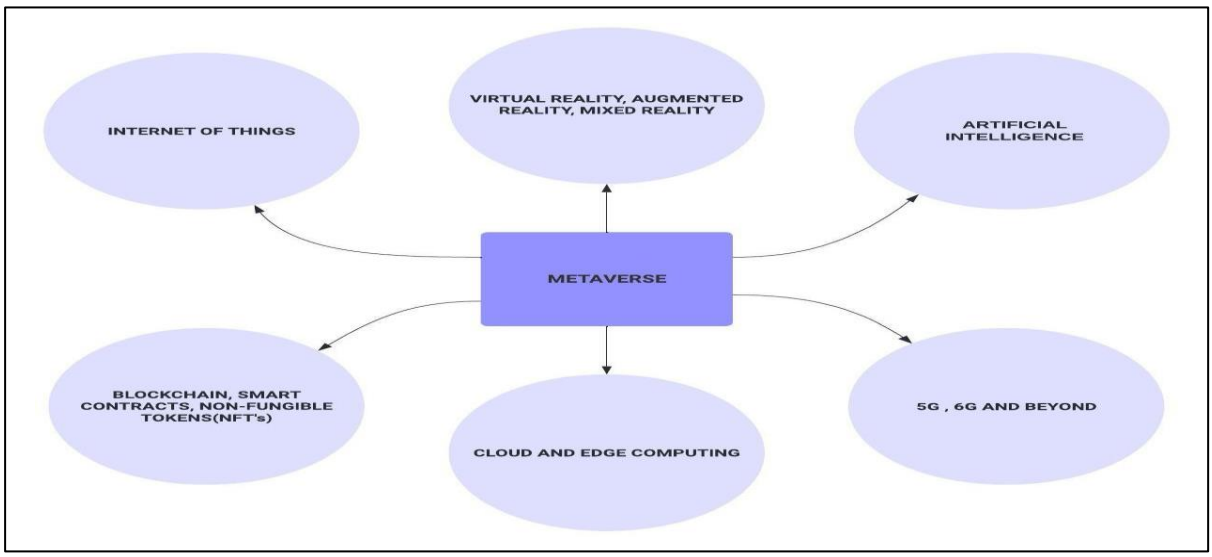


Figure 3: Key Technologies for Metaverse

These technologies are pivotal to provide features that are pivotal to metaverse and beyond reality experience. These key metaverse features are based on literature review with brief description and presented in the table 1.

Feature	Description
Immersion	Allowing users to interact and engage through multiple channels, these mediums can be virtual reality, augmented reality, chatbots, 3d-game objects. This can be obtained using AI, AR/VR, haptic technology
Realistic	The experience of metaverse should be close to real world. Use of game physics simulations should be implemented.
Digital Ownership	the system should be able to have protocols and data structures enabling digital ownership in decentralized manner. This can be achieving through blockchain technology
Scalable	The system should be able to support the growing number of avatars, virtual worlds, 3d-objects and users
Interoperable	Able to communicate with other virtual environments and systems, Avatars or digital user should be able to switch between virtual environments with same virtual identity

Table 1: Description of metaverse features

5. Current progress in Metaverse

Currently Metaverse is in developmental stage and big companies are betting on its success. Notably "Meta" is extensively working on building virtual environments and better hardware for accessing the Metaverse. This hardware includes better AR and VR headsets as well as new haptic based hardware like haptic gloves, haptic suits etc. for better interaction with the metaverse.
Games like Roblox, Fortnite and Minecraft are working on there VR versions as a step forward towards metaverse. Today most major companies have a VR or AR headset offerings. Microsoft HoloLens [31], Apple's Vision pro [32] and Meta quest [33] are some prominent examples. Sandbox [34], decentraland [35] and somium space [21] are some prototypes of metaverse.

6. Case Study

Considering the upcoming changes towards metaverse the team analyzed "Gucci Visions" and as precursors towards the new iteration of internet. The team has presented a use case diagram, flow diagram and a sample algorithm which can serve as intuitive models for creation of a more immersive and experiential internet that is the metaverse.

Gucci Visions is virtual model of Gucci's premium store in Italy which aims to portray the company's vision for the future, Metaverse and virtual reality visits of the store. it gives a

digital/virtual tour of "Palazzo Della Mercanzia", it sets a precedence for Metaverse tourism, metaverse marketing and E- commerce. The model navigates through the store and provide an immersive experience of the store through virtual reality. This serves as a precursor to the realm of digital tourism through metaverse. The team studied the virtual environment and have presented a use case diagram in figure 2 and a flow diagram in figure 3. These virtual tours can be beneficial for smart cities.

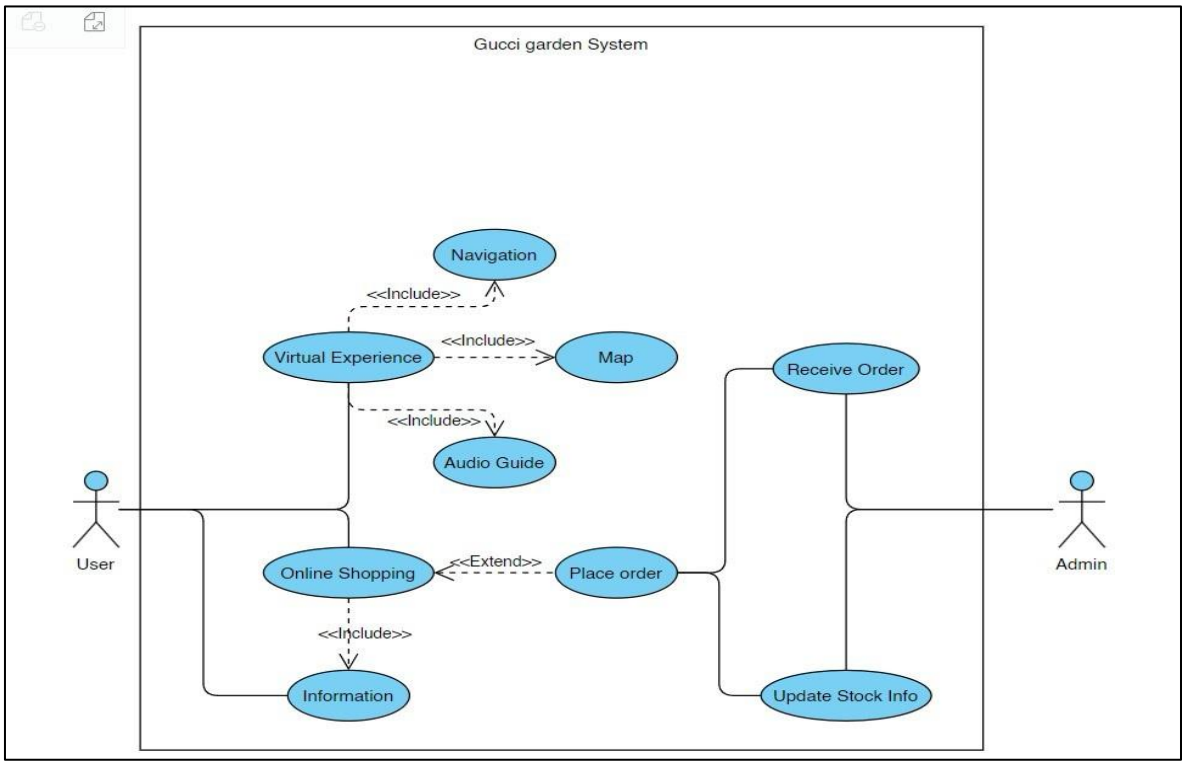


Figure 4: use case diagram of Gucci Garden

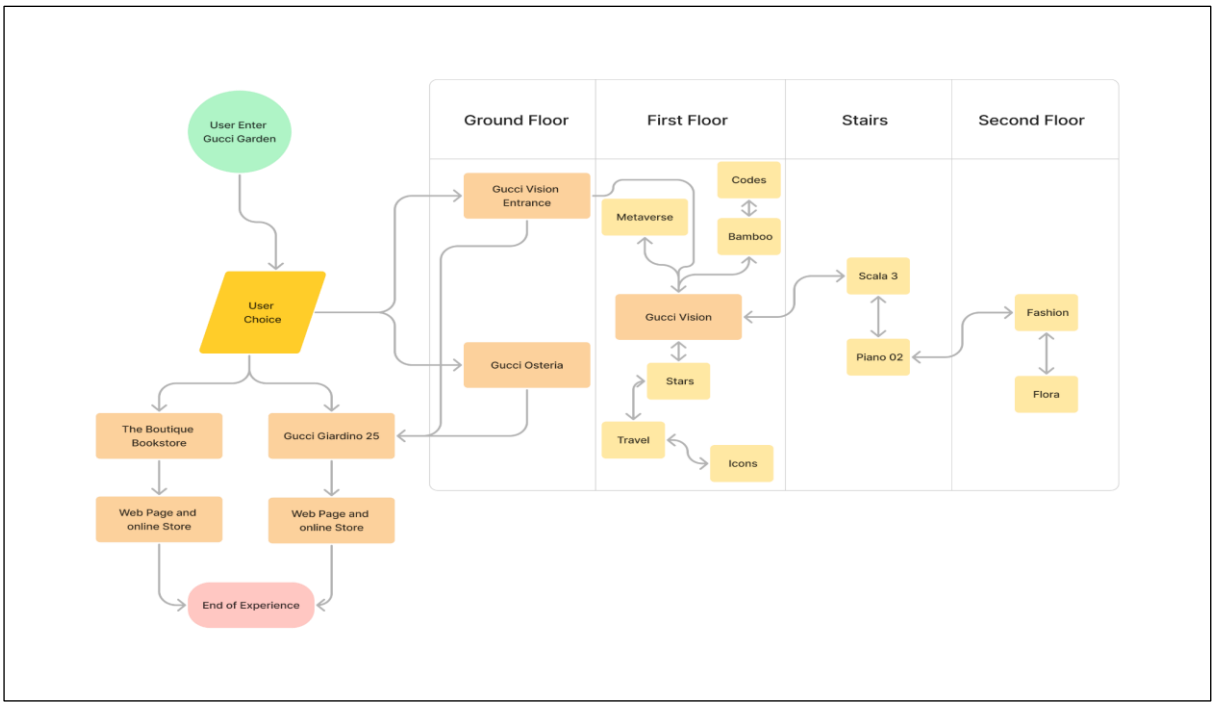


Figure 5: flow diagram of Gucci Garden

Algorithm : Metaverse Algorithmic framework for Gucci Garden Case Study
[URL : https://guccipalazzo.gucci.com/#/en/360/piano-1/travel]

Notation

1: S_r – Server
2: B_r – Metaverse Broker; where B_r is with in S_r
3: F_r – Frame
4: S_f – Frame per second
5: A_v – Avatar
6: A_v [ ] – Array of Avatar where every $A_v \in A_r$
7: VA_s [ ] - Values Table of Avatar at Server
8: VW_s [ ] - Virtual Walkthrus
9: M_q – Manenquin
10: M_q [ ] – Array of Manenquins in Studio
11:G_s - Gucci Stores
12: G_v – Gucci Visions
13: G_o – Gucci Osteria

14: G_g – Gucci Giardino
15: $V\ G_{vl}$ [ ] – Glasses/Lenses available in Gucci Visions
16: $V\ G_{lp}$ [ ] – Array of products available in Gucci Osteria
17: $V\ G_{gp}$ [ ] – Array of products available in Gucci Giardino
18: C_r – Customer
19: C_{rb} – Customer Bucket

Pre-requisite :
Customer enters Gucci Store(G_s)

Trigger :-
20: A_v welcomes C_r
21: C_r set up S_f
22: A_v prompts C_r for valid G_s ; where in valid $G_s \in [G_v \,||\, G_v \,||\, G_v]$
23: if G_s is G_v
24: {
25: C_r interacts with M_q during VW_s [ ] in G_v
26: $C_{rb} \leftarrow V\ G_{vl}$ [ ] }
27: elseif (G_s is is G_{lp})
28: {
29: $C_{rb} \leftarrow V\ G_{lp}$ [ ] }
30: else
31: {
32: $C_{rb} \leftarrow V\ G_{gp}$ [ ]
33: }
34 : else
35: {
36: print ("Not a Valid G_s");
37: print("Try again");
38: }
39: Sold Item information updated with $B_r(S_r)$;

7. Challenges to Metaverse

Currently development of Metaverse faces several challenges for its full realization. These challenges span across multiple dimensions. These challenges include technological, social, ethical and economic challenges. Table 2 presents the various challenges to Metaverse based on the reviewed literature.

Challenge	Description
Scalability	Creating a Metaverse that can handle millions of users parallelly across vast virtual spaces is a prominent challenge. Metaverse requires vast amount of 3d rendering, blockchain mining and implementation of AI models.
Hardware Limitation	Current hardware for virtual reality and augmented reality is limited in there capabilities. Currently these devices have limited battery life and are bulky. Hardware that supports Haptics is currently in early developmental stage.
Network speed and Latency	Metaverse needs high network bandwidth and low latency, even better than the currently available 5G mobile internet and Wi-Fi 6
Content Creation and Management	Metaverse relies on vast amount of 3D models, Digital environments, Animations and Avatars. Creation of these high-quality objects are time consuming and expensive. Managing this content is complex across platform. It requires help of generative AI and increased skilled workers.
Privacy and Security	This expanded iteration of internet is expose more public to new security threats and higher privacy concerns.
Digital Governance and regulations	With increased use of NFTs, Cryptocurrencies and an expanded digital environment, Metaverse require some oversite and policy changes to ensure citizens safety from cyber harassment, cyber frauds and prevent any occurrence of Criminal activity.

Table 2: Challenges to Metaverse

8. Conclusion

Metaverse represents a shift in how we will perceive internet and web based services in times to come. Its goal is to create a scalable virtual and augmented reality based ecosystem which is more immersive. It aims to blend the physical and digital world. As highlighted by this paper Metaverse converges technologies like artificial intelligence, blockchain, IoT, Virtual reality(VR), Augmented reality (AR). Frameworks “gucci garden” are trying to leverage the given idea and form a precursor to Metaverse.

Smart Cities can leverage Metaverse for citizen centric services, real time information sharing and promote digital tourism. They can use metaverse for virtual city council meetings and monitor various services through a synergy of physical and digital world. The aforementioned Case study showcase early implementations of Metaverse that point towards a broader future where virtual environments become essential to industries like retail, tourism, urban planning, and entertainment.

References

[1] Park, Sang-Min, and Young-Gab Kim. "A metaverse: Taxonomy, components, applications, and open challenges." *IEEE access* 10 (2022): 4209-4251.

[2] Lee, Lik-Hang, Tristan Braud, Pengyuan Zhou, Lin Wang, Dianlei Xu, Zijun Lin, Abhishek Kumar, Carlos Bermejo, and Pan Hui. "All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda." *arXiv preprint arXiv:2110.05352* (2021).

[3] Gadekallu, Thippa Reddy, Weizheng Wang, Gokul Yenduri, Pasika Ranaweera, Quoc-Viet Pham, Daniel Benevides da Costa, and Madhusanka Liyanage. "Blockchain for the Metaverse: A review." *Future Generation Computer Systems*143 (2023): 401-419.

[4] Ahsani, Vesal, Ali Rahimi, Mehdi Letafati, and Babak Hossein Khalaj. "Unlocking Metaverse-as-a- Service The three pillars to watch: Privacy and Security, Edge Computing, and Blockchain." *arXiv preprint arXiv:2301.01221* (2023).

[5] Kye, Bokyung, Nara Han, Eunji Kim, Yeonjeong Park, and Soyoung Jo. "Educational applications of metaverse: possibilities and limitations." *Journal of educational evaluation for health professions* 18 (2021).

[6] Mustafa, Bahaa. "Analyzing education based on metaverse technology." *Technium Soc. Sci. J.* 32 (2022): 278.

[7] Hwang, Gwo-Jen, and Shu-Yun Chien. "Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective." *Computers and Education: Artificial Intelligence* 3 (2022): 100082.

[8] Sim, Jing Kai, Kaichao William Xu, Yuyang Jin, Zhi Yu Lee, Yi Jie Teo, Pallavi Mohan, Lihui Huang et al. "Designing an Educational Metaverse: A Case Study of NTUniverse." *Applied Sciences* 14, no. 6 (2024): 2559.

[9] Chengoden, Rajeswari, Nancy Victor, Thien Huynh-The, Gokul Yenduri, Rutvij H. Jhaveri, Mamoun Alazab, Sweta Bhattacharya, Pawan Hegde, Praveen Kumar Reddy Maddikunta, and Thippa Reddy Gadekallu. "Metaverse for healthcare: A survey on potential applications, challenges and future directions." *IEEE Access* (2023).

[10] Dilibal, Cinay, and Yalcin Tur. "Implementation of developed esantem smart healthcare system in metaverse." In *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pp. 1027-1031. IEEE, 2022.

[11] Koo, Huilyung. "Training in lung cancer surgery through the metaverse, including extended reality, in the smart operating room of Seoul National University Bundang Hospital, Korea." *Journal of educational evaluation for health professions* 18 (2021).

[12] Kim, Jooyoung. "Advertising in the metaverse: Research agenda." *Journal of Interactive Advertising* 21, no. 3 (2021): 141-144.

[13] Samarnggoon, Keattikorn, Supara Grudpan, Noppon Wongta, and Konlawat Klaynak. "Developing a Virtual World for an Open-House Event: A Metaverse Approach." *Future*

*Internet*15, no. 4 (2023): 124.

[14] Wang, Fei-Yue, Rui Qin, Xiao Wang, and Bin Hu. "Metasocieties in metaverse: Metaeconomics and metamanagement for metaenterprises and metacities." *IEEE Transactions on Computational Social Systems* 9, no. 1 (2022): 2-7.

[15] Lim, Wei Yang Bryan, Zehui Xiong, Dusit Niyato, Xianbin Cao, Chunyan Miao, Sumei Sun, and Qiang Yang. "Realizing the metaverse with edge intelligence: A match made in heaven." *IEEE Wireless Communications* 30, no. 4 (2022): 64-71.

[16] Liu, Guangyuan, Hongyang Du, Dusit Niyato, Jiawen Kang, Zehui Xiong, Abbas Jamalipour, Shiwen Mao, and Dong In Kim. "Fusion of Mixture of Experts and Generative Artificial Intelligence in Mobile Edge Metaverse." *arXiv preprint arXiv:2404.03321* (2024).

[17] Huynh-The, Thien, Quoc-Viet Pham, Xuan-Qui Pham, Thanh Thi Nguyen, Zhu Han, and Dong-Seong Kim. "Artificial intelligence for the metaverse: A survey." *Engineering Applications of Artificial Intelligence* 117 (2023): 105581.

[18] Aslam, Anjum Mohd, Rajat Chaudhary, Aditya Bhardwaj, Ishan Budhiraja, Neeraj Kumar, and Sherali Zeadally. "Metaverse for 6G and beyond: The next revolution and deployment challenges." *IEEE Internet of Things Magazine* 6, no. 1 (2023): 32-39.

[19] Wang, Yuntao, Zhou Su, Ning Zhang, Rui Xing, Dongxiao Liu, Tom H. Luan, and Xuemin Shen. "A survey on metaverse: Fundamentals, security, and privacy." *IEEE Communications Surveys & Tutorials* 25, no. 1 (2022): 319-352.

[20] "Gucci visions", Gucci,last accessed on 27th june 2024 , https://guccipalazzo.gucci.com/#/en/.

[21] Somnium Space documentation, "Somnium space", last accessed on 30th august 2024 https://somniumspace.com/documentatio

[22] Yaqoob, Ibrar, Khaled Salah, Raja Jayaraman, and Mohammed Omar. "Metaverse applications in smart cities: Enabling technologies, opportunities, challenges, and future directions." *Internet of Things* (2023): 100884.

[23] Allam, Zaheer, Ayyoob Sharifi, Simon Elias Bibri, David Sydney Jones, and John Krogstie. "The metaverse as a virtual form of smart cities: Opportunities and challenges for environmental, economic, and social sustainability in urban futures." *Smart Cities* 5, no. 3 (2022): 771-801.

[24] Bansal, Gaurang, Karthik Rajgopal, Vinay Chamola, Zehui Xiong, and Dusit Niyato. "Healthcare in metaverse: A survey on current metaverse applications in healthcare." *Ieee Access*10 (2022): 119914-119946.

[25] "Official Site: Second Life - Virtual Worlds, Virtual Reality, VR, Avatars, and Free 3D Chat." Official Site | Second Life - Virtual Worlds, Virtual Reality, VR, Avatars, and Free 3D Chat. Accessed October 13, 2024. https://secondlife.com/.

[26] Entertainment, Blizzard. "World of Warcraft®: The War WithinTM Now Live!" World of Warcraft. Accessed October 13, 2024. https://worldofwarcraft.blizzard.com/en-us/.

[27] "Home - Roblox." Roblox Corporate. Accessed October 13, 2024. https://corp.roblox.com/.

[28] "Welcome to the Official Site of Minecraft." Minecraft.net. Accessed October 13, 2024. https://www.minecraft.net/en-us.
[29] "The Future of Decentralized Learning." Studyum. Accessed October 13, 2024. https://www.studyum.org/.
[30] "HINTVRTM Training and Engagement." 8chili. Accessed October 13, 2024. https://www.8chili.com/.
[31] "Microsoft Hololens: Mixed Reality Technology for Business." Microsoft HoloLens | Mixed Reality Technology for Business. Accessed October 13, 2024. https://www.microsoft.com/en-in/hololens.
[32] "Apple Vision Pro." Apple. Accessed October 13, 2024. https://www.apple.com/apple-vision-pro/.
[33] Meta Quest MR, VR Headsets & Accessories. Accessed October 13, 2024. https://www.meta.com/quest.
[34] "The Sandbox Game - Own the Future. Play, Create, Earn." The Sandbox. Accessed October 13, 2024. https://www.sandbox.game/en/.
[35] "Decentraland Player Documentation." Decentraland Documentation. Accessed October 13, 2024. https://docs.decentraland.org/player/.
[36] "WebXR Device Api." W3C, September 27, 2024. https://www.w3.org/TR/webxr/.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-21] Optimizing Dehusked Arecanut Quality Segregation: CNN-Based Approach with Contrast Enhancement and Data Augmentation

Sameer Patil[1ab,†], Aparajita Naik[2,†], Marlon Sequeira[3,†], Sulaxana Vernekar[4,†] and Jivan Parab[5,*,†]

[1a]*Electronics Programme, School of Physical and Applied Science, Goa University, Goa, India*
[1b]*Department of Electronics, Dnyanprassarak Mandal's College and Research Centre, Goa, India*
[2]*Electrical Engineering, Cambridge University, Cambridge, UK*
[3]*Electronics Programme, School of Physical and Applied Science, Goa University, Goa, India*
[4]*GVM'S GGPR College of Commerce and Economics, Goa, India*
[5]*Research Supervisor, Electronics Programme, School of Physical and Applied Science, Goa University, Goa, India*

Abstract

In the production process of Areca nut, the segregation stage is of prime importance. As of now, most commercial retailers use skilled workers for quality segregation, which means a lot of time is required for finalizing the product costing. Based on the inputs received from marketing executives, it was observed that if any method for automatic segmentation has to be meaningful, then the quality segregation should not have errors of more than 5% standard deviation. In this study, we propose a methodology based on 10-fold cross-validation training of Convolutional Neural Network (CNN) using contrast enhancement and no data augmentation of the images. Also, in this paper, we compare the results attained on the quality segregation using numerous processing methods, for instance, data augmentation for images with and without cropping and also for images with and without contrast enhancement. The database developed here uses Areca nut cultivated in the Western Ghats region of the Indian Peninsula, particularly focused on the Konkan belt. In our paper, we achieved the lowest standard deviation of 4.1% for cropped images with contrast enhancement.

Keywords

Areca nut, Segregation, Convolutional Neural Networks (CNN), Contrast Limited Adaptive Histogram Equalization (CLAHE)

1. Introduction

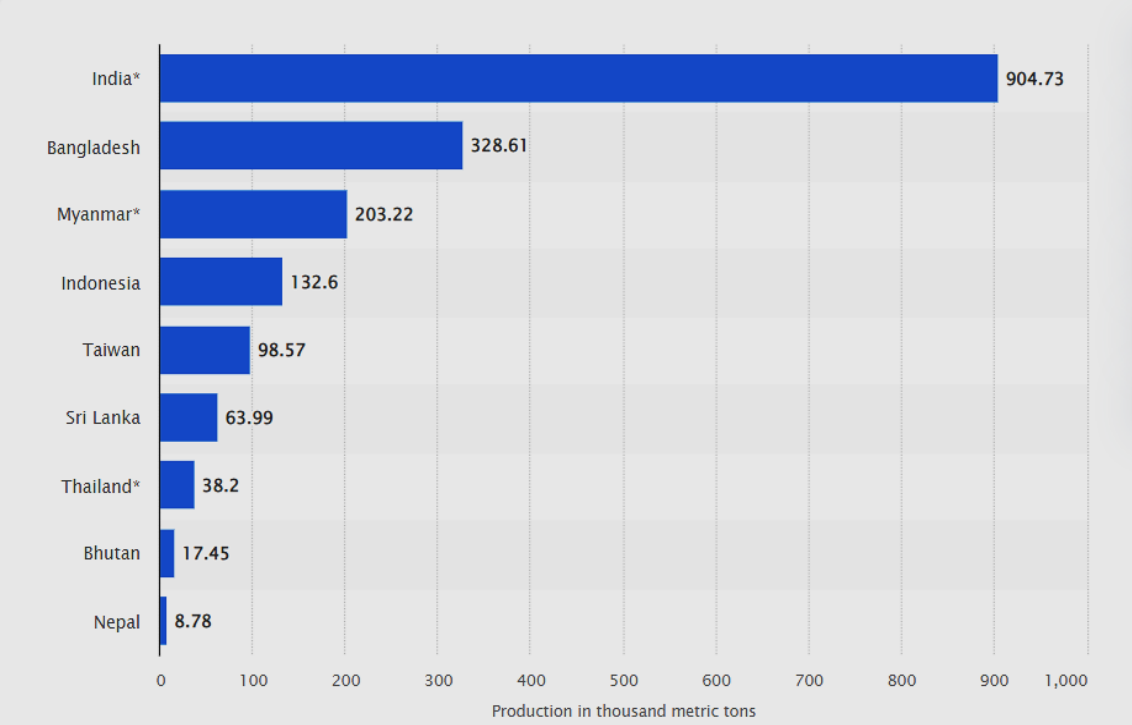
Areca palm (*Areca catechu L.*) is grown for its kernel, popularly known as Areca nut (or Betel nut or Supari) in India. It is grown commercially along the western coast of India (Maharashtra, Goa, Karnataka), Tamil Nadu, Kerala, Assam and West Bengal [1].

* Corresponding author.
† These authors contributed equally.
sameer@dmscollege.ac.in (S. Patil); an656@cam.ac.uk (A. Naik); marlon@unigoa.ac.in (M. Sequeira); sulaxgoa@gmail.com (S. Vernekar); jsparab@unigoa.ac.in (J. Parab)
0009-0007-3674-1942 (S. Patil); 0009-0002-0332-1654 (A. Naik); 0000-0002-1444-5428 (M. Sequeira); 0000-0002-0848-7349 (S. Vernekar); 0000-0002-7462-3492 (J. Parab)

It is a tropical crop grown commercially in Southeast and South Asian countries, parts of East Africa, and some tropical Pacific nations [2,3]. Since ancient times, it has been used as a masticator wherein a slice of Areca nut, along with slaked lime (aqueous Calcium hydroxide paste) and some aromatics (clove, cardamom, saffron) for extra flavor, is wrapped in a betel leaf [4]. In ancient Indian scripts, it is mentioned that Areca nut can be utilized as a therapeutic agent for leucoderma, leprosy, anemia, and obesity and has de-worming properties due to which it is used in gastrointestinal disorders [5]. Areca nut has many pharmacological properties and is extensively used in medicine. It has anti-allergic, anti-parasitic, anti-microbial, and anti-aging properties. Areca nuts are also used to produce adhesives, non-woven fabrics, textile dyes and building materials. Hence, due to its high economic significance, Areca nut has become an important cash crop.

As per the latest studies, India tops at the global level, contributing to approximately 904 thousand metric tons in 2020[6]. The top ten Areca nut producing countries over the globe are shown in Figure 1.1.

Fig 1.1: Areca nut production in Asia Pacific in 2020 by country (in 1000 metric tons) [6]



The Areca nut kernel is hard from outside with the inner endosperm marbled in dark brown and white [7]. The crucial steps in the areca nut production process are listed below.

Harvesting

1. Drying
2. De-husking

3. Nut segregation based on its quality.

Nut segregation is the most labor-intensive and time-consuming of the production process's aforementioned steps. In Goa (India), Goa Bagayatdar, a cooperative organization, is a leading Areca nut collector. At their collection centres, nuts are classified on the basis of texture, colour and the quality. Here, nuts are segregated in seven different categories (Supari, Safed, Laal, Vench, Kharad, Tukda and Baad) [8]. But, due to the shortage of skilled laborers for the above said work, it is essential to develop a unit of segregation based on its quality. This will not only solve the issue of scarcity of laborers but also will save farmer's time.

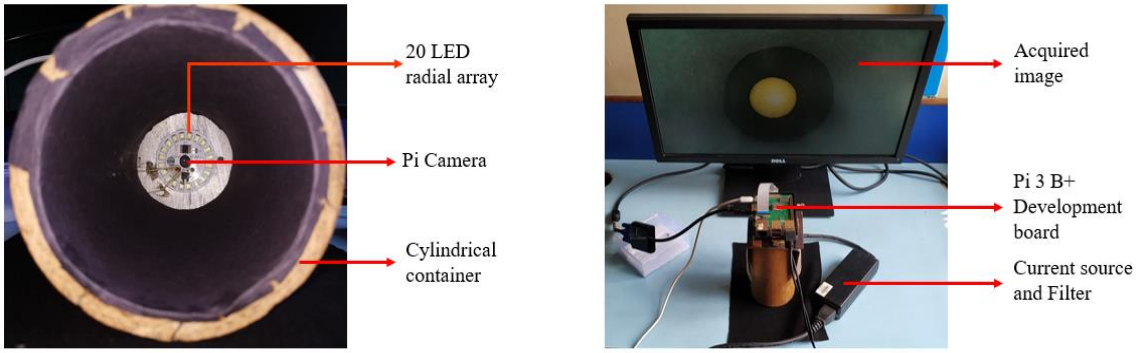
2. Literature Review

Much work is being done in machine learning and image processing to identify, categorize, and grade agricultural products. S. Siddesha et al., in their study of the texture-based classification of Areca nut, extracted different texture features using Wavelet, Gabor, Gray Level Difference Matrix, Local Binary Pattern (LBP), and Gray Level Co-Occurrence Matrix features. The Nearest Neighbor classifier was used to classify Areca nuts. A classification rate of 91.43% is achieved with Gabor wavelet features [9]. Mallaiah Suresha et al. have proposed diseased and undiseased classification of Areca nut using texture features of LBP, Haar Wavelets, GLCM, and Gabor. They achieved a 92.00% success rate [10]. T. Liu et al. have tried to achieve automatic classification by extracting the color, shape, and texture features of de-husked Areca nut [11]. Huang K.Y. used Image processing techniques and Neural Networks for quality detection and classification of areca nuts. Six geometric features, 3 color features, and defects were used for the classification process. This method of classification attained an accuracy of 90.9% [12].
Deep Learning (DL) approaches are increasingly important in machine learning because of their high degrees of abstraction and capacity to automatically identify image patterns [13]. Convolutional Neural Network (CNN) is the most frequently applied deep learning architecture for image processing among the numerous designs employed [14,15,16]. Convolution operations are used by CNN, a kind of Artificial Neural Network (ANN), in a minimum of one of its layers [14].
To the best of our information, very little research has been done on the classification of dehusked Areca nuts using CNN.

3. Data Acquisition Setup

This paper deals with the quality classification of Areca nut from the Konkan belt of India, particularly from the state of Goa. Since there are no publicly available database of the Areca nut images, A unique setup was created to create an initial database. The setup consists of a top-mounted camera with a sample table below at a distance of approximately 14cm. Surrounding the camera are radially arranged 20 white LEDs evenly illuminating the sample. A hollow cylinder coated with black paper on its inner sides is placed around the sample and camera to shield the

stray light entering the acquisition setup. The black paper prohibits light reflection from the inner walls and creates a glare on the camera lens. The power source for the setup is an AC source of 220V, 50 Hz, which is then converted to a DC constant current source coupled with a high voltage capacitor of 220 µF/ 450 V connected in parallel to reduce flicker in the illumination. In this setup, we have used a 5MP lightweight Pi camera module, which communicates with the Raspberry Pi 3 B+ board using the MIPI camera serial interface protocol. At the base of the hollow cylinder, a black cloth is placed over which an Areca nut whose image is to be acquired is kept for the reasons described above. Figure 2.1, shows the data acquisition setup designed for capturing images of Areca nuts.



(a)Top view of imaging setup (b)Components of imaging setup

Figure 2.1: Data acquisition setup
©Journal of Applied and Natural Science 15(4) ----2023

4. Methodology

4.1. **Convolutional Neural Network**

Convolutional Neural Networks (CNN) have been extensively studied in recent literature [17,18]. CNN is a class of deep learning algorithms that is incredibly efficient in classifying data by recognizing patterns in an image. A CNN is a feed-forward network consisting of basic building blocks like a convolutional layer, pooling layer, and activation layer, which are stacked with varying permutations and combinations. This varying arrangement of convolutional layer, pooling layer, and activation layer together form the feature extraction segment of a CNN [19]. Within the classification segment, the extracted features are fed into the fully connected layer and the classification layer [20]. The details of the various layers used in our custom CNN model are detailed in Figure 3.1.

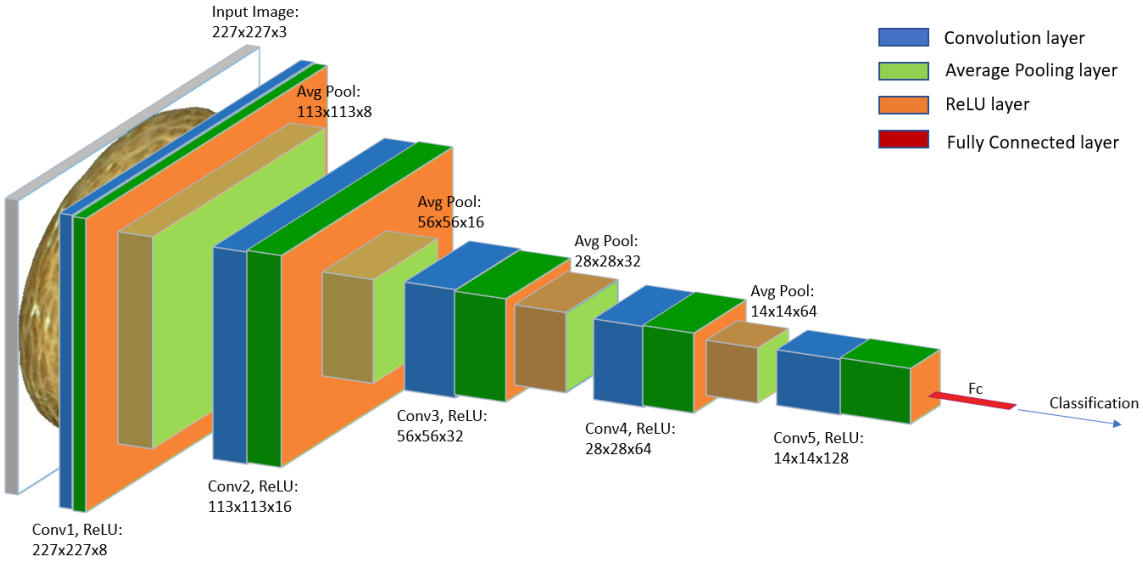
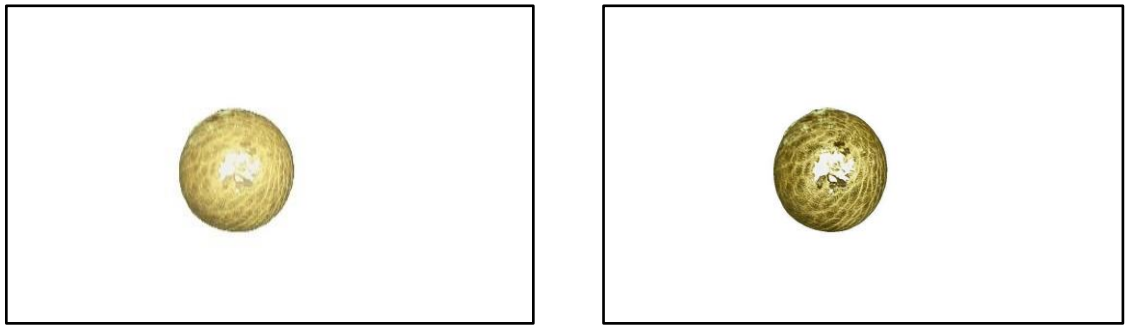


Figure 3.1: The custom CNN model

It should be emphasized that all of the photos used in this study show the Areca nut from the top. This is because, the very shape of Areca nut, which normally stabilises with its flat surface at the bottom. Also, we wanted to study the accuracy of segregation based on the top view to design an algorithm that will take reduced time for classification and thus increase the speed of segregation. In this experiment, we have performed different image processing operations as detailed below to get a better understanding of which operations will yield the best outcomes with the CNN network.

1. Areca Nut image has only been segmented and not cropped to a Region of Interest (ROI) closest to its edges. This database is labeled as NoCrop_NoContrast.
2. The Areca Nut image has been cropped to ROI closest to its edge. This database is labelled as Crop_NoContrast.
3. Areca Nut image has only been segmented and not cropped to a ROI closest to its edges and has been contrast-enhanced using Contrast -limited Adaptive Histogram Equalization (CLAHE). This database is labelled as NoCrop_Contrast.
4. Areca Nut image has been cropped to ROI closest to its edge and has been contrast-enhanced using CLAHE. This database is labelled as Crop_Contrast.

Thus, in this experiment, we are working with four distinct databases. The images of each database have been illustrated in Figure 4.1below.



(a) Uncropped and Segmented Areca nut image. (b) Uncropped and Segmented Areca nut image with contrast enhancement using CLAHE

(c) Cropped Areca Nut image. (d) Cropped Areca Nut image with contrast enhancement using CLAHE.

Figure 4.1: Images of databases of Areca nut used for the classification.

4.2 Contrast-limited Adaptive Histogram Equalization (CLAHE)

CLAHE is an algorithm used to enhance the contrast between unprocessed images. It performs histogram equalizations on non-overlapping sections of a given image and is called tiles. The surrounding tiles are then blended using bilinear interpolation to prevent introducing false borders [21].

We have also tried to find the outcome of data augmentation with each database on the final classification accuracy and the standard deviation. Therefore, with each database, we aimed to determine the classification accuracy and standard deviation with CNN, using data augmentation (with and without data augmentation).

4.3 Data Augmentation

Data augmentation is a technique in CNN, and is normally applicable, when the training samples are limited. Thus, we can produce more training examples for a network by leveraging existing images. This is accomplished by applying image processing techniques such as scaling, rotation about an axis, translation, and reflection about an axis. This results in a significantly bigger training sample size from the existing data [22].

To evaluate our CNN, we use 10-fold cross-validation. In 10-fold cross-validation, the database is split into 10 distinct folds, of which 9 folds will be used in training, and the 10th fold will be used for testing. This means that each sample used for testing is now comprised of one in the training set, and one from the training set is used for testing. Thus, the procedure is repeated 10 times, with every iteration having a new fold from one of the 10 folds for testing [23].

5. Results and Analysis

The current section presents the classification accuracy for all four databases. Specifically, in the current section, we train a CNN with 10-fold cross-validation with and without data augmentation for each database.

Table 5.1: Classification accuracy and standard deviation for uncropped images with no contrast enhancement.

	Model	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Average	Std. Deviation
No Augmentation	CNN	92.86 %	80.95 %	92.86 %	80.95 %	85.71 %	80.95 %	80.95 %	78.57 %	80.95 %	83.33 %	83.81 %	5.12%
Augmentation	CNN	71.43 %	64.29 %	57.14 %	61.90 %	71.43 %	73.81 %	54.76 %	52.38 %	66.67 %	61.90 %	63.57 %	7.36%

Table 5.1 gives the result of uncropped images with no contrast enhancement. It is observed here that no augmentation gives better results as compared to augmentation with 5.12% of standard deviation.

Table 5.2: Classification accuracy and standard deviation for cropped images with no contrast enhancement.

	Model	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Average	Std. Deviation
No Augmentation	CNN	83.33 %	66.67 %	80.95 %	80.95 %	78.57 %	78.57 %	80.95 %	69.05 %	83.33 %	83.33 %	78.57%	5.94 %
Augmentation	CNN	64.29 %	73.81 %	66.67 %	71.43 %	73.81 %	66.67 %	83.33 %	71.43 %	76.19 %	69.05 %	71.67%	5.55 %

Table 5.2 gives the result of cropped images with no contrast enhancement. The standard deviation here for no augmentation has increased compared to the augmentation for uncropped images, and the results are the opposite of those given in Table 5.1. However, both methods have a standard deviation higher than 5%.

Table 5.3: Classification accuracy and standard deviation for uncropped images with contrast enhancement.

	Model	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Average	Std. Deviation
No Augmentation	CNN	92.86 %	83.33 %	78.57 %	78.57 %	83.33 %	80.95 %	90.48 %	78.57 %	76.19 %	80.95 %	82.38%	5.41%
Augmentation	CNN	88.10 %	61.90 %	57.14 %	73.81 %	66.67 %	61.90 %	42.86 %	73.81 %	71.43 %	52.38 %	65.00%	12.75 %

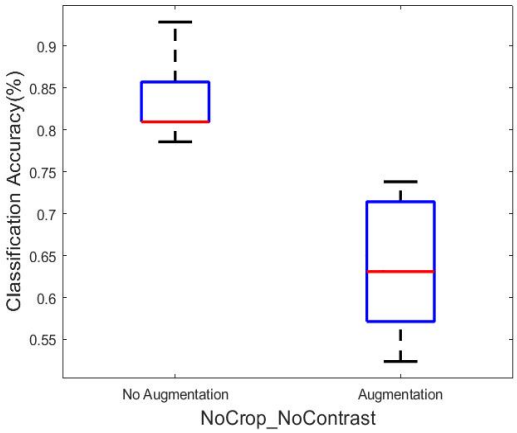
Table 5.3 shows the results of uncropped images with contrast enhancement. The results indicate that these methods do not improve significantly over the earlier two methods, whose results are listed in Table 5.1 and Table 5.2. The augmentation process gives the worst result, with a standard deviation of more than 10%.

Table 5.4: Classification accuracy and standard deviation for cropped images with contrast enhancement.

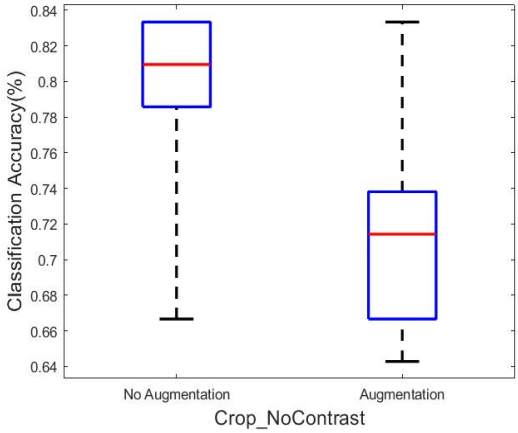
	Model	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Average	Std. Deviation
No Augmentation	CNN	76.19 %	80.95 %	73.81 %	83.33 %	83.33 %	73.81 %	80.95 %	83.33 %	73.81 %	80.95 %	**79.05 %**	**4.17 %**
Augmentation	CNN	71.43 %	73.81 %	52.38 %	59.52 %	71.43 %	59.52 %	69.05 %	66.67 %	73.81 %	64.29 %	66.19 %	7.17 %

The Table 5.4 gives the result of cropped images with contrast enhancement. Here, it may be seen that no augmentation with our custom CNN model gives a standard deviation close to 4%, which defends our claim that the top view can alone be used for the segregation process. It may be noted that the cropped image with no augmentation worked quite well, but it did not fare so well when

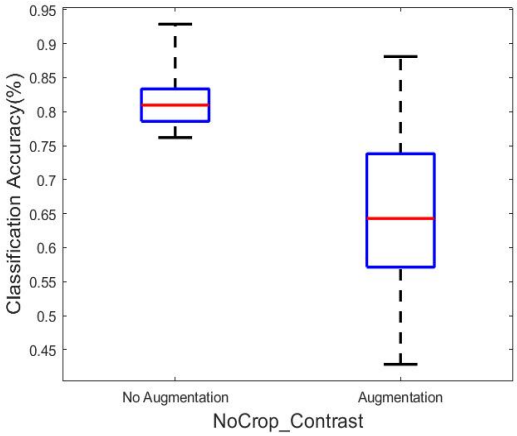
the augmentation process was utilized on the samples. All the above results and analysis have been shown in the boxplot in Figure 5.1.



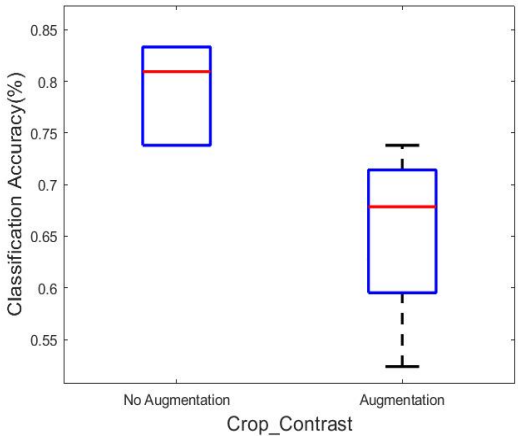
(a) Classification for uncropped images with no contrast enhancement.



(b) Classification for cropped images with no contrast enhancement.



(c) Classification for uncropped images with contrast enhancement



(d) Classification for cropped images with contrast enhancement.

Figure 5.1: Boxplot of the results and analysis of the image classification

Here, Figure 5.1, as above, shows the Boxplot of the results and analysis of the image classification. From the boxplot (a), it may be seen that, in the case of no augmentation, the accuracy is close to 80% for most trials. Whereas, in augmentation, it widely varies with the least going almost close to 50%, which is not desirable

The boxplot (b) also has accuracy for both models (for cropped and no contrast enhancement) varying widely from 66% to 84%, therefore casting doubt on the process of classification. The same is true in augmented images, wherein the accuracy varies from 64% to 84%.

As discussed in boxplot (b), boxplot (c) (for uncropped and contrast enhancement) also has a similar behavior wherein the accuracy varies widely over 75% to 95% (for no augmentation) and 45% to 90% (for augmentation). Thus, signifying that they are not consistent.
In boxplot (d), the accuracy for non-augmented images is centered around 82% with a small deviation from 74% to 84% for our custom CNN model. Thus, suggesting this method is more reliable for classification of Areca nuts. However, the same is not true in case of augmentation. The accuracy in the case of augmentation varies from 50% to 75%.

6. Conclusion

In the above article, we have carried out four diverse classification methods based on 10-fold cross-validation training of a custom CNN model using contrast enhancement and data augmentation of the images. The results indicate that using the custom CNN model, the classification method using no augmentation and contrast enhancement for cropped images, has yielded the best outcomes with a standard deviation of less than 5%. The standard deviation of less than 5% is a significant number for agriculturalists for the segregation of Areca nuts, considering we have used only the top view. Single-image segregation can greatly increase the speed of segregation; thus, the payments to the farmers can be given on the spot, and the loss of revenue due to human fatigue can be reduced. The above experiments suggest that the algorithm can be implemented and a machine can be manufactured to segregate Areca nuts automatically. This article provides concept validation for the manufacturing of automated Areca segregation units.

Acknowledgement

Authors acknowledge the help extended by the skilled segregators for classification and officials of Goa Bagayatdar for providing many samples of Areca nut used for this work.

Statements and Declarations

Data Availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Funding and/or Conflicts of Interests/Competing Interests

Funding: The authors declare that they have not availed any funding from any agency for the above-carried work.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

[1] Areca Nut - an overview | ScienceDirect Topics. (2024). https://www.sciencedirect.com/topics/neuroscience/areca-nut

[2] Origin | Arecanut. (2024). https://arecanut.org/arecanut-1/origin/

[3] Areca nut, *Wikipedia*. (2024). https://en.wikipedia.org/w/index.php?title=Areca_nut&oldid=1247804910

[4] A. Kumar *et al.*, Assessment of areca nut use, practice and dependency among people in Guwahati, Assam: a cross-sectional study, *ecancer*, vol. 15, (2021), doi: 10.3332/ecancer.2021.1198.

[5] M. S. Amudhan, Begum V Hazeena, and H. K. B. , A REVIEW ON PHYTOCHEMICAL AND PHARMACOLOGICAL POTENTIAL OF ARECA CATECHU L. SEED, *IJPSR*, vol. 3, no. 11, pp. 4151–4157,(2012), https://www.researchgate.net/publication/264710991_A_review_on_phytochemical_and_pharmacological_potential_of_Areca_catechu_L_Seed

[6] APAC: areca nut production by country 2022, Statista. (2024). https://www.statista.com/statistics/657902/asia-pacific-areca-nut-production-by-country/

[7] V. Raghavan and H. K. Baruah, Arecanut: India's popular masticatory —history, chemistry and utilization, *Econ Bot*, vol. 12, no. 4, 315–345, Oct. 1958, doi: 10.1007/BF02860022.

[8] Goa Bagayatdar Bazar – One-stop-shop for all. (2024). https://goabagayatdar.com/

[9] S. Siddesha, S. K. Niranjan, and V. N. Manjunath Aradhya, Texture based classification of arecanut, in *2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, (2015), 688–692. doi: 10.1109/ICATCCT.2015.7456971.

[10] S. Mallaiah, Ajit Danti, and N. S. K , Classification of Diseased Arecanut based on Texture Features , *International Journal of Computer Applications*, vol. NCRAIT 3, 1–6, (2014), [Online]. Available: /proceedings/ncrait/number3/15152-1419/

[11] T. Liu, J. Xie, Y. He, M. Xu, and C. Qin, An automatic classification method for betel nut based on computer vision, *International Conference on Robotics and Biomimetics (ROBIO)*, 1264–1267, (2009), doi: 10.1109/ROBIO.2009.5420823.

[12] K.-Y. Huang, Detection and classification of areca nuts with machine vision, *Computers & Mathematics with Applications*, vol. 64, no. 5, 739–746, (2012), doi: 10.1016/j.camwa.2011.11.041.

[13] J. Naranjo-Torres, M. Mora, R. Hernández-García, R. J. Barrientos, C. Fredes, and A. Valenzuela, A Review of Convolutional Neural Network Applied to Fruit Image Processing, *Applied Sciences*, vol. 10, no. 10, 3443, (2020), doi: 10.3390/app10103443.

[14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. in Adaptive computation and machine learning. Cambridge, Massachusetts: The MIT Press, (2016).

[15] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature*, vol. 521, no. 7553, 436–444, (2015), doi: 10.1038/nature14539.

[16] M. D. Zeiler and R. Fergus, Visualizing and Understanding Convolutional Networks, in *Computer Vision – ECCV 2014*, vol. 8689, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., Cham: Springer International Publishing, (2014), 818–833. doi: 10.1007/978-3-319-10590-1_53.

[17] W. Jia, Y. Tian, R. Luo, Z. Zhang, J. Lian, and Y. Zheng, Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot, *Computers and Electronics in Agriculture*, vol. 172, 105380, (2020), doi: 10.1016/j.compag.2020.105380.

[18] X. Mai, H. Zhang, X. Jia, and M. Q.-H. Meng, Faster R-CNN With Classifier Fusion for Automatic Detection of Small Fruits, *IEEE Trans. Automat. Sci. Eng.*, 1–15, (2020), doi: 10.1109/TASE.2020.2964289.

[19] M. Khoshdeli, R. Cong, and B. Parvin, Detection of nuclei in H&E stained sections using convolutional neural networks, in *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Orland, FL, USA: IEEE, (2017), 105–108. doi: 10.1109/BHI.2017.7897216.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM*, vol. 60, no. 6, 84–90, (2017), doi: 10.1145/3065386.

[21] S. Aboshosha, O. Zahran, M. I. Dessouky, and F. E. Abd El-Samie, Resolution and quality enhancement of images using interpolation and contrast limited adaptive histogram equalization, *Multimed Tools Appl*, vol. 78, no. 13, 18751–18786, (2019), doi: 10.1007/s11042-018-7022-1.

[22] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, Random Erasing Data Augmentation, *AAAI*, vol. 34, no. 07, 13001–13008, (2020), doi: 10.1609/aaai.v34i07.7000.

[23] T.-T. Wong and P.-Y. Yeh, Reliable Accuracy Estimates from *k* -Fold Cross Validation, *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 8, 1586–1594, (2020), doi: 10.1109/TKDE.2019.2912815.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-22] Sentiment Analysis for Citizen Feedback in Smart Cities with XLNet-BiLSTM: Delhi Metro as a Case Study

Vaibhav Shukla[1,*,†], Dibyalochan Kuanr[2,†], Shallu Juneja[3,†] and Ajit K. Sharma[4,†]

[1]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India
[2]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India
[3]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India
[4]Department of Computer Science and Engineering ,Maharaja Agrasen Institute of Technology, Delhi, India

Abstract

In recent years, smart cities have increasingly recognized the importance of citizen input in enhancing public services and optimizing urban infrastructure. As urban populations grow and services become more complex, understanding resident sentiments and opinions is crucial for effective governance. Sentiment analysis, a technique rooted in natural language processing (NLP), serves as a powerful tool for gauging public opinion on urban services, particularly public transportation. This paper presents a sentiment analysis framework using an advanced XLNet-Bidirectional Long Short-Term Memory (BiLSTM) model, developed with a custom dataset of citizen reviews related to the Delhi Metro, a key element of India's public transportation. The dataset was meticulously scraped from various platforms and manually labeled for accuracy. Initially, the model was trained on the IMDb dataset, achieving an impressive accuracy of 93.1%. It was then evaluated on the Delhi Metro dataset, yielding an accuracy of 1.00. However, this high accuracy may indicate overfitting due to the small dataset size, suggesting the findings are exploratory. This study highlights how sentiment analysis can improve decision-making and enhance public transportation services. By analyzing feedback on the Delhi Metro, city planners can identify areas for improvement and address citizen concerns. In conclusion, the paper underscores the potential of advanced sentiment analysis techniques in understanding public opinion and calls for further research with larger, more diverse datasets and refined models to assess citizen sentiment in smart cities comprehensively.

Keywords

sentiment analysis, smart cities, NLP, XLNet, BiLSTM, Delhi Metro, public transportation

1. Introduction

Urbanization has led to the rapid development of **smart cities**, which depend on advanced technologies and citizen engagement to enhance public services and infrastructure. As urban populations increase, the need for efficient and responsive public transportation systems becomes more critical. Citizen feedback plays a pivotal role in improving these services by providing insights into user experiences, satisfaction levels, and areas requiring enhancement.

3rd International Symposium on Smart Cities Challenges, Technologies and Trends (SCCTT-2024)
*Corresponding author.
†These authors contributed equally.
vaibhavs00788@gmail.com (V. Shukla); maildibyalochan@gmail.com (D. Kuanr); shallujuneja@mait.ac.in (S. Juneja); ajitsharmas567@gmail.com (A. K. Sharma)

Sentiment analysis is a powerful tool to gauge public opinion, enabling city planners and poli- cymakers to make data-driven decisions that address the needs and concerns of urban residents. By leveraging natural language processing (NLP) techniques, sentiment analysis extracts subjective in- formation from textual data, transforming unstructured feedback into quantifiable insights. Several algorithms have been employed for sentiment classification, each with its respective strengths and weaknesses. Traditional machine learning approaches, such as **Support Vector Machines (SVM)** and **Naive Bayes**, are frequently used due to their simplicity and effectiveness. For instance, Ajmera [1] employed SVM for sentiment analysis of IMDb movie reviews, achieving an accuracy of **82.2%**, showcasing the model's capability in handling real-world sentiment classification tasks.

As the field advances, **deep learning techniques** have gained prominence, with models such as **Long Short-Term Memory (LSTM)** networks demonstrating superior performance in various sentiment analysis applications. Abdirahman et al. [2] highlighted the effectiveness of LSTM, achieving an accuracy of **88.58%** in sentiment classification for Somali language texts. These advancements demonstrate that deep learning architectures can significantly improve sentiment analysis models by learning hierarchical representations of data.

Hybrid models that combine the strengths of multiple approaches have also emerged, pushing the boundaries of sentiment analysis further. Garg and Sharma [3] explored text preprocessing techniques alongside machine learning and deep learning algorithms, emphasizing the importance of feature extraction for improving classification accuracy. Their study demonstrated that integrating various methodologies could enhance performance, particularly in diverse, multilingual datasets.

The introduction of **transformer-based models**, such as **BERT (Bidirectional Encoder Repre- sentations from Transformers)**, has revolutionized sentiment analysis. Sousa et al. [4] achieved an accuracy of **82.5%** in stock market sentiment analysis using BERT, demonstrating its superior ability to understand context and semantics in language compared to previous models. However, despite these advancements, there remains a need for models that can effectively capture nuanced sentiments expressed in citizen feedback, particularly in smart city contexts.

This study proposes a novel sentiment analysis framework utilizing the **XLNet-BiLSTM model**, focusing on citizen reviews of the **Delhi Metro**. XLNet, an improvement over traditional transformer architectures, enhances contextual understanding by using a permutation-based training approach. By integrating this with a BiLSTM architecture, the proposed framework captures both contextual information and sequential dependencies in textual data.

To assess the effectiveness of this approach, we created a custom dataset comprising citizen reviews of the Delhi Metro, which were manually scraped and labeled. Initial results from training the model on the IMDb dataset indicated a high accuracy of **93.1%**, demonstrating the model's effectiveness in sentiment classification. Additionally, the model achieved perfect accuracy (**1.00**) on the custom dataset, underscoring the exploratory nature of this research as a proof of concept rather than a definitive evaluation.

This paper contributes to the evolving field of sentiment analysis in smart cities by presenting an innovative framework leveraging state-of-the-art techniques. By focusing on the Delhi Metro case study, we provide insights into citizen sentiment and highlight the potential of sentiment analysis for enhancing urban transportation systems. Our findings not only advance the theoretical understanding of sentiment analysis but also offer practical recommendations for improving public services through effective citizen engagement.

2. Literature Survey

Sentiment analysis has become a pivotal area of research, driven by the exponential growth of social media and online platforms filled with user-generated content. **Bonta et al.** [5] conducted a com- prehensive study on lexicon-based approaches, utilizing tools like **NLTK**, **TextBlob**, **VADER**, and **SentiWordNet**. Their study found

that VADER achieved a classification accuracy of **78.46%**, a recall of **85.0%**, and an F1 score of **81.60%**, demonstrating the effectiveness of lexicon-based methods, especially in classifying short texts prevalent in social media.

Grana [6] explored several machine learning models, including **Naïve Bayes**, **SVM**, and **RNN**, reporting that their system achieved an F1 score of **0.62** and a recall of **0.55**. This variability in performance highlights the importance of algorithm selection to improve sentiment classification outcomes. Similarly, **Drus and Khalid** [7] conducted a systematic review of sentiment analysis techniques applied to social media, advocating for a hybrid approach that combines lexicon-based methods and machine learning to improve sentiment classification, particularly in handling noisy data from social platforms.

Yogi et al. [8] performed a comparative analysis of classification algorithms, including **K-Nearest Neighbor (KNN)**, **Multinomial Naive Bayes (MNB)**, and **SVM**. Their study concluded that SVM outperformed the others with an accuracy of **89.46%**, further emphasizing the importance of algorithm selection based on dataset characteristics. In a similar context, **Al-mashhadani et al.** [9] analyzed sentiment across different social media platforms using hybrid feature extraction techniques, reporting that optimized feature sets can achieve accuracy as high as **90%**.

The impact of text preprocessing techniques on sentiment analysis was examined by **Garg and Sharma** [3]. Their study focused on methods like tokenization and stop word removal, along with machine learning and deep learning algorithms, achieving an F1 score of **47%** with SVM and **83%** with LSTM. Their findings underscore the crucial role of preprocessing in enhancing model performance, particularly in multilingual contexts. **Han et al.** [10] also demonstrated the effectiveness of SVM combined with probabilistic latent semantic analysis for Twitter sentiment analysis, achieving an accuracy of **87.20%** and a recall rate of **88.30%**.

On the deep learning front, **Srinivas et al.** [11] explored the performance of LSTM models in sentiment analysis on Twitter datasets, achieving a training accuracy of **87.4%**, showcasing the growing trend of using deep learning techniques for sentiment analysis. Additionally, **Abbas et al.** [12] applied Multinomial Naive Bayes on movie reviews, attaining an accuracy of **86%** and an F1 score of **0.85**, reinforcing the model's efficiency in text classification tasks.

A hybrid approach combining SVM and lexicon-based methods, as explored by **Muhammadi et al.** [13], yielded promising results in Twitter sentiment analysis, with a precision of **78.68%** and an F1 score of **79.60%**. Similarly, **Abdirahman et al.** [2] compared traditional machine learning with deep learning architectures for Somali sentiment analysis, with LSTM outperforming other models with an accuracy of **88.58%**.

Further advancements in sentiment analysis methodologies were showcased by **Mulyo and Widyan- toro** [14], who employed a **convolutional neural network (CNN)** for aspect-based sentiment analysis, achieving an F1 score of **0.71**, demonstrating CNN's capacity to handle context-specific sentiment tasks. Similarly, **Sultana et al.** [15] analyzed product reviews using multiple algorithms, including Naive Bayes, which achieved an accuracy of **89.85%**, highlighting the broad applicability of sentiment analysis techniques.

Mahadevaswamy and Swathi [16] focused on **Bidirectional LSTM networks**, achieving an accu- racy of **90.14%** on Amazon product reviews. **Muhammada et al.** [17] applied **Word2vec embeddings** with LSTM for analyzing hotel reviews in Indonesia, achieving an accuracy of **85.96%**, showing the efficacy of advanced word embeddings in sentiment analysis.

Lastly, **Imran et al.** [18] applied deep learning techniques to analyze COVID-19-related tweets, achieving a sentiment classification accuracy of **81.83%**. This adaptability to different contexts illus- trates the potential of deep learning models in sentiment analysis. Overall, the literature presents a diverse range of approaches, from traditional machine learning techniques to advanced deep learning models, with a growing trend towards hybrid methods that integrate multiple techniques to improve classification accuracy. The ongoing evolution of

methodologies underscores the need for continued research to enhance sentiment analysis performance, especially in domains like smart cities and public transportation systems.

3. Proposed Model

3.1. Overview

This research introduces a novel sentiment analysis framework based on an XLNet-BiLSTM model, which integrates the advanced capabilities of the XLNet architecture with the sequential processing strengths of a Bidirectional Long Short-Term Memory (BiLSTM) network. The primary objective of this model is to enhance the understanding of sentiments expressed in complex and opinionated texts, such as citizen reviews and social media posts.

XLNet is a state-of-the-art transformer-based model that addresses limitations in traditional trans- former architectures. Unlike conventional models relying on fixed context windows, XLNet employs a permutation-based training method, allowing it to capture dependencies among all words in a se- quence more effectively. This capability is particularly beneficial for sentiment analysis as it enables the generation of contextualized word embeddings that reflect the nuanced meanings of words based on their surrounding context. By considering multiple permutations of word sequences during training, XLNet learns richer representations of language, crucial for understanding subtleties in sentiment.

Once contextualized embeddings are generated by XLNet, they are fed into a Bidirectional Long Short-Term Memory (BiLSTM) network. The BiLSTM architecture processes sequential data in both forward and backward directions, enabling it to capture information from both past and future contexts. This bidirectional processing is advantageous for sentiment analysis, where the meaning of a word can be influenced by the words that precede and follow it. By leveraging this dual context, the BiLSTM enhances the model's ability to discern complex sentiment nuances and relationships within the text. The integration of XLNet with BiLSTM is crucial in overcoming challenges commonly faced in sentiment analysis, such as ambiguity and contextual variability. For instance, in opinionated texts, the same word may convey different sentiments depending on its context. The XLNet-BiLSTM model's architecture effectively handles such complexities by using contextualized embeddings to capture dynamic word meanings, while the BiLSTM interprets these embeddings sequentially.

The proposed model is trained using a custom dataset consisting of citizen reviews, providing a rich source of opinionated content. The training process involves optimizing the model to minimize the loss function and maximize the accuracy of sentiment classification. By focusing on real-world data, the model is trained not only to recognize generic sentiment patterns but also to understand specific sentiments expressed by citizens regarding public services and transportation systems.

In summary, the XLNet-BiLSTM model presents a sophisticated approach to sentiment analysis, combining advanced contextualized embeddings with robust sequential processing capabilities. This innovative architecture aims to provide deeper insights into sentiments expressed in complex texts, facilitating more informed decision-making by city planners and policymakers in smart cities.

3.2. Training on IMDb Dataset

The training of the XLNet-BiLSTM model began with the IMDb dataset, a well-established benchmark for sentiment analysis. The dataset includes 50,000 movie reviews with an equal distribution of positive and negative sentiments, providing a comprehensive and diverse training set.

The IMDb dataset's wide variety of reviews allows the model to encounter multiple contexts and sentiment expressions. By training on this data, the model learns to identify subtle nuances in sentiment, such as sarcasm, humor, and emotional complexity, commonly present in human-written texts.

During training, the XLNet-BiLSTM model utilized the rich contextual embeddings generated by XLNet, which effectively capture intricate relationships between words and phrases within the reviews. The training process involved optimizing the model to minimize the loss function and adjust its parameters over multiple epochs, improving its performance progressively. Techniques such as dropout regularization and gradient clipping were employed to prevent overfitting and enhance the model's generalizability.

Upon completing the training phase, the model achieved an accuracy of **93.1%** on the IMDb dataset.

This high accuracy highlights the model's ability to classify sentiment effectively across various contexts and expressions. The successful performance on the IMDb dataset demonstrates that the XLNet-BiLSTM model generalizes well, making it a strong candidate for real-world sentiment analysis tasks, especially those involving more nuanced opinionated texts.

The insights gained from training on the IMDb dataset validate the model architecture's effectiveness and lay the foundation for further evaluation on a custom dataset of citizen reviews. By establishing strong baseline performance in a controlled environment, the model's potential to analyze and under- stand citizen sentiment in practical applications, such as public transportation feedback, is enhanced.

.

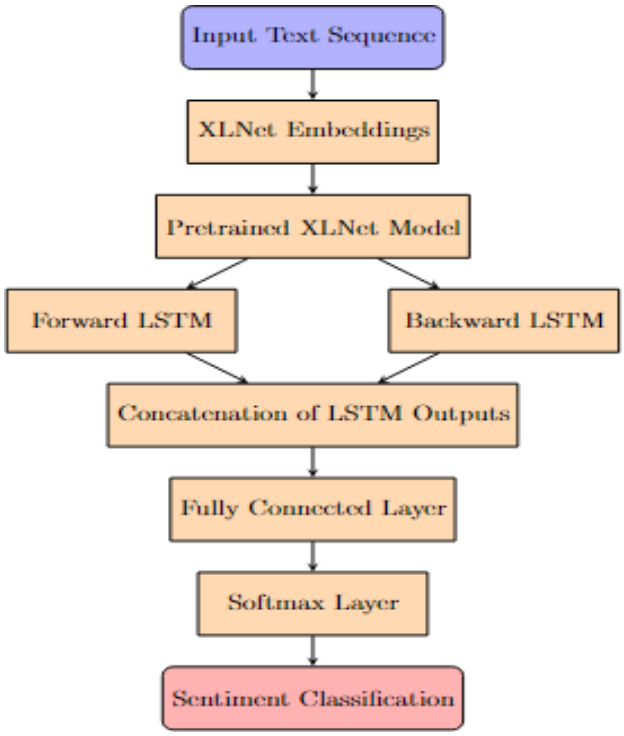


Figure 1: Overview of the XLNet-BiLSTM model architecture. This figure illustrates the integration of XLNet and BiLSTM for processing contextual embeddings and sequential data to perform sentiment analysis.

4. Delhi Metro Sentiment Dataset

4.1. Dataset Description

The custom **Delhi Metro dataset** consists of approximately 50 rows of citizen reviews, sourced from a YouTube video discussing user experiences with the Delhi Metro. This dataset encapsulates a variety of opinions and sentiments reflecting individuals' interactions with the transit system. Each entry in the dataset includes two essential columns:

- **Cleaned_Comment**: This column contains preprocessed user comments detailing their experiences with the Delhi Metro. Preprocessing was performed to standardize the text, making it suitable for sentiment analysis.
- **Sentiment:** This column represents the manually labeled sentiment of each comment, categorized as either positive or negative. Careful manual classification was applied to ensure accuracy in capturing the sentiment conveyed by the user reviews.

Despite the limited size of the dataset, rigorous manual labeling and preprocessing have been conducted to maximize data quality for both training and evaluation purposes.

4.2. Data Preprocessing

To optimize the performance of the XLNet-BiLSTM model on the Delhi Metro dataset, a detailed series of preprocessing steps was systematically applied to transform the raw text into a suitable format for analysis. These steps not only ensure the removal of irrelevant information but also help in aligning the preprocessing of the Delhi Metro dataset with the IMDb dataset, thus maintaining consistency in data handling across different datasets. By ensuring the integrity and quality of the input data, the preprocessing phase plays a pivotal role in improving the overall model performance. The following preprocessing operations were performed:

- **Lowercasing**: All text data was converted to lowercase to maintain consistency and eliminate discrepancies caused by case sensitivity. This step ensures uniform treatment of words regardless of capitalization. Words like "Metro" and "metro," for instance, are treated the same, helping the model focus on semantic meaning rather than variations in text presentation.
- **Removing URLs**: URLs present in the comments were removed, as they typically do not contribute meaningful sentiment information. These hyperlinks could distract the model and introduce noise into the dataset. By removing them, the model's focus is redirected to more sentiment-relevant features of the text.
- **Removing Special Characters**: Along with URLs, special characters (e.g., punctuation marks, hashtags) were also removed, as they often do not carry meaningful sentiment. This step ensures that the remaining text is clean and more interpretable by the model, reducing noise.
- **Tokenization**: The cleaned comments were then tokenized using the **XLNetTokenizer**, which breaks down the text into tokens. Tokenization is crucial for preparing the text for input into the XLNet model, allowing the model to process each word and sentence structure individually. Proper tokenization helps the model capture linguistic nuances and sentiment patterns more effectively.
- **Removing Stop Words**: Common stop words such as "the," "is," and "and" were removed as they do not contribute significantly to the overall sentiment. This ensures the model focuses on more

sentiment-rich parts of the text, improving the relevance of the processed data.

These preprocessing steps are crucial in reducing noise and standardizing the dataset, helping the model accurately capture the nuances of sentiment expressed in the input data. By applying a consistent preprocessing strategy across both the IMDb and Delhi Metro datasets, the model can better generalize its learning and perform more effectively on unseen data.

4.3. Model Evaluation on Delhi Metro Dataset

Upon evaluating the **XLNet-BiLSTM** model on the custom Delhi Metro dataset, the model achieved an outstanding accuracy of **1.00**. This perfect accuracy suggests that the model classified all sentiments correctly. However, it is important to approach this result with caution. The small size and limited diversity of the dataset may have significantly contributed to this outcome. With only 50 reviews, the model may have learned specific patterns that do not generalize well to broader datasets or varied sentiments. Thus, while the accuracy reflects the model's performance on this particular dataset, it may not necessarily indicate its effectiveness in real-world scenarios.

4.4. Future Work and Limitations

The primary limitation of this evaluation lies in the **small size of the dataset**, which raises concerns about overfitting. Overfitting occurs when a model performs exceedingly well on the training data but struggles to generalize to new, unseen data. In addition, the dataset shows a significant imbalance, with a much higher proportion of positive reviews compared to negative ones, which could skew the model's predictive capabilities. To develop a more robust and generalizable model, future research should consider the following approaches:

- **Collecting Larger Datasets**: A larger and more diverse dataset should be gathered from various sources, including user-generated reviews from social media platforms, online forums, and public discussion boards concerning the Delhi Metro and urban transportation experiences.
- **Enhancing Dataset Diversity**: Future datasets should include reviews from different demographic groups, geographic regions, and user experiences to provide a richer dataset. This will allow the model to learn more generalized sentiment patterns, improving its predictive capabilities.
- **Addressing Dataset Imbalance**: Given that negative reviews are significantly less represented than positive ones, future work should explore techniques to address this imbalance. Strategies such as oversampling the minority class (negative reviews), undersampling the majority class (positive reviews), or employing advanced methods like Synthetic Minority Over-sampling Technique (SMOTE) can be implemented to ensure the model does not become biased toward the majority class.
- **Implementing Cross-validation**: Future evaluations should employ cross-validation techniques to assess the model's robustness and ability to generalize across different data subsets. Cross-validation will help detect overfitting and ensure the model performs well on a variety of datasets.

By addressing these limitations and expanding the dataset, future research can enhance the effectiveness of sentiment analysis models applied to urban transit systems, providing better insights and enabling improvements in public transportation services.

5. Results and Visualizations

5.1. IMDb Dataset Results

The **XLNet-BiLSTM** model achieved the following performance metrics on the **IMDb dataset**:

- **Accuracy**: 93.1%
- **Precision**: 0.93
- **Recall**: 0.93
- **F1-score**: 0.93

5.2. Delhi Metro Dataset Results

The **XLNet-BiLSTM** model achieved the following performance metrics on the **Delhi Metro dataset**:

- **Accuracy**: 100%
- **Precision**: 1.00
- **Recall**: 1.00
- **F1-score**: 1.00

5.3. Visualizations

The following visualizations provide additional insights into the model's performance:

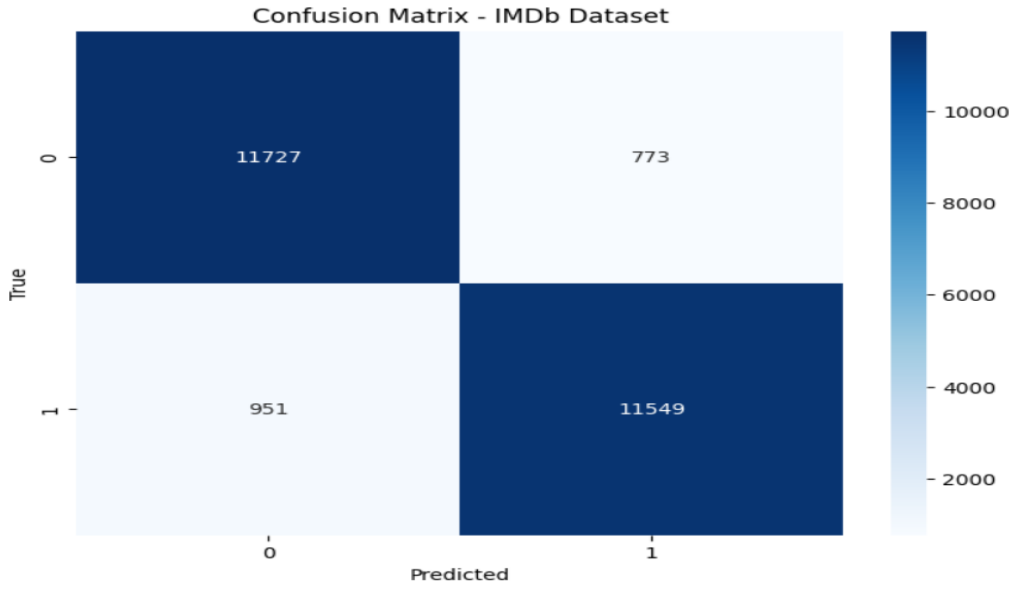


Figure 2: Confusion Matrix for IMDb Dataset. This matrix shows how well the model classified positive and negative reviews in the IMDb dataset, with correct classifications along the diagonal.

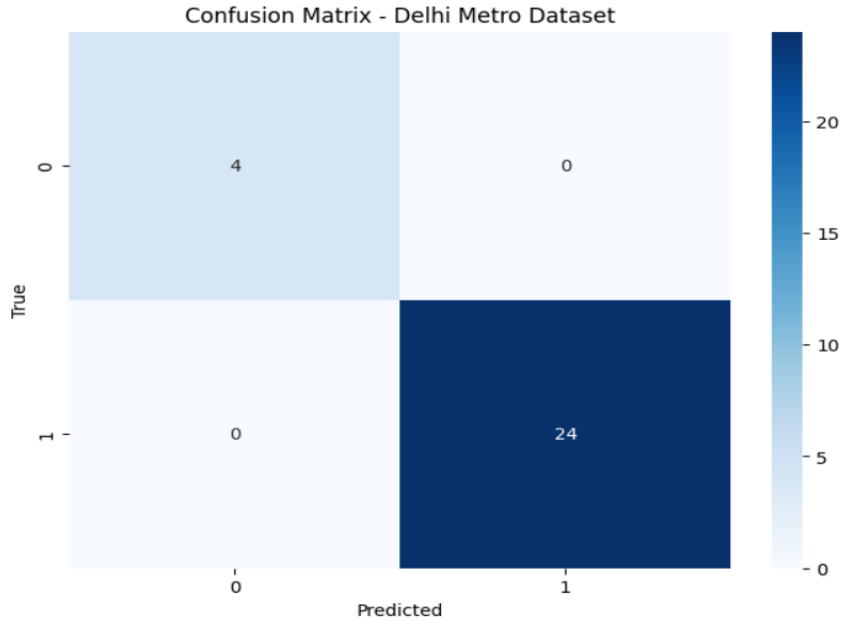


Figure 3: Confusion Matrix for Delhi Metro Dataset. This matrix shows perfect classification of positive and negative reviews, with no misclassifications.

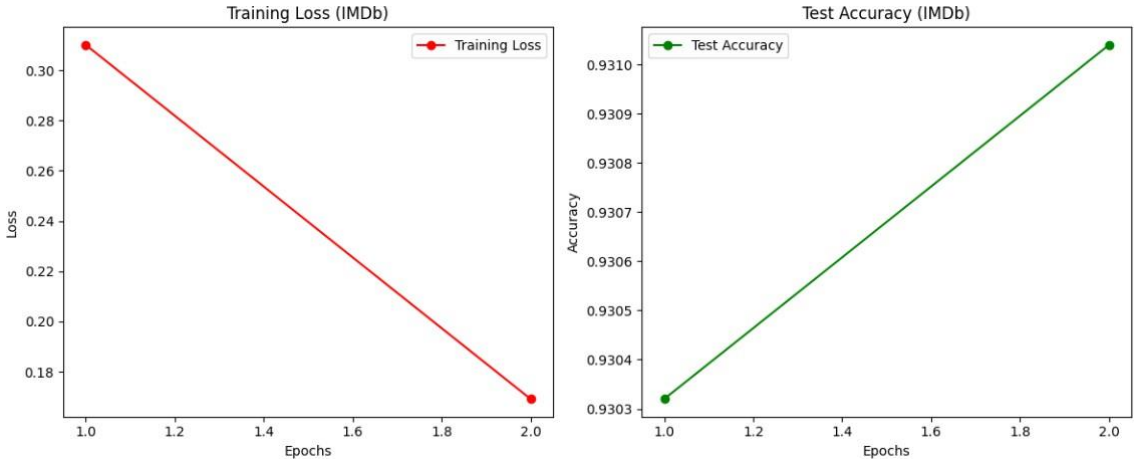


Figure 4: Accuracy and Loss Curves During Training on IMDb Dataset. The curves show a steady increase in accuracy and a decrease in loss over several epochs, indicating good model training progress.

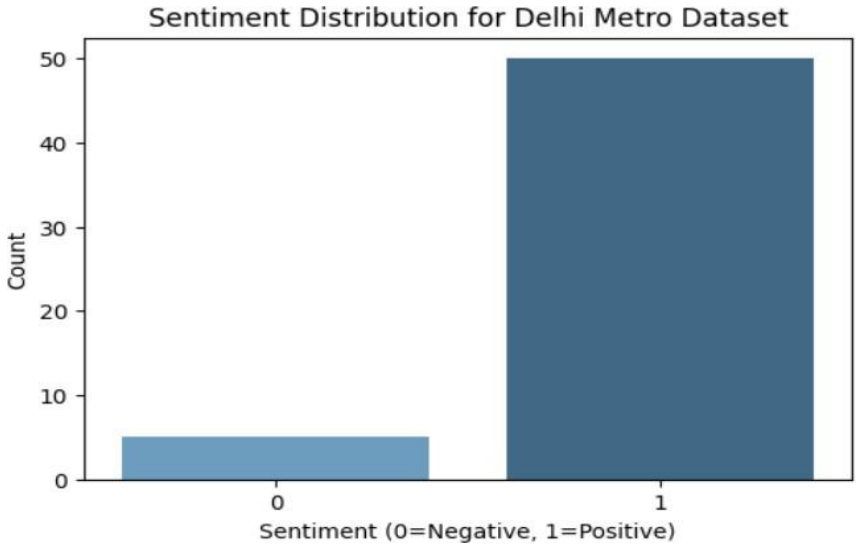


Figure 5: Sentiment Distribution for Delhi Metro Dataset. This bar chart displays the proportion of positive and negative reviews in the dataset, with positive reviews slightly outnumbering negative ones.

6. Discussion

6.1. Model Performance

The **XLNet-BiLSTM** model exhibited remarkable performance on the **IMDb dataset**, achieving an accuracy of **93.1%**, which underscores its effectiveness in processing and classifying sentiment in complex textual data. This performance aligns with existing literature on sentiment analysis models, demonstrating that advanced architectures like XLNet, combined with BiLSTM, can significantly enhance sentiment classification accuracy compared to traditional methods. The XLNet's ability to generate contextualized embeddings, coupled with BiLSTM's capability to understand sequential data, allowed the model to capture nuanced sentiments expressed in movie reviews.

In contrast, the model's testing on the **Delhi Metro dataset** resulted in an extraordinary accuracy of **1.00**, indicating perfect classification of sentiments within this limited dataset. While such results are highly encouraging, they also raise concerns regarding potential overfitting. The small size of the dataset—comprising only 50 reviews—limits the diversity of the input data, which can lead the model to memorize specific examples rather than generalizing from them. This phenomenon is a common pitfall in machine learning, particularly in NLP tasks where context and variability are crucial. To achieve more robust and generalizable results, it is imperative to validate the model against larger datasets that capture a broader spectrum of sentiments and opinions. Future studies should focus on augmenting the Delhi Metro dataset with additional reviews and possibly integrating data from other sources, such as social media, to enhance the model's training process.

6.2. Uses for Smart Cities

Sentiment analysis presents a powerful tool for understanding public sentiment and enhancing services within the framework of **smart cities**. By employing sentiment analysis techniques like the one demon- strated with the

XLNet-BiLSTM model, city planners and decision-makers can gain invaluable insights into the feelings and opinions of the public regarding various urban services, including transportation systems like the Delhi Metro. This information can be used to assess public satisfaction and identify specific areas that require improvement, such as service efficiency, safety, and accessibility.

For instance, analyzing sentiments from user-generated comments can reveal patterns in public opinion, highlighting both positive feedback and areas of concern. If the sentiment analysis indicates a consistent negative sentiment towards certain aspects of the transit system, decision-makers can prioritize these areas for enhancement. Furthermore, sentiment analysis can facilitate real-time moni- toring of public reactions to new policies or changes in service, allowing for quicker responses to public concerns.

In the context of smart cities, where the integration of technology and data analysis plays a pivotal role, sentiment analysis can drive data-informed decision-making. By continuously gathering and analyzing feedback from citizens, urban planners can create more responsive and adaptable transit systems that not only meet current needs but also anticipate future demands. Ultimately, the application of sentiment analysis in smart cities can lead to improved public services, enhanced citizen engagement, and a higher overall quality of urban life.

7. Conclusion

This study presents a novel sentiment analysis model that effectively combines **XLNet** and **BiLSTM** architectures to analyze citizen feedback on urban services, specifically focusing on the **Delhi Metro**. The model demonstrated exceptional performance on the widely recognized **IMDb dataset**, achieving an impressive accuracy of **93.1%**. This high level of accuracy indicates the model's capability to understand and classify sentiments in complex, opinionated texts, reinforcing the effectiveness of integrating advanced natural language processing techniques.

In addition to its success with the IMDb dataset, the model was further evaluated using a custom dataset comprising citizen reviews related to the Delhi Metro. The model performed flawlessly, at- taining a perfect accuracy of **1.00**. While such results are undoubtedly encouraging, it is crucial to approach these findings with caution. The limited size of the Delhi Metro dataset—consisting of only 50 reviews—raises concerns regarding the model's potential overfitting to this small and specific set of data. Overfitting occurs when a model learns to recognize patterns in the training data but fails to generalize these findings to new, unseen data. As a result, while the model's perfect accuracy on this dataset is promising, it should not be construed as definitive proof of its robustness in real-world applications.

To address these concerns, future research should prioritize the collection and analysis of larger and more diverse datasets. By expanding the range of inputs, researchers can better assess the model's generalizability and reliability across different contexts and settings. Gathering feedback from various sources, such as social media platforms, public forums, and other transportation systems, will provide a more comprehensive understanding of public sentiment and allow for a more robust evaluation of the model's performance.

Moreover, exploring the implications of sentiment analysis for smart city initiatives is a promising avenue for further investigation. The insights gleaned from citizen feedback can significantly inform urban planning and decision-making processes, leading to improved public services and enhanced citizen engagement. By continuously monitoring and analyzing public sentiment, city planners can make data-driven decisions that address the needs and concerns of their constituents, ultimately fostering a more responsive and adaptive urban environment.

In conclusion, this research not only demonstrates the potential of the XLNet-BiLSTM model in sentiment analysis but also underscores the importance of validating findings with broader datasets to ensure the model's effectiveness in real-world applications. Future studies will play a critical role in advancing our understanding

of sentiment analysis within the context of smart cities, paving the way for innovative solutions that enhance urban living and promote citizen satisfaction.

Acknowledgments

The authors would like to express their sincere gratitude to Dr. Shallu Juneja for her invaluable guidance and support throughout the research process. This work was conducted as part of a minor project for the 7th semester, as outlined in the syllabus of the Department of Computer Science and Engineering at Maharaja Agrasen Institute of Technology. The authors also acknowledge the resources and facilities provided by the department, which significantly contributed to the completion of this project.

References

[1] A. Ajmera, Sentiment analysis of imdb movie reviews, International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN 10 (2022) 2321–9653. doi:10.22214/ijraset.2022.47795.

[2] A. A. Abdirahman, A. O. Hashi, M. A. Elmi, O. E. R. Rodriguez, Comparative analysis of machine learning and deep learning models for sentiment analysis in somali, SSRG International Journal of Electrical and Electronics Engineering 10 (2023). doi:10.14445/23488379/IJEEE-V10I7P104.

[3] N. Garg, K. Sharma, Text pre-processing of multilingual for sentiment analysis based on social network data., International Journal of Electrical & Computer Engineering (2088-8708) 12 (2022). doi:10.11591/ijece.v12i1.pp776-784.

[4] M. G. Sousa, K. Sakiyama, L. de Souza Rodrigues, P. H. Moraes, E. R. Fernandes, E. T. Matsubara, Bert for stock market sentiment analysis, in: 2019 IEEE 31st international conference on tools with artificial intelligence (ICTAI), IEEE, 2019, pp. 1597–1601. doi:10.1109/ICTAI.2019.00231.

[5] V. Bonta, N. Kumaresh, N. Janardhan, A comprehensive study on lexicon based approaches for sentiment analysis, Asian Journal of Computer Science and Technology 8 (2019) 1–6. doi:10.51983/ajcst-2019.8.S2.2037.

[6] P. A. Grana, Sentiment analysis of text using machine learning models, International Research Journal of Modernization in Engineering Technology and Science 04 (2022).

[7] Z. Drus, H. Khalid, Sentiment analysis in social media and its application: Systematic literature review, Procedia Computer Science 161 (2019) 707–714. doi:10.1016/j.procs.2019.11.174.

[8] T. N. Yogi, N. Paudel, Comparative analysis of machine learning based classification algorithms for sentiment analysis, International Journal of Innovative Science, Engineering & Technology 7 (2020) 1–9.

[9] M. I. Al-mashhadani, K. M. Hussein, E. T. Khudir, et al., Sentiment analysis using optimized feature sets in different facebook/twitter dataset domains using big data, Iraqi Journal For Computer Science and Mathematics 3 (2022) 64–70. doi:10.52866/ijcsm.2022.01.01.007.

[10] K.-X. Han, W. Chien, C.-C. Chiu, Y.-T. Cheng, Application of support vector machine (svm)

in the sentiment analysis of twitter dataset, Applied Sciences 10 (2020) 1125. doi:10.3390/app10031125.

[11] A. C. M. V. Srinivas, C. Satyanarayana, C. Divakar, K. P. Sirisha, Sentiment analysis using neural network and lstm, in: IOP conference series: materials science and engineering, volume 1074, IOP Publishing, 2021, p. 012007. doi:10.1088/1757-899X/1074/1/012007.

[12] M. Abbas, K. A. Memon, A. A. Jamali, S. Memon, A. Ahmed, Multinomial naive bayes classification model for sentiment analysis, IJCSNS Int. J. Comput. Sci. Netw. Secur 19 (2019) 62.

[13] R. H. Muhammadi, T. G. Laksana, A. B. Arifa, Combination of support vector machine and lexicon-based algorithm in twitter sentiment analysis, Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika 8 (2022) 59–71. doi:10.23917/khif.v8i1.15213.

[14] B. M. Mulyo, D. H. Widyantoro, Aspect-based sentiment analysis approach with cnn, in: 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), IEEE, 2018, pp. 142–147. doi:10.1109/EECSI.2018.8752857.

[15] N. Sultana, P. Kumar, M. Patra, S. Chandra, S. Alam, Sentiment analysis for product review, ICTACT J Soft Comput 9 (2019). doi:10.21917/ijsc.2019.0266.

[16] U. Mahadevaswamy, P. Swathi, Sentiment analysis using bidirectional lstm network, Procedia Computer Science 218 (2023) 45–56. doi:10.1016/j.procs.2022.12.400.

[17] P. F. Muhammad, R. Kusumaningrum, A. Wibowo, Sentiment analysis using word2vec and long short-term memory (lstm) for indonesian hotel reviews, Procedia Computer Science 179 (2021) 728–735. doi:10.1016/j.procs.2021.01.061.

[18] A. S. Imran, S. M. Daudpota, Z. Kastrati, R. Batra, Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets, Ieee Access 8 (2020) 181074–181090. doi:10.1109/ACCESS.2020.3027350.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper Id-6] Development of an interoperable Web Application to work in areas with limited connectivity in the Admissibility Processes of Activities, as a complement to Smart Cities★

Maycol Carhuascusma-Mescua[1,†], Roman Vargas-Huanac[1,†] and Wilver Auccahuasi[1,†*]
[1] *Universidad Continental, Huancayo, Perú*

Abstract
Smart cities are characterized by comprehensive connectivity, for which a level of integration is needed, for which integration must mature. In this sense, we find connectivity problems that occur in the areas of the Peruvian jungle, with interferences that cause cuts in the different services that use the Internet connection, such as web systems, for which the National Commission for Development and Life without Drugs (DEVIDA) is a public entity, has its headquarters in a district of the Peruvian Amazon jungle. The admissibility of activities is the registration of all activities to be carried out. These activities are characterized by controlling coca leaf plantations, as well as their eradication campaigns, being important to have the information in real time. Carrying out an analysis of similar works, we found works where XML messaging is used to achieve interoperability between the various systems, which is adopted in the present work. A web system has been designed that integrates connectivity detection functionalities in order to send information through an alternate mechanism, based on the exchange of information via XML messaging. The results indicate that it is possible to have information in real time, in all locations, which achieves interoperability between different systems.

Keywords
Smart Cities, web system, interoperability, processes, connectivity, real-time access.

1. Introduction

Analyzing the state of the art on the application of Smart Cities concepts, we found works where there is a need to optimize the design of buildings to improve energy efficiency and the quality of the urban environment, for which a simulation analysis of the wind environment and heat transfer

3rd International Symposium on Smart Cities Challenges, Technologies and Trends (SCCTT-2024)

* Corresponding author. Wilver Auccahuasi; wauccahuasi@continental.edu.pe
† These authors contributed equally.
mcarhuasuma@continental.edu.pe (Maycol Carhuasuma); rvargas@continental.edu.pe (Roman Vargas); wauccahuasi@continental.edu.pe (Wilver Auccahuasi)
0000-0001-8820-4013 (Wilver Auccahuasi)

was carried out using a CFD model, with the help of the Ecotect software together with the WinAIR plugin to simulate the air flow and thermal conditions around the building, having as results in the simulation that the distribution of wind and temperature were aligned with real conditions, validating the accuracy of the model, concluding that the simulations provided a reliable basis to guide architectural design and improvements in energy efficiency, contributing to sustainability goals in smart cities [1].

With the need to evaluate the impact of smart city policies on urban resilience in Chinese cities, a multi-period double difference model was applied to mitigate selection biases and evaluate the effect of policies, defining urban resilience indicators and a mediation analysis with the update of the industrial structure; Finding that smart city policies significantly improve urban resilience, with a mediating effect of industrial upgrading, concluding that smart city policies are effective in increasing urban resilience, and their expansion and strengthening of industrial upgrading are recommended [2].

Cybersecurity of physical devices in cyber-physical systems is vulnerable to attacks, for which an Intrusion Detection System (IDS) was integrated with a Digital Twin (DT) to simulate and monitor attacks in a controlled environment, using Eclipse Ditto as a DT platform and Snort as IDS, together with a Raspberry Pi and a Kali Linux virtual machine to carry out cyber attacks, resulting in the IDS effectively detecting Hping3 flood attacks, but showed limitations in identifying NMAP scans, with a significant increase in CPU and memory usage during attacks, concluding that the integration of DT and IDS represents a valuable innovation for cybersecurity in the IoT field, although improvements are required in the detection of certain types of attacks and in the recovery of system resources [3].

The need for a secure and efficient system for data processing in smart city environments, facing cyber threats and limitations of current IoT frameworks, for which the BFLIoT system was developed, which combines federated learning and blockchain to decentralize data processing and improve security, ProVerif was used for formal verification of the security of the BFLIoT protocol, ensuring the integrity and confidentiality of communications; A system was achieved that maintains high transaction rates and low energy consumption, with a robust anomaly detection framework, the BFLIoT represents a significant advance in IoT technology, offering a scalable, secure and efficient solution, with the potential to be applied in various sectors beyond smart cities [4].

The integration and visualization of BIM (Building Information Modeling) and GIS (Geographic Information Systems) models on web platforms for smart cities is complex and requires effective visualization solutions. A 3D web client based on Cesium was developed, using HTML5 and WebGL technologies, allowing the visualization and interaction with 3D models of cities in web browsers without the need for additional plugins, using the Cesium visualization engine and the virtualcityMAP platform, together with photogrammetric techniques to create 3D models and BIM-GIS integration tools, obtaining that the web client managed to visualize 3D models of cities and

infrastructure data, improving the interoperability between BIM and GIS and allowing users to interact with data effectively in real time; it is concluded that the solution is effective for visualizing 3D models on web platforms, but greater interoperability between BIM and GIS is required to improve analysis such as disaster management, energy analysis and indoor navigation [5].

The integration of drones into urban transportation systems presents significant challenges such as air traffic management, conflict detection and obstacle resolution, as well as the need for safe landing protocols in complex urban environments, an Intelligent Total Transportation Management System (ITTMS) was developed that incorporates drone operations, using emerging technologies such as the Internet of Things (IoT) to optimize urban mobility and traffic management, drone tracking models, real-time monitoring, advanced air traffic management and detailed landing protocols were implemented to ensure the safety and efficiency of drone operations in smart cities, as a result of the simulations showed that the proposed drone tracking models and landing protocols improved drone traffic efficiency and safety, enabling effective management of multiple drones in urban airspace; Concluding that the proposed solutions proved effective in simulations, it was concluded that further research is needed to refine these systems and address the complexities of drone management in urban environments, ensuring safe and efficient operations [6].

The need for an assessment framework for smart building (SB) integration in smart cities (SC) that considers multiple performance and sustainability factors, for which a two-round Delphi survey approach was used to validate and improve the assessment framework, collecting opinions from experts in the field, where generative artificial intelligence models, specifically ChatGPT-3 and Google Bard, were employed to assess the impact of different factors on smart city performance; a consensus was reached among experts on the importance and impact of various criteria in assessing SB integration in SC, as well as identifying areas for improvement in the proposed framework; concluding that combining AI methodologies with human expertise can address bias challenges in generative language models, and the validated framework provides an effective tool to assess smart building integration in the context of smart cities [7].

Carrying out a literature review on the different systems that work with different interoperability methodologies between systems and devices, we find in the integration of legacy systems in an industrial environment, proposing a solution divided into four main areas. The first part focuses on the incorporation of the Open Platform Communication Unified Architecture (OPC UA) protocol, widely used in the industrial environment. The second sub-section suggests the implementation of an OPC UA—Message Queue Telemetry Transport (MQTT) wrapper to enhance current digital transformation trends. The third part highlights the obsolete practices of the HTTP protocol and the legacy systems based on them, and finally, an event-based approach for data acquisition and storage without a transmission protocol is presented. The study concludes that the integration of these legacy systems can result in a significant cost reduction, especially in the data acquisition phase, which is based on obsolete technologies. Furthermore, the integration of appropriate artificial

intelligence models is suggested to apply customized levels of logic, allowing the incorporation of contemporary modules [8].
In solutions based on the integration and operation of the Publish@Ionio ecosystem, focusing on the experience of participants and their roles within the system. The importance of the module-based data structure is highlighted, which allows for efficient and compartmentalized management of access rights and responsibilities. In addition, the need to establish clear guidelines for publishing on social networks and managing content requests is mentioned, in order to avoid misunderstandings and internal conflicts [9].

The importance of standardization and collaborative integration in the implementation of Industry 4.0. In this context, the Open Communications Platform Unified Architecture (OPC UA) standard plays a crucial role by enabling the development of heterogeneous systems and facilitating the fluid exchange of data between devices. To take full advantage of OPC UA capabilities, it is necessary to unlock other application services, such as cloud computing, allowing for greater flexibility and efficiency in data management [10].

The integration of OWL reasoners into frameworks containing probabilistic reasoners written in Prolog, such as TRILL, to manage the non-determinism of tableau methods implemented by Semantic Web reasoners. Furthermore, the importance of semantic annotations of processes to enable interoperability is highlighted, citing works that show how digital transformation plays a strategic role in simplifying relationships with citizens and businesses, and in growing the community and the economy. The need to redesign processes or create new ones to ensure that a public service responds to the specific needs of different citizens is mentioned using a semantic approach to BPMN annotation with domain ontologies. The paper also reviews several works that combine OWL ontologies with more powerful logic languages such as FOL, and presents recent methods that use Prolog directly within OWL reasoners. An example is the extension of the Nova Hybrid Reasoner (NoHR) that answers queries about hybrid theories composed of an OWL ontology. In addition, OWL-S, an OWL ontology for the description of Semantic Web Services, is described, which allows the declarative specification of the semantics of web services described syntactically with WSDL. OWL-S describes web services from three points of view: the service profile, the service model and the service ontology, thus facilitating the discovery, use and composition of web services [11].
An innovative architecture for a Home Energy Management System (HEMS) is presented, which aims to improve energy efficiency and reduce environmental impact. The proposed architecture includes a combination of flexible hardware and open middleware, allowing the implementation of new energy-efficient policies and strategies. Furthermore, proofs of concept have been performed to validate the integrity of the system and its ability to adapt to future trends and technological evolutions [12].

In comparing the use of CDA R2 and FHIR standards in electronic medical record exchange in Taiwan, CDA R2, defined in 2005, is based on the XML format and is used for clinical document retrieval between institutions. However, this standard has limitations, as it only supports clinical document exchange and requires packaging all data into a single XML file, which can be a slow and cumbersome process. On the other hand, FHIR offers greater flexibility and compatibility with modern web technologies, using formats such as XML, JSON, and Turtle. FHIR allows medical records to be divided into multiple resources, making it easier to select and access specific data. In addition, FHIR supports mobile devices and resource-constrained systems, and its implementation and testing are more accessible and faster thanks to its RESTful API and the availability of test servers [13].

The need to standardize the structure and content of electronic health records (EHRs) to facilitate information exchange between medical institutions. Existing standards such as openEHR, HL7, and CEN TC251 EN 13606 are mentioned, which seek to achieve data independence and semantic interoperability. These standards use archetype-based technology to define clinical knowledge, and the study examines suitable formalisms to describe, represent, and reason about these archetypes. Furthermore, the paper highlights the importance of Semantic Web technologies in managing clinical knowledge related to EHRs. The use of platforms such as ResearchEHR is mentioned to practically apply these standards. Although it is acknowledged that information exchange is limited to specific partners, it is suggested that future research could explore intelligent and semi-built knowledge graph frameworks in the context of e-health, addressing issues such as data insufficiency, explainability, and inconsistencies [14].

In related work, the application of an Ocean Digital Twin aims to monitor environmental conditions, such as water quality and biological events. Using a combination of fixed and mobile sensors, including autonomous underwater vehicles (AUVs), the aim is to integrate and analyse data in real time to provide useful information to researchers, policy makers and industry. The implementation of a medallion data model allows the transition of raw data to formats optimised for analysis, facilitating interoperability and access through standardised APIs [15].

We found works related to the modular use of hardware in heterogeneous environments. Most of the components of this solution were developed using the ROS framework, known for its wide range of drivers, algorithms, libraries and useful features such as APIs for parameters. The architecture of the solution defines a clear division of areas within the ROS domain of the gateway, which is essential to define the purpose of each component and promote reuse. The analysis of the current literature reveals a notable absence of relevant information, especially regarding hardware modularity in robots. The only comparison found in other articles concerns the physical adaptability of the robot to its mission, rather than managing the required payloads on demand. This makes the proposed solution a promising representation of possible improvements in the field of robotics that can be solved based on interoperability between the different APIs [16].

In organizing design patterns for privacy, we present a taxonomy of pattern types that can be used to create robust and versatile systems. Patterns such as Device Gateway and Edge Computing are highlighted, which improve the effectiveness of data acquisition and local processing, respectively. In addition, the dependencies and interactions between these patterns are explored, providing a comprehensive mapping that is essential for architects and designers in applying pattern languages to real-world scenarios, which enables the use of different interoperability techniques [17].

We find work related to efficient communication between WebAssembly modules (WASM) by implementing a new interface description language (IDL) and its own communication protocol, known as Karmem. This approach seeks to improve efficiency and interoperability in distributed applications built on top of WASM, allowing effective communication between modules written in different languages. The paper details the development of Karmem, its features, and the tests performed to assess its ability to address the challenges related to efficiency and interoperability in data communication [18].

We found papers presenting an innovative methodology for the development of a hybrid machine tool system that integrates machining and surface heat treatment processes. In situ experiments and techniques such as optical microscopy, mechanical and X-ray methods were used to determine residual stresses, as well as a profilometer to measure shape deviations and surface roughness. The results indicate that the implementation of this system can increase productivity by 1.9 times and eliminate the possibility of waste during finishing [19].

In the blockchain-based framework for electronic health records (EHR) management, highlighting its ability to improve security, efficiency, and interoperability in handling patient data. It focuses on overcoming the limitations of centralized systems, such as the risk of a single point of failure, by using smart contracts and distributed storage through IPFS. In addition, a cost analysis is performed to assess the economic viability of the proposed framework [20].

Among the growing challenges in urban mobility management as urban populations increase, an AI-based Decision Support System (DSSU) is proposed that enables planners and policy makers to assess and prioritize urban mobility modifications using city-specific multi-criteria criteria. Through case studies in Helsinki, Amsterdam, Messina and Bilbao, the transformative potential of the recommendation engine to improve mobility policies and foster more livable and resilient urban environments is demonstrated, with a smart city approach, allowing the exploitation of different architectures based on interoperability [21].
We found works presenting the LAAFFI protocol, designed for the authentication and authorization of IoT devices in federated environments, such as those involving civil and military organizations. The design requirements, formal security validation, and the resilience of the protocol against various attacks are discussed. In addition, a prototype is implemented to evaluate performance metrics such as latency and throughput in terms of operations per second [22].

The problem is related to the poor connectivity in the Peruvian jungle areas, which is why computer systems have difficulty saving and updating data in their databases, causing the data to not be updated. The National Commission for Development and Life without Drugs (DEVIDA) is a public entity, attached to the sector of the Presidency of the Council of Ministers, responsible for guiding the national policy against drugs to the year 2030 and becoming the national counterpart of all international funding destined for the fight against drugs. Currently, it has its headquarters in a district of the Peruvian Amazon jungle.

This problem is solved in the present work, through the implementation of a procedure based on XML messaging, which allows to always have updated information, so when an attempt is made to perform an information registration process, and the lack of connectivity is detected, the system collects all the information to be sent, packages it in an XML file and sends it by alternate means, in order to solve the problem and always have access to the information.

2. Methodology

We begin the description of the methodology with the analysis of the functional requirements that we present below:

2.1. Target population

The web system is aimed at solving the problem of intermittent connectivity and access to the Internet, in web systems that require real-time information on the different financing requests made by executing entities belonging to the districts located in the departments that are in the areas located in the Peruvian Amazon. In this way, the efforts of the activity are directed towards solving the problem of illicit crops that are part of the drug trafficking value chain, through the availability of real-time information.

2.2. Description of the development methodology

The description of the proposal is based on being able to explain the technological tools used for the development of the web system, considering the development mode, the configuration of the web system, the messaging services and the working mode of the interoperability function, which allows the information to always be available in real time.

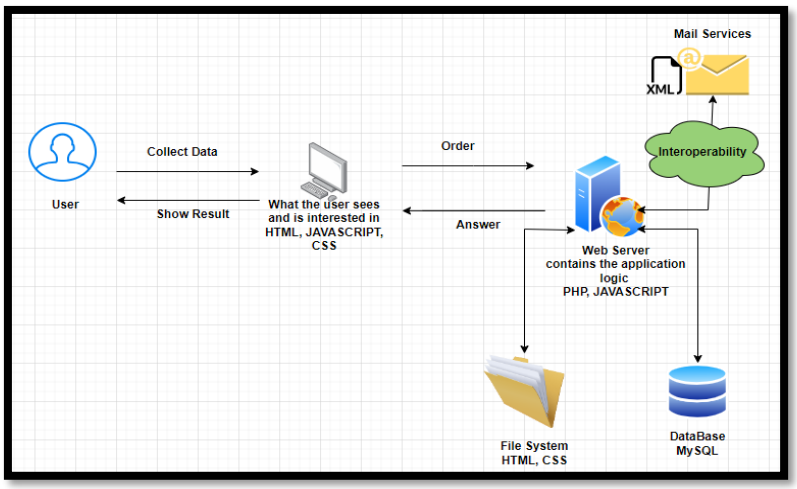


Figure 3: Description of the architecture of the development proposal.

Figure 1 shows the architecture of the system, based on the flow of information between the server and the user for the storage of the database with the characteristic of interoperability, based on the exchange of XML messages.

The development of the web application for the eligibility of activities in the investment area of DEVIDA, based in the city of Pichari, was carried out using standard web technologies, such as HTML, CSS, JavaScript, PHP, and additional frameworks or libraries as needed. Finally, extensive testing was carried out to verify the correct functioning of the web application, as well as to ensure its quality and usability. Next, we are going to present the screenshots of the web application development process, according to the following detail:

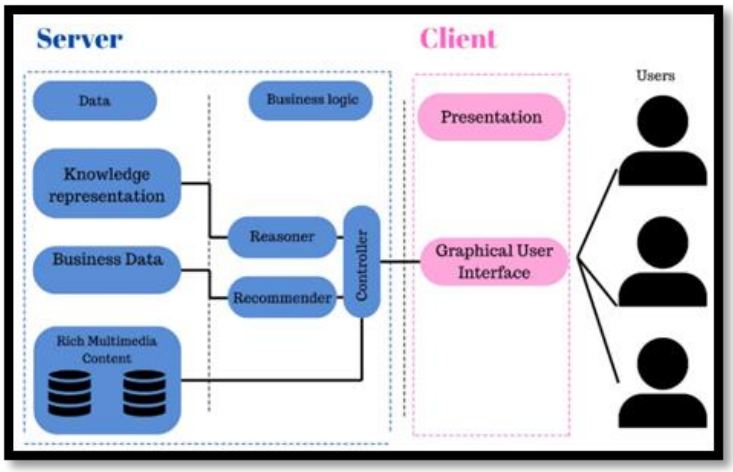


Figure 4: Connectivity diagram.

Figure 2 presents the connectivity model between the different components where the model is verified from the server and client views. In the server view, the presence of the database, the business model and the information containers, as well as the business model, can be seen. For the client view, the presentation mode can be seen, which is based on the representation of the GUI, which is the graphical user interface that is the means of communication between the different users and the system.

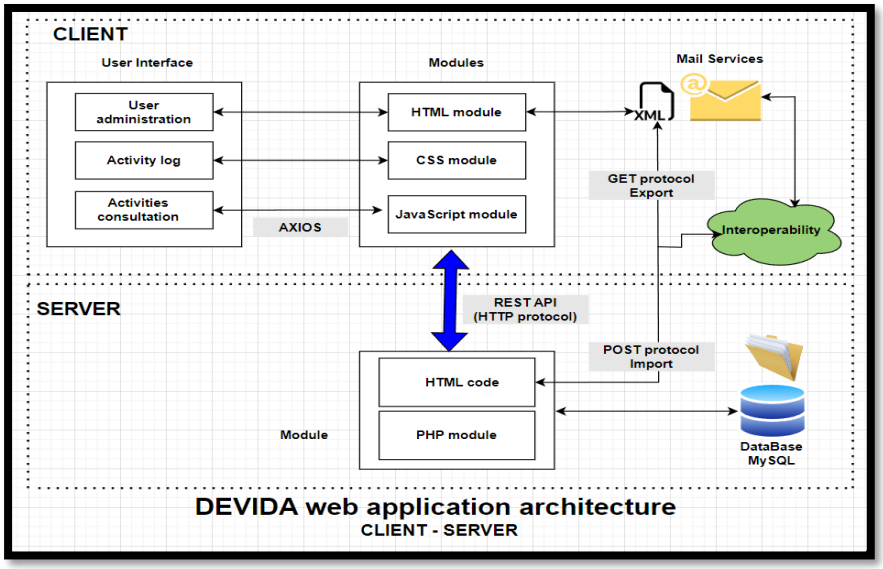


Figure 5: Connectivity tools.

Figure 3 shows the development of the web application. It was carried out using the client-server architecture, supported by the Laravel VueJS Stack, which provides the necessary and easy-to-use tools for the construction of both the client side (Front end) and the server side and the database (Back end).

On the server side, we have worked with the PHP programming language, guided by its Laravel framework, with the intention of taking advantage of the security tools, speed in the execution of tasks and the working mode that is very orderly and sequential.

On the client side, we have worked with the HTML tag language, CSS cascading style sheets and JavaScript accompanied by its VueJS framework, in order to give greater dynamism to its components and show an elegant style.

MySQL has been used for database management, because there are currently more servers that support MySQL and it is better suited to systems developed with the PHP language. In the communication between the Back end and the Front end, a REST Application Programming Interface (API) (with HTTP protocol) has been used, which was developed in the Front end and the

queries are made from JavaScript using Axios, which is a very necessary library to make operations easier.

We adapted the Extensible Markup Language (XML) in the development of the system, with the intention of exchanging data in rural areas where there is no internet connectivity. Thus, we stored a button in the Front end, which has the function of making a Get request to the Back end, sending the ID of the activity that we want to export, then the Back end receives the ID through said Get protocol, to extract all the information of the selected activity that is in the Database, to be exported in XML format. Likewise, to import an activity, a file-type button was inserted, which allows uploading an XML file, using the POST protocol that carries all the information of said activity, proceeding to structure an HTML that can be understood by the Back end and store it in our Database, without any inconvenience.

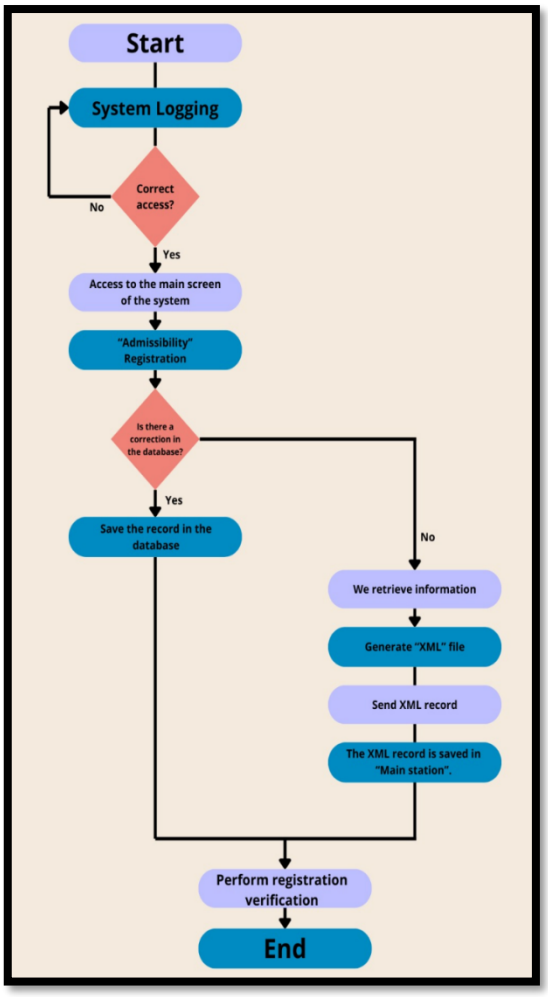


Figure 6: Flowchart of the we system.

In Figure 4, the flow chart of the interoperability solution is presented, where there are two control points, the first related to access to the system, which limits access to the system, the second is related to the functionality of the system in order to decide whether to send the information through the system itself when connectivity is present, and when connectivity is not present, the XML message is created to be able to send it through alternate means.

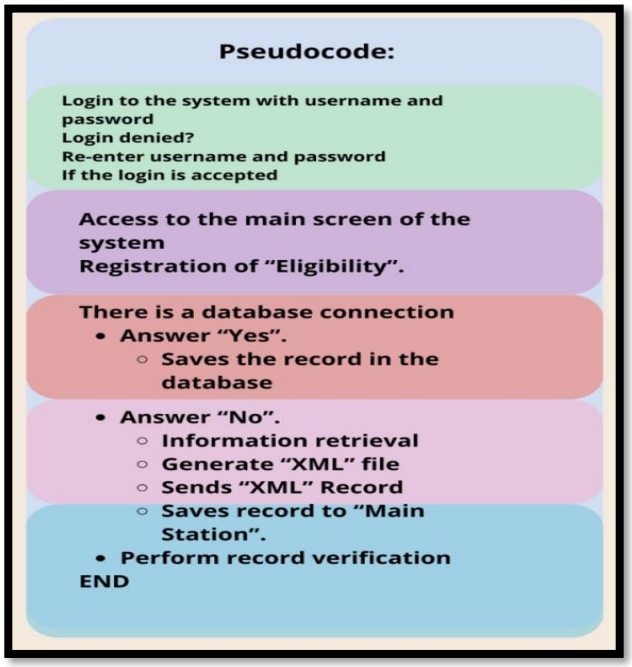


Figure 7: Pseudocode of the we system.

In figure 5, the pseudocode corresponding to the flow chart is presented, where we describe the processes and functionalities so that they can be implemented at the time of programming, the two decision factors are considered, the access to the system and the functionality of sending information through XML messaging.

3. Results

The results we present are dedicated to presenting the mode of use of the implemented web system, with the functionalities of interoperability through the use of XML messaging.

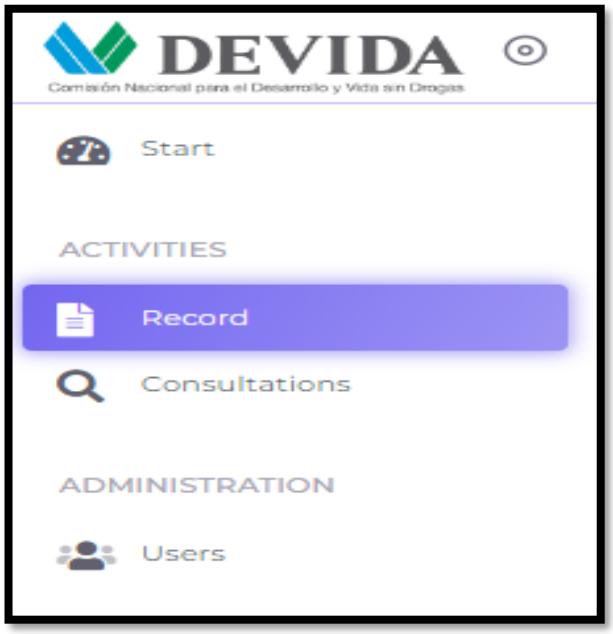


Figure 8: Main screen of the web system.

Figure 6 shows the main screen of the developed web system, where the main processes are displayed. It is an essential requirement to have a username and password to access the system.

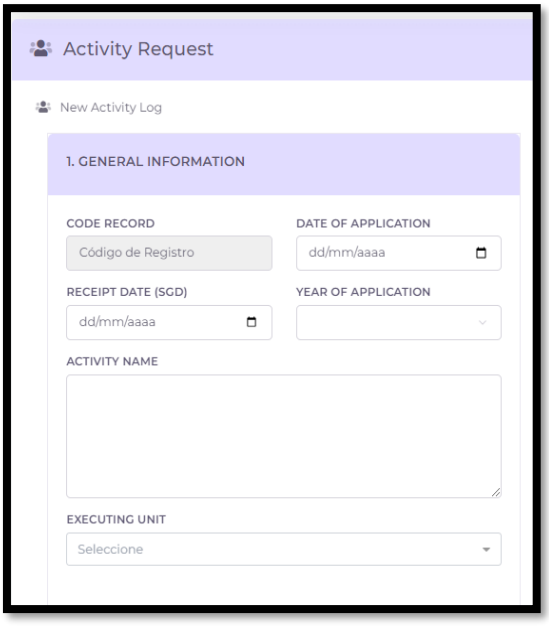


Figure 9: Main screen of the activity request.

In Figure 7, we present the screen where a request for an activity is made, where the main data, typology and location are shown. This information is important, as it is a requirement to have the information of the place where the registration is made and the registration date, in order to have relevant information for decision-making.

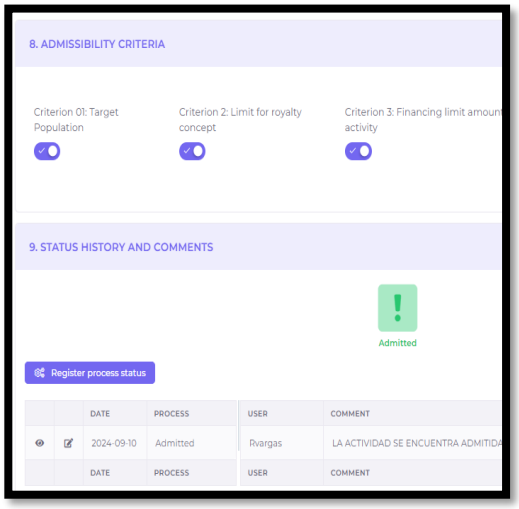


Figure 10: Record of the processes.

Figure 8 presents the eligibility criteria, which is the starting point for having a record. The form allows the selection of a criterion, it also allows the status of the record to be shown and the report of the records made.

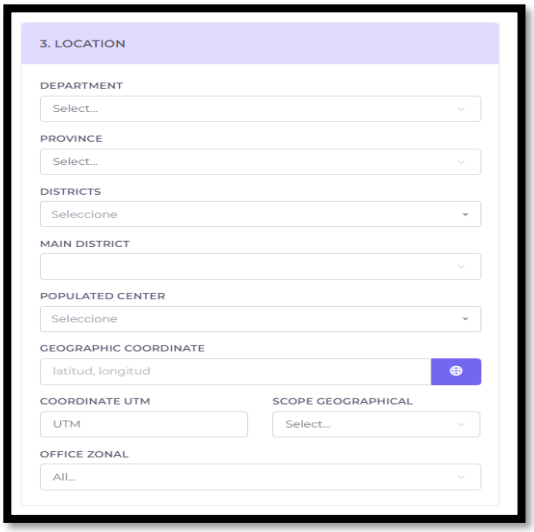


Figure 11: Location record.

Figure 9 shows the form for registering the location. This information is important to know where in the Amazon rainforest the registration is being carried out.

Figure 12: Record of the processes.

Figure 10 shows the completion of entering information for a complete record. It shows that the information is in the database, so it can be accessed from anywhere as long as there is connectivity and access to the system.

Figure 13: Interoperability option.

Figure 11 presents the interoperability option. This option is used when it is not possible to save directly in the database. With the option of exporting via XML file, we can send the information via alternative means, in order to update the system.

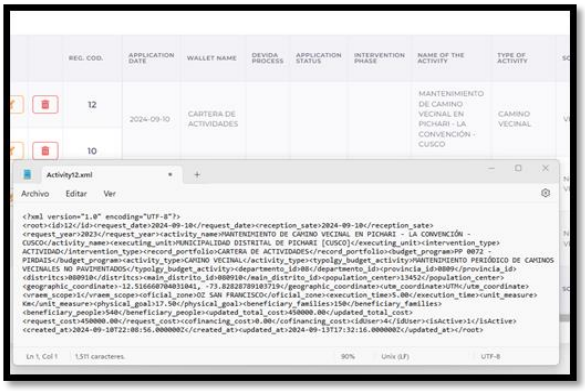


Figure 14: XML file.

Figure 12 shows the XML file containing the record information to be exported and sent via alternate means for inclusion in the system.

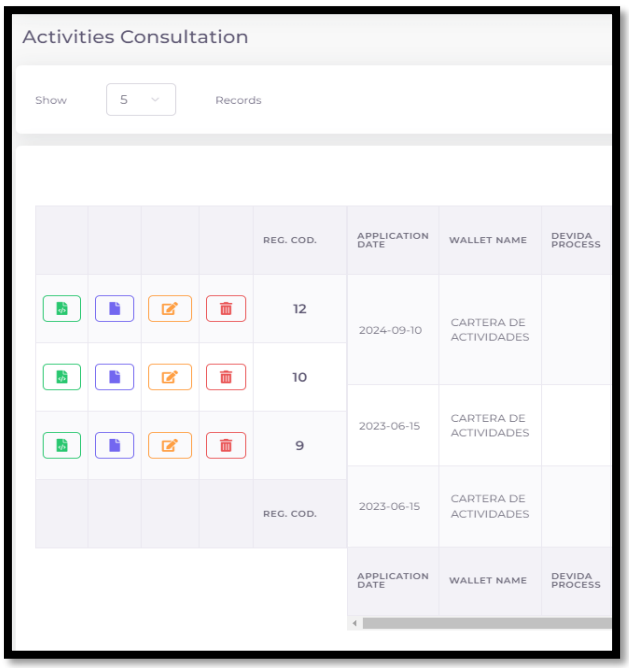


Figure 15: Search for records.

Figure 13 shows the details of the information search, after having updated the exported files through XML messages. With the update of the files, the records can be accessed from any location that has connectivity with the system.

4. Discussions

The discussions we present are related to the analysis of the use and exploitation of the system. In the tests carried out, it was possible to demonstrate that the interoperability functionality eliminates downtime in the search processes, due to the lack of connectivity in the registration process.

By exporting the information of each record in an XML file, the system can be updated. The alternate means described can be email, shared files, file exchange platforms, among others, using mobile or alternate connectivity. By carrying out the analysis from the central office, it can be demonstrated that the information is always available, which optimizes the search and analysis process of the activities carried out in the Peruvian Amazon rainforest, so that decisions can be made from the highest levels of government.

Regarding the performance of the system, it is observed that it works according to the design, with which the system works on the available computers without delays, because it has a local database, which improves the performance in search time mainly in local records, functioning as a decentralized database.

5. Conclusions

The conclusions we reached at the end of the investigation are determined to be able to demonstrate the use of interoperability between systems, based on the use of XML messages, which allows communication between systems that are in different physical locations. It was shown that alternative solutions can be considered in the absence of connectivity. In a normal situation, one has to wait for a connection to be available to update the records. These downtimes can cause a failure in decision making, affecting the tasks related to the control of coca leaf crops in the Peruvian jungle.

We recommend replicating the proposed methodology for situations where communication between different systems is required, using the different interoperability techniques. Many alternatives are available, including commercial solutions, libraries, and even proprietary solutions based on the XML language. Using a communication protocol at the time of writing and reading the files, we can achieve communication between different systems.

In the process of developing the user interface of the web application, moderate usability tests were carried out remotely, in which together with the users we found small details in the creation of new users (functions and roles), chronological disorder in the list of activities when queries were made and inconveniences in rural areas, where there are low levels of internet connectivity. These findings allowed us to better adapt the code by simulating the real process carried out in the investment area of DEVIDA, then two buttons were implemented to import and export activities through the extensible markup language (XML), in order to always have updated information on the files that are subject to constant evaluation by the area in charge of DEVIDA. With the latest version of the web application, users showed comfort in handling the system, who indicated that the activities they carried out physically, they can now do digitally, in an automated way, having the files to be evaluated at any time, which allowed them to spend less time in the admissibility of activities. This technological implementation also contributed significantly to the productivity of the company DEVIDA, being the beginning of an appropriate escalation to achieve effectiveness in the work of the DEVIDA staff based in Pichari.

References

[1] Li, Y.; Chen, H.; Yu, P. Green Campus Transformation in Smart City Development: A Study on Low-Carbon and Energy-Saving Design for the Renovation of School Buildings. *Smart Cities* 2024, *7*, 2940-2965. https://doi.org/10.3390/smartcities7050115

[2] Feng, Y.; Wang, J.; Zhang, T. The Impact of Smart City Policies on City Resilience: An Evaluation of 282 Chinese Cities. *Sustainability* 2024, *16*, 8669. https://doi.org/10.3390/su16198669

[3] El-Hajj, M. Leveraging Digital Twins and Intrusion Detection Systems for Enhanced Security in IoT-Based Smart City Infrastructures. *Electronics* 2024, *13*, 3941. https://doi.org/10.3390/electronics13193941

[4] Sefati, S.S.; Craciunescu, R.; Arasteh, B.; Halunga, S.; Fratu, O.; Tal, I. Cybersecurity in a Scalable Smart City Framework Using Blockchain and Federated Learning for Internet of Things (IoT). *Smart Cities* 2024, *7*, 2802-2841. https://doi.org/10.3390/smartcities7050109

[5] Varlık, A.; Dursun, İ. Three-Dimensional Web-Based Client Presentation of Integrated BIM and GIS for Smart Cities. *Buildings* 2024, *14*, 3021. https://doi.org/10.3390/buildings14093021

[6] Nguyen, D.-D.; Dang, Q.-D. Development of an Intelligent Drone Management System for Integration into Smart City Transportation Networks. *Drones* 2024, *8*, 512. https://doi.org/10.3390/drones8090512

[7] Shahrabani, M.M.N.; Apanaviciene, R. An AI-Based Evaluation Framework for Smart Building Integration into Smart City. *Sustainability* 2024, *16*, 8032. https://doi.org/10.3390/su16188032

[8] Korodi, A.; Nițulescu, I.-V.; Fülöp, A.-A.; Vesa, V.-C.; Demian, P.; Braneci, R.-A.; Popescu, D. Integration of Legacy Industrial Equipment in a Building-Management System Industry 5.0 Scenario. Electronics 2024, 13, 3229. https://doi.org/10.3390/electronics13163229

[9] Giannakoulopoulos, A.; Pergantis, M.; Lamprogeorgos, A. User Experience, Functionality and Aesthetics Evaluation in an Academic Multi-Site Web Ecosystem. Future Internet 2024, 16, 92. https://doi.org/10.3390/fi16030092

[10] Molina, A.; Vargas, D.; Rodas, A. Implementation of A Data-Acquisition System and Its Cloud-Based Registration Using the Unified Architecture of Open Platform Communications. Eng. Proc. 2023, 47, 20. https://doi.org/10.3390/engproc2023047020

[11] Redavid, D.; Ferilli, S. Semantic Web Services Ingestion in a Process Mining Framework. Electronics 2023, 12, 4767. https://doi.org/10.3390/electronics12234767

[12] Motta, L.L.; Ferreira, L.C.B.C.; Cabral, T.W.; Lemes, D.A.M.; Cardoso, G.d.S.; Borchardt, A.; Cardieri, P.; Fraidenraich, G.; de Lima, E.R.; Neto, F.B.; et al. General Overview and Proof of Concept of a Smart Home Energy Management System Architecture. Electronics 2023, 12, 4453. https://doi.org/10.3390/electronics12214453

[13] Liu, T.-J.; Lee, H.-T.; Wu, F. Building an Electronic Medical Record System Exchanged in FHIR Format and Its Visual Presentation. Healthcare 2023, 11, 2410. https://doi.org/10.3390/healthcare11172410

[14] Sachdeva, S.; Bhalla, S. Using Knowledge Graph Structures for Semantic Interoperability in Electronic Health Records Data Exchanges. Information 2022, 13, 52. https://doi.org/10.3390/info13020052

[15] Vasilijevic, A.; Brönner, U.; Dunn, M.; García-Valle, G.; Fabrini, J.; Stevenson-Jones, R.; Bye, B.L.; Mayer, I.; Berre, A.; Ludvigsen, M.; et al. A Digital Twin of the Trondheim Fjord for Environmental Monitoring—A Pilot Case. J. Mar. Sci. Eng. 2024, 12, 1530. https://doi.org/10.3390/jmse12091530

[16] Carreira, R.; Costa, N.; Ramos, J.; Frazão, L.; Pereira, A. A ROS2-Based Gateway for Modular Hardware Usage in Heterogeneous Environments. Sensors 2024, 24, 6341. https://doi.org/10.3390/s24196341

[17] Azeez, H.H.; Sharbaf, M.; Zamani, B.; Kolahdouz-Rahimi, S. Applying Pattern Language to Enhance IIoT System Design and Integration: From Theory to Practice. Information 2024, 15, 595. https://doi.org/10.3390/info15100595

[18] Silva, L.; Metrôlho, J.; Ribeiro, F. Efficient Data Exchange between WebAssembly Modules. Future Internet 2024, 16, 341. https://doi.org/10.3390/fi16090341

[19] Skeeba, V.; Ivancivsky, V.; Chernikov, A.; Martyushev, N.; Vakhrushev, N.; Titova, K. Design and Technological Aspects of Integrating Multi-Blade Machining and Surface Hardening on a Single Machine Base. J. Manuf. Mater. Process. 2024, 8, 200. https://doi.org/10.3390/jmmp8050200

[20] Tahir, N.U.A.; Rashid, U.; Hadi, H.J.; Ahmad, N.; Cao, Y.; Alshara, M.A.; Javed, Y. Blockchain-Based Healthcare Records Management Framework: Enhancing Security, Privacy, and Interoperability. Technologies 2024, 12, 168. https://doi.org/10.3390/technologies12090168

[21] Shulajkovska, M.; Smerkol, M.; Noveski, G.; Bohanec, M.; Gams, M. Artificial Intelligence-Based Decision Support System for Sustaina

[22] Jarosz, M.; Wrona, K.; Zieliński, Z. Distributed Ledger-Based Authentication and Authorization of IoT Devices in Federated Environments. Electronics 2024, 13, 3932. https://doi.org/10.3390/electronics13193932

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-25] Enhancement of Accuracy in Soil Urea Estimation Using Machine Learning Tools

Sulaxana Vernekar[1, †], Marlon Sequeira[2,†,*], Sameer M. Patil[3, †], Jivan Parab[2, †] and Prof. Gourish Naik[2, †]

[1] *GVM'S GGPR College of Commerce and Economics*
[2] *Electronics Programme, School of Physical and Applied Sciences,Goa University, Goa, India*
[3] *Dnyanprassarak Mandals College & Research Centre, Mapusa, Goa India*

Abstract
Soil health is vital for getting a good crop yield. Analysis of available soil nutrients done at the right time not only helps in conservation of soil fertility but can also help in getting a good crop yield by limiting the usage of external inputs to the soil such as fertilizers, water etc. The use of AI in agriculture is being explored in recent times to optimize the crop yield. Machine Learning techniques are used in developing smart soil sensing systems to provide accurate soil nutrients distribution. In this study, a sample of 40 spectral data in the frequency range of 500MHz to 1000MHz was passed to the ParLeS software. The PLSR cross validation in ParLeS gave us an RMSE of 2.87. However, when Ridge regression based on machine learning was applied, we obtained a RMSE of 1.02 with parameter alpha set to 0.005. Thus, we can say conclusively that, machine-learning methods yield better results than traditional methods. In addition, implementation of ParLeS needs LabVIEW type of environment and needs external graphics support, whereas, Ridge regression can be implemented using simple Python environment, which is now a day most often used programming language. The implementation does not require compulsory graphics support.

Keywords
ParLes, PLSR, Ridge Regression.

[3]* Corresponding author.
† These authors contributed equally.
sulaxgoa@gmail.com (S. Vernekar); marlon@unigoa.ac.in (M. Sequeira); sameer@dmscollege.ac.in (S. Patil); jsparab@unigoa.ac.in (J. Parab); gmnaik@unigoa.ac.in (G Naik)
(S. Vernekar); 0000-0002-7462-3492; 0000-0002-1444-5428 (M. Sequeira); 0000-0002-0848-7349, 0009-0007-3674-1942 (S. Patil); (J. Parab)

1. Introduction

Agriculture is the backbone of any thriving economy. Advancements in technology has seen lot of influence on the way agriculture is practiced. Smart farming has paved a way for sustainable agriculture and increasing the crop productivity. Soil fertility plays an important role in crop production [1]. The available nutrients in the soil can highly influence the crop yield. Cultivating crops constantly without proper analysis of the soil can deteriorate its health leading to the soil becoming arid. Smart farming techniques are based on micromanagement of the farm taking into consideration the spatial and temporal variability exhibited by soil. This enables the proper management of external inputs such as fertilizer and pesticide application etc. to the soil [2]. Proper understanding and knowledge about the soil can enable the farmers to take proper decisions in crop management thus enhancing the crop productivity [3]. There are several issues that need to be tackled in agriculture such as lack of digitization, food safety issue, ecological problems and inefficient agri-food supply chain. Integration of Industry 4.0 in agriculture can greatly influence productivity, agri-food supply chain efficiency, food safety, and the sustainable use of natural resources [4].

Artificial Intelligence (AI) is the most rapidly growing technology embedded into all aspects of human life. In agriculture AI technologies can be used in precision farming for soil and irrigation management, weather forecasting, plant growth, disease prediction, and animal management [5]. With the exponential growth and development of data processing, information technology, and artificial intelligence, smart farming makes use of cutting-edge innovations to boost productivity and reduce labor stress and automating soil and crop management with AI [6]. Smart soil prediction is a low-cost method of forecasting a soil's performance over a wide range of crops.

Digital soil mapping (DSM) is used to generate digital maps of the type and quality of soil by combining soil sensing data with environmental factors [7]. Recent years have seen a significant rise in the use of DSM in soil science, which can be attributed to the integration of several ideal factors, including, but not limited to, tremendous interest in quantitative and spatial soil information, the buildup of databases of predicted or interpreted soil properties together with thoroughly known environmental factors, and the development of computational methods combined with computer resources to extract these stores of soil data [8]. Obtaining exact data on soil nutrient composition is a critical step in the implementation of precision agriculture and DSM is providing a potential breakthrough [9]. Artificial Intelligence tools such as fuzzy systems, decision trees, expert knowledge, machine learning algorithms, deep learning methodologies, and other artificial intelligence technologies can be used to provide more accurate forecasts and solutions in DSM [10].

AI models and DSM have been utilized in soil fertility prediction, offering a decision-making tool capable of forecasting the best crop based on soil pH, soil nutrients, soil moisture, environmental variables, and other components [11]. It was observed from a study conducted on prediction of soil nutrients using spectroscopic data that using Machine Learning (ML) techniques greatly improves the accuracy of soil nutrient prediction [12]. ML algorithms were used in a study to find the relationship between independent variables and dependent variables for soil data analysis. The independent variables were moisture, temperature, soil pH, Cation Exchange Capacity(CEC) and the dependent variables were Nitrogen, Phosphorus and Potassium (NPK). This study showed that there exist relationships between Phosphorus, Potassium, soil pH and CEC; Nitrogen and soil moisture and temperature using ML algorithms [13].

In another review study on using machine learning methods for predicting soil properties, agricultural yield, and soil fertility, it was observed that for soil prediction, Random Forest (RF) and deep learning techniques surpass traditional ML algorithms. Depending on the model's inputs, the RF and deep learning techniques can reliably forecast soil conditions and crop to be grown. It was also found from the study that inaccurate data has the ability to reduce forecasting precision. Variations in geographical elements, meteorological circumstances, and farming techniques can hamper the process of generalizing models. Furthermore, selecting relevant characteristics from numerous influencing factors necessitates subject expertise and testing [1].

2. Methodology

To obtain the RF spectra of various samples a cell is designed based on the principle of dielectricity. The design details of the cell are discussed in [14].

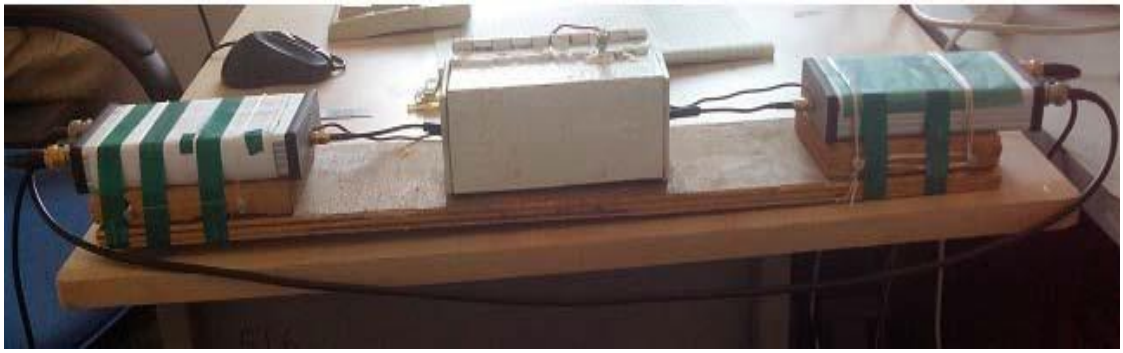


Figure1: Experimental Setup

The experimental setup consists of a cell which is placed inside the iron box at the centre as shown in figure 1. A Signal Hound tracking generator USBTG44A and a Signal Hound spectrum

analyzer USB-SA124B were used for obtaining the RF response. The sample is placed in the cell and RF signal from the tracking generator is injected into the cell through the central copper wire. The strength of the signal reduces due to dielectric loss offered by the sample solution as the signal propagates towards the receiver end.

The RF spectrum analyzer connected at the receiver end of the cell captures signal proportional to the radiation loss due to the sample solution. The cell has a capacity of holding 15ml of liquid. Soil samples were prepared in the laboratory by mixing 5 different components urea, potash, sodium chloride, calcium carbonate and phosphate in distilled water. Molar solutions for each of the component was prepared and for 15ml of water the amount of each component required to be added was calculated. It was found that amount of urea required was 225mg/15ml. Similarly, for the remaining components the amount required to be added for 15ml of water was calculated. The amount of each component to be added is shown in table 1.

Table 1: Concentrations denotation table

Concentrations denotation	Concentration(mg/15ml)				
	Urea	Potash	Phosphate	Lime	Salt
0.5	112.5	139.7	1890	187.5	109.87
1	225	279.4	3780	375	219.74
1.5	337.5	419.1	5670	562.5	329.61
2	450	558.8	7560	750	439.48
3	675	838.2	11340	1125	659.25

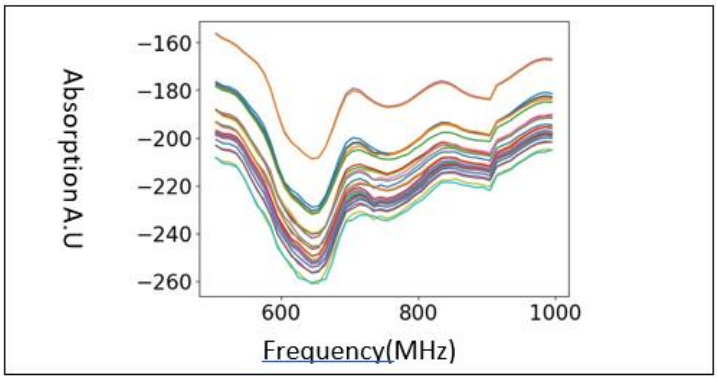


Figure 2: RF Spectra of 40 samples in the frequency range 500MHz-1000MHz

Figure 2 shows the RF spectra of 40 samples which were used in building the model for soil urea estimation. These samples were prepared by adding different concentrations of the components taken as per the table 1.

The soil urea estimation using this spectral data of 40 samples was done using two methods.

The first method was using the ParLeS software based on Partial Least Squares Regression(PLSR) model. The second method was using Machine Learning Algorithm i.e. Ridge Regression.

ParLeS is a chemometrics software for multivariate modelling and prediction. It provides users with various algorithm options to transform, preprocess and pretreat spectra. It may be used to implement principal components analysis (PCA); partial least squares regression (PLSR) with leave-n-out cross validation; and bootstrap aggregation-PLSR (bagging-PLSR). ParLeS facilitates the implementation of a large number of preprocessing techniques as well as bagging-PLSR, which can improve the robustness and accuracy of PLSR models. Other unique features of ParLeS include the provision of a number of assessment statistics and graphical output as well as a user-friendly interface and functionality [15].

In this study, a sample of 40 spectral data in the frequency range of 500MHz to 1000MHz was passed into the ParLeS software. The PLSR cross validation technique was used for soil urea estimation where the n=5 was chosen for the cross validation. Using this an RMSE of 2.87 was obtained. A screenshot of the ParLeS software is as shown in Figure 3.

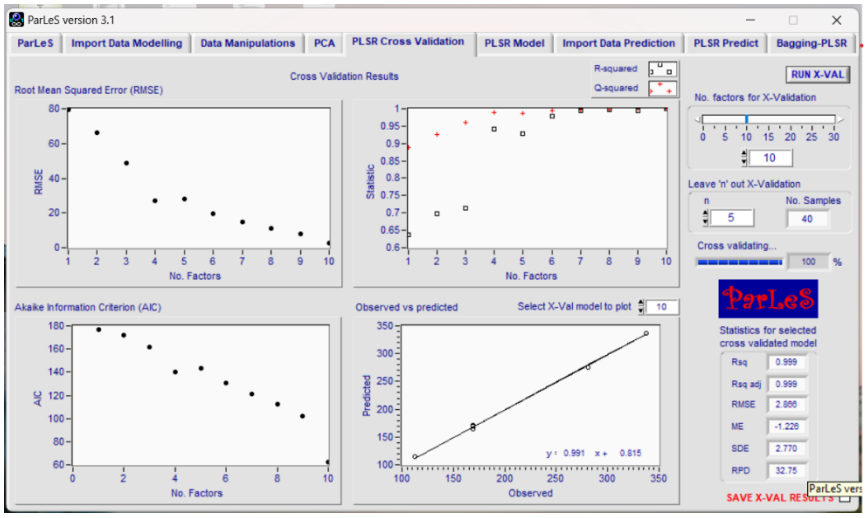


Figure 3: Screenshot of ParLeS software

Ridge regression, also known as L2 regularization, is one of the many regularization techniques applied to linear regression models. Regularization is a statistical method used to prevent errors due to the overfitting of training data. Ridge regression is specifically tailored to address multicollinearity in regression analysis, which is crucial when developing machine learning models with many parameters, particularly when these parameters are significantly weighted.
A standard, multiple-variable linear regression equation is:

$$Y = X_0 + B_1X_1 + B_2X_2 + \cdots + B_nX_n$$

In the above equation, Y represents the expected value or the dependent variable, X is the predictor or independent variable, B denotes the regression coefficient linked to that independent variable, and X0 is the value of the dependent variable when the independent variable is zero, also referred to as the y-intercept. It's important to observe how the coefficients illustrate the relationship between the dependent variable and a specific independent variable. The best-fitting line for a given dataset is obtained by calculating coefficients for each independent variable that result in the smallest residual sum of squares (also called the sum of squared errors).
The Residual sum of squares (RSS) represents how well a linear regression model matches the training data and is represented by the formula:

$$RSS = \sum_{i=1}^{n} (Y_i - \hat{Y}_l)^2$$

This formula is used to calculate the accuracy of model predictions against the expected values in the training data. If the Residual Sum of Squares (RSS) is zero, it indicates that the model perfectly predicts the dependent variables. If two or more variables have a strong linear correlation, high-value coefficients are generated, causing the model's output to be sensitive to minor changes in the input data. This indicates that the model overfitted on a single training dataset and is unable to correctly generalise to new test datasets. This causes the model to be unstable.

Multicollinearity exists when two or more predictors have a near-linear relationship or are highly correlated, which results into unreliable and unstable estimates of regression coefficients. Ridge regression is a procedure for eliminating the bias of coefficients and reducing the mean square error by shrinking the coefficients of a model towards zero in order to solve problems of overfitting or multicollinearity that are normally associated with ordinary least squares regression.

Ridge regression corrects for high-value coefficients by introducing a regularization term (often called the penalty term) into the RSS function. This penalty term denoted as L2, is the sum of the squares of the model's coefficients. It is represented in the formulation:

$$RSS_{L2} = \sum_{i=1}^{n} (Y_i - \hat{Y}_l)^2 + \lambda \sum_{j=1}^{P} B_j^2$$

The L2 penalty term reduces all coefficients to balance the high ones. This process is utilized in ridge regression, to calculate new coefficients that minimize the residual sum of squares (RSS) for a model, thereby reducing overfitting.

Ridge regression doesn't reduce all coefficients equally and is proportional to their original magnitude. As the lambda (λ) parameter increases, coefficients with higher values diminish more rapidly than those with lower values, resulting in a greater penalty for the former [16].

In machine learning, ridge regression is used to reduce overfitting that results from model complexity. Model complexity can be due to a model possessing too many features and features possessing too much weight. Feature weight refers to a given predictor's effect on the model output.

In machine learning terms, ridge regression amounts to adding bias into a model for the sake of decreasing that model's variance. Bias measures the average difference between predicted values and true values and variance measures the difference between predictions across various realizations of a given model. As bias increases, a model predicts less accurately on a training dataset. As variance increases, a model predicts less accurately on other datasets. Bias and variance thus measure model accuracy on training and test sets respectively. To reduce the model bias and variance, ridge regression technique can be used [16].

Using Ridge regression technique allows control over the bias-variance trade-off. Increasing the value of λ increases the bias but reduces the variance, while decreasing λ does the opposite. The goal is to find an optimal λ that balances bias and variance, leading to a model that generalizes well to new data.

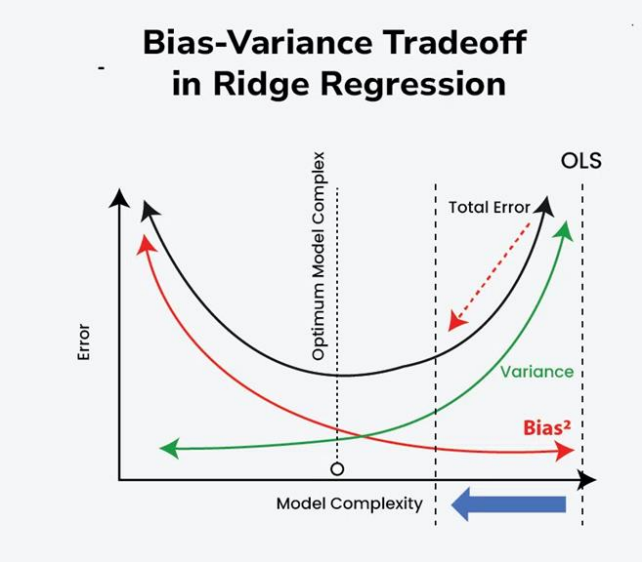


Figure 4: Bias Variance Tradeoff

Selection of an appropriate value for the ridge parameter k is crucial in ridge regression, as it directly influences the bias-variance tradeoff and the overall performance of the model. There are several methods for the selection of ridge parameter:

1. Cross-Validation

Cross-validation is one of the most popular method used in the selection of the ridge parameter. In this method, the dataset is divided into multiple subsets, and the model is trained on some subsets while being validated on the remaining ones. The process is repeated over multiple iterations, and the average performance across all iterations is used to determine the optimal value of λ.

- **K-Fold Cross-Validation**: The dataset is divided into K subsets (folds). The model is trained on K- folds and validated on the remaining fold. This process is repeated K times, with each fold being used as the validation set once. The average performance across all folds is used to select λ.
- **Leave-One-Out Cross-Validation (LOOCV)**: A special case of K-fold cross-validation where K equals the number of observations. Each observation is used as a validation set once, and the model is trained on the remaining observations. This method is computationally intensive but provides an unbiased estimate of the model's performance.

2. Grid Search: This method defines a grid of possible values for λ and the ridge regression model is trained for each value of λ. The performance of the model is evaluated for each value of λ from the grid and the one with the best performance is then selected as the ridge parameter.

3. Bayesian Optimization: Bayesian optimization is used to efficiently explore the space of possible λ values and find the optimal value. This method can be more efficient than grid search for large search spaces.

4. Information Criteria:

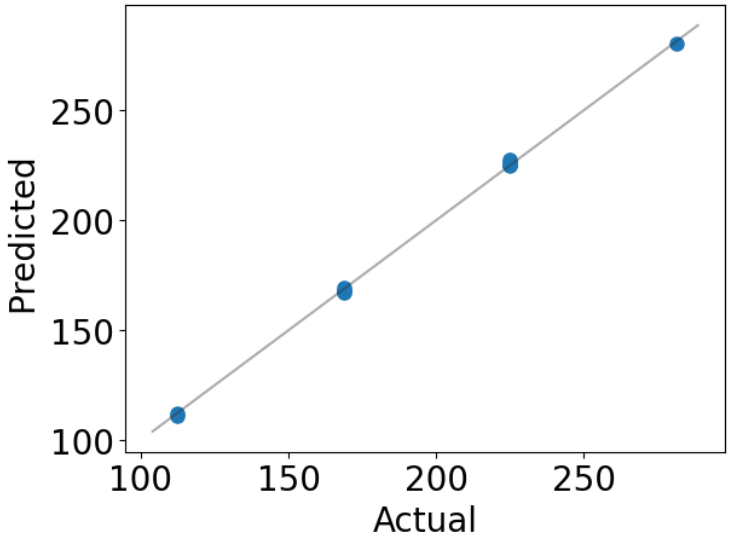
Use information criteria like Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) to select the optimal value of λ. These criteria balance model fit and complexity.

5. Domain Knowledge:

- Incorporate domain knowledge about the problem to guide the choice of λ. For example, if you know that overfitting is a significant concern, you might choose a larger value of λ [17].

Ridge regression was implements using the sklearn python package in the python programming language. The sklearn package includes the Application Programming Interface (API) interface to implement the same. The linear_model.Ridge() API is used to implement the ridge regression. The only parameters supplied to the API is alpha with a value of 0.005 . Here it may be noted that the alpha is equivalent to λ specified above. The others parameters have default values. The parameter with the default values are copy_X=True, fit_intercept=True, tol=0.0001, max_iter=None, positive=False, solver='auto', and random_state=None.

The dataset consisting of 40 samples was used for training and testing the ML model using Ridge regression. With the parameter alpha set to 0.005, the RMSE obtained using this technique was found to be 1.02.



(mg/15ml)

Figure 5: Typical graph showing the actual and predicted urea values

3. Result and Discussion

It may be seen that the analysis using Ridge technique (Which is a machine learning based tool for regression analysis) gives excellent performance with error as low as 1.02. Whereas, the error in traditional technique of ParLeS regression is 2.87, which is nearly three times more than that of Ridge technique. As mentioned earlier that, in addition to the advantage of less error, the implementation of the algorithm can be done in a simpler computational platform, not necessarily requiring complicated LabVIEW back end. This is reflected in the table 2. The regression graph shown in figure 5 show good agreement between the actual and predicted values using the ridge regression technique.

Table 2: Result obtained using various methods

Model Name	**PLSR**	**Ridge Regression**
RMSE	2.87	1.02

4. Conclusion

In this article we studied the application of Ridge Regression Technique for analysis of urea in the soil for better productivity of the crops. In past we had done such analysis using ParLeS (which is propriety and not an open source software). The results obtained were encouraging with errors as low as 1.02mg/15ml. Therefore, we conclude here that Ridge Technique is far superior to the traditional technique of regression analysis.

References

[1] Folorunso, O.; Ojo, O.; Busari, M.; Adebayo, M.; Joshua, A.; Folorunso, D.; Ugwunna, C.O.; Olabanjo, O.; Olabanjo, O. 2023. Exploring Machine Learning Models for Soil Nutrient Properties Prediction: A Systematic Review. Big Data Cogn. Comput. 7, 113. DOI: 10.3390/bdcc7020113

[2] Andreas Kamilaris , Francesc X. Prenafeta-Boldú. 2018. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90.DOI: 10.1016/j.compag.2018.02.016.

[3] Jain, S., Sethia, D. 2023. A Review on Applications of Artificial Intelligence for Identifying Soil Nutrients. In: Saini, M.K., Goel, N., Shekhawat, H.S., Mauri, J.L., Singh, D. (eds) Agriculture-Centric Computation. ICA 2023. Communications in Computer and Information Science, vol 1866. Springer, Cham. DOI: 10.1007/978-3-031-43605-5_6

[4] Liu, Y., Ma, X., Shu, L., Hancke, G. P., & Abu-Mahfouz, A. M. 2020. From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges. IEEE Transactions on Industrial Informatics, 1–1. DOI:10.1109/tii.2020.3003910

[5] Shaikh, F.K.; Memon, M.A.; Mahoto, N.A.; Zeadally, S.; Nebhen, J. 2021. Artificial intelligence best practices in smart agriculture. IEEE Micro , 42, 17–24.

[6] Chen, Q.; Li, L.; Chong, C.; Wang, X. AI-enhanced soil management and smart farming. Soil Use and Management. 2022, 38, 7–13. DOI: 10.1111/sum.12771

[7] Dobos, E. 2006. Digital Soil Mapping: As a Support to Production of Functional Maps; Office for Official Publication of the European Communities: Luxembourg

[8] Wadoux, A.M.C.; Minasny, B.; McBratney, A.B. 2020. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. Earth-Sci. Rev. , 210, 103359

[9] Dong, W.; Wu, T.; Sun, Y.; Luo, J. 2018. Digital mapping of soil available phosphorus supported by AI technology for precision agriculture. In Proceedings of the 2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics), Hangzhou, China, pp. 1–5.

[10] Khaledian, Y.; Miller, B.A. Selecting appropriate machine learning methods for digital soil mapping. Appl. Math. Model. 2020, 81, 401–418.

[11] Shahare, Y.; Gautam, V. Soil Nutrient Assessment and Crop Estimation with Machine Learning Method: A Survey. In Cyber Intelligence and Information Retrieval; Springer: Berlin/Heidelberg, Germany, 2022; pp. 253–266.

[12] Trontelj ml., J.; Chambers, O. Machine Learning Strategy for Soil Nutrients Prediction Using Spectroscopic Method. Sensors 2021, 21, 4208. DOI:10.3390/ s21124208

[13] Umm E Farwa, Ahsan Ur Rehman, Qasim Khan, S. ., & Khurram, M. (2020). Prediction of Soil Macronutrients Using Machine Learning Algorithm. International Journal of Computer (IJC), 38(1), 1–14.

[14] S. R. Vernekar, I. A. P. Nazareth, J. S. Parab and G. M. Naik, "RF spectroscopy technique for soil nutrient analysis," 2015 International Conference on Technologies for Sustainable Development (ICTSD), Mumbai, 2015, pp. 1-4. doi: 10.1109/ICTSD.2015.7095878

[15] Raphael A. Viscarra Rossel, “ParLeS: Software for chemometric analysis of spectroscopic data”, Chemometrics and Intelligent Laboratory Systems, Volume 90, Issue 1, 2008, Pages 72-83, ISSN 0169-7439, https://doi.org/10.1016/j.chemolab.2007.06.006.

[16] Jacob Murel, Eda Kavlakoglu, “What is ridge regression?”, URL: www.ibm.com/topics/ridge-regression.

[17] “What is Ridge Regression?” URL: www.ibm.com/topics/ridge-regression.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-28] Crime Investigation Using Lip Reading

Mansimran Kaur [1, *, †], Dhruv Rastogi [2,†] Anjali Sharma [3,†], Pratham Mahajan[4,†], Anushka Dahiya [5,†], Preeti Nagrath[6,†]

[1,2,3,4,5,6] Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Abstract

The use and development of lip reading techniques have been revolutionized through the incorporation of deep learning architectures and visual speech analysis which has greatly helped crime investigations whereby investigators can understand conversations using surveillance videos without sound. This capability has been found to be very useful in solving many cases with a lot of complications. In contrast to prior studies using the Grid Corpus dataset that relied on known models such as LipNet that have had high accuracy, we aim to train a new lip-reading model from scratch using a mix of Gated Recurrent Unit (GRU) networks and Convolutional Neural Networks (CNN). This innovative methodology has made 86.17% of accuracy in the given problem. The conclusion that can be drawn from this work points to the significance of this project in playing a positive role in criminal investigations, specifically within the areas of enhanced lip reading analysis using deep learning technology and helping the law enforcement agencies in improving their ability to understand visual speech from security cameras.

Keywords

Lip-reading, Deep Learning, Convolutional Neural Networks, GRU, Crime Investigation, Forensic Analysis, Evidence Analysis

1. Introduction

Computer vision and language understanding technology have for instance been used in lip reading to enhance the interaction between man and computer. This technology has brought a great change in the lives of the people with hearing impairments since it offers them a device that enhances better means of communicating. Besides its usage in personal life, lip reading also brings additional enhancements to security systems and the quality of surveillance and evidence. Also, it has a significant function in the creation of the assistive technologies which assist in filling communication gaps in different areas, including police-community-police interaction, patient-doctor interaction, and so on; thereby making such interactions more effective and efficient.

* Corresponding author.
† These authors contributed equally.
mansimran2703@gmail.com (Mansimran Kaur); dhruvrastogi43@gmail.com (Dhruv Rastogi); anjali2609@gmail.com (Anjali Sharma); prathammahajan103@gmail.com (Pratham Mahajan); dahiya0502@gmail.com (Anushka Dahiya); preeti.nagrath@bharatividyapeeth.edu (Preeti Nagrath);

0000-0002-5863-3113 (Preeti Nagrath)

Lip reading is important not only in facilitating communication but also useful for people with hearing impairments. It lets them understand what has been said through oral communication by the gestures, lip movements, facial expressions, etc. Thus, in the conditions of acoustic interference or when all audible signals are distorted, lip reading acts as useful signal in addition to auditory information. It also leads to better interaction with people as the clients can be made to attend to conversations better than before. Apart from its use to the deaf and the hard of hearing, lip reading is important in areas such as law enforcement and security, where speech without sound is important in several practical scenarios.

The progress in the last few years has been significantly boosted by the large datasets such as the LRW (Lip Reading in the Wild) and the Grid Corpus. Specifically, the research activities related to the LRW dataset have been boosted in the past years. Many researches have shown good performances and high hit rates, for instance, LipNet model that has been proven to be efficient in recognizing basic sentence by detecting lips movement with an accuracy of more than 95%. These achievements highlight a lot of progress that has been made in the application of existing paradigms for the accurate interpretation of visual speech gestures.

However, our approach to solving the problems in this area is significantly different from the typical practice of using ready-made models. The main aim of this study is to investigate the dynamic relationship between visual speech input and the factors associated with sentence prediction. As such, we have started with the challenging goal of building a new lip-reading model from the ground up. This new model is designed based on both Convolutional Neural Networks (CNNs) and Bidirectional Gated Recurrent Unit (GRU) networks, so that we can explore the variations in the architecture and layers to better understand the working of lip reading.

In this paper, the major emphasis is placed on the Grid Corpus dataset, and while our goal goes beyond simply comprehending the nuances of visual speech, we strive to decipher the intricacies of the most advanced sentence construction. Over the course of this project, we have been able to hit an accuracy rate of 86.17%. This success is due to our efforts to respond to the peculiarities of the Grid Corpus dataset, with a keen emphasis on the nuances that matter for lip reading.

By building a new model tailored to the unique characteristics of this dataset, we aspire to enhance the capabilities of lip-reading systems and further the understanding of visual speech processing. Our research not only highlights the potential for improved accuracy in lip-reading applications but also sets the stage for future advancements in this exciting area of study. As we continue to delve into the nuances of visual speech recognition, we hope to contribute to the broader goal of developing more robust and reliable communication technologies.

Lip-reading technologies have the potential of improving the investigation of crime processes through the following reasons. When the video recording has no clear audio, or the audio is missing, low quality, or has been altered by someone with a malicious intent, or contains background noise, these technologies come handy. Lip-reading systems also work by detecting vocal gestures and kinematics involving mouth movements, facial expressions, and gestures to translate spoken words

and discover important information that may be lost. This capability is especially useful in situations where it is impossible to take an audio capture, for instance, in espionage or important protection surveillance. In other words, lip-reading technology provides law enforcement with a tool to get valuable dialogue from mute videos that in turn helps them to assemble the evidence, comprehend the perpetrators' motives, and solve crimes more efficiently. As the technology progresses, this integration into investigations could change how LEA's deal with cases where traditional audio doesn't suffice.

The research work of this paper is intended to solve one of the biggest problems in the area of criminal investigation to help make a change by improving the efficiency of the police. To this end, we have carefully developed a novel model that employs deep learning architectures, including both CNN and GRU models. The aim of this work is to identify speech from the video data that is captured in surveillance videos where audio may be missing or of poor quality.

This revolutionary method has the potential of greatly boosting the efficiency of investigation in police work since visual data, which is usually silent, can be turned into useful information. As a result of translating visual signals into comprehensible language, our model can reveal important information, find suspects, and provide other information that is essential for solving crimes. The significance of this study is thus significant to law enforcement organisations as it provides a state-of-art solution that can identify key conversational and exchange occurrences in video footage and enhance crime solving. By way of this work, we hope to provide investigators with the means to do their work better and, in so doing, bring more safety and security into society.

2. Related Work

This section gives an outline of all the different forms of automatic lip reading that are in practice in the present times. In the past, there has been limited application of deep learning methods to automated lip reading; most of the previous work employed different strategies. Such approaches were often accompanied by the need to extract image features from frames or video features from video sequences. Some of the previous techniques used were optical flow, movement detection and other manually designed vision pipelines [7], [8], [9], [10], [11], [13].
Due to the vast amount of literature available on automated lip reading, it is impossible to discuss all the areas in detail in this context. For the more detailed information, the literature review of the current state of art in lip reading technologies will help to understand the development and the current state of the technologies.

Goldschen was the first to work on visual sentence-level speech recognition using HMMs on a small hand segmented phones dataset [1]. This was an important step in beginning to investigate the possibility of visual speech. On this basis, Neti further extended the work by developing the first audio-visual speech recognition system that operates at the sentence level using HMMs with specific features that were trained on the IBM ViaVoice dataset [2]. Their revolutionary work combined both, visual and audio data which enhanced the speech recognition in noisy environment. However, it is necessary to point out that the dataset used in this research, including 17,111 utterances from 261 people, or approximately 34.9 hours, is not publicly accessible. Another

important point is that their visual-only results do not contain individual visual recognition outputs, but are used to rescore noisy audio-only predictions.

Potamianos continued this work, obtaining WER of 82.31% and 91.62 % for speaker adapted and speaker independent models respectively using the same IBM ViaVoice dataset [3]. For speaker adapted models, on the connected DIGIT corpus – a dataset consisting of phrases with digits, he got WERs of 38.53% and for speaker independent models, WERs of 16.77%. These improvements demonstrate the advancements in the methodologies used for AVSR, where specific advancements have been made in speaker adapted and speaker independent cases across different datasets.

Furthermore, Gergen used a speaker-dependent training with the GMM/HMM system on the mouth regions only after applying LDA on the DCTs. This approach resulted in outstanding speaker-dependent accuracy of 86.4% which set a new benchmark for GRID corpus dataset. However, some issues were still present, such as the ability to generalize performance to other speakers and in extracting motion features [16]. To overcome these limitations, LipNet was considered as the best solution that provides improvements in the generalization of the speaker and in the extraction of motion related features in lip reading tasks.

In recent years, deep learning has gained increased attention in the lip reading area, but most of the work has still been centered on word or phoneme recognition. LipNet, however, is different from other methods as it goes further to predict the whole phrases in the sequence rather than predicting single words. Current approaches in this area tend to focus on learning multimodal audio-visual features [18] [20] [21] and incorporating video into conventional speech recognition models including GMM-HMM and HMM for word and character recognition [24] [26]. Sometimes, these approaches use one or more than one technique [4].

Malek Miled and his research team document a holistic algorithm for lip-reading that adequately combines advanced image processing methods with deep learning methodologies. With this end, the study advocates an innovative hybrid model in which an edge-based filter is utilized to separate the mouth region, thus enhancing the accuracy of lip movement detection. The combination of CNN with Bi-GRU ends up in a robust model that is highly sensitive to the spatio-temporal mechanisms of lip movement. The algorithm thus achieved an excellent accuracy of 90.38% in testing, which indicates significance improvement in efficiency compared with traditional methods. This research not only pushes forward the domain of lip reading but has also opened up potential applications in silent communication as well as speech recognition technologies [34].

Mini-3DCvT is a newly designed lipreading technique focusing on the complexity needed to adequately extract visual spatial characteristics, temporal dynamics, and at the same time maintain a lightweight model structure [33]. The technique combines visual transformers with 3D convolution for apt capturing of spatiotemporal local and global attributes in a sequence of continuous images. Weight transformation and distillation strategies come into play within the architectures of both convolution and transformer that make the model more streamlined to drastically improve its efficiency. The method manifests itself in a high recognition accuracy,

scoring 88.3% on the LRW dataset and 57.1% on LRW-1000, with good computational complexity and a minimal number of parameters.

Co-adaptation of feature detectors in neural networks is a problem that Hinton et al. address in their paper. They propose a method known as dropout which during the training phase, certain neurons are switched off at random. This approach also helps the network to create better feature representations in a way that it doesn't overly depend on one neuron. The authors provide convincing proof of dropout improving accuracy for different architectures and tasks of neural networks. Furthermore, they explain the theory behind dropout and, as a result, improve the understanding of how neural networks work and become popular in deep learning areas such as computer vision and natural language processing [27].

In their paper, Gergen et al. introduce a dynamic stream weighting technique for turbo-decoding-based audiovisual automatic speech recognition (ASR) [25]. This method enhances the integration of audio and visual information, allowing for more accurate speech recognition in challenging environments. They demonstrate that adapting the weighting of audio and visual streams dynamically can improve performance significantly. Meanwhile, Haliassos et al. (2020) present a robust approach to face forgery detection, emphasizing that "lips don't lie." Their work focuses on developing a generalizable model capable of accurately identifying manipulated facial features, contributing to advancements in security and integrity verification in visual media [26].

In another paper published in 2024, Robin Anburaj B contributes to the field of lip reading by introducing a vision-based system that effectively combines a convolutional neural network (CNN) with an attention-based Long Short-Term Memory (LSTM) architecture. By leveraging pre-trained CNN models, the study enhances feature extraction from processed video frames, which are crucial for understanding temporal characteristics of lip movements. The system achieves an impressive 80% accuracy using TensorFlow and ensemble learning techniques, demonstrating its potential for practical applications. Furthermore, the research highlights the importance of integrating machine learning with visual speech analysis to advance communication accessibility, particularly for individuals with hearing impairments [28].

This paper contributes a lot towards lip reading since it filled up a gap of similar such Turkish-language datasets [30]. For that reason, video data were recorded from 72 different people pronouncing 71 different words involving audio streams as much as possible to the greatest extent visual information was focused. The replication was done through the Camtasia application for increasing the dataset size and diversity. The percentages used to test this proposed model are as follows: adjectives with 71.8%, nouns with 71.88%, and verbs with 79.69%. This work is of help to fill gaps in Turkish lip-reading resource while achieving an enhancement on communication aids for the hearing impaired.

Chung and Zisserman were able to contribute by using spatial and spatiotemporal convolutional neural networks based on the VGG architecture for classifying words [5]. The authors in their

studies, on a dataset from BBC TV with 333 and 500 classes, showed that ST models are less accurate than S models by an average of 14%. Furthermore, their models were restricted in modeling sequences of different lengths, and they did not consider the sequence prediction at the sentence level, which could be explored in more detail in other tasks.

In a different approach, Garg used a pre-trained VGG model that targets faces to classify words and phrases, although their study was carried out on a small MIRACL-VC1 dataset, which has only 10 words and expressions each [6]. Garg's strongest recurrent model used a training method that froze the VGGNet parameters while training RNN, a method that deviated from joint training methods used in other models. Even when working with a very simple dataset and the classification tasks were restricted to only 10 classes, the authors' model obtained fairly low, yet still reasonable, accuracies of 56.0% for word classification and 44.5% for phrase classification.

In contrast to these efforts, LipNet represents a major breakthrough in optical speech recognition by offering a fully end-to-end model capable of predicting sequences at the sentence level. LipNet's approach is distinct in that it generates sequences of tokens directly from an input series of images, eliminating the need for explicit alignments. This is achieved through its use of Connectionist Temporal Classification (CTC) during training, which allows the model to learn sequences without requiring precise frame-by-frame labeling. LipNet's end-to-end structure marks a significant advancement in the field, showcasing the potential for more robust, comprehensive, and context-aware visual speech recognition systems.

A novel paper makes important contributions to the literature of lip reading because it proposes a novel deep learning model that maps directly a video sequence of lip movements into text transcriptions [29]. The system works pretty well using this end-to-end architecture combining 3D convolutional neural networks with bidirectional Long Short-Term Memory networks, effectively interpreting visual cues within motions. For benchmark datasets, the model shows remarkable performance with character error rate being 1.54% and the word error rate being 7.96%. Such headways not only offer the hearing-impaired more precise lip-reading technologies but also promote accessibility for them to have unobstructed communication in challenging auditory environments.

A 2024 paper contributes quite a lot to lip-reading research as the first large-scale Korean dataset on lip reading is incorporated, which comprises over 120,000 utterances derived from diverse TV broadcasts, including news, documentaries, and dramas [31]. The article is designed to help fill the gap in existing resources for Korean lip reading, which had previously been substantially underexplored compared to English. However, the authors suggest a strong preprocessing method to extract a consistent region of interest for facial parts and introduce a transformer-based model concentrated on grapheme units for efficient sentence-level analysis. Experimental results validate

the effectiveness of this dataset and model and hence pave the way for further possible developments in Korean lip-reading technology.

Another recent research makes major contributions in the area of lip reading by proposing a comprehensive Cantonese sentence-level lip-reading dataset with over 500 unique speakers and more than 30,000 samples [32]. It especially appeals to the relatively low number of Cantonese datasets compared to the mushrooming Mandarin ones. The research boasts a rich pipeline of dataset collection and construction, including a new visual frontend: the 3D-visual attention net, a combination of convolutional and self-attention mechanisms for the detailed representation of lip-region features. Coupled with an effective backend conformer in modeling temporal sequences, this laid foundation for highly valuable future research into dialect-specific lip reading.

Analysis of some publicly available datasets, such as LRW, OuluVS, CUAVE, and SSSD, is conducted in the study [35]. Advanced deep learning models are elaborately examined for lip reading at the word level. Observations made on various state-of-art architectures during this study resulted in achieving new accuracy while lip-reading, significantly on the LRW dataset, a surprise from 66.1% to 94.1%. The conducted research combines well-established models, using which the effectiveness of ResNet, WideResNet, EfficientNet, MS-TCN, and ViViT were improved by using alternative modified variants of feature extractors and classifiers. According to the results, for feature extraction, settings of 3D-Conv + ResNet18 as well as the MS-TCN model selection for inference enable generalization over various datasets and lead to better performance in tasks of lip reading.

The evaluation of LipNet leverages the GRID corpus, chosen for its sentence-level structure and large dataset. The phrases within this corpus follow a well-defined grammatical pattern, consisting of six distinct word categories: command (4 options), color (4 options), preposition (4 options), letter (25 options), digit (10 options), and adverb (4 options). Each category contains a specific set of possible words, such as {bin, lay, place, set} for commands, {at, by, in, with} for prepositions, and {blue, green, red, white} for colors. The letter category ranges from A to Z, with W included separately, while digits span from zero to nine, and the adverbs are drawn from {again, now, please, soon}. This structured combination results in a total of 64,000 possible sentence configurations. Example sentences from the dataset include statements like "place red at C zero again" and "set blue by A four please," illustrating the variety and complexity of potential phrases.

3. Methodology

The main objective of this study is to improve lip reading through the use of deep learning approach with the use of Conv3D and GRU networks.

The study uses GRID Corpus dataset not only because it is available but also because of the variety of speakers and the variety of lighting conditions. These characteristics make it ideal for the research objective which is to construct a generalized lip-reading model that is effective in all conditions.

In the case of video data, pre-processing is done in the following steps to ensure that the input data is ready for model training. First, frames are captured from the video stream then the frames are converted to the right color space. The face region of interest, particularly the lip region, is then extracted from each frame, and then the illumination is corrected for any variation. The frames are also uniformly resized to a standard resolution. Next, label encoding is performed on the dataset followed by the tokenization and converting into numerical form which is suitable for the model output layer format.

The proposed model architecture integrates several key components: Conv3D layers to learn spatial and temporal features, Activation layers to inject non-linearity and MaxPooling3D layers to down sample the feature maps. Time Distributed and Flatten layers are used to pass the data for sequential model and Bidirectional GRU layers are employed to capture both past and future contexts in lip movements. To reduce overfitting, Dropout layers are used and the final layer is an output layer with Dense layer to give the final predictions. All of them are important in the process of making the final decision on lip movements and the corresponding sentence in order to have the best performance of the model.

This combination of methods and the systematic pre-processing of the GRID corpus is intended to generate a sound and general lip-reading model. role in capturing spatial and temporal features crucial for lip reading. The Connectionist Temporal Classification (CTC) loss function is applied with the Adam optimizer and 0.01 as the learning rate to train the model, and early stopping mechanisms to prevent overfitting. The dataset is divided into training and validation sets, besides accuracy is chosen as the evaluation metric.

This research addresses the gap in lip-reading methodologies by proposing a deep learning model that combines Conv3D and GRU networks. The study employs the Grid Corpus dataset for its diversity and aligns with ethical considerations.

3.1 Data Collection:

This work also benefits from the Grid Corpus created by Oxford University as a substitute for the datasets that are not available to the public, including LR2 and LRW. This feature-rich dataset contains a diverse set of linguistic material, which includes speakers with various accents and speaking styles, which is particularly important for training effective lip-reading models.

Furthermore, the Grid Corpus is intended to cover different lighting conditions, so the model will be able to learn lip movements under different visual situations. Notably, the dataset is accompanied

by fine-grained annotations associating spoken words with lip movements, which are instrumental in training and testing models intended for analyzing VSC information. In this way, the Grid Corpus greatly improves the functionality of the automated lip reading tools for researchers and developers in the field, and thus helps to advance the use of lip reading in assisting technologies and criminal investigations.

Data	Audio	Front Video	Side Video	Alignment	Meta Data
Size (MB)	651.4	837.1	870	2	0.062

Table 1:
Structure of Grid Corpus Dataset

3.2 Video Pre-processing:

Only frames are extracted evenly from the video sequences, which synchronously and effectively samples the dynamic lip movements necessary for lip reading. Once extracted, these frames are converted to grayscale since the analysis does not require all the colors, only the necessary information in order to carry out the analysis. Next, lip regions are extracted by face landmark detection which locates a set of important facial landmarks and allows the model to concentrate on the areas of interest only.

In addition to reduce the effect of illumination on the model, lighting normalization is performed on the dataset. This step assists in making the dataset more uniform since issues such as variations in lighting will be minimized from the actual video clips. Further, the frames are scaled to a fixed size to standardize the format of the data set and to ease the computational processing during the training of the models. All these preprocessing steps are very important in the preprocessing of the dataset and fine tuning of the dataset to enhance the efficiency of training and enhancing the lip-reading model.

3.3 Label Encoding:

Since text processing is always pronounced on text, it is very important to divide spoken phrases in preparation for processing, by converting these phrases into handy tokens which are referred to as text tokens. Such approach can let the model analyze the components of the speech separately, which, in turn, makes the task easier.
After tokenization, a mapping from characters to numbers takes place to reduce these tokens to numbers. This encoding process maps each of them to a specific integer so that the model can understand and comprehend it well.

The adoption of numerical representations is critical due to the fact that the result of most machine learning models, specifically the output layer, frequently employs numerical data for processing. First and foremost, encoding helps to define a better mapping from characters to numbers which helps the learning step better in finding patterns and relations in the data It also creates a solid base for the further work of the model, which can predict the desired accuracy based on the set of encoded spoken phrases.

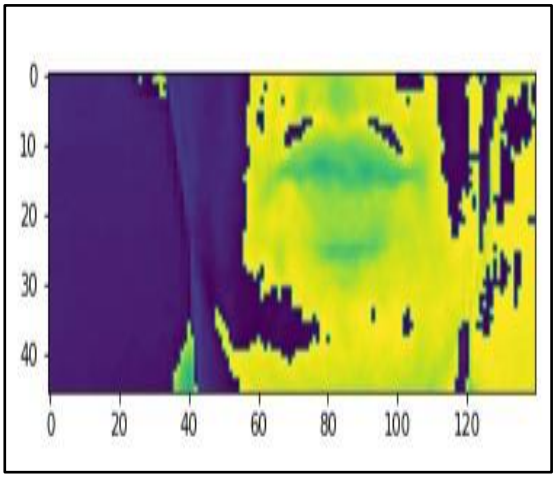


Figure 1: Grey-scaled and normalized frame extracted from video

3.4 Model Architecture:

The proposed architecture integrates Conv3D layers for spatial-temporal features, Activation layers for non-linearity, MaxPooling3D layers for down sampling, Time Distributed and Flatten layers for temporal processing, Bidirectional GRU layers for sequential modelling, Dropout layers for regularization, and a Dense output layer for predictions.

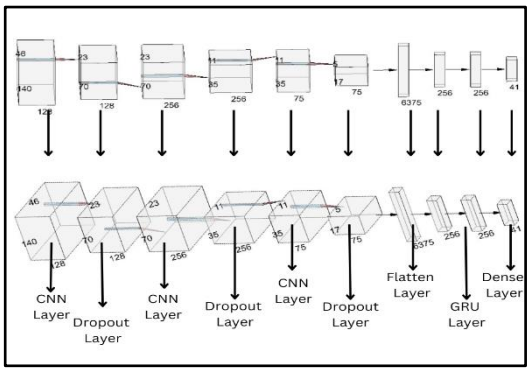


Figure 2: Lip Reading Model Architecture

3.5 Model Training:

The Grid Corpus data set is divided systematically by the `tf.data` module of TensorFlow into training and validation. In training, it engages 10-epoch training with the training set and further configures several important parameters aiming to optimize performance. Specifically, it is evident that the loss function Connectionist Temporal Classification aptly fits its usage on sequence-based tasks like lip reading due to its flexibility with respect to aligning input sequences and output labels. As for optimization, the Adam optimizer is also used with a learning rate of 0.01, which is a good compromise between convergence speed and stability.

The mechanisms for early stopping added in the training procedure monitor how the model is performing and stops when improvements become stagnant. This is an efficient technique for preventing overfitting, whereby the model would generalize well for data yet to be seen. Generally, these strategies will make a potent training framework that will contribute favorably to boost the ability of the model in learning and predicting visual speech patterns based on the Grid Corpus dataset.

4. Result and Analysis

4.1. Result

The Connectionist Temporal Classification (CTC) loss function used plays a crucial role in the training process of the introduced model. This particular loss function is specifically tailored to cope with the difficulties of the task, which is the sequence based tasks like lip reading. Another advantage of using CTC is wanting in flexibility in managing different output lengths, which is essential in lip reading since the duration by which words are spoken may not necessarily correspond to the frame found in the video. CTC helps the model to learn the alignments between the input sequences, which are the frames of lip movement video and the related target labels while the timing for each phoneme does not need to be annotated accurately. This flexibility is especially helpful in those cases where it is challenging to provide accurate labels for speech data.

After the training process carried out for 10 epochs, the desirable accuracy of the model was reached and it constituted 86.17%. This outstanding performance goes on confirming the efficiency of the architecture in perceiving the spatial and temporal dynamics related to lip movements. The capacity to perceive these subtleties is important so that the real picture of the information contained in speech can be seen. The CTC loss function was particularly significant in this success since it allowed the model to match predictions with target label values.

CTC effectively deals with the problems connected with the temporal relation of phoneme occurrences, which in natural speech may occur at different time points. The fact is that traditional methods that rely on the strict correspondence of time points to certain events may fail in this case, while the CTC approach enables the model to deal with the shift between the lip movements and phoneme production. CTC in a way improves the generalization of the model by allowing the model to predict the most probable sequence of outputs given the sequence of inputs, hence the lip-reading

predictions will be accurate even if the speaking style, speed etc. changes. The effectiveness of the model at the level of accuracy is quite gratifying, yielding a high rate of 86.17%. The reason for such success of the network can be explained by the very structure of the network, which is a combination of convolutional layers and recurrent layers. This architectural design allows the model to learn local spatial features which are necessary for identifying lip motion within a particular temporal frame and, at the same time, learn long range temporal dependencies which are also important for recognizing lip movements.

Also, it is necessary to draw attention to the fact that the convolutional layers play the most essential part in the model for feature extraction. These layers perform well in their role by applying successive convolutional filters on the input matrix which makes it easier to extract features from the video frames. Consequently, they can gather details such as the shape of lips and their movements during the entire video, which will help the model to distinguish small peculiarities connected with definite phonemes. This capability greatly improves the model's performance and is a major contribution to the efficiency of the lip-reading process. With this, it is easier for the model to learn and interpret visual features that are essential for speech recognition.

On the other hand, Gated Recurrent Unit (GRU) makes the model more resistant because prediction is done using the previous and the subsequent frames. This capability is especially relevant when the information may be sporadic or intermittent as is the case with lips movement of the speaker as well as variations that may be noted with time. Therefore, the GRU improves the temporal information from both directions and expands the scope of where the lip movements occur. This temporal awareness is helpful for the sequential aspects related to visual speech in a way that even in unfavorable conditions, better prediction is provided.
In totality, the proposed model is a comprehensive and efficient model that incorporates CNNs and GRUs for the purpose of extracting local spatial features and for managing long-term dependencies, respectively. This dual use is beneficial to the model's operation in various circumstances and indicates a high level of resilience. This shows that the architecture is viable in practice where the observer may not have all the information necessary to perform the task as this study has shown that it is capable of performing well even under these circumstances will without doubt make it a worthwhile tool not just in lip reading but in many other related applications. With further advancements in research and development, this model could potentially lead to even more significant improvements in visual speech recognition systems.

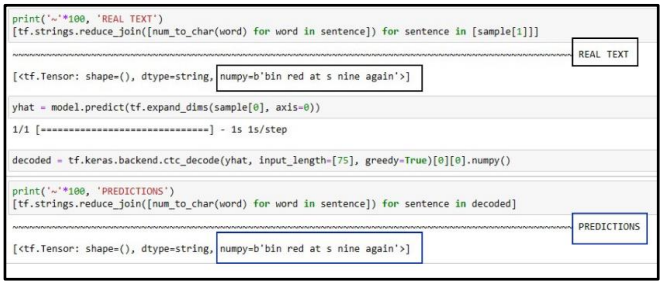


Figure 4: Predicted Result of the Model

Table 2: Performance of the model compared to BiLSTM

Criteria	BiLSTM	GRU
Accuracy	75%	86.17%
Validation Accuracy	34%	77.5%
Loss	0.69	0.5

4.2 Analysis and Future Prospects

The visual speech model presented in this paper is a stack of 3D convolutional and bidirectional GRU layers, thereby boosting the results using the CTC loss function. The results produce a magnificent accuracy of 86.17% in predicting visual speech cues.

5. Conclusion

A review of the literature has been done where more focus has been placed on the importance of lip reading systems in enhancing crime investigation. This paper emphasises that the above systems can significantly contribute to improving the investigative process.

The combination of CNNs and GRUs has been vital to the creation of a vigorous deep learning model for lip reading. All these architectures play an independent part in enhancing the accuracy and efficiency of the given model. CNNs are especially useful for the lip image to extract the facial features which help the model to understand and learn the complicated visual patterns and details implicitly linked with speech. On the other hand, GRU networks perform well in capturing temporal dependencies that are crucial for lip reading since the model learns the temporal dependencies of lips movements.

This framework using the best of both CNNs and GRUs provides a good starting point to future development of lip reading. This integration not only improves the credibility of the model but also paves way for further advancements in the relation between human and computer interaction and especially in the field of accessibility. Therefore, this research creates a foundation for new developments that can enhance knowledge sharing and reduce misunderstanding in different fields.

The findings of this project prove that the CNN-GRU model can detect visual speech through lip reading with considerable accuracy across the various data sets. The high performance of this model can be considered as its applicability in real-world scenarios, especially in crime investigation to collect a large amount of evidence. It also has potential for the purpose of/communication between the hearing impaired people and provides them with a tool to improve interaction in their

environment. However, this paper has demonstrated some significant improvements in lip reading technology although the researcher wishes to acknowledge that there is so much more that has not been explored yet in this field.

To sum up, the relevance of deep learning in the field of lip reading – with CNNs and GRUs in particular — means enormous potential in the sphere of changing communication systems. Indeed, as we progress further in this area, future research and development will not only improve the efficiency of the current models but also increase the number of possible uses. Its focus will help shape a better world in which more people have equal opportunities due to elimination and or reduction in barriers with emphasis on communication, and the general use of technology that acts as a key to unearthing commonalities of the disabled to the rest of society.

Acknowledgements

This research would not have been possible without the unwavering support and encouragement of numerous individuals. A deep appreciation is extended to Dr. Preeti Nagrath and Dr. Dharmender Saini, whose expertise and thoughtful guidance have been invaluable throughout this entire journey. Their insightful suggestions, constructive feedback, and steadfast support played a crucial role in shaping the direction and quality of this research, making it a truly enriching experience.

Gratitude is also owed to the colleagues at Bharati Vidyapeeth's College of Engineering, whose stimulating discussions and collaborative efforts greatly enriched the research process. Their diverse perspectives and shared knowledge fostered an environment of innovation and creativity that contributed significantly to the outcomes of this study. Special thanks are extended to [specific individuals, if applicable] for their technical assistance and generous advice, which were instrumental in navigating various challenges encountered during the research.

Finally, heartfelt thanks are offered to family and friends for their constant understanding and encouragement. Their unwavering support and belief in this endeavor provided invaluable motivation throughout this journey. This research is not just a culmination of academic effort but also a reflection of the collective support and inspiration received from all those involved. Their contributions have made this achievement possible, and for that, deep appreciation is expressed.

References

[1] J. Goldschen, O. N. Garcia, and E. D. Petajan, "Continuous automatic speech recognition by lipreading," in Motion-Based recognition, pp. 321–343, Springer, 1997.

[2] C. Neti, G. Potamianos, J. Luettin et al., "Audio visual speech recognition," Technical report, IDIAP, 2000.

[3] G. Potamianos, C. Neti, G. Gravier et al., "Recent advances in the automatic recognition of audiovisual speech," Proceedings of the IEEE, vol. 91, no. 9, pp. 1306–1326, 2003.

[4] H. Ninomiya, N. Kitaoka, S. Tamura et al., "Integration of deep bottleneck features for audio-visual speech recognition," in International Speech Communication Association, 2015.

[5] J. S. Chung and A. Zisserman, "Lip reading in the wild," in Asian Conference on Computer Vision, 2016a.
[6] A. Garg, J. Noyola, and S. Bagadia, "Lip reading using CNN and GRU," Technical report, Stanford University, CS231n project report, 2016.
[7] Matthews, G., Zeidner, M., & Roberts, R. D. Boston "Review. Emotional intelligence: Science and myth", 2002
[8] G. Zhao, M. Barnard, and M. Pietikäinen, "Lipreading With Local Spatiotemporal Descriptors," IEEE TRANSACTIONS ON MULTIMEDIA, 2009.
[9] M. Gurban and J. -P. Thiran, "Information Theoretic Feature Extraction for Audio-Visual Speech Recognition," in IEEE Transactions on Signal Processing, Dec. 2009.
[10] G. Papandreou, A. Katsamanis, V. Pitsikalis et al., "Adaptive multimodal fusion by uncertainty compensation with application to audiovisual speech recognition," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 3, pp. 423–435, 2009.
[11] V. Pitsikalis, A. Katsamanis, G. Papandreou., "Adaptive multimodal fusion by uncertainty compensation," in Interspeech, 2006.
[12] Z. Zhou, G. Zhao, X. Hong , "A review of recent advances in visual speech decoding," in Image and vision computing, vol. 32, no. 9, pp. 590–605, 2014.
[13] P. Lucey, S. Lucey, and S. Sridharan, "Using a Free-Parts Representation for Visual Speech Recognition," in Digital Image Computing: Techniques and Applications (DICTA'05), Queensland, Australia, 2005.
[14] Goldschen, Alan & Garcia, Oscar & Petajan, Eric.. "Continuous Automatic Speech Recognition by Lipreading", 1997.
[15] J. S. Chung and A. Zisserman, "Out of time: automated lip sync in the wild," in Workshop on Multi-view Lip-reading, ACCV, 2016b.
[16] Z. Zhou, G. Zhao, X. Hong et al., "A review of recent advances in visual speech decoding," in Image and Vision Computing, vol. 32, no. 9, pp. 590–605, 2014.
[17] C. Feichtenhofer, A. Pinz, and A. Zisserman, "Convolutional two-stream network fusion for video action recognition," in Proc. CVPR, 2016.
[18] J. Ngiam, A. Khosla, M. Kim et al., "Multimodal Deep Learning," in Proceedings of the 28th International Conference on Machine Learning, ICML 2011, pp. 689-696, 2011.
[19] Y. Assael, B. Shillingford, S. Whiteson et al., "LipNet: Sentence-level Lipreading," 2016.
[20] S. Petridis and M. Pantic, "Deep complementary bottleneck features for visual speech recognition," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 2016, pp. 2304-2308, doi: 10.1109/ICASSP.2016.7472088, 2016.
[21] Y. Fu, S. Yan, and T. S. Huang, "Classification and feature extraction by simplexization," in IEEE Transactions on Information Forensics and Security, vol. 3, no. 1, pp. 91–100, 2008.
[22] I. Almajai, S. Cox, R. Harvey et al., "Improved Speaker Independent Lipreading using Speaker Adaptive Training and Deep Neural Networks," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016.

[23] D. Hu, X. Li et al., "Temporal multimodal learning in audiovisual speech recognition," in IEEE Conference on Computer Vision and Pattern Recognition, pp. 3574–3582, 2016.

[24] A. Takashima, I. Bakker, J. G. van Hell et al., "Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words," in Brain Lang., vol. 167, pp. 44-60, 2017.

[25] S. Gergen, S. Zeiler, A. H. Abdelaziz et al., "Dynamic stream weighting for turbo-decoding-based audiovisual ASR," in Interspeech, pp. 2135–2139, 2016.

[26] A. Haliassos, K. Vougioukas, S. Petridis, and M. Pantic, "Lips Don't Lie: A Generalisable and Robust Approach to Face Forgery Detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[27] Hinton, Geoffrey & Srivastava, Nitish & Krizhevsky, Alex & Sutskever, Ilya & Salakhutdinov, Ruslan. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint. arXiv, 2019.

[28] B, Robin. (2024). Lip Reading using Deep Learning. INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT. 08. 1-5. 10.55041/IJSREM36177.

[29] Bs, Prashanth & M V, Manoj & Puneetha, B. & Lohith, R & Gowda, V & Chandan, V & Sneha, H.. (2024). Lip Reading with 3D Convolutional and Bidirectional LSTM Networks on the GRID Corpus. 1-8. 10.1109/NMITCON62075.2024.10699241.

[30] Pourmousa, Hadi & Özen, Üstün. (2024). Lip reading using deep learning in Turkish language. IAES International Journal of Artificial Intelligence (IJ-AI). 13. 3250. 10.11591/ijai.v13.i3.pp3250-3261.

[31] Cho, Sunyoung & Yoon, Soosung. (2024). Korean Lip-Reading: Data Construction and Sentence-Level Lip-Reading. Journal of the Korea Institute of Military Science and Technology. 27. 167-176. 10.9766/KIMST.2024.27.2.167.

[32] Xiao, Yewei & Liu, Xuanming & Teng, Lianwei & Zhu, Aosu & Tian, Picheng & Huang, Jian. (2024). Cantonese sentence dataset for lip-reading. IET Image Processing. 18. n/a-n/a. 10.1049/ipr2.13123.

[33] Wang, Huijuan & Cui, Boyan & Yuan, Quanbo & Pu, Gangqiang & Liu, Xueli & Zhu, Jie. (2024). Mini-3DCvT: a lightweight lip-reading method based on 3D convolution visual transformer. The Visual Computer. 1-13. 10.1007/s00371-024-03515-y.

[34] Miled, Malek & Messaoud, Mohammed & Bouzid, Aicha. (2022). Lip reading of words with lip segmentation and deep learning. Multimedia Tools and Applications. 82. 1-21. 10.1007/s11042-022-13321-0.

[35] Arakane, Taiki & Saitoh, Takeshi. (2023). Efficient DNN Model for Word Lip-Reading. Algorithms. 16. 269. 10.3390/a16060269.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-29] Smart Learning Ecosystems: The Influence of Online Skill-Based Education on Employability Outcomes

Nimisha Bansal[1], Dr. Garima Sainger[2]

[1] *Institute of Business Management, G.L.A University, Mathura, India*
[2] *Institute of Business Management, G.L.A University, Mathura, India*

Abstract
The report analyzes how online courses elevate skill proficiency in smart learning platforms and affect job placement outcomes. With the rise of skill demands in a fluctuating work environment online learning is essential for developing essential job skills. This study examines the way online skill-driven education improves skills by offering flexibility and accessibility and boosting job prospects. The study shows that disparities in program standards and employer perceptions of online education qualifications are challenges. This analysis provides recommendations to teachers and learning organizations on how to improve virtual education. The conclusions point out that educational organizations must cooperate, and require the implementation of strict guidelines for fair use and effective teaching strategies in skill courses. By so doing, people can acquire relevant skills that are important in enhancing employment opportunities through smart educational frameworks.
Keywords: Online Courses, Skill Proficiency, Smart Learning Platforms, Job Placement Outcomes, Virtual Education, Skill-Driven Education, Equitable Access

1. Introduction

In this decade technology has advanced rapidly turning instructions and learning locations into Smart Learning Environments. The cooperating strategies and innovative technologies with the learner need to create an interesting, personalized education process (Cai et al., 2020). The challenges of current employment require new and integrated systems which are important in addressing the issues of interdisciplinary workers.

An Overview of Smart Learning Ecosystems

Smart Learning Environments can personalize and meet industry and learner needs by using data analysis and teamwork support. These networks often revolve around electronic platforms that enhance knowledge sharing and encourage active learning delivering timely comments to foster improved content engagement (Freigang *et al.*, 2018). Using big data analytics and artificial intelligence improves the enjoyment and value of learning (Visvizi *et al.*, 2023).

nimisha1411@gmail.com(Nimisha Bansal); Garima.sainger@gla.ac.in(Garima Sainger)

As the labor market changes dramatically SLEs offer critical education specifically designed for industry demands. These platforms offer significant and rapid learning prospects that are progressively vital as established learning approaches have difficulty aligning with advancements and changing vocational roles (Hoel *et al*, 2018).

Importance of Online Skill-based education

Online skill-based education is popular because it meets diverse learning needs with its easy accessibility and adaptability. Through different formats such as virtual simulations and interactive tutorials, learners obtain practical skills in this education approach (Bansode & P, 2016). As remote work and distance education increase due to the COVID-19 outbreak, a constant requirement for personal skill enhancement emerges. Skill-oriented online learning effectively fills the persistent skill deficiencies present in numerous fields (Bozkurt *et al*., 2020). By highlighting practical applications and specific skills these programs supply learners with the resources to fulfill employer expectations. The connection of schools to industry expectations is important since evidence reveals that hiring managers place a premium on competent individuals instead of formal degrees (Rakowska *et al*., 2021). People who participate in online skill-oriented education usually find it easier to land jobs and achieve success in their professions.

Objectives of the Review

We intend to assess the role of skill-oriented learning provided through web-based resources in Smart Learning Ecosystems in shaping job prospects. Specifically, the objectives are as follows:

1. Our goal is to review the design of Smart Learning Ecosystems and its relevance for virtual learning.
2. Evaluate how online learning benefits and challenges job skill improvement.
3. The project centers on revealing the connection between online learning for skills enhancement and workforce performance while listing major contributors to effective job attainment.

2. Conceptual Framework

The concept of Smart Learning Ecosystems presents a systems view of their functioning and impact on learning and career prospects. In the current world of learning, SLEs combine traditional practices with technology and direct attention to competencies. Smart Learning Ecosystems is a unique learning environment that is integrated and versatile learning spaces that use technology to improve learning experiences and achievement (Stavropoulos et al., 2021). These systems allow the integration of different types of assets including digital resources and analytics to foster a core learning space. The greatest aim of SLEs is to ensure there is continuous learning and realization of skills development in order to meet the emerging challenges in the workplace (Jeladze et al., 2017). SLEs understand the need to work with teachers, business partners as well as students. Ample cooperation ensures that learning opportunities are both large and fun also due to resource sharing (Merdian et al. 2019).

Significant Elements of Smart Learning Systems

The main components of Smart Learning Ecosystems include:

[1] Key Components of Smart Learning Ecosystems (SLEs)	Description
Learner-Centric Design	Concentration on aligning what students desire and require to create unique learning directions based on their strengths and ambitions. (Martins *et al*., 2020).
Technology Integration	By merging technologies such as LMSs and mobile applications with AI capabilities we optimize learning environments (Rienties *et al*., 2023).
Collaboration and Networking	Enhancing support among students educators and industry leaders via web-based systems discussions and cooperative projects that boost the exchange of ideas and collective solutions. (Tabuenca *et al*., 2021).
Data-Driven Insights	Evaluation of learners' progress occurs through analytics which yields targeted advice and action to boost skills (Arifah *et al*., 2024).
Continuous Learning and Adaptability	Support for lifelong learning by delivering training materials and activities that can be modified promptly to keep participants informed in an adapting job market. (Uskov *et al*., 2020).

Role of Smart Learning in Technology

The effectiveness of Smart Learning Ecosystems heavily relies on technology. It functions as the essential support for developing stimulating and adaptive learning activities. Key aspects of technology's role include:

Adaptive Learning Technologies	Description
Adaptive Learning Technologies	Systems that observe how users behave and score modify learning tools and help foster the perfect mix of challenge and support (Silva *et al*., 2013).
Virtual and Augmented Reality	Create authentic experiences that help students gain skills in a safe space (Mohapatra *et al*., 2019).
Cloud Computing	Enables learners and educators to access resources anytime and anywhere, as these platforms easily store and distribute learning materials (Polin *et al*., 2023).
Artificial Intelligence	AI-developed tools offer fast assessment and measure how students perform while highlighting problematic areas so learning can improve effectively (Akintayo *et al*., 2024).
Collaboration Tools	By using video and sharing content material students and educators create teamwork and engage in conversations which enhance learning in a community setting (Markkula *et al*., 2015).

3. Online Skill-Based Education

The current learning system also encompasses online learning based on skills that provide the relevant knowledge required for individuals to succeed in the current world employment market. Since various digital platforms meet user needs for skill acquisition, this kind of education benefits learners and employers.

Definition and Characteristics

Practical skills learning programs occur on the internet and focus on the improvement of skills appropriate for certain sectors. This learning style focuses on the application of knowledge that is acquired from modeling and emulation as well as the packaging of skills (Fajaryati et al., 2020).

Key characteristics of online skill-based education include:

Feature	Description
Flexibility	Individuals can pick their study schedules that facilitate their online learning alongside work and individual obligations (Stone *et al*., 2019).
Personalization	In digital learning platforms, adaptive methods change content delivery and speed according to learners' specific circumstances (Salman *et al*., 2024).
Interactivity	Using Quizzes Forums And Team Work Improves Both Engagement And Retention In Online Skill Learning (Sun *et al*., 2013).
Industry-Relevant Curriculum	To align with the current job market demands courses receive input from skilled advisors (Sistermans & I.J., 2020).

Different Online Skill-Based Programs

Skill-based courses offered online are available in various formats to address various sectors and learnings. Some common types include:

Type of Online Skill-Based Program	Description
Certification Programs	Courses made for immediate learning of vital skills in certain sectors leading to a certificate (Atienza & J.B., 2024).
Massive Open Online Courses (MOOCs)	Universities and organizations offer courses to many people for little to no cost while instructing a mix of skills (Matsuda *et al.*, 2015).
Professional Development Workshops	Short courses focus on improving particular skills for workers while fitting in with hectic timetables (Reaves & J, 2019).
Micro-credentials	In-depth teachings produce recognized patches for improving useful skills to enhance career prospects (Hunt *et al.*, 2020).
Bootcamps	Programs that deliver swift and in-depth education in major technical skills including programming and data science (Williams *et al.*, 2021).

Advantages of Online Skill-Based Education

Online skill-based education offers several advantages, including:

[1] Benefit	Description
Accessibility	Provides access to superior learning content for individuals from different regions and experiences without being tied to their location (Heron *et al.*, 2013).
Cost-Effectiveness	By offering online programs instead of physical classes learners save money and can acquire valuable skills without the hefty costs of travel and housing (Maloney *et al.*, 2015).
Self-Paced Learning	Users reach the course information at a velocity that aligns with their learning styles (Offutt *et al.*, 2017).
Immediate Feedback	Different digital platforms deliver rapid replies on assessments and quizzes that assist students in recognizing their shortcomings and changing their learning practices (Belmar *et al.*, 2023).
Networking Opportunities	Through virtual courses, participants have the opportunity to collaborate with other learners and industry specialists which benefits their career paths (Rutten *et al.*, 2016).

Challenges and Limitations
Despite its advantages, online skill-based education also faces several challenges and limitations:

[1] Challenge	Description
Limited Hands-On Experience	Though virtual education features simulations it frequently doesn't supply essential hands-on education for particular fields (Schwichow *et al.*, 2016).
Digital Divide	Lack of dependable technology and robust internet often prevents some students from success which intensifies gaps in education (Van Deursen *et al.*, 2011).
Quality Variability	Online courses have increased in number but varied in quality creating issues for students to pinpoint safe and efficient programs. (Chow & J.Y., 2013).
Self-Motivation Requirements	In the absence of rigid class structures, online courses might cause students to find it hard to develop and sustain motivation and self-discipline. (Kebritchi *et al.*, 2017).
Perception Issues	Some employers continue to value classic degrees more than their digital counterparts modifying the views on online learning in job opportunities. (Hew *et al.*, 2014).

4. Employability Outcomes

The assessment of employability acts as an important measure of readiness for individuals in securing employment. This chapter looks into the definition of employability and reviews the significance of skill-oriented training for elevating employability results.

Definition of Employability
The term employability indicates the ability of people to find and keep a job by being able to accommodate modern employment demands and satisfy employer expectations. It includes various skills along with knowledge and traits that help people effectively manage their career paths. Finding job outcomes in employability is just one aspect; it entails securing meaningful employment; achieving career growth; and ensuring long-term career continuity as shown in Figure 1.

Employability can be understood through several dimensions:

[1] Employability Factor	Description
Skills and Competencies	Businesses across various sectors require specialized technical and interpersonal skills (Mahajan *et al.*, 2022).
Experience	Practical skills gained from internships and previous employment enhance a candidate's appeal (Mahajan *et al.*, 2022).
Personal Attributes	Qualities such as flexibility and resilience contribute to an individual's professional success (Byrne *et al.*, 2022)

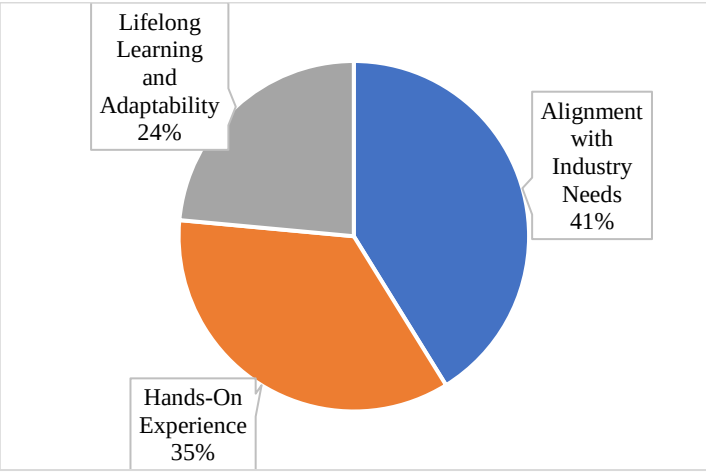


Fig 1. Employability factors

Figure 1 shows the various factors that play a role in determining employability as shown on the pie chart and each factor has its influence percentage. The key contributor is Integral to Market Trends, comprising 41% of the entire impact which demonstrates the significant need for academic efforts to align with what businesses require. A 35% portion of the Make Waves initiative shows how Hands-On Experience adds value to skill applications in actual settings. Furthermore, skills enhancement and versatility gain 24 percent representation and indicate that employees needed to adjust their skills for survival in a flexible labor market. When combined these factors demonstrate the varied aspects of employability and stress the necessity for education programs to include them effectively.

Relevance of skill training

Must haves in job performance are gained through skill-based programs that increase employability as shown in Figure 2. The importance of this education can be highlighted through several key points:

[1] Impact on Employability	Description
Alignment with Industry Needs	Skill-based programs are designed with input from industry professionals to ensure that taught skills are practical and in demand, helping graduates meet employer requirements and improve hiring prospects (Mardis *et al.*, 2018).
Hands-On Experience	Many skill-oriented programs include practical tasks that allow learners to apply their knowledge in real-world contexts, strengthening their skill sets and providing concrete examples for potential employers (Wettaka & J, 2020)
Lifelong Learning and Adaptability	Skill-based education fosters a mindset of continuous learning, helping individuals stay current with new developments and adapt to changing industry demands in a dynamic job market (POPOVSKI *et al.*, 2013)
Improved Confidence and Job Readiness	Tailored training and hands-on experiences in skill-driven education boost learners' confidence and enhance their readiness for the workforce (Ristiani *et al.*, 2022)

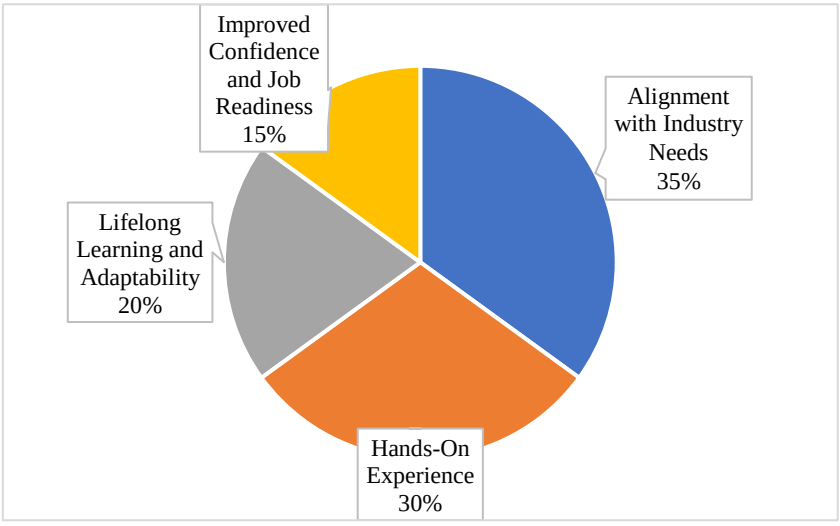


Fig 2. Relevance of skill training

Figure 2 reveals the effect of different factors on the preparation for employment. At 35%, industry expectations are crucial for education programs when preparing graduates for employment. Enhancing vocational skills through direct training reaches 30% and highlights whether they should be experienced in real applications. Moreover, learning throughout life and flexibility add 20%, demonstrating the need for people to refresh their skills for job relevancy. Additionally, 15% views improved confidence and employment preparedness as key elements that help graduates feel more

certain in their jobs. The combined elements emphasize the necessity for an integrated method of education that satisfies business requirements and builds hands-on skills.

Measuring Employability Outcomes

Evaluated results on employability are necessary to analyze the success of educational offerings and their effect on professional achievements as shown in Figure 3. Common methods for assessing employability outcomes include:

[1] Method for Measuring Employability Outcomes	**Description**
Surveys and Questionnaires	Educational institutions use surveys to assess how well their programs prepare graduates for their current jobs (Glerum *et al.*, 2021)
Longitudinal Studies	Ongoing studies tracking graduates provide insights into their career progression and the long-term impact of skill-based education on employability (Di Fabio & A, 2017)
Employer Feedback	Feedback from employers regarding the skills of new hires helps institutions assess how well their programs meet industry needs (Neroorkar & S, 2022)
Performance Metrics	Institutions analyze job placement rates and salary data to evaluate the effectiveness of their skill-based training programs (Neroorkar & S, 2022)

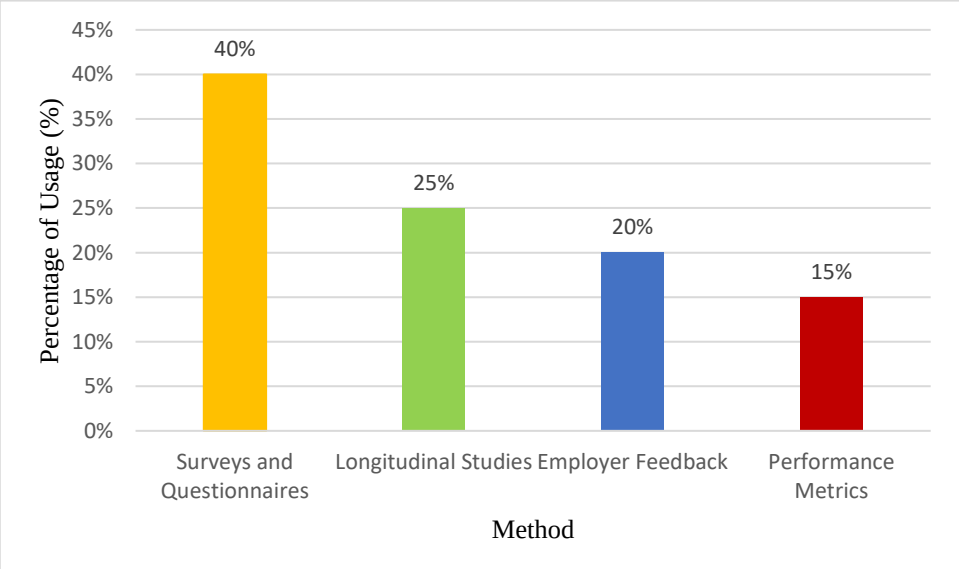


Fig 3. Method for Measuring Employability Outcomes

Figure 3 displays techniques for evaluating job readiness outcomes and stresses their significance in regulatory assessments. At 40%, post-surveys and questionnaires are the most frequently used instruments to measure how programs equip graduates for jobs. In 25% of cases, longitudinal studies reveal insights into job advancement and the sustained consequences of education. At 20%, feedback on worker abilities comes from employers. By using these methods together institutions optimize their programs to meet market demands.

Influencers of Career Success

Several factors influence employability outcomes, including:

[1] Factor Influencing Employability	**Description**
Educational Quality	The effectiveness of teaching methods and the alignment of courses with industry demand influence the acquisition of employment-related skills (Goh *et al*., 2021)
Networking Opportunities	Establishing connections within the industry improves job prospects and provides valuable industry insights (Heine & V, 2020)
Personal Attributes	Attributes such as motivation, communication, and adaptability significantly impact employability, as employers seek candidates who can manage workplace challenges effectively (Cordeiro *et al*., 2017)
Market Conditions	Economic conditions and demand for specific skills affect job opportunities, with more openings available during prosperous times compared to economic downturns (Ho *et al*., 2023)
Experience and Internships	Gaining practical experience through internships or volunteer work enhances employability by demonstrating the ability to apply skills in real-world settings (Chia & Z.Q, 2022)

The Influence of Online Skill-Based Education on Employability

Online skill-oriented education greatly affects job readiness outcomes and influences career opportunities for students positively and negatively. This segment analyzes the good and bad results of online training while presenting a comprehensive outlook on its impact on job readiness in the current market.

Positive Impacts

Many positive aspects in connection with employability are discovered via online skill-based education. Thus, online skill-based education is also effective in enhancing contemporary learned skills that are suitable to different industries. Skills are generally taught in courses in a manner that reflects actual job requirements and enables students to develop ideal abilities that fit employers'

requirements. Coursera and edX offer educational programs that offer industry-recognized certifications and allow employers to view students' skills. Regarding the skill-based education that is delivered online, industry specialists often collaborate with educators to update the content, which be as relevant and valuable as possible (Mittal et al., 2021). By so doing, learners acquire important information about the current practices and trends within the industry hence making them more marketable to employers. Most of the courses are project-based, which gives the learners a chance to solve real problems and gain experience that is desirable by employers (Manjushree et al., 2021).

Online skill-based education is very flexible and accessible which in turn increases employment opportunities. Students can engage in programs anywhere they wish which is suitable for balancing between studies, work, and household chores. This opportunity enables a lot of people to benefit from learning irrespective of their geographical location or ability to pay for the services.

Negative Impacts

Opportunity diversity in employment is limited by the numerous problems that accompany online skill-based education. An important issue that can be associated with the use of distance education in skills is the absence of practical experience. Though many programs focus on theoretical comprehension and skill development they often lack practical experiences necessary for some positions. Variation in practice can limit the learners from showcasing their skills in real-life situations implying that it becomes more challenging for them to compete with traditional hands-on trainees. Online education might not be more appreciated by employers resulting in some hindrance to new graduates looking for employment. Some companies prefer conventional diplomas from recognized schools to digital qualifications that these groups consider inferior or not genuine. This prejudice may adversely affect candidates who finished skill-oriented online courses (Stephany *et al*., 2022).

Online education's fast advancement has produced considerable disparities in the quality of programs available. Different online courses exist with various standards; thus students could have difficulty recognizing reliable programs that deliver proficient education and significant credentials. Mediocre programs can weaken the relevance of the skills gained and might yield job results for grads which leads to suspicion about the entire success of online education in enhancing job opportunities (Okeleke & P.K, 2017). While online learning for skills enhances career opportunities through skill acquisition and practical training it confronts issues related to empirical experience and evaluations from employers. Students and educators need to acknowledge these effects along with employers while they journey through the developing field of skill-focused education (Hassan *et al*., 2020).

5. Emerging future directions and recommendations

As education develops further online skill-focused education will significantly impact how employability is affected. This section describes new trends and policy effects along with advice for educators and institutions.

Trends in Online Skill-Based Education

Several trends are shaping the future of online skill-based education:

[1] Future Trends in Online Skill-Based Education	Description
Increased Personalization	Advancements in AI and machine learning are enabling more personalized learning experiences, with tailored learning paths that cater to individual needs and learning styles (Tavakoli & M, 2023)
Integration of Soft Skills	While technical skills remain crucial, there is a growing emphasis on developing soft skills such as communication and teamwork, which are becoming a focus in online education alongside technical competencies (Tan *et al.*, 2021)
Hybrid Learning Models	The blend of digital and traditional classroom instruction is becoming more popular, allowing students to combine the flexibility of online learning with the benefits of hands-on experiences in a physical environment (Mishra *et al.*, 2023).
Micro-Credentials and Badges	The rising popularity of micro-credentials offers students recognition for specific skills or competencies, providing employers with a clearer understanding of a job applicant's skillset and improving employability (Gurjar *et al.*, 2023)

Policy Implications

As online skill-based education continues to expand, several policy implications must be considered:

Policy Implication	Description
Quality Assurance Standards	Policymakers must establish robust quality control measures to ensure online educational programs meet high standards, reducing variability in quality and enhancing the credibility of online qualifications (Protopsaltis *et al.*, 2019).
Equitable Access	Educational authorities must address the digital divide by investing in technology, providing financial aid, and implementing digital education initiatives to ensure equal access to online skill-based training (Bashay & M, 2020)
Collaboration with Industry	Schools should collaborate with industry partners to align educational programs with current job market demands, enhancing graduates' employability (McKee *et al.*, 2020)

Recommendations for Educators and Institutions

To maximize the effectiveness of online skill-based education in improving employability, educators and institutions should consider the following recommendations:

Recommendation	Description
Emphasize Practical Learning	Best practices indicate that integrating real-world projects and simulations into online programs enhances learners' practical knowledge (Boettcher *et al*., 2021)
Focus on Continuous Improvement	Institutions should regularly assess and update their programs based on feedback from learners and changes in industry standards to ensure relevance and quality (Isa *et al*., 2024)
Foster Lifelong Learning	Encourage a commitment to lifelong education by offering diverse opportunities for professional development and skill enhancement beyond foundational schooling (Lim *et al*., 2024)

Conclusion

Employability results are significantly affected by online skill-based education in the following ways. This section gathers the main findings and provides final opinions on the impact of innovative learning environments on employability. In this review, some of the findings that have been made on online skill-based education and employability outcomes have been presented. Online training in skills enhances the amount of knowledge that learners acquire and provides training that is relevant to the employers Online training encourages different types of participation in learning and promotes equal opportunities for all learners. Some of the comments remain key challenges such as lack of relevant working experience, and employment prejudices which hamper the achievement of positive employment outcomes.

Smart learning ecosystems can be regarded as promising to enhance the quality of online learning and improve employment outcomes. Technology and organization of teachers and industry make it easier for us to provide for enhanced and diverse learning environment for learners in preparation for the competitive job market world. With the growing need for skilled workers on the horizon advancing the quality and availability of online learning will be essential for strengthening individuals and refining tomorrow's workforce.

References

[1] Arifah, N., Bafadal, I., & Sumarsono, R. B. (2024). Change Leadership in The Development of Digital Learning Ecosystem: A Case Study at Brawijaya Smart High School. *Management of Education: Jurnal Manajemen Pendidikan Islam*, *10*(2).

[2] Atienza, J. B. (2024). Assessment of Online Teaching for Skill Based-Topics in Online Physical Education Classes. *Community and Social Development Journal*, *25*(3), 58-69.

[3] Bansode, P. (2016). Education and Skill Development: A Systematic Review. *Technology*, *9*(48).

[4] Bashay, M. (2020). Digital Skills for an Equitable Recovery: Policy Recommendations to Address the Digital Skill Needs of Workers Most Vulnerable to Displacement. *National Skills Coalition*.

[5] Belmar, F., Gaete, M. I., Durán, V., Chelebifski, S., Jarry, C., Ortiz, C., ... & Varas, J. (2023). Taking advantage of asynchronous digital feedback: development of an at-home basic suture skills training program for undergraduate medical students that facilitates skills retention. *Global Surgical Education-Journal of the Association for Surgical Education*, *2*(1), 32.

[6] Boettcher, J. V., & Conrad, R. M. (2021). *The online teaching survival guide: Simple and practical pedagogical tips*. John Wiley & Sons.

[7] Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., ... & Paskevicius, M. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education*, *15*(1), 1-126.

[8] Byrne, C. (2022). What determines perceived graduate employability? Exploring the effects of personal characteristics, academic achievements, and graduate skills in a survey experiment. *Studies in Higher Education*, *47*(1), 159-176.

[9] Cai, Y., Ma, J., & Chen, Q. (2020). Higher education in innovation ecosystems. *Sustainability*, *12*(11), 4376.

[10] Chia, Z. Q. (2022). *Influencers of trainee satisfaction during an internship in organizations in Malaysia* (Doctoral dissertation, UTAR).

[11] Chow, J. Y. (2013). Nonlinear learning underpinning pedagogy: evidence, challenges, and implications. *Quest*, *65*(4), 469-484.

[12] Cordeiro, H. T. D., & Albuquerque, L. G. D. (2017). Career profiles of Generation Y and their potential influencers. *BAR-Brazilian Administration Review*, *14*, e170013.

[13] Di Fabio, A. (2017). A review of empirical studies on employability and measures of employability. *Psychology of career adaptability, employability, and resilience*, 107-123.

[14] Fajaryati, N., Budiyono, Akhyar, M., & Wiranto. (2020). The employability skills needed to face the demands of work in the future: Systematic literature reviews. *Open Engineering*, *10*(1), 595-603.

[15] Freigang, S., Schlenker, L., & Köhler, T. (2018). A conceptual framework for designing smart learning environments. *Smart Learning Environments*, *5*, 1-17.

[16] Glerum, D. R., & Judge, T. A. (2021). Advancing employability: applying training evaluation to employability development programs. *Career Development International*, *26*(3), 363-390.

[17] Gurjar, P., Ringo, J., Waller, M., & Gilliland, S. (2023). Skill Badging: Using Micro-Credentials to Improve Holistic Student Development and Career Readiness. *College Teaching*, 1-9.

[18] Hassan, M. U., Iqbal, Z., & Shakir, K. (2020). Impact of ICT training and education on women's employability and entrepreneurial skills: achieving sustainable development goals. *International Journal of Business Forecasting and Marketing Intelligence*, *6*(2), 157-166.

[19] Heine, V. (2020). Exploring social media influencers' career construction: an inductive inquiry.

[20] Heron, M., Hanson, V. L., & Ricketts, I. (2013). Open source and accessibility: advantages and limitations. *Journal of Interaction Science*, *1*, 1-10.

[21] Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational research review*, *12*, 45-58.

[22] Ho, C., Stewart-Smith, C., & Gunaratne, D. (2023). How career influencers can promote sustainable careers and the well-being of underrepresented students. In *Handbook of research on sustainable career ecosystems for university students and graduates* (pp. 407-427). IGI Global.

[23] Hoel, T., & Mason, J. (2018). Standards for smart education–towards a development framework. *Smart Learning Environments*, *5*(1), 1-25.

[24] Hunt, T., Carter, R., Zhang, L., & Yang, S. (2020). Micro-credentials: The potential of personalized professional development. *Development and Learning in Organizations: An International Journal*, *34*(2), 33-35.

[25] Isa, M., Neliwati, N., & Hadijaya, Y. (2024). Quality Improvement Management in Teacher Professional Development. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, *5*(2), 136-147.

[26] Jeladze, E., Pata, K., & Quaicoe, J. S. (2017). Factors Determining Digital Learning Ecosystem Smartness in Schools. *IxD&A*, *35*, 32-55.

[27] Kebritchi, M., Lipschuetz, A., & Santiague, L. (2017). Issues and challenges for teaching successful online courses in higher education: A literature review. *Journal of Educational Technology Systems*, *46*(1), 4-29.

[28] Lim, Z. Y., Yap, J. H., Lai, J. W., Mokhtar, I. A., Yeo, D. J., & Cheong, K. H. (2024). Advancing Lifelong Learning in the Digital Age: A Narrative Review of Singapore's SkillsFuture Programme. *Social Sciences*, *13*(2), 73.

[29] Mahajan, R., Gupta, P., & Misra, R. (2022). Employability skills framework: a tripartite approach. *Education+ Training*, *64*(3), 360-379.

[30] Maloney, S., Nicklen, P., Rivers, G., Foo, J., Ooi, Y. Y., Reeves, S., ... & Ilic, D. (2015). A cost-effectiveness analysis of blended versus face-to-face delivery of evidence-based medicine to medical students. *Journal of medical Internet research*, *17*(7), e4346.

[31] Manjushree, D. L., Varsha, T. L., Arvind, W. K., & Laxman, D. N. (2021). Performance analysis of the impact of technical skills on employability. *International Journal of Performability Engineering*, *17*(4), 371.

[32] Mardis, M. A., Ma, J., Jones, F. R., Ambavarapu, C. R., Kelleher, H. M., Spears, L. I., & McClure, C. R. (2018). Assessing alignment between information technology educational opportunities, professional requirements, and industry demands. *Education and Information Technologies*, *23*, 1547-1584.

[33] Markkula, M., & Kune, H. (2015). Making smart regions smarter: smart specialization and the role of universities in regional innovation ecosystems. *Technology Innovation Management Review*, *5*(10).

[34] Martins, V. F., Tomczyk, Ł., Amato, C., Eliseo, M. A., Oyelere, S. S., Akyar, Ö. Y., ... & Silveira, I. F. (2020). A smart ecosystem for learning and inclusion: an architectural overview. In *Computational Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part I 20* (pp. 601-616). Springer International Publishing.

[35] Matsuda, N., Furukawa, T., Bier, N., & Faloutsos, C. (2015). Machine Beats Experts: Automatic Discovery of Skill Models for Data-Driven Online Course Refinement. *International Educational Data Mining Society*.

[36] McKee, S., & Gauch, D. (2020). Implications of Industry 4.0 on skills development. *Anticipating and preparing for emerging skills and jobs: Key issues, concerns, and prospects*, 279-288.

[37] Merdian, I., Tullius, G., Hertkorn, P., & Burgert, O. (2019). Interactive learning in smart learning ecosystems. In *The Interplay of Data, Technology, Place, and People for Smart Learning: Proceedings of the 3rd International Conference on Smart Learning Ecosystems and Regional Development 3* (pp. 15-22). Springer International Publishing.

[38] Mishra, J., & Dholakia, K. (2023). The Future of Hybrid Learning Models and Sustainable Education in the Post-Pandemic Era. In *Sustainable Practices in Higher Education: Finance, Strategy, and Engagement* (pp. 53-70). Cham: Springer International Publishing.

[39] Mittal, P., & Raghuvaran, S. (2021). Entrepreneurship education and employability skills: the mediating role of e-learning courses. *Entrepreneurship Education*, *4*(2), 153-167.

[40] Mohapatra, M., Das, S. C., Patnaik, B., & Satpathy, I. (2019). Role of information technology in education and skill-based learning for employability. *International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN*, 2278-3075.

[41] Neroorkar, S. (2022). A systematic review of measures of employability. *Education+ Training*, *64*(6), 844-867.

[42] Offutt, J., Ammann, P., Dobolyi, K., Kauffmann, C., Lester, J., Praphamontripong, U., ... & White, L. (2017, April). A novel self-paced model for teaching programming. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale* (pp. 177-180).

[43] Okeleke, P. K. (2017). *Impact of vocational skills acquisition on self-employability readiness and potentials among unemployed graduates in Delta State, Nigeria* (Doctoral dissertation, University of Lagos (Nigeria)).

[44] Polin, K., Yigitcanlar, T., Limb, M., & Washington, T. (2023). The making of smart campus: A review and conceptual framework. *Buildings*, *13*(4), 891.

[45] POPOVSKI, V., POPOVSKA, M. B., & TOPUZOVSKA, M. (2013). LIFELONG LEARNING FOR BEST ADAPTATION TO THE 21. CENTURY CHANGES IN THE WORLD OF WORK. *Economic Development,(1-2)*, 58-70.

[46] Protopsaltis, S., & Baum, S. (2019). Does online education live up to its promise? A look at the evidence and implications for federal policy. *Center for Educational Policy Evaluation*, 1-50.

[47] Rakowska, A., & de Juana-Espinosa, S. (2021). Ready for the future? Employability skills and competencies in the twenty-first century: The view of international experts. *Human Systems Management*, *40*(5), 669-684.

[48] Reaves, J. (2019). 21st-century skills and the fourth industrial revolution: a critical future role for online education. *International Journal on Innovations in Online Education*, *3*(1).

[49] Rienties, B., Divjak, B., Iniesto, F., Pažur Aničić, K., & Žižak, M. (2023). Online work-based learning: A systematic literature review. *International review of education*, *69*(4), 551-570.

[50] Ristiani, R., & Lusianingrum, F. P. W. (2022). The effect of self-confidence on job readiness. *Journal of Applied Business, Taxation and Economics Research*, *1*(5), 450-461.

[51] Rutten, M., Ros, A., Kuijpers, M., & Kreijns, K. (2016). The usefulness of social network sites for adolescents' development of online career skills. *Journal of Educational Technology & Society*, *19*(4), 140-150.

[52] Salman, O., Khasawneh, Y., Alqudah, H., Alwaely, S., & Khasawneh, M. (2024). Tailoring gamification to individual learners: A study on personalization variables for skill enhancement. *International Journal of Data and Network Science*, *8*(2), 789-796.

[53] Schwichow, M., Zimmerman, C., Croker, S., & Härtig, H. (2016). What students learn from hands-on activities. *Journal of research in science teaching*, *53*(7), 980-1002.

[54] Silva, A. P., Lourtie, P., & Aires, L. (2013). Employability in online higher education: A case study. *International Review of Research in Open and Distributed Learning*, *14*(1), 106-125.

[54] Sistermans, I. J. (2020). Integrating competency-based education with a case-based or problem-based learning approach in online health sciences. *Asia Pacific Education Review*, *21*(4), 683-696.

[55] Stavropoulos, P., Korfiati, A. M., Panopoulos, N., & Mourtzis, D. (2021, June). A Smart Learning Model targeting regional SME ecosystems: a conceptual framework. In *Proceedings of the Conference on Learning Factories (CLF)*.

[56] Stephany, F. A. B. I. A. N. (2022). Closing the Digital Skill Gap: The Potential of Online Platform Data For Active Labour Markets Policies. *Zenodo*.

[57] Stephany, F. A. B. I. A. N. (2022). Closing the Digital Skill Gap: The Potential of Online Platform Data For Active Labour Markets Policies. *Zenodo*.

[58] Stone, C., Freeman, E., Dyment, J. E., Muir, T., & Milthorpe, N. (2019). Equal or equitable?: The role of flexibility within online education. *Australian and International Journal of Rural Education*, *29*(2), 26-40.

[59] Sun, J. N., & Hsu, Y. C. (2013). Effect of interactivity on learner perceptions in Web-based instruction. *Computers in Human Behavior*, *29*(1), 171-184.

[60] Tabuenca, B., Serrano-Iglesias, S., Martin, A. C., Villa-Torrano, C., Dimitriadis, Y., Asensio-Pérez, J. I., ... & Kloos, C. D. (2021). Affordances and core functions of smart learning environments: A systematic literature review. *IEEE Transactions on Learning Technologies*, *14*(2), 129-145.

[61] Tan, C. Y., Abdullah, A. G. K., & Ali, A. J. (2021). Soft Skill Integration for Inspiring Critical Employability Skills in Private Higher Education. *Eurasian Journal of Educational Research*, *92*, 23-39.

[62] Tavakoli, M. (2023). Hybrid human-AI driven open personalized education.

[63] Uskov, V. L., Howlett, R. J., & Jain, L. C. (Eds.). (2020). *Smart Education and E-learning 2021* (Vol. 188). Singapore: Springer.

[64] Van Deursen, A., & Van Dijk, J. (2011). Internet skills and the digital divide. *New media & society*, *13*(6), 893-911.

[65] Visvizi, A., Troisi, O., & Grimaldi, M. (2023). Exploring the drivers for digital transformation in smart education: An ecosystems approach. In *ITM Web of Conferences* (Vol. 51, p. 01003). EDP Sciences.

[66] Wettaka, J. (2020). *Improving hands-on training for skill acquisition of engineering Students at the Department of Mechanical and production engineering, Kyambogo University* (Doctoral dissertation, Kyambogo University (unpublished work)).

[67] Williams, J., Newton, B., Allen, A., Lanceley, L., Garner, O., Cook, J., ... & Clark, L. (2021). Skills Bootcamps process evaluation.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-10] Intelligent Web system for the management of pre-professional practices in a public university in Lima-Peru, using fuzzy logic in a Smart City enviroment*

Carlos Claros-Vasquez[1,†], Romero Espinoza-Bustamante[1,†] and Wilver Auccahuasi[1,†*]

[1] *Universidad Continental, Huancayo, Perú*

Abstract

In the university educational processes, one of the final processes is the completion of professional internships, which allow the student to carry out activities specific to his/her specialty. At the end of the procedure, a report is sent which is evaluated by 3 professors. The evaluations are mostly carried out qualitatively, which does not allow a final result for the student. The present research is carried out in order to analyze the qualitative results and transform them into a quantitative value, related to the approval or disapproval of the report, for which the fuzzy logic technique is used, with which we obtain a final grade. The proposal also includes a web system that can support the processes of student registration, entry of documents and final results, in such a way that they can be accessed without having to be on the university campus. The documentary processing processes are based on the ISO 25000 standard that ensures compliance with the standards of the University and the degrees and titles office, which is in charge of issuing conformity in the completion of professional internships. The proposal can be replicated and scaled to other processes related to the use of fuzzy logic.

Keywords

Smart Cities, university, process, standard, management, practice, management, practices.

3rd International Symposium on Smart Cities Challenges, Technologies and Trends (SCCTT-2024)

* Corresponding author. Wilver Auccahuasi; wauccahuasi@continental.edu.pe
† These authors contributed equally.
cclaros@continental.edu.pe (Carlos Claros-Vasquez); romero Espinoza-Bustamente@continental.edu.pe (Romero Espinoza-Bustamante); wauccahuasi@continental.edu.pe (Wilver Auccahuasi)
0000-0001-8820-4013 (Wilver Auccahuasi)

1. Introduction

Managing smart city infrastructures in multi-tenant environments presents challenges in identifying problems and dysfunctions, as well as in integrating data and services accessible to different stakeholders. A unified knowledge model was proposed that allows smart city operators to manage and analyze data more efficiently, facilitating scenario creation and informed decision making, using data analytics processes, microservices, "what-if" analysis tools, and data storage to integrate contextual information on traffic, weather, and air quality, as well as critical events, achieving better traffic prediction and management, allowing operators to suggest alternative routes and act on road restrictions in real time. In addition, the communication of alerts and changes was improved through mobile applications and variable message boards; concluding that the proposed model improves the scalability and maintainability of smart city infrastructures, allowing more effective collaboration between different entities and optimizing the use of resources. This lays the foundation for the development of smarter and more reactive applications in the urban context [1].

The lack of comprehensive studies comparing the implementation of business models in various smart cities and the advantages they offer in each context, for which a narrative literature review and a comparative case analysis were used to investigate business models in smart cities, which are published in academic journals, conferences, books and reports; in addition, cases from London, Amsterdam and Berlin were analyzed, where several business models were identified, such as public-private partnerships (PPP), performance-based contracts, community-centered models, innovation centers, and asset monetization strategies, with variations in their implementation and advantages in each city. Concluding that the study highlights the importance of adapting business models to the specific context of each city to maximize their effectiveness. Directions for future research are suggested, including the impact of sustainability policies and emerging technologies [2].

The lack of practical application of Smart Heritage in heritage sites at area level, using open access data, for which a case study was conducted in Melbourne's Chinatown, using open access data and global best practices for Smart Heritage, where open data from the City of Melbourne were used, such as 3D models, pedestrian counting, and smart tourism practices from the European Smart Tourism Competition (ECST), where ways in which open data can support smart heritage transformation were identified, improving urban heritage management and visitor experience, concluding that open access data is essential for heritage conservation in urban contexts and can foster the sustainable development of historic areas [3].

The vulnerability of urban communities to natural disasters due to climate change and the lack of international standards for resilient infrastructure, for which existing ISO standards were reviewed and analyzed and highlighted the need for a holistic framework for smart and

resilient community infrastructure, they employed an analysis of ISO and IEC standards, including those on emergency management, smart infrastructure and resilience, where gaps in standards coverage were identified in areas such as financing, information management, and lack of a comprehensive framework for resilient smart cities, concluding that international standardization can improve urban resilience to disasters, and the development of new standards can close critical gaps in disaster preparedness and response [4].

The need to better understand urban lifestyle patterns using mobile network data, while maintaining user privacy, for which the LEAF framework was created, which analyzes anonymous mobility data and integrates it with geographic and ontological information to model urban lifestyles, where mobility data from mobile networks, points of interest (PoI), and spatial analysis using a vector model were used, having as results the LEAF framework allowed to accurately model lifestyle patterns in urban areas, showing high consistency with survey data and an RMSE of 5.167, concluding that LEAF is a robust and accurate tool to understand the dynamics of urban lifestyles and support informed decision making in urban planning and resource allocation [5].

Concentrated loading areas in urban areas present challenges in logistics, such as the lack of adequate and optimized spaces for loading and unloading in cities, for which the integration of smart devices in loading areas was analyzed and proposed, including reservations, real-time monitoring and temporary storage, to improve efficiency and sustainability, with the incorporation of examples of technologies such as reservation systems, occupancy and load monitoring devices, and lockers for temporary storage in loading areas, resulting in the implementation of these technologies in urban loading areas improving space management, reducing unauthorized use and optimizing loading and unloading, favoring sustainability, concluding that the integration of smart devices in loading areas is key to improving urban logistics, but a unified regulatory framework and support in data collection are needed to optimize its operation [6].

The need to validate smart city interfaces from a user-centered design perspective, comparing low and high fidelity prototypes in physical and virtual reality environments, for which low fidelity physical prototype field tests and high fidelity tests in virtual reality environments were compared to evaluate realism, interactivity, presence and task difficulty, through the implementation of low and high fidelity prototypes, qualitative interviews and presence and interactivity analysis tools were used, involving design students and experts to obtain insights, resulting in that virtual reality improves visual realism and discussion of user flows, although it has limitations to replicate realistic interactions and social contexts, compared to field tests, concluding that virtual reality is useful in the design process to visualize procedures and test complex user flows, but must be complemented with field tests to capture the context and real user interaction [7].

By reviewing various publications that focus on the use of the fuzzy logic technique for different problematic situations, we describe below the uses and applications, such as the application in the area of commercial law and corporate governance in Mexico, as well as in other international contexts. Key works are mentioned such as "Derecho Mercantil. Parte General y Sociedades" by Paredes and Meade, and "Ley General de Sociedades Mercantiles" by Macedo, which provide a solid foundation on Mexican commercial legislation. In addition, corporate governance issues are discussed, including the importance of internal control and best corporate practices, the research includes analysis on strategic diversification and business performance in Mexico, highlighting the relevance of corporate governance for risk mitigation in private companies. International principles and standards are cited, such as the OECD Principles of Corporate Governance and the Auditing Standards and Procedures of the Mexican Institute of Public Accountants, which underline the importance of transparency and responsibility in business management [8].

The application of the fuzzy model to assess the cost overruns in the healthcare sector due to treatment delays, in a context of the COVID-19 pandemic, where delays in healthcare have been significant. The proposed model uses fuzzy logic to handle the uncertainty and ambiguity inherent in cost assessment in complex and variable situations. The model describes the implementation of the model, such as the need for experts for quantification and the difficulty of obtaining accurate data. To mitigate these problems, a final stage is suggested in which the data is automatically updated each time the model is applied. This proposal seeks to improve accuracy and reduce the costs associated with the involvement of experts, thus facilitating a more efficient and continuous assessment of the cost overruns in the healthcare sector due to treatment delays [9].

In the study of algorithm efficiency and challenges in evaluating technological solutions due to the possible inclusion of closed source code and the need for laborious research, a basic model of a virtual environment used for decision making is proposed for specific experimental studies. In evaluating technological solutions, the importance of classifying resource consumption into qualitative categories such as "high", "medium" or "low" is highlighted instead of focusing on specific metrics such as bytes or microseconds. This significantly simplifies the evaluation by grouping technological solutions into a reduced number of classes based on resource consumption. In addition, various methodologies and formal models for software component optimization and selection are mentioned, underlining the relevance of multi-objective approaches and evolutionary optimization in cloud service composition [10]

.

The study discusses smart grids (SG) and their associated technologies, including distributed energy resources (DER), power electronics components, electric vehicles (EV), and communication and cybersecurity issues. Artificial intelligence (AI) techniques such as fuzzy logic, knowledge-based systems, and artificial neural networks, as well as Internet of Things

(IoT) architecture for applications in SG, are highlighted. Furthermore, IoT and blockchain (BC)-enabled services, such as secure and traceable digital transactions, and the improvement of SG services through AI-based analytics, are discussed. The importance of promoting local DERs, especially renewable energy resources (RER), to achieve stable, reliable, sustainable, and affordable electricity is proposed. The progress of DERs in the global SG context and the support for emerging EV technology to reduce the transportation sector's dependence on oil are discussed. With digital advancements, SG services have been significantly enhanced, especially through the use of AI, IoT and BC, enabling automated services and real-time monitoring of the power grid in terms of reliability, availability, resilience, stability, security and sustainability [11].

In software requirements risk assessment, the goal is to minimize maintenance time and promote customer value. The novelty of the work lies in improving requirements risk assessment and handling developers' subjective judgments on multiple conflicting criteria to provide robust solutions within the cloud computing framework that includes information about responsible persons, their rationale, their assumptions, and their initial and final decision values, to support software requirements reuse in distributed/global development [12].

In the transformation of maintenance strategies in recent decades, moving from a corrective approach to a proactive or predictive one. This evolution is crucial due to the increasing complexity of industrial systems and their impact on the competitiveness and productivity of companies. The research proposes the inclusion of fuzzy logic in the mining industry. In addition, a new approach to risk-based maintenance is presented, detailing an algorithm for risk assessment and estimation, exploring various implementation possibilities of the developed method and discussing the results obtained and their implications [13].

In the evaluation of waste sorting systems, a multi-criteria model based on fuzzy logic is identified for evaluating municipal waste treatment systems, defining parameters such as membership functions, inference rules and fine-tuning functions, based on expert opinion. An evaluation model is implemented in a case study of the Wrocław sorting facility for selectively collected waste (metal and plastic), analyzing the evaluation of the sorting line in seven selected cases. The system evaluation level and sorting efficiency are presented as results, a two-step evaluation method based on fuzzy theory is developed for evaluating the efficiency of the waste sorting system and its impact on the energy quality of the RDF and sorting efficiency is analyzed [14].

Decision support systems for assessing agricultural vulnerability and risk of sugarcane to climate change using a multi-agent model. The proposal involves a dynamic model of industrial systems and the digitalization of the sugarcane field in Mexico to achieve precision agriculture of sugarcane. The implementation applies fuzzy logic and its implementation in MATLAB, highlighting the importance of these tools in the research and development of agricultural models [15].

In assessing the risk in power transformer fleets, the technical condition is considered as the strategic importance of the units, the analysis was performed on a fleet of 19 units and the results demonstrated the viability of the proposed approach. In addition, a new method based on fuzzy logic was developed to improve the accuracy and consistency in the estimation of transformer insulation, the approach based on fuzzy logic showed an accurate assessment of the state of the insulating paper, although it presented some inelasticity that can be reduced by using a larger number of membership functions or replacing the trapezoidal distributions with triangular or Gaussian ones [16].

Integrating COVID-19 pandemic-inspired behaviors into agent-based modeling (ABM) to respond to pollution in water distribution systems. The study uses a modeling framework developed in MATLAB, which is coupled in real-time with the EPANET hydraulic simulation software, allowing instant synchronization and optimization problem solving during simulation. The research focuses on a medium-sized water network model, Net3, to demonstrate the capture of emergent phenomena and the effectiveness of the proposed approach. It details how consumer behaviors, influenced by the pandemic, can be modeled deterministically, eliminating the need for uncertainty analysis. The results show the actions taken by different types of societies over 24 hours, evenly distributed due to the lack of geographic information on pollution [17].

In this work, we developed a method to evaluate pre-professional internship reports in a public university in Peru, where evaluations are often carried out in a qualitative manner, such as good, average or low, which makes it difficult to determine whether the report is approved. We developed the traditional evaluation method through a web system and how this can also be complemented with the use of the Fuzzi Logic model, which allows us to evaluate whether the student passed their pre-professional internship report.

2. Methodology

Pre-professional internships are an essential component of the training of students at public universities in Peru, who begin their insertion into the labor market for a specific period of time, thus building a bridge between theory and practice, between the training stage and entry into the labor market. In this sense, pre-professional internships are a critical aspect for students at public universities, since they allow them to acquire work experience in their area of study and establish important contacts in the working world. In the public university, where there are a large number of students and companies involved in this process, the management of pre-professional internships can be complex. When implementing a web-based system to manage the pre-professional internship process, a series of challenges and problems arise that hinder its effective implementation and proper use.

Consequently, if we consider that public universities in the Lima region have a varied student population between the ten semesters of study, it is estimated that everyone at some point will take

the pre-professional internship course, since this course is included between the ninth and tenth cycle of the study plan and the school management must be prepared to be able to generate a process in the most perfect way possible. In this sense, we should mention that for different reasons the professional schools do not have a system that automates the management process of pre-professional practices, which is why it becomes an object of study in this thesis.

The materials and methods are described by the presentation of the processes necessary to be able to replicate the methodology.

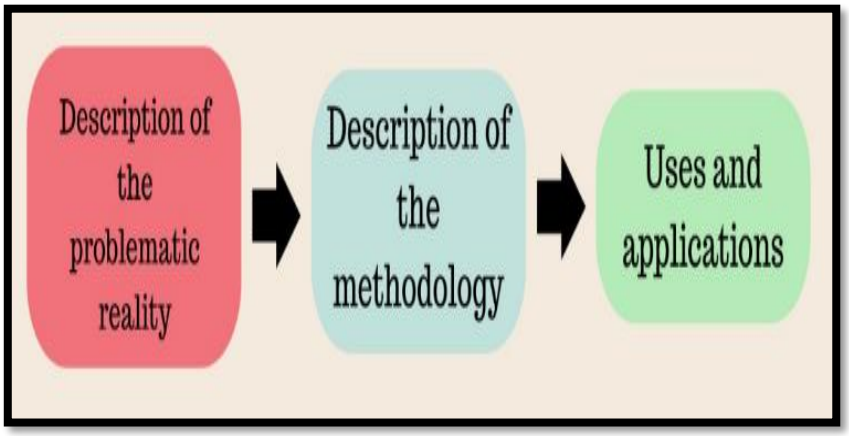


Figure 16: Block diagram of the proposal.

Figure 1, presents the block diagram that represents the method of the proposal, starting with the description of the problem, followed by the description of the methodology and ending with the uses and applications.

2.1. Description of the problematic reality

The description of the problematic reality is described in the conditions that are needed to be able to start and to be able to conclude the pre-professional practices processes. We begin by indicating that the pre-professional practices are an indispensable requirement to be able to obtain the bachelor's degree for the students, which is why there is a need for the students to be able to pass this evaluation, which consists of the review of the practice report that the student made and that must be approved.

The mechanism of review and approval of the practice report consists of the detailed review of the report by three teachers, each of them evaluates the report and according to their recommendations, the report is approved or disapproved. The report is sent to 3 teachers, after the student registers in the Web System for the exclusive use of students and the degree and title office.

Below we present the main processes of the pre-professional internship system:

1. **Management of pre-professional internships**
   The management of pre-professional practices encompasses the development, programming, coordination, control and assessment of internship plans, with the purpose of ensuring an adequate education for students and an effective connection with the work environment.
   Pre-professional internship management involves the planning, coordination and evaluation of practical learning opportunities for students, in order to combine theory and practice, boost the development of skills and competencies, and facilitate the transition to employment.

2. **Selection process**
   The selection process comprises a series of organized and consecutive stages that are carried out for the purpose of collecting data on candidates, analyzing their capabilities and competencies, and making informed decisions about their suitability for a given position. The selection process encompasses a set of methods and tools used to recognize and assess suitable candidates, with the goal of determining who will be hired to fill a position, thereby ensuring a fair and unbiased selection.

3. **Number of students practicing**
   Refers to the number of students who are carrying out internships in an organization or company as part of their academic or pre-professional training. These numbers may vary depending on the context and policies of the university and the organization receiving the internship.

4. **Percentage of students practicing**
   Refers to the percentage of students who are carrying out internships in an organization or company as part of their academic or pre-professional training. These percentages may vary depending on the context and policies of the university and the organization receiving the internship.

5. **Supervision**
   Supervision encompasses the act of observing, guiding, and leading the work of subordinates for the purpose of ensuring that established objectives are achieved and required performance standards are met.

6. **Number of supervisory visits**
   Refers to the number of times a supervisor visits or inspects the internship center to monitor and evaluate its performance. These supervisory visits are part of a monitoring and

follow-up process designed to ensure that quality standards, established objectives, and university policies and procedures are being met.

7. **Communication**
   Communication refers to the process by which information and meaning are transferred and exchanged between two or more individuals, for the purpose of sharing knowledge, ideas, feelings, and coordinating actions effectively. Communication involves conveying information, ideas, emotions, and meanings using both verbal and nonverbal symbols for the purpose of exerting influence on other people's thinking and behavior.

8. **Number of students who recommend taking the pre-professional internships**
   Refers to the number of students who recommend to other students to carry out an internship in an organization or company as part of their academic or pre-professional training. These numbers may vary depending on the context and policies of the university and the organization receiving the internship.

2.2. Description of the methodology

The description of the methodology is related to the description of the tools used to implement the web system and the intelligent part, which consists of working through fuzzy logic, for which the Matlab tool has been used. Below we describe the architecture of the application.

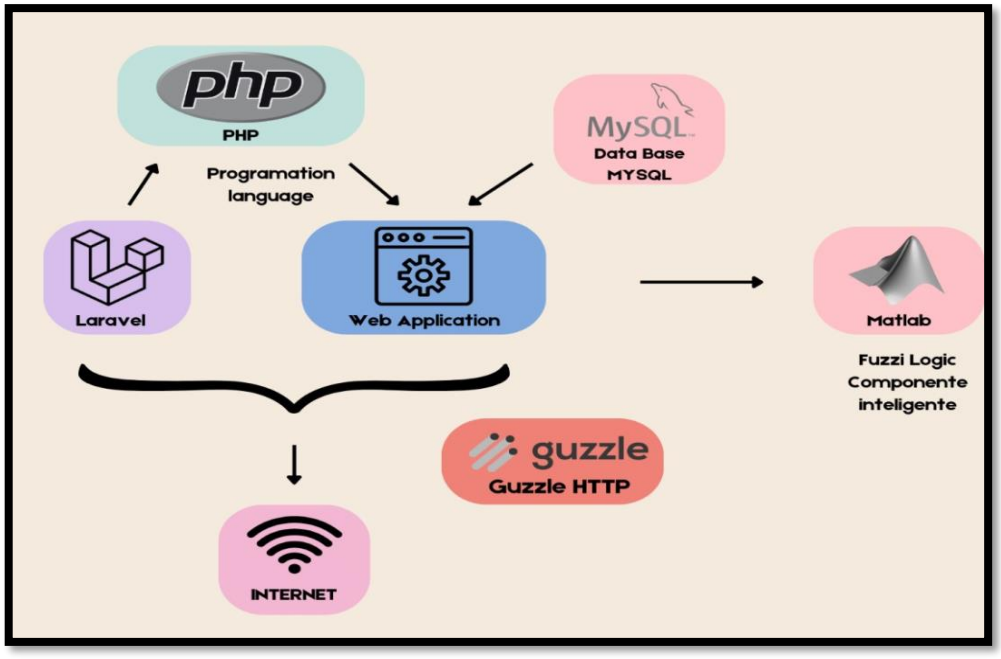


Figure 17: Description of the architecture of the development proposal.

Figure 2 presents the architecture of the implemented solution; for the implementation of the web system, where the use of the PHP programming language in version 8 with Laravel in version 9.19 is used, the chosen database is MySQL and as the HTTP client Guzzle for PHP, the intelligent component is developed in Matlab using the Fuzzi Logic tool.

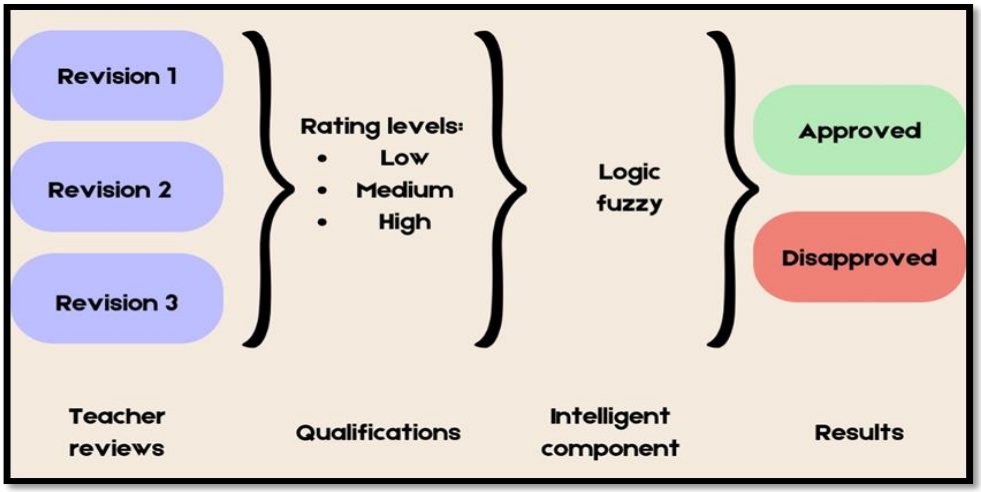


Figure 18: Description of the evaluation method for pre-professional practice reports.

In Figure 3, the evaluation mode of the practice reports is presented, the process begins with the sending of the report to three teachers, each one evaluates according to their perception and experience, the result of the evaluation is a grade that in many cases is a qualitative grade, which can have the values of "High", "medium" or "Low", which makes it difficult to obtain the final result of the evaluation that is approved or failed; For this task, we resort to the use of the intelligent component that is the use of fuzzy logic, for which the values of the teachers' grades are evaluated, and the final evaluation value is returned.

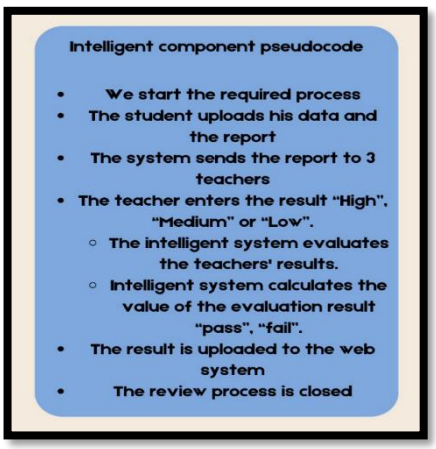


Figure 19: Pseudocode of the we system.

In Figure 4, the pseudocode of the intelligent component is presented, where it is described how the web system and the Matlab tool work, in order to have the final result of the evaluation, this process is carried out manually, as a demonstration mode, where the evaluation value of each teacher is taken, loading it into the Matlab application and the result is loaded into the web system, with which the process is closed, with the final evaluation values which are "Passed" or "Failed".

2.3. Uses and applications

The uses and applications are related to the way of using the implemented web system, where the intelligent component that performs the calculation of the evaluation of the partial results stands out; in the implemented test model the process is carried out manually, with the intention of being able to carry it out automatically by presenting a function that converses the Matlab function with PHP.

3. Results

The results we describe are related to being able to describe how the web system and the intelligent component work, analyzing the main processes of the system, such as the student registration process and the evaluation with the Matlab tool.

Figure 20: Implementation design.

In Figure 5, the registration screen of the proposed system is presented, where the entry is evidenced through a user registration, having to enter with its corresponding password.

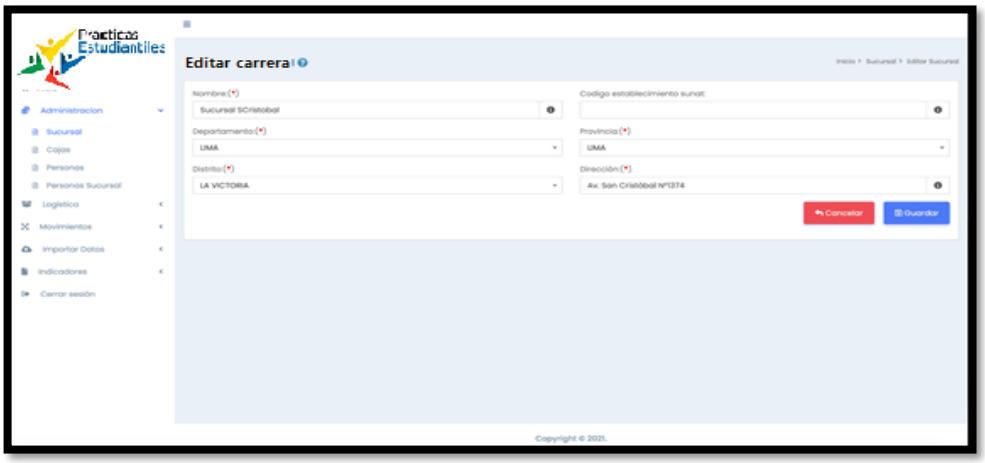


Figure 21: Data entry form.

Figure 6 shows the main form for data entry by students. The data entered includes the specialty that the student is studying, the place where he/she did his/her internship, and most importantly, the entry of the report made, where the information on the internship process is found.

3.1.1. Smart component

When quantitative results are obtained, evaluation is carried out using Fuzzi Logic techniques, with which a report is evaluated by 3 teachers, so its result must be approved and disapproved, the evaluators' review values are "good", "regular", and "bad". Below we present the generated model, considering an output with "approved" and "failed" values.

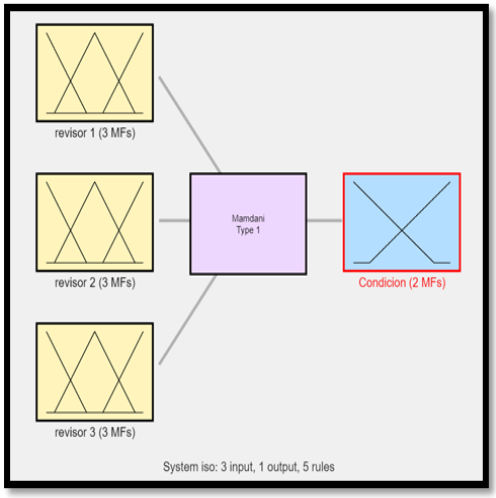
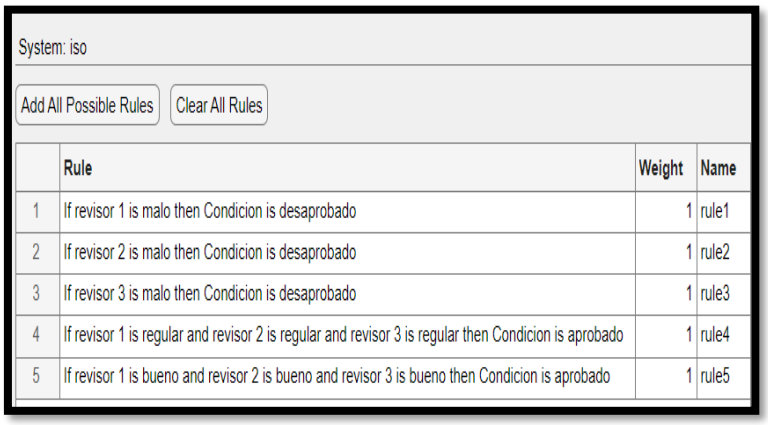


Figure 22: Fuzzi Logic Model.

In Figure 7, the fuzzy logic model is presented where there are 3 inputs that correspond to the evaluators and an output that corresponds to the result of the evaluation.



	Rule	Weight	Name
1	If revisor 1 is malo then Condicion is desaprobado	1	rule1
2	If revisor 2 is malo then Condicion is desaprobado	1	rule2
3	If revisor 3 is malo then Condicion is desaprobado	1	rule3
4	If revisor 1 is regular and revisor 2 is regular and revisor 3 is regular then Condicion is aprobado	1	rule4
5	If revisor 1 is bueno and revisor 2 is bueno and revisor 3 is bueno then Condicion is aprobado	1	rule5

Figure 23: Interpretation of the rules.

Figure 8 presents the roles considered in the evaluation, where there is a general condition, the student is approved if at least the reviewers' evaluations correspond to the regular value, higher values guarantee approval and a bad value ensures failure.

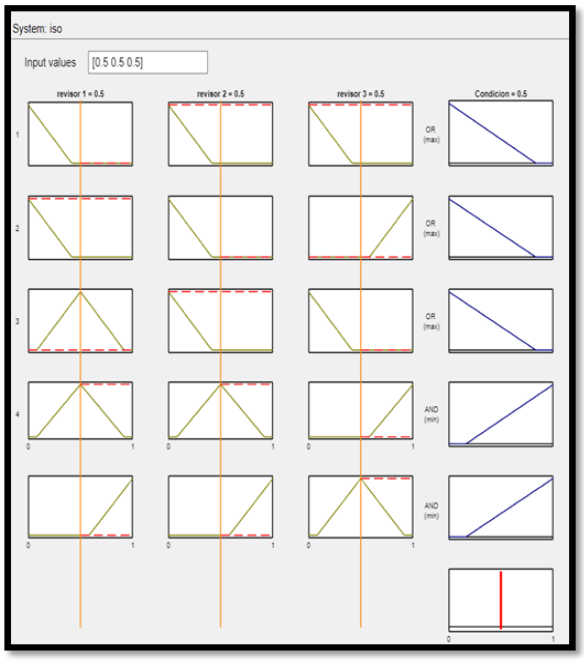


Figure 24: Analysis of results.

In Figure 9, the interpretation of the rules is presented, considering a value of bad ensures disapproval, and a minimum approval value is when the 3 evaluators consider it as regular, higher combinations ensure approval of the report.

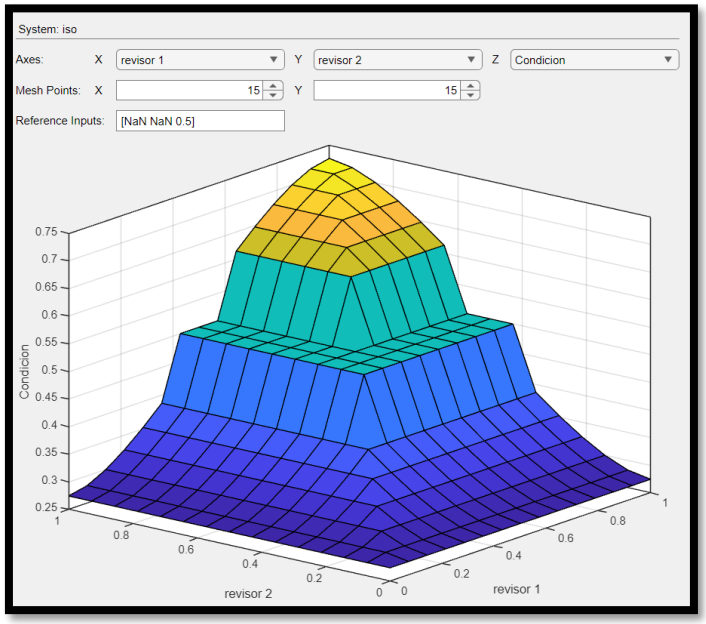


Figure 25: Fuzzi logic model 3D.

Figure 10 shows the 3D model, where the minimum approval values are evident; values greater than 0.5, which corresponds to the "regular" value of the evaluation, ensure approval and lower values are considered as failure.

4. Conclusions

In conclusion, with the use of fuzzy logic, embedded in a web application, an increase from 40.52% to 82.76% was achieved, which allowed validating that the use of fuzzy logic, implemented in a web system, is favorable for the indicator number of practicing students.

It is concluded that for the indicator number of supervised visits to the pre-professional internship center there was an increase of visits to the pre-professional internship centers by the pre-professional internship supervisors quite considerable equivalent to 42.24% since at the beginning the supervised visits were on average 40.52% per semester and after the implementation of the software the sales were on average 82.76%.

It is concluded that the indicators number of students practicing and the number of supervised visits to the center of pre-professional practices allow a favorable measurement for the management of pre-professional practices validating that these are concluded in a correct way and that they increase in quantity and therefore there will be more graduates for the university.

It is recommended to optimize the internship computer system by developing a new project solution following the MVC (Model View Controller View) software architecture pattern that will help to separate the data and business logic of the internship system, thus facilitating maintenance, component reuse and scalability of the application.

It is also recommended to improve the organization of the employability workshops to be more precise the workshop of Induction of pre-professional practices, it could be considered the use of the google calendar for the scheduling of the workshops mentioned on the website and other media that the Communications office uses for the dissemination of our event thus informing students the schedule of the workshops in a timely manner for their registration and participation.

It is recommended to continue conducting periodic studies on the satisfaction of students who are doing internships, since in many surveys the suggestions presented by the students show which changes or problems are the most critical in order to address them in a timely manner, even more so if they are related to the internship computer system.

Finally, it is also recommended to make an evaluation about the programming language to be used in a future new version of the internship computer system, since the public universities can use the reculate framework for the projects that are developed in the corporation can also adopt the Python programming language that has additional advantages to the language used in the present pre-professional internship system (PHP) since most of the web applications used in the corporation and also in the development industries that use Python with R.

When evaluators provide qualitative values, we resort to the use of evaluation tools using the fuzzy logic technique, resulting in a result that can be worked on, as is the case with the evaluation of reports, where a report result is required, for which the fuzzy logic technique is very helpful.

References

[1] Bellini, P.; Bologna, D.; Nesi, P.; Pantaleo, G. A Unified Knowledge Model for Managing Smart City/IoT Platform Entities for Multitenant Scenarios. *Smart Cities* 2024, *7*, 2339-2365. https://doi.org/10.3390/smartcities7050092

[2] Wolniak, R.; Gajdzik, B.; Grebski, M.; Danel, R.; Grebski, W.W. Business Models Used in Smart Cities—Theoretical Approach with Examples of Smart Cities. *Smart Cities* 2024, *7*, 1626-1669. https://doi.org/10.3390/smartcities7040065

[3] Geng, S.; Chau, H.-W.; Jamei, E.; Vrcelj, Z. Demystifying the Use of Open-Access Data in Smart Heritage Implementations. *Tour. Hosp.* 2024, *5*, 1125-1150. https://doi.org/10.3390/tourhosp5040063

[4] Nguyen, D.N.; Usuda, Y.; Imamura, F. Gaps in and Opportunities for Disaster Risk Reduction in Urban Areas Through International Standardization of Smart Community Infrastructure. *Sustainability* 2024, *16*, 9586. https://doi.org/10.3390/su16219586

[5] Moghari, S.; Fallah, M.K.; Gorgin, S.; Shin, S. LEAF: A Lifestyle Approximation Framework Based on Analysis of Mobile Network Data in Smart Cities. *Smart Cities* 2024, *7*, 3315-3333. https://doi.org/10.3390/smartcities7060128

[6] Bóna, K.; Sárdi, D.L.; Büki, A.; Domaniczki, V. The Future Role of Smart Devices in Systems of Urban Concentrated Loading Areas. *Eng. Proc.* 2024, *79*, 15. https://doi.org/10.3390/engproc2024079015

[7] Joundi, J.; Baccarne, B.; Robaeyst, B.; Bombeke, K.; De Marez, L.; Saldien, J. Comparing Virtual and Real-Life Rapid Prototyping Methods for User Testing Smart City Interfaces: A Case Study. *Appl. Sci.* 2024, *14*, 9918. https://doi.org/10.3390/app14219918

[8] Mancilla-Rendón, E.; Lozano, C.; Torres-Esteva, E. Fuzzy Governance Model. Mathematics 2021, 9, 481. https://doi.org/10.3390/math9050481

[9] Brotons-Martínez, J.M.; Sansalvador-Sellés, M.E. Proposal for a Fuzzy Model to Assess Cost Overrun in Healthcare Due to Delays in Treatment. Mathematics 2021, 9, 408. https://doi.org/10.3390/math9040408

[10] Nikulchev, E.; Ilin, D.; Gusev, A. Technology Stack Selection Model for Software Design of Digital Platforms. Mathematics 2021, 9, 308. https://doi.org/10.3390/math9040308

[11] Kumar, N.M.; Chand, A.A.; Malvoni, M.; Prasad, K.A.; Mamun, K.A.; Islam, F.R.; Chopra, S.S. Distributed Energy Resources and the Application of AI, IoT, and Blockchain in Smart Grids. Energies 2020, 13, 5739. https://doi.org/10.3390/en13215739

[12] Gupta, C.; Gupta, V. Are These Requirements Risky: A Proposal of an IoT-Based Requirements Risk Estimation Framework. Mathematics 2022, 10, 1210. https://doi.org/10.3390/math10081210

[13] Tubis, A.; Werbińska-Wojciechowska, S.; Sliwinski, P.; Zimroz, R. Fuzzy Risk-Based Maintenance Strategy with Safety Considerations for the Mining Industry. Sensors 2022, 22, 441. https://doi.org/10.3390/s22020441

[14] Giel, R.; Kierzkowski, A. A Fuzzy Multi-Criteria Model for Municipal Waste Treatment Systems Evaluation including Energy Recovery. Energies 2022, 15, 31. https://doi.org/10.3390/en15010031

[15] Badillo-Márquez, A.E.; Aguilar-Lasserre, A.A.; Miranda-Ackerman, M.A.; Sandoval-González, O.O.; Villanueva-Vásquez, D.; Posada-Gómez, R. An Agent-Based Model-Driven Decision Support System for Assessment of Agricultural Vulnerability of Sugarcane Facing Climatic Change. Mathematics 2021, 9, 3061. https://doi.org/10.3390/math9233061

[16] Zaldivar, D.A.; Romero, A.A.; Rivera, S.R. Risk Assessment Algorithm for Power Transformer Fleets Based on Condition and Strategic Importance. Algorithms 2021, 14, 319. https://doi.org/10.3390/a14110319

[17] Kadinski, L.; Ostfeld, A. Incorporation of COVID-19-Inspired Behaviour into Agent-Based Modelling for Water Distribution Systems' **Contamination** Responses. Water 2021, 13, 2863. https://doi.org/10.3390/w13202863

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

[Paper ID-31] Smart Cities and AI: Leveraging User Data for Big Data Analytics Insights

Ruchika Pharswan[1,*,†], Rohit Goel[2], Kajal Kansal[3]

[1]Department of Information Technology : Delhi Technological University

[2,3]Department of Computer Science and Engineering : Maharaja Agrasen Institute of Technology

Abstract

Artificial intelligence (AI) has gotten progressively significant in everyday life, from navigation systems to intelligent assistance, prompting initiatives like Smart City. Thus, the successful adoption of the smart city is directly or indirectly linked to the acceptance of advanced technologies. Researchers have mainly explored the technological framework of smart city adoption. However, the impact of the emerging technologies and their acceptance on smart city adoption yet largely unexplored and continues to be an abstract idea on various grounds, which leaves a gap. This paper aims to determine the influencing factors of smart city adoption by analyzing the AI application user's experience and acceptance from the lens of various stakeholders. In this paper, we have extracted the data from the Social Media Platform, then performed sentiment analysis and network analysis to identify the significant constructs that can impact the Smart cities adoption. Then we use stepwise regression analysis with permutation testing approach via SVM classifier to determine the influence of constructs on the user in AI application experience. The study establishes that perceived innovation, value, compatibility, ease of use, enthusiasm, performance, efficiency, social influence, and user expertise significantly increase the stakeholder's AI experience. The current study can be an excellent source for various stakeholders' reference while adopting AI or comparative Technology deployed in the smart city.

Keywords

Artificial Intelligence, Smart Cities, Support Vector Machine, Social Media Platform

1. Introduction

The prominence of smart cities emanates from the changes and challenges that emerge due to the growing urban population. The rapid rates of urbanization and the fast-growing population in urban areas present many concerns like a strain on existing infrastructure, traffic congestion, environmental changes, security issues, inefficacious medical facilities, and outmoded Technology and governance wherein smart cities intended to contribute and sustain a high quality of life [1].

CEUR Workshop Proceedings : 3rd International Symposium on Smart Cities Challenges, Technologies and Trends (SCCTT-2024)

*R. Pharswan
ruchikapharswan2024@gmail.com (R. Pharswan); rohitgoyal8178@gmail.com (R. Goel); kansalkajal.30@gmail.com (K. Kansal)
0000-0003-0478-8168 (R. Pharswan)

The smart city is amalgamating advanced technologies to overcome modern problems to develop a more sustainable, substantial, and compelling city. However, the impact of these emerging technologies on smart city adoption is yet largely unexplored and continues to be an abstract idea on various grounds. Significantly, the smart cities ideas and development were inextricably associated with the emergence of new information technologies. So, it becomes imperative to pay attention to the factor directly and indirectly associated with the usage and applications of those technologies that can impact smart city services[2].

Though numerous researchers have examined the possible smart city enablers such as technological frameworks and evaluate their impacts on the adoption process, the linkage betweentechnology adoption and smart city development is yet to be explored[3]. To bridge this knowledge gap, the current study aims to understand the impacts of user's experience of AI applications on smart city developments to accelerate its acceptance and adoption.

Although new technology, infrastructure, and industries can build smart cities, its adoption is more likely to depend upon technology stakeholder's participation and experience Smart city's idea can be implemented and adopted extensively only when users have familiarity and prior experience with the constituent technologies[4]. Consequently, there is an essential missing link or insufficiently addressed issue: how the stakeholder's AI application experience may be critical to the smart cities' adoption[5]. There needs to be a study that identifies and connects all the determinants of the stakeholder's AI applications experience associated with the smart cities. Our research will be essential to bridge this literature gap and be a leader for future research[6]–[8].

This study's primary objective is to focus on the information system (IS) context of smart cities from various stakeholder lenses. The undertaken research addresses the following research questions aiming to identify the linkage between AI adoption, stakeholder's AI experience, and smart cities adoption:

***RQ**: How is stakeholder's AI application experience associated with the successful adoption in the smart city?*

This paper has proposed a robust framework that collects data from social media platforms (SMP) (i.e., Twitter) based on hashtags and keywords related to the AI application and smart cities, followed by preprocessing and training the data. We have used social media analytics-based sentiment analysis. After that, we performed network analysis. The second phase is the validation phase. Stepwise regression and support vector machines are used to confirm the significance of identified constructs that can impact the user's AI application experiences and, consequently, smartcity adoption.

The rest of the paper is organized as follows: We consolidated available literature related to smart city adoption in the second section. The third section portrays the research methodology followed for the undertaken research. The fourth section outlines the findings along with the hypothesis development that is validated in a later section. The fifth section discourses the implications of our results. Simultaneously, the sixth section has concluded our research's learnings, followed by limitations and future research scopes in the seventh section.

2. Literature Review

2.1 AI and Technology adoption

AI is considered an essential component of digital transformation and is globally viewed as one of the core competitive technologies by organizations in every sector. AI's proven potential is extremely capable of developing efficient and smart cities. Despite the propelling advantages of AI, evidence indicates that AI's adoption is still in the nascent stage[9], [10]. There is a need for a study that uncovers the factors influencing AI adoption. Numerous works of literature discuss IT adoption, focusing more on organizations' points of view. The old theories discussed some characteristics like the complexity of Technology, compatibility, the relative advantage to use it, observability, and last one is trialability. They considered these factors are essential for the diffusion of any technology. Various studies explore AI in multiple domains. Some studies consider the theoretical aspect of AI and other deal with AI adoption in different sectors. Many studies have explored AI's pervasive nature, and not much literature exploring the AI adoption from the stakeholder's lens leaves a gap. Studies on AI adoption from technical, organizational, and environmental aspects are available so far[11]. Therefore,there is a need for research that identifies the factors influencing AI adoption from various stakeholders' viewpoints, such as AI services provider, governance, AI end-users. Our study is intended to fill this gap with our proposed work.

2.2 Stakeholders Participation and Experience

Due to the pertinence of smart cities to various stakeholders, it is crucial to consider the opportunities and barriers associated with its adoption from the stakeholders' lens.

2.2.1 AI Application Providers

The experience of AI vendors with these technologies also impacts AI adoption. Adopting any new technology, among others, is also influenced by the marketing strategies of the vendors, so for the overall adoption of AI, it is crucial to consider the factors that influence the vendor's AI experience. Since not everyone is very versed with new technologies, it requires a certain level of skills to use the technologies like AI, so vendor's expertise is one can be one of the factors that can influence one's AI usage experience and hence the opinion regarding its adoption Studies validate the significant impact technology vendors have on technology adoption, directly or indirectly associated with the vendor's AI experience.

2.2.2 Government

Administrating a smart city government must become smart and incorporate technology-based decision systems for better planning and policymaking. The smart Government envisions improving the governing strategies, procedures and modifying the community services to provide the best. The Government's readiness to promote the new Technology will encourage growth and adoption[12]. Government targets accomplishing transparency in administrative procedures, policies-making systems, governance systems, and decision-making systems to improve the planning and deployment of public services and accessible assets. These factors can be considered the factors influencing the Government AI experience and their AI acceptance and adoption decision.

2.2.3 Citizen

Residents are the smart services users; thus, it is pivotal to address the factors driving or hindering smart cities' adoption from the user's perspective to ensure the successful and substantial adoption of smart cities.

The participation and engagement of citizens is a measure that impacts smart city adoption. Numerous research has effectively been conducted on citizen participation. If citizens find AI-based products and services easy to use, safe, and secure, they might like to continue with more technologies or solutions like smart cities[10]. There are notion of technology adoption from the user's viewpoint relies on their experience. Various research explored citizen participation, yet being an emerging technology, AI user experience is less explored. Citizen's AI experience relies upon factors like ease of use, security of their data, their prior experience or expertise, as discussed in some literature[13].

3. Methodology

The whole methodology is separated into two portions. The first is scrutinizing stage, where information encompassing AI and smart cities adoption is extracted from SMP. Different exploratory investigations such as sentiment analysis and network analysis are performed to recognize the potential constructs. The subsequent stage is the validation stage. Wherein factors influencing the AI user's experiences are again determined and validated, however, by statistical techniques.

3.1 Phase one: Scrutinizing phase

3.1.1 Research setting and data collection

Twitter is one of the largest SMPs in terms of both number of active users online and the scale ofreach and penetration to focused groups of users. Hence mining the data from Twitter helps us analyze the opinions and understand the signals that could lead to theoretical constructs that can be tested later. The data collection framework consists of four modules. Firstly, the data extraction module is built on the python toolkit utilizing the Twitter streaming API with the results stored in a .csv file[14]. The keywords and hashtags used for data collection are [''#ArtificialIntelligence,' ''#AI,' ''#Smartcity,' ''#DataScience,' ''#Citizens,' ''#Applications,' ''#Smartgrid,' ''Sustainability,' ''Automatic'']. We downloaded over one million tweets.

As part of the next step, we start preprocessing the data by cleaning, stemming, tokenization, and normalization. It also includes removing duplicate texts, denoising, and spell-checks. Normalizing the data caters to interpreting the standard terms from slang jargon linguistics. Stemming overall normalizes all the tenses to present tense and tokenization, creating tokens to words with usage in the sentences, i.e., noun, adjective, adverb. Data summarization is critical to reducing a large volume of unstructured text to manageable form while keeping the essential signal information athand.

3.1.2 Sentiment and communities of topics

We use sentiment scores to identify the polarity of the expressed opinion in each of the tweets. Wethen classify the tweets into positive, negative, and neutral sentiment groups and then repeat the procedure on extreme positive and negative sentiment groups to understand the prominent themes[13].

3.2 Phase two: Validation phase

In the second phase, we again extracted the data from the SMP and performed preprocessing. Thenwe use the permutation testing approach to identify the significant constructs that impact the user is AI application experience, for which we use an SVM classifier. As a result of the permutation testing approach, we now have a set of significant factors, and now we validate the same by usingthe stepwise regression analysis

4. Finding

4.1 Insights from the Phase one analysis: Scrutinizing Phase

In this paper, data was analyzed from Twitter to provide insights into the conversation around smartcities and AI adoption. When we have closely analyzed the conversations, we learned that some leading experiential keywords illustrate the positive and negative discussions from the various stakeholders, as shown in figure 1.

Fig 1: Word cloud representing positive and negative conversations group around AI adoption and Smart cities.

For instance, keywords like "efficient", "solution", "readiness", "governance", "technology", "confidence", "interaction" have been sighted from the positive conversations. Moreover, these keywords are outwardly related to the theoretical presumptions we have undertaken for the hypothesis development while theory building. Applying a similar approach to the negative conversations, we observe keywords like "fear," "e-attack," "failure," "privacy," "inconsistency," "emotions," "identity- theft" came into the picture. Furthermore, these keywords are also associated with the hypothetical theory we have constructed for AI adoption and smart cities.

4.2 Insights from Network analysis and clustering

In our study, words are the unit of analysis, and we have effectively sighted some keywords emerging from the conversations. However, it is imperative to determine how the extracted keywords are associated with each other as a part of a more extensive conversation. To recognize this, we develop a network diagram of leading keywords for each sentiment group. Figures 2 and 3 represent the network diagrams of positive and negative conversations, respectively. Scrutinizing both network diagrams concurrently, we figured out that although both the network diagrams are around smart cities and AI adoption, however, exhibit different perspectives. One network diagram entails more development, automation, governance, reliability, and success, whereas another reflects the severe concern to cost, privacy, complexity, latency, and security.

Capturing signals and analyzing those using social media Analytics, we tracked down some emerging keywords from the conversation around the AI adoption and smart cities. Since in this study dependent variable is the stakeholder's AI adoption experience that directly and indirectly depends upon individuals' AI usage experience.

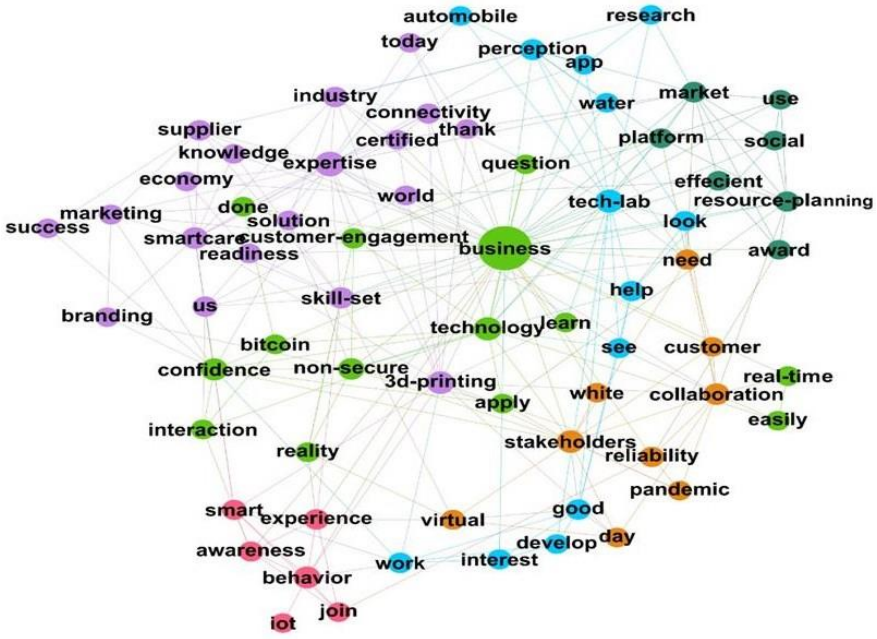


Fig 2: Network Diagram for positive group conversation
5.

Thus, in this research study, we have considered stakeholder's AI usage experience and emotions as a proxy for further AI and smart city adoption experience[5]. We attempt to correlate constructs identified from apprehending literature and analyzing SMPs and factors like Perceived *Innovation, Perceived Value, Perceived compatibility, AI Expertise, Efficiency, Performance, Perceived Security, Perceived Ease of Use, Perceived Risk, Perceived Enthusiasm, Perceived Trust, Perceived Readiness, Social Influence, User's literacy, User's expertise, comes* out to be potential common factors that can influence the adoption of AI and smart cities.

4.3 Insights from the Phase two analysis: Validation phase

In phase two, we selected stepwise regression analysis to identify the most significant constructs out of all the potential constructs. Since we have high dimensionalities in our data and many potential constructs, so prune the list, we conducted the iterative regression analysis and the support vector machine (SVM) to scrutinize all the possible set of most significant factors that impact the stakeholder's AI usage experience[15]. To obtain the most significant set of constructs, we have used the permutation testing approach. In this technique, random labels are assigned to the constructs, and then we use an SVM classifier to train the model. After that, we test the significance of each construct by using the "one left out "cross-validation strategy. We use this procedure in iteration until we separated the most significant constructs. Finally, the permutation testing

approach is presented in box and whisker plot in figure 4. We can see that four constructs are discovered to be insignificant, having a p-value more prominent than 0.05.

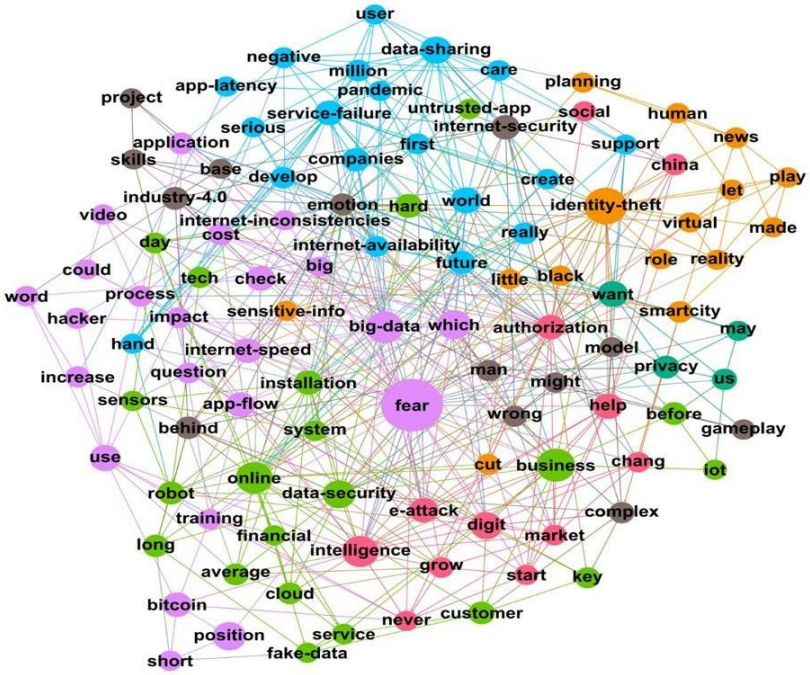


Fig 3: Network Diagram for Negative group conversation

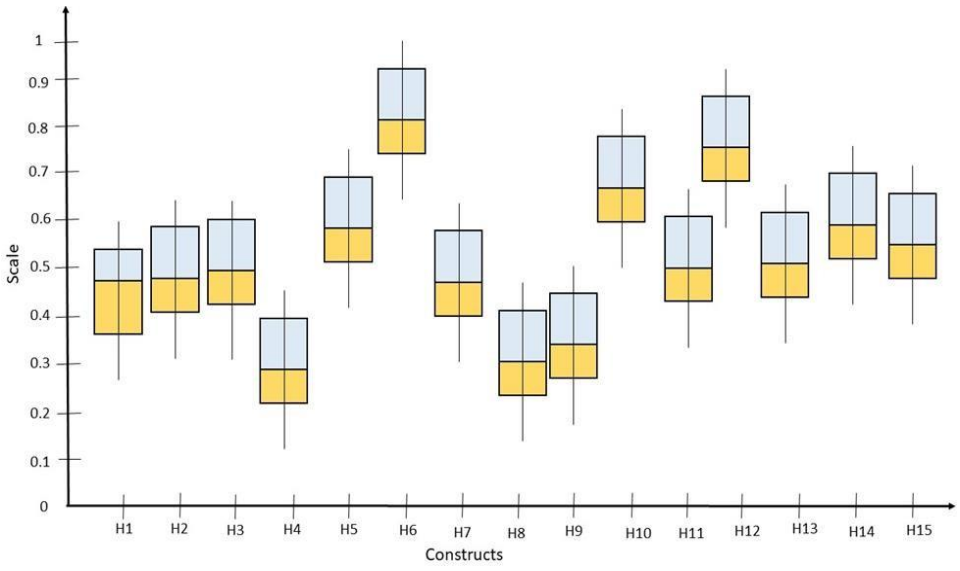


Figure 4: Construct's Significance testing using SVM

We have identified fifteen constructs and tested each of these hypotheses individually using Pearson's chi-square to determine their P-values. Those who are having p-values less than or equivalent to 0.05 are considered as the most significant factors. However, the constructs AI expertise, perceived compatibility, perceived security, and perceived risk was not accepted due to its higher p-values. Out of 15 potential constructs identified from the literature and our Prior analysis, 11 contracts were eventually found to be significantly associated with the AI usage experience of various stakeholders. . The result of regression analysis is represented in Table 1.

Table 1: Result of stepwise regression

Index	Coef.	Std. Err.	T	P> |t|	[0.025	0.975]
Constant	0.3272	0.001	269.096	0	0.325	0.33
Perceived Value	0.1393	0.037	3.765	0	0.067	0.212
Perceived Innovation	0.1368	0.029	4.784	0	0.081	0.193
Perceived Readiness	1.444	0.141	-10.265	0	-1.72	-1.168
AI Expertise	0.0254	0.026	-0.961	0.331	-0.077	0.026
Efficiency	0.0248	0.022	-1.144	0.252	-0.067	0.018
Performance	0.1201	0.068	1.762	0.078	-0.013	0.254
User's literacy	0.3487	0.128	-2.731	0.006	-0.599	-0.098
Perceived Compatibility	-0.0083	0.02	0.413	0.68	-0.031	0.048
Perceived Security	0.1847	0.26	-0.71	0.478	-0.694	0.325
Perceived Ease of Use	0.337	0.133	-2.535	0.011	-0.598	-0.076
Social Influence	2.0008	0.048	-41.432	0	-2.095	-1.906
User Expertise	1.4483	0.515	2.812	0.005	0.439	2.458
Perceived Risk	0.0039	0.03	0.131	0.896	-0.054	0.062
Perceived Enthusiasm	1.2314	0.162	7.596	0	0.914	1.549
Perceived Trust	0.1029	0.039	1.564	0.057	0.014	0.312
R-Square	56.2					
Adjusted R-Square	52.2					

5. Conclusion

The entire concept of the smart city is to make people, Government, and Technology smart. Emerging Technology like AI is the keystone of smart cities. Thus, to understand how users perceive smart city initiatives, it is imperative to identify how users perceive the AI applications, how positive their user

experience with AI is, and the factors influencing the AI usage experience.The entire study is divided into two segments. First is scrutinizing phase where data surrounding AI and smart cities adoption is extracted from SMP and various exploratory analysis is performedto identify potential construct. The second phase is the validation phase. Wherein data constructs are again identified but by statistical techniques. As the finding of the first phase, we identified two groups having upbeat sentiments regarding AI and smart city adoption. We identified two network diagrams for each sentiment group. Mapping all the findings of phase one, we determined fifteen constructs that, according to the phase one analysis, influences the user's AI application experience. The second phase is the validation phase. We try to validate the step one findings by using statistical techniques. We use stepwise regression analysis along with the permutation testing techniques wherein SVM is used as a classifier. Multiple set of significant constructs has been made and tested using SVM and stepwise regression. Mapping the output of phase one and phasetwo, we find out of 15 constructs, 11 constructs were significant and influenced the AI and smartcity adoption. This will be the first study that identified the crucial factors for accelerating the user's AI application experience and hence smart cities adoption via the highlighted approaches inthis research.

References

[1] E. Araral, “Why do cities adopt smart technologies? Contingency theory and evidence from the United States,” *Cities*, vol. 106, no. July, p. 102873, 2020, doi: 10.1016/j.cities.2020.102873.

[2] T. Fontes, V. Costa, M. C. Ferreira, L. Shengxiao, P. Zhao, and T. G. Dias, “Mobile payments adoption in public transport,” *Transp. Res. Procedia*, vol. 24, pp. 410–417, 2017, doi: 10.1016/j.trpro.2017.05.093.

[3] P. Zhao, Y. Yen, E. Bailey, and M. T. Sohail, “Analysis of urban drivable and walkable street networks of the ASEAN smart cities network,” *ISPRS Int. J. Geo-Information*, vol. 8, no. 10, 2019, doi: 10.3390/ijgi8100459.

[4] N. Staletić, A. Labus, Z. Bogdanović, M. Despotović-Zrakić, and B. Radenković, “Citizens’ readiness to crowdsource smart city services: A developing country perspective,” *Cities*, vol. 107, no. July, p. 102883, 2020, doi: 10.1016/j.cities.2020.102883.

[5] Y. Lim, J. Edelenbos, and A. Gianoli, “Identifying the results of smart city development: Findings from systematic literature review,” *Cities*, vol. 95, no. October 2018, p. 102397, 2019, doi: 10.1016/j.cities.2019.102397.

[6] M. Zhang, P. Zhao, and S. Qiao, “Smartness-induced transport inequality: Privacy concern, lacking knowledge of smartphone use and unequal access to transport information,” *Transp. Policy*, vol. 99, no. January, pp. 175–185, 2020, doi: 10.1016/j.tranpol.2020.08.016.

[7] R. K. R. Kummitha and N. Crutzen, “How do we understand smart cities? An evolutionary perspective,” *Cities*, vol. 67, no. May, pp. 43–52, 2017, doi: 10.1016/j.cities.2017.04.010.

[8] Q. Liu, Z. An, Y. Liu, W. Ying, and P. Zhao, “Smartphone-based services, perceived accessibility, and transport inequity during the COVID-19 pandemic: A cross-lagged panel study,” *Transp. Res. Part D Transp. Environ.*, vol. 97, no. June, p. 102941, 2021, doi: 10.1016/j.trd.2021.102941.

[9] A. Shamsuzzoha, J. Niemi, S. Piya, and K. Rutledge, "Smart city for sustainable environment: A comparison of participatory strategies from Helsinki, Singapore and London," *Cities*, vol. 114, no. March, p. 103194, 2021, doi: 10.1016/j.cities.2021.103194.

[10] S. EL HILALI and A. AZOUGAGH, "A netnographic research on citizen's perception of a future smart city," *Cities*, vol. 115, no. March, p. 103233, 2021, doi: 10.1016/j.cities.2021.103233.

[11] H. Jiang, "Smart urban governance in the 'smart' era: Why is it urgently needed?," *Cities*, vol. 111, no. June 2020, p. 103004, 2021, doi: 10.1016/j.cities.2020.103004.

[12] G. Masik, I. Sagań, and J. W. Scott, "Smart City strategies and new urban development policies in the Polish context," *Cities*, vol. 108, no. June 2019, p. 102970, 2021, doi: 10.1016/j.cities.2020.102970.

[13] P. Monachesi, "Shaping an alternative smart city discourse through Twitter: Amsterdam and the role of creative migrants," *Cities*, vol. 100, no. March, p. 102664, 2020, doi: 10.1016/j.cities.2020.102664.

[14] P. Neirotti, A. De Marco, A. C. Cagliano, G. Mangano, and F. Scorrano, "Current trends in smart city initiatives: Some stylised facts," *Cities*, vol. 38, pp. 25–36, 2014, doi: 10.1016/j.cities.2013.12.010.

[15] S. A. Bagloee, M. Heshmati, H. Dia, H. Ghaderi, C. Pettit, and M. Asadi, "Blockchain: The operating system of smart cities," *Cities*, vol. 112, no. May 2019, p. 103104, 2021, doi: 10.1016/j.cities.2021.103104.

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

About MAIT

Maharaja Agrasen Institute of Technology, Delhi is affiliated to Guru Gobind Singh Indraprastha University, Delhi, India (NBA Accredited, ISO 9000:2015 Certified, Graded as 'A' by the National Assessment and Accreditation Council, Bangalore). It was established in the year 1999 and since then it has taken long strides in providing quality Engineering and Technical education to students. The institute offers Bachelor's Degree in 5 disciplines of Engineering - Computer Science and Engineering (240 students' intake), Electronics and Communication Engineering, Electrical and Electronics Engineering, Information Technology, Mechanical and Automation Engineering (180 students' intake each) and Mechanical Engineering (60 students' intake) and Postgraduate degree in Master of Business Administration (180 students' intake). The Institute is approved by All India Council for Technical Education and affiliated to Guru Gobind Singh Indraprastha University, Delhi. Keeping in view the values ushered in by Maharaja Agrasen and the objectives envisioned by the Maharaja Agrasen Technical Education Society, the Institute aims to provide well qualified manpower to this nation.

About CSE

The Department of Computer Science and Engineering at Maharaja Agrasen Institute of Technology, Delhi was established in year 1999. It was formed to provide an outstanding research environment complemented by excellence in teaching. The Department offers B.Tech. degree affiliated to Guru Gobind Singh Indraprastha University, Delhi. The CSE department at MAIT has a team of experienced and highly qualified faculty who are committed to providing quality education to the students. the Computer Science and Engineering department at Maharaja Agrasen Institute of Technology offers a comprehensive learning experience, combining theoretical knowledge with practical skills. It prepares students for a successful career in the field of computer science and equips them with the necessary skills to adapt to the evolving technology landscape. The curriculum is designed to meet the current industry demands and trends, and department offers specializations in following areas:

1. Artificial Intelligence
2. Artificial Intelligence and Machine Learning
3. Machine Learning and Data Analytics
4. Full Stack Development

Department works with the vision "To be centre of excellence in education, research and technology transfer in the field of computer engineering and promote entrepreneurship and ethical values".

3rd International Symposium
on
Smart Cities Challenges, Technologies and Trends
(SCCTT-2024)
29th November 2024

Organized by

Department of Computer Science & Engineering
Maharaja Agrasen Institute of Technology
Agrasen Chowk, Sector-22, Rohini, New Delhi

Publication Partners

CEUR Workshop Proceedings
ceur-ws.org
ISSN 1613-0073

Partners

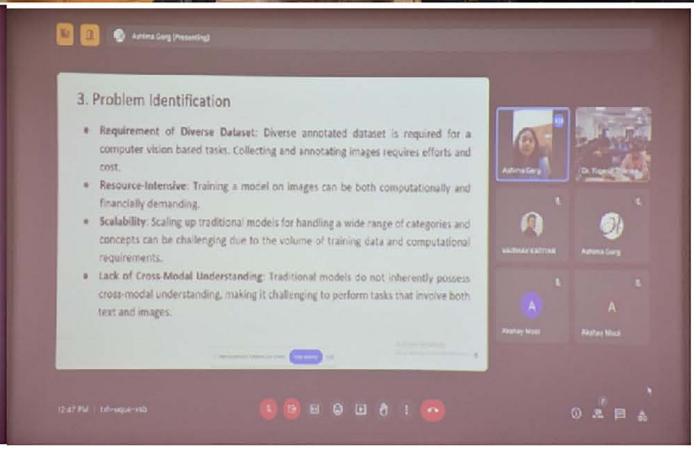
INSTITUTION'S INNOVATION COUNCIL
(Ministry of HRD Initiative)

MAIT-IIC
Technical Partner

Financial Partner


Maharaja Agrasen Institute of Technology
Department of Computer Science & Engineering
Welcomes
All Participants & Distinguished Guests
In
2nd International Symposium
On
Smart Cities Challenges Technologies and Trends (SCCTT-2023)
Date: 11th January 2024
Venue: MAIT Campus, Block-1, Lab 131/132
Organizing Chair
Mr. Ashish Sharma, Dr. Yogesh Sharma, Dr. Deepak Gupta, Dr. Sandeep Tayal


IMPORTANCE OF CYBER SECURITY IN SMART CITIES
Prof Anurag Jain
GGS Indraprastha University, Delhi
anurag@ipu.ac.in
9871739730


3. Problem Identification
Requirement of Diverse Dataset: Diverse annotated dataset is required for a computer vision based tasks. Collecting and annotating images requires efforts and cost.
Resource-Intensive: Training a model on images can be both computationally and financially demanding.
Scalability: Scaling up traditional models for handling a wide range of categories and concepts can be challenging due to the volume of training data and computational requirements.
Lack of Cross-Modal Understanding: Traditional models do not inherently possess cross-modal understanding, making it challenging to perform tasks that involve both text and images.

MAHARAJA AGRASEN INSTITUTE OF TECHNOLOGY
MAHARAJA AGRASEN CHOWK, SECTOR -22, ROHINI, DELHI-86
Ph.: 011-27582095, 8448186931 | E-mail: mait@mait.ac.in | Web.: www.mait.ac.in