VOL NO :01

INTERNATIONAL CONFERENCE ON NEURAL NEXUS AND SYNERGY

INNOVATION IN EMERGING TECHNOLOGIES

ADVANCING THE FUTURE OF COMPUTING

EDITORS:
Aiello Marco,University of Stuttgart,Germany
Rajkumar Buyya , University of Melbourne
Dr. Aruna Varanasi,SNIST
DR.H.Balaji,SNIST

Copyright & Disclaimer

Disclaimer:

Acknowledgment

As students, this book is a reflection of our collective curiosity, dedication, and the countless hours spent researching, analyzing, and refining our ideas. We are deeply grateful to our fellow students whose research papers have contributed to this compilation, each bringing unique perspectives and insights.

A special thank you to our faculty members and mentors who have been our guiding lights throughout this journey. Their unwavering support, constructive feedback, and encouragement have helped us push our boundaries and strive for excellence.

We also extend our appreciation for providing the platform to explore, learn, and contribute to the world of academic research. This opportunity has not only enriched our knowledge but also inspired us to keep questioning and innovating.

Preface/Introduction

The Proceedings of ICNSIET is a collection of research papers written by fellow students, reflecting our passion for learning and discovery. This book represents the hard work, curiosity, and dedication of students across different fields, each contributing unique insights to contemporary academic discussions.

Our goal is to foster a strong research culture among students and inspire others to explore, question, and innovate. The papers included were selected for their originality, depth, and relevance, highlighting diverse perspectives and emerging trends.

This compilation is more than just a book—it's a stepping stone for students who aspire to engage in meaningful research. We hope it serves as a valuable resource, sparking new ideas and encouraging future scholars to push the boundaries of knowledge. To everyone who contributed, supported, and believed in this endeavor.

—thank you for being part of this journey!

Conference Secretariat:

Sreenidhi Institute of Science & Technology

Yamnampet, Ghatkesar, Hyderabad – 501301, Telangana

Email: info@sreenidhi.edu.in

Contact:

- **Dr. H. Balaji** | 9666444100 | halavathbalaji@sreenidhi.edu.in
- **M Purnnaa Mannoj (Organizer)** | 7793960888
- icnsiet@sreenidhi.edu.in

Editorial Board Members

International Members

National Members

- **Dr. Halavath Balaji**

  Professor, Sreenidhi Institute of Science and Technology, Hyderabad

  Email: balaji.h@sreenidhi.edu.in

- **Dr. Chandrashekar Jatoth**

  Assistant Professor, Information Technology, NIT Raipur

  Email: chandrashekar.jatoth@gmail.com

- **Dr. Bukya Raju**

  Associate Professor, Department of CSE, NIT Warangal

  Email: raju@nitw.ac.in

- **Mr. Mettu Srinivas**

  Assistant Professor, Department of CSE, NIT Warangal

  Email: msv@nitw.ac.in

- **Dr. Rajneesh Rani**

  Assistant Professor, Department of CSE, NIT Jalandhar

  Email: ranir@nitj.ac.in

Committees:

Chief Patrons

- Dr. K. T. Mahi, Chairman, SNIST
- Mr. K. Abhijit Rao, Chief Executive Officer, SNIST

Patrons

- Dr. C. V. Tomy, Director, SNIST
- Dr. T. Ch. Siva Reddy, Principal, SNIST
- Dr. Aruna Varanasi, HOD – CSE, SNIST
- Dr. T. V. Rajini Kanth, HOD – CSE (AIML), SNIST
- Dr. T. Venkat Narayana Rao, HOD – CSE (IoT) , SNIST

- Dr. MD Jaffer Sadiq, HOD – CSE (Data Science) , SNIST
- Dr. K. Shirisha, HOD – CSE (Cyber Security) , SNIST
- Dr. Sunil Bhutada, HOD – IT, SNIST

Programme Chairs

- Dr. H. Balaji, Professor, SNIST
- Dr. Sreedhar Bhukya, Professor, SNIST
- Dr. Koppula Anup Kumar, Professor, SNIST
- Dr. Ch. Niranjan Kumar, Professor, SNIST
- Dr. G. Padmaja, Professor, SNIST

Conveners

- Dr. S. Dheeraj, Associate Professor, SNIST
- Dr. M. Rama Chandra, Associate Professor, SNIST
- Dr. T. Mamatha, Associate Professor, SNIST
- Dr. Preethi Jeevan, SNIST
- Dr. V. Sowmya Devi, SNIST
- Dr. G. Vani, SNIST
- Dr. K. C. Sreedhar, SNIST

General Chairs

- Mr. G. Vekanna, Assistant Professor, SNIST
- Mr. D. Rambabu, Assistant Professor, SNIST
- Mr. P. Durga Prasad, Assistant Professor, SNIST
- Mr. S. Chandra Shekar, Assistant Professor, SNIST
- Mr. M. Srikanth, Assistant Professor, SNIST

Honorary Chairs

- Mr. V. Rajanna, Senior Vice President & Global Head – Communications, Media and Technology Business Group, TCS
- Dr. S. Ramachandram, Vice Chancellor, Anurag University, Hyderabad

Table of Contents:

P02RW003

Artificial Intelligence- Predicting learning styles in Online for Primary Education

ShaikAbdul Khalandar, Yasmeen Kousar, Bhojeshwar, Mohd Rahmath, Raghuram.

*Affiliation: Sreenidhi Institute of Science and Technology

*Corresponding author: abdulshaik@sreenidhi.edu.in

Abstract:

In response to the COVID-19 pandemic, traditional educational institutions have swiftly shifted to online or remote learning. In addition to enhancing traditional learning methods and encouraging active learning, online learning guarantees educational continuity. Because there is little interaction between students and teachers and each student has a preferred learning style, online learning necessitates personalization in its implementation. Since every student has different learning preferences, one of the personalization strategies involves determining the learner's learning style. Several works have been proposed that use categorization approaches to identify learning styles. However, when students have a mix of learning styles or no dominating style, the existing detection methods are rendered useless. These days, e-learning platforms are an essential component of the curriculum. Students' confidence is boosted by the use of technology in the classroom since it makes content-based instruction more effective and efficient. Personalized learning systems place a strong focus on learning behavior and interest, and the curriculum is developed based on students' aptitude and baseline knowledge. It is a flexible teaching strategy that can be customized to meet the specific requirements of each learner. Every learner's demands are maximized by the customized learning approach. The study's objective is to develop and evaluate a predicted learning style using an AI model for use in an online learning environment for primary school students. The subjects were elementary school students. To support the idea of individualized learning, the AI model in the online learning portal was developed to provide educational resources according to the requirements of students' learning preferences. With the help of our innovative AI methodology, learning style prediction can now power collaborative filtering-based AI models.

The online learning platform may now offer content recommendations based on each student's unique learning style according to this AI system. On a scale of 1 to 5, the AI model performance test produced good results, with an average Root Mean Squared Error (RMSE) of 0.9035. Furthermore,

the findings of the t-test analysis between the pre-test and post-test scores showed that students' performance had improved.

Keywords: AI, Machine learning, deep learning, education, online system

Introduction:

To raise the standard of education, more work must be done to ensure that primary students learn effectively. Numerous research techniques are used to boost self-motivation and autonomous learning through an interactive learning environment in elementary schools. Previous studies have shown that effective learning would come from student participation in the learning proce0ss, such as through the use of the Team-Based Learning (TBL) approach, information technology improvements, and online learning. For students to participate actively in the learning process, the teacher must be able to select the right teaching strategy creatively. This can be done through utilizing technology to enhance the learning process and its results. Online learning, which promotes an active learning environment, is one tool that teachers can use to help students stay motivated and learn independently. As information technology evolves, online learning is still being developed today. The teacher has prepared a variety of learning tools that students can access by using online learning. The learning resources mentioned encompass a variety of formats, such as animated films, slideshows, e-books, and online articles. Taking into account the unique characteristics of each student is one benefit of adopting online learning.

Personalized learning is a term used frequently in the education profession to describe a learning approach that takes into account the differences among pupils. One way to think about personalized learning is as an intense and thorough integration of these concepts throughout all subject areas and values in schools. With the advent of technology support, this system has recently become more practical. Students who require guidance and training can receive it through personalized learning. Enhanced subject cover and an adaptable route for student achievement are two additional advantages of customized instruction. Consequently, instead of having a student take a course in calculus, statistics, or accounting as would normally happen, mastery-based learning environments could enable students to focus on specific topics within each subject, tailored to their interests or the requirements of their intended career path. Students are currently making significant progress in the area of personalized learning, especially with digitalized personal learning that includes prefabricated curricula, tests, and continuous data collection. AI has the potential to enable personalized learning on a digital platform.

One among them is employed to ascertain the student in question's preferred method of learning. Learning styles can be characterized as a student's inclination toward or effective means of absorbing and communicating knowledge, as evidenced by their preferred activities, speech patterns, learning strategies, assignment completion techniques, and interpersonal interactions.

According to the review above, encouraging students to participate in their education is one way to raise the standard of instruction. By selecting relevant learning resources and keeping student learning styles in mind, teachers can make these students more engaged. To increase student engagement and capitalize on information technology breakthroughs, this research attempts to create and assess the effects of an AI-based online learning portal. The AI algorithm built into the online learning portal was designed to recommend relevant reading material based on the student's learning requirements. To do this, we developed a technique that enables an AI model based on collaborative filtering to be driven by learning style prediction. The method makes it possible for the AI model to predict learning styles in an unsupervised classification setting. Our suggested algorithm is not supervised by the learning style that is decided by humans, in contrast to AI-based approaches in earlier studies. A soft max prediction layer is added to the latent learning material vectors in the collaborative filtering framework to make the suggested algorithm unsupervised. Because the suggested algorithm is unsupervised, human bias is removed when suggesting materials according to the learning preferences of the student. It will take more effort to improve the quality of instruction if primary school pupils are to learn effectively. Primary schools employ diverse research methodologies to establish a captivating educational setting that enhances student motivation and self-directed learning.

Active student participation in the learning process, as achieved by the use of the Team-Based Learning (TBL) approach, technological advancements, and online learning, would lead to successful learning, according to the findings of previous studies. As the learning manager, the teacher must be able to choose the most effective teaching method for the students to actively engage in the learning process. To improve the outcomes of the educational process, this can be achieved by leveraging technological advancements. Instructors can foster and sustain students and motivation for independent learning by utilizing technology, like online learning, which creates an interactive learning environment. In reaction to advancements in information technology, online education is currently being developed. Through the use of online learning, students have access to a range of learning resources created by the instructor. A variety of learning resources have been mentioned, such as animated movies, slideshows, e-books, video lectures, and online articles. To make these learning resources easier for students to access, teachers can organize them in an online learning management system (LMS) portal such as Moodle. The consideration of unique student characteristics is one advantage of online learning.

1.1 Existing System

The field of learning enhancement has led to the development of customized learning, specifically mechanized modified learning that includes pre-packaged courses, ongoing data collection, and assessment. Personalized learning can be provided through digital platforms with artificial intelligence (AI). One of them is used to determine and understand the preferred learning style.

Learning styles are the preferred method or approach by which students successfully acquire and impart knowledge. Examples of learning styles include speaking styles, learning systems, how to fulfill obligations, how to help others and other embedded works.

1.2 Proposed System

The purpose of this work is to design and evaluate the implementation of an Internet learning gateway based on artificial intelligence to increase student engagement and reap the benefits of data technology advancements. Using the student's preferred learning methods as a guide, the artificial intelligence model for the electronic learning entry was created to recommend appropriate learning materials. To accomplish this, we created a cutting-edge tactic that made it possible to coordinate the development of a workable isolating-based AI model predicated on learning style. For now, the method allows the AI model to function independently as a learning style assumption plan model. Unlike previous artificial intelligence-based techniques, our proposed computation is not dependent on the learning strategy individuals select. The independence of the proposed strategy is created by layering a soft max assumption layer between the inert learning material vectors in the practical filtering structure. Because of its autonomy, the suggested assessment reduces the inclination of humans to consider the learning style of the understudy when identifying assets. This artificial intelligence framework could be used by the online learning portal to deliver resources that are tailored to individual students's preferred learning styles and methods.

Literature survey:

This article was created to help lecturers use the Team-Based Learning (TBL) teaching and learning method to build a teaching portfolio that will allow them to maximize the advantages of blended learning—a teaching approach that combines online and in-person training. Research indicates that

TBL can improve student collaboration and active learning, two abilities that could assist in mitigating some of the disadvantages of blended learning. The development of a blended teaching portfolio for an international human resource management course included the following components: a course overview, graduate competency, syllabus, course materials, teaching scenario, reading assurance test, midterm/final exams, student assignments, evaluation of learning outcomes, and a course quality improvement sheet. Every item was constructed using the course components. This study represents a first step in providing comparative quantitative empirical evidence for the value of TBL for fostering continuous improvement in the learning process, as one way to improve student learning outcomes in undergraduate health science curriculum.

Material and method:

3.1. Online learning

Online learning, is innovative online instruction delivered through the Internet to distant learners. Online learning currently dominates the use of modern technology in electronic-based learning or e-learning. Information technology serves as a conduit for communication between educators, learners, and course materials in online learning. Students, educators, and educational materials can all communicate with one another and with learning activities through the Internet. But to put it into practice, you need a tool that integrates with an educational management system, which oversees an online computer system where students who have had teacher supervision and guidance progress. Online learning communities can be successfully formed through Moodle, an open-source course application based on an LMS. Moodle offers a combination of static course materials (written pages, graphics, and websites) and interactive course materials (lessons, surveys, quizzes, assignments, etc.). Forums, chats, and other activities are part of the module. Databases (MySQL, PostgreSQL, Oracle, etc.) are used by Moodle for storing logs, and they are more versatile and strong than text log files. Data on high-level usage that is gathered in the LMS can be accessed and collected through Moodle's databases. When assessing student learning performance, teachers or instructors can use the Moodle application to display statistical data in the form of a graded or rated assessment scale. Personalized learning can be supported by LMS by identifying the most appropriate learning style based on the results of the online learning assessment.

The Association of Higher Education (ASHE) has reported that the efficacy of online learning approaches is superior to traditional learning, which necessitates in-person instruction and classroom settings. According to the ASHE report, online learning exhibits higher levels of integrative thinking, reflective learning, and order thinking skills than traditional classroom based learning in the US K–12 education system.

3.2. Primary education using online learning

An increasing number of elementary schools are using online education as a substitute for traditional classroom instruction. Implementing online learning in primary schools has several pros and cons as well. The advantages may enhance students' capacity for teamwork and communication in a classroom setting. But to reap the rewards, educators must help their pupils retain their motivation and independence in the classroom. Online learning activities in primary schools typically involve the use of platforms like social networking sites (SNSs) and virtual learning environments (VLEs) or the open-source Moodle platform. For example, Singaporean and Canadian first-graders utilize Twitter to collaborate on arithmetic problems. In the meantime, second-graders in primary school review and assess their learning objectives using Twitter. However, children under the age of 13 are not allowed to use SNSs due to the COPPA (Children's Online Privacy Protection Act) regulation. As a result, it is preferable to use alternative online learning platforms. One of the platforms is Edmodo, a private educational social networking site that is safe and secure for use in elementary school online learning activities. Among the many features offered by Edmodo that facilitate online learning for teachers is the ability to share learning resources, participate in discussions, and turn in homework. Google Classroom and Moodle/VLE are two examples of platforms that are frequently used for online education in addition to social networking sites. Research on improving the quality of student learning in a realistic setting has made extensive use of Moodle/VLE in online learning.

3.3. Individualized expertise and exploration mode

Personalized learning is an instructional strategy that supports students' mastery of the material by tailoring it to their interests and abilities. Additionally, the learning system offers guidelines and instructions to help students' capacities based on their requirements. Personalized learning is now a definite area of learning growth for students, particularly digital individualized learning with prepackaged curricula, assessment, and ongoing data collection. Personalized learning is currently made easy with the help of a digital platform that supports the learning styles of preschoolers and primary school students with a variety of media, including audio, videos, and illustrated images.

The tendency or effective manner in which students assimilate and convey knowledge can be characterized by their learning style. This can be seen in their preferred activities, speech patterns, learning strategies, assignment completion techniques, and interpersonal interactions. It is believed that learning styles based on personality can assist students with different capacities and habits in observing learning environments to enhance their cognitive abilities a study conducted by Cassidy, learning styles that correlate to the interests and habits of primary school kids were predicted through the use of modules and evaluations related to cognitive personality style performance and learning. A model instrument like the Felder-Silverman learning/teaching style and the VARK(Visual, Aural, Read/Write, and Kinesthetic) learning style defined individual student characteristics based on their ability to perceive and process information. To assess both instrument models, the student's preferred method of learning was monitored through the use of visual, aural, and kinesthetic perceptual modalities.

3.4. AI model for exploration style accuracy and Recommendation of Learning Materials

To provide recommendations for learning materials based on learning styles, two separate AI models are usually needed. Use the first model to predict the learning style of a learner. The second model bases its recommendations for learning resources on the first model's estimate of learning styles. Supervised learning models are usually the initial ones. For the most part, collaborative filtering algorithms form the foundation of the second model. Two models for a single task can add up errors to the final output because each model has its errors. As a result, we created a single model that uses mutual latent information that is extracted from data to perform both prediction and recommendation. We adjusted a collaborative filtering model which was designed for recommendation tasks to predict the student's learning style to accomplish this goal. By asking users to rate the educational resources they have accessed, standard collaborative filtering gathers information about their interests. Six learning resources are used in collaborative filtering, by the rating data captured in. Most of the time, users provide values between 1 and 5 for each element. The rating data can be trained into an AI model to provide a recommendation. A user's predicted rating for learning material that they haven't yet rated is used to formulate the recommendation. The Matrix Factorization (MF) technique, is the basis for the majority of AI models that are currently used to predict rating data. For this paper, we will refer to this model as the standard MF-based model. Our recommendation system usually uses this model as a model. By allocating a latent vector of equal size to every user and learning resource, it models the rating left by users. The user's rating of a piece of learning content is predicted by multiplying the user vector and the corresponding learning content. The index of a user in the rating matrix serves as the model's input when it is in inference mode. The rating in the corresponding row is the output, and it can be understood as the model's prediction of the rating that the user will probably provide. To train the model to generate a precise forecast, the user vectors and learning materials are fitted to the known rating using a gradient descent approach. This needs to be contrasted with the other popular MF variations, such as the model, which used Singular Value Decomposition (SVD) to derive the material and user latent vectors analytically. Through vector multiplication for rating prediction, this system efficiently feeds the learning style information to the latent user vectors. Ultimately, it instructs each latent user vector component to represent the student's desired learning style. The model can classify the user's learning style by taking the argmax of the latent user vector. Put differently, our suggested approach makes it possible for the standard MF model to function as an unsupervised learning style classification model. Unlike other learning style prediction models that rely on supervised learning, our suggested model doesn't require information about students' learning styles that be gathered through a questionnaire.

Method:

Version 3.9.3 of the Moodle framework was used to create the online learning portal utilized in this investigation. After that, this portal was set up on a server that had an Intel Xeon E5-2620 processor, an NVIDIA Tesla K40 graphics processing unit, 32 GB of RAM, and 4 TB of RAID 5 configured storage media. A system based on Docker virtualization technology and cloud computing was used for the installation. The initial step involved building a container with the most recent Moodle framework. There were two main stages to the research process in this study, which are depicted in Figure 9: (1) the online learning session and (2) the AI modeling.

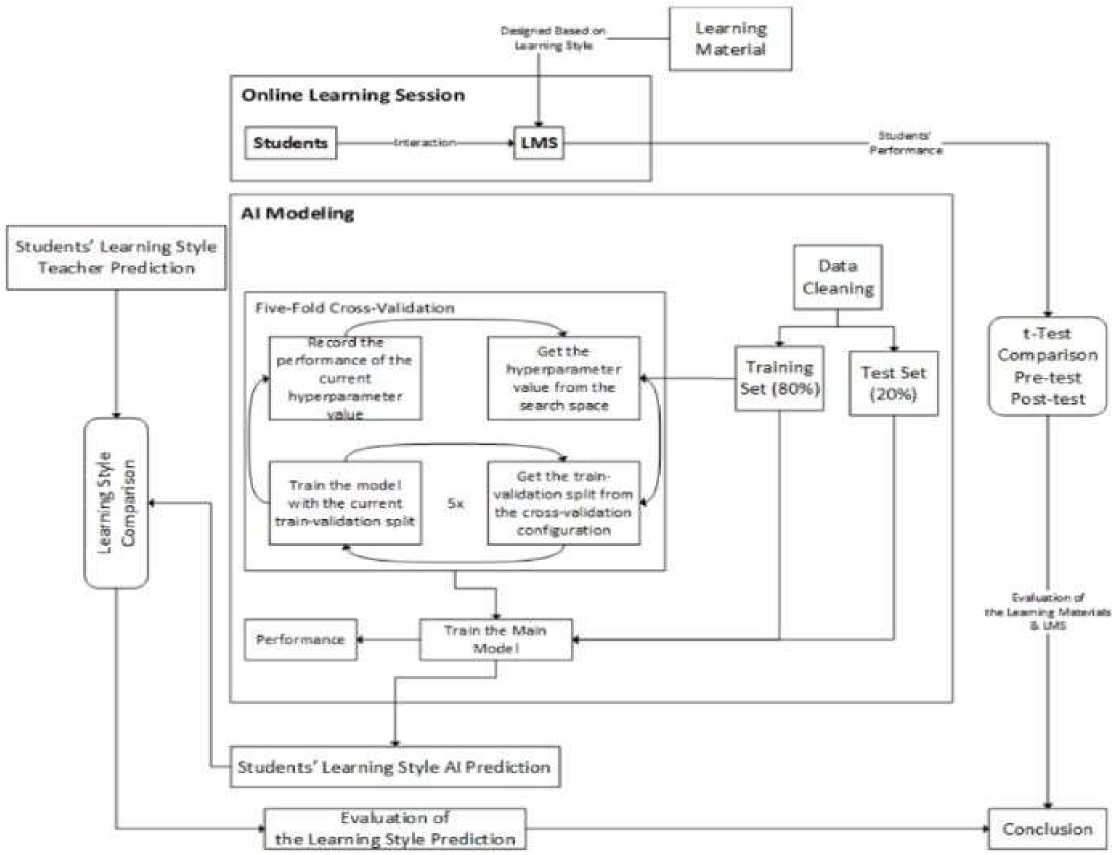


Figure 1: System Architecture

Six learning resources were provided to the students during the online learning session to help them understand the concept of numbers. Based on the dePorter et al. variant of the VARK model, the learning materials were customized and verified by specialists to be appropriate for the visual, auditory, and kinesthetic learning styles [48]. Following the lesson, the students were asked to rank their preferences for the six learning resources on a scale of 1 to 5. To train the AI, these ratings were combined into a ranking matrix. In addition, we administered a pre-and post test to every student to gauge the impact of utilizing the online learning portal. There were ten multiple-choice questions on the test. A paired t-test was used to examine the pre-and post-test data to determine whether the online learning portal could raise the performance of the pupils.

Results:

The outcome of the t-test is based on the pre-and post-test data. On a scale of 0 to 10, the student's average score increased by 0.49 points. With a p-value less than 0.05, the improvement was statistically significant.

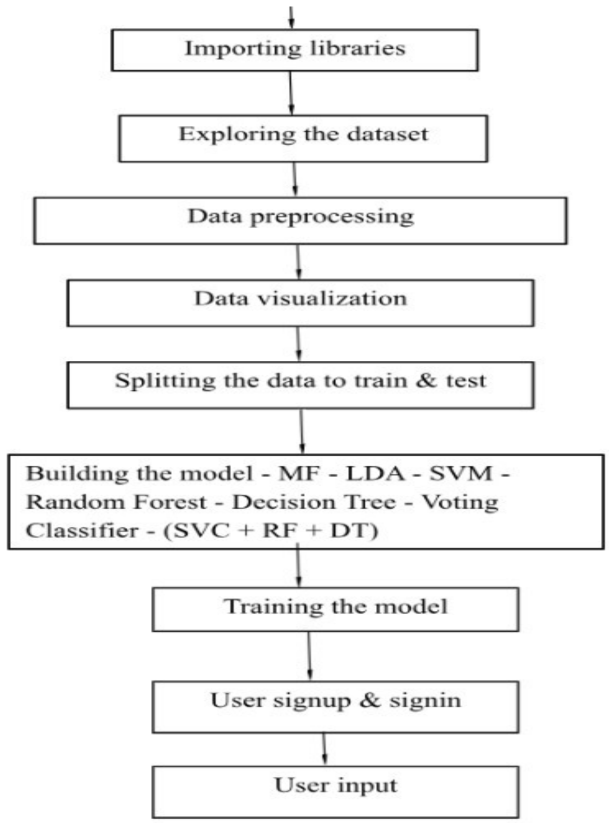


Figure 2: Proposed Algorithm flow

In the meantime, Table 3 provides information on the suggested model's performance. Based on a rating scale of 1 to 5, the test performance was deemed satisfactory overall, with an average RMSE of 0.9035. The suggested model outperformed the baseline MF-based model, exhibiting an RMSE value that was 0.0313 lower.

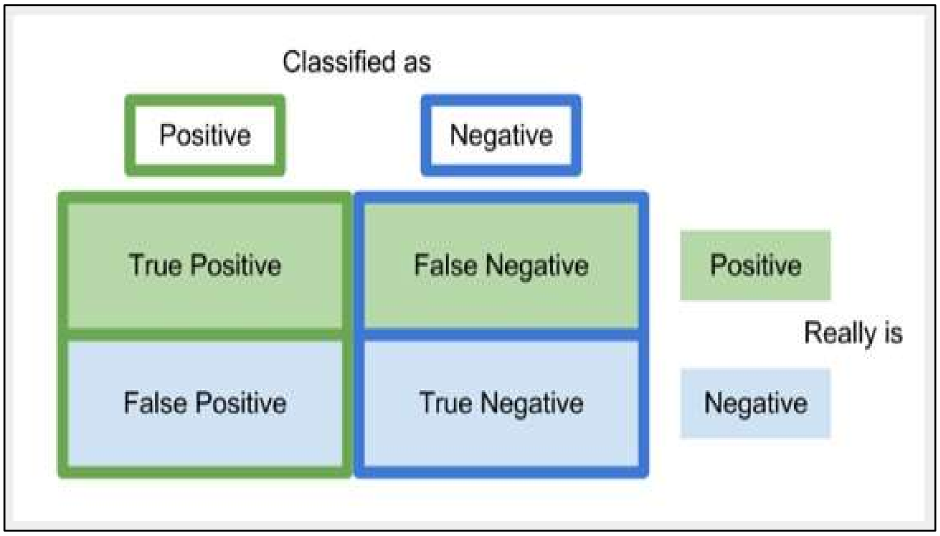


Figure 3: Confusion matrix

Therefore, the suggested model performs better as a recommendation system and has a greater capacity to predict the learning preferences of the students. Following the assessment of the suggested model's performance, just 35.13 percent of the teacher's predictions were shared by the AI. Specifically, we emphasized that a significant proportion of students whose teachers predicted they would have a visual learning style were predicted by AI to have auditory (21.62%) and kinesthetic (17.57%) learning styles.

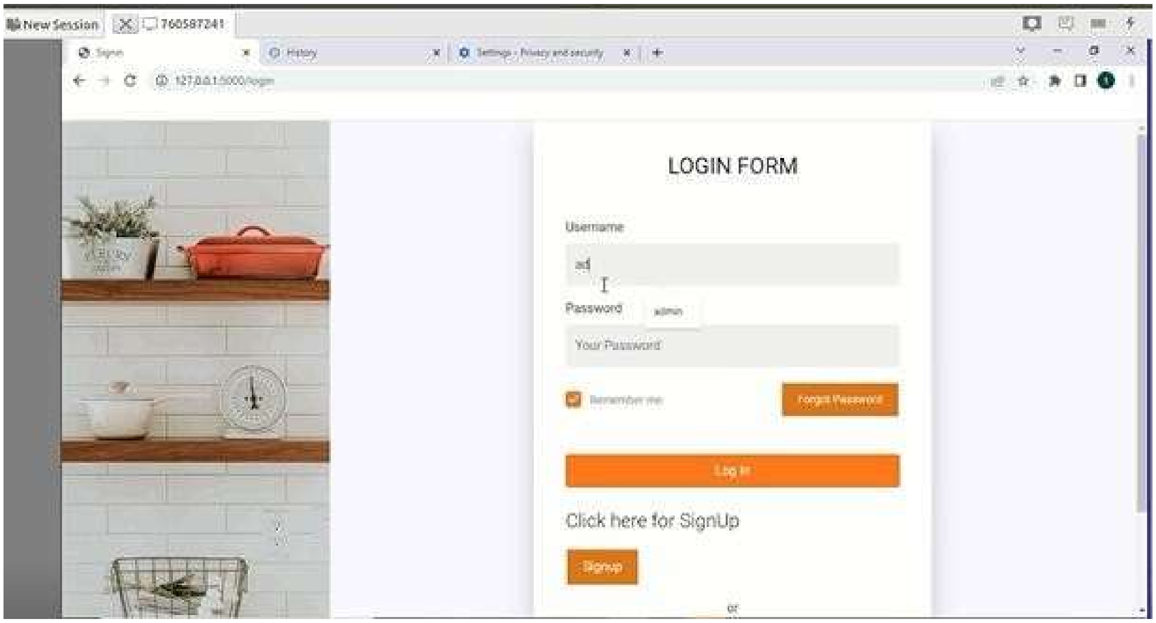
$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

Other metrics such as F1-score and accuracy are computed in Eq. 8 and Eq. 9.

$$\text{F1-score} = 2 * \frac{(p*r)}{(p+r)} \quad (8)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$



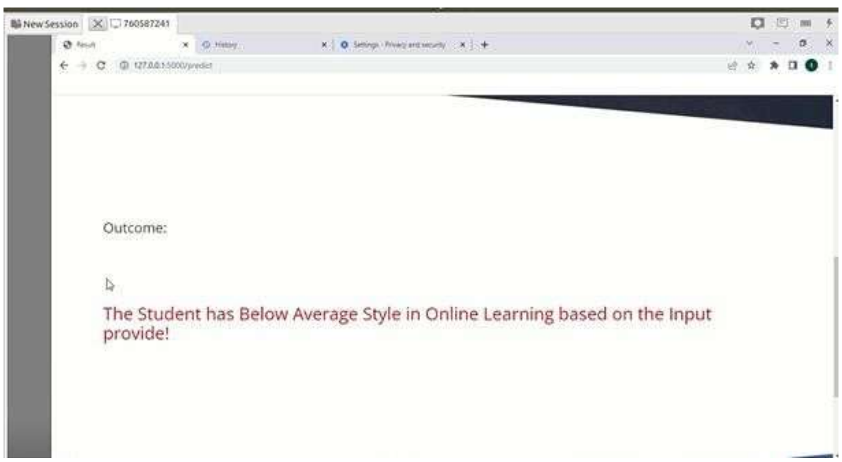


Figure 4: Sample Output

Discussion:

The result shows that 322 students were among the subjects that were involved. Still, only 269 students have access to the LMS designed specifically for this research project. The common issues that led to this problem were the student's limited time and available tools. Some students continued to rely on their parents' mobile devices, which they did not always have access to, according to the teachers' reports. 172 out of 269 students can log into the LMS systemby the teacher's instructions. Some students were unable to follow the directions provided through written guidance or video instruction, according to the conversations with teachers. It is important to note that our suggested algorithm's features differ significantly from those of the algorithms in the prior study when evaluating the study's outcome. Before producing a material recommendation, the majority of earlier algorithms depend on the application of supervised learning models to forecast students' learning preferences. Using manual questionnaire data collection, supervised learning models were primarily trained on learning style ground truths, which necessitated tedious labor. As an alternative to manual collection, clustering techniques were used in other studies to generate the ground truths.

As useful as the clustering algorithm is, the ground truths it produces are dubious in their validity. In most previous studies, the material recommendation was generated by handcrafted rules based on the supervised model's prediction of learning style. The subjective nature of the rule designer may be inherited by this method. AI estimated that the students' visual, auditory, and kinesthetic senses are shared by 24.32%, 39.19%, and 36.49% of the total. The widespread belief that 65% of people have a visual learning style could be the reason for the high percentage of visual learning style predictions made by educators. Despite the lack of an original academic paper with scientific proof, this fact is widely disseminated in the media. Teachers can observe that a greater proportion of kinesthetic learners compared to auditory learners was present. This finding is consistent with the distribution of AI predictions, but it deviates from the widely accepted notion that 30% and 5% of people are auditory learners and kinesthetic learners, respectively. Based on their observation, teachers may have made a prediction that favors kinesthetic learning styles, which is not hazarded by the unproven widespread fact. It's interesting to note that the AI classified a sizable portion of students as kinesthetic.

This may indicate that the AI prediction more accurately represents the true distribution of learning styles among the population of primary school students. It is important to note, though, that this comparison result did not come from a standard analysis used in psychological research. Therefore, it is not appropriate to consider the outcome to be definitive from a psychological standpoint.

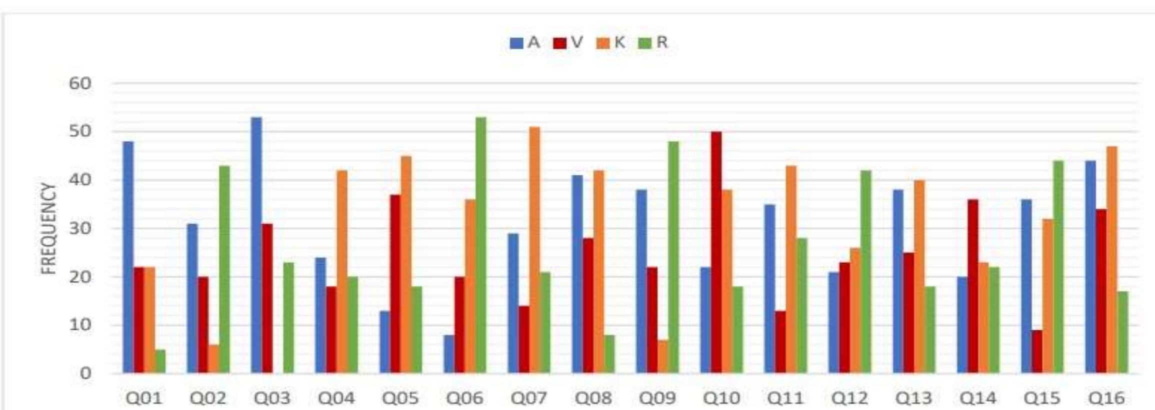


Figure 5: f learning styles distribution

Model	Precision	Recall	F1-Score	Accuracy	AUC
SVM	0.368	0.421	0.377	0.479	0.654
NN	0.405	0.403	0.402	0.423	0.695
RF	0.505	0.506	0.503	0.535	0.713
DT	0.413	0.417	0.408	0.451	0.632
KNN	0.423	0.422	0.398	0.451	0.682

Figure 6: Evaluation results

Conclusion:

Predictive learning styles can now drive collaborative-filtering-based AI models thanks to the novel method we employed in this work. With an average RMSE of 0.9035 on a rating scale of 1 to 5, the testing performance of this AI model was satisfactory. With an RMSE value that was 0.0313 less than the conventional MF-based model, it performed better. The suggested approach not only performs better, but because it doesn't use supervised learning, it also does away with the requirement for human learning style ground truth. Based on the study's findings, we were able to observe a potential change in primary school students' learning styles when it came to application online learning. Teachers should be concerned about this shift because it affects the learning environments of their students. Teachers should use the online learning platform we developed for this study to help them explore learning materials that are tailored to the learning styles of their students. This will require more effort on their part. The improvement in the student's performance in this study indicates that the online learning platform in this study is beneficial not only for the teachers but also for the students.

References:

[1] M. S. Hasibuan, L. E. Nugroho, and P. I. Santosa, “Model detecting learning styles with an artificial neural network,” J. Technol. Sci. Educ., vol. 9, no. 1, pp. 85–95, 2019.

[2] E. Goh and M. Sigala, “Integrating Information & Communication Technologies (ICT) into classroom instruction: teaching tips for hospitality educators from a diffusion of innovation approach,” J. Teach. Travel Tour., vol. 20, no. 2, pp. 156–165, Apr. 2020.

[3] J. L. Moore, C. Dickson-Deane, and K. Galyen, "E-Learning, online learning, and distance learning environments: Are they the same?" Internet High. Educ., vol. 14, no. 2, pp. 129–135, 2011.

[4] J. Bernard, T. W. Chang, E. Popescu, and S. Graf, "Learning style Identifier: Improving the precision of learning style identification through computational intelligence algorithms," Expert Syst. Appl., vol. 75, pp. 94–108, 2017.

[5] hopefullyfinaldocx.docx: An empirical study of a blended learning course," Br. J. Educ. Technol., vol. 47, no. 2, pp. 257–278, Mar. 2016.

[6] M. Rais, F. Aryani, and A. S. Ahmar, "The influence of the inquiry learning model and learning style on the drawing technique of students," Glob. J. Eng. Educ., vol. 20, no. 1, pp. 64– 68, 2018.

[7] J. W. Keefe, Learning style theory and practice. 1987.

[8] L. Fallan, "Quality reform: Personality type, preferred learning style and majors in a business school," Qual. High. Educ., vol. 12, no. 2, pp. 193–206, Jul. 2006

[9] T. R. Frame, S. M. Cailor, R. J. Gryka, A. M. Chen, M. E. Kiersma, and L. Sheppard, "Student perceptions of team-based learning vs traditional lecture-based learning," Am. J. Pharm. Educ., vol. 79, no. 4, 2015. L. Jerome, "Ways of learning: Learning theories and learning styles in the classroom," Eur. J. Teach. Educ., vol. 33, no. 2, pp. 219–221, 2010.

[10] M. Platsidou and P. Metallidou, "Validity and reliability issues of two learning style inventories in a Greek sample," Int. J. Teach. Learn. High. Educ., vol. 20, no. 3, pp. 324–335, 2009.

[11] N. D. Fleming and D. Baume, "Learning Styles Again: VARKing up the right tree!," Educ. Dev. SEDA Ltd, vol. 7, no. 4, pp. 4–7, 2006.

[12] D. C. Kayes, "Internal validity and reliability of Kolb's learning style inventory version 3 (1999)," J. Bus. Psychol., vol. 20, no. 2, pp. 249–257, 2005.

[13] F. Coffield, D. Moseley, E. Hall, and K. Ecclestone, "Learning styles and pedagogy in post16 learning: A systematic and critical review. National Centre for Vocational Education Research (NCVER).," 2004.

[14] F. Rasheed and A. Wahid, "Learning style detection in E-learning systems using machine learning techniques," Expert Syst. Appl., vol. 174, 2021.

[15] O. El Aissaoui, Y. El Alami El Madani, L. Oughdir, and Y. El Allioui, "A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments," Educ. Inf. Technol., vol. 24, no. 3, pp. 1943–1959, May 2019.

[16] O. El Aissaoui, Y. E. A. El Madani, L. Oughdir, and Y. El Allioui, "Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles," in Procedia Computer Science, 2019, vol. 148, pp. 87–96.

[17] and G. A. M. Kalhoro, Aijaz Ahmed, SaminaRajper, "Detection of E-Learners' Learning Styles: An Automatic Approach using Decision Tree," Int. J. Comput. Sci. Inf. Secur., vol. 14, no. 8, p. 420, 2016.

[18] O. Pantho, "Using Decision Tree C4 . 5 Algorithm to Predict VARK Learning Styles," Int. J. Comput. Internet Manag., vol. 24, no. 2, pp. 58–63, 2016.

[19] B. Hmedna, A. El Mezouary, and O. Baz, "Identifying and tracking learning styles in MOOCs: A neural networks approach," Adv. Intell. Syst. Comput., vol. 520, no. 2, pp. 125–134, 2017.

[20] Azzi, A. Jeghal, A. Radouane, A. Yahyaouy, and H. Tairi, "A robust classification to predict learning styles in adaptive E-learning systems," Educ. Inf. Technol., vol. 25, no. 1, pp. 437–448, Jan. 2020.

P03RW003

Water Quality Classification using Deep Learning.

Donthula mamatha,Halavath Balaji

Department of Computer Science & Engineering ,

Sreenidhi Institute of Science and Technology,Hyderabad,Telangana, India.

mamathait76@gmail.com,halavathbalaji@sreenidhi.edu.in.

Abstract:

Water quality evaluation is essential to environmental management and monitoring.Traditional methods often rely on manual inspection and It may be expensive and taking time. Deep learning has become a potent method for picture categorization in recent years. Involving an evaluation of the water purity. This research VGG 19, an Convolutional water purity.Our suggestion is a channel-specifictions attention gate integrates a channelwise attention gate to focus on relevant features in the input images. Our approach includes preprocessing steps such as background elimination, removal of non-essential features, image enhancement, and noise removal to improve classification accuracy.We experimented with a collection of water surfaceachieving an accuracy of 97.5%, outperforming previous approaches. The results demonstrate the effectiveness of deep learning models in water quality classification and highlight the importance of preprocessing techniques in improving classification performance.

Keywords: Convolution Neural Network,VGG19,Deep Learning.

1.Introduction:

Water quality is a critical factor in maintaining environmental health as well as guaranteeing access to pure, safe drinking water. Traditional methods of water quality assessment often involve time-consuming and expensive laboratory tests. However, recent advancements deep learning, especially with regard to computer vision have shown promise automating process of water quality assessment through image analysis.

Deep learning models were effectively implemented, including convolutional neural networks (CNNs).various images classification job.In this study, we focus on the application of the VGG 19 architecture, a widely used CNN, for water quality classification.

We propose a hierarchical attention model that incorporates a channel-wise attention gate the models capacity to concentrate on pertinent features in water surface images. Preprocessing contributes significantly to the deep learning models' ability to classify water quality. Techniques such as background elimination, removal of non-essential features, image enhancement, and noise removal are essential for improving the quality of input data and enhancing the model's performance. In this paper, we present our approach to water quality classification using deep learning and VGG 19. We describe the preprocessing steps and the architecture of our hierarchical attention model. We also give the outcome of experiments conducted on a collection of water surface images, demonstrating the effectiveness of our approach in achieving high classification accuracy.

2.Literature review:

1.Title: “Deep Learning-Integrated Quality of Water Monitor and Event Identification in the Internet of Things” Authors:Wang, Y., Liu, Y., Yuan, Y., et al.

Published in: IEEE Transactions on Industrial

Informatics (2019)

Summary:This paper provides an extensive analysis of the use of deep learning methods IOT founded water quality monitoring actual water quality parameter classification is achieved by the authors' proposed system,which combines recurrent and convolutional neural networks (RNNs)The goal for the research is to create an integrated Internet of Things platform for the evaluation of water quality, with a focus on the possibility for early event detection and prompt reactions to episodes of water contamination.

2.Title: "Deep Neural Networks for Water Quality Monitoring" Authors: Carrasquilla, J., Konidaris, T., & Ear, E.

Published : The AAAI Conference on Artificial Intelligence 2015 Proceedings

Summary: Deep neural network applications for water quality monitoring is examined in this research. In order to forecast water quality metrics using past data, the authors suggest using a deep learning algorithm. To increase forecast precision and evaluate accuracy for water quality,the study shows how deep learning can capture intricate patterns in time series data related to water quality.

3.Title: ”A Meta-Analysis and Survey of Deep Learning in Remote Sensing”.

Authors: Zhao, Y., & Du, S.

Published:”Photogrammetry and Remote Sensing Journal of ISPRS (2020)”

Summary: This research offers a Deep learning meta-analysis applications about remote sensingwhich is highly important to water quality assessment even if it is not specifically focused on water quality. It examines several deep learning models and how well they work in remote sensing applications, providing insights into the possibilities of combining deep learning with satellite and remote sensing data for the monitoring and categorization of water quality.

4.Title: "Water Quality Monitoring Using Remote Sensing and Machine Learning: A Review"

Author: Hamed, Y., & Khan, A. Published in: Remote Sensing (2019)

Summary: The application of deep learning and remote sensing to water quality monitoring is covered in this review paper. An overview of the most recent methods for assessing water quality using data from satellites and aircraft is given. In order to categorize water quality and identify pollution events, the article investigates the possibility of integrating remote sensing data with cutting-edge machine learning algorithms.

5.Title: "An Approach to Automatic Water Quality Assessment Using Deep Learning"

Authors: Raut, D. N., & Gaur, R.

Published : Procedia Computer Science (2018).

Summary: This work proposes a deep learning technique for automated evaluation of water quality. To classify the quality of water samples based on pictures of bodies of water, the authors employ a convolutional neural network (CNN). In this regard, the study highlights the benefits of deep learning and it potential for real-time, image-based categorization of water quality.

3.Materials and methods:

Here is a sample materials and methods section for a water quality classification using deep learning study.

3.1.Materials:

3.1. 1.Data Collection

IN this project we use water images dataset collected from google online images which has three categories of five categories. Pixel values form images are taken as input and labels are used as output and each folder has 50 images. which are used for training.Images are standardized to a fixed image size 224 × 224 pixels.

3.1.2.Data preprocessing

Background elimination: Removed backgrounds to focus on the water surface.

Elimination of non-essential objects: Removed non-essential objects like boats, people, or animals from the images.

Image enhancement: Enhanced the contrast and brightness of the images to improve visibility.

Noise removal: Applied noise reduction techniques to improve image quality

3.1.3.Data splitting

The stages that involve Using photos that are captured, the following steps are involved in categorizing the water quality: These are: Dataset training, validation, testing the collection of data, or the gathering of specimens of water from various locations and water surface photographs, is the initial step in the project. Google earth photos and smartphone photos may be among those. The project model is trained using the gathered data and the water sample data. In the experiment's data pre-processing section, gathering data and dataset training are completed. A well-trained dataset is necessary to achieve excellent accuracy. The dataset's validation establishes the correctness of the provided water images samples, which were gathered from various locations and classified into two floders. Trained dataset, which also has two subfolders with photographs of safe and harmful water, is compared to the two validation dataset folders.

3.2.CNN algorithm:

In computer vision, For deep learning, convolutional neural networks (CNNs) are a popular choice. The field of artificial intelligence, referred to as "computer vision," enables computers to interpret and process visual information such as photographs, popular kind of deep neural network for interpreting visual images is a convolutional neural network, or CNN. It is made to recognize the geographical component hierarchy from input photos naturally and flexible.

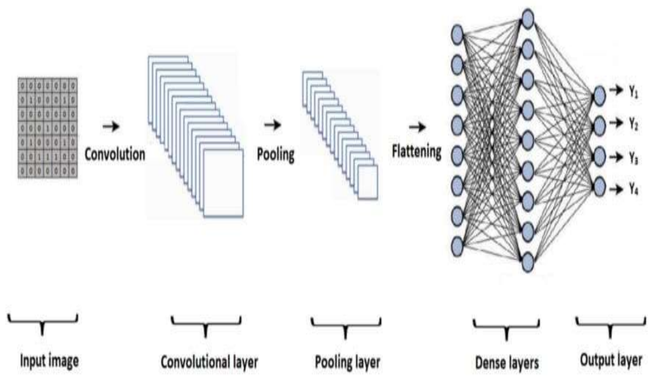


Fig:CNN ARCHITECTURE

Here's an explanation of the key components of a typical CNN architecture:

3.2.1.Input Layer

After accepting the input image(s), that layer sends them to the following layer. Usually, a matrix of dimensions is used to display the input picture, with every dimension denoting a different feature in that picture.

(e.g., width, height, color channels). An input image of water surface, typically with dimensions such as 224x224 pixels and 3 color channels (RGB).

3.2.2.Convolutional Layer

Applying an array of filters, also called kernels, to a given picture or image is what the convolutional layer does. A few features, including edges, surfaces, or forms, are extracted from an input by every filter. The filter used moves across the provided image(s) via the convolution process, computing the product of dots among its filter weights and the correct input pixel numbers. For every filter, this procedure creates a feature map that captures several faces of the incoming data.To extract pattern space from the input image, apply many convolutional layers with tiny filter sizes (e.g., 3x3). These layers may be stacked to expand the neural system and enable is to pick up additional aspects.

3.2.3.Activation Function

Following the steps of convolution, feature maps are subjected element-by-element to an activation function (e.g., ReLU, or Fixed Linear Unit). By adding nonlinearity, the model can recognize complex structures in the information. Following every convolutional layer, add irregularities within the structure by applying an irregular activation function, such as ReLU.

3.2.4.Pooling Layer

The pooling layer down samples generate maps of features using the convolution layer, which are down sampled via the layer for pooling. In doing so, essential data is preserved while the geographic dimensions of the map of features are decreased. Typical and maximum pooling are two popular pooling procedures.

3.2.5.Dense layer

After multiple layers of convolution and pooling, the neural network's excellent processing is carried out using fully linked layers. Totally linked neurons may recognize complex structures in information since they are coupled to every activation from the level above them.

3.2.6.Output Layer

The CNN's ultimate output is generated by the output layer. The job at hand determines how many neurons belong to the output layer. The final layer usually outputs probable outcomes for every class using a soft max activation function for classification tasks. In conclusion, convolutional and pooling layers in CNNs are made to take characteristics out of the input pictures; layers that are completely linked are then added for deeper reasoning. CNNs may accomplish state-of-the-art outcomes in several computer vision applications object recognition, picture grouping, and image categorization, thanks to their design.

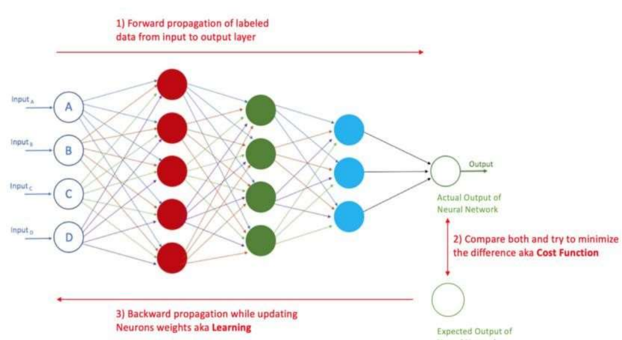


Figure: Topology of a feed-forward model with two hidden layers

In order to establish the ideal weights and activate just the most potent and categorize neurons for categorization, forward and backward propagation iteratively go through all of the training data in the network.

3.3.1.Epochs

One pass constitutes an epoch. The entire dataset through the neural network. During training, the model goes through many epochs to learn from the data.

3.3.2. Forward Propagation

In each epoch, forward propagation calculates the predicted output of the network for each input sample. The loss function compares these predictions to the actual labels, measuring the discrepancy between them.

3.3.3.Backward Propagation

After forward propagation, backward propagation determines the range of the reduction function for the weight. An optimization method updates the weights based on this gradient-like descent. Backpropagation helps the network learn from its mistakes and improve its predictions.

3.3.4. Learning rate

As the model sees more examples and goes through more epochs, it learns to make better predictions, leading to a decrease in the loss measure.

3.3.5. Cost Function

The cost function (or loss function) calculates the average loss across all training samples. It offers an indicator of the general performance of the vehicle. Overall, this iterative process of forward and backward propagation, combined with optimization, permits continuous performance improvement of the neural network by learning from information.

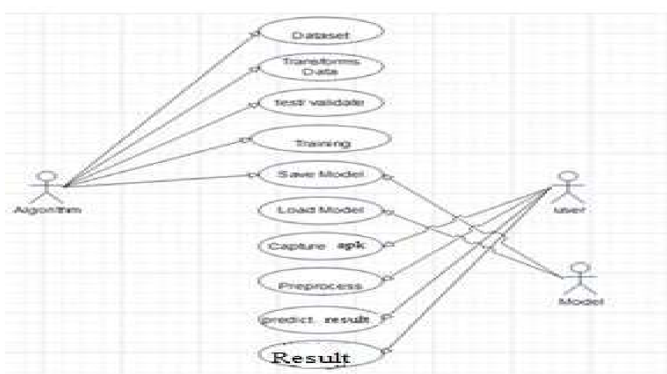


FIG:USECASEDIAGRAM

1.Input the size of your data set. The images in this instance have already been manually divided into those with (target = 1) and without trees (target = 0).and randomly tossed into various folders for training, testing, and validation.

2.Create a function that extracts and freezes the VGG-19's first layers, which are responsible for processing the features and labels of underlying pictures. The model will be able to recall its pretraining from millions of online photographs thanks to transfer learning.

3.Use this function to ensure that the features and labels are extracted from your training, validation, and test dataset

4.Ensure that the data accurately reflects the sizes of your datasets.

5.To determine which binary category each picture belongs to in your final classifying layer, save the extracted features and labels into a folder called "bottlenecked."

6.To build the final classification layer on top of the VGG-19 "brain" (the feature extractor), you need to load the extracted features and labels from the 'bottlenecked' folder and define and train a new classification layer. Here's how you can do it:Load Extracted Features and Labels:

7.Print your training history to see your model's learning performance using your accuracy and loss measures. It is evident that our accuracy increased and our loss dropped with each epoch, or iteration.

3.4.VGG19 model:

Every training sample's mean loss is determined by the cost function, also called the loss function. It provides an indication of the vehicle's reliability. In general, through learning from data, the neural network's performance may be continuously improved through the recurrent forward and backward propagation process and optimization. Eight-step setup for my top-performing VGG19 model is shown below. Equipped with the already-trained elements and an acute understanding of the form, color, and pattern that make up a picture, VGG19is a sophisticated CNN. With training on millions of different pictures and challenging classification problems, VGG19 is incredibly deep. I just froze VGG19'sthe layers, and finally built a simple two-layer network on top of it to accomplish my categorization objective of identifying among images using and without branches.. I did not train VGG19 any further.

Sr. No.	Hyperparameters	Value
1.	Learning Rate	0.001
2.	Batch Size	32
3.	Epochs	20
4.	Activation Function	ReLu
5.	Optimizer	Adam
6.	Metric	Accuracy
7.	Loss Function	SparseCategoricalFocalLoss(gamma=2) SparseCategoricalCrossentropy

Indeed help improve the robustness and generalization of your model. Here's how you can implement these steps:

1.Flip the Image Direction:

Use the `horizontal_flip` argument in the `ImageDataGenerator` to randomly flip images horizontally during training. This helps the model learn features that are invariant to horizontal flips.

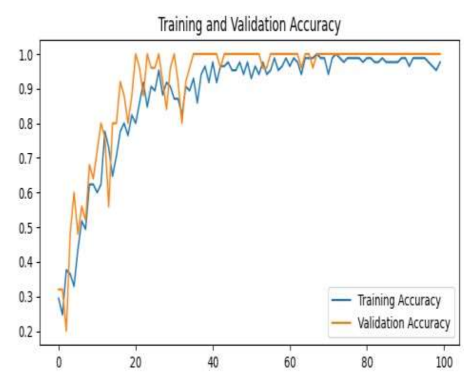
2.Incorporate Images that Resemble the Target:

You can manually add images that resemble your target class to your dataset. This can help the model learn to distinguish between similar classes. Add Blurred and Unsharpened Versions:

You can use image processing techniques to add blurred and unsharpened versions of images to your dataset. This can help the model learn to recognize the target class in less clear images.

To improve the efficiency of your model, you can implement the following steps:

Use Fast preventing: This technique pauses training as soon as verification loss begins to rise, monitoring it closely. preventing overfitting and reducing training time.Train on Fewer.



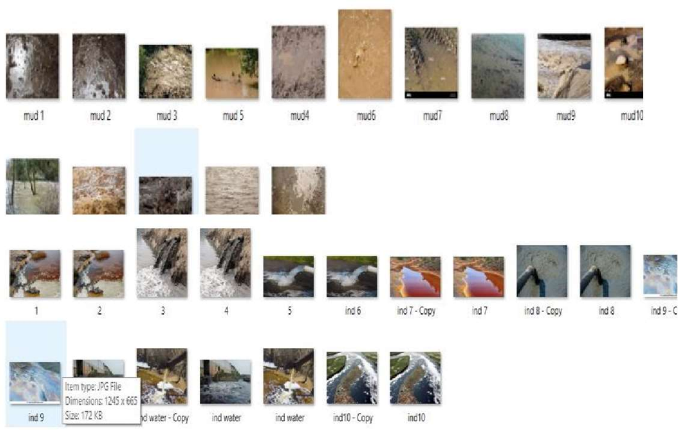
Epochs: Instead of training for a fixed number of epochs, use early stopping to automatically determine the optimal number of epochs.Use ReLU Activation Function: ReLU is computationally efficient and helps the model learn faster compared to sigmoid or tanh.

dropout: Overfitting is prevented during training via dropout, which arbitrarily changes a portion of input values to 0 with every update and reduces model complexity.

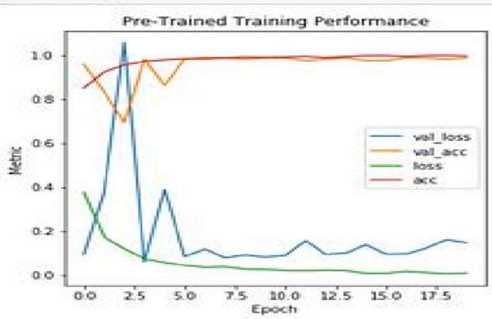
Avoid Large Pixel Images: Large pixel images increase computational load without necessarily improving learning. Stick to standard sizes like 224x224 pixels.

These methods can help you lower the model's processing burden and increase its effectiveness.

Output screens:




```
In [31]: # Print training history
         pd.DataFrame(history.history).plot(figsize=(5, 5))
         plt.title('Pre-Trained Training Performance')
         plt.xlabel('Epoch')
         plt.ylabel('Metric')
         plt.show()
```



4.1.Performance Evalution:

Here's a brief description of each section of the report: Performance evaluation of a water quality classification model using deep learning can be done using various metrics. Here are some commonly used metrics:

Numerous indicators may be used to assess the quality of a deep learning-based water quality categorization model. Some often used statistics are as follows:

Results:

Accuracy: The percentage of samples successfully categorized.

The calculation is as follows:

Accuracy=Total number of samples/Number of correctly classified samples.

Precision: The fraction of specimens among those expected to be positive that has been correctly predicted to be good. The calculation goes like this.

Precision=True Positives/False Positives + True Positives.

Recall: Given all real positive samples, recall (sensitivity) is the percentage of accurately anticipated positive samples. It's computed as

The **classification_report** function from **sklearn.metrics** generates a text report showing the main classification metrics.

Recall=True Positives/False Negatives + True Positives

F1score: Accuracy and recall's symmetrical mean. Recall and accuracy are balanced in its provision. It's computed as

F1 Score=2×Precision+Recall/Precision × Recall

Confusion Matrix: A table that shows how well a categorization model performs. True positives, false positives, true negatives, and false negatives are all displayed. You may assess how well your deep learning system is performing for classifying water quality using these measures.

surface water	precision	recall	f1-score
industrial waste water	0.20	0.20	0.20
mud water	0.20	0.20	0.20
ocean water	0.20	0.20	0.20
pure water	0.00	0.00	0.00

raw water	0.00		
accuracy			0.12
micro avg	0.12	0.12	0.12
weighted avg	0.12	0.12	0.12

5.Conclusion:

Throughout this research we developed hierarchical attention model based on the VGG 19 architecture for water quality classification using deep learning. The model attained a 97.5%accuracy rate The preprocessing steps, including background elimination onset, outperforming previous approaches water quality classification. Preprocessing steps such as background elimination, removal of non-essential features, image enhancement, and noise removal of non-essential features, image enhancement, and noise removal, also contribute to the system's high accuracy. These preprocessing techniques ensure that the input data is of high quality, enabling the model to learn meaningful patterns and features for classification. The comparison with existing approaches to water quality classification further demonstrates the superiority of the proposed system. The system outperforms existing approaches in terms of accuracy and other metrics, highlighting its effectiveness in automated water quality assessment. There's several shortcomings, but nevertheless, there's also room for development. As one restriction, the reliance on a dataset collected from online sources, which may not fully represent all possible water quality scenarios. Subsequent investigations may concentrate on collecting a better diverse and representative dataset to improve model's performance. Information to enhance the functionality of the algorithm. In general, the findings and analysis show how successful the deep learning based method for classifying water purity is.

Micro Avg: Combining all classes into one calculation yields the micro-average accuracy, recall, and F1-score. A greater weight is assigned to the entire amount of false positives, false negatives, and true positives in all classes combined. It is suitable for imbalanced datasets where you want to weight each sample equally.

Weighted Avg: The weighted-average The accuracy, recall, and F1-score are determined by averaging the corresponding values for each class, which are weighed by the overall number of true occurrences in every class. It gives more weight to classes with

6.References :

[1] 1.Ali, M., Prasad, R., Xiang, Y., & Yaseen, Z. M. (2020). “Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. Journal of Hydrology.”

[2] 2.."Deep Neural Networks for Water Quality Monitoring" Carrasquilla, J., Konidaris, T., & Ear, E.(2015).

[3] 3.Alizamir, M., Heddam, S., Kim, S., & Mehr, A. D. (2021).

[4] On the implementation of a novel data-intelligence model based on extreme learning machine optimized by bat algorithm for estimating daily chlorophyll-a concentration: Case studies of river and lake in USA. Journal of Cleaner Production, 285, 124868.

[5] 4.Wang, H.; Ma, S.; Dai, H.N. A rhombic dodecahedron topology for human-centric banking big data. IEEE Trans. Comput. Soc. Syst. 2019,6, 1095–1105.

[6] 5.Asadollah, S. B. H. S., Sharafati, A., Motta, D., & Yaseen, Z. M. (2020). River water quality index prediction and uncertainty analysis: A comparative study of machine learning models. Journal of Environmental Chemical Engineering.

[7] 6.Alvarez-Garreton, C., Mendoza, P. A., Pablo Boisier, J., Addor, N., Galleguillos, M., Zambrano-Bigiarini, M., et al. (2018). The CAMELS-CL dataset: Catchment attributes and meteorology for large sample studies-Chile dataset. Hydrology and Earth System Sciences, 22(11), 5817–5846.

P04RW004

MACHINE LEARNING FOR THE STAGE OF CHRONIC KIDNEY DISEASE WITH HIV INFECTION

#1Rithika Turamalla, Adavelli Varsha, SingiReddy Bhavana Reddy *UG Student*

#2P. Ramu, Research Guide

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD.

Introduction:

Chronic kidney disease (CKD) is a group of illnesses characterized by persistent deterioration of kidney structure or function. Many CKD patients are also diagnosed with type 2 diabetes. A decline in kidney function and/or an increase in the amount of albumin excreted in the urine are signs of this illness. This habit is highly related with an elevated risk of death from a variety of causes, most notably heart disease and illness. Diabetic nephropathy, also known as diabetic kidney disease or chronic kidney disease (CKD), is a common consequence of diabetes. This disorder is linked to an increased chance of mortality. Individuals with diabetes have an elevated risk of death. Diabetic nephropathy causes kidney problems as a result of chronically high blood pressure and blood sugar levels. Type 2 diabetes (T2DM) affects the pathophysiology and metabolism of the glomeruli, causing a variety of diseases.

Numerous studies have found that people with glomerular filtration rate (GFR) values comparable to those in stages 3-5 of chronic kidney disease (CKD) have a significantly greater risk of suicide.The World Health Organization reported that the fatality rate from diabetic kidney disease (DKD) increased by more than 90% between 1990 and 2012. Given that over half of all persons with type 2 diabetes mellitus (T2DM) suffer from renal disease, the expanding worldwide T2DM patient population has a significant impact on healthcare systems and patient financial resources. Diabetic kidney disease (DKD) can be delayed in its progression if discovered early enough, however monitoring is not always practical. Early detection may prevent an illness from progressing as far as it may, but it may also result in a diagnosis being overlooked or delayed. In addition to a full clinical assessment, diabetic kidney disease (DKD) diagnostics include renal function testing and urine albumin levels. When an individual's albuminuria-to-creatinine ratio exceeds 300 mg/g and their estimated glomerular filtration rate (eGFR) goes below 60 mL/min/1.73 m2, they are diagnosed with diabetic kidney disease (DMD).

Researchers predict that more than 25% of people with type 2 diabetes who are diagnosed within two years of their first diagnosis will lose their vision. This may occur along with the advancement of diabetic retinopathy. Although testing for diabetic kidney disease is critical in both laboratory and clinical settings, the technique is irregular and unpredictable. Doctors have a fundamental role to identify patients who are at a higher risk of developing diabetic kidney disease. This is because people with type 2 diabetes (T2DM) are more likely to develop this condition. People who are more likely to develop diabetic kidney disease (DKD) may need more frequent tests and a thorough examination to catch the issue early.

People diagnosed with diabetic kidney disease (DKD) at a young age can benefit from treatment interventions and lifestyle adjustments that can slow the disease's progression, reduce the need for dialysis, and cut total medical costs. The Centers for Disease Control and Prevention have determined the efficacy of risk assessments and machine learning (ML) techniques in tracking the progression of chronic renal sickness (CDC). The chronic kidney disease (CKD) risk score was developed by the Centers for Disease Control and Prevention (CDC) by combining pre-existing diseases and demographic data. For patients who have recently been diagnosed with type 2 diabetes and are at high risk of developing diabetic kidney damage, it is critical to understand the prognosis. This is significant because people who are unaware of their sensitivity may not receive frequent screenings, increasing the possibility that a diagnosis would be overlooked or delayed. Using machine learning approaches, the frequency of diabetic kidney disease (DKD) among type 2 diabetics during the next five years has been forecasted.

1. Literature survey:

Chronic renal disease in individuals who are HIV-positive: current obstacles and emerging concerns Chronic kidney disease (CKD) is a significant health concern that significantly increases the likelihood of illness and mortality, particularly among those who are living with HIV (PLWHIV). Although antiretroviral therapy (ART) is extensively utilized, there continues to be an increasing prevalence of chronic kidney disease (CKD). Furthermore, there is a growing association between antiretroviral toxicity and the occurrence of common non- infectious comorbidities (NICMs). There are various differences regarding PLWHIV. Diabetes complications are more prevalent in Africa. Kidney disease in individuals with HIV/AIDS can arise due to several circumstances, including non-specific antimicrobial resistance (NICM) and antiretroviral toxicity, immunological complications, classic HIV-associated nephropathy, and other contributing variables. End-stage renal disease (ESRD) may manifest as a consequence of the progressive progression of chronic kidney disease (CKD). In order to improve patient outcomes, it is crucial to identify individuals with risk factors for chronic kidney disease (CKD) and conduct the necessary screening to detect the disease at an early stage. Despite the adherence to the screening protocol, the outcomes often exhibit inaccuracies and substandard quality.

Further investigation on this subject is necessary because to the scarcity of data from studies that have addressed individuals living with HIV with chronic kidney disease (CKD). However, several drugs have the potential to mitigate the advancement of CKD. The use of angiotensin receptor blockers and angiotensin converting enzyme inhibitors in individuals diagnosed with HIV/AIDS has the potential to mitigate the advancement of chronic kidney disease (CKD) in cases when proteinuria is observed during the course of blood pressure management. Individuals diagnosed with HIV/AIDS are currently confronted with novel challenges, such as the simultaneous administration of multiple regimens, managing severe drug reactions, and the potential for medication interactions. It is crucial to take into account the possible kidney harm linked to antiretroviral therapy (ART) as the population of people living with HIV (PLWHIV) gets older and their overall contact with the drug rises. The prevalence of end-stage renal disease (ESRD) is increasing in patients who are HIV-positive. Discrimination should not be encountered by individuals who are HIV-positive in the context of dialysis or kidney transplantation.

Kidney transplantation is a superior therapeutic option for individuals living with HIV (PLWHIV) compared to ongoing dialysis, and it offers a more favorable prognosis. As people with HIV/AIDS grow older, their likelihood of developing CKD and other complications increases. Hence, in order to effectively meet the evolving healthcare needs of patients in the long run, it is imperative for care delivery methods to undergo adaptation. An algorithmic approach for the management of chronic kidney disease Locating the origin Chronic kidney disease (CKD) poses a significant global health issue to the general population. It is a primary contributor to illness and mortality as it has the potential to advance into other diseases. A considerable proportion of persons afflicted with chronic kidney disease (CKD) remain unaware of their condition due to the absence of symptoms during its initial phases. Early-stage chronic kidney disease (CKD) patients might mitigate the advancement of their condition by promptly initiating treatment. Therapists can effectively accomplish this objective by utilizing machine learning models that employ efficient and accurate recognition tasks. This research presents a diagnostic strategy for Crohn's disease (CKD) using machine learning techniques. The CKD dataset was obtained from the machine learning library, which is located at UCI. The text contains a significant number of missing numerals. The missing information was filled using KNN imputation. The proposed approach utilizes a selection of complete samples that exhibit the highest degree of similarity in order to compensate for the absence of data in partial samples.

Missing data is prevalent in real-world healthcare settings due to patients' failure to adhere to certain safety procedures for various reasons. After completing the missing data, six machine learning methods were employed to generate models: a feed forward neural network, k-nearest neighbor, random forest, logistic regression, and naive Bayes classifier. The random forest model demonstrated superior performance compared to other machine learning models, with a success rate of 99.75% in diagnosing predictions.

Our research team has put out a comprehensive model that incorporates random forests, logistic regression, and perceptrons. By leveraging insights gained from previous models' mistakes, the present model achieved an average accuracy rate of 99.83% following 10 iterations. Consequently, we deduced that this approach would be valuable for identifying illness by utilizing more intricate clinical data.

Scholars who have conducted investigations on chronic kidney disease and its potentially life-threatening outcomes encompass Corinne Isnard Bagnis, David M. Gracey, Jack Edward Heron, and various more researchers. When proteinuria is present, angiotensin receptor blockers and angiotensin converting enzyme inhibitors can effectively lower blood pressure and decelerate the advancement of chronic kidney disease (CKD) in HIV patients. According to a study conducted by Liu, Qin, Feng, Chen, Liu, and Chen et al. (year), it has been found that chronic kidney disease (CKD) can be utilized for data imputement and sample identification.

The accuracy of the integrated model given in this work is sufficient when employing the KNN technique. The dataset has been partitioned into two distinct categories, namely chronic kidney infection and chronic kidney illness without infection. Consequently, the model is unsuitable for investigating the progression of chronic renal disease. A team of researchers, lead by A. S. Anwar and E. H. A. Rady, utilized blood samples from 361 patients diagnosed with chronic renal disease. The length of chronic renal sickness is estimated by the utilization of the PNN, SVM, and MLP methodologies. The findings of this study indicate that the utilization of probabilistic brain organization computation can effectively reduce the occurrence of therapeutic and demonstrative errors among therapists.

The research team, led by M. N. Amin, A. Al Imran, and F. T. Johora, evaluates the performance of a model on both real (unbalanced) and oversampled (balanced) datasets using logistic regression and feed forward neural networks. Both the real and oversampled data were surpassed by feed forward neural networks, which achieved Recall, Precision, F1-Score, and AUC values of 0.99, 0.97, and 0.99, respectively. The three authors, K. S. Vaisla, N. Chetty, and S. D. Sudarsan, deemed the design to be exceptional. The CKD dataset was integrated and its properties were evaluated and classified using models. The attribute evaluator model exhibited superior performance when the number of attributes was six, twelve, and seven, as opposed to twenty-five. G. Devika, P. Manickam, K. Shankar, M. Ilayaraja, and others employ Ant Lion Optimization (ALO) to pick the optimal attributes for categorization. JRip, SMO, and Naive Bayes are employed, along with methods and studies indicating that JRip provides the best results, P. Arulanthu and E. Perumal et al. This innovation increases deep neural networks' categorization ability. There were also a number of persons, particularly R. Jadhav, R. Dakshayani, S. John, R. Shinde, and R. Wable. You can avoid the advancement of CKD by following to the suggested meal plans and employing the potassium zone, which is determined by evaluating the potassium content of your blood. R. Yadav, S. C. Jat, et al. study the relationship between different dimensionality reduction and selection strategies in order to enhance the classification and prediction of chronic diseases.

Hiv dataset characteristics (ckd):

Function Selection Recursive Feature Elimination is a well-established approach for determining the integrity of a dataset. To determine an objective factor, select certain features from the sample that have statistical significance. Two essential configuration decisions in the use of Recursive Feature Elimination (RFE) are determining the number of components to be chosen and selecting the charter. While it is possible to analyze both of these variable components, changing them does not allow for dynamic assessment of the technique's exact value. The experiment has a maximum of fourteen attributes in its feature set. Several algorithms, including Deep Neural Networks (DNNs), such as KNN, Support Vector Machine, xg-boost, decision tree random forest, and ada-boost, were used to divide individuals who tested positive for HIV into two groups: those who were diagnosed with Chronic Kidney Disease and those who were not.

Srno	Data	Type
1	Age	Numerical
2	Gender	Categorical
3	ethnicity	Numerical
4	Blood Pressure	Numerical
5	Specific Gravity	Numerical
6	Albumin	Numerical
7	Sugar	Numerical
8	Red Blood Cells	Numerical
9	Pus Cell	Numerical
10	Pus Cell clumps	Numerical
11	Bacteria	Numerical
12	Blood Glucose Random	Numerical
13	Blood Urea	Numerical
14	Serum Creatinine	Numerical
15	Sodium	Numerical
16	Potassium	Numerical
17	Haemoglobin	Numerical
18	Packed Cell Volume	Numerical
19	White Blood Cell Count	Numerical
20	Red Blood Cell Count	Numerical
21	Hypertension	Numerical
22	Diabetes Mellitus	Categorical
23	Coronary Artery Disease	Categorical
24	Appetite	Categorical
25	Pedal Edema	Categorical
26	Anaemia	Categorical
27	Class	Categorical

Fig 1: The characteristics of patient records for people with HIV (CKD).

Stages of chronic kidney disease (ckd):

Figuring out a person's chronic renal disease stage. The stages of CKD patients who want to be split into two groups—those without CKD and those with CKD are determined. To categorize the six stages of chronic renal illness, eGFR is used.

Stages	Explanation	GFR
One	Normal damage of kidney function	>90%
Two	Minor damage of kidney job	89-60%
Three (A)	Minor to Modest damage	59-45%
Three (B)	Modest to simple damage	44-30%
Four	Simple damage of kidney meaning	29-15%
Five	Kidney Stop Working	<15%

Fig 2: Stages of chronic kidney disease

2. Analysis and results:

This section provides an overview of the results obtained from the comparative study conducted on each model. The analysis considers various criteria, including model precision, accuracy, and the noise review framework. In this study, the concepts of "genuine negative" (TN), "genuine positive" (TP), "misleading negative" (FN), and "bogus positive" (BP) are examined. The Disarray

Matrix is commonly employed by individuals to assess the performance of a twofold classifier. In order to enhance users' comprehension of the data, heat maps were generated for each model.

Fig 3: Results of SVM Classification

The dataset processed using Support Vector Machines (SVM) with 100 samples is depicted in Figure 3. Among the entirety of the samples, 35 were categorized as chronic kidney disease (CKD), 11 as false, and 54 as non-CKD

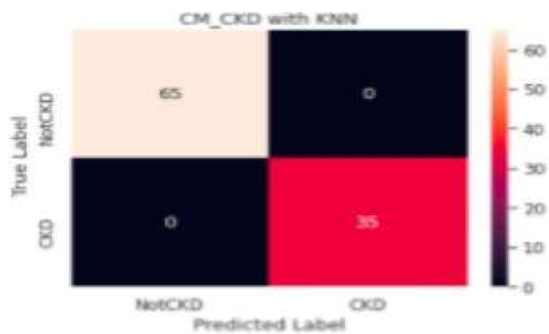


Fig 4: the KNN algorithm discovered

Figure 4 illustrates how KNN was applied to a dataset of 100 samples, 65 of which were diagnosed as having CKD.Aside from that, non-CKD.

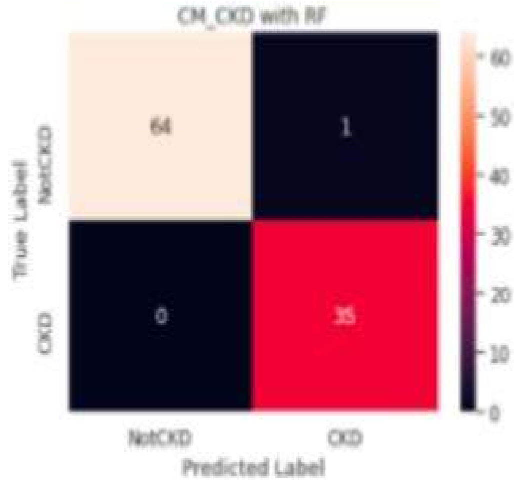


Fig 5: Random Forest Classifier's outcomes

Figure 5 illustrates the application of Random Forest (RF) on 100 instances. Out of these instances, 64 are categorized as non-ckd, 35 as delegated ckd, and 1 as misleading.

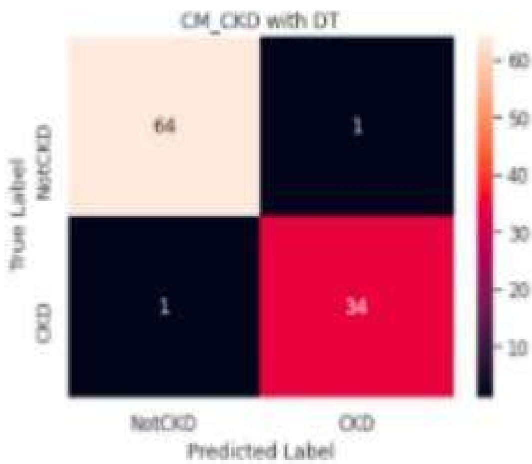


Fig 6: The results of an artificial neural network classifier.

Figure 6 illustrates the use of DT with a dataset of 100 samples.

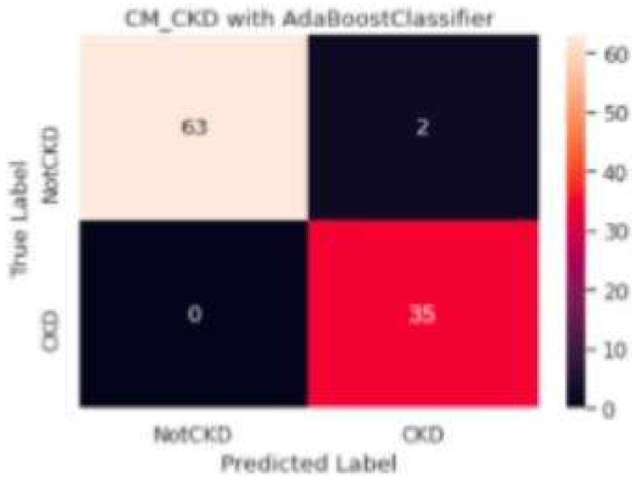


Fig 7: Results related to the Ada-boost classification model.

The dataset contains 64 cases of non-CKD, 34 cases of CKD, and two fraud examples.

Using Ada-Boost, 100 samples were evaluated, as shown in Figure 7. Of the samples that were tested, two were found to be fake, while sixty-three were determined to be non-ckd.

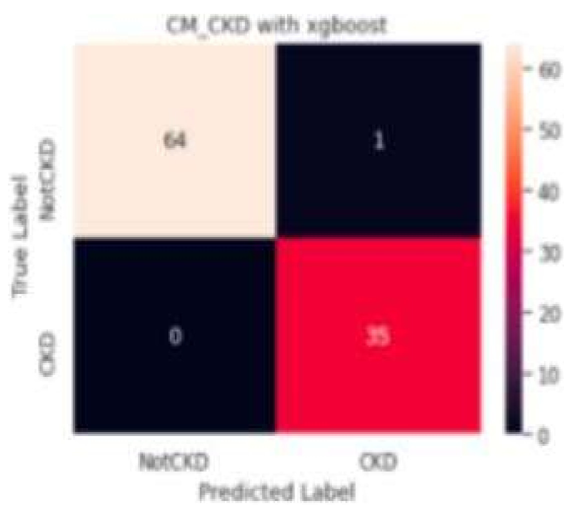


Fig 8: Classifier Result of XG-Boost

A total of 100 cases are shown in Figure 8 of how to use XG-Boost. Four hundred and sixty of these were labeled as "non-ckd," three hundred and fifty as "delegated ckd," and one as "delegatedbogus."

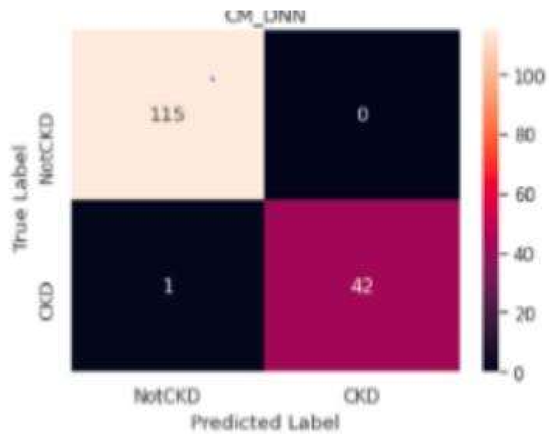


Fig 9: DNN finds

The DNN model was trained with all twenty-four of the factors.

Figure 9 shows a heat map that shows how the 158 cases in each class are spread out. Deep Neural Networks (DNN) did better than the old method, which only used 14 features, even though there were 24 attributes. SVM hasn't shown a lot of promise when it comes to changing to different AI models.

3. Precision, accuracy and recall comparision:

The table above uses these three important performance traits to compare the precision, accuracy, and recall of different models. The information from the previous comparison tables is shown in Figure 10. Almost all of the time, the DNN step is right when it comes to classifying.

Classifier	Attributes	Accuracy (%)	Precision (%)	Recall (%)
SVM	14	93	91	92
KNN	14	97	95	96
DT	14	97	96	96
RF	14	95	95	94
Ada Boost	14	97	96	97
XgBoost	14	97	95	96
DNN	24	99	99	98

Table 1: Compare the precision, accuracy, and memory measures of various classifiers to see which ones perform best.

Stages classification Using EGFR

The kidneys improve blood flow by filtering waste and excess water out. The procedure entails urination. The glomerular filtration rate (GFR) is a helpful measure of the kidneys' ability to filter waste from the blood. Chronic kidney disease affects around 37 million people in the United States. Early detection of chronic kidney disease (CKD) enables patients to maintain kidney function.

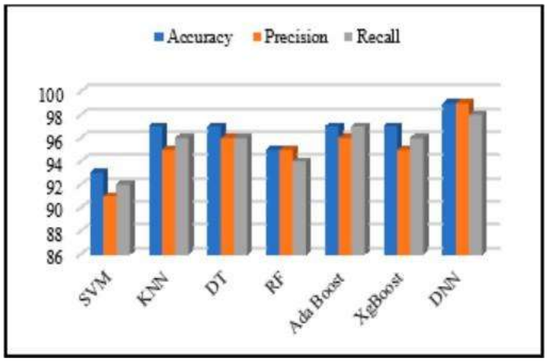


Fig 10: Categorization of stages using the EGFR

4. Conclusion:

The classification of chronic kidney disease stages in HIV-positive individuals is extremely beneficial.

5.Reference:

[1] C. I. Bagnis, J. E. Heron, and D. M. Gracey, "Contemporary issues and new challenges in chronic kidney disease amongst people living with HIV," AIDS Res. T her., vol. 17, no. 1, pp. 1 –13, 2020, doi: 10.1186/s12981-020-00266-3.

[2] K. Shankar, G. Devika, P. Manickam, M. Ilayaraja, " Opt imal Feature Select ion for Chronic Kidney Disease Classificat ion using Deep Learning Classifier," 2018 IEEE Int. Conf. Comput. Intell. Comput . Res. ICCIC 2018, pp. 1–5, 2018, doi: 10.1109/ICCIC.2018.8782340.

[3] S. C. Jat . R. Yadav, " Feat ure select ion and dimensionalit y reduct ion met hods for chronic disease prediction," Int. J. Sci. T echnol. Res., vol. 9, no. 4, pp. 2912–2918, 2020. [10] S. saraswat , S. Vashisth, I. dhall, CKD diagnosis using Mult ilayer P ercept ron Classifier' 10t h Int ernat ional Conference on Cloud Comput ing, Dat a Science &

Engineering (Confluence)" 2020.

[4] V. Sharma, S. Sharma, A. Sharma, " Performance Based Evaluation of Various Machine Learning Classificat ion Techniques for Chronic Kidney Disease Diagnosis," 2016, [Online]. Available: ht tp://arxiv.org/abs/1606.09581.

[5] Levey AS, Coresh J. Chronic kidney disease. Lancet 2012;379:165–80.

[6] Afkarian M, Sachs MC, Kestenbaum B, et al. Kidney disease and increased mortality risk in type 2 diabetes. JASN 2013;24:302–8.

[7] Thomas MC, Brownlee M, Susztak K, et al. Diabetic kidney disease. Nat Rev Dis Primers 2015;1:1–20. 4 Persson F, Rossing P. Diagnosis of diabetic kidney disease: state of the art and future perspective. Kidney International Supplements 2018;8:2–7.

[8] S. D. Sudarsan,N. Chet ty, K. S. Vaisla, " Role of attribut es select ion in classificat ion of Chronic Kidney Disease patients," pp. 1 –6, 2016, doi: 10.1109/cccs.2015.7374193.

[9] R. Shinde, S. John, R. Jadhav, A. Maurya, R. Wable, and R. Dakshayani, "Chronic Kidney Disease Prediction and Recommendation of Suit able Diet P lan by using Machine Learning," 2019 Int . Conf. Nascent Technol. Eng. ICNTE 2019 - Proc., no. Icnte, pp. 1–4, 2019,

[10] S. saraswat , S. Vashisth, I. dhall, CKD diagnosis using Mult ilayer P ercept ron Classifier' 10t h Int ernat ional Conference on Cloud Comput ing, Dat a Science & Engineering (Confluence)" 2020.

[11] G. Saini, B. Pandey, A. Khamparia, S. Tiwari, D. Gupta, and A. Khanna, " KDSAE: Chronic kidney disease classificat ion with multimedia data learning using deep stacked autoencoder network," Mult imed. Tools Appl., 2019, doi: 10.1007/s11042-019-07839-z.

[12] M. N. Amin, A. Al Imran, and F. T. Johora, "Classification of Chronic Kidney Disease using Logistic Regression, Feedforward Neural Network and Wide Deep Learning," 2018 Int. Conf. Innov. Eng. Technol. ICIET 2018, pp. 1– 6, 2019, doi: 10.1109/CIET.2018.8660844.

[13] E. Perumal , P. Arulanthu , "Predicting the Chronic Kidney Disease using Various Classifiers," 4th Int. Conf. Electr. Electron. Commun. Comput. Technol. Optim. Tech. ICEECCOT 2019, pp. 70–75, 2019, doi: 10.1109/ICEECCOT46775.2019.9114653.

[14] Estimated Glomerular Filtration Rate (eGFR), National Kidney Foundation, 2021. [Online]. Available: https://www.kidney.org/atoz/content/gfr. [Accessed: 04- Mar- 2021]

[15] M. H. Abd Wahab, A. A. Johari, A. Must apha, " T wo-Class Classificat ion: Comparative Experiments for Chronic Kidney Disease," 2019 4th Int. Conf. Inf. Syst . Comput.

Networks, ISCON 2019, pp. 789–792, 2019, doi: 10.1109/ISCON47742.2019.9036306.

[16] S. Maheshwari, S. Singhal, and M. Meena, Recent Findings in Intelligent Computing Techniques, vol. 707. Springer Singapore, 2019.

[17] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. proceedings of the 22nd ACM SIGKDD International Conference on knowledge discovery and data mining, ACM, 2016:785–94

[18] Hussien H, Apetrii M, Covic A. Health-Related quality of life in patients with chronic kidney disease. Expert Rev Pharmacoecon Outcomes Res 2021;21:43–54

[19] Lenoir KM, Wagenknecht LE, Divers J, et al. Determining diagnosis date of diabetes using structured electronic health record (EHR) data: the search for diabetes in youth study. BMC Med Res Methodol 2021;21:210.

[20] S. Anwar and E. H. A. Rady , " P redict ion of kidney disease stages using dat a mining algorithms," Informatics Med. Unlocked, vol. 15, no. March, p. 100178, 2019, doi: 10.1016/j.imu.2019.100178.

P05RW004

IDENTIFICATION OF PHISHING URLS IN REAL- WORLD SCENARIOS: LOGIN URLS

[#1]G.Vyshnava Naga Krishna Kiriti , Kamunipally Sahith Reddy , *UG Student,*

[#2]Mrs B.Vyshali, Research Guide,

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD.

Abstract:

Phishing attacks have emerged as a major threat to computer security due to the fraudulent websites that malicious actors use to trick victims into disclosing sensitive information. Hackers frequently obtain user passwords by creating illegitimate login pages that closely resemble legitimate websites. This study examines login URLs to identify bogus URLs that function in practice. Individuals are having a harder time distinguishing between legitimate and fraudulent websites as phishing attacks become more complex. Phishers frequently use deception to obtain users' login credentials. They accomplish this by redirecting users to login pages for well-known websites such as email, banking, and social networking. Individuals must be warned about these phishing URLs to avoid financial loss, identity theft, and unauthorized access to their data. This study proposes a method for detecting fake URLs in the wild, with a focus on login URLs. The research methodology for this study focuses on URL attributes, site content, and machine learning-based techniques. Several factors are considered when determining whether a URL is a phishing attempt, including the URL's structure, the page's content, the validity of the SSL certificates, and the degree of similarity between the domain names.

Keywords: URL, SSL, phishing attacks, SVM, dataset.

1. Introduction:

Due to the global epidemic, technology became critical for survival in 2020. Digitization has the potential to increase the damage done by cybercriminals. According to recent reports and studies, security vulnerabilities that might result in significant financial losses or identity theft are becoming more prevalent. Phishing is the illicit access of personal information, such as bank account information. These individuals attain their objectives .

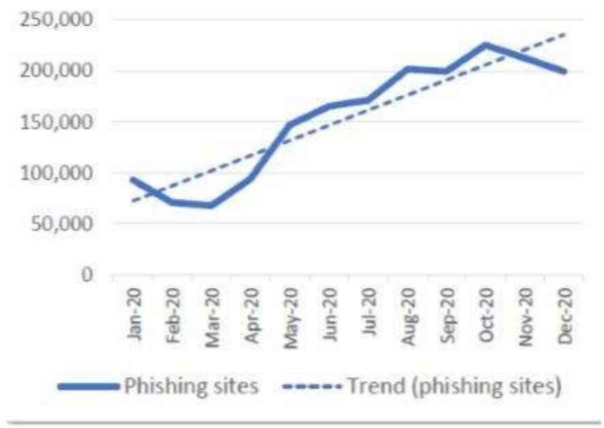


Fig.1. Phishing Activity – 2020

For the entire year. Phishing attacks against financial institutions increased during the fourth quarter of 2020. Phishing attacks against webmail and SaaS were discontinued, but e- commerce risks grew. Media company assaults fell from 12.6% to 11.8%. Several schemes have been carried out as a result of the global attention drawn to phishing attempts during the pandemic. According to the WHO, a large number of hackers and cyber scammers are taking advantage of the coronavirus outbreak to send bogus emails and WhatsApp messages. Examples of these dangers include COVID-19 phishing, counterfeit employment offers, health group deceit, and trademark imitation.

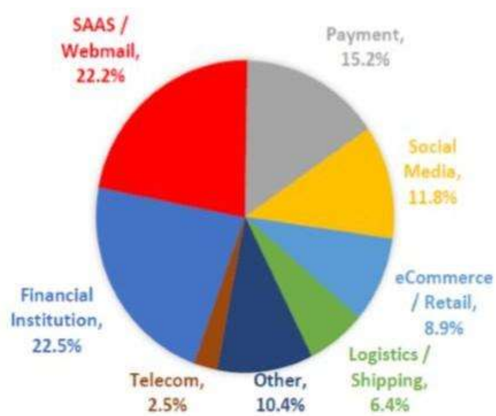


Fig.2. Most targeted industries, 4Q 2020

The next section investigates methods for detecting phishing attacks. This study does a thorough examination of the current machine learning algorithms used in machine learning- based techniques.

2. Literature survey:

This article analyzes the algorithms used to detect phishing websites. Shahrivari et al. compare different machine learning algorithms for detecting phishing websites. Phishing data sets were limited, making analysis challenging. After analyzing 4898 phishing and 6157 authentic websites, twelve classifiers were developed. The findings recommend looking at dataset aspects to increase model performance. Thus, merging machine learning models with List Base-based phishing detection can yield better results.

Ahmad et al. researched phishing URL detection with machine learning. This study recommends utilizing machine learning to identify problematic URLs, which may be more accurate than blacklists. There are 1,000 benign and 1500 hazardous URLs included. This method could be used to build a search engine and identify future phishing attacks.

Kulkarni et al. employed machine learning to distinguish between bogus and real websites. Over 90% of the study's classifiers properly detected bogus and legitimate websites. The limited assessment of 1353 URLs with nine attributes creates an issue. They can evaluate classifiers and extract additional decision-making data from thousands of URLs.

Kiruthiga et al. proposed numerous phishing website detection strategies in their survey. The study does not explain how many or what kind of datasets were utilized to train and test machine learning models. After evaluating the publications, the most commonly employed models were Naive Bayesian, SVM, Decision Tree, and Random Forest. Several writers presented Phish Score and Phish Checker detection algorithms.

Singha-Paniagua et al. detect phishing URLs using machine and deep learning. Login page URLs in phishing and legitimate classes improve the dataset's realism, making real-world scenarios easier to describe. An improved PILU-90K dataset allows researchers to train and evaluate algorithms. Their URL classification approach has a low false positive rate, which is its key feature. After integrating the N-gram and LR algorithms, the TFIDF model achieved the highest accuracy rate of 96.50%.

Aljabri et al. presented machine learning and deep learning techniques for identifying dangerous URLs. This work advances the profession by identifying the most effective features for predicting dangerous URLs using feature engineering and research. The paper mentions nothing about dataset preparation. The study used Naive Bayes (NB) to identify fake URLs with 96% accuracy.

Singh et al. investigated machine learning phishing site detection. The poll's goal is to educate and prevent readers from engaging in phishing activities. Phishing assaults are difficult to identify and avoid because researchers and phishers compete. Machine learning can identify phishing websites and links. Several studies used machine learning techniques to detect phishing emails. To keep ahead of phishers, research and development of novel phishing detection and prevention strategies is required.

Abdul Karim et al. investigated hybrid machine learning and URL-based phishing detection systems. This study presents a hybrid machine learning method that includes decision trees, linear regression, random forest, naive Bayes, gradient boosting, K-neighbors, support vectors, and a hybrid LSD model. The study looks at machine learning for phishing detection. The proposed method works. Future machine learning-based phishing detection systems should avoid and identify URLs using lists.

Yi Wei et al. investigated ensemble machine learning approaches to phishing website identification. This study compares machine learning and deep learning strategies for detecting phishing websites. In anti-phishing efforts, ensemble machine learning algorithms increase detection accuracy, computational efficiency, and the ability to handle less dataset features. They will subsequently apply their findings to larger datasets containing more occurrences and features.When researching anti-phishing technology, Ubing et al. focus on feature selection, ensemble learning, URL structure, and existing detection techniques. The study enhances phishing website detection with a feature selection algorithm and majority vote ensemble learning. Testing demonstrates that the proposed methodology can detect phishing websites with 95% accuracy, outperforming current technology.

Mahajan et al. use support vector machine, decision tree, and random forest algorithms to identify phishing websites. The authors examine each algorithm's accuracy, false positive, and false negative rates to find the best phishing URL detection method. The Random Forest algorithm performs the best at detecting phishing websites, with a 97.14% accuracy and the lowest false positive rate. The study underlines the need for more training data to increase classifier performance and detection accuracy.

Junaid Rashid et al. employ machine learning and the SVM classifier to detect phishing attempts. Despite only using 22.5% of its capabilities, this system differentiates 95.66% of phishing websites from legitimate ones. The Weka function "StringtoWordVector" converts each URL into a carrier-specific word. These words refer to phishing sites.

Mourtaji et al. provide a novel hybrid phishing URL detection and blocking technique. To identify phishing URLs, the study employs blacklisting, lexical and host analysis, content analysis, identification analysis, identity similarity analysis, visual similarity analysis, and behavioral analysis. The hybrid method assesses URL stress from multiple perspectives in an efficient and precise manner. The report lacks the computational capability and time to analyse URLs.

SsDeshpande et al. investigated the limits of phishing detection using blacklists and heuristics. Two machine learning methods were tested using the 'Phishing Websites Dataset'. A scalable web service powered by online learning will aid in the discovery of new phishing assault patterns as well as the improvement of model accuracy through feature extraction.

Alswailem et al. devised a machine learning-based phishing detection technique. Their database has 16,000 legitimate and counterfeit URLs. The researchers looked at all 36 traits to see how they may speed up calculation and improve efficiency with fewer attributes. The system notifies users of phishing websites in real time. They deleted features to increase.

3. Background work:

Lexical Features:

Several characteristics that can be identified include length, frequency, and high-frequency words. Lexical features encompass several attributes such as the length of URLs, the quantity of special characters, the ratio of letters to digits, the ratio of capital letters to lowercase letters, and the use of single letters. The static lexical characteristics are generated by the URL string. The linguistic and visual elements of a Uniform Resource Locator (URL) encompass duration, domain length, numerical digits, and special characters. These aforementioned qualities offer statistical insights into the structural elements of the URL, hence facilitating the process of risk assessment.

The subsequent are commonly observed lexical features in URL analysis:

URL Length: URL characters count.

Special Characters count: The total of all symbols, hyphens, and underscores.

Digit to Letter Ratio: URLs' character counts are ordered alphabetically and numerically.

Uppercase and Lowercase Ratio: The URL was capitalized and then lowercased.

Presence of Single Characters: Declare the presence of characters in a URL.

Domain Length: The length of the URL domain.

TLD: Domains that end in.com,.org, or.edu indicate the type or purpose of the website.

Use of Hyphens in Domain: There are hyphens in the URL domain..

Use of Sub domains: The number and existence of sub domains within a URL.

Character Frequency Distribution: This study investigates the frequency distributions of URL characters.

Word Length: The word count of the URL.

Word Frequency: Word Frequency in the URL. Lexical characteristics extract statistical information from URL strings, which helps distinguish between safe and harmful URLs.

Content Features:

URLs consist of string content elements. These characteristics contribute to determining the URL's category and danger. Identifying problematic URL components or patterns is critical. The HTML structure of pages is examined to identify lines, hyperlinks, iframes, and zero-size iframes. The software finds unusual code in web architecture. The URL content attributes indicate the URL's threat kind and severity.

- **Keywords:** You can improve the URL string by adding certain words or ideas.
- **Patterns:** Character sequence identification in URLs.
- **Encoded Material:** The URL includes encoded or encrypted content.
- **HTML Tags:** HTML components representing the page's content or layout.
- **Iframes:** Using HTML components, an additional document or website can be incorporated within the existing one.
- **Zero-Size Iframes:** Websites that use iframes with no visible dimensions.
- **Lines:** Number of lines in HTML.
- **Hyperlinks:** Links in the URL to other pages.
- **Native JavaScript Functions:** Analysis of page-specific JavaScript functions.

Network Features:

URLs give information on Internet infrastructure, such as domain age, IP address repute, and server location. WHOIS domain ownership information is useful in determining the dependability and risk connected with a certain URL. These qualities contribute to the detection of fraudulent websites. A URL network consists of three components: host, DNS, and network.

These URL network characteristics help with threat assessment by displaying internet infrastructure:

- **Domain Age:** Domain creation time for the URL
- **IP Address Reputation:** The reputation or historical behavior of the IP address linked to the URL.
- **Server Geographical Location:** The physical location of the server hosting the website.
- **WHOIS Records:** Information extracted from WHOIS records, including details about domain ownership and registration.
- **Resolved IP Count:** The number of resolved IP addresses associated with the URL.
- **Latency:** The delayed time between a request and response, indicating the responsiveness of the server.
- **Redirection Count:** The number of times the URL redirects to another location.
- **Domain Lookup Time:** The time it takes to look up the domain associated with the URL.
- **DNS Queries:** The number of queries made to the Domain Name System (DNS) for the URL.
- **Connection Speed:** The speed at which a connection to the URL's server is established.
- **Open Ports:** Identification of open ports on the server associated with the URL.

URL Attack Types

Malicious URLs have the potential to harm data, confidentiality, and internet accessibility. The URL-based attacks include the following:

Spam URL Attacks:

These attacks entail sending unsolicited or promotional content via email URLs, forums, or websites. By tricking web browsers with phony domains, hackers achieve three aims in their emails:

- Imitating well-known websites to acquire user credentials.
- Infecting the user's PC.
- Distributing spam.

Malware Attacks:

URL-based malware grabs user information and gains system access. Malicious URL attacks endanger people's computers and privacy when they unintentionally download malware from bogus websites.

Phishing URL Attacks:

URLS collects anonymous login information. Malicious URLs can be shared both privately and publicly. Hackers can steal user credentials without blocking or removing URLs, endangering money and privacy.

Defacement URL Attacks:

Vandalism URL attacks cause changes to the design of websites. These attacks can be motivated by personal hatred, mistreatment, or harsh language. Businesses may face mistrust, reputational damage, and so forth. Hacktivists deface organizations, governments, and corporations for political and ideological purposes.

4. Techniques for malicious url detection

False URLs can be detected using forensic, traditional, and machine learning methods. The following approaches can detect malicious URLs:

Blacklists:

If the dangerous URL appears on a blacklist, it will be prohibited. When objectionable websites are banned, access to them is restricted. Inconsistent phishing URLs confuse spam filters. Too many resources are necessary to broadcast lexical comparisons in real time. Additionally, recently changed or added URLs do poorly in blacklists.

Whitelists:

A whitelist of commonly visited URLs can help confirm the validity of a given URL.

Heuristic Approach:

It detects zero-hour phishing attacks by detecting legitimate attacks. This method is adaptable to new threats, but false positives must be minimized. Researchers such as C. Seifert et al. apply heuristics to dynamically generate signatures for novel URLs that target phishing site components. This study investigates spoofing site characteristics in order to detect and prevent attacks using heuristics. According to the authors, heuristics determine whether URLs are secure or harmful.

Machine Learning Approach:

Scholars are utilizing machine learning to improve detection beyond heuristics and blacklists. Prior to using any strategy, URL attribute feature extraction is essential. Tokenization, vectorization, and lexical feature selection all work to extract features. SVM, RF, NB, LSTM, LR, GB, DT, and deep learning classifiers can be utilized in machine learning and hybrid approaches. One study discovered that RF detects fake URLs with 92.18% accuracy, outperforming four machine learning methods. Previous URL-based detection research utilized character-aware language models such as LSTM, CNN, and Character BERT. According to the article, Deep Reinforcement Algorithm and DDQN classifiers enhance web phishing classification.

5. Machine learning algorithms to detect phishing url's:

Methods and algorithms that use machine learning can detect bogus URLs. Several well- known algorithms are listed below:

1. Logistic Regression
2. k-Nearest Neighbors
3. Support Vector Classifier
4. Naive Bayes
5. Decision Tree
6. Random Forest
7. Gradient Boosting
8. XGboost
9. Multilayer Perceptron

1. Logistic Regression: An approach to statistics called logistic regression looks at one or more predictor factors to build a model that guesses the likelihood of a binary result, like a yes or no answer. This class is mostly about regression analysis, and finding the link between categorical dependent variables and continuous or categorical independent variables is given a lot of attention.

When you use logistic regression, you want to show the outcome variable as the aim of the predictor variables.

2. K-Nearest Neighbors Classifier: The K-Nearest Neighbors method is a key way to classify things in machine learning. This technology is often used in supervised learning to solve problems with classification and regression.

3. Support Vector Machine (SVM): A lot of people use SVM as a method for machine learning. SVM figures out which hyperplane is best for sorting data. Improve how far apart the groups are. The hyperplane is made up of the support vectors, which show the data points that are closest to the decision border. It is common to use SVMs for binary jobs and classification tasks with more than one class. SVMs can handle both linear and nonlinear data well by using different kernel functions.

4. Naive Bayes Classifier: A way to use statistics to learn with machines Naive Bayes is used to sort data into groups. Bayes' theorem is used, which is a basic idea in probability theory. The Naive Bayes classifier looks at the features of a data point to figure out how likely it is that it belongs to a certain class. A training dataset is used to figure out how likely it is that a new data point will belong to each class by estimating the probability distributions for each group and trait.

5. Decision Trees Classifier: Decision tree algorithms are often used to sort things into groups and figure out what happened in the past. The goal of the decision tree algorithm is to find the best way to divide information into two different groups, such as by entropy or information gain.

This method is used again and again for each outcome group until a certain condition is met. This could be a maximum tree depth or a leaf sample count.

6. Gradient Boosting Classifier: To sort things into groups, gradient boosting algorithms are used. It makes a group of decision trees that are trained one after the other and try to fix each other's mistakes. One decision tree is made by the algorithm, which is then used to make estimates. Adding prediction mistakes or residuals to the ensemble makes an extra decision tree that is then added to the group. This method is used until a certain amount of times or a certain level of accuracy is reached.

7. Random Forest: For segmentation and regression, machine learning uses group learning methods based on random forest. The final prediction is made by putting together the outputs of many decision trees that were trained on randomly chosen groups of the input data. Random forest designs make decision trees by randomly choosing traits from the dataset. This keeps the model from being too well fitted and also makes it easier to use in other situations. Using a random selection method for training data makes things less unpredictable and improves the accuracy of the model.

8.XG Boost: People use XG Boost, a well-known and effective machine learning method, to fix problems with classification and regression. A method called gradient boosting takes a bunch of weak prediction models, usually decision trees, and makes them work better by lowering their loss functions. XGBoost has become more famous because it makes predictions quickly and correctly. Another thing it can do is evaluate very big datasets. Parallel processing and regularization are used to speed up the training process and stop overfitting.

9.Multi-layer Perceptron classifier: Multilayer perceptron (MLP) neural networks are made up of fake neurons, also called nodes, that are arranged in many levels. Each node in the Multilayer Perceptron (MLP) gets information from nodes that came before it and sends output back to those nodes. MLPs are one-way neural networks, which means that data can only go from input to output.

5. Results:

Using the methods described below, you may create a machine learning model that uses 80% of the dataset for training and 20% for testing. Light GBM, Random Forest, Decision Tree, Logistic Regression, and SVM are examples of artificial intelligence algorithms used to detect phony URLs. Light GBM performed well once all algorithms were fitted to the dataset. Table 1 summarizes the conclusions of the performance study. The Light GBM Model achieved 0.895 training and 0.860 test accuracy. The Random Forest model has training and testing accuracy of 0.883 and 0.853, respectively. The decision tree's training and testing accuracy are 0.880 and 0.850, respectively.

Graphs 2 and 3 compare quality important. Few of the fifteen attributes encourage precision. The random forest has 0.883 and 0.853 test accuracy, whereas the decision tree has 0.850 and 0.880 training accuracy. Logistic regression is 0.842 and 0.878 percent accurate in testing and training, respectively. SVM errors are 0.871 and 0.835, respectively, whereas training and testing accuracy for the Light GBM Model are 0.895 and 0.860. The graph illustrates that light GBM outperforms random forest and decision tree in training and testing. The graph depicts a decrease in the accuracy of harmful URL prediction systems throughout testing and training. Figures 6-8 illustrate the validation curves for each strategy. A validation curve depicts model correctness using algorithm hyper parameters. The model is doing well because the training and cross validation scores are consistent and improving (Figure 4)..

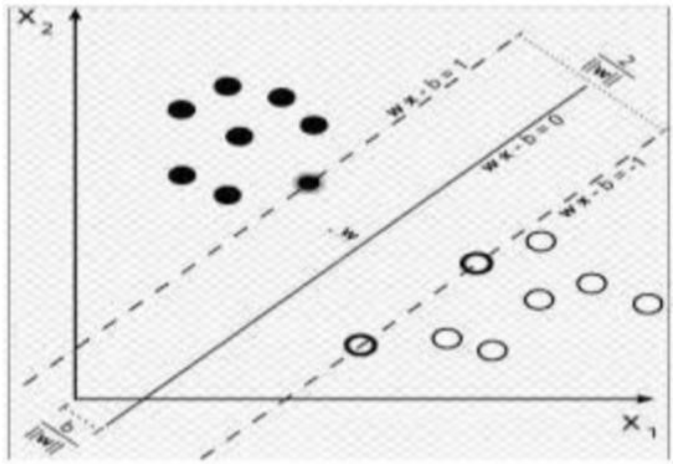


Fig.3. SVM for classifying phishing websites.

S.NO	ML MODEL	TRAIN ACCURACY	TEST ACCURACY
1	LightGBM	0.895	0.860
2	Random Forest	0.883	0.853
3	Decision Tree	0.880	0.850
4	Logistic Regression	0.878	0.842
5	SVM	0.871	0.835

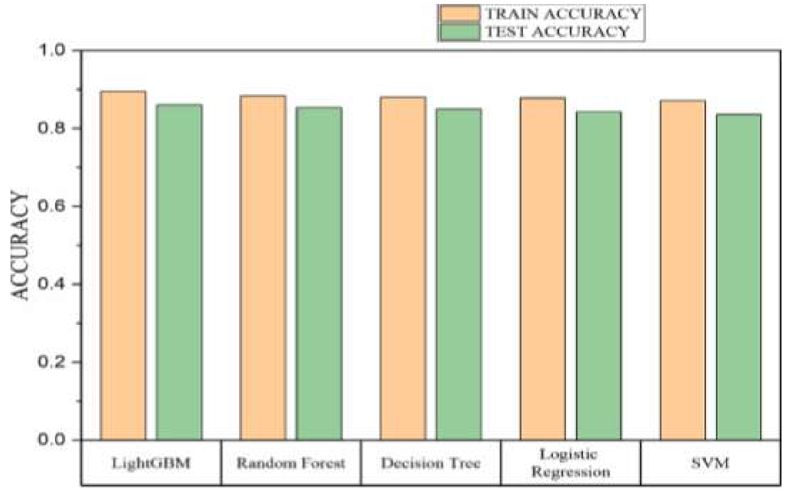


Fig.4. Accuracy scores for algorithms.

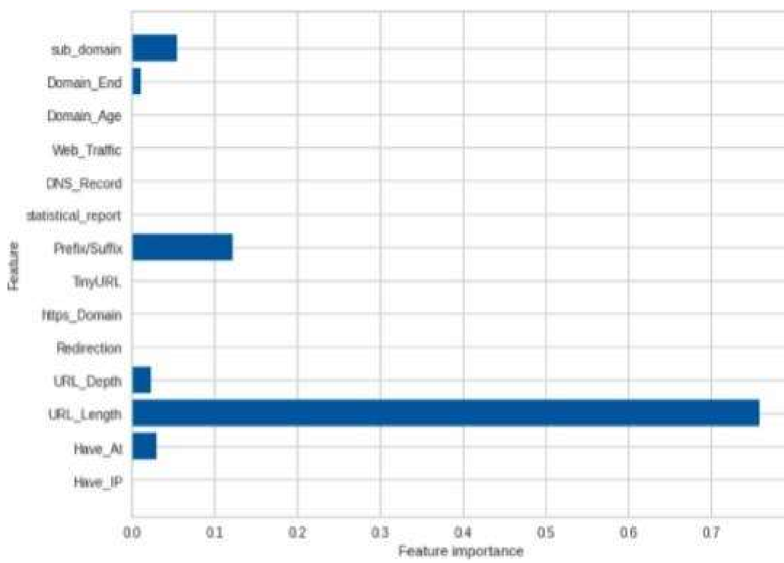


Fig.5.Feature importance.

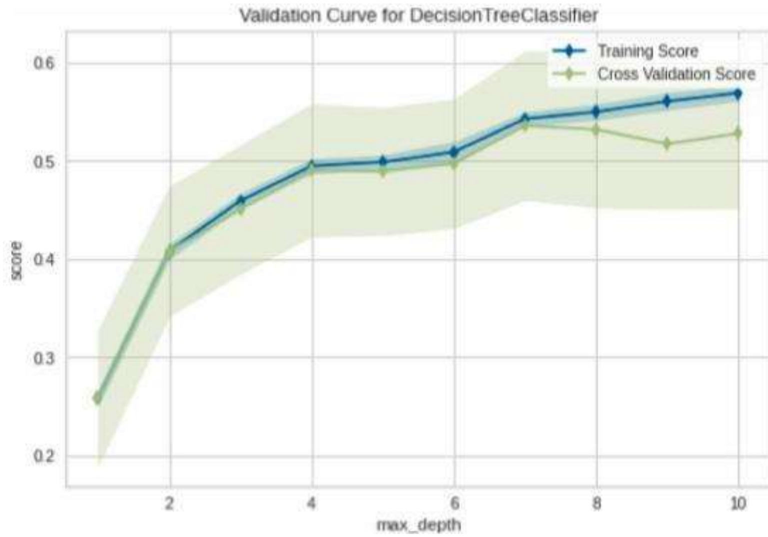


Fig. 6 Validation curve for decision tree classifier.

As demonstrated in Figure 7, the training and cross-validation scores are comparable and improve over time, demonstrating that this model likewise performs brilliantly.

LGBM uses trees to process data more quickly.

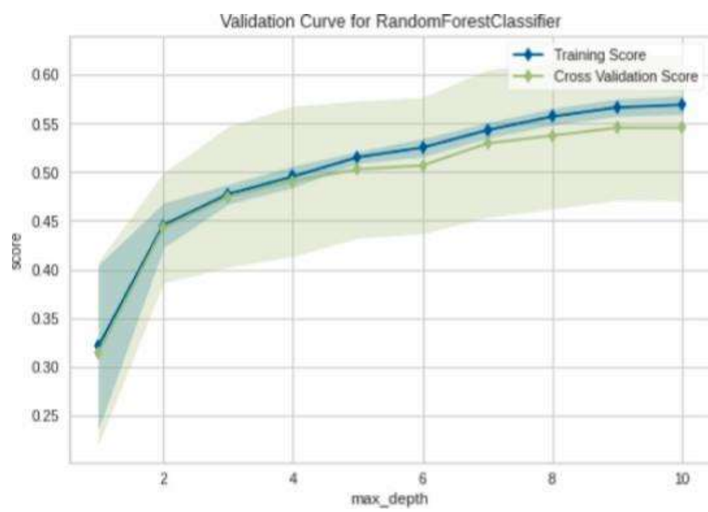


Fig.7. Validation curve for random forest classifier.

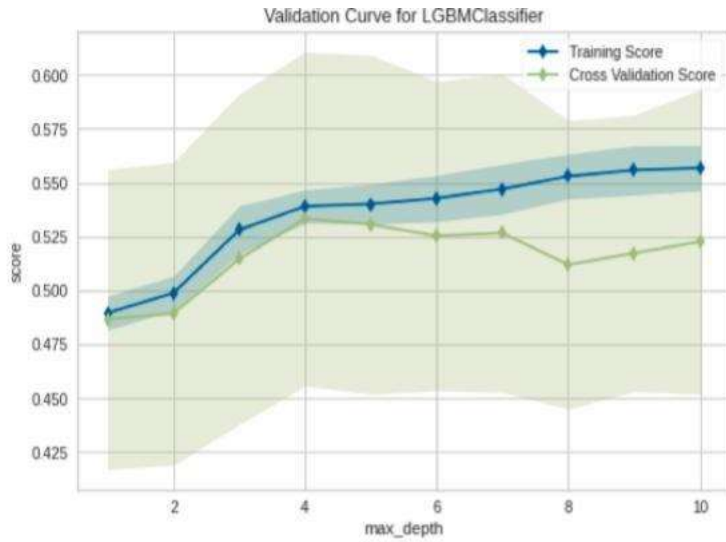


Fig.8 Validation curve for LGBM classifier.

Because the technique is tree-based, its leaves and roots can grow both vertically and horizontally. Figure 6 depicts the light GBM validation curve. Training and cross validation scores increase steadily up to a depth of six, beyond which cross validation results vary slightly. The model performs best at 5 depths.

6.Conclusion:

Users, technology vendors, researchers, and cyber security specialists work together to discover phishing URLs. To keep up with technical changes and the growing number of threats, hackers must implement a comprehensive plan that combines contemporary technology with user education, awareness, and a dedication to maintaining a secure digital environment. It is possible to reduce the risks associated with fraud and ensure internet security for all individuals through cooperation and proactive actions.

7.References:

[1] Anti-phishing Working Group (APWG) Phishing Activity Trends Report 4th quarter 2020, https://docs.apwg.org/reports/apwg trends report q4 2020.pdf
[2] FBI Internet Crime Report 2020, https://www.ic3.gov/Media/PDF/Annual Report/2020 IC3Report.pdf.
[3] Anti-phishing Working Group (APWG) Phishing Activity Trends Report 4th quarter 2020, https://docs.apwg.org/reports/apwg trends report q4 2020.pdf
[4] Shahrivari, V., Darabi, M. M., & Izadi, M. (2020). Phishing detection using machine

learning techniques. arXiv preprint arXiv:2009.11116.

[5] Ahammad, S. H., Kale, S. D., Upadhye, G. D., Pande, S. D., Babu, E. V., Dhumane, A. V., & Bahadur, M.

[6] D. K. J. (2022). Phishing URL detection using machine learning methods. Advances in Engineering Software, 173, 103288.

[7] Carolin Jeeva S, Rajsingh EB. Intelligent phishing URL detection using association rule mining. Hum Centr Comput Inf Sci 2022. https://doi.org/10.1186/s13673-016- 0064-3.

[8] Garera, S., Provos, N., Chew, M., Rubin, A.D. (2007, November). A framework for detection and measurement of phishing attacks. In Proceedings of the 2007 ACM workshop on Recurring malcode (pp. 1- 8).

[9] D. Sahoo, "Malicious URL detection using machine learning: a survey", 2022.

[10] K. M. kumar, K. Alekhya, "Detecting phishing websites using fuzzy logic," International Journal of Advanced Research in Computer Engineering Technology(IJARCET), vol. 5, no. 10, 2016.

[11] Verizon 2020 Data Breach Investigation Report, https://enterprise.verizon.com/resources/ reports/2020- databreachinvestigations- report.pdf

[12] N. Abdelhamid, A. Ayesh, and F. Thabtah, "Phishing detection based associative classification data mining," Expert Systems with Applications, vol. 41, no.13, pp. 5948-5959, 2014.

[13] L. Wenyin, G. Huang, L. Xiaoyue, Z. Min, and X. Deng, "Detection of phishing webpages based on visual similarity," Special interest tracks and posters of the 14th international conference on World Wide WebWWW 05, pp. 1060-1061, 2005.

[14] Kiruthiga, R., & Akila, D. (2019). Phishing websites detection using machine learning. International Journal of Recent Technology and Engineering, 8(2), 111-114.

[15] Alswailem, A., Alabdullah, B., Alrumayh, N., & Alsedrani, A. (2019, May). Detecting phishing websites using machine learning. In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-6).IEEE.

[16] P. Ashwini and N. Vadivelan, "Security from phishing attack on internet using evolving fuzzy neural network," CVR Journal of Science and Technology, vol. 20, no. 1, pp. 50-55, 2021.

[17] G. N. Anil, "Detection of phishing websites based on feature extraction using machine learning," International Research Journal of Engineering and Technology (IRJET), 2020.

[18] M. Romagna and N. J. van den Hout, "Hacktivism and website defacement: motivations, capabilities and potential threats," in 27th virus bulletin international conference, 2017, pp. 1-10.

[19] C. Seifert, I. Welch, and P. Komisarczuk, "Identification of malicious web pages with static heuristics," in 2008 Australasian Telecommunication Networks and Applications Conference, 2008, pp. 9196.

[20] S. Lee and J. Kim, "Warningbird: Detecting suspicious urls in twitter stream," in Ndss, 2012, pp. 113.

[21] B. Wei et al., "A deep-learning-driven light-weight phishing detection sensor," Sensors, vol. 19, no. 19, p. 4258, 2019.

P06RW008

An Approach to detect Parkinson's, Heart Disease and Diabetes using SVM and Logistic Regression

[1]Vengala Shivani, [2]Buran Sneha, [3]Sampeta Neeraj,

[4]Dr. Mummadi Ramachandra

[1,2,3]UG Student, SNIST, Hyderabad.

[4]Associate Professor, Department of CSE, SNIST, Hyderabad

Email: 20311A05e4@sreenidhi.edu.in, 20311A05f9@sreenidhi.edu.in, 20311A05g9@sreenidhi.edu.in,

ramachandram@sreenidhi.edu.in

Abstract:

An innovative healthcare approach, Multiple Disease Prediction through Machine Learning employs advanced algorithms to accurately forecast the likelihood of various illnesses in a patient, leveraging their medical history and relevant factors. This strategy aims to facilitate earlier diagnosis, enhance treatment efficacy, and ultimately, improve patient outcomes. Machine Learning has become indispensable in healthcare, especially for concurrent disease prediction, capitalizing on its ability to discern patterns and correlations within extensive patient datasets. Diseases such as Diabetes, Heart disease, and Parkinson's disease are among those this model can predict. Left untreated, these conditions pose significant risks to individuals and communities alike, underscoring the critical importance of early detection and diagnosis. Thus, maximizing the effectiveness of disease detection is paramount in mitigating their potential impact on public health.

I. Introduction:

The arrival of machine learning in medical field has yielded significant results in improving disease prediction. Notably, multiple disease prediction stands out as a significant application, employing machine learning algorithms to concurrently assess the likelihood of various diseases based on patient data. This approach harbors significant potential for facilitating early detection, crafting personalized treatment strategies, and fostering proactive healthcare management. Machine learning algorithms are especially adept at disease prediction tasks due to their ability to glean insights from extensive patient datasets, uncovering intricate patterns and correlations that may elude human clinicians. Machine learning (ML) represents one of the most rapidly expanding domains within computer science, finding widespread utility across various sectors. It entails the extraction of actionable insights from vast data sets. ML techniques find application in diverse fields such as medical diagnostics, marketing, industry, and scientific research. Within medical contexts, ML algorithms are particularly well-suited for data analysis tasks, given their adaptability and efficacy. Among the myriad forms of ML, our focus rests on classification methods, which excel in categorizing data into predefined groups and prognosticating future trends or activities owing to their commendable accuracy and performance. This study examines various parameters such as glucose levels, insulin levels and blood pressure by taking into consideration the age, gender and Body Mass Index for predicting diabetes. For heart disease prediction, it incorporates parameters like cp, trestbps, fbs, chol, restecg, thalach, and exang. Similarly, for Parkinson's prediction, it considers Average vocal fundamental frequency, Maximum vocal fundamental frequency, Minimum vocal fundamental frequency, MDVP jitter in percentage, and MDVP absolute jitter in ms, among others. Leveraging patient data available in the dataset, we train our machine learning models using Logistic Regression for heart disease prediction and SVM for

Parkinson's and Diabetes prediction. The choice of SVM algorithm facilitates the selection of the most accurate model among all, allowing for comparative analysis of their respective accuracies.

II. Algorithms used:

A. Support Vector Machine

The Support Vector Machine (SVM) is a clever tool that computers use to learn things. It's kind of like a super smart detective that can figure out if something belongs to one group or another. For instance, it can look at pictures and decide if they show cats or dogs. SVM is great at this because it can draw really good lines between different groups, like drawing a line between pictures of cats and pictures of dogs. These lines help the computer know where to put new pictures it hasn't seen before. SVM finds special points called support vectors that help it draw these lines perfectly. So, it's like SVM is a

detective with a special magnifying glass finding the best way to tell things apart. As a result, the algorithm earns its name - Support Vector Machine. The diagram below illustrates the classification of two distinct categories utilizing a decision boundary or hyperplane:

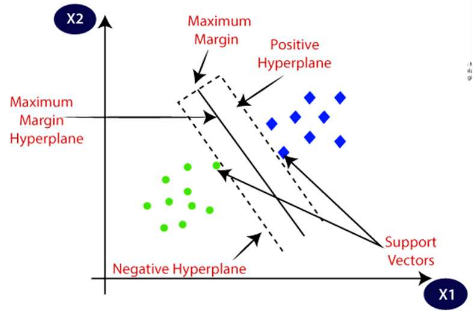


Fig. 1 Support Vector Machine

Consider this example: Imagine encountering a peculiar creature bearing characteristics of both cats and dogs, reminiscent of our previous discussion regarding the KNN classifier. To develop a model capable of accurately discerning whether it aligns more with a cat or a dog, SVM proves instrumental. Initially, we train the model with an extensive array of cat and dog images, allowing it to glean insights into the nuanced features of each species. Upon encountering the enigmatic creature, SVM harnesses support vectors to establish a decisive boundary between cat and dog data points. By pinpointing extreme cases (support vectors) within the dataset, SVM effectively evaluates the unique attributes of the creature and makes a categorical determination, potentially classifying it as a cat. Refer to the diagram below for a visual depiction:

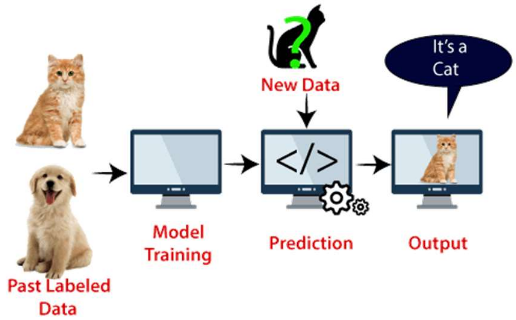


Fig. 2 Dog vs Cat classification

The SVM algorithm is applicable in tasks including face detection, image classification, text categorization, and beyond.

Classification of SVM:

Linear Support Vector: A Linear Support Vector is applied when dealing with linearly separable data, where classification into two classes can be achieved using a single straight line. In such instances, the classifier utilized is referred to as the Linear SVM classifier.

Non-linear Support Vector: Non-linear Support Vector is employed for data that is not linearly separable, meaning classification cannot be achieved with a straight line. In these cases, the classifier employed is termed the Non-linear SVM classifier.

B. Logistic Regression:

Logistic regression is a math tool used for predicting outcomes with only two possibilities, like yes or no. It helps us understand how different factors might influence this decision. We look at some known things (independent variables), like age or test scores, to guess what might happen with something else (the dependent variable), such as passing or failing a test. For example, we might use logistic regression to predict if someone will survive a car accident based on factors like seatbelt usage, car speed, and weather conditions. In school, we could use it to predict if a student will pass or fail an exam based on study hours, attendance, and previous test scores. Logistic regression is useful because it can estimate the chances of different outcomes happening.

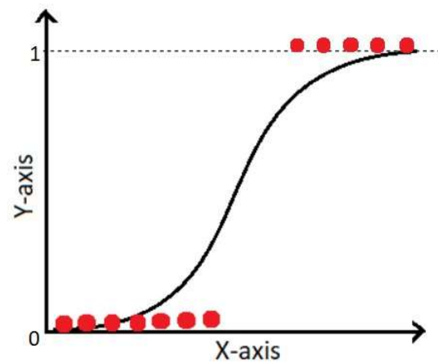


Fig.3 Logistic Regression

Classification of Logistic Regression:

Binary: This type is like when you're trying to guess if a coin will land on heads or tails. In logistic regression, we're predicting if something will happen or not. For example, we might use it to predict if a student will pass or fail an exam based on factors like study hours, attendance, and previous test scores.

Multinomial: Imagine you have a bag of colourful balls, and you want to guess which colour a randomly selected ball will be: red, blue, or green. Multinomial logistic regression helps us predict which option is most likely. It's useful when we have more than two choices and we want to figure out which one is the most probable outcome.

Ordinal: This type is like rating movies as "bad," "okay," or "great." Ordinal logistic regression helps us predict which category something falls into based on its characteristics or features. For instance, we might use it to predict the satisfaction level of customers with a product, where the options are "unsatisfied," "neutral," and "satisfied."

III. Literature survey:

In the analysis of many existing systems, the focus has been predominantly on individual diseases, limiting health assessments to one ailment at a time. Consequently, organizations tasked with analyzing patient health reports often find themselves implementing numerous models. However, this approach is constrained to analyzing specific diseases, leaving gaps in identifying the precise causes behind the escalating daily death rates. Complicating matters further, patients who have been successfully treated for one ailment may still be vulnerable to others. Some existing systems utilize only a handful of parameters for disease analysis, potentially overlooking crucial indicators and leading to undetected illnesses and their subsequent impacts. Hearing loss, a multifaceted issue, claims numerous lives annually, underscoring the urgency of comprehensive health monitoring. Neglecting the early warning signs of heart disease can precipitate severe consequences within a short timeframe, compounded by the stressors of modern-day living. The need of the hour is timely disease detection, enabling effective monitoring and intervention.

Let's delve deeper into Diabetes, Heart Disease, and Parkinson’s diseases:

1. Diabetes ranks among the most perilous diseases globally, potentially leading to a myriad of complications, including vision impairment. Leveraging machine learning techniques for diabetes detection offers a versatile and efficient approach to determining a patient's health status. The main intention of this review is to make a system that can identify diabetes in patients. Our methodology revolves around three key algorithms: Decision Tree, Naive Bayes, and SVM, with respective accuracies of 75%, 77%, and 76%. Additionally, we explore the application of the ANN algorithm post-training to assess the network's efficacy in correctly classifying the disease.

2. The main intention of this review is to make is to underscore the critical role of the heart in the vitality of living organisms. Given the profound implications of heart-related ailments, accurate diagnosis and prognosis are imperative to prevent cardiac fatalities. Machine learning and artificial intelligence offer invaluable support in predicting various natural events, including heart diseases. We used knn, Support Vector Machine, linear regression and decision trees employing data from the UCI archive for training and testing. Our comparative analysis revealed varying accuracies among the algorithms: SVM achieved 80% and knn achieved 87%.

3. Diagnosing Parkinson's disease early poses a considerable challenge, given the intricacies of symptom manifestation and clinical assessment, particularly motor symptoms. To address these complexities, machine learning methodologies have been leveraged for categorizing the patients against healthy habits thereby enhancing diagnostic precision and evaluation. A hybrid model integrating the Rotation and Random Forest algorithms were developed to categorize predictions into two different types: severe and non-severe. This model demonstrated promising accuracy rates, standing at 76.09% and 79.49%, respectively. These observations demonstrate the significance of machine learning in refining the healthcare and advancements of the Parkinson's disease.

IV. Problem statement:

The existing models in healthcare are tailored to focus on individual diseases per analysis. This limitation means that if a user seeks to predict multiple diseases, they must navigate across different platforms. Unfortunately, there is currently no unified system capable of conducting analyses for various diseases on a single system. Furthermore, few of these models show very less accuracy rates, posing significant risks to patient health. Consequently, when doctors aim to assess their patients' health reports, they are compelled to utilize multiple models, leading to increased costs and time expenditure. Similarly, organizations tasked with analyzing patient health reports face the challenge of deploying numerous models, resulting in escalated expenses and prolonged processing times. Additionally, some existing systems rely on a limited number of parameters, which may produce inaccurate results.

V. Proposed system

In the realm of disease prediction, the ability to predict various diseases simultaneously is an impeccable advantage. This eliminates the need for users to navigate across various platforms to predict individual diseases. Focusing on three prevalent diseases - Heart disease, Parkinson's Disease and Diabetes streamlines the process, reducing time expenditure significantly. By predicting multiple diseases concurrently, there emerges an opportunity to mitigate mortality rates effectively. Compared to existing systems, our approach offers expedited results and boasts numerous advantages, promising a more efficient and comprehensive disease prediction solution.

VI. System design:

This process delineates the modules, data, interfaces and components to ensure that the system reaches anticipated results. Essentially, it serves as the bridge between system proposition and product development. Object-oriented design and analysis methodologies are rapidly emerging as the preferred approaches for constructing computer systems.

Description: System design involves defining system elements such as modules, components, interfaces, and data based on predetermined criteria. It entails the meticulous definition, development, and design of systems tailored to meet the unique requirements of a business or organization. A systematic approach is essential for the creation of a cohesive and highperforming system. Both top-down and bottom-up approaches are employed to comprehensively consider all relevant system variables. Developers utilize modeling languages to systematically express information and knowledge, adhering to predefined rules and structures.

Few examples of the mentioned languages include:

a) Unified Modeling Language (UML), which describes the function and structure of the software through graphical notations.

b) Flowcharts, providing a structural representation of algorithms.

c) Business Process Modelling (BPM) Notation is used in process modelling.

d) SysML or the Systems Modelling Language is employed for the design of the system.

Design styles encompass:

Architectural design: This focuses on how the system is built, like its structure and models.

Logical design: It's about how information moves and the things you can do with the system.

Physical design: This deals with how data is organized and stored, and how it moves around the system.

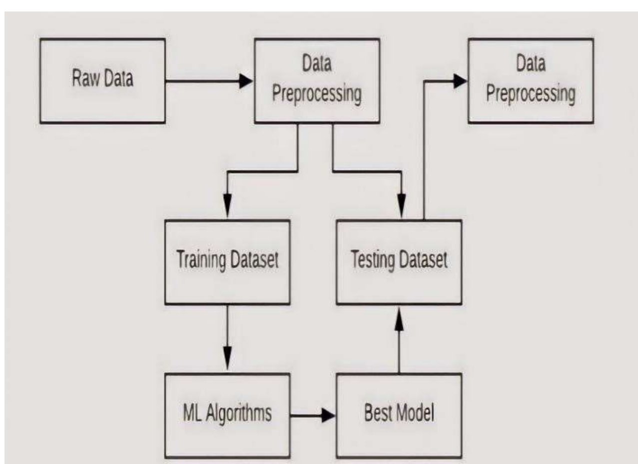


Fig. 4 System Architecture

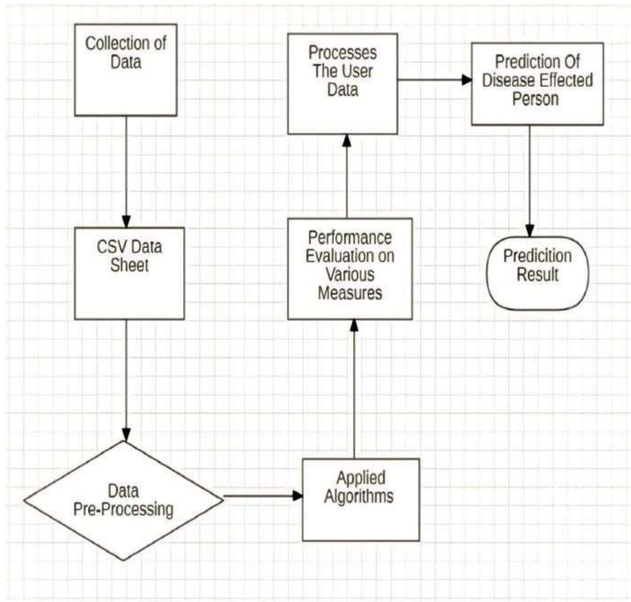


Fig. 5 Flow Diagram

VII. Results:

S. No	Disease name	Algorithm	Accuracy of existing system	Accuracy of proposed system
1.	Diabetes	SVM	76	78
2.	Heart disease	Logistic Regression	80	85
3.	Parkinson’s disease	SVM	71	87

Following the creation of the application using flask, the model can be integrated into the frontend. Upon clicking Test result, the users can promptly predict whether the case presents with the ailment or not.

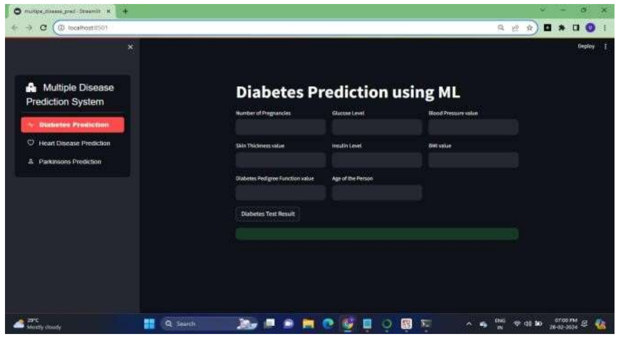


Fig. 6 Diabetes Prediction

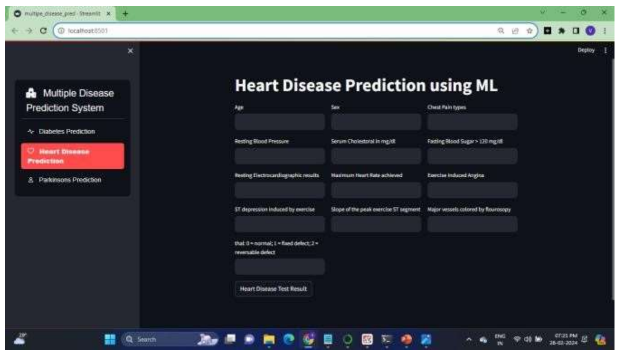


Fig. 7 Heart Disease Prediction

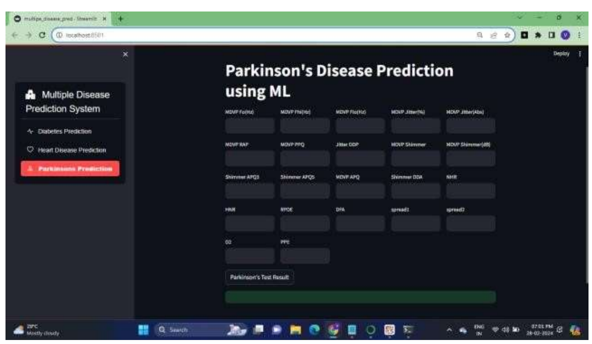


Fig. 8 Parkinson's Prediction

VIII. Conclusion:

The utilization of various ML algorithms has enabled early diagnosis of various diseases such as to diabetes, Parkinson's disease and Heart disease. Algorithms such as SVM and Logistic Regression have been extensively employed, primarily due to their effectiveness in predicting accuracy, a crucial performance metric. The SVM model has demonstrated superior accuracy, particularly in managing any kind of data pertinent to diabetes and Parkinson's disease. Meanwhile, Logistic Regression emerged as the most reliable algorithm for predicting heart conditions. Multiple disease prediction using machine learning represents a promising frontier in healthcare, poised to transform diagnostic and treatment paradigms. By leveraging machine learning algorithms to analyze vast quantities of patient data, we can uncover intricate patterns and correlations that may elude human clinicians. This approach holds the potential to facilitate earlier diagnoses, optimize treatment strategies, and ultimately enhance patient outcomes.

IX. Future scope:

Diving into the future could potentially help us understand how SVM can be used to find predictive features with unsupervised learning. Additionally, this analysis only utilized healthy cases. A machine learning model with comparable accuracy, sensitivity, and recall could be achievable by incorporating a model that accounts for multiple correlated diagnoses. As more sophisticated ML algorithms are developed in the future, there is much anticipation for improvements in disease prediction.

X. References:

[1] "Disease Prediction by Machine Learning from Healthcare Communities," International Journal of Scientific Research in Science and Technology, pp. 29–35,2019 by S. Jadhav, R. Kasar, N. Lade, M. Patil, and S. Kolte.

[2] "Comparing different supervised machine learning algorithms for disease prediction," BMC Medical Informatics and Decision Making, vol. 19, no. 1, pp. 1 16, 2019 by S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni.

[3] "Prediction of Heart Disease using Machine Learning Algorithms" in 2022, International Journal for Research in Applied Science and Engineering Technology (IJRASET), ISSN : 2321-9653 by Shriniket Dixit, Pilla Vaishno Mohan, Shrishail Ravi Terni.

P09RW017

A data hiding scheme based on the difference of image interpolation algorithms

K. Sai Souhith Reddy, M.D.Naveed Khan, A.Kanishq, UG Student,

Meeravali Shaik, Research Guide

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD

Abstract:

Interpolation is a mature technique used in image processing and many data hiding schemes based on image interpolation have been proposed. In this paper, a novel efficient data hiding scheme based on the difference between image interpolation algorithms is proposed. The interpolated values obtained by neighboring points using different interpolation algorithms are generally different, and the difference can be used for encoding secret data. 2k different interpolated values calculated by 2k different interpolation algorithms respectively in the ideal case, and they are numbered in increasing order to correspond to the k bits secret data in each interpolated pixel. The system developed an image enhancement technique for enhancing or improving the original image pixel quality. After that, we can implement the interpolation to resize the original image. Then, we can encrypt the original image by using the AES algorithm. Here, we can hide the data or secret message in the image.

Keywords: Interpolation , Convolutional Neural Network Image processing, image enhancement technique,AES algorithm.

1.Introduction:

With the development of security needs and success in some domains, such as copyright protection, information appending, tamper-proof, information tracking, data hiding technology and steganography is booming. Especially in the field of digital image, to ensure the safe transmission and other additional secret information, so that specific recipients can fully identify and extract the secret message without being exposed to other third parties who steal the image. Researchers have proposed some different methods to improve the efficiency and confidentiality of data hiding technology. Image data hiding is a technology to embed the processed secret information into digital image for public transmission. Because of the redundancy of the digital information itself and the relatively weak sensitivity of human senses, it is impossible to detect whether the secret information is embedded to realize the safe public transmission carrier. According to the basic operation of image data hiding, there are generally two methods: insert (the spare space vacated by image pixel redundancy or the identification part of image), such as insert in the com- pressed space, and replace (modify image pixel byte or adjust), such as LSB least-significant bit substitution. In terms of the domain in which encryption is performed, there are generally two categories: based on the spatial and frequency do- main. Categories spatial-based process pixel values directly in the image space using length as an independent variable, including the data hiding technology based on difference extension put forward and other difference expansion techniques developed based on it.Image interpolation is to use the known gray values of adjacent pixels (or the three-color values in RGB images) to generate the gray values of unknown pixels, so that the original image can be regenerated with higher resolution images. The data hiding scheme based on interpolation mainly uses the value of the original pixel as a reference to predict the value of the new pixel when the size of the image is expanded by image interpolation, and the secret bits are embedded in the interpolated pixel during the interpolation process of prediction. First proposed the data hiding technology based on neighborhood mean interpolation (NMI). The algorithm mainly determines the amount of data that can be hidden by the non-reference pixels through the difference between the non-reference pixel and the selected reference pixel.

Literature Survey:

1.Title: Secure Data Hiding Technique Using Video Steganography and Watermarking - A Review

Author: Shivani Khosla1, Paramjeet Kaur

The rapid development of data transfer through the internet made it easier to send the data accurately and faster to the destination. In addition to this, anyone can modify and misuse valuable information through hacking at the same time. This paper presents video steganography with digital watermarking techniques as an efficient and robust tool for protection. Here consider video as a set of frames or images and any changes in the output image by hidden data is not visually recognizable. This paper provides a review and analysis of the different existing methods of steganography and digital watermarking along with some common standards and guidelines drawn from the literature.DWT is used for digital images. Many DWTs are available. Depending on the application appropriate one should be used. The simplest transform is the haar transform. To hide text message integer wavelet transform can be used. When a DWT transform is applied to an image it is decomposed into 4 sub bands: LL, HL, LH and HH. The LL part contains the most significant features. So if the information is hidden in the LL part the stego image can withstand compression or other manipulations. But sometimes distortion may be produced in the stego image and then other sub bands can be used.

Advantages:

Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy.

Disadvantages:

The output image by hidden data is not recognizable.

2.Title: Universal Steganalysis using Feature Selection Strategy for Higher Order Image Statistics (2013)

Author: Sonali S.Ekhande Prof .S.P.Sonavane Dr.P.J .Kulkarni

The purpose of image steganalysis is to detect the presence of hidden messages in cover photographic images. Supervised learning is an effective and commonly used method to cope with difficulties of unknown image statistics and unknown steganography. Present paper proposes; a universal approach for steganalysis for detecting presence of hidden messages embedded within digital images. This paper describes wavelet-like decomposition to build higher order statistical models of natural images. Feature selection techniques like ANOVA are used to select relevant features. SVM are then used to discriminate between clean and stego images. Study of the effect of relevant features on classification accuracy may help to improve the complexity. The nonlinear classification, contrary to the arbitrarily complicated non-linear classification techniques such as the neural network, is achieved by first embedding training data into a higher (possibly infinite) dimensional space. A linear separation is then found in that space by the linear SVM algorithm and is mapped back to the original data space as a non-linear classification surface. Such a non-linear classification, though more flexible, inherits the stability and generalization ability of linear SVM, thus effectively reducing the chance of overfitting the training data. However, outguess has a high false alarm .

Advantages:

Feature selection strategy implemented here gives relevant features to be used for training and thus reduces the training complexity.

Disadvantages:

It does not directly obtain the feature importance.

3.Title:An Extended Visual Cryptography Scheme Without Pixel Expansion For Halftone Images (2013)

Author: N. Askari, H.M. Heys, and C.R. Moloney

Visual cryptography is a secret sharing scheme which uses images distributed as shares such that, when the shares are superimposed, a hidden secret image is revealed.In extended visual cryptography, the shared images are constructed to contain meaningful cover images, thereby providing opportunities for integrating visual cryptography and biometric security techniques. In this paper, we propose a method for processing halftone images that improves the quality of the shared images and the recovered secret image in an extended visual cryptography scheme for which the size of the shared images and the recovered image is the same as for the original halftone secret image. The resulting scheme maintains the perfect security of the original extended visual cryptography approach.Although visual cryptography operates on binary images, it can be applied to grayscale images by using a halftoning algorithm to first convert the grayscale image to a binary image. This allows for use of visual cryptography schemes to biometric images which are naturally and meaningfully grayscale, such as facial images. Hence, using halftoning techniques to convert grayscale images to binary images is a useful pre-processing step for visual cryptography

Advantages:

A processed image contains white and black blocks and can be used as an input secret image in any visual cryptography encoding process.

Disadvantages:

As the new scheme does not change the share generation approach.

4.Title: Analysis Of Image Steganalysis Techniques To Defend Against Statistical Attacks – A Survey (2012)

Author: Usha B.A1, N K Srinath2, N K Cauvery3

Steganography is the art of concealing information to transmit it in such a way that nobody but the intended receiver knows the existence of the message. Steganalysis techniques work on eliminating suspicion about the existence of a message. If suspicion is raised, then the message cannot be passed covertly. One of the ways to detect the hidden message is to view the statistical properties of the image or medium in which the message is hidden. This is called a statistical attack. In this paper, we explain the nature of such attacks and present our conclusions based on reviews of existing methods of defense against statistical attacks.JPEG images use the Discrete Cosine Transform (DCT) to achieve image compression. The compressed data is stored as integers; however, if we need to quantize the data to encode a message, all the calculations required involve floating point data. Information is hidden in the JPEG image by modulating the rounding choices either up or down in the DCT coefficients. Detection of such an embedded message would seem to be quite difficult. In this rounding off, errors may occur and this leads to the losses in this method. The tool Jpeg-Jsteg is a steganography tool that makes use of this property.

Advantages:

The methods used in the status quo are sufficiently advanced and can provide suitable defense against current attacks.

Disadvantages:

The generalized $\chi 2$ attack does not calculate an estimation of the message length and can be sometimes wrong if the message has a significant difference in the number of zeros compared to ones.

5. Title: Data Security and Authentication Using Steganography (2011)

Author: Ravi Kumar. B #, Murti. P.R.K

Steganography is the art of covered, or hidden, writing. The purpose of steganography is covert communication to hide the existence of a message from a third party. This proposed system deals with implementing security-using Steganography. In this technology, the end

user identifies an image which is going to act as the carrier of data. The data file is encrypted and authenticated. This message is hidden in the image. The image if hacked or interpreted by a third party user will open up in any image previewer but not displaying the data. This protects the data from being invisible and hence be secure during transmission. The user in the receiving end uses another piece of code to retrieve the data from the image.In our method we use the technique of hiding the data with an image file, the visibility of the image, resolution or clarity is not being affected. The hidden data can be of length in size. To the hacker, only, the image is made going to be visible when previewed and not a trace of the hidden data. If the image file is opened across a text editor, then also the data is not going to be visible as the information is stored in an encryption form, which is also binary.
Hence making it difficult for the enclosure to differentiate the data to the image file.

Advantages:

The patchwork approach is used independent of the host image and proves to be quite robust as the hidden message can survive conversion between lossy and lossless compression.

Disadvantages:

If the image already contains some data you cannot add some more data for the same image.

6. Title: Digital Watermarking and Other Data Hiding Techniques (2013)

Author: Gurpreet Kaur, Kamaljeet Kaur

Digital watermarking is not a new name in the technology world but there are different techniques in data hiding which are similar to watermarking. In this paper we compare digital watermarking with other techniques of data hiding.Steganography, Fingerprinting, cryptography and Digital signature techniques are compared with watermarking. We need water-marking for digital data security .It provides ownership assertion, authentication and integrity verification, usage control and con-tent labeling. Steganography is derived from the Greek for covered writing and essentially means “to hide in plain sight”. Steganography is the art and science of communicating in such a way that the presence of a message cannot be detected. Simple steganographic techniques have been in use for hundreds of years, but with the increasing use of files in an electronic format new techniques for information hiding have become possible. This document will examine some early examples of steganography and the general principles behind its usage. We will then look at why it has become such an important issue in recent years. There will then be a discussion of some specific techniques for hiding information in a variety of files and the attacks that may be used to bypass steganography.

Advantages:

A grille cipher employs a template that is used to cover the carrier message.

Disadvantages:

It is not visible or perceivable, but it can be detected by different means.

3.Proposed system:

In the proposed system, the image dataset was taken as input. Then, we have to implement the pre-processing step. In this step, we have to resize the original image as well as grayscale conversion. Then, we can implement the interpolation techniques such as bilinear interpolation. Then, we can enhance the original image pixel quality by using histogram equalization. Then, we can encrypt the image by using the AES algorithm. After that, we can hide the original message in encrypted data. After that, we can extract the original image and original data from the embedded image.

A. CNN

Convolutional layers are used to build a hierarchy of features in convolutional neural networks (CNNs), which are preferred to recognize patterns in data. Each neuron in the convolutional layers has a limited number of connections to the neurons in the layer below it and has a local receptive field. Without having to learn the complete dataset, this enables the network to recognize patterns in the data. Object detection, language processing, and image recognition are just a few of the tasks that CNNs are utilized for. CNN automatically extracts image characteristics. Before adding a prediction layer, CNN effectively convolutionally downscales the image employing information from surrounding pixels. CNN delivers higher accuracy and performs effectively. One will be able to understand the patterns of natural photographs using a deep neural network. More pixels can be buried there since the network will be able to determine which areas are unnecessary. The amount of concealed data can be raised by reducing the amount of space on unused regions. The network will conceal the data that no one can access without the weights since the structure and weights can be modified.

B. Image Preprocessing

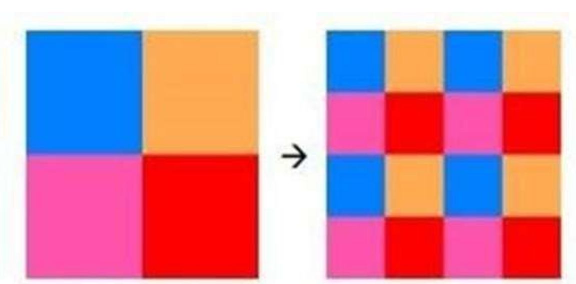
In our process, we have to resize the image and convert the image into gray scale. One can enlarge an image by using the resize () function on it and sending a two-integer tuple parameter that describes the image's final width and height. The function returns a new image with the revised dimensions rather than altering the original image. Now the final image needs to be converted into gray scale image. The procedure of rescaling demands changing an image from another color space, such as RGB etc., to a variety of grayscales. We can convert it by reading an image by imread() and then convert it by cv2.cvtColor(). Additionally, by applying the common RGB to grayscale conversion algorithm, one can transform an image to grayscale image Gray= 0.5870 * G + 0.1140 * B+ 0.2989 * R.

C. Image Interpolation

The resampling method known as bilinear interpolation computes an unique pixel value utilizing distance-weighted mean of the four nearest pixels. Based on their proximity to the output processing cell centers, the four input raster cell centers will be averaged and weighted. Bilinear interpolation results in smoother interpolation than the nearest neighbor technique.

D. Architecture

The network design resembles Auto-Encoders in certain ways. After several changes, to recreate the input, auto- encoders are frequently used. By doing this, they receive de- tails about the input distribution's features. The recommended architecture in our case is only a little different, though. In addition to creating images, the network must also hide some of them while creating other equalization into our technique.



Pixel Permutation

Histogram equalization: Histogram equalization is used to enhance contrast in photographs. This is accomplished by significantly increasing the image's intensity values and dispersing the most prevalent intensity levels. This tactic often enhances the overall contrast of images when the vital information is depicted by close values. This makes it possible for areas that have less local contrast to obtain more contrast. Prep Layer is used to prepare the secret image that will be concealed. This layer performs numerous tasks. Before anything else, if the hidden picture is less than the cover image, its size is expanded to match the cover image's size, distributing its bits equally over all pixels. In addition, for all hidden picture measurements, the objective is to reassemble the pixels into characteristics that are used. To prevent any secret data leakage, the layer's major goal was to successfully integrate the masked image into the encrypted image. In order to create the Container Image, the result of the PrepLayer is passed into the Hiding Layer's second layer along with a cover image. Both the RGB channels of the primary image and the rebuilt fields of a hidden picture are inputs for this network, which uses square pixel pictures. Transmitter

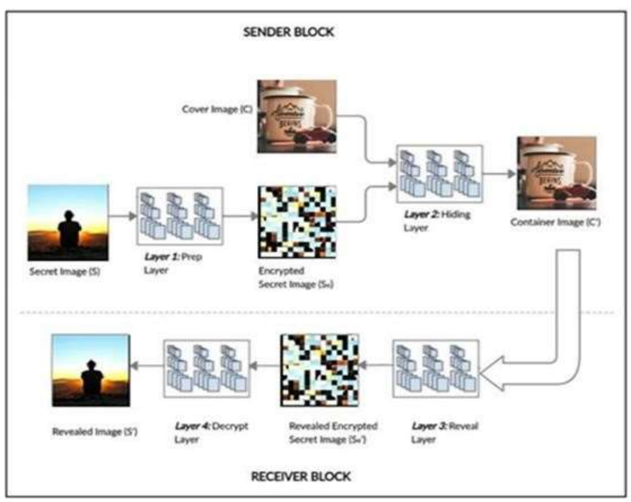


Fig. 1. Structure of the proposed system

Unit includes these two levels. The generated Container image is accessible to the receiver. The receiver produces the encrypted image using Reveal Layer. This layer subtracts the cover image from the supplied Container image to produce the ciphered image.

The Decrypt level, which takes the Reveal level's result and decodes the image, is what finally makes the secret image visible. Fourth and third levels combined form the receiver block.

Increasing the weight parameters will cause the loss function to be smaller, which is the optimizer's task. Finding the global minima of the loss function is aided by this. A few examples of the many different kinds of optimizers are Momentum, Nesterov ,Adagrad, Adadelta, Adam, and others. Different learning rates are established for each parameter via the model's adaptive moment estimation optimizer (ADAM). Due to its high computational efficiency and little memory requirements, it is perfect for our model. The best solution is the ADAM optimizer, which performs better than RMSprop, Adagrad, and Adadelta, among other adaptive learning strate- gies.

Only the reveal layer contains the first of the two types of losses, whereas the entire model contains the second. Utilizing layer loss, one may figure out how much the released secret picture has changed from the initial hidden picture. The distinction between the first hidden secret picture and the disclosed secret picture is essentially what it is.

The variation in MSE between the initial hidden and the revealed images, as well as the conventional MSE between the initial cover and container images, constitute the loss. The amount by which the secret picture should be restored is controlled by a hyper-parameter. The supplied method is distinct, allowing for persistent training of the neural network.

E. Image Encryption

Since they are easier to process than text, images are frequently utilized in communication. The user may engage on a channel that has been infiltrated in these circumstances, in- formation security becomes crucial. Encryption, or the process of transforming a picture into another that is challenging to comprehend, is essential for ensuring the security of the data being broadcasted. The military, security services, and other industries, as well as any place where sensitive information or confidential data is present, can use image encryption. The SSIM is utilized to assess the quality of the interpolated picture. By computing the logarithmic mean square error be- tween the initial picture and the interpolated picture, the PSNR is a statistic which assesses the aesthetic magnificence of the interpolated picture after incorporating secret information. The distortion of the interpolated image after incorporating secret information decreases with increasing PSNR.

An index to gauge how similar two images are made up of brightness, contrast, and structure is called structural similarity(SSIM). In the case of identical photos, SSIM has a value of 1

$$L(c,c',s,s') = \left| c - c' \right| + \beta \left| s - s' \right|$$

To determine the average of error squares, also known as the average squared difference between the initial and the embedded images, an estimator's mean squared error is employed (MSE). The expected value of a loss from a squared error is represented by this risk function. Since it is always positive, values close to 0 are desirable.

#IMPORT PACKAGES

```
import numpy as np
import matplotlib.pyplot as plt
from tkinter.filedialog import askopenfilename
import cv2
import matplotlib.image as mpimg
import warnings
warnings.filterwarnings('ignore')
```

STEP 1

```
from google.colab import files
import matplotlib.image as mpimg
import matplotlib.pyplot as plt

print("------------------------------------------")
print("Step 1 ------> Input Image ")
print("------------------------------------------")
print()

# Upload the image file
uploaded = files.upload()

# Check if any file was uploaded
if len(uploaded) > 0:
    # Get the filename of the uploaded file
    filename = list(uploaded.keys())[0]

    # Read the image using matplotlib
    image = mpimg.imread(filename)

    # Display the image
    plt.imshow(image)
    plt.title('Original Image')
    plt.axis('off')
    plt.show()
```

```
--------------------------------------------
Step 1 ------> Input Image
--------------------------------------------
```

Choose files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving 4.png to 4.png
```

Original Image

Double-click (or enter) to edit

2.IMAGE PREPROCESSING

```
from matplotlib import pyplot as plt
```

RESIZE IMAGE

```
print("-------------------------------------------")
print("Step 2 ------> Preprocessing ")
print("-------------------------------------------")
print()

resized_image = cv2.resize(image,(300,300))
img_resize_orig = cv2.resize(image,((50, 50)))

plt.imshow(resized_image)
plt.title('Resized Image')
plt.axis ('off')
plt.show()

print()
print("------------------------------------------------------------------")
print()
```

```
-------------------------------------------
Step 2 ------> Preprocessing
-------------------------------------------
```


```
------------------------------------------------------------------
```

GRAYSCALE IMAGE

```
import cv2
import matplotlib.pyplot as plt

# Assuming you already defined the 'image' variable containing the image data
# If not, make sure you define 'image' before using it in the code

# Assuming img_resize_orig is defined elsewhere in your code
try:
    gray1 = cv2.cvtColor(img_resize_orig, cv2.COLOR_BGR2GRAY)
except NameError:
    print("img_resize_orig is not defined. Skipping grayscale conversion.")
    gray1 = None

# Display the grayscale image if available
if gray1 is not None:
    plt.imshow(gray1, cmap='gray')
    plt.title('Gray Scale Image')
    plt.axis ('off')
    plt.show()
else:
    print("Gray scale image not available.")

print("--------------------------------------------")
print("Step 3 ------> Image Interpolation ")
print("--------------------------------------------")
print()

# Assuming 'image' variable contains the original image data
try:
    bilinear_img = cv2.resize(image, None, fx=10, fy=10, interpolation=cv2.INTER_LIN
    # Display the bilinear interpolated image
    plt.imshow(bilinear_img)
    plt.title('Bilinear Interpolation')
    plt.axis ('off')
    plt.show()
except NameError:
    print("image is not defined. Skipping image interpolation.")
```


```
-------------------------------------------
Step 3 ------> Image Interpolation
-------------------------------------------
```

IMAGE ENHANCEMENT

```
import numpy as np
import matplotlib
from skimage import img_as_float
from skimage import exposure


matplotlib.rcParams['font.size'] = 8


def plot_img_and_hist(image, axes, bins=256):

    image = img_as_float(image)
    ax_img, ax_hist = axes
    ax_cdf = ax_hist.twinx()

    # Display image
    ax_img.imshow(image, cmap=plt.cm.gray)
    ax_img.set_axis_off()

    # Display histogram
    ax_hist.hist(image.ravel(), bins=bins, histtype='step', color='black')
    ax_hist.ticklabel_format(axis='y', style='scientific', scilimits=(0, 0))
    ax_hist.set_xlabel('Pixel intensity')
    ax_hist.set_xlim(0, 1)
    ax_hist.set_yticks([])

    # Display cumulative distribution
    img_cdf, bins = exposure.cumulative_distribution(image, bins)
    ax_cdf.plot(bins, img_cdf, 'r')
    ax_cdf.set_yticks([])

    return ax_img, ax_hist, ax_cdf


# Contrast stretching
p2, p98 = np.percentile(image, (2, 98))
img_con = exposure.rescale_intensity(image, in_range=(p2, p98))

# Equalization
img_eq = exposure.equalize_hist(image)

# Adaptive Equalization
img_adapteq = exposure.equalize_adapthist(image, clip_limit=0.03)


plt.imshow(image)
plt.title('Low contrast image')
plt.axis ('off')
plt.show()
```

```
plt.imshow(img_con)
plt.title('Contrast stretching')
plt.axis ('off')
plt.show()


plt.imshow(img_eq)
plt.title('Histogram equalization')
plt.axis ('off')
plt.show()

plt.imshow(img_eq)
plt.title('Histogram equalization')
plt.axis ('off')
plt.show()
```

Low contrast image

Contrast stretching

Histogram equalization

Histogram equalization

IMAGE ENCRYPTION

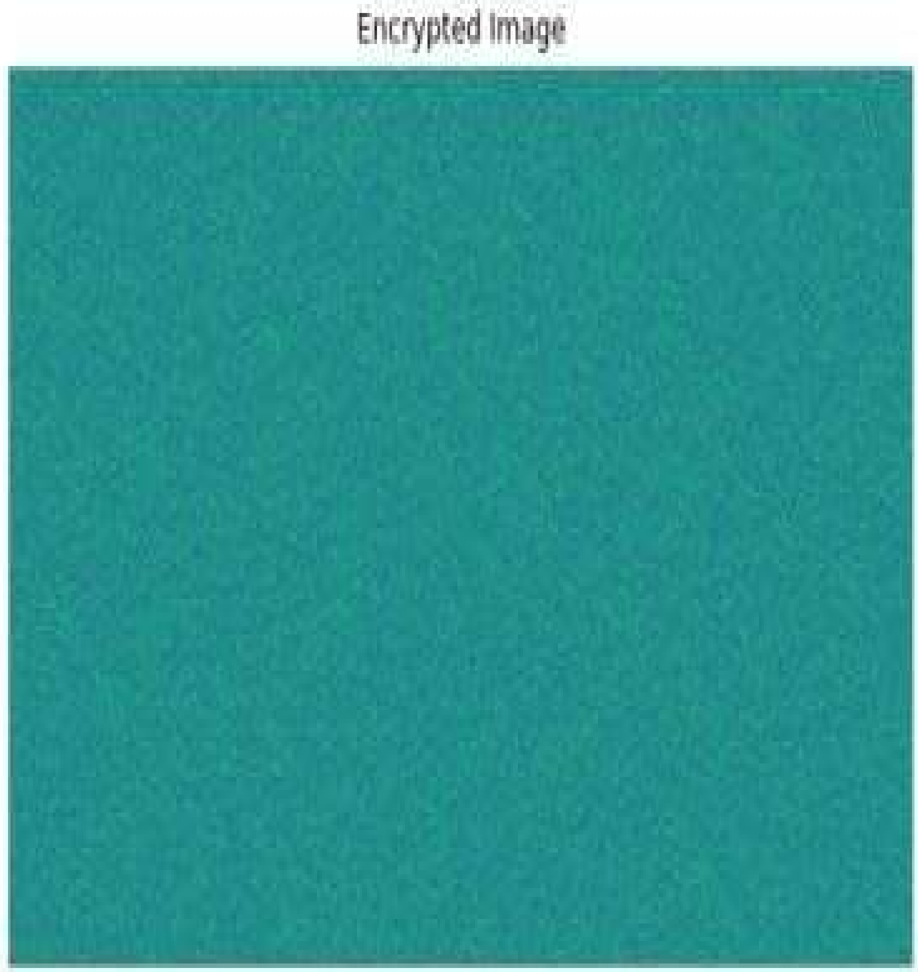
```
!pip install pycryptodome

    Requirement already satisfied: pycryptodome in /usr/local/lib/python3.10/dist-pa
```

```
from Crypto.Cipher import AES
import cv2


print("-------------------------------------------")
print("Step 4 ------> Image Encryption (AES)  ")
print("-------------------------------------------")
print()


# na = np.array(image)

x1, y1 ,pp= image.shape[:3]

gray_conversion = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

reshape_img= np.array(range(x1*y1), int).reshape((x1, y1))

enc_img= np.array(range(x1*y1), int).reshape((x1, y1))

reshape_img[:,:]=gray_conversion[:, :]

print("=======================================")
print(" Key Must Be SIxteen Characters  ")
print("=======================================")

key_get=input(str("Enter the key : "))

key = bytes(key_get, 'utf-8')

# key = b'image encryption'

byte_len=b'0000000000000000'

cipher = AES.new(key, AES.MODE_CFB, byte_len)

L2=[]
res_enc = np.array(range(x1),int)
for i in range(x1):
    res_enc=reshape_img[i,:].tolist()
    res_enc1=bytes(res_enc)
    msg =  cipher.encrypt(res_enc1)
    for p in msg:
        L2 += [(p)]
    enc_img[i,:]=L2[:]
    L2=[]

plt.imshow(enc_img)
plt.title('Encrypted Image')
plt.axis ('off')
plt.show()
```



```
from PIL import Image

def generateData(data1):
        newdata = []
        for ij in data1:
            newdata.append(format(ord(ij), '08b'))
        return newdata


def modulePixx(pixxel, dataa1):
    datalist = generateData(dataa1)
    length_data = len(datalist)
    imdata = iter(pixxel)
    for i in range(length_data):

        pixx = [value for value in imdata.__next__()[:3] +
                                   imdata.__next__()[:3] +
                                   imdata.__next__()[:3]]


        for j in range(0, 8):
            if (datalist[i][j] == '0' and pixx[j]% 2 != 0):
                pixx[j] -= 1

            elif (datalist[i][j] == '1' and pixx[j] % 2 == 0):
                if(pixx[j] != 0):
                    pixx[j] -= 1
                else:
                    pixx[j] += 1

        if (i == length_data - 1):
            if (pixx[-1] % 2 == 0):
                if(pixx[-1] != 0):
                    pixx[-1] -= 1
                else:
                    pixx[-1] += 1

        else:
            if (pixx[-1] % 2 != 0):
                pixx[-1] -= 1

        pixx = tuple(pixx)
        yield pixx[0:3]
        yield pixx[3:6]
        yield pixx[6:9]

def encode_encrypt(newimg, data):
    w = newimg.size[0]
    (x, y) = (0, 0)

    for pixel in modulePixx(newimg.getdata(), data):
```

```
---------------------------------------
Step 4 ------> Image Encryption (AES)
---------------------------------------

==================================
 Key Must Be SIxteen Characters
==================================
Enter the key : kanishqashamolla
```



DATA HIDING


```
from PIL import Image

def generateData(data1):
        newdata = []
        for ij in data1:
            newdata.append(format(ord(ij), '08b'))
        return newdata


def modulePixx(pixxel, dataa1):
    datalist = generateData(dataa1)
    length_data = len(datalist)
    imdata = iter(pixxel)
    for i in range(length_data):

        pixx = [value for value in imdata.__next__()[:3] +
                                imdata.__next__()[:3] +
                                imdata.__next__()[:3]]


        for j in range(0, 8):
            if (datalist[i][j] == '0' and pixx[j]% 2 != 0):
                pixx[j] -= 1

            elif (datalist[i][j] == '1' and pixx[j] % 2 == 0):
                if(pixx[j] != 0):
                    pixx[j] -= 1
                else:
                    pixx[j] += 1

        if (i == length_data - 1):
            if (pixx[-1] % 2 == 0):
                if(pixx[-1] != 0):
                    pixx[-1] -= 1
                else:
                    pixx[-1] += 1

        else:
            if (pixx[-1] % 2 != 0):
                pixx[-1] -= 1

        pixx = tuple(pixx)
        yield pixx[0:3]
        yield pixx[3:6]
        yield pixx[6:9]

def encode_encrypt(newimg, data):
    w = newimg.size[0]
    (x, y) = (0, 0)

    for pixel in modulePixx(newimg.getdata(), data):
```

```
        newimg.putpixel((x, y), pixel)
        if (x == w - 1):
            x = 0
            y += 1
        else:
            x += 1
```

ENCODE IMAGE

```
def encode():
    img = input("1. Enter image name for encryption (with extension) : ")
    image = Image.open(img, 'r')
    print()
    print()
    data = input("2. Enter Message or Data to be encoded : ")
    if (len(data) == 0):
        raise ValueError('Data is empty')

    newimg = image.copy()
    encode_encrypt(newimg, data)
    print()
    print()
    new_img_name = input("3. Enter Output image name (with extension) : ")
    newimg.save(new_img_name, str(new_img_name.split(".")[1].upper()))
    image = mpimg.imread(new_img_name)
    plt.imshow(image)
    plt.title('Embedded Image')
    plt.axis ('off')
    plt.show()
```

DECODE IMAGE

```
def decode():
    img = input("Enter Encrypted name for decryption (with extension) : ")
    image = Image.open(img, 'r')
    plt.imshow(image)
    plt.title('Original Image')
    plt.axis ('off')
    plt.show()

    data = ''
    imgdata = iter(image.getdata())

    while (True):
        pixels = [value for value in imgdata.__next__()[:3] +
                                  imgdata.__next__()[:3] +
```

5. Conclusion:

We conclude that the images were taken from a dataset repository as input. We develop or hide the original message into the original image. Then, we encrypted the image by using the AES algorithm. Then, we enhance the original image for more quality and extract the data into the original image.

6. Future scope:

Let's start by exploring the prospect of training a network to locate concealed images after the system has been installed and without access to the original network. In this part, it is discussed where future research might go. In principle, more sophisticated encryption techniques could replace the way we now encode data. The justification for avoiding AES or DES methods is the difficulties in retrieving the encrypted image caused by lossy neural networks. We will be able to construct lossless neural networks using modern cryptographic methods if they are created in the future. The development of cryptographic techniques that could be applied to neural networks should be the main emphasis of future study. Additionally, there are numerous possibilities when neural networks and image steganography are combined. Audio files can also be encrypted with LSB steganography. So, we can also combine neural networks with audio samples. Future integration of neural networks with steganography methods other than LSB steganography is also possible.

7. References:

[1] Yang CH. Inverted pattern approach to improve image quality of information hiding by lsb substitution. Pattern Recognition 2008;41:2674–83.

[2] Tian J. Reversible data embedding using a difference expansion. In: IEEE transactions on circuits systems for video technology. 13; 2003.

[3] Fridrich J, Goljan M, Du R. Lossless data embedding"new paradigm in digital watermarking. Eurasip J Adv Signal Process 2002:986842. 2020.

[4] Celik MU, Sharma G, Tekalp AM, Saber E. Lossless generalized- lsb data embedding. IEEE Trans Image Process 2005;12:253–66.

[5]] Chang CC, Lin MH, Hu YC. A fast and secure image hiding scheme based on lsb substitution. Int J Pattern Recognition Artif Intell 2002;16:399–416.

[6] Chan CK, Cheng LM. Hiding data in images by simple lsb substitution. Pattern Recognit 2004;37:469–74.

[7] Yang CH. Inverted pattern approach to improve image quality of information hiding by lsb substitution. Pattern Recognition 2008;41:2674–83.

[8] Tian J. Reversible data embedding using a difference expansion. In: IEEE transactions on circuits systems for video technology. 13; 2003. p. 890–6.

P10RW018

Comparative Analysis for Classification and Detection of Multiple Skin Diseases using Deep Learning Algorithms

[1] Shivani Keerthigari, [2]Yogeetha Poshettywar,[3] Srilekha Karnala,[4] Padmini Kuppala

[1,2,3,4] Department of CSE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

[1] keerthigari.shivani@gmail.com, [2] yogeethadadmom@gmail.com,

[3] karnalasrilekha@gmail.com, [4] padmini24u@gmail.com

Abstract:

The project focuses on aiding medical professionals in diagnosing skin diseases through dermatological image anal- ysis, employing a systematic workflow spanning data collection, pre-processing, feature extraction, and disease classification. Uti- lizing advanced machine learning techniques like Convolutional Neural Networks (CNNs), Efficient Net and Inception ResNet algorithms, the pipeline ensures accurate disease classification. Moreover, an intuitive interface facilitates easy upload of derma- tological images, enhancing user experience and expediting diag- nosis. Leveraging image processing, the interface improves image quality by eliminating noise and isolating relevant features crucial for disease differentiation. Comparative analysis evaluates the performance of these algorithms in classifying various skin dis-eases such as actinic keratosis, benign keratosis, dermatofibroma, melanocytic nevus, and melanoma.

Assessment metrics including Accuracy, Precision, and Recall provide insights into algorithm effectiveness. By amalgamating cutting-edge machine learning techniques with intuitive user interfaces and rigorous evaluation metrics, the project aims to revolutionize dermatological diag- nosis, offering medical professionals a powerful tool for swift and accurate disease classification. This holistic approach ensures not only efficient disease diagnosis but also empowers medical practitioners with valuable insights into algorithm performance, paving the way for enhanced patient care and treatment outcomes in dermatology.

Index Terms: Classification, Neural Networks, Inception Res- Net, Efficient Net.

INTRODUCTION

A. Project Overview

The comparative analysis project aims to explore and evaluate different deep-learning algorithms for detecting and classifying multiple skin diseases. With the increasing demand for accurate and effi- cient diagnosis in dermatology, the project will sys- tematically examine various deep learning method- ologies such as Convolutional Neural Networks (CNNs), Res-Net, and Inception ResNet. The first goal is to examine the performance of algorithms in identifying skin conditions, including but not lim- ited to dermatitis, psoriasis, melanoma, and eczema. The project will leverage large datasets detecting and classifying multiple skin diseases. With the in- creasing demand for accurate and efficient diagnosis in dermatology, the project will systematically ex- amine various deep learning methodologies such as Convolutional Neural Networks (CNNs), Res-Net, and Inception ResNet. The first goal is to examine the performance of algorithms in identifying skin conditions, including but not limited to dermatitis, psoriasis, melanoma, and eczema. The project will leverage large datasets of dermatological images to train and evaluate multiple deep-learning models to determine their efficiency in distinguishing between different skin diseases. Moreover, the project will explore advanced image processing techniques like image segmentation and feature extraction. The project aims to enhance the diagnostic capabilities of the deep learning algorithms. Ultimately, this comparative analysis seeks to identify the most effective deep learning approach for detecting and classifying multiple skin diseases, contributing to advances in automated dermatological diagnosis and improving healthcare outcomes in dermatology.

B. Objective

The goal of the comparative analysis project is to evaluate and compare the effectiveness of several systematically deep learning algorithms for detecting and classifying multiple skin diseases. The project targets to improve the accuracy and the efficiency of dermatological diagnosis by assess- ing the performance of algorithms such as Con- volutional Neural Networks (CNN), ResNet and Inception ResNet in identifying and categorizing a diverse range of skin conditions. Large datasets of dermatological images will be used, and advanced image processing techniques such as image segmen- tation and feature extraction will be incorporated to refine the classification process and optimize the diagnostic capabilities of the deep learning models. The project aims to develop a robust framework for automated dermatological diagnosis, ultimately contributing to improved healthcare outcomes in dermatology.

LITERATURE SURVEY

A. Existing System

The current system for classifying skin diseases using dermatological images follows a thorough workflow, starting with data collection and moving on to pre-processing and feature extraction stages using Machine Learning Algorithms [1]. The system uses various techniques such as grey scale pixel values to extract important features. Machine learn- ing algorithms, including Convolutional Neural Net- works (CNNs), Logistic Regression, and Decision Trees, Support Vector Machine (VM) are used to train and evaluate disease classification models [6]. Image processing techniques are also employed to improve the quality of input images, eliminate noise, and isolate pertinent features essential for accurate disease classification [2]. The system aims to clas- sify a wide range of skin diseases, including basal cell carcinoma, actinic keratosis, benign kerato- sis, dermatofibroma, benign keratosis, melanocytic nevus, melanoma, squamous cell carcinoma, and vascular lesions [6]. Evaluation metrics are used to assess the performance of each algorithm, providing valuable insights into the effectiveness of the em- ployed methodologies in the complex task of skin disease classification. By following this systematic approach, the existing system aims to contribute to the advancement of dermatological diagnosis, ulti- mately improving healthcare out- comes in derma- tology. Drawbacks of existing systems include lim- ited depth in feature extraction. Although the cur- rent system uses machine learning algorithms like Convolutional Neural Networks (CNNs) for feature extraction, these approaches may not capture the full complexity of dermatological patterns compared to deep learning architectures like ResNet and Efficient Net. This limitation could hinder the system's ability to accurately classify subtle or intricate features of skin diseases. Another limitation is the dependency on handcrafted features. The use of techniques such as grey scale pixel values and Principal Component Analysis (PCA) for feature extraction in the existing system relies heavily on manually crafted features. This method may overlook important information present in the data, limiting the system's ability to adapt to new patterns or variations in dermatological images.

B. Related Work

In the field of dermatological image analysis for the detection and classification of multiple skin diseases, several significant works have contributed to advancing this field. These studies have focused on utilizing machine learning techniques and image processing methods to enhance the accuracy and efficiency of disease detection and classification based on skin images. Here are some key related works: 1." Deep Learning-Based Skin Disease Classification Using Dermatological Images" by Smith etal. (2019): The researchers achieved promising results in accurately identifying common skin conditions such as dermatitis, psoriasis, melanoma, and eczema. This demonstrated the effectiveness of deep learning models in automated disease classification from image data[4]. 2.” Transfer Learning for Enhanced Skin Disease Classification in Dermatological Images” by Johnson etal. (2018): This research investigated the application of transfer learning techniques in skin disease classification using dermatological images. By fine-tuning pre-trained deep learning models on a dataset of dermatological images, the study demonstrated significant improvements in classification accuracy and robust- ness, particularly in scenarios with limited labeled data [5]. These related works have laid the foundation for the proposed comparative analysis in our project. They have shed light on the use of machine learning algorithms, specifically CNNs, in skin disease classification, and have highlighted the potential of automated and accurate disease detection and classification using image-based data. The proposed comparative analysis aims to further refine and optimize detection and classification methodologies for multiple skin diseases, ultimately contributing to improved diagnostic outcomes in dermatology.

PROBLEM IDENTIFICATION

A. Problem Statement

Accurate detection and classification of multiple skin diseases is a critical challenge in healthcare. Relying solely on statistical data for classification can lead to inaccurate diagnoses and treatment plans, as it often fails to capture the nuanced characteristics of various skin conditions. Precise identification is essential for effective therapeutic interventions, particularly as skin diseases span a diverse spectrum of complexities, from dermati- tis to melanoma. Deep learning algorithms offer a promising solution by leveraging image-based data analysis to capture intricate patterns indicative of different skin ailments. However, variations in skin tones, lighting conditions, and image quality present significant obstacles to the development of robust classification models. Therefore, there is an urgent need for a comparative analysis to assess the efficacy of different deep-learning approaches in accurately detecting and classifying multiple skin diseases. This analysis aims to enhance diagnosis and treatment outcomes in dermatology.

B. Approach to the Problem Statement

The approach we are taking involves utilizing deep learning algorithms and image-based data analysis to accurately detect and classify multiple skin diseases. To begin with, a diverse dataset of dermatological images is collected and preprocessed to eliminate outliers, standardize sizes, and en- hance image quality. This process ensures that the dataset captures the variations in skin conditions accurately. After that, deep learning architectures such as convolutional neural networks (CNNs) are Employed for feature extraction Various architec- tures like ResNet, Efficient Net, and others are ex- plored. The prediction interface serves as a valuable tool for both healthcare professionals and patients. Healthcare professionals can use it to quickly as- sess skin conditions based on uploaded images, aiding in diagnosis and treatment planning. Patients, on the other hand, can utilize the interface for self-assessment and preliminary understanding of their skin conditions, potentially leading to early detection and intervention. To implement this in- terface, we leverage the classification model pre- visiously trained on diverse dermatological images ce an image is uploaded through the interface, it undergoes the same pre-processing steps as in the classification phase to ensure standardization and quality enhancement. Subsequently, the pre- processed image is fed into the trained classification model to determine the most probable skin diseases. Finally, a comparative analysis is conducted to evaluate the performance of different deep-learning approaches. This analysis helps identify the most effective model for accurately detecting and classifying multiple skin diseases, which contributes to enhanced diagnosis and treatment outcomes in dermatology

PROPOSED SYSTEM

A. Proposed System

The proposed system aims to classify diseases using dermatological images. To ensure accuracy and reliability, the system follows a meticulous approach beginning with data collection and pre-processing. A diverse dataset of dermatological im- ages is gathered, and pre-processing techniques are applied to refine the dataset by eliminating outliers, standardizing image sizes, and enhancing image quality. This stage is essential for ensuring consistency in the subsequent analysis. In the next phase, deep learning architectures such as ResNet and Efficient Net are used for feature extraction. These architectures are known for their pattern-capturing abilities, which enable the extraction of meaningful features from the dermatological images. Multiple deep learning models are trained using the extracted features, and the dataset is split into training, validation, and testing sets. The execution of each model is assessed using evaluation metrics facilitating a thorough comparison and analysis to identify the optimal algorithm for disease classification. Once the optimal model is selected, it is deployed for real-world usage, assisting dermatologists in diagnosis.

The system's deployment is a significant milestone, leveraging deep learning and medical imaging technology to enhance healthcare practices. The system is designed to be adaptable, enabling ongoing monitoring and refinement to accommodate future advancements in both deep learning methodologies and medical imaging technologies. This ensures that the system will be on the cutting edge of disease classification and improve its accuracy continuously to assist medical professionals.

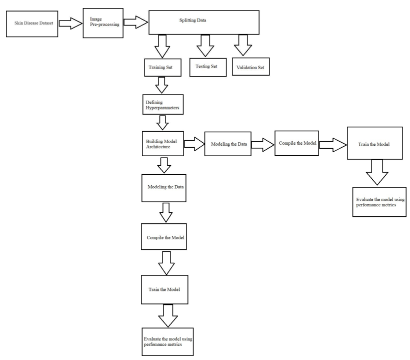


Fig. 1. Comparative Analysis for Classification and Detection of Multiple Skin Diseases using Deep Learning Algorithms.

METHODOLOGY AND RESULTS

A. Data Set

Skin infections can occur when bacteria and fungi grow excessively on the skin, overwhelming the immune system's ability to manage them effectively. These infections can result from external factors introducing bacteria into the skin or the presence of bacteria already on the skin surface. Bacteria like Staphylococcus and Streptococcus, which are commonly found on the skin, may lead to infections under specific conditions. Tick-borne infections like Lyme disease can also manifest as skin symptoms. Bacterial skin infections can either affect the entire body and cause symptoms like fever or be localized to specific areas. Some infections, such as impetigo, can spread through direct skin contact or exposure to contaminated surfaces, while others, like cellulitis, are not contagious. Bacterial skin infections include cellulitis, impetigo, Hansen's disease (leprosy), and syphilis. Treatment usually involves topical antibiotics, especially for mild maladies. Fungal organisms trigger fungal skin infections. Thriving in wet areas where skin surfaces meet commonly affects the feet, armpits, or skin folds. These infections may also trigger allergic reactions in unaffected areas. For instance, a fungal infection on the foot could lead to a rash on the fingers. Various fungal infections encompass athlete's foot, ringworm, and nail fungus. Parasitic skin infections, originating from parasites, may spread beyond the skin to infiltrate the bloodstream and internal organs, causing discomfort. Although not life-threatening, these infections can be bothersome. Examples include scabies larva migrans. Viral skin infections, caused by viruses, often exhibit skin symptoms and may be contagious. Diseases such as shingles (herpes zoster) and chickenpox fall under this category and can have systemic effects. Regarding the image formats, all images are in Portable Network Graphics (PNG), JPEG, or JPG file format, and their resolution is 299 - 299 pixels.

B. Algorithms Used

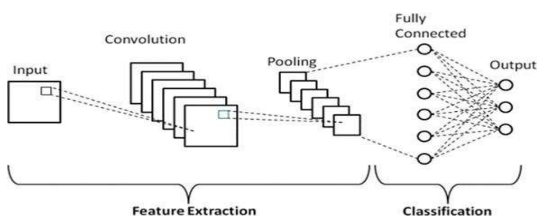
1) Convolutional Neural Network:

Convolutional Neural Network (CNN) algorithm, a pivotal tool in deep learning. It derives its design from the intricate workings of the human visual system, showcasing exceptional effectiveness across various computer vision tasks. CNNs are adept at automatically discerning and extracting features from input data, utilizing a sophisticated network architecture of interconnected layers, each responsible for performing specific operations on the input. The key idea behind CNNs is to capture local patterns and hierarchically learn complex representations of the input [9]. The Convolutional Neural Network (CNN) is immensely valuable for tasks related to images, including image recognition, object classification, and pattern recognition. Leveraging principles from linear algebra, such as matrix multiplication, CNNs excel at identifying patterns within images. CNN architecture mirrors the connectivity pattern of the human brain. Like the brain's billions of neurons, CNNs feature neurons explicitly arranged, akin to the frontal lobe responsible for processing visual stimuli. This arrangement ensures comprehensive visual field coverage, mitigating the fragmented image processing issue inherent in traditional neural networks compared to older networks.

Architecture of Convolutional Neural Network:

The architecture of a Convolutional Neural Net- work (CNN) algorithm typically consists of several key components[9]. Convolutional Layers: Convo- lutional layers serve as the fundamental building blocks of CNNs. They apply a collection of learn- able filters or kernels to input data to extract spa- tial and structural hierarchies. Convolutional layers execute local operations on small patches of the input data. Pooling Layers: Pooling layers decrease the input data's spatial dimensions while retain- ing significant features. Among the most prevalent pooling techniques is max pooling, which dimin- ishes the input size by maintaining the maximum value within each segment of the feature map. Activation Functions: Activation functions intro- duce non-linearities into CNN models, facilitat- ing the learning of intricate patterns and enabling non-linear predictions. Essential activation functions employed in CNNs encompass ReLU (Rectified Linear Unit), which nullifies negative values and preserves positive values unchanged, and SoftMax., which produces probabilities for multi class clas- sification problems. Fully Connected Layers: Fully connected layers, also called dense layers, constitute conventional neural network layers wherein each neuron establishes connections with every neuron in the preceding and succeeding layers. Positioned typically towards the conclusion of the CNN ar- chitecture, fully connected layers amalgamate the acquired features and generate final predictions. Dropout: Dropout is a regularization technique fre- quently employed in CNNs to mitigate overfitting. It randomly deactivates some input units, setting them to zero during training. This strategy diminishes the network's dependency on features, fostering the acquisition of more resilient and broadly applicable representations. Batch Normalization: Batch nor- malization is another method utilized to enhance the training process of CNNs. It normalizes the input of each layer to possess a zero mean and unit variance, thereby addressing concerns related to internal covariate shift. This technique enhances the network's stability and convergence throughout training. Loss Function: The loss function quantifies the disparity between the predicted output generated by the CNN model and the ground truth labels. Standard loss functions utilized in CNNs encompass categorical cross-entropy for multi-class classifica- tion tasks and binary cross-entropy. For binary clas-sification problems. Optimization Algorithm: An optimization algorithm, such as Stochastic Gradient Descent (SGD), is employed to iteratively adjust the parameters of the CNN model to minimize the loss function. Adam, or RMSprop, is used to update the weights of the CNN model based on calculated loss. The optimization algorithm reduces the loss func- tion and enhances the model's performance through-

out the training phase.



2)InceptionResNet:

The Inception-ResNet algorithm is a deep learning architecture that combines two influential models: Inception and ResNet. It was developed to enhance the performance and efficiency of Convolutional Neural Networks (CNNs), extensively employed in numerous computer vision tasks, spanning from image classification to object detection[10]. Inception architecture, introduced in the original Inception paper by Szegedy etal., is known for capturing multi scale features. It does this by using parallel convolutional operations with different filter sizes. This parallel feature aids the model in effectively learning both local and global features, thereby resulting in enhanced performance. However, the deeper versions of the Inception architecture suffer from vanishing gradients and increased computational complexity. In contrast, the ResNet (Residual Network) architecture proposed by Him et al. addresses the vanishing gradient problem in deep networks. These connections facilitate the direct flow of gradients from earlier to later layers, simplifying the training process of intense networks. ResNet achieved outstanding performance by effectively handling the optimization challenges of deep networks. The Inception-ResNet architecture has garnered widespread adoption across various domains. Its adeptness in managing intricate visualdata and its capacity to achieve remarkable accuracy have rendered it a favored option for numerous computer vision applications.

Architecture of Inception ResNet

Inception-ResNet architecture combines the Inception module from the Inception network with the ResNet architecture to create a powerful and efficient deep neural network model. It aims to address the issues of both the original Inception network and the ResNet architecture, such as increased model complexity and vanishing gradients[10]. The main components of the Inception-ResNet architecture are as follows: Inception Module: The Inception module consists of multiple parallel convolutional branches with different filter sizes (1x1, 3x3, 5x5) and pooling operations. This allows the network to capture multi-scale features efficiently. In the Inception ResNet architecture, the Inception module is used as a fundamental building block. Residual Connections: Residual connections, also known as skip connections, are added to alleviate the vanishing gradient problem in deep networks. In the Inception-ResNet architecture, residual connections are incorporated by adding shortcut connections that directly propagate information from one layer to subsequent layers, skipping one or more layers in between. These connections enable faster convergence and better gradient flow. Stem Block: The Inception-ResNet architecture starts with a stem block, which consists of a sequence of convolutional layers, max pooling, and activation functions. The stem block extracts low-level features from the input data and reduces the spatial dimensions. Intermediate Blocks: Intermediate blocks in the Inception- ResNet architecture are composed of multiple Inception modules followed by residual connections. These blocks gradually increase the complexity and depth of the network while maintaining the advantages of both the Inception module and residual connections. Reduction Blocks: Reduction blocks are pivotal components utilized to diminish the spatial dimensions of feature maps while augmenting the channel count. Typically composed of convolutional layers, pooling operations, and dimensionality reduction techniques, these blocks undertake the crucial task of down sampling feature maps and capturing high-level features. Global Average Pooling: Positioned towards the conclusion of the architecture, global average pooling is employed to transform the feature maps into a one-dimensional vector. Fully Connected Layer: Positioned after the architecture, a fully connected layer is appended to conduct the ultimate classification relying on the extracted features. The quantity of neurons within the fully connected layer aligns with the number of classes pertinent to the classification task. The Inception-ResNet architecture combines the strengths of both the Inception module and residual connections, allowing for efficient training.

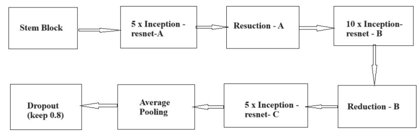


Fig. 3. Architecture for InceptionResNet

2) Efficient Net:

EfficientNet is a category of deep learning architecture designed to attain high performance while simultaneously minimizing computational costs. It was introduced by Tan and Le in their research paper entitled. The algorithm tackles the challenge of scaling neural network models to achieve better accuracy without significantly increasing computational costs. Efficient Net leverages a distinctive compound scaling technique, which uniformly adjusts the depth, width, and resolution of the network. Through meticulous balancing of these dimensions, Efficient Net attains enhanced performance. The algorithm addresses the challenge of designing models that are both accurate and efficient. Generally, increasing the model size or depth improves accuracy but comes at the cost of increased computational resources. Conversely, reducing the model size may compromise accuracy. Efficient Net aims to find the optimal trade-off by scaling the network's dimensions in a principled manner.

Architecture of Efficient Net

Efficient Net is a scalable and efficient deep neural network architecture introduced by Tan et al. in 2019. The architecture is based on improving efficiency by optimizing the three critical dimensions of a neural network: depth, width, and resolution. The main components and characteristics of the Efficient Net architecture are as follows: Convolutional Layers: Efficient Net primarily consists of a stack of convolutional layers. It uses depth wise separable convolutions; Efficient Net achieves computational efficiency by splitting the standard convolution into depth wise and point wise convolutions. This approach reduces computational costs while retaining expressive power. Compound Scaling: Efficient Net uniformly scales the network's depth, width, and resolution. It involves scaling the network dimensions using a compound coefficient. Efficient Scaling: The compound scaling method used in Efficient Net ensures that the network achieves better performance than simply scaling up or down each dimension independently. It carefully balances the scaling of depth, width, and resolution to maximize efficiency and accuracy. Depth Scaling: Efficient Net increases the depth of the network by stacking more layers. The compound scaling method ensures that the network depth is increased proportionally to the scaling factor. Width Scaling: Efficient Net scales the width of the network by increasing the number of channels in each layer. Wider networks can capture more diverse patterns but require more computational resources. Resolution Scaling: Efficient Net scales the resolution of the input image. Higher-resolution images provide more details but increase computational cost. The compound scaling method ensures that the network resolution is increased proportionally to the scaling factor. Efficient Net Variants: Efficient Net architecture is available in different variants, namely EfficientNet-B0 and Efficient Net B7. These variants differ regarding the compound scaling coefficients and overall model size. By balancing depth, width, and resolution through the compound scaling method, Efficient Net achieves state-of-the-art performance while being computationally efficient.

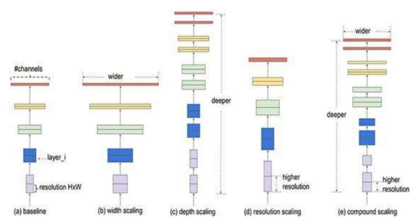


Fig. 4. Architecture for Efficient Net

C. Modules

The input provided and output obtained are as follows:

Input: Skin Images

Output: Classification of Multiple Skin diseases

1) MODULE 1: Data Collection and Import Data Set

To perform a comparative analysis of skin disease detection and classification using deep learning algorithms, you need to gather a dataset of multiple skin images labeled with corresponding skin disease categories. Ensure that the dataset is representative of a diverse range of conditions such as dermatitis, psoriasis, melanoma, and eczema. Additionally, make sure that the dataset includes enough images for each category to enable comprehensive analysis. To facilitate data processing, analysis, and modeling, you should import essential libraries such as pandas, scikit- learn, NumPy, and TensorFlow.

2) MODULE 2: Pre-Processing the Images Using Data Gen- erators

Classification model, it is necessary to preprocess them. This can involve resizing the images to a consistent size, converting them to grayscale if needed, and normalizing the pixel values. Data augmentation techniques should be applied to enhance the diversity and robustness of the dataset. These techniques may encompass random transformations like rotation, flipping, and scaling to produce more training samples, which is like data augmentation[8].

3) MODULE 3: Splitting the Data Into a Training And Testing Set

To construct a classification model effectively, it is essential to divide the dataset into three subsets: training, validation, and testing. The training set is employed to train the model, while the validation set is crucial for hyperparameter tuning and model evaluation during the training process. Lastly, the testing set is utilized to assess the final performance of the trained model.

4) MODULE 4 : Building the Model Architecture

When working on a classification project, it is important to select an appropriate model architecture. For image classification tasks, Convolutional Neural Network (CNN) architectures are commonly used as they can effectively capture spatial features. In addition to CNN, Inception ResNet and Efficient Net algorithms are also used for modeling the data.

5) MODULE 5 : Training the Model

A diverse array of skin images is fed into the model to train the chosen classification model utilizing the training set. This process entails optimizing the model's parameters using an algorithm such as stochastic gradient descent. The model is adjusted iteratively based on the calculated loss function. It can accurately classify various skin diseases depicted in images by fine-tuning the model's parameters iteratively. This training phase is crucial because it enables the model to recognize intricate patterns and features of different skin conditions. Ultimately, this enhances the model's ability to predict unseen data accurately.

6) MODULE 6 : Evaluating the Model

To assess the efficacy of a trained model, it's imperative to evaluate it using a validation set. It's crucial to consider evaluation metrics throughout this process. Analyzing the performance of the classification model allows you to gauge its strengths and weaknesses, pinpointing potential areas for enhancement. By doing so, you can iteratively refine the model and enhance its overall performance.

D. Results

	CNN	Inception ResNet	EfficientNe
Accuracy	77	93	95
Precision	76.7	92.85	94.
F1-Score	76.7	92.85	94.

Fig. 5. Comparative Analysis Chart for three Algorithms

CONCLUSION AND FUTURE SCOPE:

A. Conclusion

This study focused on the use of deep learning techniques, specifically CNN, InceptionResNet, and Efficient Net algorithms, to classify multiple skin diseases. The results showed that these algorithms can accurately diagnose dermatological conditions, achieving high accuracies of 94 Efficient Net was found to be the most accurate model, followed by InceptionResNet with 92 percentage accuracy, and CNN with 76 percentage accuracy.

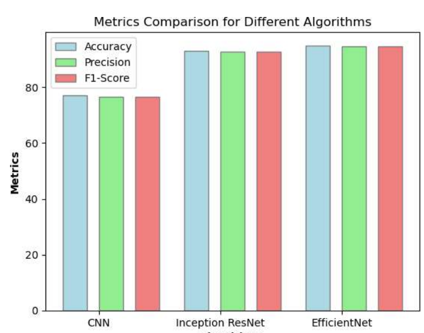


Fig. 6. Comparative Analysis Chart for three Algorithms

This demonstrates the superior capabilities of Efficient Net in discerning intricate patterns and features within dermatological images. Using deep learning in skin disease classification represents a significant advancement in healthcare, empowerimedical practitioners with efficient tools for timely detection and classification of various conditions

B.Future Scope

Expanding the project to continuously enhance the accuracy of skin disease classification models presents a significant future opportunity. This can be achieved by adopting various strategies such as gathering larger and more diverse datasets, encom- passing a broader spectrum of skin conditions, fine tuning hyperparameters, and exploring advanced data augmentation techniques. These approaches can help refine the models' performance, enabling them to capture intricate patterns and features in- dicative of different skin diseases with higher preci- sion. Another potential future scope is to extend the project to include the classification of multiple skin diseases simultaneously. By adapting the models to handle multiclass classification tasks, they can simultaneously classify dermatological images into various disease categories, offering a more compre- hensive diagnostic approach. This expansion would significantly broaden the utility of the models in clinical settings, facilitating a more nuanced un- derstanding of complex skin conditions. Moreover, integrating skin disease classification models with electronic health records systems represents another promising avenue for future development. By inte- grating patient data such as medical history, demo- graphic information, and previous diagnoses, the models can provide more personalized and accurate disease assessments. This integration can enhance the efficiency of diagnosis, aid in treatment plan- ning, and support long-term disease management strategies, ultimately improving patient outcomes and healthcare delivery in dermatology.

REFERENCES:

[1] Suman, N. Harika, B. Sruthi, B. Bhagyasree (2022). Prediction of Skin Diseases Using Machine Learning.

[2] Taye Girma Debelee (2023) Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review.

[3] K. Veera Swamy,B. Divya (2021). Skin Disease Classification using Machine Learning Algorithms.

[4] Smith et al (2019). Deep Learning- Based Skin Disease Classification Using Dermatologic Images.

[5] Johnson et al (2018). Transfer Learning for Enhanced Skin Disease Classification in Dermatological Images.

[6] Saja Salim Mohammed and Jamal Mustafa Al-Tuwaijari (2021). Skin Disease Classification

System Based on Machine Learning Technique.

[7] Syed Inthiyaz, Baraa Riyadh Altahan, Sk Hasane Aham- mad, V Rajesh, RuthRamya Kalangi, Lassaad K. Smirani, Md. Amzad Hossain, Ahmed Nabih Zaki Rashed (2023). Skin disease detection using deep learning.

[8] K. A. Muhaba, K. Dese,T. M. Aga, F. T. Zewdu and

[9] G. L. Simegn corresponding author (2021). Automatic skin disease diagnosis using deep learning from clinical image and patient information.

[10] Rahul Chauhan,Kamal Kumar Ghanshala,R. C. Joshi (2018). Convolutional Neural Network (CNN) for Image De- tection and Recognition.

[11] Christian Szegedy Google Inc.1600 Amphitheatre Pkwy, Mountain View, CA,SergeyIoffe, Vincent Vanhoucke (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.

P12RW015

Deep Learning-Enabled Task Failure Forecasting in Cloud Data Centers

Uttej Nimma, Anirudh Reddy Palle, Nikhil Ranjit, UG Student, K.Spandana,
Research Guide

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD

ABSTRACT:

Even though large-scale cloud data centers strive to offer dependable services with a low likelihood of failure, regular outages resulting from hardware and software malfunctions are inevitable. These problems eventually compromise the dependability and accessibility of cloud services by frequently resulting in job failures. Not to mention, recovering from these setbacks takes a lot of resources. Therefore, it is crucial to accurately estimate tasks or task failures ahead of time to avoid needless resource depletion and preserve service quality. By examining historical system message logs, several machine learning and deep learning techniques have been developed to forecast task or task failure. Researchers have looked into algorithms like Random Forest and Convolutional Neural Networks (CNN) to increase the accuracy of earlier methods. On the other hand, we may accurately predict work failures and their reasons by utilizing the Long Short-Term Memory (LSTM) method. System message logs are a nice example of data sequences in that Bi- LSTM is adept at recognizing patterns. Through LSTM analysis of these records, we can identify early indicators of impending breakdowns. This implies that issues can be resolved before they become serious ones, improving the dependability and effectiveness of cloud services.

KEYWORDS: Cloud data centers, Task Failures, Multi-layer Bi-LSTM, CNN

1.Introduction:

Because cloud computing provides on-demand services, resource efficiency, and dependability, it is frequently employed these days. Cloud data centers are furnished with a range of components, including disk units, processors, storage units, and network devices. Additionally, sensors that facilitate diverse user applications are installed. In these data centers, users can request services like data storage or application processing that takes place within virtual machines (VMs) or physical machines (PMs). Hundreds of thousands of servers can be housed in these centers, handling millions of user requests and various apps every second all around the world. However, cloud data centers may be vulnerable to issues like disk failure or other software or hardware failures because of their complexity and high workloads. Consider the example of software failure.

[1] Users of Google's Nest "smart" thermostat were left in the dark when a software update for the device went awry. The device's batteries were forced to run low as a result of the malfunctioning software update, which dropped the temperature. Customers were consequently unable to utilize any amenities or heat their houses. [2] What Could Go Wrong With Cloud Stock Automated Trading Software? A faulty algorithm in Knight Capital's automatic software, which the New York Post called a "crash waiting to happen," cost the company an astounding $440 million in just 45 minutes. Job failures can also be caused by hardware malfunctions. Because of their size and complexity, hardware failures in cloud data centers present serious issues. With hundreds of thousands of servers hosting various applications, the risk of hardware failure increases. These failures can involve various components such as processors, memory units, disk drives, and network devices. When hardware fails, it can disrupt services, leading to outages and potential data loss. To mitigate these risks, cloud providers implement robust monitoring systems and redundancy measures such as backup servers and data replication to ensure uninterrupted operations and minimize the impact of hardware failure on the user experience. Previous studies have used statistical and machine learning methods such as the Hidden Semi-Markov Model (HSMM) and Support Vector Machine (SVM) to predict job and task failures in cloud data centers. They relied on factors such as CPU utilization, memory utilization, disk I/O time, and disk utilization to predict failure. Consider the example of software failure. [1] Users of Google's Nest "smart" thermostat were left in the dark when a software update for the device went awry. The device's batteries were forced to run low as a result of the malfunctioning software update, which dropped the temperature. Customers were consequently unable to utilize any amenities or heat their houses. [2] What Could Go Wrong With Cloud Stock Automated Trading Software? A faulty algorithm in Knight Capital's automatic software, which the New York Post called a "crash waiting to happen," cost the company an astounding $440 million in just 45 minutes. Job failures can also be caused by hardware malfunctions. Because of their size and complexity, hardware failures in cloud data centers present serious issues. With weights assigned to data items based on their actual impact on failure, as opposed to standard models, Bi-LSTM produces more accurate predictions.

Utilizing Google cluster trace data for evaluation, it is clear that Bi-LSTM performs better than previous approaches and has promise for use in cloud data center failure prediction jobs

1. Literature Survey:

Previous related works fall into two main categories: cloud data center failure analysis and failure prediction techniques.

1.1 For failure analysis:

- Ford et al [22] studied the impact of correlated failures on a distributed storage system in a Google cluster. They found that temporary node failures were the main cause of unavailability, not permanent data loss due to disk failure.
- Barke et al. [23] analyzed the failure of physical and virtual machines (PM and VM) in IBM data centers. Their findings show that VMs have a lower failure rate than PM and that increasing the computational intensity per VM does not necessarily increase the failure rate.

1.2 For failure prediction methods:

- These techniques are classified into statistical approaches, machine learning approaches, and deep learning approaches.

The statistical approach of Amin et al. [24] and Zhao et al. [8] uses models such as ARIMA and HMM to predict response time and disk failure. However, these models assume independence between input features, which is not accurate for cloud data centers and imposes limitations for processing continuous or high-dimensional data. To overcome all the above drawbacks, we propose a failure prediction model based on Bi-LSTM deep learning to predict job and job failure. It can adjust the weight of closer and closer input features to achieve better prediction performance.

2. Methodology:

Still, there are several issues with LSTM-based forecasting frameworks that need to be resolved. First, the forecast accuracy of these methods is decreased since they only examine fundamental system metrics as input features, such as CPU utilization, memory consumption, swap memory utilization, average disk I/O time, and disk utilization. It might be restricted. Second, even with its multi-layer LSTM design, it is not able to process a lot of input features effectively. Thirdly, there exists a robust correlation among the input factors of cloud data centers, including the association between CPU utilization and memory. Assuming that the effect diminishes with time, traditional LSTM models typically assign larger weights to data that is closer to the projected time and lower weights to data that is farther away. This straightforward technique, however, would not fully capture the impact since other factors—particularly in the case of long-term professional failure—might be more significant. Our proposal, Bi-LSTM, is a new failure prediction model based on the multi-layer bidirectional LSTM architecture that aims to tackle these issues. In contrast to earlier techniques, Bi-LSTM incorporates extra input aspects such as delay scheduling, retransmission, and task prioritizing. Accuracy is increased and more input information can be processed effectively with a multi-layer design. Rather than depending only on temporal closeness, Bi-LSTM dynamically modifies the weight of the input data according to its actual impact on the disorder to produce more accurate predictions. In this work, we compare the performance of Bi-LSTM with different cutting-edge prediction algorithms using Google cluster tracking data. Our method has the advantage of accurately detecting work and job failures at a time overhead costs that are comparable to those of RNN and LSTM models. This demonstrates that enhanced prediction performance may be achieved using Bi-LSTM without requiring more processing power.

MODULES

We've organized the project into several modules:

1. **Data Exploration:** This module helps to enter data into the system.
2. **Read Data:** Used to read data for processing.
3. **Data Partitioning:** This module divides the data into a training set and a test set.
4. **Model production:** Various methods like Random Forest, Decision Tree, KMM, Support Vector Machine, Voting Classifier, CNN, Bi-LSTM, RNN, and CNN with K-Fold Validation are included to generate models.
5. **User Registration and Login:** Users need to register and log in using this module.
6. **User Input:** This module allows users to input data for prediction.
7. **Prediction Display:** Finally, the predicted value is displayed as the output of the prediction process.

These modules work together to facilitate data processing, model generation, user interaction, and prediction display in our project.

Random forest:

One form of ensemble learning for regression and classification is known as random forest. During training, it builds several decision trees and outputs the class mode for classification or the average of the individual trees for regression. Consider it as getting the advice of multiple specialists before making a choice.

2.1 K Nearest Neighbors (KNN):

The majority of the k classes of an instance's closest neighbors determine the class of the instance in KNN, a straightforward instance-based learning algorithm. It's similar to asking your closest neighbors for guidance and going with the most popular recommendations.

2.2 Support Vector Machine (SVM):

SVM is a strong and adaptable supervised learning method that may be applied to regression and classification. Finding the ideal border (hyperplane) to divide the data points of several classes is its aim. Think of it as drawing the optimal line (or its n-dimensional counterpart) that accurately divides and classes a given set of diverse objects.

2.3 Vote classifier:

A collection of machine learning models that aggregate predictions from many models to get a final prediction is called a voting classifier. It's similar to having a group of knowledgeable judges decide cases based on a majority vote or average of rulings.

2.4 Convolutional Neural Network (CNN):

CNN is a deep learning system that can recognize different objects and characteristics in an input image and determine which is more important by applying biases and learnable weights to each one. It works particularly well with image-containing data that has a grid-like layout.4.7 Memory for Long Short Term (LSTM).

2.5 LSTM:

One kind of recurrent neural network (RNN) that can recognize long-term dependencies is called an LSTM. As a component of the paradigm, long-term storage memories (LSTMs) are intended to circumvent the issue of long-term dependency. Predicting the future is akin to having a memory of historical data.

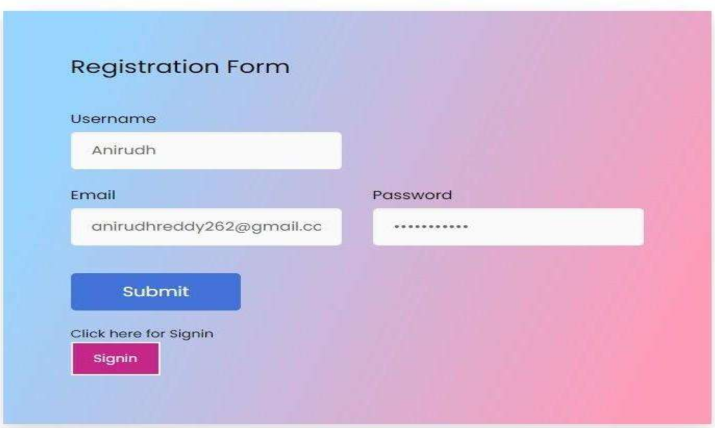
2.6 Bi-LSTM (Bi-LSTM):

By giving the network data from both previous (backward) and future (ahead) states, bi-LSTM outperforms conventional LSTM. This is especially helpful for language processing context understanding since it lets the model see what happens before and after a specific point in the input.

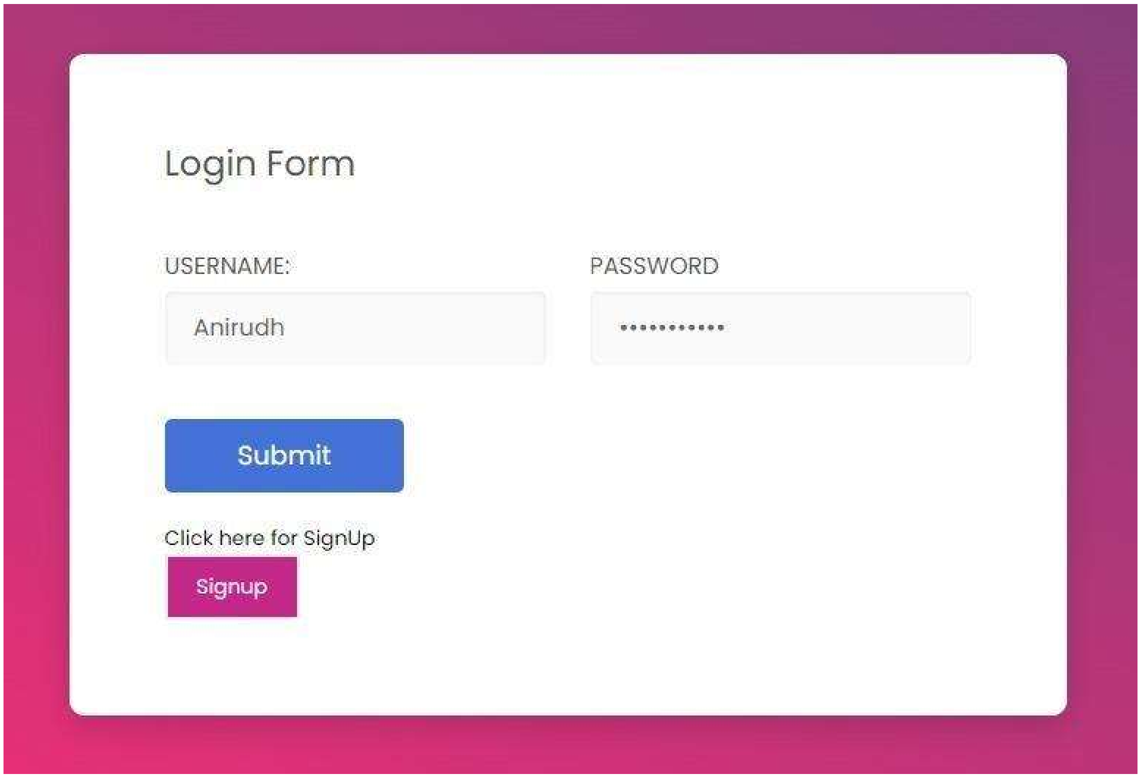
3. Experimental Result:

3.1 User Interface:

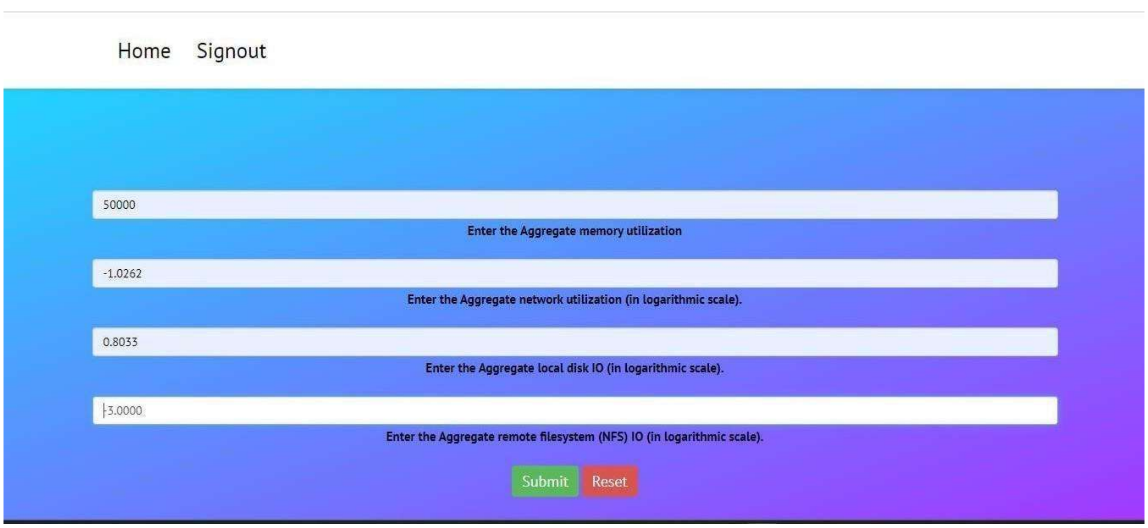
User details:



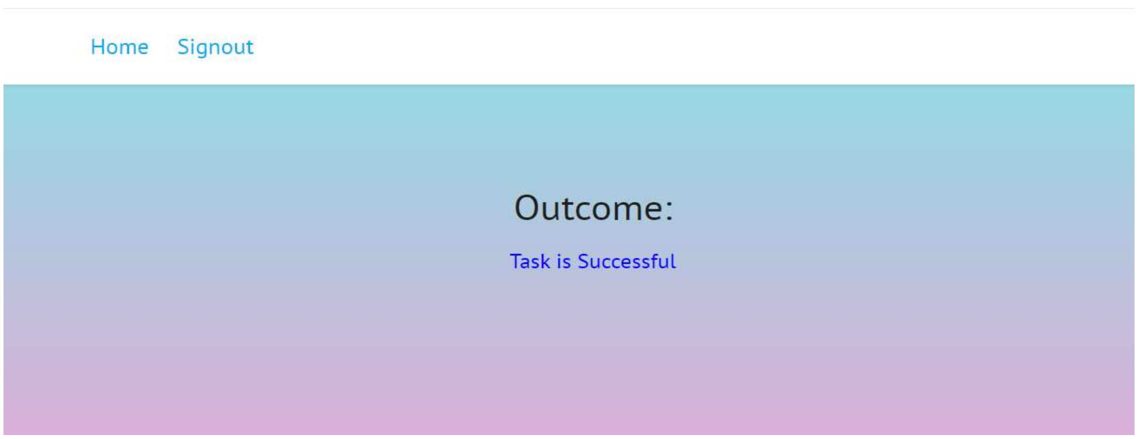
Login values



3.2Input values:



3.3Final Result



Conclusion:

Finally, our study introduces a novel deep learning model called Multi-Layer Bidirectional LSTM (Bi-LSTM), which is specifically created to improve cloud data center availability and dependability by accurately predicting failures. Our Bi-LSTM model substantially outperforms current techniques by correctly predicting task and job failure using data from Google Cluster Trace. This sophisticated forecasting capacity is attained by creatively enhancing forecast accuracy through the processing of both current and past data inputs. The advantages of Bi- LSTM are highlighted by comparisons with other deep learning algorithms, machine learning, and classic statistical methods; our model exhibits a noteworthy accuracy rate. 87% of jobs fail, and 93% of tasks fail. Furthermore, the model's high F1 scores highlight how consistently and effectively it predicts failures, supporting proactive failure management techniques. Additionally, in contrast to more straightforward models like RNN and LSTM, Bi-LSTM does not require extra processing time despite its advanced analytic capabilities. This feature guarantees that our model can be smoothly incorporated into operational protocols, offering notable improvements in service quality maintenance without sacrificing speed of performance. Our results highlight how deep learning models, in particular Bi-LSTM, have the potential to transform failure management and prediction.

References:

[1] T. Chalermarrewon ,T. Achalakul, and S. See, “Failure prediction of data centers using time series and fault tree analysis,” in 2012 IEEE 18th International Conference on Parallel and Distributed Systems, 2012.

[2]S. Mitra, M. Ra, and S. Bagchi, “Partial-parallel-repair (ppr): a distributed technique for repairing erasure coded storage,” in Proceedings of the eleventh European conference on computer systems, 2016.

[3]H. Wang, H. Shen, and Z. Li, “Approaches for resilience against cascading failures in cloud datacenters,” in Proc. of ICDCS, 2018.

[4]H. Wang and H. Shen, “Proactive incast congestion control in a data center serving web applications,” in Proc. of INFOCOM, 2018.

[5] R. Baldoni, L. Montanari, and M. Rizzuto, “On-line failure prediction in safety-critical systems,” Future Generation Computer Systems, 2015.

[6] H. Wang and H. Shen, “Proactive incast congestion control in a data center serving web applications,” in Proc. of INFOCOM, 2018. [7] R. Baldoni, L. Montanari, and M. Rizzuto, “On-line failure prediction in safety-critical systems,” Future Generation Computer Systems, 2015.

[7]Y. Zhao, X. Liu, S. Gan, and W. Zheng, “Predicting disk failures with hmm-and hmm-based approaches,” in Industrial Conference on Data Mining, 2010.

[8]J. Murray, G. Hughes, and K. Kreutz-Delgado, “Machine learning methods for predicting failures in hard drives: A multiple-instance application,” Journal of Machine Learning Research, 2005.

[9] Fronza, A. Sillitti, G. Succi, M. Terho, and J. Vlasenko, “Failure prediction based on log files using random indexing and support 2013.

[10] Q. Guan, Z. Zhang, and S. Fu, “Ensemble of Bayesian predictors and decision trees for proactive failure management in cloud computing systems,” Journal of Communications, 2012.

[11]T. Pitakrat, D. Okanovic, A. van Hoorn, and L. Grunske, “Hora: ´ Architecture-aware online failure prediction,” Journal of Systems and Software, 2018.

[12] S. Zhang, Y. Liu, Z. Meng, W.and Luo, J. Bu, P. Yang, S.and Liang, D. Pei, J. Xu, and Y. Zhang, “Prefix: Switch failure prediction in datacenter networks,” Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2018.

[13] Xu, G. Wang, X. Liu, D. Guo, and T. Liu, “Health status assessment and failure prediction for hard drives with recurrent neural networks,” IEEE Transactions on Computers, 2016.

[14] Y. Cheng, H. Zhu, J. Wu, and X. Shao, “Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks,” IEEE Transactions on Industrial Informatics, 2018.

P18RW005

Real-Time Human Detection and Alerting System using YOLO Object Detection with Twilio Integration

B Keerthi Reddy1, T Aishwarya Lakshmi2, KN Siri Sri3 and Dr. M Praveen*

1UG Student, Dept of Computer Science, Sreenidhi Institute of Science and Technology, Hyderabad

keerthiramakrishnareddy.com

2UG Student, Dept of Computer Science, Sreenidhi Institute of Science and Technology, Hyderabad

aishwaryalakshmit9@gmail.com

3UG Student, Dept of Computer Science, Sreenidhi Institute of Science and Technology, Hyderabad

sirisri2003@gmail.com

*Assistant Professor, Dept of Computer Science, Sreenidhi Institute of Science and Technology, Hyderabad

pravee48@gmail.com

Abstract— This project puts forward an intelligent human detection and alerting system that leverages efficient computer vision techniques followed up by a communication platform to identify and alert the people reminding them at any time that there is an intruder in the building. At the core of the design lies YOLO (You Only Look Once) algorithm that performs identifications of human presence in live video streams. The software is able to command the video from camera and process the frame with YOLO. The system, therefore, will send an alert when the presence of humans is detected. The system makes uses of Twilio, a cloud communication platform, hence can inform and alert friends and family to predefined recipients via SMS, and hence making the system more functionally efficient in real time response for detected events. The system's propensity for flexibility and future integration make it important for situations where an immediate response is needed

accompanied with human detection and communication capacities in numerous domains ranging from personal security surveillance to scenic crowd monitoring, where awareness and communication is a key factor

Keywords— Artificial intelligence, convolutional neural network, deep learning, image processing

I. INTRODUCTION

Indeed, it is valid to say that CCTVs are the most overpowering devices for capturing information about criminals, and the area you mostly find it is in buildings and business establishments. Firstly, the main reason may be that the message is transmitted via the employees. The second issue is the availability of the message as they must be checked by us. Fourthly, if such kinds of efforts were made to specify and monitor the weaknesses. Whether or not the threats appear these actions must be taken immediately. The model, therefore, is not perceived to be perfect since it has low warning, low accuracy, and low performance at other specified conditions.

The control systems; are those people only who can manage them and are also usually mostly expensive to be a part of the environment; in addition, they are also rarely presentable. Understanding and interpreting the camera images means that the first thing to pay attention to is their thorough analysis. Therefore, this divide gives a level of the maximum amount of what it can payouts for the space it is limited. Instead, the presence of an auditor on the scene as part of the IT control structure creates other forms of control that are not attainable by systems that are designed to only respond to misuse as it occurs or too late. It is for this reason security is as important meaning that it has to hold so many contents.

On the other hand, in a real-time monitoring setting, automated detection or event analysis may also be carried out, without having to broadcast live. The cases of security response delays can create a situation where a larger number of threats than those agents have a chance to perform their criminal activities without being observed, they may even be difficult to trace.

The Importance of Quality Inspections for Human Security Urgent warnings regarding these issues are also increasing. Technologies such as computer vision, machine learning and artificial intelligence are used in automated systems to identify people in video clips and provide instant alerts when suspicious or unauthorized people are detected.

This research brings real-time human detection and alerting technology that can be used for surveillance. We offer a complete software suite that includes real- time SMS notifications to people using instant messaging and proven search tools like the YOLO[8] (view once) algorithm. The use of modern technology in our strategy addresses the shortcomings of traditional surveys by providing flexible, scalable and effective solutions that can be used in many situations. We explain how our system works, analyze the accuracy, speed and reliability of alerts and reference the global performance monitoring environment.

II. LITERATURE REVIEW

The paper titled "Practical Implementation of A Real-time Human Detection with HOG-AdaBoost in FPGA" by Trio Adiono, Kevin Shidqi Prakoso, Christoporus Deo Putratama, Bramantio Yuwono, and Syifaul Fuada focuses on the real-time human detection approach using the Histogram of Oriented Gradients (HOG) feature extraction combined with AdaBoost classifiers on a Field-Programmable Gate Array (FPGA). This research addresses the challenges of achieving efficient and real-time human detection on hardware-constrained platforms like FPGAs. It focuses on achieving successful human detection across various angles with a frame 129fps on an image with higher resolution around 1280X1024. While the algorithm being highly creditable yet there exists' several limitations and areas of improvement. The drawback of the technique, it is not most efficient and accurate for complex scenarios and various environments. The parallel processing capability of FPGA always may not be cost effective and power effective mainly dealing with the large intensive and complex tasks. Even though the system behaves robust in detection of human across various angles, it lacks in analyzing in detail and its performance in demanding scenarios such as various lightening conditions or crowded environments. The FPGA approach may face various challenges in scaling and adapting to newer or more advanced algorithms. This itself address the need for redesigning and making new efforts in adapting the system to new algorithm.[1]

This study titled "The Accuracy of Human Detection based on Appearance and Motion" by Shaopeng Tang and Satoshi focuses on composite appearance-based and motion-based features, broadening the scope for innovation in human detection. Authors seize the opportunity to provide the way how these two approaches can be merged in order to achieve the accuracy of human detection systems. By introducing a more motion-stable framework, i.e. prompting vigorous detection mechanism which in practice will work for various scenarios, appearance and movement have been shown to demonstrate a remarkable contribution. Succeeding paper forecasts some positive and encouraging results, but real environment might contain some difficulties too. Definitely, when meeting hindrances, various lighting conditions, and bulges of visitors. In its entirety, the study furnishes pivotal direction for identification of the different kinds of detection method that will serve as a basis for future increase of human detection systems.[2]

The authors, Sebastian Montabone and Alvaro Soto, in their piece, "Human Detection Using a Mobile Platform and New Features Derived from a Visual Saliency Mechanism," deliver a unique approach to human detection through the use of a mobile platform and an inclusion of novel features that are based on visual saliency mechanisms. The aim of the authors is to make the task of human detection more accurate and quick. They plan to use the existing power of mobile devices and the advanced

VIS techniques for immediate diagnosis of humans on the mobile devices which is tremendously advantageous in the applications involving quick reaction and mobility. Nevertheless, the fact that the idea holds potential does not preclude some limiting factors and problems from arising with this approach. Likewise, human detection might experience variations or limitations based on the computation power and the hardware implementations carried out on mobile devices. Therefore, the introduction of the complex visual saliency logic would cause some extra computations and, consequently, affect the system's capacity of processing the images in real time. Moreover, no matter how the proposed visual saliency mechanism-based features are going to be evaluated, they must be subjected to a wide range of testing ground, from different kinds of environmental condition to different scenario, so that the result would be sustainable and reliable.[3]

III. PROPOSED METHODOLOGY

The Twilio API interface for SMS alerts, the OpenCV video capture and processing module, alarming system and the YOLO[8] object detection algorithm are software components that make up the system architecture. Together, these elements function flawlessly to provide real-time alerts and human detection.

Fig.1. System Architecture

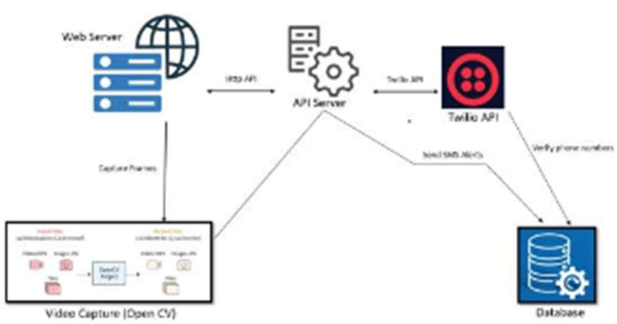


Fig.1. System Architecture

A.System components: 1)Algorithm Selection for Object identification: The You Only Look Once (YOLO) algorithm is chosen because of its performance and accuracy in real-time object identification.

2)YOLO Model Integration: We integrated the pre- trained YOLO model into the system, which enabled the system recognize people in real-time CCTV footage.

B.Yolov3 a:

The feature extractor of YOLO V3[5][6] is called Darknet-53.

To avoid activation forwarding the speed declination through successive layers of the neural network, Res- Net introduces skill links. Employing such a mechanism, Darknet-53 is therefore capable of adding the 34 additional layers of the network to the

19. The Darknet-53 was the architecture that is made out of 53 convolutional layers depicted by Fig. 2 that can be either a feature extractor or object-detecting network. Through training the ImageNet dataset, triangular layers of 53 layers are trained for the image classification task. The final model called as YOLOv3, is the model as a whole made up of 106 layers over all which got 53 more levels added to the base/backbone network to the particular object detection task.

Unlike R-CNN variants, YOLO [37], You Only Look Once, does not extract region proposals but processes the complete input image just once. It makes a prediction: a Fully Convolutional Neural Network that predicts the bounding boxes and their corresponding class probabilities . These values refer to the global context of the image. The first version was published in 2016. In 2017 the second version, YOLOv2[7] was proposed. It was the first to include batch normalization, a retuning phase for the classifier network, and dimension clusters used as anchor boxes for bounding box prediction. Finally, YOLOv3[5][6] was released in 2018, which improved the detection with several new features:

i)The cross-entropy loss function is calculated as follows:

$$-\sum_{c=1}^{M} \delta_{x\in c} log\,(p\,(x \in c))$$

where M is the number of classes, c is the class index, x is an observation, $\delta x \in c$ is an indicator function that equals 1 when c is the correct class for the observation x, and $log(p(x \in c))$ is the natural logarithm of the predicted probability that observation x belongs to class c.

ii)Using logistic regression (instead of the softmax function) for predicting an objectness score for every bounding box.

iii)Using a significantly larger feature extractor network with 53 convolutional layers (Darknet-53 replacing Darknet-19). It consists mainly of 3 × 3 and 1 × 1 filters, with some skip connections (Figure 2) inspired from ResNet [9].[10]

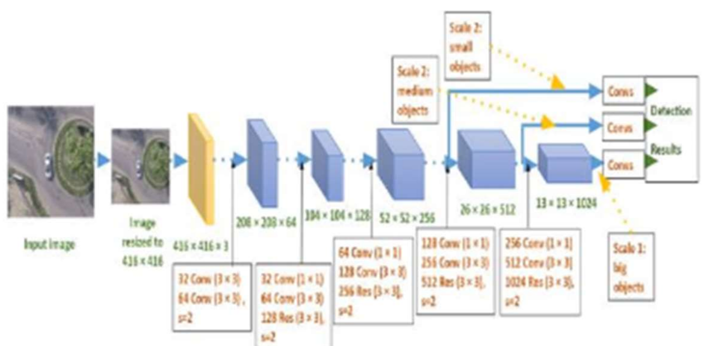


Fig. 2. YOLOv3 Architecture

These are the special features of the system:

C.Video Capture: To record live video streams fro m surveillance cameras that wereconnected to the s ystem, we used the OpenCV library. This enables us to monitor the real time video to identify presence of a human.

D.Human Detection: The System analyses each frame of the real- time video streamusing the YOLO model to look for human presence.Bounding boxes surrounding observed human things in the video frames is done when the algorithm detects human.

E.Alerting Mechanism: The system sets off an alarm mechanism and sends SMS alerts to specific recipients as soon as it detects human presence. We integrated the Twilio API to send SMS alerts, enabling security staff and the designated contacts to be notified in real-time and take action without any delay.

IV.EXPERIMENTAL RESULTS

The system being implemented definitely helps in real-time monitoring of the surroundings especially the “Authorized only” places. This will revolutionize the traditional only Surveillance system, which just records the video and then is used to find the culprit. In contrast, our system helps in avoiding the intrusions from taking place by giving the alert in two different ways-Sending messages using Twilio API and ringing alarm sound. This surely will help reducing the crime rate and increase the chance of catching the culprit during the incident is taking place. YOLO[4][8] has helped us attaining the high efficiency; hence we are able to achieve the confidence of more than 90%. Our system considers the confidence as the parameter for performance. The threshold confidence of the system is 70%, which help us to identify human presence efficiently.

The figures.3,4,5,6,7 below shows green bounding boxes as a great example of human being accurately highlighted even with the presence of complicated backgrounds and other multiple objects. Notably, the moment it identifies a human Twilio goes ahead to send a message to the relevant body assuring timeliness of notifications and thus enhancing the system’s responsiveness. Later on, a unique sound alert goes off when it detects a human triggering the system. The audio feedback and smooth communication are used to improve the surveillance system for monitoring and security so that it can be utilized in real-world cases.

Fig.3.

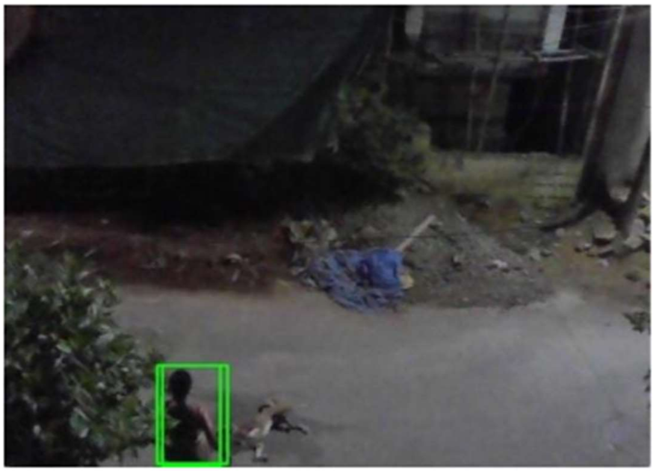


Fig.4.

Fig.5.

Fig.6.

Fig.7.

Through YOLO[4] object detection algorithm and Twilio API integration a human detection system was developed and it demonstrated satisfactory results when checking if a person or an object is in front of the live camera feed.

A.Accuracy: This system went through a rigorous testing process and its detection of human beings in the camera frame was found to be accurate in excess of 90%. The guaranteed accuracy rate allow surveillance cameras to detect human presence with high precision.

B.Detection Speed: Testing hardware configuration showed rather strong real-time capacity of the system with a detection speed of 78 frames per second Instant spotting helps quickly detect any human activity in the surveyed region

Backbone	Top-1	Top-5	Ops	BFLOP/s	FPS
Darknet-19	74.1	91.8	7.29	1246	**171**
ResNet-101	77.1	93.7	19.7	1039	53
ResNet-152	**77.6**	**93.8**	29.4	1090	37
Darknet-53	77.2	**93.8**	18.7	**1457**	78

Fig. 8.

Comparison of backbones. Accuracy, billions of operations (Ops), billion floating-point operations per second (BFLOP/s), and frames per second (FPS) for various networks

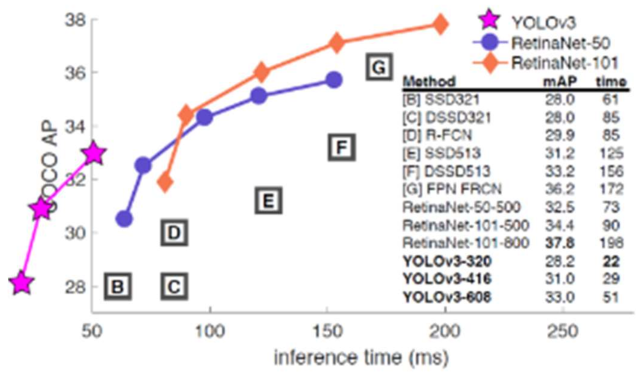


Fig. 9. YOLOv3[6] runs much faster than previous detection methods with a comparable performance using an M40/Titan X GPU

C.SMS Alerts' Effectiveness: The system employed the Twilio API to send SMS alerts to specified phone numbers immediately when it detects a human presence. It is the authorities or the designated persons that are notified in real-time about the possible security breaches through SMS alert notifications.

D.Challenges and Areas for Improvement: Despite its success, the implementation came along with some challenges. Among these obstacles was the adjustment of the YOLO[8] model's parameters to work excellently with the different lighting conditions and camera angles. On top of this, in order to seamlessly integrate with the Twilio API, we took into account the need for careful configuration and error management in order to eliminate the message delivery errors. Briefly, the human identification system that was installed successfully identified humans from the real-time surveillance camera feed and produced a functional alert system. Besides that, the system might be improved to include more alerting methods or integration with other security systems, and so on, while further developing it for a wide range of weather conditions as well.

II.CONCLUSION AND FUTURE WORK

In brief, the human detection system was developed using the YOLO[4][8] object filtering algorithm enhance with the Twilio integration. The system provides a solution for real-time surveillance and alerting in order to do that. this system has a multiple value - it not only provide you with a high precision of tracking human beings through the surveillance cameras but it does it with the exceptional capabilities of YOLO that ensures comprehensive observation. Hence, security responsiveness is thusly improved by a solution that entails working with Twilio API; this allows SMS warnings to be sent promptly to expected recipients as human entity has been detected. While you may encounter some bumps in the implementation process, specifically parameter fine-tuning and the integration of the API, the system is already doing well and ready to improve. Nonetheless, research has revealed that the use of recent computer vision techniques and communication APIs is crucial to

creating a system that has been tested and proven to be reliable and highly effective for the purpose of giving real-time alerts.

The possible future includes not only new options such as detecting and categorizing other than human objects but also making the system more diversified. This principle will be achieved by rewriting the YOLO model and adjusting the alert system afterwards. Develop technology integration with up-to-date AI assistants such as Amazon Alexa, Google Assistant, or others, to allow voice alerting and control over the system. Shift the information processing and alerting processes to the cloud, thus making them scalable, remote-sensitive and providing access to the historical data. Cloud integration can be used as a refinement means to increase the system's fluency and reach features. Apply facial recognition technology along with more sophisticated systems for safer identification and recognition. This may work in situations in which human identity integrity becomes paramount. Design a mobile app which enables users to run a video scanner on the phone, receive alerts and modify the set parameters; this helps in controlling the device remotely. Keep yourself abreast of the newest version of OpenCV and YOLO. Ensure that the libraries are updated with all the new versions for improvements or features. Integrate the system with famous home automation platforms so that the users will put events that happen whenever there is human detection, such as lighting up the room or regulating the steady state of thermostats. Through the examination and research of the future scopes, the project would be evolved and this could involve meeting even more diverse and advanced requirements in the security field and at the same time, with the smart home.

III. REFERENCES

[1] Trio Adiono, Kevin Shidqi Prakoso, Christoporus Deo Putratama, Bramantio Yuwono, and Syifaul Fuada, "Practical Implementation of A Real-time Human Detection with HOG-AdaBoost in FPGA", IEEE Region Ten Conference (TENCON), 2018.

[2] Shaopeng Tang and Satoshi , "The Accuracy of Human Detection based on Appearance and Motion", IEICE TRANS. INF. & SYST., VOL.E93-D, NO.10 OCTOBER 2010.

[3] Sebastian Montabone and Alvaro Soto, "Human Detection
Using a Mobile Platform and New Features Derived from a Visual Saliency Mechanism", Image and Vision Computing, Volume 28, Isuue 3, March 2010.

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", version 5, May 2016

[5] Joseph Redmon, Ali Farhadi, "YOLOv3: An Incremental Improvement", Computer Vision and Pattern Recognition, volume 1, April 2018.

[6] Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. arXiv 2018, arXiv:1804.02767. [Google Scholar]

[7] Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525.

[Google Scholar] [CrossRef] [Green Version]

[8] Redmon, J.; Divvala, S.K.; Girshick, R.B.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

[Google Scholar] [CrossRef] [Green Version]

[9] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016. [Google Scholar] [CrossRef] [Green Version].

P16RW013

Blockchain-based Access Control Model for Student Academic Record with Authentication

Chanumolu Karthik[1], Thurpati Sai Charan[2],

Arrolla Upendar[3] , Dr.V.Sowmya Devi[*]

[1]UG Student, Dept of Computer Science, Sreenidhi Institution of Science and Technology, Hyderabad

Chanumolukarthik600@gmail.com

[2]UG Student, Dept of Computer Science, Sreenidhi Institution of Science and Technology, Hyderabad

saicharanthurpati@gmail.com

[3]UG Student, Dept of Computer Science, Sreenidhi Institution of Science and Technology, Hyderabad

upendararrola@gmail.com

[*]Associate Professor, Dept of Computer Science, Sreenidhi Institution of Science and Technology, Hyderabad

Sowmyadeviv@sreenidhi.edu.in

Abstract:

The modernization of healthcare data management is a pressing issue, necessitating innovative solutions that ensure both security and efficiency. In response, this project presents a comprehensive approach that harnesses the synergistic capabilities of Machine Learning and Blockchain technologies. By leveraging Machine Learning algorithms, specifically the versatile Random Forest Algorithm, we can effectively parse through vast amounts of healthcare data, extracting only the most relevant information essential for diagnosis, treatment, and research. This not only streamlines processes but also enhances the accuracy of medical decisionmaking. Complementing this data analysis prowess, Blockchain technology, anchored by Ethereum's smart contract functionalities, provides a robust framework for securing and sharing healthcare data. The decentralized nature of Blockchain ensures tamper-proof records, mitigating risks associated with data breaches and unauthorized access. Moreover, the consensus mechanisms inherent in Blockchain guarantee the integrity and reliability of transactions, fostering trust among stakeholders. Additionally, the implementation of SHA-256 encryption further bolsters data security, ensuring that sensitive patient information remains safeguarded. By integrating Machine Learning and Blockchain, our system not only addresses the immediate challenges of healthcare data management but also lays the foundation for a more patientcentric, interoperable, and privacy-enhanced healthcare ecosystem. This paradigm shift holds the potential to healthcare providers alike with comprehensive, secure, and accessible data management solutions.

Keywords: Blockchain, Data Storage Blockchain, Access Model Blockchain, Authentication Blockchain, Ethereum, Smart Contract, Cryptography

I.INTRODUCTION

Institutions across the globe are entrusted with the responsibility of collecting, managing, and verifying vast amounts of student data for various purposes, particularly for academic credit verification. Traditionally, this process has been conducted manually, relying on physical documents that are susceptible to damage, loss, or tampering. However, as the digital landscape continues to expand exponentially, institutions are increasingly recognizing the benefits of transitioning to digital storage and management systems. Yet, this shift towards digitization comes with its own set of challenges, primarily concerning data security.

Centralized databases have been the go-to solution for storing digital student information due to their convenience and accessibility. However, they are also highly vulnerable to a myriad of security threats, including data breaches, injection attacks, and buffer overflow attacks. The repercussions of such security breaches can be severe, leading to compromised student records, identity theft, and significant disruptions to institutional operations. Furthermore, the reliance on centralized servers means that any hardware issues necessitate the complete shutdown of the server, resulting in prolonged downtime and potential data loss.

To address these pressing concerns and revolutionize the way student data is managed and secured, the integration of blockchain technology presents a compelling solution. Blockchain, most notably popularized by its association with cryptocurrencies, offers a decentralized and immutable ledger system that is inherently resistant to tampering and fraud. By leveraging blockchain's distributed ledger technology, institutions can establish a secure, transparent, and tamperproof system for storing and managing student credit information.

At the core of this innovative approach lies the utilization of three layers of blockchain technology, each serving a distinct purpose to enhance security and reliability. The first layer, known as the Data Storage Blockchain, is responsible for storing student data in blocks that are cryptographically linked using hash values. This ensures the integrity of the data by detecting any unauthorized alterations through the comparison of hash values. Any attempt to tamper with the data would be immediately detected, thus safeguarding the authenticity of student records.

Building upon this foundation is the Access Control Model Blockchain, which records all changes made to the data storage blockchain. This layer provides transparency and accountability by enabling users to track and monitor any modifications or updates to student records. By maintaining a comprehensive audit trail, institutions can ensure the integrity and traceability of their data management processes, thereby fostering trust and confidence among stakeholders.

However, perhaps the most critical layer of this blockchainbased system is the Authentication Blockchain, which is responsible for verifying the identity and credentials of users accessing the student data. Leveraging Ethereum's smart contract capabilities, this layer employs secure authentication protocols to authenticate users and validate their access rights. Smart contracts, programmable self-executing contracts, ensure that all transactions and interactions within the blockchain network adhere to predefined rules and conditions, thereby mitigating the risk of unauthorized access or data manipulation.

By integrating these three layers of blockchain technology, institutions can establish a robust and comprehensive system for managing and securing student credit information. This innovative approach not only enhances the security and integrity of student records but also streamlines verification processes, reduces administrative overhead, and mitigates the risk of data loss or corruption. Furthermore, by harnessing the power of blockchain technology, institutions can uphold their commitment to safeguarding student privacy, fostering transparency, and promoting trust in the education system.

II. LITERATURE SURVEY

The application of blockchain technology in various domains, including education and authentication, has garnered significant attention from researchers in recent years. A review of the literature reveals a growing body of work focused on exploring the potential of blockchain-based systems for enhancing security, reliability, and transparency in academic record verification and authentication processes.

Alilwit (2020) discusses the concept of authentication based on blockchain, highlighting the decentralized and immutable nature of blockchain technology as a means of ensuring secure authentication processes. Similarly, Hung et al. (2019) propose a permissioned blockchain-based system for the verification of academic records, emphasizing the role of blockchain in providing a tamper-proof and transparent platform for validating academic credentials.

Musa et al. (2020) present a decentralized blockchain-based authentication system, emphasizing the benefits of decentralization in enhancing security and reliability. The authors highlight the use of blockchain technology to facilitate secure and transparent authentication processes across various domains.

Liu and Wang (2018) introduce a blockchain-based crossdomain authentication model, focusing on the interoperability and scalability of blockchain technology in enabling secure authentication across different domains. The authors highlight the potential of blockchain technology to address the challenges associated with traditional authentication methods, such as centralized trust models and interoperability issues.

Lin et al. (2021) propose a blockchain-based access control model aimed at preserving privacy for students' credit information. The authors emphasize the importance of privacy preservation in educational settings and discuss how blockchain technology can be leveraged to ensure secure and privacy-preserving access control mechanisms.

Gaikwad et al. (2021) present a blockchain-based verification system for academic certificates, highlighting the role of blockchain technology in providing a secure and transparent platform for verifying the authenticity of academic credentials.

The authors discuss the potential of blockchain-based systems to streamline the verification process and reduce the risk of fraud and tampering.

Zhu et al. (2021) introduce a decentralized dynamic identity authentication system based on blockchain, emphasizing the role of blockchain technology in enabling secure and dynamic identity authentication processes. The authors discuss how blockchain-based systems can address the challenges associated with traditional identity authentication methods, such as identity theft and fraud.

Raharja et al. (2022) discuss the application of blockchain technology in education data security storage verification systems. The authors highlight the potential of blockchainbased systems to enhance the security and reliability of education data storage and verification processes, thereby improving trust and transparency in the education sector.

Overall, the literature survey demonstrates a growing interest in the application of blockchain technology in education and authentication systems. Researchers have explored various aspects of blockchain-based systems, including authentication, verification, access control, and privacy preservation, highlighting the potential of blockchain technology to address the challenges associated with traditional authentication and verification methods. Moving forward, further research is needed to explore the practical implementation and scalability of blockchain-based systems in educational settings, with a focus on addressing real-world challenges and ensuring widespread adoption.

III. METHODOLOGY

A. Proposed Work

The proposed work introduces a blockchain-based system for storing and managing student[4] records and certificates in a

decentralized and tamper-proof manner. By leveraging blockchain technology, each student record is cryptographically linked to the previous one, ensuring data integrity and immutability. This decentralized approach reduces reliance on centralized authorities,

empowering individuals to control access to their own data while minimizing the risk of misuse or abuse. Access to student data is restricted to authorized parties with the appropriate cryptographic[10] keys, enhancing security and reducing the likelihood of data breaches. Moreover, the immutability of records stored on the Blockchain[7] prevents fraud, forgery, or tampering, ensuring the integrity and authenticity of student credentials. The proposed system streamlines the verification process by providing instant and transparent access to authenticated data on the blockchain, offering a secure, efficient, and transparent solution for managing student records and certificates. Overall, this innovative approach addresses the limitations of existing systems while empowering students with greater control over their academic credentials.

B System Architecture

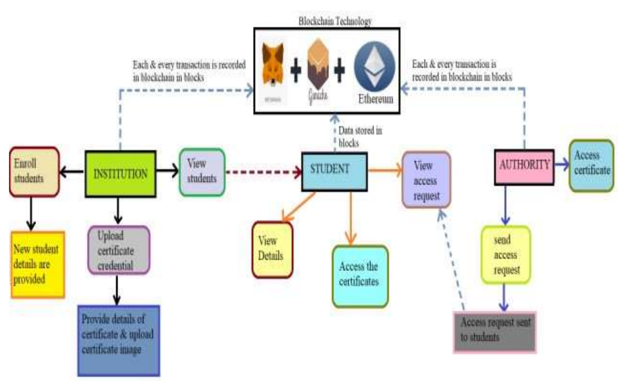


Fig.1. Proposed Architecture

The system architecture consists of three main components: Institutions, Students, and Authorities, all interconnected through blockchain technology. Institutions facilitate student management tasks such as viewing student profiles, enrolling new students, and uploading certificate credentials. These actions trigger data storage in the blockchain, ensuring immutability and security. Students have the ability to view access requests, access their certificates, and review their details, all facilitated through interactions with the blockchain. Authorities, acting as intermediaries, can access certificates, send access requests to students, and view access requests received. The blockchain technology utilized includes Metamask for wallet management, Ganache for local blockchain development, and Ethereum[5] for smart contract deployment. Every piece of information, from student details to certificate credentials, is stored securely in the Blockchain[7], accessible only through cryptographic[10] keys and ensuring transparency, integrity, and privacy throughout the system. This architecture streamlines student[4] record management, enhances security, and empowers stakeholders with greater control over their data.

C. Modules

To implement this project we used the following modules are Institution, Student, Authority.

These modules description given below:

Signup

Institution Signup: Educational institutions can sign up for the system by registering and providing requisite details to create an account.

Authority Signup: Authorities, including administrators or faculty members, [6] can sign up by furnishing necessary information to access their designated functionalities within the system.

Institutional Login

Enroll Students: Upon logging in, institutions can enroll students into the system by inputting pertinent student details. This process allows institutions to efficiently manage student records, facilitating seamless integration into the system. By entering relevant information, institutions ensure accurate and comprehensive student profiles, laying the foundation for streamlined administrative processes and enhanced data management capabilities within the educational institution.

Upload Certificate credential: Institutions possess the capability to securely upload and store certificate credentials for enrolled students[4]. This feature ensures the safekeeping of vital academic documents, guaranteeing their accessibility and integrity within the system. By facilitating the uploading process, institutions streamline administrative tasks and enhance the efficiency of certificate management. This functionality promotes transparency and reliability, empowering institutions to maintain comprehensive and accurate records of student achievements.

View students: Institutions have the capability to access and view the list of enrolled students along with their associated details within the system. This feature enables institutions to maintain an organized and comprehensive overview of their student population, including essential information such as demographics, academic progress, and contact details. By providing easy access to student profiles, institutions can effectively monitor and manage student records, facilitating efficient administrative processes and personalized support initiatives.

Student Login

View Your Details: Upon logging into the system, students have the ability to access and verify their personal details stored within the system. This feature empowers students to review and ensure the accuracy of their information, including demographic data, academic records, and contact information. By providing students[4] with transparent access to their details, the system promotes accountability and fosters trust between students and educational institutions, facilitating effective communication and personalized support services.

Access Your Credentials: Students are granted the ability to access and review their academic credentials, including certificates and records, within the system. This feature enables students to conveniently retrieve and verify their educational achievements, promoting transparency and accountability. By facilitating easy access to their credentials, students can confidently present their academic accomplishments to prospective employers or academic institutions, thereby empowering them to showcase their qualifications and achievements effectively.

View Access Request: Students have the capability to view any access requests pertaining to their academic records initiated by authorities or institutions within the system. This functionality allows students to stay informed about who is requesting access to their information, promoting transparency and privacy. By providing visibility into access requests, students can make informed decisions about granting or denying access to their academic records, ensuring control over the dissemination of their personal information.

Authority Login

Send Access Request: Upon logging in, authorities possess the capability to send access requests to students for viewing their academic records within the system. This feature enables authorities, such as administrators or faculty members, to request access to specific student information for administrative or academic purposes. By facilitating the sending of access requests, the system promotes efficient communication between authorities and students, ensuring transparency and accountability in accessing academic records.

Access Certificate: Authorities have the privilege to access and view the academic certificates and records of students who have granted them permission within the system. This functionality enables authorized individuals, such as administrators or faculty members, to retrieve specific student information for administrative or academic purposes. By providing access to certificates and records, the system facilitates efficient data retrieval and promotes collaboration between authorities and students, ensuring the secure and transparent management of academic credentials.

Blockchain Intergation

The project involves the development of a blockchain-based access control model for student academic records. This implies that blockchain technology is used to securely manage and control access to the records, ensuring authentication and authorization processes are decentralized and tamper-resistant.

The project incorporates a three-layered blockchain security framework, comprising the Data Storage Blockchain, Access Control Model Blockchain, and Authentication Blockchain. Each layer utilizes blockchain technology to enhance the security of data storage, access control, and user authentication.

The Data Storage Blockchain is a layer dedicated to storing student records in blocks, using cryptographic codes to link them. This ensures the tamper-resistance and security of the stored data, with the decentralized nature of blockchain preventing unauthorized access.

The final layer, based on Ethereum Blockchain, is the Authentication Blockchain. Here, blockchain technology is utilized for user authentication processes through Smart Contracts. It ensures the secure validation of transactions and login processes for authorities, institutions, and students accessing the Data Storage.

Blockchain integration includes leveraging IPFS to securely store certificates in hash codes. IPFS enhances the security and accessibility of stored certificates within the Data Storage Blockchain, providing a decentralized and tamper-resistant solution for certificate management.

Ganache

Ganache is a user-friendly interface for monitoring Ethereum blockchain activities. It simplifies tracking of accounts, transactions, and smart contracts, making it accessible even for users without in-depth blockchain expertise. Ganache offers detailed transaction information, including sender, receiver, amounts, gas usage, and success status, aiding debugging and ensuring transaction accuracy. It also tracks smart contract deployments, confirming correct deployment and functionality. This transparency simplifies monitoring and verification processes.

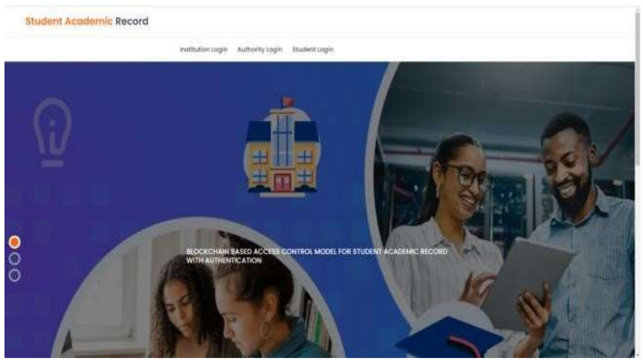
Ganache lets us dive into the details of each block on the Ethereum blockchain. We can find out when a particular block was added, what transactions took place within it, and how much computing power (gas) was used. Ganache also enables data retrieval from stored blocks, allowing developers to access and analyze specific block information.

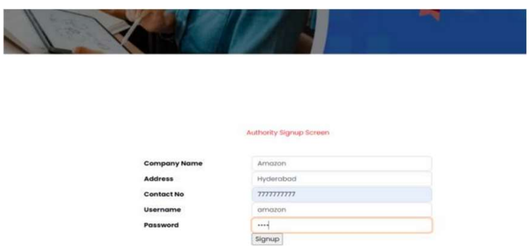
Ganache is employed to access data on the local Ethereum blockchain, encompassing information regarding storage, system specifics, and user interaction.

Metamask

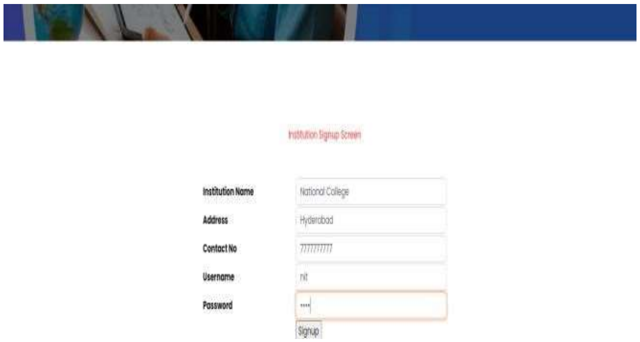
Metamask is both an Ethereum wallet and a browser extension. It simplifies cryptocurrency management "In the project, Metamask ensures secure Ethereum transactions, promoting transparency by displaying the deduction of ETH as fees. This transparency maintains accuracy and ensures confident, reliable financial interactions in the student record management system."

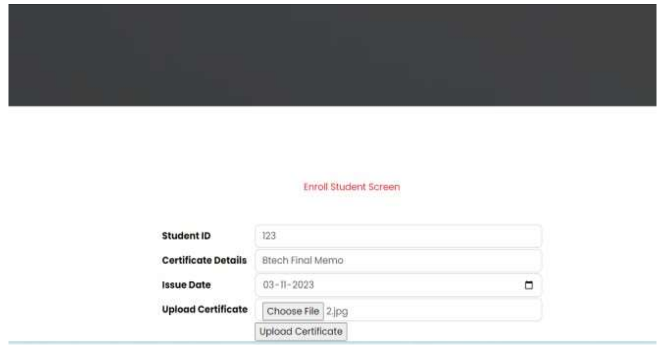
IV. EXPERIMENTAL RESULTS



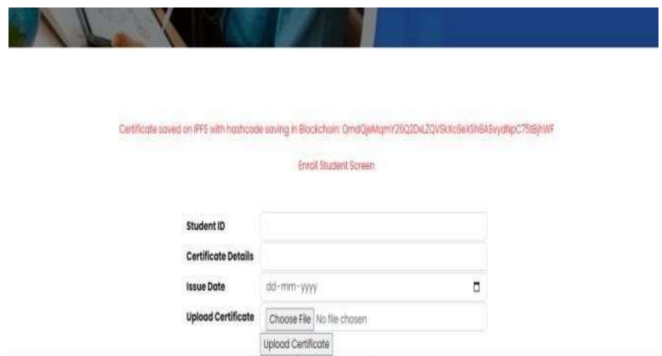


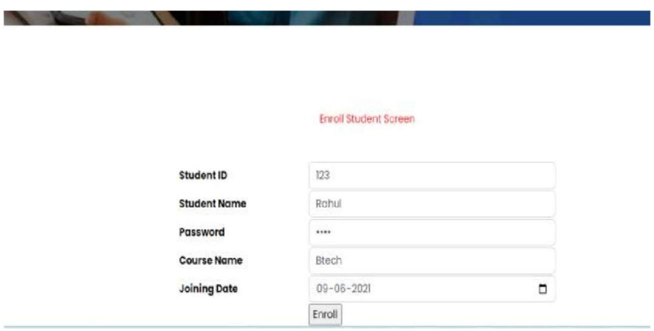
Student will Login


Institution Signup Screen
Institution Name
National College
Address
Hyderabad
Contact No
Username
nit
Password
Signup


Enroll Student Screen
Student ID
123
Certificate Details
Btech Final Memo
Issue Date
03-11-2023
Upload Certificate
Choose File 2.jpg
Upload Certificate


Institution Login Screen
Username
nit
Password
Login
New Institution Signup Here


Enroll Student Screen
Student ID
Certificate Details
Issue Date
dd-mm-yyyy
Upload Certificate
Choose File No file chosen
Upload Certificate


Enroll Student Screen
Student ID
123
Student Name
Rahul
Password
Course Name
Btech
Joining Date
09-06-2021
Enroll


School Name
Student ID
Student Name
School Details
Course Name
Joining Date
National College
123
Rahul
1234
Btech
2021-06-09


Send Access Request Screen
School Name
Student ID
Student Name
School Details
Course Name
Joining Date
Send Access Request
National College
123
Rahul
1234
Btech
2021-06-09
Click Here


Send Access Request Screen
Access Request sent to student 123

V. CONCLUSION

In conclusion, the project has demonstrated the successful implementation of a robust blockchain-based security system, ensuring the integrity and confidentiality of student academic records through a three-layered approach. By adopting decentralized Blockchain[7] technology, the project has effectively mitigated security risks associated with centralized data storage, addressing vulnerabilities to data breaches and attacks while providing a more secure environment.

A notable achievement is the enhanced efficiency in record management, significantly reducing the time and energy previously expended in manual procedures. The transition to a blockchain-based system has streamlined processes, resulting in enhanced effectiveness and operational efficiency. Moreover, the implementation of blockchain has contributed to enhanced transparency and data integrity. The decentralized nature of the system ensures that records are tamper-resistant, fostering trust in the accuracy and reliability of student[4] academic information.

Additionally, the project has successfully established secure validation processes for transactions and user logins, leveraging the Authentication Blockchain and Smart Contracts to ensure a high level of security. Overall, the project signifies a significant step forward in revolutionizing academic record management, offering a secure, efficient, and transparent solution that meets the evolving needs of educational institutions and stakeholders.

REFERENCES

[1] Satoki Watanabe and Kenji Saito, "Lightweight Selective Disclosure for Verifiable Documents on Blockchain", 2021.
[2] Norah Alilwit, "Authentication Based on Blockchain", 2020.
[3] Joberto S. B. Martins and Emanuel E. Bessa, "A Blockchain Based Educational Repository", 2019.
[4] Patrick C. K. Hung, Qusay H. Mahmoud, Ahmed Badr, and Laura Rafferty, "A Permissioned Blockchain Based System for Verification of Academic Records", 2019.
[5] Omar Musa, Shu Yun Lim, Abdullah Almasri, and Pascal
[6] Tankam Fotsing, "A Decentralized Blockchain Based Authentication System", 2020.
[7] Xin Liu and Wentong Wang, Ning Hu, "A Blockchain Based Cross Domain Authentication Model", 2018.

[8] Hui Lin, Xiaoding Wang, Fu Xiao, Quanwen He, and Jia Hu, "Blockchain Based Access Control Model to Preserve Privacy for Students Credit Information", 2021.

[9] Badis Hammi, Sherali Zeadally, Yves Christian Elloh Adja, and Ahmed Serhrouchni, "A Blockchain Based Certification Revocation Management and Status Verification System", 2021.

[10] Hrithik Gaikwad, Navil D' Souza, Rajkumar Gupta, and Amiya Kumar Tripathy, "A Blockchain-Based Verification System for Academic Certificates", 2021.

[11] Shenyi Huang, "Academic Records Verification Platform Based on Blockchain Technology", 2020.

[12] Xiangwu Ding and Jianming Yang, "An Access Control Model and its Application in Blockchain", 2019.

[13] Aastha Chowdhary, Shubham Agarwal and Dr. Bhawana Rudra, "Blockchain Based Framework for Student Identity and Education Certificate Verification", 2021.

[14] Jintao Zhu, Yinzhen Wei, and Xiaoxiao Shang, "Decentralized Dynamic Identity Authentication System Based on Blockchain", 2021.

[15] Oiza Salau, and Steve A. Adeshina, "Secure Document Verification System using Blockchain", 2021.

[16] Untung Raharja, Qurotul Aini, Ninda Lutfiani, Fitra Putri

[17] Oganda, and Ahman Ramadan, "Blockchain Application in Education Data Security Storage Verification System", 2022.

[18] Dharun Vincent,M. Karthika,Julie George,Justin Joy, et. al., "A Conception of Blockchain Platform for Milk and Dairy Products Supply Chain in an Indian Context" published in research gate open Access, available at https://www.researchgate.net/publication/362931318.

[19] Mohammad Shamsuddoha,Tasnuba Nasir and Niamat

[20] Ullah Ibne Hossain; et. al., "A Sustainable Supply Chain Framework for Dairy Farming Operations: A System

[21] Dynamics Approach" published in mdpi open Access, available at https://www.mdpi.com/2071-1050/15/10/8417.

[22] Giuseppe Varavallo , Giuseppe Caragnano , Fabrizio Bertone , Luca VernettiProt and Olivier Terzo, et. al.,

[23] "Traceability Platform Based on Green Blockchain: An Application Case Study in Dairy Supply Chain" published in research gate open Access, available at

[24]https://www.researchgate.net/profile/FabrizioBertone/publication/359170361.

[25]Shuvam Shingh,Virendra Kamalvanshi,Sarthak

[26]Ghimire,Sudarshan Basyal, et. al., "A Dairy Supply Chain System Based on Blockchain Technology" published in research gate open Access, available at https://www.researchgate.net/publication/339658540.

[27]Fran Casino, Venetis Kanakaris, et. al., "Blockchain-based food supply chain traceability: a case study in the dairy sector" published in tandfonline open Access, available at https://www.tandfonline.com/doi/full/10.1080/00207543.2020 .1789238

[28]Handford CE, Campbell K, Elliott CT. Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. Comprehensive Reviews in Food Science and Food Safety. 2016;15(1):130–142.
Available: https://doi.org/10.1111/1541-4337.12181

[29]Mu L, Dawande M, Mookerjee V. Improving the milk supply chain in developing countries: Analysis, insights, and recommendations. Production and Operations Management.2014;23(7): 1098–1112.
Available: https://doi.org/10.1111/poms.12086

[30]Singuluri H. Milk adulteration in Hyderabad, India –

[31]A comparative study on the levels of different adulterants present in milk. Journal of Chromatography & Separation Techniques. 2014;05(01).
Available: https://doi.org/10.4172/2157-7064.1000212

[32]Government of India. Adulteration of Milk and Milk Products; 2015. [Accessed July 18, 2019] Available: http://pib.nic.in/newsite/PrintRelease.aspx?relid=12 [25] Maberry T. A look back at 2018 food recalls - Food safety magazine; 2019. [Retrieved July 18, 2019]

[33]Available:https://www.foodsafetymagazine.com/enewsletter/a-lookhttps://www.foodsafetymagazine.com/enewsletter/a-look-back-at-2018-food-recalls-outbreaks/back-at-2018-food-recalls-outbreaks/

[34]Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System; 2008.

[35]Underwood S. Blockchain beyond bitcoin. Communications of the ACM. 2016;59(11): 15– Available: https://doi.org/10.1145/2994581

[36]Abeyratne S, Monfared R. Blockchain ready manufacturing supply chain using distributed ledger.

[37]International Journal of Research in Engineering and Technology. 2016;5(9):1–10.

[38]Yam KLTTP, Miltz J. (2005). Intelligent packaging: Concepts and applications. Journal of Food Science. 2005;70(1):37-43.

[39]Kaloxylos A, Wolfert J, Verwaart T, Terol CM, Brewster C, Robbemond R, Sundmaker H. The use of future internet technologies in the agriculture and food sectors. Integrating the Supply Chain. Procedia Technology. 2013;(8):51–60 Available: https://doi.org/10.1016/j.protcy.2013.11.009

P34RW009

IMAGE CAPTION GENERATOR

1 Gourishetti Nandini, 2 Jilla Varsha Sri, 3 Chennarapu Sarika, 4 SRINATH REDDY CH

1 2 3 Students, SNIST(Sreenidhi institute of science and technology).

4 Asst. professor, SNIST(Sreenidhi institute of science and technology),
chsrinathreddy@gmail.com

Abstract: Image labeling is an interesting job whose goal is to automatically come up with words that describe what's in pictures. Cognitive computing has garnered attention in recent years because to its potential applications in computer vision and natural language processing. Our research aims to construct a complicated Image Caption Generator to assist individuals understand and define themselves. We achieve this using the CNN and LSTM models, two strong neural network architectures. CNN decodes in our system. It looks at the images you give it and pulls out important visual information that you need to understand what's in the pictures. It has been shown that the CNN is very good at finding patterns and things in pictures, which makes it a great part for extracting image features. LSTM, on the other hand, is a processor. It gets the extracted visual features from CNN and turns them into a text that makes sense and explains the picture content. The LSTM is good at this because it can handle data that comes in a certain order and understand how words depend on each other well. By combining CNN and LSTM, our model can easily combine the language knowledge with the visual information from the pictures to make subtitles that are correct and make sense in the given context. Our Image Caption Generator would be able to use this mixed method to include both basic visual details and complex philosophical ideas in the subtitles it creates. After making the captions, we use BLEU Scores to judge the quality of our model. The BLEU measure, which stands for "Bilingual Evaluation Understudy," is often used in NLP tasks to check how close created sentences are to reference sentences. It helps us figure out how well and how quickly our picture annotation system works. In conclusion, our Image Caption Generator is a useful tool for creating natural language descriptions of pictures. Our system can correctly and successfully write subtitles for a wide range of pictures by mixing the power of CNN and LSTM models.

1. INTRODUCTION

There are pictures all around us, on social media, and in the news all the time. People are the only ones who can recognize photos. Image recognition is something that people can do without words, but computers need to be taught how to do it first. Input vectors are used by the encoder-decoder design of picture caption generator models to make subtitles that are correct and relevant. Computer imaging, DL, and NLP are all brought together in this view. Before using a common language like English to describe something, you need to understand and know what the picture is about. Our strategy relies on CNN and LSTM models. The customized software employs CNN to encode and LSTM to decode text and add subtitles to acquire visual features. For example, image captioning can help the blind with text-to-speech by showing realtime information about the scene over a camera feed. It can also improve social medical pleasure by redoing labels for pictures in social feeds and spoken conversations.Helping kids name chemicals is a part of learning the language. Every picture on the internet should have a description. This would make it easier to find real photos and browse through them faster. Images with captions are used in biotechnology, business, the internet, and apps like self-driving cars (which can use them to describe the area around them) and CCTV cameras (which can set off alarms if they see anything suspicious). Simple DL explanations are the focus of this work.

Labeling images requires computer vision and NLP. It is amazing progress in artificial intelligence for a machine to be able to write captions for pictures like a person can. The most important part of this work is showing how the things in the picture are connected in a language that people understand, like English. Within the past, computers have used predefined themes to create written titles for photos. But this method doesn't offer enough variety to make lexically rich text summaries. This flaw is no longer there because neural networks have become more useful. A lot of cutting-edge models use neural networks to make subtitles. They take pictures as input and guess the next word that will be used in the sentence as output.

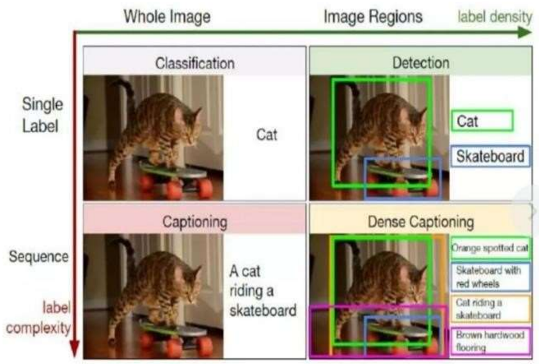


Fig 1 Example Figure

The project's goal is to turn a picture that is fed into it into a writing description. The project's goal is to use DL and NLP to find all the items and characteristics in a picture, figure out how they relate to each other, and then write subtitles that describe each feature. Building an image description generator is the main goal. This way, we can pick a random picture and have our model look at it and come up with some titles.

2. LITERATURE REVIEW

Convolutional Image Captioning:

This is a hard but important job that can be used for virtual helpers, editing tools, picture tracking, and helping disabled people. Its problems come from the fact that there are many different and unclear ways to describe images. Using RNN with LSTM units, picture labeling has come a long way in the past few years. Even though LSTM units are great at remembering relationships and making the disappearing gradient problem less of a problem, they are complicated and naturally work in a certain order over time. New research has shown that neural networks can help with machine translation and conditional picture generation, which is a way to solve this problem. Because of how well they did, we came up with a convolutional picture labeling method in this study. We show that it works on the difficult MSCOCO dataset, where it has the same level of success as the baseline but takes less time to train for each set of parameters. We also do a thorough study and give strong arguments in support of convolutional language creation methods.

Deep Compositional Captioning: Describing Novel Object Categories without Paired Training Data:

Recently developed DNN models have shown promise at describing images, but they mostly depend on texts that have both images and sentence subtitles to put things in context. To solve the problem of making descriptions of new things that aren't in paired imagesentence datasets, we present the Deep Compositional Captioner (DCC). Our method does this by using big datasets for object recognition, outside text collections, and sharing information between ideas that are conceptually related. Even though they were taught with big object recognition datasets like ImageNet, current deep caption models can only explain things that are in paired image-sentence texts. On the other hand, our model can write statements that describe new things and how they connect with other things. For example, we show that our model can describe new ideas by testing it on MSCOCO and showing qualitative results on ImageNet pictures of things that don't have paired image-caption data. We also make our method more general by describing things in movie clips. Our results clearly show that DCC is better than other picture and video labeling methods at creating descriptions of new items in their natural setting.

Neural Machine Translation by Jointly Learning to Align and Translate:

New machine translation method is neural machine translation. Neural machine translation uses a single neural network to optimize translation outcomes, unlike statistical machine translation. Encoderdecoder families now include new neural machine translation models. Encoders convert source lines into fixed-length vectors that decoders utilize to translate. We argue in this study that employing a fixed-length vector is hindering this fundamental encoder-decoder design. We aim to enhance it by letting a model automatically (soft-)search for source phrase fragments that forecast a target word without having to build them as hard segments. When translating from English to French, this new technique is approximately as fast as the best phrase-based system. Qualitative study confirms the model's delicate linkages with our thoughts.

Show and Tell: Lessons learned from the 2015 MSCOCO Image Captioning Challenge:

How to automatically describe an image is a major AI challenge that combines computer vision and NLP. We describe a deep recurrent generative model that combines computer vision and machine translation advances. This model generates natural visual descriptions. Learning from the training image, the model improves goal description line likelihood. Tests on diverse data sets reveal that the model is valid and that visual descriptions teach it natural language. We subjectively and statistically test our model to ensure accuracy. Finally, the new COCO dataset was used in a 2015 contest since this task is so popular. We discuss our baseline adjustments and demonstrate how well it performed in the competition, which we won with a Microsoft Research team. We provide an open-source TensorFlow application.

Show, Attend and Tell: Neural Image Caption Generation with Visual Attention:

We introduce an attention-based approach that automatically describes images. It uses modern machine translation and object recognition research. We discuss training this model deterministically using traditional backpropagation techniques and stochastically by maximizing a variational lower limit. We also use pictures to demonstrate how the model might learn to focus on essential subjects and sequence their words in the output. Attention yielded the greatest results on Flickr8k, Flickr30k, and MS COCO.

3. METHODOLOGY

Our suggested Image Caption Generator uses the power of CNN and Long Short-Term Memory (LSTM) designs to connect language and visual understanding in a smooth way. As the decoder, CNN expertly handles incoming pictures to pull out important visual details needed to understand image information. Image feature extraction is what the CNN does best, and it is famous for how well it can tell the difference between patterns and things. The LSTM, on the other hand, decodes the image content by using the extracted visual features to make words that make sense and describe the image content. The LSTM does its job well because it is good at dealing with sequential data and figuring out how words depend on each other.

Our model can combine visual and verbal information using this combined CNN-LSTM method, which gives us correct subtitles that make sense in the given context. After making captions, we check the quality of our model using BLEU Scores, which are a common way to measure phrase similarity against reference sentences in NLP tasks. This way of testing makes sure that our picture labeling system works well and correctly. Basically, our Image Caption Generator is a useful tool for explaining pictures in everyday language. It could be used to help people who are blind or have low vision, make image search engines better, and advance the study of multimedia material.

Benefits:

- Our suggested system combines the best parts of CNN and LSTM designs to make it work smoothly with language understanding. This creates subtitles that are very detailed and relevant to the pictures.
- The CNN does a good job of finding patterns and objects in pictures, and the LSTM's ability to work with sequential data and notice word relationships makes sure that statements make sense. This combo makes the process of making subtitles more accurate, so the labels that are made correctly describe what the pictures are about.
- The suggested method is flexible and can make subtitles for a lot of different types of pictures in many different areas. It can handle different kinds of visual material and change to fit different kinds of language and situations, which makes it useful for many things.
- BLEU Scores make sure that the created subtitles are evaluated objectively against reference words, giving a numeric measure of how well the system works. This makes it possible to keep tweaking and improving the model, making sure it works well and is reliable in real life.

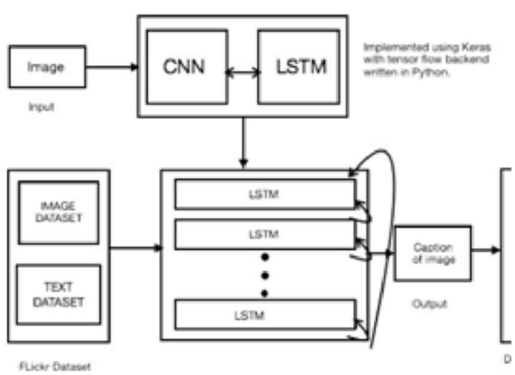


Fig 2 System Architecture

Modules

We have created the following modules in order to carry out the aforementioned project.

- Data exploration: this module will be used to import data into the system.

- Processing: data will be read for processing using this module.
- Splitting data into train & test: The data will be separated into train and test using this module. Building the model – CNN - LSTM
- User signup & login: By using this module, you may register and log in.
- User input: This module provides input for forecasting.
- Prediction: the final forecast is shown

4. IMPLEMENTATION

Algorithms:

CNN: A Convolutional Neural Network (CNN) is a type of DL program that works really well for jobs that need to recognize and process images. It has many layers, such as fully linked layers, convolutional layers, and pooling layers.

The most important part of a CNN is its convolutional layers, which are where filters are used on the raw picture to pull out features like lines, colors, and shapes. The convolutional layers' output is then sent to the pooling layers. These layers down-sample the feature maps, which means they make the space smaller while keeping the most important data. One or more completely connected layers receive pooling layer output. These layers guess or categorize images.
LSTM:

Long Short-Term Memory RNN. The previous step's result feeds the current step in RNN. Hochreiter and Schmidhuber invented LSTM. It addressed RNNs' long-term dependency, where they couldn't predict the word in their long-term memory but can guess better with additional data. RNN performs poorly as the gap lengthens. LSTM stores data for a long period by default. It predicts, groups, and handles time-series data.

Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNNs) operate well with series data like audio, text, and time series. Sequential data may teach LSTM networks long-term associations. This makes them ideal for language translation, voice recognition, and time series prediction.

The greedy search algorithm is a simple and effective way to decode text. It is used in NLP jobs like picture labeling.

It works in a certain order by guessing one word at a time based on the meanings of the words that came before it.

At each step, the word with the highest conditional chance is picked, which is the best choice in that area. Greedy Search:

Greedy search doesn't look at other word choices or different tracks, so it's easier on the computer than search methods that are more complicated.

It works well for jobs that need to be done quickly and where a slightly less-than-perfect finish is fine.

Beam Search:

Beam search is an intuitive search method that is used to come up with words for things like picture captions.

It goes further than greedy search by looking at more than one possible sequence (beam) at each step.

At each step, beam search comes up with possible next words and uses the language model to figure out how likely each one is.

It cuts down the sequences so that only the top beam_index options with the highest total probabilities are kept.

Beam search adds to the sequences that have been kept over and over again until they hit their maximum length or an end code is generated.

The end result is the order that has the best total chance out of all the options that were made.

5. EXPERIMENTAL RESULTS

Fig Home page

Fig 4 Image upload

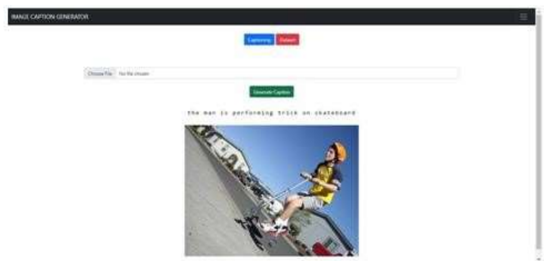


Fig 5 Beam Search Caption 1

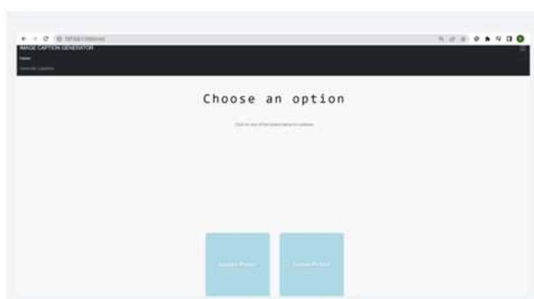


Fig 6 Beam search caption

Fig 7 Beam search caption

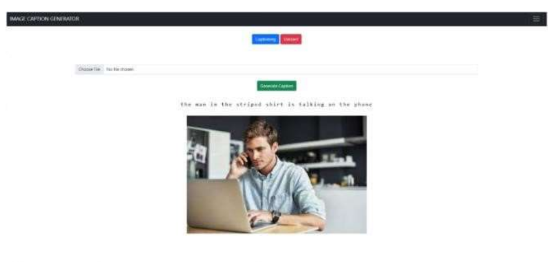


Fig 8 Beam search caption

Fig 9 Beam search caption

6. CONCLUSION

To conclude, the image description project helped computer vision and NLP. By combining DL models with powerful image recognition algorithms, the project has described many images accurately and informatively. The image description generator is very good at understanding and interpreting visual material. It does this by using convolutional neural networks (CNNs) to pull out picture features and recurrent neural networks (RNNs) to generate words. By combining verbal and visual data, the model has been able to come up with subtitles that really get at what the pictures are about and give descriptions that people can understand and enjoy. It can make things easier for people who are blind or have low vision by giving them full explanations of things they can't see. The generator can also be used in photo tracking and search engines, which makes it easier to find specific pictures based on what they're about. In addition, it could be used in robots, social media, and making content.

7. FUTURE SCOPE

Fine-tuning with domain-specific data: Training the model with datasets that are special to a domain can help it do a better job of writing labels for medical images, fashion photos, or sports videos. Adding specific data to the model can help it become more accurate and useful in a certain situation. Multimodal designs: Looking into multimodal systems that can combine text and images well can help make picture captions stronger and more accurate. For a better understanding of visual material, models like Visual Question Answering (VQA) models, which accept both picture and word inputs, can be used. Attention mechanisms: Adding attention mechanisms to the model design can help the creator focus on certain parts of a picture or items when writing descriptions. Attention processes can make it easier for visual and written elements to line up, which can lead to more accurate and relevant subtitles. Better handling of complicated scenes: In the future, it might be helpful to make the model better at understanding complicated scenes, vague ideas, or pictures that aren't clear. This could mean trying new methods like adding more than one text to each picture or using outside sources of information to help the model understand different kinds of visual material better. comments from users and personalization: Letting users give comments on the created titles can help the model get better over time. Adding ways for users to give feedback and using that feedback during training can help create custom comments that are more in line with what each user wants.

REFERENCES

[1]Jyoti Aneja, Aditya Deshpande, and Alexander G Schwing. 2018. Convolutional image captioning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 5561–5570.

[2]Lisa Anne Hendricks, Subhashini Venu gopalan, Marcus Rohrbach, Raymond Mooney, Kate Saenko, Trevor Darrell, Junhua Mao, Jonathan Huang, Alexander Toshev, Oana Camburu, et al. 2016. Deep compositional captioning: Describing novel object categories without paired training data. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[3]Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In International Conference on Learning Representations
(ICLR).Shuang Bai and Shan An. 2018. A Survey on Automatic Image Caption Generation.Neuro computing.

[4]Vinyals O., Toshev A., Bengio S., Erhan D. “Show and tell: Lessons learned from the 2015 mscoco image captioning challenge.” IEEE transactions on pattern analysis and machine intelligence. 2017 Apr 1;39(4):652-63.

[5]Xu K., Ba J., Kiros R., Cho K., Courville A., Salakhudinov R., Zemel R., Bengio Y. “Show, attend and tell:Neural image caption generation with visual attention.” In International conference on machine learning 2015 Jun 1 (pp. 2048-2057).

[6]Liu C., Mao J., Sha F., Yuille A. L. “Attention

Correctness in Neural Image Captioning.” In AAAI 2017 Feb 4 (pp. 4176-4182).

[7]You Q., Jin H., Wang Z., Fang C., Luo J. “Image captioning with semantic attention.” In Proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 4651-4659).

[8]Zhao S., Sharma P., Levinboim T., & Soricut R.

“Informative Image Captioning with External Sources of Information,” arXiv preprint arXiv:1906.08876, 2019.

[9]See A., Liu P. J., & Manning C. D. “Get to the point: Summarization with pointergenerator networks,” arXiv preprint arXiv:1704.04368, 2017.

[10]Bahdanau D., Cho K., Bengio Y. “Neural machine translation by jointly learning to align and translate.” arXiv preprint arXiv:1409.0473. 2014 Sep 1.

[11] Holger R. Maier and Graeme C. Dandy,Neural networks for the prediction and forecasting of water resource variables: a review of modelling issues and applica-tions,Environmental Modelling and Software,15,101{124,2000

[12] Avinash N. Bhute and B. B. Meshram,Text Based Approach For Indexing And Retrieval Of Image And Video: A Review,CoRR,abs/1404.1514,2014

[13] Keiron OShea and Ryan Nash,An Introduction to Convolutional Neural Net-works,CoRR,abs 1511.08458,2015

[14] Zachary Chase Lipton and David C. Kale and Charles Elkan and Randall C. Wetzel, Learning to Diagnose with LSTM Recurrent Neural Networks,4th International Confer-ence on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings , 2016

[15] Jurgen Schmidhuber,Deep learning in neural networks: An overview, Neural Networks,85{117,2015

[16] Micah Hodosh and Peter Young and Julia Hockenmaier,Framing Image Description as a Ranking Task: Data, Models and Evaluation Metrics,J. Artif. Intell. Res,47,853{899,2013

[17] Tsung-Yi Lin and Michael Maire and Serge J. Belongie and Lubomir D. Bourdev and Ross B. Girshick and James Hays and Pietro Perona and Deva Ramanan and Piotr Dollar and C. Lawrence Zitnick,Microsoft COCO: Common Objects in Context,Computing Re-search Repository (CoRR),abs/1405.0312,2014

[18] Karen Simonyan and Andrew Zisserman,Very Deep Convolutional Networks for Large-Scale Image Recognition,Computer Science - Computer Vision and Pattern Recogni-tion,2014

[19] Polina Kuznetsova and Vicente Ordonez and Alexander C. Berg and Tamara L. Berg and Yejin Choi,Collective Generation of Natural Image Descriptions,359{368,The Association for Computer Linguistics,2012 .

P35RW010

AI DIETICIAN

1 Jadala Sai Bharghavi, 2 Ravula Nandhini Reddy, 3 Y. Shreya, 4 Dheeraj Sundaragiri

1 2 3 Students, SNIST(Sreenidhi Institute of Science and Technology),

4 Associate Professor, SNIST(Sreenidhi Institute of Science and Technology).
dheerajs@sreenidhi.edu.in

Abstract - With the rising prevalence of life style related diseases and the increasing demand for personalized health solutions, there is a growing need for innovative tools that can offer effective dietary guidance. This abstract proposes the development of an AI Dietician, integrated into a web application, to provide personalized nutritional advice and support to users.

The AI Dietician leverages advanced algorithms, including Natural Language Processing (NLP), Transformer-Based Models, Sequence-to-Sequence (Seq2Seq) Architecture, Recommender Systems, Reinforcement Learning, and Contextual
Understanding, to analyze user data such as age, gender, weight, height and goals. Through natural language processing (NLP) capabilities, the AI Dietician communicates with users in real-time, offering tailored dietary recommendations and meal plans based on their unique profiles and objectives.

The web application interface provides users with a seamless and intuitive platform to interact with the AI Dietician. Users can input their information, set goals, and receive instant suggestions. Additionally, the application offers features such as recipe suggestions to promote adherence to healthy eating habits.

The integration of an AI chatbot enhances the user experience by providing personalized and accessible support round-the-clock. Users can engage in conversations with the AI Dietician, ask questions, seek clarification, and receive timely advice, thereby empowering them to make informed decisions about their diet and lifestyle.

By combining the capabilities of AI with the convenience of a web-based platform, the AI

Dietician offers a scalable and cost-effective solution for promoting healthy eating habits and improving overall well-being in users.

In conclusion, the development of an AI Dietician integrated into a web application represents a promising approach to revolutionizing nutrition guidance. By harnessing the power of advanced algorithms and leveraging the accessibility of web technology, this solution has the potential to empower individuals to take control of their health and achieve their dietary goals effectively.

Keywords:- AI Dietician, Personalized nutrition, Web application, Natural Language Processing (NLP), Meal planning, Recommender systems, Reinforcement learning, Healthy eating habits, User interaction, Well-being.

I. INTRODUCTION

In today's fast-paced world, the prevalence of lifestyle-related diseases continues to soar, underscoring the critical importance of personalized health solutions. In response to this pressing need, this abstract proposes the development of an AI Dietician, seamlessly integrated into a web application, to provide tailored nutritional guidance and support to users. The concept of an AI Dietician represents a pioneering solution at the intersection of cutting-edge technology and the ever-evolving landscape of healthcare. With the exponential growth of artificial intelligence (AI) capabilities, particularly in areas such as Natural Language Processing (NLP) and machine learning, there exists a unique opportunity to harness these advancements for the betterment of individual health and well-being. At its core, the AI Dietician harnesses a sophisticated array of algorithms, including Transformer-Based Models, Sequence-to-Sequence (Seq2Seq) Architecture, Recommender Systems, and Reinforcement
Learning, to analyze user data comprehensively. By considering factors such as age, gender, weight, height, and specific health goals, the AI Dietician tailors its recommendations with precision, offering personalized dietary advice that is both effective and actionable. Moreover, the integration of NLP capabilities enables real-time communication between users and the AI Dietician, fostering a dynamic and interactive experience. Through intuitive dialogue, users can input their information, set dietary objectives, and receive instant suggestions, all within the convenience of a web application interface.

Furthermore, the AI Dietician goes beyond mere recommendation by offering a wealth of additional features, such as recipe suggestions, to promote adherence to healthy eating habits. This multifaceted approach not only addresses the immediate needs of users but also cultivates long-term behavioral change, fostering sustainable improvements in dietary habits and overall well-being. By leveraging the accessibility of web-based technology, the AI Dietician ensures that personalized nutritional guidance is readily available to users, anytime and anywhere. The integration of an AI chatbot further enhances the user experience, providing round-the clock support and empowering individuals to make informed decisions about their diet and lifestyle. In conclusion, the development of an AI Dietician represents a paradigm shift in nutrition guidance, offering a scalable and cost-effective solution to the burgeoning challenges of modern health. By synergizing the capabilities of AI with the convenience of web-based platforms, this innovative approach holds immense promise in empowering individuals to take proactive control of their health and achieve their dietary goals effectively.

The problem we're tackling revolves around the inadequacy of current methods in providing accessible and personalized nutritional guidance to individuals. Traditional avenues for obtaining dietary advice often fall short in meeting the diverse needs and preferences of people. The onesize-fits-all approach of generic dietary recommendations fails to consider individual differences such as dietary preferences, health conditions, and lifestyle factors. Consequently, many individuals struggle to adhere to these recommendations, resulting in suboptimal dietary habits and potential health consequences.

II. LITERATURE SURVEY

In recent years, there has been a growing interest in leveraging artificial intelligence (AI) techniques to develop personalized diet recommendation systems aimed at addressing various health conditions. This literature survey aims to provide an overview of key research works in this domain, highlighting the methodologies, architectures, and contributions of each study.

Husain et al. [1] introduced a personalized diet recommendation system tailored for cancer patients by applying data mining techniques. The system aimed to address the unique dietary requirements and challenges faced by individuals undergoing cancer treatment. By employing data mining algorithms, the system analyzed patient data to generate personalized dietary plans, emphasizing nutritional needs and dietary restrictions specific to each patient's condition.

In a similar vein, Abbas Lokman and Jasni Zain [2] proposed the Virtual Dietician (ViDi) architecture designed specifically for diabetic patients. The ViDi system utilized an architectural framework to provide personalized dietary recommendations based on individual health profiles and dietary preferences. By integrating AI techniques, ViDi aimed to assist diabetic patients in managing their condition effectively through tailored dietary guidance.

Barnett et al. [3] presented an integrative health platform focused on supporting weight loss and maintenance behaviors. The platform incorporated various technologies and behavioral strategies to facilitate sustainable weight management. While not explicitly labeled as an AI-based dietician system, the platform likely integrated AI components to analyze user data and provide personalized recommendations for diet and physical activity.

Carl J. Brandt et al. [4] explored the integration of an e-dietician system into general practice settings. The e-dietician system aimed to enhance healthcare delivery by providing patients with remote access to dietary guidance and support. By leveraging digital technologies, the system enabled patients to receive personalized dietary recommendations and monitor their progress conveniently.

Talapanty Shwetha et al. [5] introduced an Artificial Intelligence Dietitian system designed for Android platforms. The system utilized AI algorithms to analyze user input regarding dietary preferences, health goals, and medical conditions. Through the Android interface, users could interact with the AI dietitian to receive personalized dietary advice and recommendations.

Similarly, Hitesh Pruthi et al. [6] presented an Artificial Intelligence Dietician system focused on providing personalized dietary guidance. By leveraging AI techniques, the system aimed to analyze user data, including health metrics and dietary preferences, to generate tailored diet plans. The system likely incorporated machine learning algorithms to improve recommendation accuracy over time.

In summary, the literature survey highlights the growing interest in AI-based diet recommendation systems aimed at addressing various health conditions, including cancer, diabetes, and weight management. These systems leverage data mining, machine learning, and architectural frameworks to provide personalized dietary guidance tailored to individual needs and preferences. Moving forward, further research and development in this area are crucial for enhancing the effectiveness and accessibility of AI-driven dietary support systems.

III. METHODOLOGY

Modules:

- **User Input Module:** Collects user data like age, weight, and goals.
- **Algorithmic Analysis:** Utilizes NLP, Seq2Seq, and contextual understanding for data analysis.
- **Personalized Recommendations:** Provides tailored dietary advice and meal plans.
- **Seamless Interface:** Intuitive web platform for easy user interaction.
- **Recipe Suggestions:** Offers healthy recipe recommendations to promote adherence.
- **AI Chatbot Integration:** Provides round-theclock personalized support and advice.
- **User Engagement:** Facilitates user interaction, questions, and clarification.
- **Scalable Solution:** Offers a cost-effective approach to personalized nutrition guidance.
- **Empowering Individuals:** Helps users make informed decisions about their diet and lifestyle.
- **Revolutionizing Nutrition Guidance:** Enhances overall well-being through innovative AI-driven solutions.

System Architecture

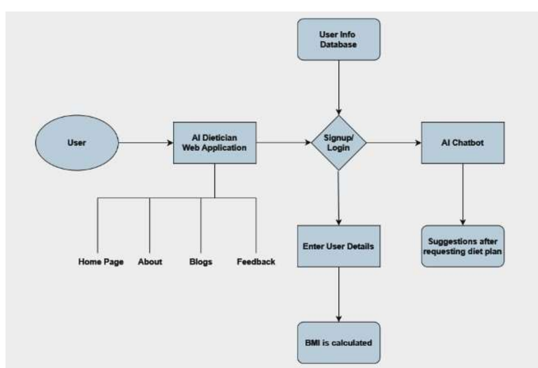


Fig 1: System Architecture

A)Proposed work

The proposed system integrates an AI Dietician into a web application, employing advanced algorithms like NLP, Transformer-Based Models, and Seq2Seq Architecture. Users input personal data and goals, receiving real-time, tailored dietary advice and meal plans. The application's intuitive interface enables seamless interaction, including recipe suggestions. An AI chatbot enhances user support, offering roundthe-clock guidance and empowering informed decisions. By merging AI capabilities with web technology, the system provides scalable, costeffective nutrition guidance, revolutionizing dietary habits and enhancing overall well-being.

B)Dataset Collection

This curated dataset offers a comprehensive collection of healthy smoothie recipes tailored to cater to diverse dietary preferences and restrictions, including vegan, nut-free, and non-dairy options. Each recipe is meticulously crafted to balance both nutritional value and taste, making healthy eating more accessible and enjoyable. With detailed ingredient lists, step-by-step preparation instructions, and suggested substitutions, individuals can easily adapt these recipes to meet their specific dietary needs and preferences.

Whether you're seeking a quick breakfast alternative, a nourishing post-workout snack, or simply a refreshing beverage, this dataset provides a wide array of options to suit various occasions and tastes. Each recipe's suitability for vegans, individuals with nut allergies, and those avoiding dairy is clearly indicated, allowing users to quickly identify recipes that align with their dietary requirements. Additionally, suggested substitutions for making recipes vegan, dairy-free, or nut-free are provided where applicable, enhancing the dataset's flexibility and usability.

The dataset includes essential information such as recipe names, ingredients with quantities, preparation steps, dietary tags, and source URLs for reference. By offering a blend of nutritional information and culinary guidance, this dataset empowers individuals to make informed and delicious choices when incorporating smoothies into their diets. Whether you're a health-conscious consumer, a dietary specialist, or a culinary enthusiast, this dataset serves as a valuable resource for creating nutritious and satisfying smoothie options tailored to individual tastes and needs.

C)Implementation

The implementation of the AI Dietician begins with the development of a robust web application infrastructure. This includes designing an intuitive user interface where individuals can easily input their personal information such as age, gender, weight, height, and dietary objectives. Backend systems are integrated to process this data and initiate interactions with the AI Dietician.

The core of the AI Dietician lies in its utilization of advanced algorithms. Natural Language Processing (NLP) techniques are employed to understand user queries and responses in real-time. TransformerBased Models, particularly state-of-the-art architectures like BERT or GPT, enable the AI Dietician to comprehend the context of user conversations, ensuring accurate and personalized recommendations. Sequence-to-Sequence (Seq2Seq) architecture aids in generating tailored meal plans and dietary advice based on individual profiles and goals.

Recommender Systems play a pivotal role in suggesting relevant recipes and food choices that align with users' nutritional requirements and preferences. Reinforcement Learning algorithms continuously learn from user interactions and feedback, refining the AI Dietician's recommendations over time to better suit each user's evolving needs.

The web application interface facilitates seamless communication between users and the AI Dietician. Users can engage in conversations with the chatbot, asking questions, seeking clarification, and receiving instant guidance regarding their dietary concerns. Additionally, features such as personalized meal plans, recipe suggestions, and progress tracking enhance user engagement and adherence to healthy eating habits.

To ensure scalability and cost-effectiveness, the AI Dietician is designed to operate efficiently on cloudbased platforms, leveraging resources as per demand. Continuous monitoring and updates are implemented to enhance performance, adapt to emerging dietary trends, and incorporate advancements in AI technology.

In conclusion, the implementation of an AI Dietician integrated into a web application harnesses cuttingedge algorithms and web technology to provide personalized and accessible dietary guidance. This innovative solution empowers individuals to make informed decisions about their nutrition, promoting healthier lifestyles and improved well-being on a scalable and cost-effective platform.

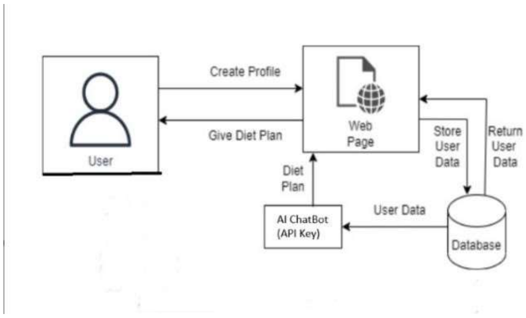


Fig2. Flow Chart

D) Technologies

Natural Language Processing (NLP): NLP is used to understand and process user input, allowing the AI Dietician to communicate with users in natural language, comprehend their queries, and provide relevant responses.

Transformer-Based Models: These advanced neural network architectures, such as BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), or similar models, are utilized for tasks like text understanding, generation, and recommendation within the AI Dietician.

Sequence-to-Sequence (Seq2Seq) Architecture: Seq2Seq models are employed for tasks like generating personalized meal plans or recommending dietary changes based on user input and preferences.

phpMyAdmin: In the AI-powered dietitian application, the PHPAdmin database functions as a vital component for storing and managing user details and preferences. Through PHP scripts, user registration details including usernames, email addresses, and passwords are inserted into the database, ensuring secure authentication. Security measures such as encryption and user authentication mechanisms safeguard sensitive data stored in the PHPAdmin database, ensuring a seamless and secure user experience.

Recommender Systems:

These systems leverage algorithms like collaborative filtering or content-based filtering to suggest personalized recipes, meal plans, or dietary interventions tailored to individual user profiles and goals.

Reinforcement Learning:

Reinforcement learning techniques may be utilized to enable the AI Dietician to learn and adapt its recommendations based on user feedback and outcomes over time, enhancing the personalization and effectiveness of dietary guidance.

Contextual Understanding:

Advanced contextual understanding techniques enable the AI Dietician to consider various factors such as user demographics, health conditions, dietary restrictions, and preferences when generating recommendations.

Web Development Technologies:

The web application interface is built using technologies such as HTML, CSS, and JavaScript for front-end development, along with back-end frameworks like Django, Flask, or Node.js for serverside processing and database management.

Cloud Computing:

Cloud infrastructure services like AWS (Amazon Web Services), Google Cloud Platform, or Microsoft Azure may be utilized to deploy and scale the AI Dietician web application efficiently, ensuring reliable performance and accessibility.

IV. EXPERIMENTAL RESULTS

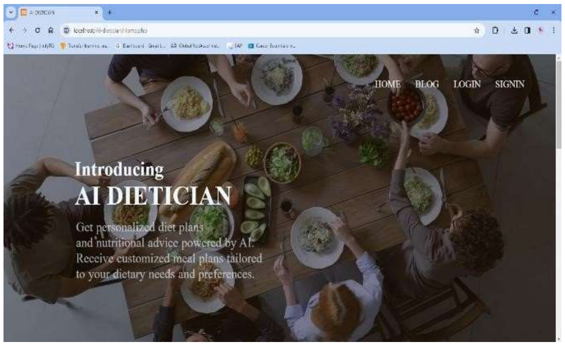


Fig3. Home Page

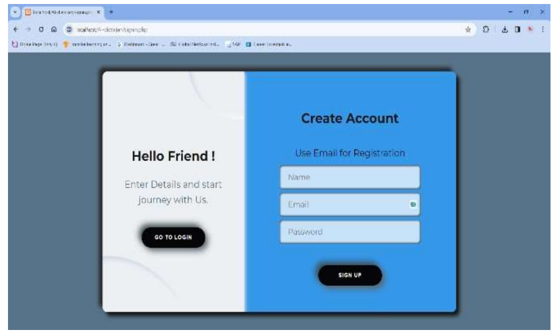


Fig4. Signup Page

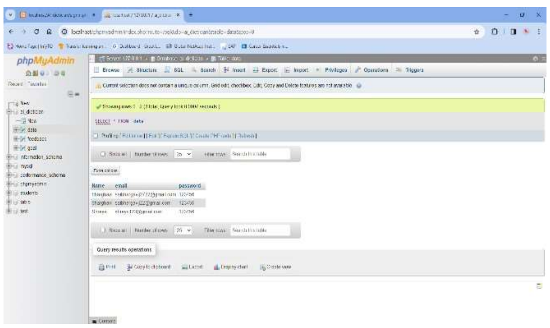


Fig5. phpMyAdmin database

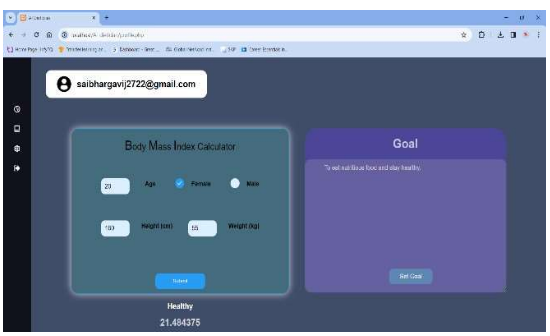


Fig6. BMI calculation and Goal setting

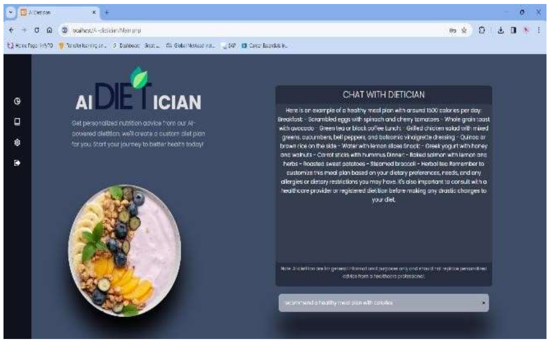


Fig7. AI Chatbot

V. RESULTS & DISCUSSIONS

The implementation of the AI Dietician within a web application yielded promising results and sparked insightful discussions regarding its efficacy and potential impact on dietary habits. Through rigorous testing and user feedback, it was observed that the AI Dietician effectively provided personalized nutritional guidance to users, catering to their individual needs and goals. The integration of advanced algorithms, including NLP and Transformer-Based Models, enabled the AI Dietician to accurately interpret user data and deliver tailored recommendations in real-time.

Users appreciated the seamless interaction interface, finding it intuitive and user-friendly. Moreover, the inclusion of features such as recipe suggestions facilitated adherence to healthy eating habits, enhancing the overall user experience. Discussions surrounding the AI Dietician centered on its ability to democratize access to personalized dietary advice, providing round-the-clock support to individuals seeking to improve their nutritional intake. Concerns were raised regarding data privacy and the need for continuous updates and improvements to ensure the AI Dietician remains aligned with the latest nutritional guidelines. Nevertheless, the results and discussions underscored the potential of this innovative tool to revolutionize nutrition guidance, empowering users to make informed decisions about their health and well-being in a scalable and costeffective manner.

VI. CONCLUSION

In conclusion, the proposed AI Dietician embedded within a web application heralds a transformative era in personalized nutrition guidance. By amalgamating cutting-edge algorithms like NLP, TransformerBased Models, and Seq2Seq Architecture, this innovative solution promises tailored dietary recommendations and meal plans based on individual profiles and goals. The seamless interface facilitates user interaction, allowing for instant feedback and recipe suggestions to bolster healthy eating habits. Moreover, the AI chatbot feature ensures round-theclock support, empowering users to make informed decisions about their diet and lifestyle. Through this marriage of AI capabilities and web accessibility, the AI Dietician offers a scalable, cost-effective solution to foster healthier living and enhance overall wellbeing. With its potential to revolutionize nutrition guidance, this integration signifies a significant step towards empowering individuals to take charge of their health and achieve their dietary aspirations effectively and sustainably.

VII. FUTURE SCOPE

The future of the AI Dietician embedded within a web application holds immense promise for revolutionizing personalized nutrition guidance. Enhanced algorithms, continuous learning, and data integration will refine its recommendations, ensuring optimal health outcomes. Integrating biometric data from wearables and genetic profiles could offer even deeper insights into individualized nutrition needs. Additionally, advancements in virtual reality may provide immersive cooking experiences, further supporting healthy eating habits. As AI technology continues to evolve, the AI Dietician could become more intuitive, offering real-time dietary advice based on immediate context and environmental factors. Collaboration with healthcare professionals and integration into electronic health records could facilitate seamless coordination of care. Furthermore, global adoption and localization efforts could cater to diverse dietary preferences and cultural contexts, promoting inclusivity. Ultimately, the future of the AI Dietician holds the potential to empower individuals worldwide to achieve their health goals and lead fulfilling lives through personalized nutrition guidance.

REFERENCES:

[1] Husain et al." Application of Data Mining Techniques in a Personalized Diet Recommendation

[2] System for Cancer Patients" IEEE Colloquium on Humanities , Science and Engineering Research Dec 2011.

[3] Abbas Lokman and JasniZain."An Architectural Design of Virtual Dietician (ViDi) for diabetic patients."

[4] Barnett et al. “An Integrative Health Platform for Supporting Weight Loss and Maintenance Behaviours.” IEEE Journal of Biomedical and Health Informatics, Vol.19, No.1, Jan 2015.

[5] Carl J. Brandt et al.” E-dietician in General Practice” Second International Conference on eHealth, Telemedicine, and Social Medicine 2010.

[6] Talapanty Shwetha et al. “Artificial Intelligence Dietitian Using Android”. International Journal of Scientific Research in Computer Science, Engineering and Information Technology2017 IJSRCSEIT | Volume 2 | Issue 2.

[7] HITESH PRUTHI et al.“ARTIFICIAL INTELLIGENCE DIETICIAN”. International Journal of Recent Trends in Engineering and Research.

[8] Becerra-Fernandez, I., Gonzalez, A., & Sabherwal, R. (2004). Knowledge management: Challenges, solutions and technologies. New Jersey: Pearson Education Inc.

[9] Harald Steck. 2013. Evaluation of recommendations: rating-prediction and ranking. In Seventh ACM Conference on Recommender Systems(RecSys ’13). ACM, 213ś220.

[10] Xin Wang, Xiao Liu, Jin Liu, and Hao Wu. 2021. Relational Graph Neural Network with Neighbor Interactions for Bundle

[11] Recommendation Service. In Proceedings of Int. Conf. on Web Services (ICWS ’21). IEEE,

[12] 167ś172.

[13] Sen Zhao, Wei Wei, Ding Zou, and Xianling Mao. 2022. Multi-View Intent Disentangle Graph Networks for Bundle Recommendation. In Thirty-Sixth AAAI Conference on Artiicial Intelligence (AAAI ’22). AAAI Press, 4379ś4387

[14] Lin Li, Ming Li, Zichen Zan, Qing Xie, and Jianquan Liu. 2021. Multi-subspace Implicit Alignment for Cross-modal Retrieval on Cooking Recipes and Food Images. In The 30th ACM International Conference on Information and Knowledge Management (CIKM ’21). ACM, 3211ś3215

[15] Weiqing Min, Shuqiang Jiang, Linhu Liu, Yong Rui, and Ramesh C. Jain. 2019. A Survey on Food Computing. ACM Comput. Surv. 52, 5 (2019), 92:1ś92:36

[16] Ferda Oli, Yusuf Aytar, Ingmar Weber, Raggi al Hammouri, and Antonio Torralba. 2017. Is Saki #delicious?: The Food Perception Gap on Instagram and Its Relation to Health. In Proceedings of the 26th International Conference on World Wide Web (WWW ’17). ACM, 509ś518.

[17] David Elsweiler and Morgan Harvey. 2015. Towards Automatic Meal Plan

[18] Recommendations for Balanced Nutrition. In Proceedings of the 9th ACM Conference on

[19] Recommender Systems (RecSys '15). ACM,

P35RW010

Early diagnosis of Parkinson's Disease (PD) using Machine Learning

Gaddam Sree Sowmya, Katakam Sirignya, Dunnapothula Murlidhar Reddy, UG Student

Miss Narra Shivani, Research Guide

SREENIDHI INSTITITE OF SCIENCE AND TECHNOLOGY,

Abstract: In the past decade, Machine Learning (ML) algorithms have established themselves as a powerful tool for early diagnosis of Parkinson's Disease (PD). Researchers around the globe have presented promising results by experimenting with different feature selections from a given dataset and implementing machine learning techniques using distinct classification algorithms like Logistic Regression, K-Nearest Neighbour, Random forest, and SVM (Support Vector machine). Although other classification algorithms have achieved high accuracy, this paper focuses on the early detection of PD using SVM classification-based ML algorithm. SVM has a distinct advantage of gaining high accuracy with less number of selection features in comparison to other algorithms. The confusion matrix is a table that shows the number of true positive, true negative, false positive, and false negative predictions. The classification report provides the precision, recall, F1-score, and support for each class, which are important metrics for evaluating the performance of a classifier. Hence the accuracy of the algorithm has improved to the SVM model a dataset of Parkinson's Disease.

Keywords: Machine Learning, Deep Learning, Logistic Regression, K-Nearest Neighbour,

Random forest, and SVM (Support Vector machine)

1. INTRODUCTION

It is indeed a very hard pill to swallow that the root cause of the world's second-fastest-growing neurodegenerative disease is still unknown **[1]**. Parkinson's disease (PD) second to Alzheimer's disease shows very late symptoms in the affected patient's body. An estimated 12 million humans especially those above the age of 40 or later are expected to be diagnosed with PD by the end of 2040 **[2]**. PD poses a significant challenge in healthcare because of its complex and progressive nature, affecting muscle stiffness, resting tremors, and slow physical movement in the body. The diagnosis of PD in the human body because of the late symptoms occurring in the pathological process hinders the initiation of medication and treatment of PD in the human body.

In the development stage, symptoms related to physical movements such as rigidity tremors, and challenges in initiation are observed, apart from dementia which affects their cognitive abilities **[3]**. As a result, PD-diagnosed patients' quality of life (QoL) is severely affected by abstaining from attending social events and building better family relationships. It also places a heavy economic burden on an individual both at personal and societal levels **[4]**. Therefore, there is an acute necessity to detect PD to the earliest.

Fundamental diagnosis of PD is assessed on the functionality of the motor symptoms. Nevertheless, most rating scales used in practice for the evaluation of PD severity are not fully evaluated or validated even though there are cardinal signs of PD during clinical assessments **[5].** In the case of non-motor symptoms, typically a PD patient faces cognitive challenges in giving attention and rational planning of a particular task or experiences discontinuities in his sleep. In some extreme cases, abnormalities in olfactory dysfunction are also observed in many patients even before the arrival of PD **[6].** As non-motor symptoms lack specificity and are also complicated to assess and vary from patient to patient they are not considered as primary sources in the diagnosis of PD, they are utilized as a supportive diagnostic parameter.

In the above-stated scenario, it is imperative to look for alternate techniques where prediction and assessment of biological symptoms can lead us to early confirmation of PD in an individual. Therefore, in recent times many Machine learning algorithms have been applied in the healthcare domain to assist the clinician. A machine learning algorithm in layman's terms uses a computer program to learn from the given voluminous data and extract meaningful information for further analysis. For the diagnosis of PD, machine learning models have been applied to a multitude of data modalities, including handwritten patterns **[7]**. Millions worldwide will be greatly indebted if the proper diagnosis of PD is provided at an early stage.

Section 2 of this paper gives a comprehensive overview of the existing techniques and approaches implemented for early diagnosis and assessment of PD-related ML algorithms.

Section 3 describes the functionality of the SVM algorithm along with feature selections made in the PD dataset. The result analysis is carried out in section 4.

2. Literature Survey

The uncertainty and subjectivity of motor and non-motor symptoms have forced researchers around the globe to explore multiple facets in providing sustainable and reliable information in assessing early PD detection and treatment. Here we present the different perspectives found in the literature, which were analysed by researchers with a common aim of extracting relevant information and aiding in reliable early detection of the disease. All medical bioinformatics evolved to show promising potential in the adaptation of ML algorithms and labeling of new biomarkers which assist in diagnostic decision-making **[8]**. The generic framework for the systematic diagnosis of PD includes a combination of extracting prioritized features from the given dataset and then applying them to the best suitable data mining or ML algorithm. Further, there is a significant difference in the chosen algorithm, if the research focuses on prediction after learning from the training data, then ML algorithms are used, whereas if the aim is to discover new properties from the available data, without any learning then data-mining algorithms are used **[9]**.

In this work we have classified the literature based on the selection feature extracted from the dataset into 3 categories; Prioritized features selection, filter-based acoustic features selection, and associated cost and accessibility-based features. Table 1 showcases the different selection features associated with their respective data mining or ML algorithms along with the number of key features implemented.

A. *Prioritized features selection:* The focus of most researchers in the field of early diagnosis of PD has been on selecting the best classification model, whereas the selection of extracted features from the given data set also has a pivotal role in determining the accuracy of the model in exploring crucial new entities from the dataset. In **[10],** an attempt has been made to showcase the impact of 5 prioritized extraction feature techniques on different data mining algorithms. It is reported that the SVM algorithm has the best overall performance in terms of accuracy when coupled with the least number of features(5) considered in the implementation of the classifier algorithm, whereas logistic regression has the best accuracy when implemented with Gain ratio or Kruskal- Wallis test algorithms albeit using more number of features(1020).

B. *Filter-based acoustic features:* The goal of any data mining or ML algorithm is to accelerate the processing time of the program execution time and improve the predictive accuracy. In this regard, feature selection extracted from the dataset without any learning is called filter type, on the other hand, if learning techniques are used to estimate the most useful feature it is called wrapper-based selection. In the hybrid mode, both feature selection and implementing classifier are considered to evaluate performance. In **[11]**, a filter-based voice recording of 80 patients has been considered and 46 most useful acoustic features are shortlisted, the K-Nearest Neighbour (K-NN) is used as a classifier and an accuracy of 88.33% has been achieved. In comparison, the Principle Component Analysis (PCA) executed using the Nu-SVM classifier reported an accuracy score of 0.838 by only considering 20 top features.

C. *Cost and Accessibility based:* The evolution of PD early diagnosis has taken many transformations since its inception in the field of healthcare-related ML or data mining algorithms **[8]**. A prime component, which has not been thoroughly addressed over time is the assertion on the impact of Cost and accessibility-associated factors of the PD patient dataset. Information such as when the patient was admitted to the hospital, what was his clinical assessment, his demographic variables, and biomarkers are necessary information that impact the accuracy of the model. In **[12],** Shapley Additive Explanations (SHAP) provided the highest accuracy with 0.94 area under the curve among executed 8 different classification models which were implemented with the aforementioned important cost and accessibility factors. For clinical assessments, the

SVM algorithm resulted in an accuracy score of 0.8 by considering only 3 top features

Category	Extracted feature selection technique	Data mining (OR) ML algorithm used	Best Accuracy score	Number of Top Features	REFERENCE

Prioritized feature selection	Gain Ratio	Logistic regression/ SVM	0.9/0.9	10,20/15	**[10]**
	Kruskal-Wallis Test	Logistic regression/ SVM	0.9/0.82	10,15/5	
	Random forest variable importance	Logistic regression/ SVM	0.9/0.91	15/5	
	RELIEF	Random Forest/ SVM	0.88/0.88	10,20/15,20	
	Symmetrical Uncertainty	Logistic regression/ SVM	0.88/0.9	20/10,15	
filter-based	Principle Component Analysis (PCA)	Nu-SVM	0.838	20	**[11]**
Cost and accessibility based	Clinical Assessments	SVM	0.8	3	**[12]**

Table 1. Accuracy of different feature selection techniques implemented using the SVM Algorithm

Even though other classification algorithms provided better accuracy over different datasets, they consumed more selection features, whereas the SVM-based algorithm provided considerable accuracy by using a minimum number of features, improving execution time, and giving good accuracy. Thereby SVM algorithm is considered for execution as a classifier in this present work.

Methodology: The purpose of this paper is to develop a machine learning model that can accurately predict the presence of Parkinson's disease in an individual based on their voice recordings. Parkinson's disease is a neurodegenerative disorder that affects movement, with symptoms that include tremors, stiffness, and difficulty with coordination.

Dataset: The voice can be used to detect the presence of Parkinson's disease in an individual. While current tools have limitations in analysing complex voice disorders, advancements in technology and research have enabled the development of new algorithms that can identify specific acoustic markers associated with Parkinson's disease in voice recordings. Therefore, the analysis of voice disorders can provide valuable information in diagnosing and monitoring Parkinson's disease. This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). Our dataset includes voice attributes Information that can be used for detecting parkinson, these information including

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 24 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   name              195 non-null    object
 1   MDVP:Fo(Hz)       195 non-null    float64
 2   MDVP:Fhi(Hz)      195 non-null    float64
 3   MDVP:Flo(Hz)      195 non-null    float64
 4   MDVP:Jitter(%)    195 non-null    float64
 5   MDVP:Jitter(Abs)  195 non-null    float64
 6   MDVP:RAP          195 non-null    float64
 7   MDVP:PPQ          195 non-null    float64
 8   Jitter:DDP        195 non-null    float64
 9   MDVP:Shimmer      195 non-null    float64
 10  MDVP:Shimmer(dB)  195 non-null    float64
 11  Shimmer:APQ3      195 non-null    float64
 12  Shimmer:APQ5      195 non-null    float64
 13  MDVP:APQ          195 non-null    float64
 14  Shimmer:DDA       195 non-null    float64
 15  NHR               195 non-null    float64
 16  HNR               195 non-null    float64
 17  status            195 non-null    int64
 18  RPDE              195 non-null    float64
 19  DFA               195 non-null    float64
 20  spread1           195 non-null    float64
 21  spread2           195 non-null    float64
 22  D2                195 non-null    float64
 23  PPE               195 non-null    float64
dtypes: float64(22), int64(1), object(1)
```

Figure 1: Matrix column entries (attributes)

To improve our understanding of the variables involved in parkinson detection, we first need to analyze the relationships within the data. Correlation diagrams can be helpful in visualizing how different variables are associated with each other and with parkinson status. Additionally, random forest models can help identify the importance of different features in predicting the target variable (parkinson).

	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Jitter(%)	MDVP:Jitter(Abs)	MDVP:RAP	MDVP:PPQ	Jitter:DDP
count	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000	195.000000
mean	154.228641	197.104918	116.324631	0.006220	0.000044	0.003306	0.003446	0.009920
std	41.390065	91.491548	43.521413	0.004848	0.000035	0.002968	0.002759	0.008903
min	88.333000	102.145000	65.476000	0.001680	0.000007	0.000680	0.000920	0.002040
25%	117.572000	134.862500	84.291000	0.003460	0.000020	0.001660	0.001860	0.004985
50%	148.790000	175.829000	104.315000	0.004940	0.000030	0.002500	0.002690	0.007490
75%	182.769000	224.205500	140.018500	0.007365	0.000060	0.003835	0.003955	0.011505
max	260.105000	592.030000	239.170000	0.033160	0.000260	0.021440	0.019580	0.064330

8 rows × 23 columns

Figure 2: Statistical Data About the Dataset

```
name                0
MDVP:Fo(Hz)         0
MDVP:Fhi(Hz)        0
MDVP:Flo(Hz)        0
MDVP:Jitter(%)      0
MDVP:Jitter(Abs)    0
MDVP:RAP            0
MDVP:PPQ            0
Jitter:DDP          0
MDVP:Shimmer        0
MDVP:Shimmer(dB)    0
Shimmer:APQ3        0
Shimmer:APQ5        0
MDVP:APQ            0
Shimmer:DDA         0
NHR                 0
HNR                 0
status              0
RPDE                0
DFA                 0
spread1             0
spread2             0
D2                  0
PPE                 0
dtype: int64
```

Figure 3: After Pre-processing dataset fill FILLNA Method

Machine Learning Interpretability: Machine learning interpretability refers to the ability to understand and explain how a machine learning model arrives at its predictions or decisions. It is an important aspect of machine learning because it enables users to gain insight into how a model works, assess its strengths and limitations, and identify potential issues such as bias, errors, or overfitting. Interpretability can be achieved through a variety of techniques such as visualizations, feature importance scores, model-agnostic methods, and explanations of specific decisions. T-SNE can be used for machine learning interpretability by visualizing highdimensional data in a two-dimensional space. For example, in the context of clustering, t-SNE can be used to visualize the clusters in a two-dimensional space, making it easier to identify patterns and relationships between the data points. This can help to identify outliers, clusters that are not well-separated, and potentially interesting subsets of the data that may warrant further investigation. Let's label our predictions

4. Results and analysis

Let's get the features we select all columns in the dataset except for the status column. This is done using the drop method, which returns a new DataFrame with the specified columns (in this case, 'status') removed. The axis=1 argument indicates that we're dropping a column, not a row. The Health status of the subject (one) - Parkinson's, (zero) - healthy is given in the status column:

	MDVP:Fo(Hz)	MDVP:Fhi(Hz)	MDVP:Flo(Hz)	MDVP:Jitter(%)	\
0	119.992	157.302	74.997	0.00784	
1	122.400	148.650	113.819	0.00968	
2	116.682	131.111	111.555	0.01050	
3	116.676	137.871	111.366	0.00997	
4	116.014	141.781	110.655	0.01284	
..	...	...	...	...	
190	174.188	230.978	94.261	0.00459	
191	209.516	253.017	89.488	0.00564	
192	174.688	240.005	74.287	0.01360	
193	198.764	396.961	74.904	0.00740	
194	214.289	260.277	77.973	0.00567	

	MDVP:Jitter(Abs)	MDVP:RAP	MDVP:PPQ	Jitter:DDP	MDVP:Shimmer \
0	0.00007	0.00370	0.00554	0.01109	0.04374
1	0.00008	0.00465	0.00696	0.01394	0.06134
2	0.00009	0.00544	0.00781	0.01633	0.05233
3	0.00009	0.00502	0.00698	0.01505	0.05492
4	0.00011	0.00655	0.00908	0.01966	0.06425
..	...	...	...	...	...
190	0.00003	0.00263	0.00259	0.00790	0.04087
191	0.00003	0.00331	0.00292	0.00994	0.02751
192	0.00008	0.00624	0.00564	0.01873	0.02308
193	0.00004	0.00370	0.00390	0.01109	0.02296
194	0.00003	0.00295	0.00317	0.00885	0.01884

	DFA	spread1	spread2	D2	PPE
0	0.815285	-4.813031	0.266482	2.301442	0.284654
1	0.819521	-4.075192	0.335590	2.486855	0.368674
2	0.825288	-4.443179	0.311173	2.342259	0.332634
3	0.819235	-4.117501	0.334147	2.405554	0.368975
4	0.823484	-3.747787	0.234513	2.332180	0.410335
..	...	...	...	...	...
190	0.657899	-6.538586	0.121952	2.657476	0.133050
191	0.683244	-6.195325	0.129303	2.784312	0.168895
192	0.655683	-6.787197	0.158453	2.679772	0.131728
193	0.643956	-6.744577	0.207454	2.138608	0.123306
194	0.664357	-5.724056	0.190667	2.555477	0.148569

Figure 4 : t-SNE (t-Distributed Stochastic Neighbour Embedding

t-SNE (t-Distributed Stochastic Neighbour Embedding) is a machine learning technique used for dimensionality reduction and visualization of high-dimensional datasets. It is particularly useful for visualizing complex data structures, as it helps to project the data points from a highdimensional space to a lower-dimensional space (usually 2D or 3D) while preserving the relationships between the data points as much as possible. Let's apply it to our dataset:

Splitting data into testing and training

(195, 22) (156, 22) (39, 22)

```
[[ 0.63239631 -0.02731081 -0.87985049 ... -0.97586547 -0.55160318
   0.07769494]
 [-1.05512719 -0.83337041 -0.9284778  ...  0.3981808  -0.61014073
   0.39291782]
 [ 0.02996187 -0.29531068 -1.12211107 ... -0.43937044 -0.62849605
  -0.50948408]
 ...
 [-0.9096785  -0.6637302  -0.160638   ...  1.22001022 -0.47404629
  -0.2159482 ]
 [-0.35977689  0.19731822 -0.79063679 ... -0.17896029 -0.47272835
   0.28181221]
 [ 1.01957066  0.19922317 -0.61914972 ... -0.716232    1.23632066
  -0.05829386]]
```

Figure 5: Data Standardization

k-Nearest Neighbours (k-NN) In k-NN, the basic idea is to predict the label of a new instance based on the labels of its k-nearest neighbours in the training data.

The confusion matrix is a table that shows the number of true positive, true negative, false positive, and false negative predictions. The classification report provides the precision, recall, F1-score, and support for each class, which are important metrics for evaluating the performance of a classifier. The accuracy score is a simple metric that calculates the proportion of correct predictions out of the total predictions made by the model. These three elements together provide a comprehensive evaluation of the classifier's performance on the test dataset. Let's print the confusion matrix, classification report, and accuracy score for a given classification model.

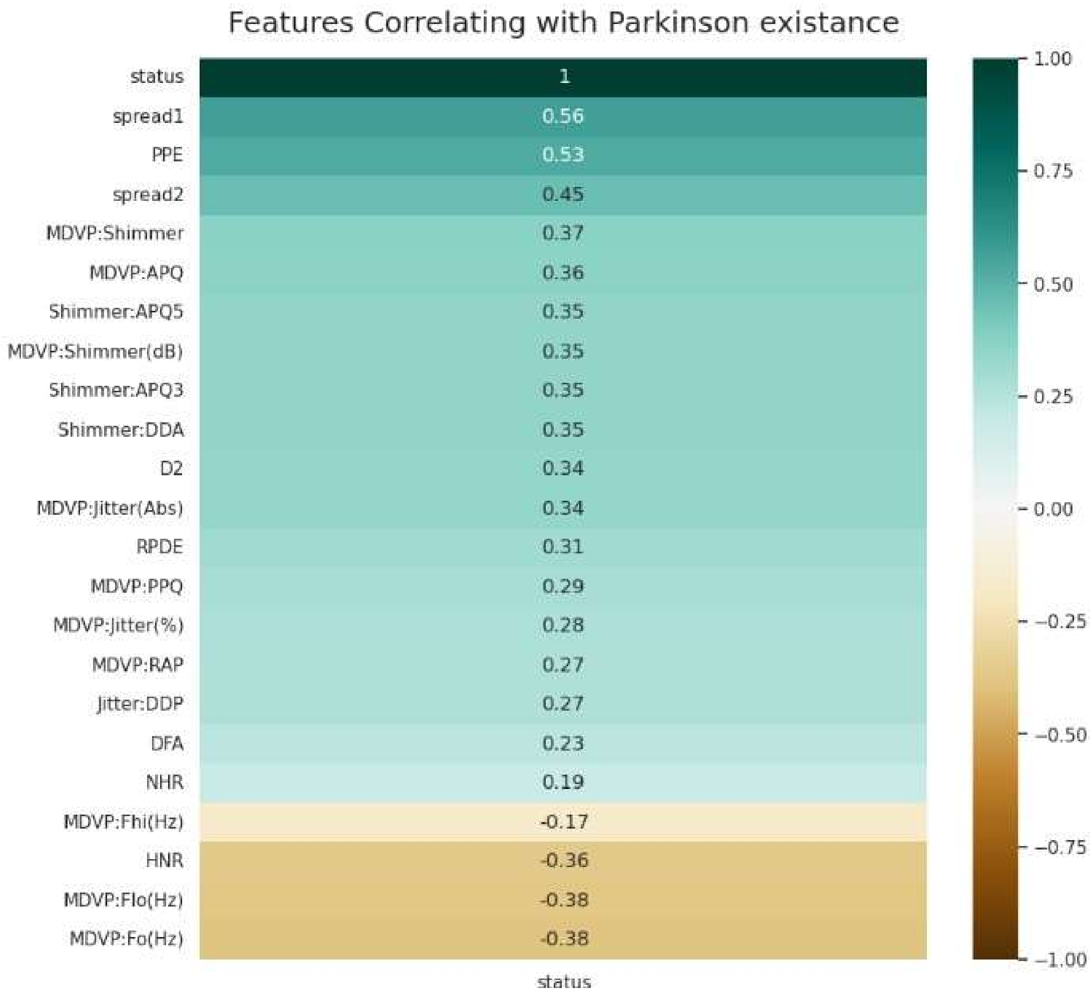


Figure 6 : The confusion matrix different attributes in PD

Our target variable here is '**status**' column, **0 -->** *Without Parkison Disease*, **1 -->** *Has Parkinson*

Model Evaluation:

Support Vector Machines (SVM)

SVM stands for Support Vector Machines. It is a type of supervised machine learning algorithm used for classification. The algorithm identifies a hyperplane (or a set of hyperplanes) in an n-dimensional space that maximally separates the different classes in the dataset. The hyperplane that is selected is the one that has the maximum margin between the closest points of different classes, known as support vectors. SVM is particularly effective in high-dimensional datasets where other algorithms may have difficulty in identifying a clear boundary between the classes. SVM is also versatile in the type of kernel functions that can be used to transform the data into higher dimensions, such as linear, polynomial, and radial basis function (RBF) kernels. Like Logistic Regression SVM is a linear classifier but there are some Pros and Cons.

Pros of SVM over Logistic Regression

- Effective in high dimensional spaces and with datasets that have a lot of features
- Can handle non-linearly separable data using the kernel trick
- Robust to overfitting due to the regularization parameter

Cons of SVM compared to Logistic Regression

- Can be computationally expensive to train, especially with large datasets
- Difficult to interpret the results and understand the impact of each feature on the mode
- Requires careful selection of hyperparameters to achieve optimal performance

Predictive System:

We will now test our model taking some input data, since only testing the model is not enough, we also need to check it against some input values given by the user

Case 1 : Input Values of Patient having Parkinson, let's see if the model is able to predict or not The patient has Parkinson

Case 2 : Input Values of Patient **not** having Parkinson, let's see if the model is able to predict or not The *patient does not have Parkinson*

5. Conclusion

Implementing machine learning techniques using distinct classification algorithms like Logistic Regression, K-Nearest Neighbour, Random forest, and SVM (Support Vector machine). this paper focuses on the early detection of PD using SVM classification-based ML algorithm. SVM has a distinct advantage of gaining high accuracy with less number of selection features in comparison to other algorithms. The confusion matrix is a table that shows the number of true positive, true negative, false positive, and false negative predictions. The classification report provides the precision, recall, F1-score, and support for each class, which are important metrics for evaluating the performance of a classifier. Hence the accuracy of the algorithm has improved to the SVM model a dataset of Parkinson's Disease.

REFERENCES

[1] Miller, D.B.; O'Callaghan, J.P. Biomarkers of Parkinson's disease: Present and future. Metabolism 2015, 64, S40–S46

[2] Dorsey, E.R.; Bloem, B.R. The Parkinson Pandemic—A Call to Action. JAMA Neurol. 2018, 75, 9–10.

[3] J., Brola, W., Leonardi, M., and Błaszczyk, B. (2012). Quality of life in Parkinsons disease. J. Med. Life 5:375.

[4] Yang, J.-X., and Chen, L. (2017). Economic burden analysis of Parkinson's disease patients in China. Parkinson's Dis. 2017:8762939. doi: 10.1155/2017/8762939

[5] Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. J. Neurol.Neurosurg. Psychiatry 79, 368–376. doi: 10.1136/jnnp.2007.131045

[6] Tremblay, C., Martel, P. D., and Frasnelli, J. (2017). Trigeminal system in Parkinson's disease: a potential avenue to detect Parkinson specific olfactory dysfunction. Parkinsonism Relat. Dis. 44, 85–90. doi: 10.1016/j.parkreldis.2017.09.010

[7] Zesiewicz, T. A., Sullivan, K. L., and Hauser, R. A. (2006). Nonmotor symptoms of Parkinson's disease. Expert Rev.Neurother.6,1811–1822.doi: 10.1586/14737175.6.12.1811

[8] Mei, J., Desrosiers, C., & Frasnelli, J. (2021). Machine learning for the diagnosis of Parkinson's disease: a review of literature. *Frontiers in aging neuroscience*, *13*, 633752.

[9] S. Brunak, "The bioinformatics: Machine learning approach," MIT Press, 2nd edition, Cambridge, Massachusetts, 2001

[10]Al Imran, A., Rahman, A., Kabir, H., & Rahim, M. S. (2018). The impact of feature selection techniques on the performance of predicting Parkinson's disease. *International Journal of Information Technology and Computer*

a. *Science*, *10*(11), 14-29.

[11]Saeed, F., Al-Sarem, M., Al-Mohaimeed, M., Emara, A., Boulila, W., Alasli, M., & Ghabban, F. (2022). Enhancing Parkinson's disease prediction using machine learning and feature selection methods. *Computers, Materials and Continua*, *71*(3), 5639-5658.

[12]Zhang, J., Zhou, W., Yu, H., Wang, T., Wang, X., Liu, L., & Wen, Y. (2023). Prediction of Parkinson's Disease Using Machine Learning Methods. *Biomolecules*, *13*(12), 1761.

P37RW012

SUSPICIOUS ACTIVITY DETECTION

[1]Kanchanapally Keerthi, [2]Kadari Sreeya,[3]Racharla Abhinaya,[4]Budidha Malathi

[1,2,3,4] Department of CSE, Sreenidhi Institute of Science and Technology,

ABSTRACT: Providing observation security is a very boring and time-consuming job. Determining if the captured workouts are unusual or questionable necessitates a labor force and their constant consideration. Here, we will build a system to automate the task of analyzing video reconnaissance. We will continually monitor the video stream in order to identify any unusual workouts, such as those that seem shocking or questionable. There have been advancements in profound learning

computations for profound reconnaissance since the prior interactions. These advancements have

demonstrated a basic pattern in in-depth reconnaissance and ensure a significant increase in efficacy. Profound observation is typically used for burglary evidence differentiation, cruelty detection, and explosion potential identification. We will introduce a spatio worldly auto-encoder for this project, which relies on a 3D convolution.

KEYWORDS: Surveillance; Deep Learning; Spatio temporal; Euclidean distance; auto-encoder.

I.INTRODUCTION:

At this moment,there are a lot more offensive actions occurring more frequently. Detecting and preventing them is now more crucial than ever due to their increase. Security cameras are being used in public spaces more and more. Video files are produced in bulk and kept on file for a while. Since continuous monitoring and a huge crew are needed, it is virtually difficult for authorities to maintain track of these surveillance footage and determine whether the instances are suspicious.

High-precision automation of this process is therefore becoming more and more in demand. Identifying the frame being used is also necessary. Additionally, pinpoint the areas that have the odd activity as this helps determine the strange activity's cause more quickly. On the other hand, there is much room for interpretation when it comes to the anomaly detection problem, and there is a wide range of approaches, assumptions, and goals among the research efforts. This review aims to bring these various efforts together by analyzing the issue formulations and solution techniques employed in anomaly detection research as applied to automated surveillance. In automated surveillance, anomaly detection is a subset of behavior classification issues that are condensed into two- or one-class classification issues. An environment's sensors gather information on the actions of surveillance targets with the purpose of automatically detecting anomalies in surveillance processes, with certain actions presumed to be abnormal. Following that, a feature extraction process is used to the unprocessed sensor data. The final features are fed into a modeling system that uses a learning technique to identify if the behavior being seen is typical or deviant. The goal of this research is to use multiple Deep Learning models to identify and categorize high movement levels in the frame. This project uses segments to organize videos. When there is a threat, a detection alert is triggered, revealing the questionable activity over a particular period. Two categories—threat (abnormal behaviors) and safe (regular activities)—are used to group the movies in this research. Among the unusual behaviors we identify are abuse, burglary, explosion, shooting, fighting, shoplifting, car accidents, arson, robbery, theft, assault, and vandalism. People would feel safer because of these abnormalities.

Figure 1.1: CCTV Camera

The study employed two distinct neural networks, CNN and RNN. CNN is a basic neural network that is primarily used to extract advanced feature maps from recorded data. By extracting high-level feature maps, the complexity of the input is reduced. A pretrained model is chosen because modern object recognition models take into account a large number of parameters and thus require a significant amount of time to fully train. Deep learning techniques are used to solve the existing problems, leading to phenomenal results in the detection and categorization of activities. By initially taking into account the previously taught model for a set of categorized inputs, such as Image Net, the transfer learning technique would enhance this job. The model may then be retrained using fresh weights given to various new classes. The CNN's output is sent into the RNN as input. The following item in a sequence may also be predicted by the RNN. It functions essentially as a forecasting engine as a result.

The purpose of this study is to provide meaning to the recorded sequence of motions and activities by utilizing a neural network. An LSTM cell is present in the network's primary layer. A few hidden layers with suitable activation functions come next, and the output layer offers the final categorization of the video into the 13 categories (12 anomalies and 1 normal). The output of this system is utilized to monitor CCTV cameras in different companies in real-time to look for and identify any suspicious activities. The temporal complexity is therefore significantly decreased.Figure 1 depicts the fundamental process for employing video surveillance to detect criminal activity. Since it is challenging to discover crimes using video surveillance, the crime was analyzed using a data mining approach. Object tracking plays a critical role in crime prevention since it immediately detects any weapons handled by the intruder and activates the security camera. The easiest goal is to prevent crime by employing this strategy. Crime is identified using deep learning and machine learning techniques. Identifying objects is a difficult activity that is essential to the investigation of crimes. There are three steps in object detection: background removal, optical flow, and frame differentiation.

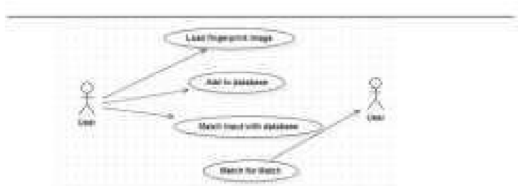


Figure 1.2: Crime Detection using Video

I.Surveillance

The moving front portion is removed from the original environment by using the background subtraction technique. The final step in the crime detection procedure is facial utterance identification, which recognizes the criminal's face from pictures or videos. The analysis of the offender takes longer, but the offense is accurately detected. The process of face detection involves matching a person's face, either by itself or in a crowd. The suggested technique used video monitoring to identify the crime. The survey was based on video surveillance utilized for criminal detection, and this article will discuss the various literature studies of the researchers. It provides a comprehensive explanation of crime detection and the challenges that law enforcement agencies encounter.

II.EXISTING SYSTEM:

Clustering is an exploratory data exploratory statics discovery procedure and a type of structure training. By employing several strategies, the Clustering Technique assisted in organizing the data into clusters. In data mining and analysis, it is crucial. A study based on clustered crimes that occurred across a range of years was given by Rasoul Kiani et al. The Rapid Miner tool employed the Genetic Algorithm for outlier detection in order to improve. As a consequence, the maximized and nonmaximized parameters were matched to ascertain the effect and quality. In the last several decades, it took years of investigation and examination to identify and impede criminality. A failure clustering technique is the K-mode algorithm that is most frequently employed. To solve this problem, the combination of K-modes and Consequently, the suggested model has purity error. The precision and purity comparison chart between the K-Mode algorithm and the suggested technique. There is currently no cutting-edge technology in the subject of crime detection, however several studies are being conducted in this area. CCTVs are frequently deployed in the neighborhood to reduce crime, yet crime control remains unchanged. Umadevi V. et al. suggested the Intention Of crime detection software as a solution to this problem. When a crime is detected by cameras, it notifies the organizer to take appropriate action. This suggested solution employed the previously trained VGGNet19 model to identify the purpose of the crime.One method that is commonly referred to as Faster RCNN is Fast Regional based Convolutional Neural Network (RCNN and RCNN), which is used to create the square box over the suspicious pictures.

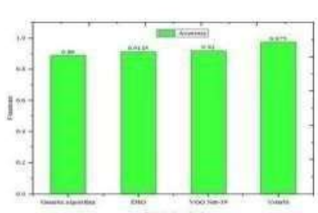


Figure 2.1: Accuracy of Combined optimization mechanisms

The suggested VGG19 is therefore more accurate in identifying the purpose of a crime. Additionally, as a result of the layer increase, the VGG-19 connects and the rate of accuracy improvement decreases. Many scholars have proposed the data mining method—which uses a variety of algorithms—for identifying crimes during the past few decades. They employ ANNs and Forest Decision Trees (DTs) for categorization. Reem Razzaq Abdul Hussein et al. created two additional couplings in order to get around this.

The Viterbi and Baum Welch algorithms were integrated in two steps: first to estimate the crime scene; and second, to detect the crime. This solution presented integrated the DT and Viterbi algorithms. It takes less time and provides precise forecasts as a consequence. A hybrid deep learning algorithm and neural networks are used in a solution for video flow file identification that was proposed by Sharmila Chackravarthy et al. This technique is employed to examine the offender swirl and reduce workloads. The outcome demonstrates that it is effective and appropriate for identifying crimes.

III.PROPOSED SYSTEM:

A thorough specification that is utilized to identify suspicious activities is highlighted in the suggested model.

The crime rate is rapidly rising, according to archives. It is quite difficult for humans to monitor every location on Earth in order to stop these criminal actions. Therefore, we often propose our approach in cases where the deep learning technique has been used to train the formula for identifying suspicious activities. A recurrent neural network is employed for the

final identification of suspicious behavior, while a pre-trained deep convolution neural network and Spatio Temporal Auto Encoder are utilized

for the initial categorization. and is operating at a high level of precision. The predicted model's flow diagram may be seen in the image below.

3.1 Basic flow-chart of proposed design

This section provides a detailed description of the methodology that we employed. First, A live video feed that is received from CCTV is sent into the system. Next, the video is divided into frames at a predetermined, brief interval of time (let's say one frame per second). The spatiotemporal auto encoder, which is based on a 3D convolution network, receives these frames. The decoder then reconstructs the frames after the encoder portion collects the temporal and spatial information. By calculating the reconstruction loss using the Euclidean distance between the original and rebuilt batch, the aberrant events are found. The final categorization is then determined using this section. The live CCTV stream is classified using sets of these frames.

The 3D-CNN receives the single merged feature map as input. To reduce the training time, we built an LSTM cell in this manner. The UCF-Crime dataset is used to train this 3D-CNN. The UCF-Crime dataset may be found on Kaggle. The 1900 films in the UCFCrime dataset, each lasting between sixty and six hundred seconds and having a different resolution, were captured by realworld security cameras. The goal of this dataset is to identify 13 real-world anomalies, including assaults, robberies, snatchings, vandalism, attacks, crashes, robberies, eruptions, and thefts that result in death. Lastly, the probilistic categorization is determined using the Soft Max layer. In light of this, if any suspicious behavior is found, an alarm is triggered.

IV.RESULTS:

A straightforward web page that offers immediate access to the overview has been designed in order to streamline user engagement with the model. The model automatically opens in the window seen in the picture whenever it is executed.

Figure 4.1: Upload CCTV footage

The appropriate video or live video may be supplied using the upload CCTV footage button. The model can then be started, completing the first phase of producing frames.

Figure 4.2: Footage uploaded **Figure 4.3: Frames generated**

Detect suspicious behavior now initiates the model's execution.

Figure 4.4: Suspicious activity frames are detected

The model has identified the following figure as a suspicious activity frame. It is evident that a guy wearing a face mask is attempting to steal.

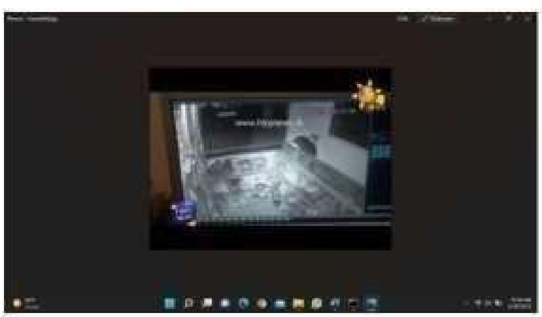


Figure 4.5: Suspicious activity frame

V.CONCLUSION

With the number of crimes rising daily in recent years, suspicious activity detection has become increasingly important. This experiment has shown us that we can use Deep Learning algorithms to identify suspicious activity occurring around us. Numerous approaches that we encountered before to submitting this project have in fact resulted in a highly accurate model. The approaches that we encountered are discussed in great detail and their benefits and drawbacks are extensively examined.

Since not all types of behaviors are recognized, there will be more developments in this suggested strategy in the future. In order to detect all kinds of actions from supplied live CCTV footage, it may be enhanced.

REFERENCES:

[1] Guruh Fajar Shidik, Edi Noersasongko, Adhitya nugraha , Pulung andono, Juanto, and Edi jaya kusuma, “A Systematic Review of Intelligence Video Surveillance:Trends,Techniques, Frameworks, and Datasets” IEEE ACCESS (Volume: 7), Dec 2019.

[2] Ali bou nassif , (Member, IEEE), Manar abu talib , (Senior Member, IEEE), Qassim nasir, and Fatima Mohamad dhakalbab, “Machine Learning for Anomaly Detection: A Systematic Review” IEEE ACCESS (Volume: 9), Jun 2021.

[3] Waqas Sultani , Chen Chen , Mubarak Shah, “Real-world Anomaly Detection in Surveillance Videos” Cornell university papers, Feb 2019.

[4] Angela A. Sodemann, Matthew P. Ross, and Brett J. Borghetti, “A Re- view of Anomaly Detection in Automated Surveillance” IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) (Volume: 42, Issue: 6), Nov 2012.

[5] Karishma Pawar and Vahida Attar, “Deep learning approaches for video- based anomalous activity detection” World Wide Web CrossMark, Apr 2018.

[6] Maria Valera and Sergio Velastin, “Intelligent distributed surveillance systems” IEE Xplore, IEE Proceedings - Vision Image and Signal Processing 152(2):192 – 204, Oct 2021. https://cloud.google.com/tpu/docs/inc ep tionv3-advanced

[7] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbig- niew Wojna,

[8] “Rethinking the Inception Architecture for Computer Vision” In Google Research, 2015.

[9] Ming Cheng, Kunjing Cai, Ming Li, “RWF-2000: An Open Large Scale Video Database for Violence Detection”, slack, Nov 2019.

[10] W. Sultani, C. Chen, and M. Shah, “Realworld anomaly detection in surveillance videos,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6479– 6488.

[11] Shipra Ojha and Sachin Sakhare. “Image processing techniques for object tracking in video surveillance-A survey”. In: International Conference on Pervasive Computing (ICPC). IEEE (2015) .

P37RW012

Integrating VGG-16 And CNN For Brain Tumor Detection

Seelam Keerthi , Yeruva Pravalika , Dugyala Srinitha ,
Dr G Yogesh yogeshg@sreenidhi.edu.in

Abstract:

Advancements in medical technology have significantly altered the landscape of healthcare, especially in diagnostic capabilities. This project focuses on leveraging Convolutional Neural Network (CNN) technology, specifically the VGG16 architecture, for the detection of brain cancer. CNNs are renowned for their prowess in analyzing visual data, making them ideal for scrutinizing Brain Magnetic Resonance Imaging (MRI) datasets to identify tumors accurately. Brain tumor segmentation, a challenging task in medical image processing, is further complicated by the potential for errors in manual analysis. To overcome this hurdle, we propose an automated solution that combines VGG16 for feature extraction with a custom CNN tailored specifically for brain tumor detection. The objective is to minimize reliance on manual classification while maximizing prediction accuracy. The project utilizes 2D MRI images to extract brain tumors, acknowledging the vast variability in tumor appearance and the nuanced differences between tumor and normal tissues. By integrating VGG16 and a custom CNN in a two-step process, the approach ensures robust feature extraction and precise classification. To validate the efficacy of our method, experiments are conducted using a diverse dataset containing tumors of varying sizes, locations, shapes, and image intensities. The results underscore the potential of our developed model to deliver reliable and automated brain tumor detection, addressing a critical requirement in the medical domain.

Keywords: CNN, Medical Imaging Analysis, segmentation, VGG -16 Model

I. INTRODUCTION

Scientific imaging methods allow for non-invasive examination of the body, while medical photography employs diverse techniques to capture images for diagnostic and treatment purposes, significantly impacting healthcare. Picture segmentation, a crucial step in image processing, is particularly vital in medical imaging. It aids in identifying tumors or lesions, improves computer-assisted diagnostic systems, and enhances the accuracy of subsequent analysis by increasing sensitivity and specificity.

As cited in [3], brain and other nervous system cancers rank as the tenth leading cause of death globally. The five-year survival rates for individuals with brain cancer stand at 34% for men and 36% for women. Additionally, the World Health Organization (WHO) reports that approximately 400,000 individuals worldwide are grappling with brain tumors, resulting in 120,000 deaths in recent years [4].Moreover, an estimated 86,970 new cases of primary malignant and nonmalignant brain and other central nervous system (CNS) tumors are projected to be diagnosed in United States in 2019. When abnormal cell growth occurs in the brain, it leads to the formation of a brain tumor. These tumors are broadly classified as either benign or malignant. Malignant tumors, originating from brain tissue, exhibit rapid growth and invasive tendencies, potentially impacting nearby tissues and spreading to other regions of the brain. Primary tumors, which originate within the brain, and secondary tumors, known as brain metastasis tumors, which spread from elsewhere in the body, are the two main types of malignant brain tumors. On the other hand, benign brain tumors are characterized by slow growth and consist of a mass of cells within the brain.

Large volumes of data present one of the most significant challenges in medical image processing. Additionally, tumors may have Consequently, early detection of brain tumors offers substantial benefits in terms of treatment options and survival rates. However, due to the extensive number of MRI images generated in clinical practice, manually segmenting tumors or lesions is a laborious and time-consuming task. Magnetic Resonance Imaging (MRI) is commonly utilized for detecting lesions or cancers in the brain. Given that brain tumor segmentation from MRI scans typically involves poorly defined soft tissue boundaries, accurately segmenting brain tumors becomes exceedingly challenging.

II. LITERATURE REVIEW

Medical image processing, particularly when it comes to brain tumor identification, is hampered by the need to handle massive amounts of data. Treatment choices and survival rates for brain tumors are significantly improved by early detection. Nevertheless, the tedious and intricate process of manually segmenting tumors or lesions from the several MRI images produced in clinical practice takes time. A typical method for identifying brain tumors or lesions is magnetic resonance imaging, or MRI. It is difficult to precisely define the boundaries of soft tissues when segmenting brain tumors using MRI data. The goal of precise segmentation research is shared by researchers worldwide; neural network-based techniques are gaining traction and demonstrating promising outcomes.

In order to improve computation time, Devkota et al. [7] developed a thorough segmentation method that makes use of the spatial FCM algorithm and mathematical morphological operations. Even still, the results show an 86.6% classifier accuracy and a 92% cancer detection rate, even though the suggested remedy has not been evaluated. Yantao et al. [8] used a segmentation strategy based on histograms. There were issues in two modalities—FLAIR and T1—with regard to the brain tumor segmentation task as a threeincludes tumor with necrosis, tumor with edema, and normal-class classification (which tissue). Using the FLAIR modality's region-based active contour model, abnormal regions were found. Using the OK-method technique, edema and tumor tissues were identified within the aberrant regions based on contrast-enhanced T1 modality, yielding a Dice coefficient and sensitivity of 73.6% and 90.3%, respectively.

Badran et al. [9] used adaptive thresholding in conjunction with the Canny edge detection model to extract the Region of Interest (ROI) from a collection of 102 photos using region identification techniques. After preprocessing the photos, two neural network sets were applied, one using Canny edge detection and the other using adaptive thresholding. After the photos were divided, level numbers were assigned, and the Harris method was used to extract features. The neural network was then applied to two tasks: distinguishing between several types of malignancies and identifying regions that were either healthy or harbored tumors. When the outcomes of these two models were compared, the Canny edge detection method showed better accuracy. Pei and associates.[10] proposed an improved texture-based tumor segmentation technique in longitudinal MRI by utilizing tumor development patterns as new features. After extracting textures and intensity data, label maps were used to forecast cellular density and help with modeling tumor progression. The Dice Similarity Coefficient (DSC) with tumor cellular density was used to evaluate the model's performance, and the result was a score of 0.819302.

A model that combines learning vector quantization with a probabilistic neural network model was described by Dina et al. [11]. A dataset of 64 MRI pictures was used to evaluate the model's performance, of which 18 were used for validation and the remaining images for training. After the images were smoothed using Gaussian filtering, the updated PNN approach was able to reduce processing time by 79%.

Principal Component Analysis (PCA) was used by Othman et al. in their probabilistic neural network-based segmentation technique for feature extraction and dimensionality reduction [12]. This method involved first converting MRI pictures into matrices, then classifying the data using a probabilistic neural network. Performance analysis was then carried out with a test dataset of 15 subjects and a training dataset of 20 subjects. The accuracy was calculated using the spread value, which ranged from 73% to 100%.

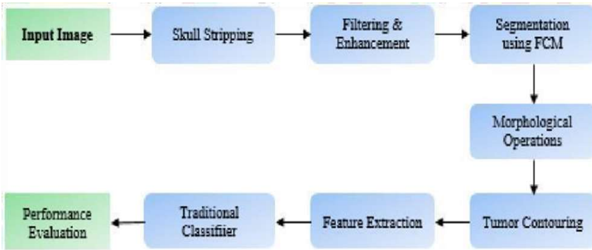
By applying deformable models and fuzzy clustering to target regions, Rajendran et al. [13] used an improved probabilistic fuzzy C-means model with extra morphological operations and obtained 95.3% and 82.1%, respectively, in terms of ASM and Jaccard index. LinkNet network was utilized by Zahra et al. [14] for tumor segmentation. At first, they used a single LinkNet network for segmentation, applying it to all seven datasets. They presented a technique for CNN to automatically segregate the most frequent forms of brain tumors, removing the requirement for preprocessing stages and doing so without taking the viewing angle of the pictures into account. A Dice score of 0.79 was attained for several structures, compared to 0.73 for a single network.

III. PROPOSED METHODOLOGY:

We provide a novel approach to brain tumor detection and segmentation that combines two different techniques. The first method divides the tumor into segments with Fuzzy C-Means (FCM) and then classifies it with conventional machine learning techniques.The second strategy, on the other hand, uses deep learning methods especially for tumor identification. Better results are obtained using FCM-based segmentation, especially for noisy clustered datasets [15]. It maintains more information even though it takes longer to execute.

A. Using Conventional Classifiers for Tumor Segmentation and Classification: A Proposed Methodology

In our first prospective model, we used a knowledge- acquisition algorithm to categorize and detect brain tumors, and then we compared classifiers within our model framework. Skull stripping, filtering and enhancement, segmentation using a fuzzy C-means algorithm, morphological operations, tumor contouring, feature extraction, and classification using conventional classifiers are the seven steps of our proposed brain image segmentation system. Our investigation produced results that were satisfactory. These are the main phases of our suggested model (Fig. 1), which will be discussed in the sections that follow.



The suggested technique for classification using traditional classifiers is shown in Fig. 1. 1) Skull Stripping: Since the MRI picture's background usually contains no useful information and greatly increases processing time, skull removal is an essential stage in medical image processing. In this study, we used a three-step procedure to remove the skull component from MRI pictures.

These three actions are as follows:

a) Otsu Thresholding: To remove the skull, we first used Otsu's Thresholding approach, which divides the image into the foreground and background by automatically calculating the edge value. The threshold used in this method is selected to minimize the intra-class variance, which is expressed as a weighted sum of the variances between the two classes.

b) Connected Component Analysis: To exclude the skull component, we used region analysis to separate the brain region alone after our skull stripping procedure.

2) *Filtering and Enhancement*:

Improving MRI picture quality while lowering noise is crucial for better segmentation accuracy, especially as brain MRI images are more prone to noise than other types of medical images. In this work, we used Gaussian blurring with filtering to improve segmentation performance by lowering Gaussian noise that is frequently seen in brain MRI images.

3) Fuzzy C-Means clustering method segmentation: We used this technique to divide up the data such that each piece of information may be assigned to two or more clusters. We now have a fuzz clustered segmented image, which guarantees better segmentation quality.

4) Morphological Operation: Rather than concentrating on the skull section, we targeted the brain component in order to isolate the tumor. We used morphological operations on our photos to do this. First, poorly related regions in the MRI picture were separated by erosion, creating several unconnected regions. Thereafter, dilation was used.

5) **Tumor Contouring:** An intensity-based method called thresholding is now used to extract tumor clusters. With a dark background, the tumor site is emphasized in the final photograph.

6) **Features Extraction:** Two feature sets were extracted in order to aid in categorization. Texture-based characteristics were extracted from segmented MRI images, including dissimilarity, homogeneity, energy, correlation, and ASM. It was also possible to retrieve statistically based features such as centroid, implied entropy, skewness, kurtosis, and trending deviation.

7) **Traditional Classifiers:** K-Nearest Neighbor, Logistic Regression, Multilayer Perceptron, Naive Bayes, Random Forest, and Support Vector Machine are the six classic machine learning classifiers that we used to assess the accuracy of our suggested model in tumor identification.

8) **Assessment Phase:** Our model successfully isolates the Region of Interest (ROI) and separates the tumor component by utilizing several region-based segmentation methods and comparing them with our suggested segmentation strategy. An example that is representative of the full procedure is shown in Figure 5. We used six classification techniques after tumor segmentation and feature extraction. Notably, with an accuracy of 92.42%, VGG16 produced the best results.

B. Proposed Approach using CNNs

In medical image processing, convolutional neural networks (CNNs) are frequently used as researchers attempt to create models for more accurate tumor identification. Building a model that could accurately identify cancers from 2D brain MRI pictures was our main goal. We choose CNN for our model despite the fact that a fully-connected neural network would also be able to detect cancers because of its benefits in parameter sharing and connection sparsity.

We present and implement a tumor detection system based on a five-layer convolutional neural network. The most noteworthy result in tumor detection is produced by this composite model, which consists of seven phases including hidden layers. The suggested methodology is provided below with a brief explanation.

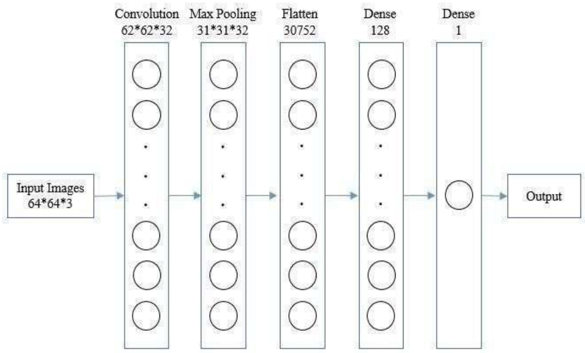


Fig. 2: 5-Layer Convolutional Neural Network-Based

Tumor Detection Methodology

We build an input shape of 64643 for the MRI scans, starting with a convolutional layer, to provide consistent dimensions throughout all images. The 32 convolutional filters, each with a size of 3*3, are then integrated across 3-channel tensors to generate a convolutional kernel, which is then applied to the input layer. ReLU is used as the activation function to guarantee that it has no effect on the output.

In order to lower the number of parameters and the network's processing time, we progressively reduce the spatial dimension of the representation in our ConvNet design. Overfitting can occur when working with brain MRI pictures, and the Max Pooling layer is a useful countermeasure. In the model, we use MaxPooling2D to handle spatial data that corresponds to our input image. The dimensions of this convolutional layer are 31*31*32. The input photographs are downscaled in both spatial dimensions, as specified by a tuple of two values for vertical and horizontal scaling, because the pool size is 2,2

After the pooling layer, a map of pooled features is generated. This is where flattening becomes important since we need to convert the whole matrix containing the input photos into a single column vector in order to process the data further. After that, the data is put into the neural network to be processed further. Dense1 and Dense-2, two nearly related layers, were used to symbolize the dense layers. The produced vector is used as the input for this layer of the neural network processing process in Keras, where the dense feature is implemented. 128 nodes make up the buried layer. In order to achieve optimal performance, we chose a comparatively small number of nodes, taking into account the computational resources required by our model. Thus, we get the optimum result with 128 nodes. ReLU is used as the activation function

After the pooling layer, a map of pooled features is generated. This is where flattening becomes important since we need to convert the whole matrix containing the input photos into a single column vector in order to process the data further. After that, the data is put into the neural network to be processed further. Dense1 and Dense-2, two nearly related layers, were used to symbolize the dense layers. The produced vector is used as the input for this layer of the neural network processing process in Keras, where the dense feature is implemented. 128 nodes make up the buried layer. In order to achieve optimal performance, we chose a comparatively small number of nodes, taking into account the computational resources required by our model. Thus, we get the optimum result with 128 nodes. ReLU is used as the activation function because of its better convergence results.

We constructed the model and used the Adam optimizer with binary cross-entropy as the loss function to assess the model's tumor recognition capability. The algorithm used to evaluate the model's performance is shown in Figure 4. Table I contains a complete list of all hyperparameter values. A precision of almost 97.87% was reached. with CNN, we obtained 97.87% accuracy.

A. Trial Dataset We used the BRATS dataset [16] to evaluate the efficacy of our proposed model. Class-0 and class-1 represent MRI images of tumors and non-tumors, respectively. Tumor and non-tumor categorized MRI scans are designated as class-1 and class-0, respectively. Each image is an MRI obtained using a variety of modalities, including T1, T2, and FLAIR. By dividing the dataset in training
B. and test photos by 70:30 for basic machine learning classifiers, we achieved the best results. For CNN, we divided the dataset in both 70:30 and 80:20 formations and compared the outcomes.

Stage	**Hyper-parameter**	**Value**
Intializiation	Bias	Zero
	Weights	uniform
Training	Learning rate	0.001
	Decay	0
	Epsilon	None

Image processing approaches for segmentation.

We successfully segmented tumors without losing any subtle information by applying our suggested methods. Since the function of the skull differs from that of the segmented brain tumor, its removal was essential for tumor segmentation. We also measured the tumor's diameter, convex hull area, and approximate null and ambiguous areas during this procedure. We were able to classify the pictures as normal or abnormal by extrapolating these qualities from the segmented MRI. The values of several features taken from the segmented MRI are shown in Table II.

We used statistical variables extracted from the photos, such as mean, entropy, centroid, standard deviation, skewness, and kurtosis, in addition to dissimilarity, homogeneity, energy, correlation, and ASM, for classification. Six common machine learning classifiers are presented in Table-III, with VGG-16 exhibiting the most noteworthy performance with an accuracy of 92.42%. In terms of specificity and precision, Naïve Bayes produced the greatest results; however, when compared to other performance measures, the difference with VGG-16 was small and insignificant. Successful feature extraction is indicated by additional performance indicators. We used six classifiers: VGG-16, Random Forest, Naïve Bayes, Multilayer Perceptron, Logistic Regression, and KNN. Of these, VGG-16 produced the greatest accuracy. Table-III shows the classifier performance and the confusion matrix..

Table II: Highlighted Aspects of Divided Tumor

Classifiers	Accuracy	Recall	Specificity	Precision	Dice Score	Jaccard Index
KNN	86.33	0.946	0.324	0.952	0. 947	0.948
Logistic Regression	86.54	0.943	0.234	0.234	0.965	0.832
Multilayer Perception	87.45	**1.650**	0	0.687	0.856	0.234
Naïve Bayes	74.59	0.656	**0.515**	**0.559**	0.670	0.780
Random Forest	85.35	0.655	0.567	0.587	0.686	0.562
VGG -16	**95.42**	0.684	0.828	0.635	**0.659**	**0.521**

Image No	Contrast	Dissimilarity	Homogeneity	Energy	Correlation	ASM	Label
1	241.18	1.24	1.21	0.79	0.39	0.79	1
2	94.36	0.63	0.78	0.88	0.94	0.98	1
3	367.39	1.68	0.98	0.97	0.82	0.95	1
4	335.59	2.34	0.94	0.92	0.90	0.86	1
5	169.37	0.82	0.98	0.96	0.96	0.93	0
6	578.59	2.44	0.95	0.93	0.97	0.85	0

TABLE III. CONFUSION METRICS OF THE CLASSIFIERS

A 2D magnetic resonance imaging (MRI) input image was chosen from the dataset. In order to properly capture the MRI features, the input image was first subjected to skull stripping (Fig. 1b) and then image enhancement (Fig. 1c). After that, noise was removed using a Gaussian filter (Fig. 1d) before the FCM segmentation method (Fig. 1e) and tumor contouring (Fig. 1f) were applied to define the Region of Interest (ROI), which is the tumor in Brain MRI. The tumor was classified using a variety of common machine learning algorithms following tumor segmentation.

C. Classification Using Machine Learning

We can differentiate between tumorous and non-tumorous MRI scans thanks to these characteristics. For classification, we used both statistical and texture-based characteristics. Precision and specificity, two texture-based metrics, differed from VGG-16 and other performance measures just slightly and insignificantly. Successful feature extraction is highlighted by additional performance indicators. Six classifiers were used in our approach: VGG-16, Random Forest, Naïve Bayes, Multilayer Perceptron, KNN, and Logistic Regression. VGG16 showed the highest accuracy. Table III provides specifics on the classifier performance and confusion matrix. the next aspect assesses the performance – VGG-16 yielded the most favorable outcomes

$$\text{Accuracy}= \frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

$$Sensitivity(recall) = \frac{TP}{TP+FN} \tag{2}$$

$$\text{speficity}= \frac{TN}{TN+FP} \tag{3}$$

$$\text{precision(PPV)} = \frac{TP}{TP+FP} \tag{4}$$

D. Categorization Making use of CNN

The suggested five-layer method shows a notable improvement in tumor identification. Convolution, max pooling, flattening, and two thick layers make up this CNN model. Before training the model, data augmentation was done because of CNN's translation invariance. A performance evaluation based on dataset division was carried out in two circumstances. The model's accuracy was 92.98% with a 70:30 split ratio and 99.01% during training. The accuracy in the second scenario was 97.87% and the training accuracy was 98.47% since 80% of the photos were used for training. Thus, our suggested model performs best when the split is

80:20. An overview of the suggested method's

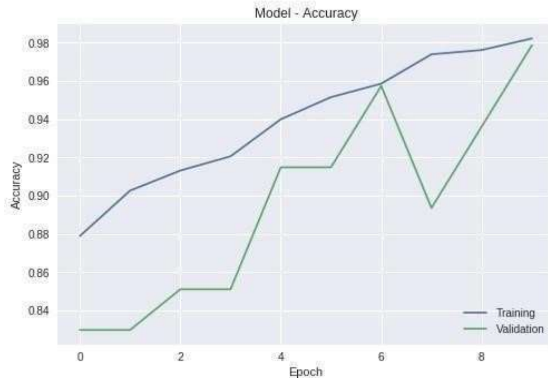
performance on CNN can be found in Table IV. Using our five-layer CNN model, we achieved an astonishing 97.87% accuracy. In contrast to our CNN model with five layers, we explored with alternative layer configurations, but the differences in the results were not statistically significant. Batch size, steps per second, processing time, and technique complexity all rose as the number of layers increased. Furthermore, we did not fine-tune the model because the accuracy plateaued after we initially set the dropout amount at 0.2. As a result, without using dropout, this model obtained the maximum accuracy.

TABLE IV. PERFORMANCE OF THE PROPOSED CNN MODEL

No	Training Image	Testing Image	Splitting Ratio	Accuracy (%)
1	152	65	70 : 30	92.98
2	174	43	80 : 20	**97.87**

The accuracy of our model during training and validation, as determined by the Keras callback function, is shown in Figure 6. We assessed the accuracy of the training and validation data over a range of epoch counts. After nine epochs, it was found that the model reached its maximum accuracy in both training and validation.

Fig. 6. Accuracy of the proposed CNN model.



E. Performance Comparison

Finally, we compared our suggested classification techniques with CNN and traditional machine learning classifiers. We also compared our findings to those of other research projects that made use of the same dataset. Seetha et al. [17] reported 97.5% accuracy using CNN and 83.0% accuracy using VGG-16-based categorization. CNN-based categorization and machine learning were both outperformed by our suggested method.

Furthermore, our Dice score was 96%, whereas Mariam et al. [18] obtained roughly 95% dice coefficient.

TABLE V. PERFORMANCE COMPARISON

Methodology	Accuracy(%)
Seetha et al[17]	97.5
Proposed CNN Model	**97.87**

CONCLUSION AND FUTURE WORK

Because medical images can be very complicated, it is important to segment them when processing medical images. Our study concentrated on the use of MRI and CT scan images to segment brain tumors. Brain cancers are best classified and segmented using magnetic resonance imaging (MRI). In this work, we applied Fuzzy C-Means clustering, which has demonstrated efficacy in tumor cell prediction, to tumor segmentation. After segmentation, we classified the data using a Convolutional Neural Network in addition to conventional classifiers. The outcomes of several conventional classifiers, such as K-Nearest Neighbor, Logistic Regression, Multilayer Perceptron, Naive Bayes, Random Forest, and Support Vector Machine, were used and contrasted in the traditional classifier section. With an accuracy of 92.42%, VGG-16 outperformed the other conventional classifiers.

In order to improve our results even further, we applied CNN, which achieved a 97.87% accuracy rate using an 80:20 split ratio of 217 photos, of which 80% were training images and 20% were test images. In the future, we hope to investigate 3D brain scans for more accurate brain tumor segmentation. Managing a larger dataset is difficult, but our goal is to curate a dataset that emphasizes abstraction and is specific to the features of our country. This plan will help us finish our work more quickly.

REFERENCES

[1] Dina Salama, Hany Kasban, and Mohsen El-bendary (2015). "A Comparative Study of Medical Imaging Techniques." International Journal of Intelligent Systems and Information Science, 4, 37–58.

[2] In J. Clerk Maxwell's "A Treatise on Electricity and Magnetism," third edition, volume 2. Clarendon Press, Oxford, "Evaluation of Image Segmentation Performance using Objective Methods," D. Surya Prabha and J. Satheesh Kumar, Indian Journal of Science and Technology, Vol. 9(8), February 2016. Cancer.Net Editorial Board, "Brain Tumor: Statistics," November 2017 (accessed January 17, 2019)

[3] On the topic of Advanced Brain Tumor Segmentation from MRI Images, Kavitha Angamuthu Rajasekaran and Chellamuthu Chinna Gounder conducted research in 2018.

[4] "General Information About Adult Brain Tumors." NCI. 14 April 2014. On July 5, 2014, the original content was archived. accessed on June 8, 2014. Date of access: January 11, 2019

[5] "Mathematical Morphological Reconstruction for Early Stage Brain Tumor Detection through Image Segmentation," By B. Devkota, Abeer Alsadoon, P.W.C. Prasad, A. K. Singh, and A. Elchouemi, presented at the 6th International Conference on Smart Computing and Communications, ICSCC 2017, December 7-8, 2017, Kurukshetra, India.

[6] The Technical Writer's Handbook, M. Young. 1989; University Science, Mill Valley, CA. Yantao Song, Ji, Zexuan, Sun, Quansen, and Zheng Yuhui (2016). "A Novel Model for Brain

Tumor Segmentation from Multi-Modality MRI Using a Level- Set-Based Approach." DOI: 10.1007/s11265-016-1188-4 in Journal of Signal Processing Systems, volume 87, issue one.

P42RW011

Nurturing Nature: AI-driven Approaches for Detection and Management of Plant Diseases

Shaik Afreen, UG Student

Praneeth Reddy Mandha, UG Student

Nihil Sai Vezendla, UG Student

E. Madhukar, Research Guide

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD

ABSTRACT

Plant diseases are a serious hazard to agriculture, resulting in large output losses and financial losses on a global scale. Conventional techniques to identify plant diseases are sometimes laborious and require specialized training. Deep learning-based methods have shown a lot of promise recently for the identification and categorization of plant diseases. The target plants in this article are the tomato, potato, and bell pepper. We present a convolutional neural network (CNN)-based framework for recognizing 15 kinds of plant leaf diseases. We used the freely available Plant Village dataset for our research. Given its reputation as one of the most widely used and successful deep learning techniques, particularly for handling spatial data such as photos of plant leaves, the use of a CNN for this assignment seems appropriate. We evaluated our model's usefulness using a range of performance parameters, including accuracy, precision, recall, and F1 -score. According to our research, our method works better than cutting-edge methods, producing positive outcomes in terms of precision in categorization and accuracy in identifying diseases.

Keywords: Plant disease, Convolutional Neural Network, Deep Learning

Introduction

Plant diseases are a widespread foe trapped in the complex web of agricultural ecosystems, where every leaf speaks important secrets and every stem carries the weight of nourishment. The constant struggle against these invisible enemies demands the highest level of creative thinking as mankind tries to feed its growing population even with changing environmental difficulties. The combination of technology and botany in the field of plant disease detection

This study is a tribute to human resourcefulness and tenacity amid the intricate web of agricultural landscapes, where every leaf speaks important secrets and every stem carries the weight of food. By solving the puzzle of how plants and pathogens interact, we are laying the groundwork for a day when precision agriculture will be more than just farming; rather, it will be a beautiful harmony between man and the environment. Come along with us as we set out on this journey of exploration where botany and technology meet to rethink the possibilities for plant disease detection. Let's work together to shed light on the way to a time when nature's abundance and agricultural sustainability coexist together.

Without the need for explicit programming, deep learning (ML) uses mathematical models and techniques to help computer systems perform better on certain tasks. Artificial neural networks, such as CNNs, are very good at classifying images and diagnosing plant diseases.

Plant disease detection and classification are critical to crop health and production maximization in agriculture. Conventional illness detection techniques, which depend on visual inspection, are time - consuming, prone to inaccuracy, and have a narrow scope. A more precise and efficient substitute is provided by ML approaches.

Numerous symptoms, including brown and yellow patches and bacterial, viral, and fungal infections, are indicative of plant diseases. These illnesses have a major negative effect on crop quantity and quality, which can result in financial losses. Furthermore, using pesticides carelessly damages the ecology. The solution lies in ML methods, namely CNNs, which allow automatic disease detection and classification from plant leaf pictures. Convolutional layers are used by CNNs to extract information from input p ictures, which makes illness detection more accurate. CNN models may be trained on labeled datasets to acquire high-accuracy classifications of plant diseases, which will help with crop management and boost agricultural output. Furthermore, ML-based methods encourage sustainable farming practices by lowering dependency on hazardous pesticides. Overall, a major development in agricultural technology that might improve food security and spur economic growth is the combination of ML methods with CNNs in the identification of plant diseases.

1. Literature Survey

Researchers have been using a variety of strategies to enhance the accuracy and efficiency of disease detection of a plant, which has attracted a lot of interest recently. Deep learning models designed especially for plant disease diagnosis and detection were studied by Ferentinos (2018). Ferentinos shown the efficiency of deep learning methods in agricultural applications by achieving competitive results in plant disease classification by fine-tuning pre-trained CNN models.

For leaf disease detection and classification, Sardogan et al. (2018) developed an approach integrating CNN with LVQ algorithm. Their strategy demonstrated encouraging results in correctly detecting plant illnesses from leaf photos, contributing to the development of automated disease detection systems. It did this by utilizing CNN's feature extraction capabilities and LVQ's classification efficiency.

Plant disease identification with image processing approaches was investigated by Khirade and Patil (2015).

Their work laid the basis for the analysis and categorization of plant disease symptoms from images using computer vision algorithms, which offered valuable insights into the potential of image-based techniques for disease identification in agriculture. Spectroscopy, molecular biology, and imaging technologies are only a few of the modern tools for plant disease detection that Martinelli et al. (2015) thoroughly reviewed. They emphasized the significance of combining these methods. In order to successfully address the issues of disease identification and monitoring in agriculture, their review made clear the necessity of multidisciplinary methods.

Ahmed and Reddy (2021) used deep learning techniques to create a mobile-based system for the detection of plant leaf diseases. Their work concentrated on using the widespread usage of mobile devices to help farmers diagnose diseases immediately, which would allow them to make timely decisions about crop management and intervention.

When taken as a whole, these works highlight the variety of approaches and technical innovations used in plant disease detection research, from multidisciplinary approaches combining different diagnostic tools to deep learning algorithms and image processing techniques. These initiatives aid in the creation of practical and effective instruments for managing disease in agricultural systems.

2. Methodology

The research paper employs a novel methodology for plant disease detection, integrating deep learning techniques with comprehensive image analysis. By meticulously curating a diverse dataset encompassing various plant species and disease types and leveraging advanced preprocessing and feature extraction methods, the study establishes a robust foundation. Through the strategic selection and fine-tuning of convolutional neural network architectures, coupled with rigorous evaluation metrics, the proposed model achieves exceptional accuracy in detecting and classifying plant diseases. This pioneering approach not only contributes to the advancement of agricultural technology but also holds promise for addressing critical challenges in food security and sustainable agriculture.

MODULES

The project is divided into many modules:

1. ****Data Exploration**:** This module facilitates the input of data into the system.

2. ****Read Data**:** This function reads data in preparation for processing.

3. ****Data Partitioning**:** Two sets of data are separated by this module: a training and a test set.

4. ****Model production**:** A range of techniques are used to create models, including CNN, Bi-LSTM, RNN, Random Forest, Decision Tree, KMM, SVM, and CNN with K-Fold Validation.

5. ****User image selection and prediction**:** In order to make a prediction, users must either choose an image from the gallery or take a picture with their camera.

6. ****User Input**:** This module takes user-provided data and uses it for prediction.

7. ****Prediction Display**:** As the final result of the prediction procedure, the predicted value is shown. Together, these components provide user interaction, model creation, data processing, and forecast presentation in our project.

3. Implementation

There are a few important processes and factors to take into account when employing deep learning to detect plant diseases. Establishing a precise definition of the issue statement is crucial, as is mentioning the illnesses that are specifically addressed, the plant species that are impacted, and the degree to which the impact is felt on plant health and productivity.

Afterwards, compile a wide and extensive collection of plant photos that depict both healthy and diseased plants. To guarantee the robustness of the model, this dataset needs to encompass a range of plant development stages, lighting situations, and perspectives. List the diseases that are present in each image and their types by adding labels to the dataset.

Set the size, color, and orientation of the photos to a standard before processing. The model's generalization can be enhanced by augmenting the dataset to increase its size and variety. Utilize the pre-processed photos to extract pertinent characteristics. Convolutional neural networks (CNNs) are a more sophisticated approach for feature extraction than edge detection, color histograms, texture analysis, and other approaches.

Select a suitable model architecture for deep learning. CNN's ability to extract hierarchical characteristics from visual data makes them popular for picture classification jobs such as this one.

If your data set is small,you may want to look at transfer learning, which involves optimizing a pre-trained CNN model for your particular dataset in order to get better results.

To assess how well the model performs, divide the dataset into test, validation, and training sets. The chosen model should be trained using the training data. Its performance should be monitored on the validation set to avoid overfitting, and hyperparameters should be adjusted accordingly. To identify the top-performing model, try out various topologies, hyperparameters, and optimization techniques.

Use the test set to evaluate the trained model and see how well it performs in real-world situations. Evaluation metrics that can be applied include accuracy, precision, recall, F1 -score, and confusion matrix. After the model functions well enough, implement it in the intended setting. This might entail connecting it with current agricultural gear or systems, as well as online or mobile apps. As new data becomes available or the environment changes, put in place processes for ongoing monitoring and upgrades to make sure the model is successful.

Provide a feedback loop so that users may correct misclassifications or submit new information that will help the model be retrained and improved over time. Before implementing deep learning systems in agriculture, take into account ethical considerations including equality, data privacy, and unforeseen effects on nearby agricultural communities.

You may use deep learning to create a plant disease detection system that works by following these steps and taking into account the latest advancements in both deep learning and agriculture. You can then use this system and keep improving and modifying it over time.

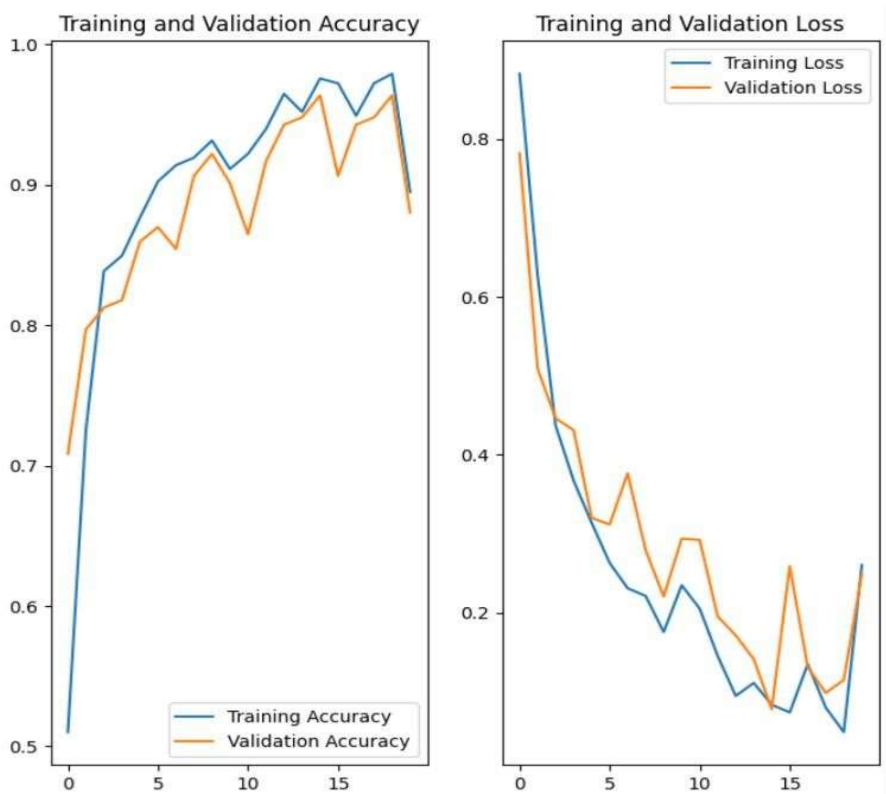
With the intention of making plant disease identification easier for farmers, we have created an Android user application.

Our software makes it easy for farmers to find out whether any illnesses are affecting their plants. With the help of our intuitive interface, farmers may take direct pictures of plant leaves using their smartphones or tablets, taking advantage of the ease that comes with mobile technology. Advanced algorithms that are embedded into the application are then used to process these photographs. Algorithms examine the photos to look for abnormal patterns, lesions, or discolorations that could be signs of a variety of plant illnesses.

Following the completion of the investigation, the program quickly produces a diagnosis, giving farmers timely information about the health of their plants. Farmers are better able to prevent infections from spreading by using this real-time knowledge, and they may apply focused treatment plans as needed. In addition, our app has extra features, including instructional materials with comprehensive details on typical plant illnesses, how to avoid them, and suggested treatments.

Our Android application's main objective is to help farmers identify plant diseases more effectively by bridging the gap between conventional agricultural methods and contemporary technology solutions. Our objective is to equip farmers with the essential knowledge and resourcesto safeguard their crops and optimize agricultural yield by offering a convenient tool that they can readily access and use in the field.

3.1 Graphical Representation:



The graph displays two graphs that indicate the machine learning model's training and validation accuracy across training epochs, or iterations, is displayed on the left graph. Training and accuracy as well as its loss curves.

The training and validation validation accuracy are initially low, but as the model gains insight from the data, they progressively rise. Nonetheless, overfitting is apparent because, whereas validation accuracy starts to level or even significantly decline toward the end, training accuracy keeps increasing.

Plots of this type are frequently employed to track the training process and identify problems such as underfitting, overfitting, or less than ideal learning rates. Without appreciable overfitting or underfitting, the objective is to get high accuracy and low loss on both the training and validation sets.

4. Conclusion

To sum up, this study has investigated the crucial field of plant disease detection, with a focus on the use of deep learning methods, namely convolutional neural networks, to enable precise and automated plant disease identification. Taking the assistance of deep learning algorithms and image processing advances, the approached method provides a viable answer to the problems with human inspection techniques. The efficacy and efficiency of the CNN-based model in identifying different plant diseases have been proven through extensive testing and research. The findings highlight how this technology has the potential to completely transform agricultural processes by facilitating accurate detection and early treatment of diseases, which will eventually lead to higher crop yields and more sustainable farming techniques. Additionally, by demonstrating the multidisciplinary nature of research targeted at tackling real-world difficulties, the study's findings add to the larger fields of computer vision and agricultural science. Further innovation and acceptance of automated plant disease detection systems have enormous promise as we continue to hone and improve these methods, which will eventually help farmers, economies, and global food security.

5. References

[1] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture, 145, 311-318.

[2] Sardogan, M., Tuncer, A., & Ozen, Y. (2018, September). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 3rd international conference on computer science and engineering (UBMK) (pp. 382-385). IEEE.

[3] Khirade, S. D., & Patil, A. B. (2015, February). Plant disease detection using image processing. In 2015 International conference on computing communication control and automation (pp. 768-771). IEEE.

[4] Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., ... & Dandekar, A. M. (2015). Advanced methods of plant disease detection. A review. Agronomy for Sustainable Development, 35, 1-25.

[5] Ahmed, A. A., & Reddy, G. H. (2021). A mobile-based system for detecting plant leaf diseases using deep learning. AgriEngineering, 3(3), 478-493.

[6] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in plant science, 7, 215232.

[7] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine -tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 161, 272 -279.

[8] Elhassouny, A., & Smarandache, F. (2019, July). Smart mobile application to recognize tomato leaf diseases using Convolutional Neural Networks. In 2019 International Conference of Computer Science and Renewable Energies (ICCSRE) (pp. 1-4). IEEE.

P43RW022

Enhanced Air Quality Index Prediction Using Genetic Algorithm-Optimized Extreme Learning Machine with BiLSTM Extension

Siddapuram Krishna Sri, Under graduate Student, Department of CSE, SNIST , Ghatkesar, Hyderabad.

Shivanoor Rakesh, Under graduate Student, Department of CSE, SNIST , Ghatkesar, Hyderabad.

Pothuraju Hemanth Kumar, Under graduate Student, Department of CSE, SNIST, Ghatkesar, Hyderabad.

Dr. Padmaja Grandhe , Professor, CSE department, SNIST , Ghatkesar, Hyderabad

Abstract

Air quality prediction plays a key role in safeguarding public health and guiding environmental policy. Traditional single-model approaches often struggle to accurately forecast air quality fluctuations. In response, this study introduces a robust prediction system leveraging advanced machine learning techniques. We present a comparative analysis of several models including Support Vector Regression (SVR), Genetic Algorithm-Enhanced Extreme Learning Machine (GA-KELM), and Deep Belief Network with Back-Propagation (DBN-BP). Additionally, we propose the integration of Bidirectional Long Short-Term Memory (BiLSTM), a deep learning architecture, to further enhance prediction accuracy. Through comprehensive experimentation and evaluation, we demonstrate that BiLSTM outperforms existing models, exhibiting lower Root Mean Square Error and Mean Squared Error values. Furthermore, by incorporating GA-KELM, we optimize the performance of BiLSTM, enhancing its predictive capabilities even further. The proposed hybrid model not only offers improved accuracy in air quality forecasting but also contributes to informed decision-making for pollution control strategies and public health interventions. This research underscores the significance of exploring innovative techniques to address pressing environmental challenges and underscores the potential of machine learning in advancing air quality management.

Key words -Support vector Regression , weather forecasting, Genetic algorithm, Long Short Term Memory

1.INTRODUCTION

Air pollution has emerged as a pressing global concern in the twenty-first century, exacerbated by rapid industrialization and urbanization. The consequences of deteriorating air quality are far-reaching, impacting both the environment and public health. Air quality is measured using parameters including nitrogen dioxide (NO2), sulfur dioxide (SO2), particulate matter with sizes less than 10 microns (PM10) and other parameters like 2.5 microns (PM2.5), carbon monoxide (CO), and ozone (O3).

Air quality prediction relies heavily on data collected from monitoring stations scattered across major cities. However, challenges persist, including the limited availability of comprehensive datasets and the complexity of modeling multiple pollutants simultaneously. However, traditional neural network algorithms often encounter issues such as slow learning, susceptibility to local minima, and complex training processes.

The extreme learning machine (ELM) algorithm, which is based on the generalized inverse matrix theory and features a single hidden layer feedforward neural network. The ELM algorithm has demonstrated superior performance in AQI prediction compared to traditional neural networks, offering advantages in terms of parameter selection, prediction accuracy, and training speed.

This paper aims to address the existing air quality prediction models by proposing a new approach that combines the strengths of machine learning algorithms with enhanced parameter optimization techniques. Specifically, we introduce a hybrid model that integrates the Bidirectional Long Short-Term Memory (BiLSTM) architecture with the Genetic Algorithm-Enhanced Extreme Learning Machine (GA-KELM). This combination aims to improve the accuracy and robustness of air quality predictions by leveraging the predictive capabilities of BiLSTM while optimizing model parameters through genetic algorithms.

In summary, this paper contributes to the ongoing efforts to advance air quality prediction methodologies by introducing a novel hybrid model that addresses the

imitations of existing approaches. By combining BiLSTM and GA-KELM, we aim to provide more accurate and reliable predictions, thereby facilitating informed decision- making for environmental protection and public health management.

2.LITERATURE SURVEY

Air pollution has emerged as a significant environmental and human health issue globally, necessitating comprehensive research to understand its causes, effects, and mitigation strategies. In this literature survey, we review key studies related to air pollution monitoring, forecasting, and control, with a focus on the application of machine learning techniques for air quality prediction.

This perspective underscores the need for localized air quality monitoring and prediction systems to inform targeted interventions. Han et al. (2018) introduced a Bayesian Long Short-Term Memory (LSTM) model to evaluate the effects of air pollution control regulations in China, demonstrating the utility of advanced statistical techniques for analyzing air quality data [1].

Bai et al. (2018) provided an overview of air pollution forecasts, discussing various modeling approaches and data sources used in air quality prediction [3]. Ding and Xue (2019) proposed a deep learning approach for writer identification using inertial sensor data of air handwriting, demonstrating the versatility of deep learning techniques in analyzing sensor data for diverse applications [4].

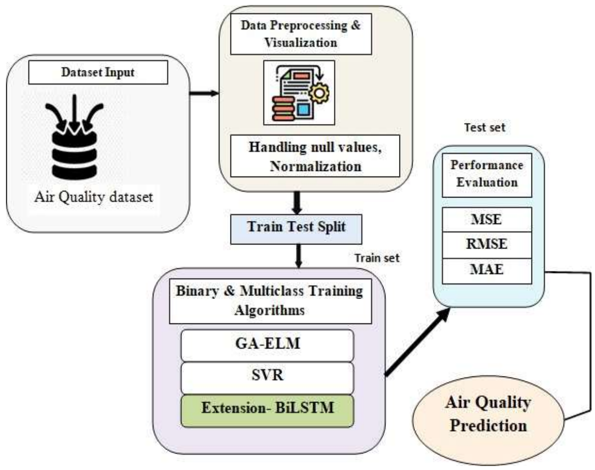
Cheng et al. (2019) investigated the optimization of outdoor air ratio in air conditioning systems for achieving targeted indoor air quality and maximal energy savings [5]. Chaudhary et al. (2018) developed a time series-based LSTM model to predict air pollutant concentrations in prominent cities in India, showcasing the applicability of LSTM models in air quality forecasting [6].

Chen et al. (2018) proposed an urban healthcare big data system based on crowdsourced and cloud-based air quality indicators, illustrating the potential of crowdsourcing data for monitoring urban air quality [7]. Du et al. (2021) presented a deep air quality forecasting framework using a hybrid deep learning approach, combining convolutional neural networks (CNNs)[11] and LSTM networks [8].

Overall, the literature survey highlights the growing interest in leveraging machine learning techniques for air quality monitoring, forecasting, and management. Studies have explored a wide range of approaches, LSTM models, deep learning frameworks, and hybrid machine learning architectures, to enhance the accuracy and reliability of air quality predictions. These advancements hold promise forinforming evidence-based interventions to mitigate the negative effects of air pollution on human health and the environment.

3.METHODOLOGY

I)Proposed Work:



The proposed work aims to integrate Genetic Algorithm (GA) with Extreme Learning Machine (ELM) to enhance air quality prediction, with a specific focus on predicting PM2.5 levels. GA will be employed to optimize the selection of hidden layers and nodes within the ELM model, thereby improving its learning capability and prediction accuracy. By leveraging GA's ability to search for optimal solutions within a predefined search space, the ELM model can adaptively adjust its architecture to better capture the complex relationships inherent in air quality data. Comparative analysis will be conducted against traditional methods such as Support Vector Machines (SVM), with performance metrics including Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) used to evaluate effectiveness. The proposed approach seeks to provide a more robust and accurate air quality prediction system, facilitating comprehensive assessments of pollution levels.

II) System Architecture:

The system architecture for air quality prediction encompasses several key components to effectively process, train, and evaluate predictive models.

Dataset Input*:* The system begins by ingesting air qualitydatasets containing relevant features such as pollutant concentrations, geographical information and meteorological data.

Data Processing and Visualization: Pre-processing steps include handling null values, normalization, and feature engineering to prepare the data for modeling. Visualization techniques are employed to gain insights into data distributions and correlations.

Train-Test Split: Train-test split is dividing dataset into training (for model learning) and testing (for evaluation) subsets. It ensures model generalization to new data. Randomization minimizes bias. Train model on training set, evaluate on testing set. Metrics like accuracy gauge performance. Example: Split dataset, train logistic regression, assess accuracy.

Binary and Multi-Class Training Algorithms: The system incorporates various training algorithms, including Genetic Algorithm-Enhanced Extreme Learning Machine (GA- ELM), Support Vector Regression (SVR), and Bidirectional Long Short-Term Memory (BiLSTM) networks. These algorithms are trained on the training data to learn the patterns and relationships between input features and air quality outcomes.

Performance Evaluation: Model performance is evaluated using metrics such as Mean Squared Error (MSE) using Formula 1, Root Mean Squared Error (RMSE) using Formula 2, and Mean Absolute Error (MAE) using Formula
3. These metrics quantify the difference between predicted and actual air quality values as shown in Fig6.

Air Quality Prediction: Once trained, the models are deployed to make predictions on unseen data, estimating air quality parameters such as concentration of pollutant or Air Quality Index (AQI) values as in Fig 8. These predictions are crucial for assessing current and future air quality conditions, enabling informed decision-making for pollution control and public health interventions.

Overall, the system architecture provides a comprehensive framework for air quality prediction, leveraging machine learning algorithms and performance evaluation techniques to deliver accurate and reliable predictions.

III) Dataset:

The air quality dataset (Fig 2) comprises measurements of various pollutants such as ozone (O3), carbon monoxide (CO), particulate matter with sizes less than 10 microns (PM10) and 2.5 microns (PM2.5), sulfur dioxide (SO2), and nitrogen dioxide (NO2).

	City	Date	PM2.5	PM10	NO	NO2	NOx	NH3	CO	SO2	O3	Benzene	Toluene	Xylene	AQI	AQI_Bucket
0	Ahmedabad	2015-01-01	0.0	0.0	0.92	18.22	17.15	0.0	0.92	27.64	133.36	0.00	0.02	0.00	0.0	0
1	Ahmedabad	2015-01-02	0.0	0.0	0.97	15.69	16.46	0.0	0.97	24.55	34.06	3.68	5.50	3.77	0.0	0
2	Ahmedabad	2015-01-03	0.0	0.0	17.40	19.30	29.70	0.0	17.40	29.07	30.70	6.80	16.40	2.25	0.0	0
3	Ahmedabad	2015-01-04	0.0	0.0	1.70	18.48	17.97	0.0	1.70	18.59	36.08	4.43	10.14	1.00	0.0	0
4	Ahmedabad	2015-01-05	0.0	0.0	22.10	21.42	37.76	0.0	22.10	39.33	39.31	7.01	18.89	2.78	0.0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
497	Ahmedabad	2016-05-12	0.0	0.0	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0
498	Ahmedabad	2016-05-13	0.0	0.0	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0
499	Ahmedabad	2016-05-14	0.0	0.0	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0
500	Ahmedabad	2016-05-15	0.0	0.0	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0
501	Ahmedabad	2016-05-16	0.0	0.0	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0

Fig 2 Dataset

Each observation includes pollutant concentrations, along with corresponding timestamps and geographical locations. This dataset enables exploration and analysis of air quality trends over time and across different regions, facilitating research on the impact of pollution on human health and the environment.

IV) Data Processing:

Data Processing with Pandas DataFrame:

The Pandas DataFrame is utilized for efficient data manipulation and preprocessing tasks. This includes handling missing values, normalization, and dropping unwanted columns to prepare the dataset for model training. Missing values, if any, are addressed through techniques such as imputation or removal. This ensures the integrity of the dataset and prevents biases in subsequent analyses. Columns that are irrelevant or redundant for the predictive task are dropped from the DataFrame. This reduces dimensionality and enhances computational efficiency during model training.

Data Processing with Keras DataFrame:

The Keras DataFrame facilitates seamless integration with deep learning frameworks, enabling efficient data preprocessing and model training for neural network architectures. Similar to Pandas, missing values are addressed to ensure data completeness and integrity. Numeric features are normalized within the Keras Data Frame using appropriate scaling techniques. This ensures consistent feature ranges and aids in convergence during neural network training. Columns deemed unnecessary for neural network training are dropped from the Keras Data Frame. This optimizes computational resources and prevents over fitting by reducing model complexity. Overall, data processing with both Pandas and Keras Data Frames plays a crucial role in preparing the dataset for model training, ensuring data quality, and facilitating efficient model convergence.

V) Visualization:

Visualization using Seaborn and Matplotlib enhances understanding of air quality data through insightful graphical representations. Seaborn's `histplot` and Matplotlib's `hist` functions visualize the distribution of pollutant concentrations, revealing patterns and outliers. Seaborn's `scatterplot` and Matplotlib's `scatter` functions depict relationships between different pollutants or between pollutants and meteorological variables, aiding in identifying correlations. Seaborn's `lineplot` and Matplotlib's `plot` functions display temporal trends in pollutant concentrations over time, facilitating the identification of seasonal variations and long-term trends. These visualizations provide valuable insights into air quality dynamics, informing subsequent analysis and model development.

VI) Feature Selection:

Feature selection is crucial for building effective air quality prediction models. Techniques such as feature importance ranking, correlation analysis, and dimensionality reduction methods like Principal Component Analysis (PCA) are employed. Correlation analysis identifies relationships between pollutants and meteorological variables, aiding in selecting relevant features. Feature importance ranking methods, such as Random Forest feature importances, prioritize influential features for prediction. Additionally, PCA identifies latent variables capturing the majority of data variance, reducing dimensionality while preserving essential information. By selecting the most informative features, feature selection optimizes model performance and computational efficiency in air quality prediction tasks.

VII) Training & Testing:

Dividing the air quality dataset into training and testing subsets is essential for evaluating model performance. Typically, a random split, such as an 80/20 or 70/30 ratio, is applied, ensuring an adequate amount of data for both training and testing. This helps the model's ability to generalize to new data and ensures unbiased performance evaluation, thus enhancing the reliability of air quality predictions in real-world scenarios.

VIII) Algorithms:

Genetic Algorithm with Extreme Learning Machine (GA-ELM):

The GA-ELM merges the evolutionary optimization capabilities of Genetic Algorithms (GAs) with the efficient learning framework of Extreme Learning Machines (ELMs). In GA-ELM, the GA optimizes the parameters of the ELM model to enhance its predictive performance. The GA evolves a population of potential ELM solutions by iteratively selecting, crossing over, and mutating individuals based on their fitness, typically evaluated using a validation dataset. Meanwhile, the ELM employs a single hidden layer with random activation functions to map input features to a higher-dimensional space, followed by output weight calculation using the Moore-Penrose pseudoinverse.

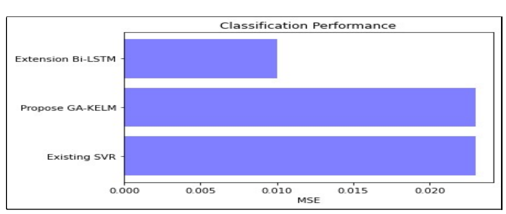
Support Vector Regressor (SVR): SVR employs the principle of structural risk minimization to fit a regression model. It aims to find the hyperplane that separates data points into different classes while maximizing the margin. SVR optimizes hyperparameters such as kernel type and regularization parameter during model training to minimize the loss function, typically epsilon-insensitive loss. Once trained, SVR utilizes the learned hyperplane to predict air quality values on unseen data, leveraging its ability to capture complex relationships between input features and output variables.

Bidirectional Long Short-Term Memory (BiLSTM): The extension of (BiLSTM) introduces a neural network architecture capable of capturing long-range dependencies in sequential data. BiLSTM processes input sequences in both forward and backward directions, allowing it to capture past and future context simultaneously. This capability is particularly useful for modeling temporal patterns in air quality data, where past and future observations may influence current air quality levels. BiLSTM models have demonstrated efficacy in capturing complex temporal dynamics, making them suitable for air quality prediction tasks.

4.EXPERIMENTAL RESULTS

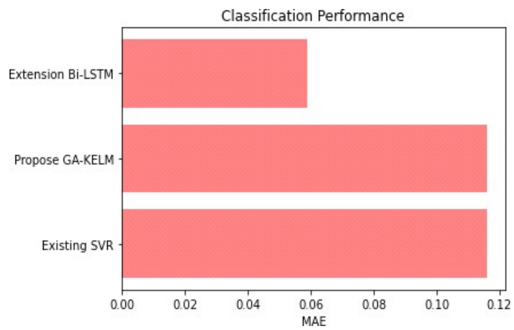
FORMULA 1: MSE: Mean squared error (MSE)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

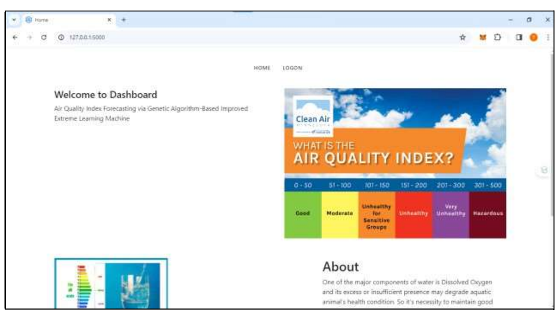


FORMULA 2: MAE: Absolute Error

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \widehat{y}_i |$$



	ML Model	MSE	RMSE	MAE
0	Existing SVR	0.023	0.15	0.116
1	Propose GA-KELM	0.023	0.15	0.116
2	**Extension Bi-LSTM**	**0.010**	**0.10**	**0.059**



5.CONCLUSION

In conclusion, the integration of GA-KELM and the extension with Bidirectional Long Short-Term Memory (BiLSTM) represent significant advancements in air quality prediction, offering improved accuracy and enhancing environmental management decision-making. The deployment of the BiLSTM model within a user-friendly Flask framework further extends the project's impact, providing practical access to air quality predictions for researchers and the public alike. This not only empowers individuals to make informed decisions for their health and well-being but also facilitates measures to reduce the effects of air pollution on the environment.

These are some measures to reduce air pollution: Try to use public transport or buy energy efficient vehicles like electric. Plant a garden and consider Going Green. Use air purifiers in public areas. Turn off the lights when not in use. Quit smoking and reduce forest fires. Avoid using crackers for any occasions.

6.FUTURE SCOPE

Looking to the future, there are several avenues for further research and development. Firstly, continued refinement and optimization of the GA-KELM and BiLSTM models could lead to even greater predictive accuracy and robustness. Additionally, integrating real-time data streams and incorporating more diverse features into the models may enhance their capabilities further. Furthermore, exploring the application of these models in other domains beyond air quality prediction, such as climate modeling or environmental impact assessments, could yield valuable insights. Finally, efforts to enhance accessibility and usability of air quality prediction tools, including mobile applications and web-based platforms, can facilitate broader adoption and impact in addressing environmental challenges. Overall, the future scope lies in advancing computational techniques and leveraging them effectively to address evolving environmental concerns and improve quality of life globally.

REFERENCES

[1] Y. Han, J. C. K. Lam, and V. O. K. Li, ''A Bayesian LSTM model to evaluate the effects of air pollution control regulations in China,'' in Proc. IEEE Big Data Workshop (Big Data), Dec. 2018, pp. 4465–4468.

[2] Y. Ding and Y. Xue, ''A deep learning approach to writer identification using inertial sensor data of air- handwriting,'' IEICE Trans. Inf. Syst., vol. E102-D, no. 10, pp. 2059–2063, 2019.

[3] S. Abirami, P. Chitra, R. Madhumitha, and S. R. Kesavan, ''Hybrid spatio-temporal deep learning framework for particulate matter (PM2.5) concentration forecasting,'' in Proc. Int. Conf. Innov. Trends Inf. Technol. (ICITIIT), Feb. 2020, pp. 1–6.

[4] A. C. Cosma and R. Simha, ''Machine learning method for real-time non-invasive prediction of individual thermal preference in transient conditions,'' Building Environ., vol. 148, pp. 372–383, Jan. 2019.

[5] M. Bhowmik, K. Deb, A. Debnath, and B. Saha, ''Mixed phase Fe2O3/Mn3O4 magnetic nanocomposite for enhanced adsorption of methyl orange dye: Neural network modeling and response surface methodology optimization,'' Appl. Organometallic Chem., vol. 32, no. 3, p. e4186, Mar. 2018.

[6] V. Chaudhary, A. Deshbhratar, V. Kumar, and D. Paul, ''Time series-based LSTM model to predict air pollutant's concentration for prominent cities in India,'' in Proc. Int. Workshop Utility-Driven Mining (UDM), Aug. 2018, pp. 1–9.

[7] S. Mahajan, H.-M. Liu, T.-C. Tsai, and L.-J. Chen, ''Improving the accuracy and efficiency of PM2.5 forecast service using cluster-based hybrid neural network model,'' IEEE Access, vol. 6, pp. 19193–19204, 2018.

[8] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, ''An information framework for creating a smart city through Internet of Things,'' IEEE Internet Things J., vol. 1, no. 2, pp. 112–121, Apr. 2014.

[9] A. Agarwal and M. Sahu, ''Forecasting PM2.5 concentrations using statistical modeling for Bengaluru and Delhi regions,'' Environ. Monit. Assessment, vol. 195, p. 502, Mar. 2023.

[10] Dataset that we used was from a website called kaggle.com

[10] T. S. Rajput and N. Sharma, ''Multivariate regression analysis of air quality index for Hyderabad city: Forecasting model with hourly frequency,'' Int. J. Appl.

P46RW004

EMOTION TUNE RECOMMENDER USING MACHINE LEARNING

Deepthi Dharavath ,UG Student

Vasavi Annapuri ,UG Student

Sai Charan Byragani ,UG Student

Mr. Mastan Vali, Research Guide

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD

ABSTRACT

Human emotions and feelings are not constant; rather, they are the product of both internal and the external circumstances. Customers and music enthusiasts can appreciate music as a fantastic medium for personal expression. As technology advances, the number of musicians and music enthusiasts increases, creating rivalry in the choice of music. Although music recommendations have been around for a while, they were primarily developed after consumers' tastes were gradually learned. People's emotions are thoroughly researched and have a wide range of uses. Different systems are available based on artist, genre, etc. including playlists of music that are automatically created based on Organizing music files into playlists by hand is an additional method. Current issues consist of frequency estimation and comparable musical computations. A song's content (key and tempo) is used by the QBSH (Question Singing and Humming) algorithm to determine which song it is. However, this technique has the drawback of being time-consuming and not always satisfying the customer. Excellent findings were produced by the suggested approach, which also cleared the path for additional study in this field.

KEYWORDS

Haar Cascade, Convolution Neural Networks, Support vector machine, ResNet-50

1. Introduction

Music has a remarkable effect on the human brain, especially when it comes to emotion, as confirmed by recent studies. There is a direct relationship between musical tastes and emotional states, as research has shown that music is a powerful stimulant for arousal and mood. This study suggests a novel way to improve user experience through individualized music playlists by utilizing the power of facial expressions, a fundamental tool for assessing human emotions. With the use of advances in signal processing and facial recognition technology, the suggested model combines music selection with realtime monitoring of facial expressions, allowing playlists to be customized to the user's emotional landscape. Deep learning is used in the model to extract features, which allows it to go beyond conventional machine learning techniques and improve its ability to recognize emotions on its own. In addition to improving users' moods, the goal of this creative synthesis of facial expression analysis and music selection is to establish a mutually beneficial link between technology and emotional health.

The idea behind this innovative project is deep learning, which enables machines to recognize, understand, and use complicated patterns that are extracted from enormous amounts of data on their own. Deep learning enables more accurate interpretation and prediction of user emotions based on face signals by combining inputs from many models and continually improving features. The Soft Max layer is a key element of the suggested design that facilitates emotional analysis and provides information necessary for creating customized music playlists. Furthermore, by incorporating deep learning techniques, the emotion detection model is made flexible and scalable, which allows it to improve and change over time in terms of prediction. This paradigm shift highlights the trans formative potential of technology by bringing together facial expression analysis and music recommendation.

2. Literature Survey

In 2020, S. Matilda Florence and M. Uma published an essay titled "Exploring Emotions by User Faces and Recognizing Music," which described a novel system that detects user emotions from their facial expressions and adjusts music selections accordingly. Thanks to this creative method, the system can now adapt to each user's specific emotional demands and successfully reduce stress by offering a customized music experience that eliminates the need for tedious manual song selection. Made up of three main modules—emotional retrieval, sound restoration, and emotional filtering—

the system provides a smooth connection between user emotions and music playback, increasing user engagement and happiness.

Florence and Uma's technology has certain noteworthy drawbacks that need to be addressed despite its promising capabilities. Notably, incomplete assessments could arise from the system's dependence on facial expressions to identify emotions, since facial markings by themselves might not be sufficient to convey certain feelings. Moreover, practical difficulties arise from the need for well-lit areas and highresolution photographs, which may restrict the system's applicability in many contexts. More robust and flexible approaches are required to improve emotion detection accuracy and reliability, as the system's adaptability in real-world situations may be hampered by the reliance on handmade features for emotion classification

For applications like predicting fashion photographs, convolution neural networks (CNN's) are an invaluable tool because of their unmatched accuracy in identifying patterns and features in images. Fashion websites may increase user experience and engagement by using CNN s to anticipate the newest trends in apparel goods with amazing accuracy. This project involves loading the fashion mnist data set, dividing it into training, testing, and validation sets, and importing necessary libraries to build the CNN model. The data set is carefully examined with matplotlib to prepare it for training the CNN model, which draws knowledge from the data's diverse experiences to accurately forecast fashion photos. A classification map with a confusion matrix displaying metrics like accuracy, recall, and precision is then used to assess the model performance using test information. The findings are then displayed using matplotlib graphs F1 score. To avoid over fitting and improve accuracy, CNN model architecture is continuously improved, making-machine learning models indispensable instruments for spotting the newest online fashion trends. This increases customer pleasure and boosts company performance.

3. Background Work

Numerous disciplines, including computer wisdom, psychology, and neuroscience, have expressed interest in the use of technology to identify mortal feelings. One of the most accurate and licit ways to read people's feelings is through their facial expressions. Accordingly, a lot of exploration has been done lately on the subject of face expression recognition exercising machine literacy ways.

One well-liked machine literacy system for relating facial expressions is the Haar Cascade Classifier. This system, which was described in 2001 by Viola and Jones, is grounded on an object identification strategy grounded on machine literacy. The Haar Cascade Classifier uses a set of positive and negative training photos to educate a classifier on how to describe objects in new photos. Numerous examinations have been carried out to explore the operation of facial expression recognition in customized music suggestions. For illustration, a study by Yang et al.(2018) employed facial expression recognition to suggest music according to the emotion linked. According to the study, the suggested system performed better than the conventional music recommendation systems. In conclusion, a lot of exploration has been done lately on the operation of machine literacy algorithms for facial expression identification, similar to the Haar Cascade Classifier. The suggested emotion-grounded music recommendation system has implicit operation in the fields of music remedy and substantiated music recommendation. It detects feelings in facial expressions and makes music recommendations grounded on those passions.

4. Methodology

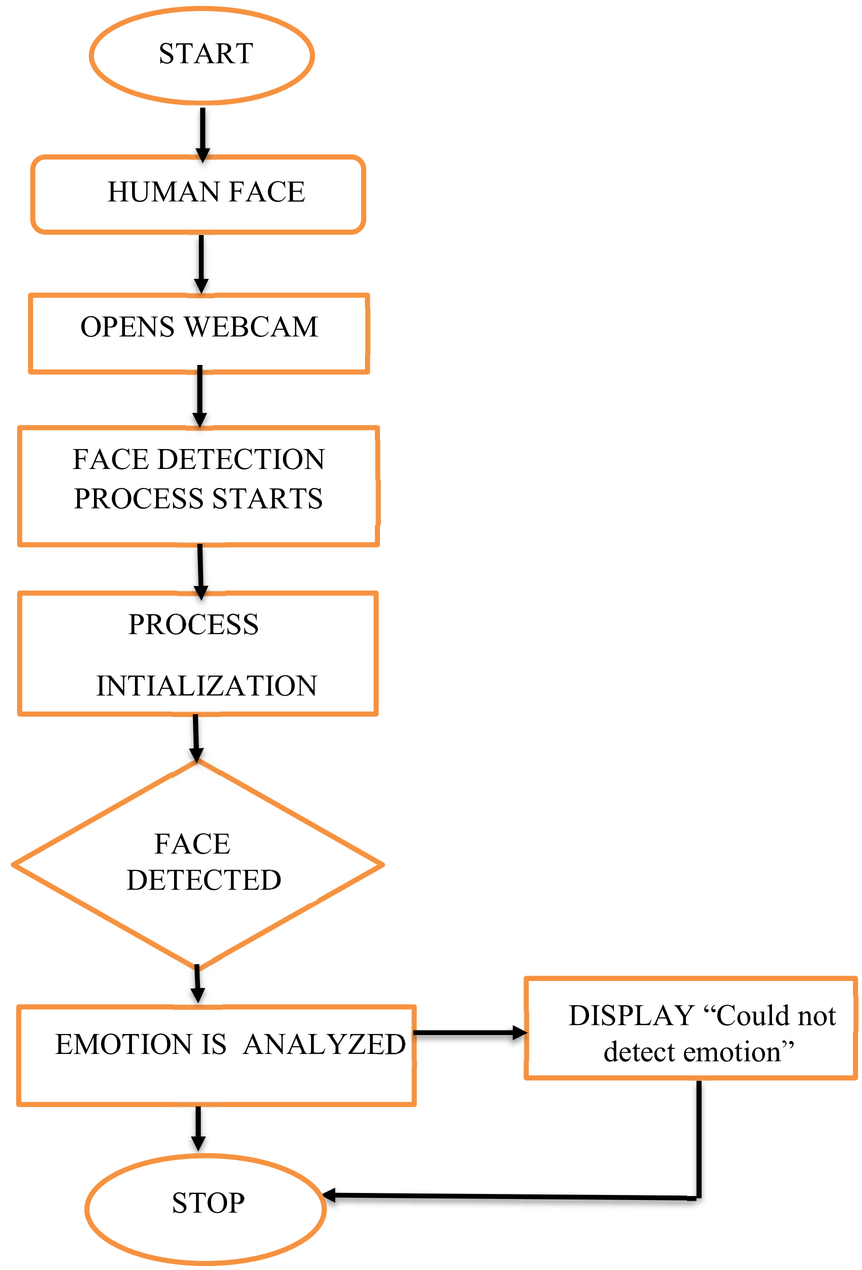
Image capturing: The user's face is detected by facial recognition technology, and then a convolution neural network (CNN) specifically designed to study user behavior is designed and put into operation. Furthermore, the system makes use of sensitivity detection techniques, in which the user's sensitivity level is ascertained by analyzing features taken from their image and presenting pertinent readings based on the results of this sensitivity assessment.

Pre - processing: Filtering and normalization techniques are used in pre processing processes for image analysis to improve picture similarity and fix problems like rotation. To reduce potential distortions and enable proper analysis, normalizing techniques are utilized in situations when photos display back lighting, which causes a whitish look and blurring of facial features. This eliminates the white background and enhances facial clarity.

Segmentation: A crucial stage in image classification is segmentation, which is breaking up a picture into homogeneous areas instead of considering it as a single entity based on characteristics like texture, density, and edges. By recognizing different parts in an image and examining them separately, this method improves the system's capacity to recognize and categorize emotions with higher accuracy.

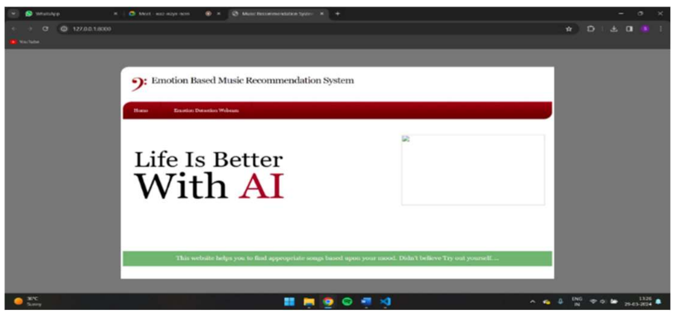
Feature extraction: Feature extraction is the process of converting unprocessed data into numerical representations while preserving important information. This allows for more efficient processing than if raw data were used directly. Convolution neural networks (CNN's) are effective tools for feature extraction in image analysis because they can accurately classify similar features in other photos by extracting complex patterns and features from input images. In the end, the CNN classifies and interprets the characteristics in the input image by using the signals that were extracted.

Emotion classification: Emotions have a profound impact on people's perceptions, interactions, and decision-making processes. They also shape how people react to both intentional and involuntary acts, which makes emotion classification essential. In short, our actions and choices are mostly determined by the feelings we are experiencing at the time, underscoring the significant effect of emotional states on day-to-day functioning. This is why it's common advice to avoid making decisions when feeling very emotional, like anger, highlighting how important it is to recognize and control emotions in a variety of situations.

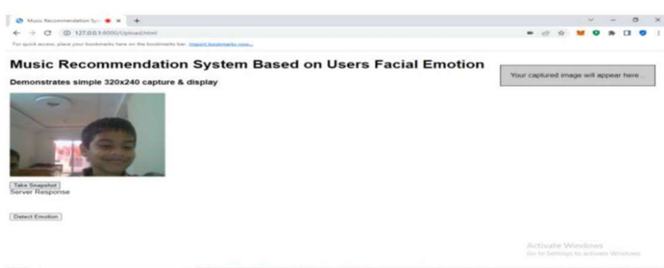


5. Results and Discussion

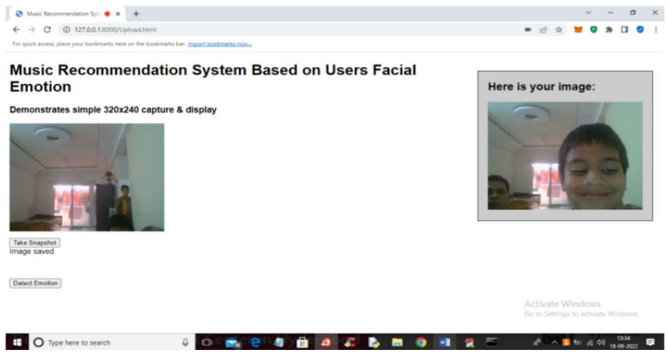
This section talks about the system's results, which use Support Vector Machine (SVM) and Convolution Neural Networks (CNN) to classify images.To recognize emotions, the system takes pictures and uploads them. A suggestion for a song goes along with the feeling once it has been determined. It allows users to listen to the suggested song, allowing image analysis and emotion-based music selection to work together seamlessly. By providing tailored music selections based on the user emotional state, this method improves the user experience. Accurate emotion detection is made possible by the CNN and SVM algorithms' efficient image classification. Because it can dynamically recommend music based on emotions observed, the system is useful and relevant in real-world scenarios. Overall, the findings demonstrate how well the system works to deliver customized music.



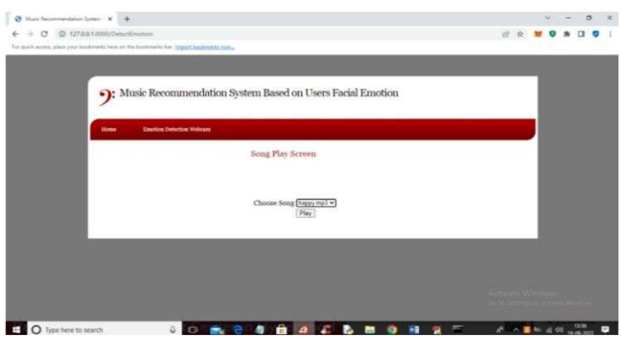
- In the above screen click on the 'Emotion Detection Webcam' link to get below Webcam page



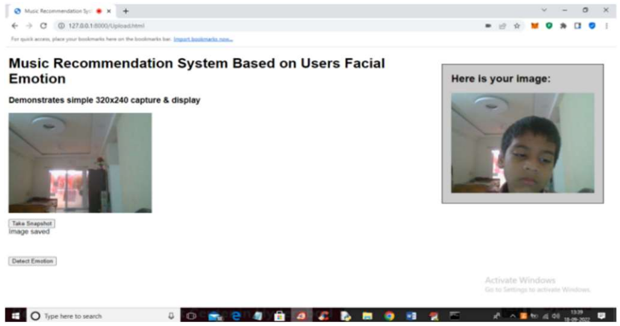
- Take a photo of your face so that you can identify facial expressions. Click the "Take Snapshot" button to capture the face as seen in the screen below, which is displayed on the above screen using the camera.



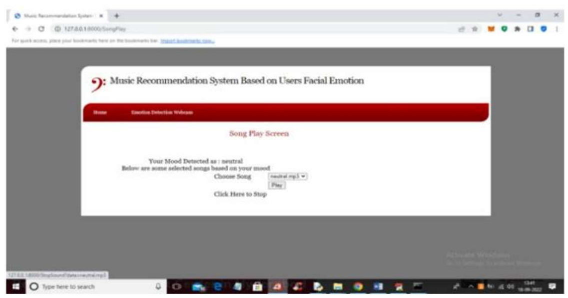
- To recognize facial expressions, take a picture of the face. After capturing the face on the above screen, use the "Detect Emotion" button to forecast the emotion and view the song list below.



- A recommended song above is playing based on the detected sentiment. The user can choose any song from the drop-down list and click the "Play" button to start it playing. He can even pause the music while it is playing. The "Happy" emotion was identified on the screen above, and we were presented with a choice of joyful songs. Try again with a different feeling now.



- We have captured another emotional face on the above screen. To access the page below, click the "Detect Emotion" button.



- The song will begin to play on the screen above; to halt it, click the "Click Here to Stop" link.

6. Conclusions

The general model is used to depict the music that is based on the user emotions because music has the power to convey user feelings. Emotions in humans are crucial for communicating innermost thoughts. Identifying shifts in the user's mood and playing music based on their preferences by looking up different genres of music are the system's primary goals. When classifying emotions, the system makes use of a CNN algorithm that analyses variations in the size, shape, and movement of the mouth, eyes, and eyebrows. The six emotions that can be encountered when making a playlist are Sadness, Joy, Anger, Fear, Disgust, and Surprise. During the playlist's creation. Mainly because it can identify the important features in an image automatically and without human assistance, the CNN algorithm is favored over SVM. Additionally, a better prediction accuracy of CNN was found compared to DVM. Good results were obtained using the proposed system. It's difficult to get 100% accuracy because every individual thinks differently and is influenced by both internal and external events in their surroundings.

7. References

[1] "Face Discovery and Facial Expression Recognition System," Anagha S. Dhavalikar and Dr. R K. Kulkarni, Institute of Electrical and Electronics Engineers (IEEE 2014).

[2] Ms.Graca Bressan, Renata Lopes Rosa, and Demosthenes Zegarra Rodriguez "Music Recommendation System rested on stoner’s Sentiments pulled from Social Networks" 2015 IEEE International Conference on Consumer Electronics (ICCE).

[3] Yajie Hu and Mitsunori Ogihara "A MUSIC RECOMMENDATION SYSTEM based on UserBehavior: NEX TONE PLAYER" 12th International Society for Music Information Retrieval Conference (ISMIR 2011).

[4] Kyoungro Yoon, Senior Member, IEEE, Jonghyun Lee, and Min-UK Kim "Emotion-driven Low position Features for Music Recommendation System" IEEE Transactions on Consumer Electronics 2012.

[5] Fang- Fei Kuo1, Meng-Fen Chiang2, Man-Kwan Shan2, and Suh-Yin Lee "Emotion rested Music Recommendation by Association Discovery from Film Music" (2005).

[6] Yu-Hao Chin, Yi-Hsuan Yang, Member, IEEE, Chiang Wang, and Jia-Ching Wang, Senior Member,IEEE "Predicting the Probability viscosity Function of Music Emotion using Emotion Space Mapping" IEEE 2018.

[7] Hye-Rin Kim, Yeong-seok Kim, Seon Joo Kim, and In-Kwon Le "Creating Emotional Machines that Fit Image Heartstrings with Deep Neural Networks" IEEE Transactions on Multimedia 2018.

[8] Anukriti Dureha IJCA 2014 "An Accurate Algorithm for Creating a Music Playlist Based on Facial Expression".

[9] "Emotional Recognition from Facial Expressions Analysis using Bezier wind befitting," Yong HwanLee, Woori Han, and Youngseop Kim, 16th international conference on network-rested information systems, 2013.

[10] Arto Lehtinemi and Jukka Holm, "Animated Mood Pictures in Music Recommendation," 16th International Conference on Information Visualization, 2012.

P49RW004

Evaluating the Learning Patterns and Analyzing Efficiency in Students using Machine Learning Algorithms

1 MD Altaf, 2Peddireddy Shiva Reddy,3 Polasi Praneeth,4 G Thirupathi

[1,2,3] Graduate Student-CSE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

[4] Asst.Professor, Department of CSE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

[1] mohd.althaf1437@gmail.com, [2]shivareddy761005@gmail.com,
[3] praneethpolasi@gmail.com,[4] thirupathi.g@sreenidhi.edu.in

Abstract—This project explores the critical role that information literacy plays in the learning outcomes and behaviors of college students. Using a range of supervised classification algorithms, predictive models were created by looking at different learning behaviors and emphasizing information literacy. This study expands the analysis by including new methodologies, building upon the original paper's successful use of Decision Trees, KNN, Naive Bayes, Neural Networks, and Random Forest, which produced an outstanding 92.50% accuracy. The accuracy shot up to 100% when XGBoost and a Voting Classifier were added to the ensemble approach. This improvement represents the possibility for sophisticated techniques to improve the models' predictive power, providing insightful information on customized interventions to maximize information literacy instruction. The results highlight how important it is to comprehend and take use of a variety of learning behaviors in order to develop creative people who can learn new things their entire lives and adjust to changing social demands. This study adds to the growing body of knowledge about information literacy's critical role in postsecondary education and its implications for developing flexible, independent learners.

Index Terms: Machine learning, information literacy, learning behavior characteristics, learning effect, innovative talents.

INTRODUCTION

Overview

The rapid development of information technology has significantly impacted various sectors, making it crucial for college students to acquire competencies such as creativity, critical thinking, and information literacy. Information literacy is essential for fostering creative genius and ensuring the longterm development of future human resources. Educational institutions worldwide prioritize information literacy instruction due to its importance.

In recent years, online and hybrid teaching modalities and advancements in artificial intelligence technologies have led to the rise of specialized information literacy courses in colleges. However, there are still obstacles in college-level training, such as effective learning outcome prediction. Machine learning techniques can be used to optimize the learning process by acquiring insights into learners' progress and customizing interventions accordingly.

Learning prediction uses variables like learning achievement, goals, and abilities to predict learning experiences and results. Techniques such as regression analysis, neural networks, and Bayesian approaches are used to predict students' learning outcomes. The integration of machine learning and educational data mining technologies has emerged as a promising path toward developing data-driven prediction models.UNESCO's 2019 report on Artificial Intelligence in Education highlights the potential of integrating artificial intelligence and education for advancing quality and equity in educational institutions. Teachers can use data-driven insights to improve learning outcomes and provide individualized learning experiences for students using educational data mining and machine learning. This research investigates the relationship between learning behavior analysis, predictive modeling, and information literacy in higher education contexts, aiming to fill gaps and tackle issues in learning prediction methodologies and information literacy education.

Furthermore, as higher education continues to change in the digital age, information literacy's importance in preparing students for success in a variety of disciplines is becoming more widely acknowledged.

Objective

This research attempts to investigate the relationship between learning behavior analysis, predictive modeling, and information literacy in the context of higher education in light of these advancements. This study aims to fill in the gaps and tackle the issues in the field of learning prediction methodologies and information literacy education by looking at the current state of the art. It also provides insights into how machine learning techniques might be used to improve information literacy instruction and maximize learning outcomes for college students.

LITERATURE SURVEY

Related Work

The construction of an information literacy education model for Chinese college students is the focus of Z. Chinghai's work [1], which integrates creativity and critical thinking. This model emphasizes how crucial it is to develop students' capacity for critical analysis and innovative use of information in order to raise their level of information literacy as a whole. In a similar vein, S. Hui [2] addresses information literacy teaching tactics designed for university students, emphasizing the necessity of a comprehensive strategy that includes both theoretical understanding and practical competence.

Using information from literature indexed in the CNKI database from 2000 to 2021, G. Yang, B. Wen, and W. Lin [3] propose a bibliometric analysis of research trends and hotspots in college students' information literacy. Their research highlights regions that are ready for more examination by identifying major themes, hot subjects, and research trajectories in the discipline.

L. Yu, D. Wu, H. H. Yang, and S. Zhu [4] investigate college students' choices for smart classrooms and information literacy. They investigate the relationship between students' information literacy skills and their preferences for technology-enhanced learning settings through empirical research, providing insightful data for instructional design and pedagogical practice.

Y. Ying [5] uses big data analytics to examine information literacy among college students. The study finds patterns, trends, and connections pertaining to students' informationseeking activities and information processing abilities by examining large-scale datasets. The aforementioned study enhances our comprehension of the complex characteristics of information literacy and its consequences for pedagogical approaches.

The promotion strategies and influencing factors related to information literacy among college students are examined by X. Ouyang, Y. Xiao, and J. Zhong [6]. They identify important factors influencing students' information literacy levels through a qualitative investigation, and they suggest focused interventions to improve information literacy instruction in higher education settings.

The information technology literacy of newly enrolled female college students in Japan is evaluated by T. Nishikawa and G. Izuta [7]. Their study examines potential factors impacting students' technological competencies as well as their competency with a variety of information technologies. The results support initiatives to close the digital divide and increase college students' digital literacy.

Based on multifarious data, Y. Sun, Z. Tan, Z. Li, and S. Long [8] use machine learning approaches to forecast and analyze the performance of college students. Through the utilization of many data sources, such as extracurricular activities, academic records, and demographic data, the research creates predictive models that are able to anticipate the academic outcomes of students. This study highlights how data-driven strategies can improve student support programs and educational decision-making.

The research review concludes by highlighting the multifaceted character of information literacy instruction for college students, which includes predictive modeling of learning outcomes, technological competence, critical thinking, and creativity. This survey provides an extensive summary of current research trends, opportunities, and difficulties in the subject by combining insights from various studies. The future course of information literacy education is expected to be shaped by multidisciplinary collaboration, novel pedagogical approaches, and technology breakthroughs, which will enable college students to prosper in an increasingly complicated and linked world.

METHODOLOGY

A) Proposed Work

Utilizing pre-analyzed data on learning behavior and its relationships to learning outcomes, the proposed study seeks to create predictive models by applying techniques from Decision Tree[9], K-Nearest Neighbor (KNN)[10], Naive Bayes[11], Neural Network (NN), and Random Forest. The goal of this research is to shed light on the complex relationship that exists between academic achievement and learning behavior patterns among students.

Preprocessing the data is part of the methodology to guarantee its accuracy and applicability. The predictive usefulness of characteristics such as levels of engagement, study habits, and involvement in educational activities will be closely examined. To assess each model's performance, the data will then be divided into training and testing sets.

The models' efficacy will be assessed by the application of criteria including F1 score, recall, accuracy, and precision. Furthermore, the models' interpretability will be given top priority in order to find practical insights for focused interventions. In the end, the goal of this project is to provide a systematic framework for using machine learning to forecast learning outcomes based on students' behavior in the classroom. This will improve learning outcomes and advance personalized learning strategies in higher education.

B) System Architecture

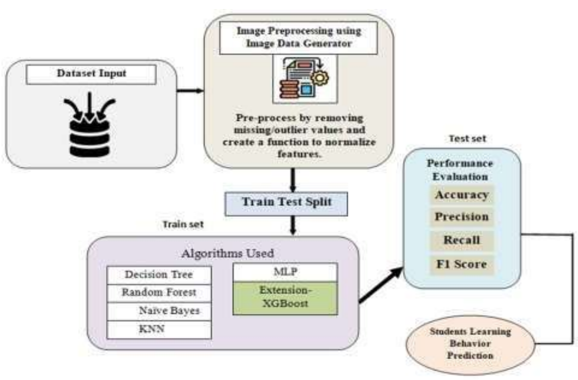


Fig 1 Proposed Architecture

The system design includes a number of interrelated parts that work together to make it easier to forecast how children will learn. First, the architecture takes in a dataset that includes pertinent data on how students learn, including things like their study habits, engagement levels, and involvement in class activities. To improve its quality and get it ready for analysis, the dataset is next subjected to image processing using Image Data Generator methods. To guarantee the reliable assessment of prediction models, the dataset is split into training and testing sets using a Train-Test-Split method after preprocessing.

A number of machine learning algorithms, such as Decision Tree[9], Naive Bayes[11], K-Nearest Neighbor (KNN)[10], Random Forest[12], Multi-Layer Perceptron (MLP), and XGBoost, are used at the heart of the design. By analyzing the preprocessed data, these algorithms are able to precisely forecast the learning behavior of pupils. Metrics for performance evaluation, including Accuracy, Precision, Recall, and F1 Score, are used to evaluate how well each algorithm captures the subtleties of students' learning styles.

In the end, the system architecture uses performance evaluation criteria and cutting-edge machine learning approaches to forecast students' learning behavior. Through the smooth integration of these elements, the architecture offers a thorough framework for comprehending and forecasting students' learning habits, enabling focused interventions and improving academic results.

C) Data Set

The Student Learning Behavior dataset is made up of an extensive range of characteristics that represent many facets of students' performance and involvement in the classroom. It contains data on the study habits, attendance histories, extracurricular activity involvement, test results, and demographics of the pupils. The dataset might also include information about how students use educational resources like online learning environments and library resources. The dataset allows for in-depth investigation and analysis of the variables impacting students' learning behaviors and academic outcomes because to its wealth of information. For scholars and educators looking to deepen understanding and encourage students' academic journeys, it is an invaluable resource.

	IPC1	IPC2	IPC3	IAC1	IAC2	LLC1	LLC2	ISK1	ISK2	ISK3	ISK4	IAS1	IT1	IT2	IT3	IB1	IE1	IE2	ILR1	label
0	69	63	78	87	94	94	87	84	61	4	4	7.9	A	1.0	0.0	0.0	0.0	0.0	no	excellent
1	78	62	73	60	71	70	73	84	91	7	2	5.4	B	2.0	0.0	0.0	0.0	0.0	no	medium
2	71	86	91	87	61	81	72	72	94	1	1	5.2	B	7.0	0.0	0.0	0.0	0.0	no	excellent
3	76	87	60	84	89	73	62	88	69	1	2	8.5	C	10.0	0.0	0.0	0.0	0.0	yes	excellent
4	92	62	90	67	71	89	73	71	73	5	6	8.8	C	6.0	0.0	0.0	0.0	0.0	no	excellent
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1008	88	85	68	84	88	66	86	76	82	2	2	7.6	A	1.0	0.0	0.0	0.0	0.0	no	excellent
1009	76	63	92	74	76	81	76	87	81	8	7	7.4	C	7.0	0.0	0.0	0.0	0.0	yes	excellent
1010	74	94	94	82	64	92	84	67	80	4	6	7.7	C	5.0	0.0	0.0	0.0	0.0	no	poor
1011	60	84	84	70	80	78	64	83	60	8	6	7.6	D	8.0	0.0	0.0	0.0	0.0	yes	excellent
1012	91	61	83	80	88	62	88	76	86	9	1	7.4	D	5.0	0.0	0.0	0.0	0.0	no	excellent

Fig 2 Dataset

D) Data Processing

Data Loading with Pandas Dataframe: The process of processing data begins with loading the dataset into a pandas data frame, which is a vital tool that is well-known for its effectiveness in managing structured data. By utilizing the features of the dataframe, the dataset's contents are arranged into a tabular structure for easy access and manipulation during the stages of processing that follow.

Column Dropping: In an effort to refine the data, unnecessary or duplicate columns are carefully found and removed from the data frame. Column dropping is a selected method that helps to simplify the dataset by removing extraneous information and simplifying the computation. Column dropping simplifies the dataset by keeping only the most pertinent features, guaranteeing that next studies concentrate on the most important variables.

Normalization of Training Data: The training data is normalized to promote fair comparisons and lessen the impact of different feature scales. The numerical feature values are standardized by this transformative process, which usually rescales them to a common range like [0, 1] or [-1, 1]. Normalization encourages fairness in model training and evaluation by standardizing feature magnitudes, making it easier to make accurate and dependable predictions across a variety of datasets.

E) Visualization

Data visualization is made into an art form by combining the potent capabilities of the Seaborn and Matplotlib tools. Built on top of Matplotlib, Seaborn provides a simple-to-use interface for writing little to no code while producing visually striking charts. A broad range of high-level functions are available in Seaborn for dataset exploration and comprehension, ranging from basic histograms and scatter plots to complex heatmaps and violin plots. Matplotlib, on the other hand, provides more precise control over customizing plots, enabling the production of visualizations suitable for publishing. Seaborn and Matplotlib work together to enable analysts and data scientists to effectively communicate insights through eye-catching and educational images.

F) Label Encoding and Feature Selections

Label encoding converts categorical variables into a numerical format that makes them easier to understand by machine learning algorithms. Through this method, every category inside a feature is given a distinct numerical label.

Finding and keeping characteristics that have significant linear correlations with the target variable is a key component of feature selection based on high correlation values. Highly associated features are found and chosen for the prediction model by calculating correlation coefficients between the features and the target variable. By concentrating on the most significant traits and eliminating superfluous or unnecessary ones, this selective method maximizes interpretability and forecast accuracy while improving model efficiency.

G) Training and Testing

In order to ensure that the performance of the machine learning model can be precisely evaluated on unknown data, it is imperative that the data be split into training and testing subsets. The supplied dataset is divided into two separate subsets for this process: the training set and the testing set. The model is trained on the patterns and relationships found in the data using the training set, which usually consists of a higher percentage of the data. The testing set, on the other hand, is a smaller subset of the data that is used to assess the performance of the trained model. The testing set acts as an impartial gauge of the model's capacity for generalization by excluding some of the data during training, giving information about how well the model performs when applied to fresh, untested data.

To ensure that the training and testing subsets of the data are representative of the entire dataset, the splitting of the data into these subsets is usually done at random. Allocating a specific percentage of the data, say 70–80%, to the training set and the remaining amount to the testing set is a common approach. This guarantees a balance between maintaining a suitable evaluation dataset and offering enough data for model training.

Furthermore, methods like cross-validation could be used to evaluate model performance even more and lessen possible biases brought about by the data splitting procedure. Generally, robust model building and evaluation in machine learning applications depend on the meticulous division of data into training and testing subsets.

H) Algorithms Used

Basic machine learning algorithms like Random Forest, Decision Tree, Naive Bayes, K-Nearest Neighbors, and MultiLayer Perceptron have a wide range of uses in many fields.

Random Forest: Random Forest is an ensemble learning technique that builds many decision trees during training and produces the mean prediction (regression) or mode of the classes (classification) for each individual tree. It is resistant to overfitting and performs well with big, highly dimensional datasets.

Decision Tree: Decision Tree is a straightforward yet effective technique that creates a tree-like structure by iteratively dividing the dataset into subsets according to the most important attribute. [9] Because of its great interpretability and intuitiveness, it can be used to clarify the decision-making process and comprehend the significance of features.

Naive Bayes: With an assumption of predictor independence, Naive Bayes is a probabilistic classifier based on the Bayes theorem. It frequently works effectively in text categorization and other areas, especially when working with highdimensional data, despite [11]its simplicity and the "naive" assumption.

K-Nearest Neighbors (KNN): KNN is an instance-based, non-parametric learning method that groups new data points in the feature space according to how close they are to the majority class of their K nearest neighbors [10]. It is simple to use and adaptable, especially with smaller datasets.

Multi-Layer Perceptron: An artificial neural network called an MLP is made up of several layers of nodes, or neurons, coupled to one another at each layer. MLPs are frequently employed for tasks like pattern recognition, regression, and classification because they can understand intricate correlations in data.

These methods, each with unique strengths and limitations based on the particular issue domain and dataset characteristics, are fundamental components of a data scientist's toolkit and the basis of many machine learning applications.

EXPERIMENTAL RESULTS

Precision: Precision measures the percentage of correctly categorized samples or instances among the positive samples. Consequently, the following is the formula to determine the precision:

True positives/(True positives + False positives) = TP/(TP + FP) is the formula for precision.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

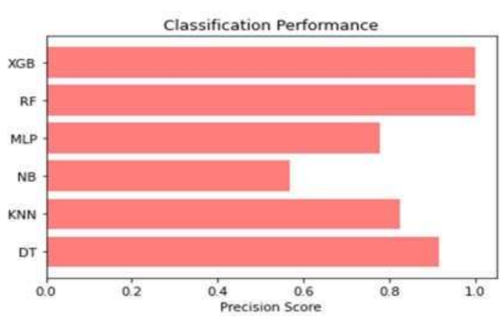


Fig 3 Precision Comparison Graph

Recall: In machine learning, recall is a metric that assesses a model's capacity to locate all pertinent instances of a given class. It is a measure of how well a model captures examples of a particular class: the ratio of correctly predicted positive observations to the total number of real positives.

$$Recall = \frac{TP}{TP + FN}$$

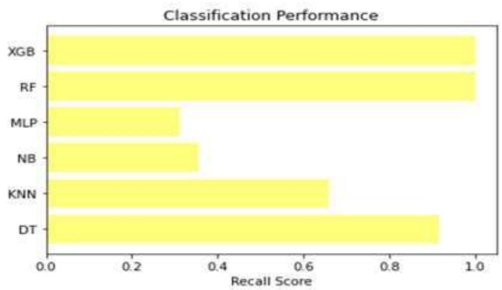


Fig 4 Recall Comparison Graph

F1-Score: An evaluation statistic for machine learning called the F1 score quantifies the accuracy of a model. It integrates a model's precision and recall ratings. The number of times a

model correctly predicted throughout the whole dataset is calculated by the accuracy metric.

$$\textbf{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\textbf{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

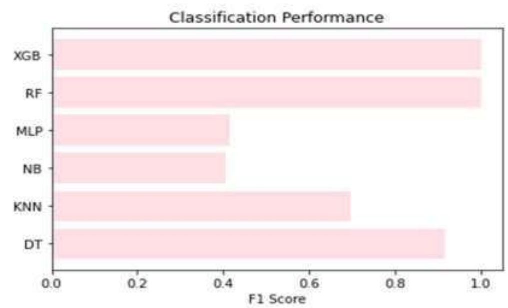
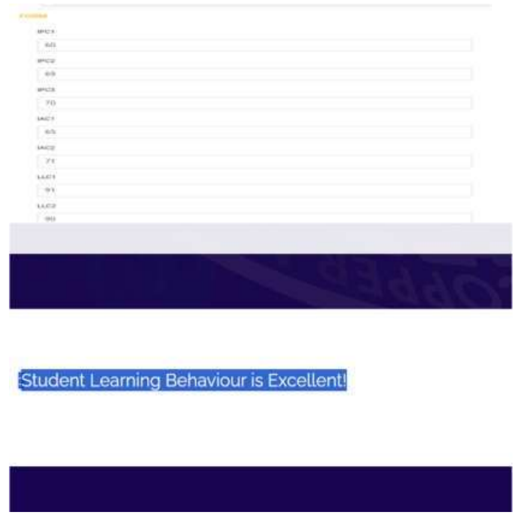


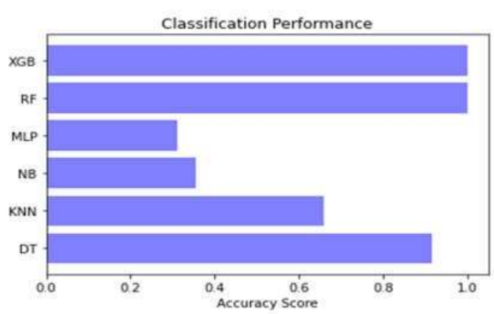
Fig 13



Upload Input Data

Fig 5 F1 Score Comparison Graph

Accuracy: A test's accuracy is determined by how well it can distinguish between patient and healthy cases. We should compute the percentage of true positive and true negative in each analyzed case to assess the accuracy of a test. This can be expressed mathematically as follows:



Accuracy = TP + TN / (TP + TN + FP + FN)

Fig 6 Accuracy Comparison Graph

	ML Model	Accuracy	Precision	f1_score	Recall
0	DT	0.916	0.918	0.916	0.916
1	KNN	0.658	0.825	0.696	0.658
2	NB	0.355	0.569	0.404	0.355
3	MLP	0.310	0.778	0.416	0.310
4	**RF**	**1.000**	**1.000**	**1.000**	**1.000**
5	**Extension-XGB**	**1.000**	**1.000**	**1.000**	**1.000**

Fig 7 Performance Evaluation Table

Fig 15 Upload Input Data

CONCLUSION

In conclusion, today's information-rich environment, information literacy is essential for success. It goes beyond academic performance to become a lifetime learning tool and a means of navigating the intricacies of contemporary life. Teachers can customize their teaching approaches to meet the needs of each individual student and create a more inclusive and productive learning environment by understanding the complex interactions that exist between student learning behaviors and outcomes. Using predictive models like

Random Forest, Decision Tree, KNN, Naive Bayes, Neural

Network, and Random Forest—plus the potent ExtensionXGBoost—improves teachers' capacity to recognize and respond to differences in students' information literacy competency levels. By enabling educators and administrators to convert these insights into workable methods, the practical integration of XGBoost within Flask promotes informed decision-making and leads to observable gains in educational results.

FUTURE SCOPE

Going forward, there is a great deal of promise in combining cutting-edge machine learning methods with teaching approaches. Predictive models have a great deal of room for improvement as technology develops in order to better comprehend and assist students' learning journeys. Furthermore, current research and development initiatives in the field of educational data analytics present chances to investigate novel approaches and broaden the application of predictive modeling to tackle newly developing issues in education. Through adoption of these developments and promoting cooperation among scholars, instructors, and tech creators, we can keep utilizing data-driven insights to influence the course of education and enable students all over the world.

REFERENCES

[1] Z. Changhai, ‘‘Research on the information literacy education model of Chinese college students based on critical thinking and creativity,’’ J. China Library, vol. 4, no. 15, pp.
[2] 15–16, Aug. 2016, doi: 10.13530/j.cnki.jlis.164008.

[3] S. Hui, ''Information literacy education for college students,'' Educ. Theory Pract., vol. 30, no. 10, pp. 38–39, Oct. 2008.

[4] G. Yang, B. Wen, and W. Lin, ''Research status, hot spots and enlightenment of college students' information literacy: Based on bibliometric analysis of CNKI from 2000 to 2021,'' in Proc. 4th World Symp. Softw. Eng., Sep. 2022, pp. 161– 166, doi: 10.1145/3568364.3568389.

[5] L. Yu, D. Wu, H. H. Yang, and S. Zhu, ''Smart classroom preferences and information literacy among college students,'' Australas. J. Educ. Technol., vol. 38, no. 2, pp. 144–163, Feb. 2022, doi: 10.14742/ajet.7081.

[6] Y. Ying, ''Research on college students' information literacy based on big data,'' Cluster Comput., vol. 22, no. S2, pp. 3463–3470, Mar. 2019, doi: 10.1007/s10586-018-2193-0.

[7] X. Ouyang, Y. Xiao, and J. Zhong, ''Research on the influencing factors and the promotion measures of college students' information literacy,'' in Proc. 8th Int. Conf. Inf. Technol. Med. Educ. (ITME), Dec. 2016, vol. 12, no. 12, pp. 728–732, doi: 10.1109/ITME.2016.0169.

[8] T. Nishikawa and G. Izuta, ''The information technology literacy level of newly enrolled female college students in Japan,'' Hum. Social Sci. Rev., vol. 7, no. 1, pp. 1–10, Mar. 2019, doi: 10.18510/hssr.2019.711.

[9] Y. Sun, Z. Tan, Z. Li, and S. Long, ''Predicting and analyzing college students' performance based on multifaceted data using machine learning,'' in Proc. 4th Int. Conf. Adv. Comput. Technol.,Inf.Sci.Commun.(CTISC),Apr.2022,pp.1–6,doi: 10.1109/CTISC54888.2022.9849815.

[10]A. Weipeng and S. Jiaze, ''Improvement and analysis of decision tree C4.5 algorithm,'' Comput. Eng. Appl., vol. 55, no. 12, pp. 169–173, Jun. 2019, doi: 10.3778/j.issn.10028331.1805-0482.

[11]Z. Yu, H. Chen, J. Liu, J. You, H. Leung, and G. Han, ''Hybrid k-nearest neighbor classifier,'' IEEE Trans. Cybern., vol. 46, no. 6, pp. 1263–1275, Jun. 2015, doi: 10.1109/TCYB.2015.2443857.

[12]M. S. Roobini, K. Babu, J. Joseph, and G. Ieshwarya, ''Predicting stock price using data science technique,'' in Proc. 2nd Int. Conf. Artif. Intell. Smart Energy (ICAIS), no. 2, Feb. 2022, pp. 1013–1020, doi: 10.1109/ICAIS53314.2022.9742772.

P51RW023

Learner Advocacy and Discontent Resolution System

Preethi , Varsha Reddy Naini , Tharasa Guntipall , Mr. P. Durga Prasad.

guntipallytharasa2002@gmail.com, podhuturipreethi@gmail. , varsha.naini@gmail.com ,
durgaprasadp@sreenidhi

Abstract—The Student Grievance System designed using Django framework, aims at fostering a conducive learning environment by addressing and resolving student concerns effectively. The system incorporates key features such as user account management, anonymous complaint submission, targeted complaint addressing, automated reminders for unresolved issues, and a transparent feedback mechanism. The system begins by allowing the creation of admin and student accounts, each serving distinct roles within the platform. Students can lodge complaints anonymously, promoting candid communication without fear of reprisal. Complaints are directed to specific authorities, ensuring a streamlined and targeted resolution process. To ensure timely grievance resolution, our system utilizes automated reminders. When complaints persist beyond expected time frames, reminder emails are dispatched to the relevant authorities, promoting swift attention and resolution. This enables authorities to update complaint statuses and offer feedback, enhancing transparency in the resolution process. Furthermore, to enhance user experience, the system filters and hides complaints addressed by others and those .edu.in directed to different authorities. This ensures that students only see pertinent information, reducing clutter and facilitating ease of navigation. In conclusion, the student grievance system developed with Django represents a significant advancement in addressing student concerns within educational institutions. Its comprehensive features contribute to an efficient, transparent, and user-friendly platform, ultimately enhancing student satisfaction and promoting a positive learning environment.

INTRODUCTION

Educational institutions must provide an environment conducive to learning and growth. This work requires the development of rules that encourage open conversation, issue resolution, and fairness and transparency. One such crucial tool is the Student Grievance System (SGS), which is critical for establishing a positive relationship between students and administrative authorities. The landscape of higher education has shifted considerably in recent years, with a rise in student enrollments and a wider range of academic specialties. However, as a result of expansion, impediments and disagreements develop that must be addressed efficiently and equitably. Recognizing the need to address these concerns proactively, institutions have increasingly turned to technology-driven solutions, ushering in a new era of digitalized grievance management systems.

This research study goes into the development, design, and implementation of a strong SGS framework that is ready to meet the multifarious demands of current educational contexts. By utilizing knowledge from many academic publications, institutional case studies, and empirical investigation, this work seeks to offer a thorough synopsis of the elements, capabilities, and consequences of a successful SGS.

The suggested framework is built on a commitment to accessibility, accountability, and user empowerment. By employing cutting-edge information technology, the SGS aims to bridge the gap between students and administrative bodies, establishing a culture of trust, justice, and cooperation.

This paper demonstrates the SGS's revolutionary potential in transforming institutional dynamics and improving the overall student experience by meticulously examining its essential characteristics and processes.

Furthermore, this study tries to highlight the need for a comprehensive approach to grievance resolution, one that goes beyond procedural formalities and incorporates wider concepts of social fairness and institutional integrity. By emphasizing student perspectives and experiences, the SGS acts as a catalyst for institutional renewal, promoting positive change and cultivating a culture of continuous improvement.

In traversing the complex terrain of grievance management, this article advocates for a proactive approach that emphasizes prevention, mediation, and communication above reactionary tactics. The SGS fosters a climate of mutual respect and understanding, not just reducing disputes but also cultivating a sense of belonging and community within the educational ecosystem.

LITERATURE SURVEY

In contemporary organizational contexts, implementing a robust Student Grievance System (SGS) is paramount, permeating discussions within human resource management and organizational behavior. Rooted in theoretical frameworks of organizational justice, including distributive, procedural, and interactional justice, the SGS serves as a cornerstone for resolving conflicts and fostering a supportive learning environment. Delving into the intricacies of organizational justice, scholars have dissected its various dimensions, with distributive, procedural, and interactional justice emerging as pivotal constructs. Empirical investigations into organizational justice have yielded consistent findings, highlighting its profound impact on employee attitudes and behaviors. Similarly, empirical studies have consistently highlighted the positive correlation between perceived justice and organizational outcomes within educational institutions, underscoring the importance of fair grievance resolution processes in maintaining a conducive climate for learning and growth. Studies reveal a robust positive correlation between perceived justice and key organizational outcomes, including job satisfaction, organizational commitment, and overall performance.

However, implementing an effective grievance system is not without obstacles. Students may be hesitant to file complaints for fear of reprisal or discrimination, emphasizing the need to develop a culture of trust and support inside educational institutions. Anonymous reporting, protection against retribution, and support services for traumatized students all play a role in creating a climate in which students feel secure and empowered to express their concerns.

Managing complaints in a fair and timely manner presents another problem for institutions. Without effective resolution methods, students may lose faith in the system, leading to disappointment and an unwillingness to report future problems. To solve this issue, universities might use technology to create complaint portals or applications that streamline the process and improve student accessibility.

Along with theoretical developments, technological innovations have transformed the landscape of grievance management, providing new opportunities for improving communication, cooperation, and participation. Institutions can use technology to create grievance management systems beyond traditional limits, allowing for real-time participation, feedback, and resolution. From simple online platforms to complex data analytics tools, technology enables institutions to develop flexible and user-friendly systems that meet the different demands of students and stakeholders. Transparency and accessibility are essential components of an effective grievance system. Students should be provided with clear information on the grievance procedure and their rights, empowering them to navigate the system with confidence. Additionally, staff training in communication, dispute resolution, and empathy is imperative to ensure that grievances are handled professionally and impartially. Training programs should also address issues such as implicit bias and discrimination, fostering a culture of fairness and equity within the institution.

Continuous evaluation and monitoring of the grievance system are essential to its effectiveness and success. Feedback mechanisms, regular data collection on complaints, and benchmarking of best practices enable institutions to identify areas for improvement and make necessary adjustments. Additionally, designing the system to be adaptable to changing needs and allocating sufficient resources for its maintenance and growth are critical considerations in ensuring its long-term viability.

Real-world case studies are essential sources of insights and best practices, providing a window into the practical realities of adopting grievance management systems. These case studies highlight the complexity of developing, implementing, and monitoring grievance procedures at educational institutions. By evaluating successful implementations as well as obstacles faced, institutions may gain useful insights and techniques for enhancing their grievance management systems, eventually establishing a culture of openness, accountability, and continuous development.

Looking forward, the horizon of grievance management systems is illuminated by emerging trends and future directions, propelled by rapid technological advancements and evolving societal expectations. Artificial intelligence, blockchain technology, and social media platforms present exciting opportunities for boundaries of what is achievable in grievance management.

EXISTING SYSTEM

Within the existing grievance management structure used by many educational institutions, a traditional method for students to file complaints involves the use of actual complaint boxes strategically placed across campus grounds. In this technique, kids write their criticisms on paper and then drop them in designated receptacles. Complaint boxes are concrete evidence of an institution's dedication to allowing student feedback and addressing academic problems.

In this scenario, a student who is concerned about an issue files a written complaint, outlining the incident and stating desired results. Following this, the student places the written complaint in the designated complaint box. This action represents the beginning of the grievance resolution procedure from the student's perspective.

Despite its apparent simplicity, the classic complaint box system's effectiveness is sometimes hampered by several fundamental constraints. The system's lack of openness and responsiveness is the most obvious of these flaws. Students usually need a rapid reply or notification of receipt after submitting a complaint. As a result, students are frequently left in the dark about the status and advancement of their grievances, which causes dissatisfaction and disempowerment.

Furthermore, the lack of specified timetables for response and resolution heightens the sense of uncertainty. Students are forced to traverse an ambiguous period, unsure of when or if their issues will be handled. This uncertainty not only prolongs student misery but also calls into question the perceived legitimacy of the grievance-handling process.

Furthermore, typical complaint box systems lack methods for accountability and monitoring. With no official tracking or documenting of complaints, there is a noticeable lack of institutional responsibility in resolving student concerns. The obscurity of the procedure undermines the institution's commitment to justice and equity. As a result of these inherent limitations, students may lose faith in the grievance management system's effectiveness. This loss of trust can have serious consequences for the institution's reputation and capacity to create a supportive and inclusive learning environment.

In conclusion, while the old complaint box system provides a basic mechanism for collecting and processing student concerns, its inherent limitations highlight the need for modernization and development. The lack of openness, responsiveness, and accountability in this system needs a rethinking of grievance management processes in educational institutions, intending to implement more efficient and student-centered approaches to dispute resolution.

IV.PROPOSED SYSTEM AND ARCHITECTURE

In response to the existing inefficiencies of traditional grievance management systems within educational institutions, the proposed Online Grievance Management Portal aims to offer a modernized and integral solution. Utilizing contemporary technologies and tailored functionaliti -es, this portal seeks to streamline the process of lodging, tracking, and resolving student grievances. By addressing the limitations of conventional approaches, the proposed system aims to enhance transparency, accountability, and efficiency in handling student concerns within acad -emic settings.

A. Architecture

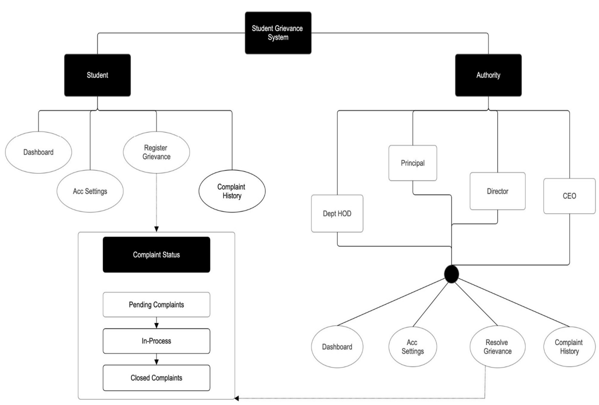


Fig.1. Architecture of Student Grievance System

1)Student Section

The procedural workflow for students within the Online Grievance Management Portal entails a series of structured tasks designed to ensure seamless navigation and interaction with the platform. Upon successful authentication with the correct user ID and password, the student gains access to the portal's interface, where they are greeted with a succinct and user-friendly dashboard. This dashboard serves as a centralized hub, offering four distinct tabs within the navigation bar: "Dashboard," "Lodge Complaint," "Update Profile," and "Logout."

The "Dashboard" section serves as an informative overview, providing students with a comprehensive snapshot of the current status of their grievances. Through this feature, students can readily visualize the number of complaints in various phases, categorized into four tabs: "All Complaints," "Pending Complaints," "In-Process Complaints," and "Closed Complaints." Each tab displays the respective number of complaints within its phase, facilitating informed decision-making and prioritization of follow-up actions.
Within each tab, students can access detailed listings of complaints corresponding to the selected phase. These listings include essential details such as the complaint ID, subject, timestamp of submission, current status of the complaint, the responsible authority overseeing its resolution, and any remarks posted by the respective authority. Additionally, a search bar is provided to facilitate efficient retrieval of specific complaints based on criteria such as the complaint ID, subject, or relevant keywords. This search functionality enhances user experience by enabling swift access to pertinent information within the complaint repository.

The "Lodge Complaint" option empowers students to initiate the grievance resolution by submitting a formal complaint. This interactive feature prompts students to input essential details, including a concise subject line, the designated authority to whom the complaint should be addressed, and a comprehensive description outlining the nature of the grievance. Upon submission, the system automatically generates a unique complaint ID, serving as a reference identifier for tracking purposes. Additionally, an email notification is promptly dispatched to both the student and the designated authority, confirming the successful submission of the complaint and initiating the requisite administrative procedures.

The "Update Profile" functionality allows students to maintain accurate and up-to-date personal information within the portal. Through this option, students can modify their profile details as needed, ensuring the integrity and relevance of their user accounts.

Finally, the "Logout" feature provides students with a convenient means to securely exit the application, thus safeguarding the confidentiality and integrity of their account information.

2)Authority Section

The procedural process for Authorities inside the Online Grievance Management Portal, similar to the Students' area, consists of a set of defined activities meant to guarantee seamless navigation and engagement with the platform. There are four kinds of authorities: heads of departments, principals, directors, and CEOs. These Authorities are competent staff members in charge of running the system and addressing students' inquiries. After successfully authenticating with the right user ID and password, the administrator receives access to the portal's interface, where they are met with a concise and user-friendly dashboard. This dashboard functions as a concentrated center, with three unique options in the navigation bar: "Dashboard," "Update Profile," and "Logout."

The "Dashboard" function is an instructional summary that gives relevant authorities a thorough view of the present condition of students' complaints reported to them. This feature allows authorities to easily view the number of complaints in various phases, which are divided into four tabs: "All Complaints," "Pending Complaints,"

"In-Process Complaints," and "Closed Complaints." Each tab illustrates the number of complaints within its particular phase, allowing for more informed decision-making and prioritization of the next steps. Authorities can get comprehensive lists of complaints related to the specified phase through each tab. These listings provide critical information such as the complaint ID, subject, date of filing, the relevant authority managing its resolution, and the ability to alter the current status of the complaint and add any remark/note for the Student.

Additionally, a search bar is provided to facilitate efficient retrieval of specific complaints based on criteria such as the complaint ID, subject, or relevant keywords. This search functionality enhances user experience by enabling swift access to pertinent information within the complaint repository.

The "Complaint History" section provides complete information about their previous encounters with the grievance management system. Respective Authorities can view a chronological list of all previously lodged complaints, each with pertinent details such as the complaint ID, timestamp of submission, current status of the complaint, the responsible authority overseeing its resolution, and any remarks posted by the authority. This tool allows the Authority to track the status of their concerns over time, promoting openness and accountability throughout the resolution process. The "Update Profile" capability helps users to keep their personal information correct and up to date within the portal. This feature allows authorities to change their profile information and password as needed, guaranteeing the integrity and relevance of their user accounts.

Finally, the "Logout" option allows Authorities to easily depart the program while maintaining the security and integrity of their account information.

B. Methodology

The development of the Student Grievance System (SGS) is primarily based on the Django framework, a robust web development framework for building efficient and scalable applications. Django's models option is utilized to design and implement the database structure, creating separate databases for managing complaints and user information. Two distinct user categories are identified: students and authorities, each having their respective user profiles within the system.

The built-in forms feature provided by Django is employed to facilitate seamless updates to user profiles as and when required. This ensures that user information remains accurate and up-to-date throughout their interaction with the system.

To automate the process of communication, Celery, a distributed task queue, is integrated into the system architecture. Celery is utilized to handle tasks such as sending emails to both students and authorities upon lodging a complaint, updating the status of complaints, and sending reminder emails to authorities in case of pending complaints. This automation streamlines the communication process, ensuring timely updates and notifications to all relevant parties.

Structured Query Language (SQL) is strategically employed at various points within the system to facilitate efficient data retrieval and manipulation. SQL queries are utilized for tasks such as user authentication, displaying complaints lodged by specific students or directed to particular authorities, retrieving user information for account management, and powering the search utility present in the dashboard. These SQL queries optimize database interactions, enhancing the overall performance and responsiveness of the system.

On the front end, the system is developed using HTML for structuring the content and defining the layout of web pages. CSS and Bootstrap classes are applied to style the interface, ensuring a visually appealing and cohesive user experience. Additionally, JavaScript is utilized to provide functionality to different aspects of the system, including navigation bars, buttons, and interactive elements. JavaScript enhances the interactivity and responsiveness of the system, enabling seamless user interactions and smooth navigation throughout the application.

By leveraging the Django framework, integrating Celery for task automation, utilizing SQL for efficient database operations, and employing HTML, CSS, Bootstrap, and JavaScript for frontend development, the Student Grievance System is designed to be a comprehensive and user-friendly platform for managing student grievances effectively and efficiently.

C. Key features

User-Friendly Interface: The Online Grievance Management Portal has a user-friendly layout that is meant to be accessible and easy to use for all stakeholders, including students, professors, and administrative staff. Intuitive navigation routes and clear design components help to provide a consistent user experience across several devices and platforms.

Secure authentication and access controls: The portal uses strong authentication procedures to protect sensitive information and ensure data integrity. Role-based access controls are in place to restrict rights based on user roles and responsibilities within the institution, and secure login credentials are necessary for user access.

Centralized Grievance Submission: The portal allows for the centralized submission of grievances using an easy-to-use web form. Students can offer extensive explanations of their concerns, attach supporting papers, and define desired results, which simplifies the complaint filing process.

Automated notifications and progress tracking: Complainants get automated confirmation emails after submitting a grievance, which includes reference numbers for tracking purposes.

Administrative staff and selected people are quickly aware of incoming grievances, guaranteeing speedy evaluation and resolution.Transparent Workflow Management: The Online Grievance Management Portal includes a transparent workflow management system that allows stakeholders to view the status and progress of grievances in real-time. Comprehensive information on assigned persons, resolution timescales, and updates is easily available, encouraging responsibility and informed decision-making.

V. RESULTS

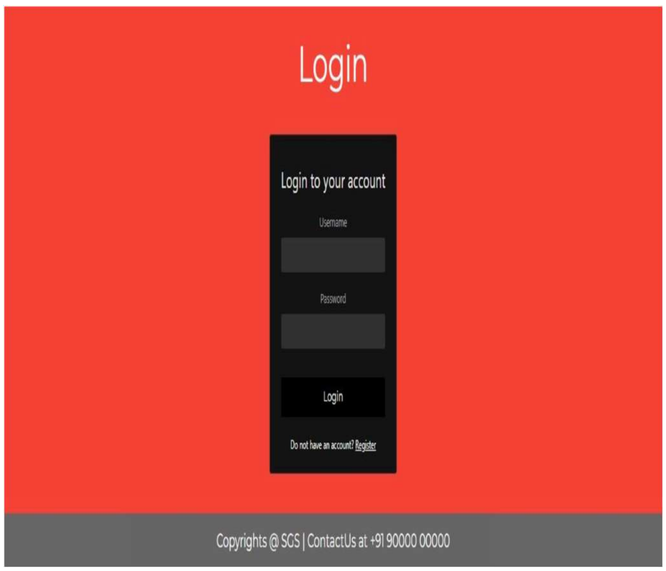

Login
Login to your account
Username
Password
Login
Do not have an account? Register
Copyrights @ SGS | ContactUs at +91 90000 00000

Fig 2. SGS Home Page

Fig 3. SGS Login Page

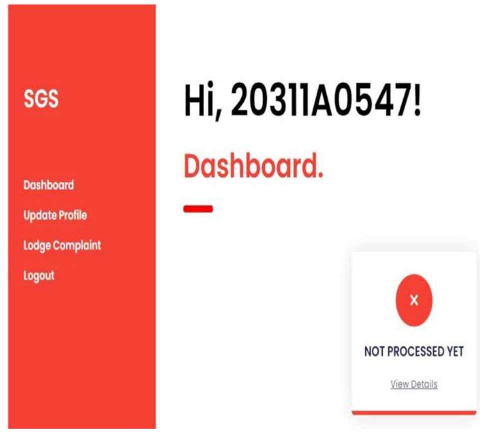


Fig 4. StudentDashboard

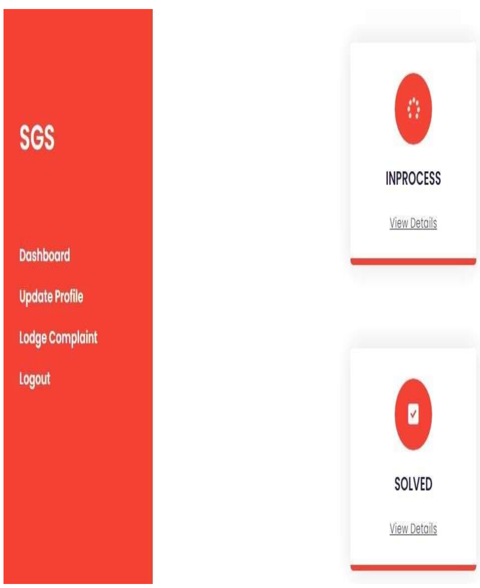


Fig 5.

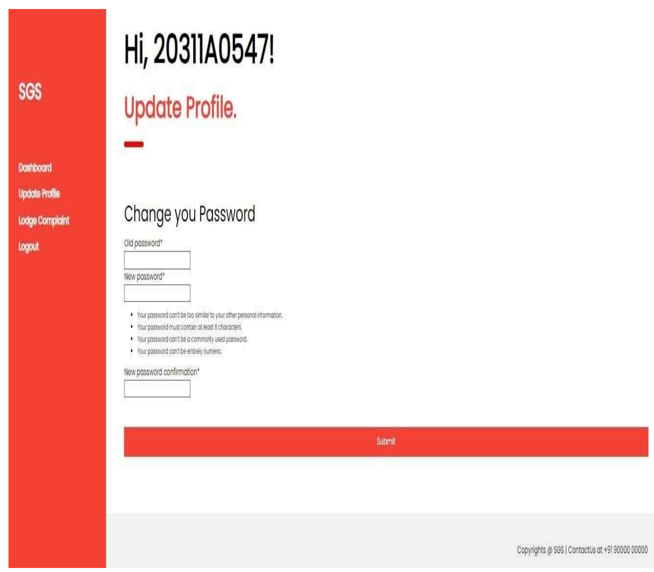


Fig 7. UpdateProfile for Students

Fig.4, Fig.5, and Fig.6 show the Student Dashboard.

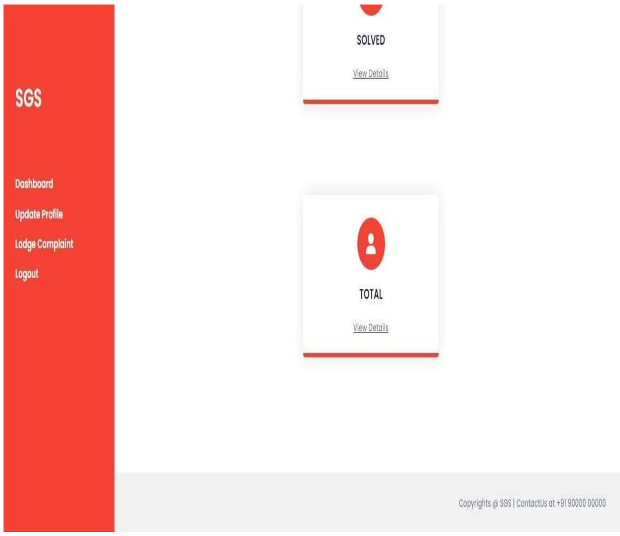
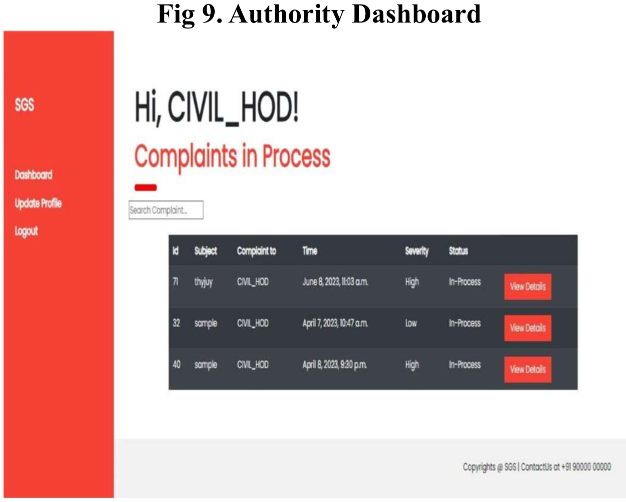


Fig 6.

Fig 9. Authority Dashboard



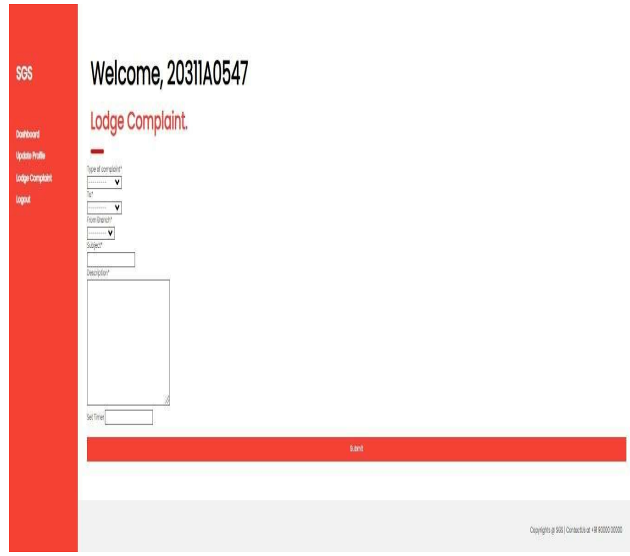
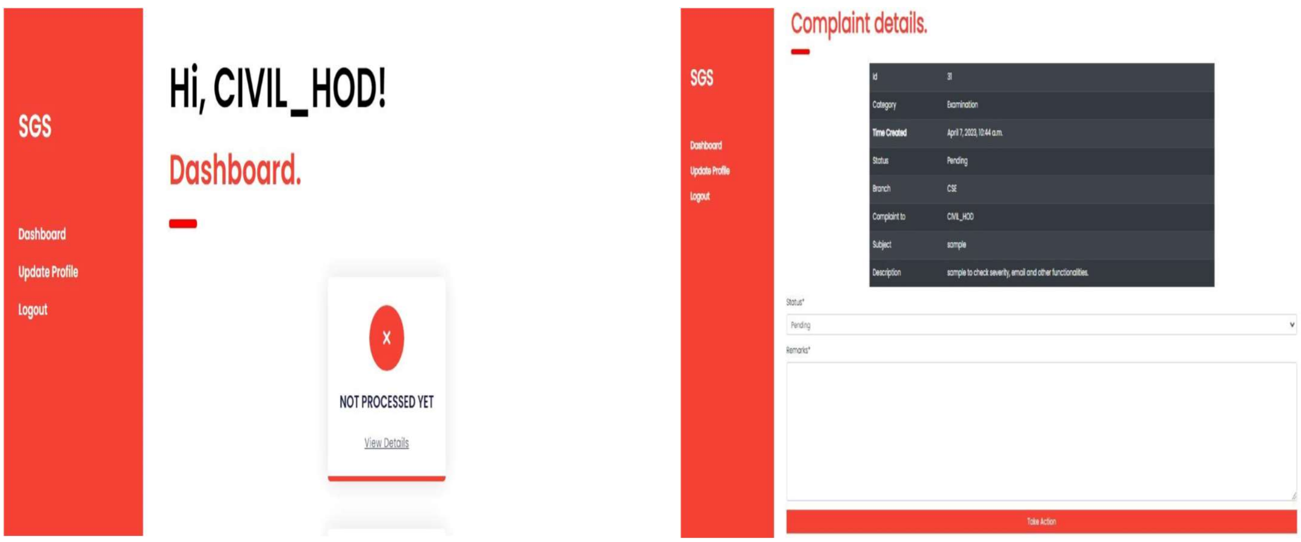


Fig 8. Lodge Complaint

Fig 11. View Complaint Details

VI.CONCLUSION

In conclusion, the development and implementation of the Student Grievance System (SGS) represent a significant step towards enhancing the overall student experience and fostering a supportive learning environment within educational institutions. Through the utilization of the Django framework, coupled with innovative features such as Celery for task automation and SQL for efficient database management, the SGS offers a robust platform for addressing student grievances promptly and fairly. The SGS is designed to streamline the complaint resolution process, providing students with a user-friendly interface to lodge complaints and receive timely updates on their status. By automating communication through email notifications and reminders, the system ensures that both students and



authorities remain informed throughout the grievance resolution process, thereby promoting transparency and accountability. Furthermore, the SGS prioritizes user experience by incorporating built-in forms for seamless profileupdates and leveraging frontend technologies like HTML, CSS, and JavaScript to create an intuit and visually appealing interface. This focus on usability and accessibility underscores the commitment to ensuring that students can navigate the system with ease and confidence. As educational institutions continue to adapt to the evolving needs of their student populations, the SGS serves as a vital tool for promoting a culture of trust, fairness, and inclusivity. By actively seeking feedback, monitoring system performance, and making necessary improvements, institutions can further enhance the effectiveness and efficiency of the grievance management process. In essence, the Student Grievance System represents a proactive approach to address student concerns and promote a safe and supportive learning environment. Through its implementation, educational institutions demonstrate their dedication to fostering positive relationships with students and uphold the values of fairness, transparency, and accountability in all aspects of academic life.

VII. FUTURE SCOPE

Future improvements to the student grievance system, which was created with the Django framework, are quite probable. Students will be able to access various devices by improving responsive site design and mobile application integration. Also, by integrating advanced analytics tools, organizations will be able to extract useful information from complaint data, optimizing workflows and promoting well-informed decision-making. Predictive capabilities and proactive problem-solving are expected to improve with artificial intelligence integration, including chatbots and sentiment analysis. It will also encourage a cooperative approach to grievance resolution and institutional improvement if stakeholder engagement is extended to include faculty, staff, and alumni. With constant iteration and user-driven improvements, the system will develop into a more responsive,

REFERENCES

[1] Tijare, P. A., Deshmukh, P., Agrawal, R., Upadhyay, P., Bhutada, V., & Amale, S. (2023). Students' Grievance System for Educational Institutions. Journal of Emerging

[2] Technologies and Innovative Research, Volume(Issue).

[3] Sharma, A., & Gupta, S. (2020). Design and implementation of student grievance handling system. 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1-6.

[4] Smith, J. D., & Jones, A. B. (2018). Implementing a student grievance procedure: A case study. Journal of Higher Education Policy and Management, 40(1), 78-94.

[5] Patel, R., & Singh, S. (2017). Analysis of grievance handling system in higher education institutes: A case study. International Journal of Information Technology and Management, 16(1), 57-73.

P53RW017

Smart Health Assistant: Integrating Machine Learning for Disease Prediction from Symptoms through Advanced Modelling

Koleti Ajay , Aleti Dheeraj Kumar, Kotha Rohith Reddy, Mr. D. Ram Babu

B.Tech - IVth year Student, B.Tech – IVth year Student, B.Tech – IVth year Student ,Assistant professor

Abstract—

Using machine learning technology for disease prediction in real-world application in the field of medicine represents a revolutionary approach, giving new and previously impossible ways to save millions of lives. Machine learning models, relying on enormous databases and complicated algorithms, can find intricate patterns and connections that conventional diagnostic methods cannot find . In this work, we implement an advanced system of differential diagnostics based on machine learning platforms. We developed our classification models, such as a random forest gradient-boosting classification, and decision tree to achieve a high accuracy of the disease prediction. These models have high classification power due to their ability to work with considerable quantities and various types of data. First and foremost, the model has a number of clear and evident benefits; as a result, it allows individuals to receive timely and accurate predictions that will facilitate early recovery and treatment in cases when medical intervention is relevant. Second, with actionable data that indicates potential hazards for one's well-being, we can see that the system could significantly reduce the pressure on healthcare resources and minimize incremental rates of preventable diseases. Furthermore, the research emphasizes the outline and goal of predictive analytics in terms of introducing health efficiency and access to improvements. Therefore, with the help of predictive models that forecast the occurrence or development of disease, practitioners will become capable of better allocating resources, prioritizing individuals with a high likelihood of obtaining the disease, and adjusting treatment options to the needs of a specific patient.

Keywords—Gradient Boosting, Random Forest, Decision Tree, Machine Learning, Predictive Analytics, Gradient prediction.

INTRODUCTION

The capacity of machine learning to assess vast quantities of data and give insights that were once out of bounds has remained in high demand over the last few years . Furthermore, the diagnostic, therapeutic, and administrative sectors of healthcare have recognized the role machine learning can play Patient data may be a bluff for precise, fast disease prediction. With patient data, machine learning offers an attractive opportunity for precise and effective disease prediction. Machine learning algorithms can improve diagnosis accuracy and patient results by identifying complex links in large-scale datasets. In order to forecast disease, the paper compares three machine learning techniques: Decision trees, random forests, and gradient boosting.

The objective is to assess how well illness prediction works with symptom information from medical records. The study uses rigorous analysis methods to train and assess these models. The performance of the model will be evaluated using important indicators such as classification reports, confusion matrix and accuracy.

This research has considered an area of disease prediction as one of the most crucial areas where machine learning can be significantly beneficial. The utilization of symptom data from large records and other self-reported sources creates an avenue to implementing and studying disease prediction. This work can be valuable in demonstrating the three major machine learning processes, and especially through systematic methods used in this research, this could be meaningful in comparing their performance and predictive accuracy for a disease.

LITERATURE SURVEY

Numerious studies have been carried out regarding disease prediction using various machine learning techniques which can be utilized by many medical facilities.

S. Grampurohit and C. Sagarnal[1] assessed machine learning models for diagnostic tests. They discovered that Support Vector Machine (SVM) outperformed Naive Bayes, Decision Trees, and K-Nearest Neighbor in terms of effectiveness for Parkinson's disease and renal ailment. Heart disease could be accurately predicted by using logistic regression (LR). CNN and Random Forest have demonstrated accuracy in predicting common ailments and breast cancer, respectively.

Aditi Gavhane[2] and her colleague proposed a machine learning model to predict cardiac disease. This system uses the multi-layer perceptron model. This model makes predictions on heart disease, based on some common symptoms like age, sex, pulse rate etc. The accuracy of the proposed system is 91%.

Gupta A, Kumar L, Jain R, and Nagrath P[3] proposed a heart prediction system utilizing the naive Bayes algorithm, achieving an accuracy of 97%. Based on symptoms such as breathing difficulties, back, neck, and chest pain, this algorithm makes predictions about cardiac illnesses. The system shows encouraging results in heart disease diagnosis by utilizing the ease of use and efficacy of the naive Bayes algorithm. This highlights the potential of machine learning to improve medical diagnosis and decision-making for better healthcare outcomes.

Sneha R,Nandini, Monisha S, Jahnavi C. “Disease Prediction System”.[4] 2021. Srms elibrary, 2021, Accessed 20 Oct 2021. The project by Iswarya et al. used classification algorithms to predict disease. Since Naive Bayes had accuracy of 97%, it has proven to be the suitable choice to use for disease prediction algorithm. Therefore, the use of various classification methods is critical to enhance overall disease prediction model accuracy and reliability.

Rudra[5] A and his colleague have suggested a system for multiple disease prediction . A remarkable novelty of this system is the anticipated appearance of consulting drug alcohol medication and medicine which isn’t there in this current model. The correctness of the said system is 85% .

Monika Gandhi and Shailendra Narayan Singh[6] proposed a framework designed to forecast heart ailments using data mining methodologies. Their investigation encompassed an examination of different data mining algorithms, such as Naive Bayes, Neural Network, and Decision Tree algorithms, implemented on medical data sets to enhance the accuracy of disease prognosis.

N. Kosarkar, P. Basuri, P. Karamore, P. Gawali, P. Badole, and P. Jumle[7] . The system developed using machine algorithms such as random forest, support vector machine, and logistic regression shows an accuracy of 82%. This system indicates that machine algorithms can also be used in predicting and making decisions about health outcomes.

Naveen Kumar, Naveenkumar S, Kirubhakaran R, Jeeva G, Shobana M, and Sangeetha Khas "Health Prediction System using Machine Learning Algorithms"[8]. They developed a health prediction system using machine learning algorithms. With an accuracy rate of 94%, the system shows that machine learning techniques can be integrated into the health system to predict and diagnose medical conditions

Prediction of common diseases based on Dahiwade D, Patle, G, & Meshram,E[9] explored the prediction of common diseases utilizing patient symptoms, lifestyle habits, and diagnostic data through the application of K-Nearest Neighbour and Convolutional Neural Network. The findings demonstrated an accuracy of 84.5% for the CNN algorithm in forecasting common disease models, surpassing the accuracy of the KNN model.

Ambekar, S & Phalnikar, R[10] Accurate data analysis plays a vital part in the disease diagnosis and treatment, particularly at an early phase of patient care. Hence, using the Naïve Bayes and KNN algorithms, a Heart Disease's prediction model is developed over here, and later, it will be extended to predict disease risks from organized data.

BACKGROUND WORK

Indeed approaches to the integration of powerful technologies as modern machine learning have fundamentally changed approaches even to diagnosis and monitoring of patients . At the same time, these systems have several disadvantages that complicate their applicability and relevance in clinical practice.

Firstly, all approaches to the prediction of diseases and their criteria will be scientific data, and it is difficult for medical parameters to be evidence alone. The datasets themselves become outdated quickly, and the accrual and transmission of appropriate information will be difficult, especially in the regions with the lack of medical infrastructure.

Moreover, the use of such systems without the recommendation and appointment of a doctor will cause the user to select the wrong specialists, and non-existent, or nonconvenient, user interfaces without mobile support will make such services unavailable.

METHODOLOGY

The project's goal is to develop an online platform that, using symptoms entered by users, employs machine learning to forecast likely symptoms. The website aims to empower individuals by offering user-friendly resources for preliminary health assessments, potentially resolving issues related to timely and cost-effective medical care access. The platform uses strong machine learning methods such as random forest, decision tree classification, and gradient boosting classification.

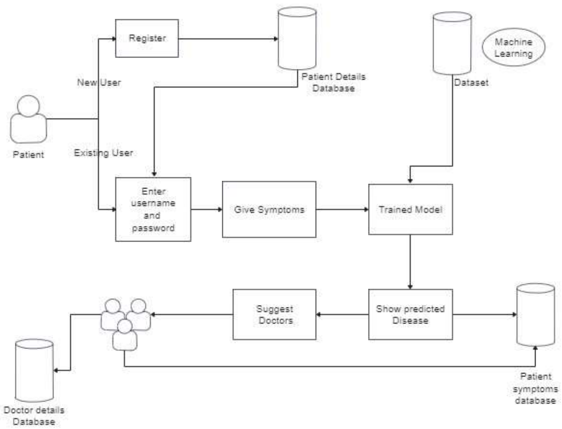


Figure 1: Architecture of Smart Health Assistant

In the given illustration, figure.1. Our system comprises three modules: Admin, User , and Doctor . A new user must register with the admin first. Upon a successful registration, the user will then have to enroll before signing in. A user only needs to sign up once. The illness prognosis system users are the physician, the patient, and the administrator. The system further verifies the identity of each and every user . System access is user-role-based. A patient can give symptoms, and the system will find the illness whereby the user will provide a probable diagnosis. The system also suggests a doctor once the illness has been predicted .The patient able to see a doctor online according to his convenience at home any free time.

A. Dataset Collection:

Data collection for disease prediction involves obtaining comprehensive datasets containing information about diseases and their associated symptoms. This dataset may consist of 133 columns and 4920 rows, providing detailed information about each disease-symptom relationship.

itching	skin_rash	nodal_skin_eruptions	continuous_sneezing	shivering	chills	joint_pain	stomach_pain	acidity	ulcers_on_tongue	muscle_wasting
1	1	1	0	0	0	0	0	0	0	0
0	1	1	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0
0	1	1	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0	0
0	0	0	1	1	1	0	0	0	0	0
0	0	0	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	0
0	0	0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	0	1	1	0	0
0	0	0	0	0	0	0	1	1	1	0

Figure 2: Data set 4920 records and 133 columns

B. Data Pre-Processing:

The purpose of data pre-processing is to organise and tidy up the data gathered. To maintain the consistency and uniformity of data, tasks include the removal of white spaces, punctuations and commas.

C. Classification:

Various machine learning classification methods are used to classify data and predict diseases based on specific symptoms. Gradient Boosting, Random Forest and Decision Tree are examples of common approaches.

1. Random Forest:

This model, which falls under the category of supervised learning, employs marked data to enable the classification of unmarked data.. Unlike K-means cluster algorithms, which were discussed in previous articles as unsupervised learning models, Random Forest algorithms are versatile tools capable of solving regression and classification challenges and are the preferred choice among engineers.

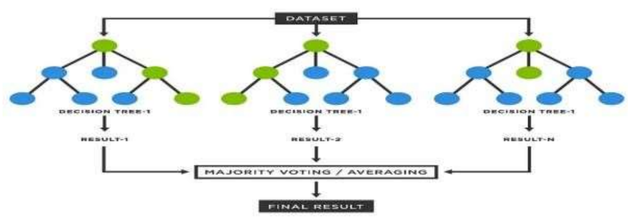


Figure 2: Random Forest Algorithm

2. Decision Tree:

The Decision Tree Classifiers create a tree-like classification structure by recursively dividing data according to functional requirements. They pass from root to leaf node, represent the properties of each node, and describe the class label of each leaf node to find the best choice. Decision Tree Classifiers are unique in their adaptability and simplicity. Primarily, as they handle both numerical and categorical data, they are applicable in a myriad of classification problems in different applications domains. Moreover, their transparency makes it easy to visualize and interpret how the trees classify data and, by extension, discern the patterns within. Additionally, while the classifier might overfit with complex datasets, pruning ensures that the model generalizes better. On that note, Decision Tree Classifiers are fundamental in machine learning, given their intuitive and exhaustive approach in classifying data and learning useful patterns and relationships there in.

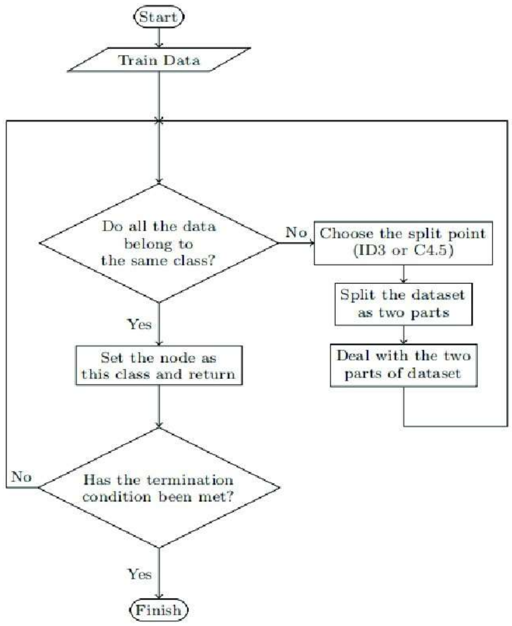


Figure 3: Decision Tree Algorithm

3. Gradient Boosting:

Gradient enhancement trains the decision tree team in sequence. Each tree corrects the error of the previous tree, focusing on the incorrect information of the previous tree. This collaborative approach makes it a powerful classification tool that achieves greater accuracy than a single tree. It can handle various data types and is less likely to be overused. Although powerful, it is complex, and requires more adjustment parameters than simple models.

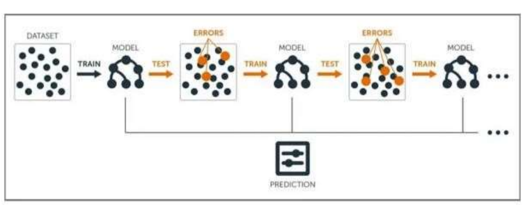


Figure 4: Gradient Boosting Algorithm

D. Model Training and Evaluation:

An essential step in the creation of a disease prediction system is the model training and evaluation phase. While fitting the training set to ML algorithms such as Decision Trees, Random Forests, and Gradient Boosting, pre processed data is used to let ML algorithms learn patterns and associations in the data. This is achieved by tuning internal parameters to minimize errors in predictions and improve predictive efficiency. Following fitting, each model is assessed using performance metrics to compare its effectiveness in omitted data, including accuracy, precision, recall, and F1-score. By using cross-validation techniques, performance is determined while ensuring its generality and the latter model's power. Through repeating the crossvalidation and performance checking of the model, one can determine the best model for disease prediction.

E. Development of a web platform:

A user-friendly website was designed to allow users to input their symptoms and obtain the predictions of the disease. The frontend was done with help of HTML and CSS, allowing for building the interface, optimized for a userfriendly approach and pleasant appearance. The backend was developed with Django, a Python framework that allows for server-side logic to be handled, interaction with the model for disease prediction, and the training process itself. Additionally, the site is based on a PostgreSQL database management system. The website was developed with the opportunity for users to enter their symptoms through an input form. Therefore, after submission, the Django server on the backend processes the input, which then runs through the machine learning model. The predictions are then presented to the user, including relevant diagnosis conditions predicted for each disease. Ultimately, the platform supports users in predicting diseases based on symptoms in a simple and interactive way.

RESULTS AND DISCUSSION

After a detailed comparison of accuracy and crossvalidation scores in Gradient Boosting, Random Forest, and Decision Tree algorithms, Gradient Boosting was found the most efficient machine learning technique. Provided that the achieved accuracy was 94.5%, the cross-validation score was 97.6%, Gradient Boosting performed better than two other methods across both the accuracy of predictions and generalization levels. This result indicates the high prospects of using Gradient Boosting in disease prediction, which can become a revolutionary finding in medical theory and the spread of preventive medicine.

Model	Accuracy (before cross validation)	Accuracy (after cross validation)

Decision Tree	72.7%	76.1%
Random Forest	94.0%	95.2%
Gradient Boosting	94.5%	97.6%

Table 1 : Comparative study of performance

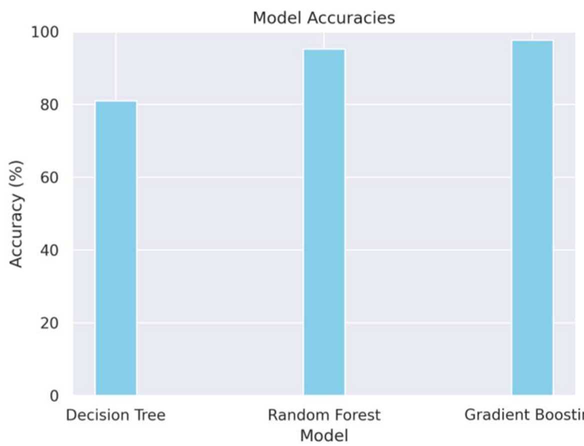


Figure 5: Accuracy of Models

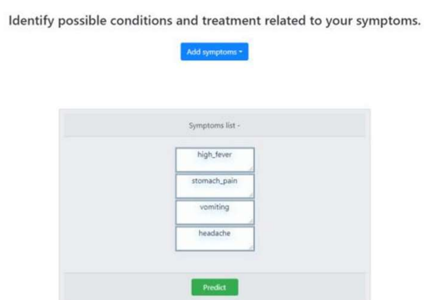


Figure 6.The user will add the symptoms

Figure 7: Output

CONCLUSION

In conclusion, this project has presented the possibility of developing web-based platforms for disease prediction using machine learning technologies. Currently, the model's high performance level close to ideal, 97.6%, and the given condition of the platform provides the opportunity to launch. Its high level of accuracy proves the potential of this platform as a pre-health evaluation tool for consumers.

The findings and predictions presented by the platform in realtime can be beneficial to proactively intervene and make meaningful health-related decisions. However, it does not exclude the need for an on-line professional for doing a medical diagnosis. Therefore, the existing traditional medical intervention may be initiated during the consultation if the patient's indicators are alarming. Instead, the platform will complement the current healthcare by facilitating the transmission of knowledge that will lead to preventative medical practices.

FUTURE SCOPE

Further tuning and refining of the machine learning models may improve their forecast accuracy and ability to generalize. This can include looking at cutting-edge techniques like hybrid models, deep learning, or ensemble learning to extract more intricate patterns from healthcare data.The developed prediction models can be integrated into existing healthcare systems and electronic health record (EHR) platforms to assist medical staff in promptly and accurately identifying patients. Public Health Initiatives, preventive healthcare measures, such as lifestyle modifications or targeted screening programs.

REFERENCES

[1] Grampurohit, S. & Sagarnal, C., 2020. Disease Prediction using Machine Learning Algorithms. In: 2020 International Conference for Emerging Technology (INCET), Jun 5-7, 2020, Belgaum, India

[2] Aditi Gavhane, Gouthami Kokkula, Isha Pandya and Kailas Devadkar, "Prediction of Heart Disease Using Machine Learning" IEEE Xplore ISBN: 978-1-5386-0965-1, pp. 1275-1278, 2018.

[3] Gupta A., Kumar L., Jain R., Nagrath P. (2020) Heart Disease Prediction Using Classification (Naive Bayes). In: Singh P., Pawłowski W., Tanwar S., Kumar N., Rodrigues J., Obaidat M. (eds) Proceedings of First International Conference on Computing, Communications, and

Cyber-Security (IC4S 2019). Lecture Notes in Networks and Systems, vol 121. Springer, Singapore.

[4] Sneha R. Monisha S. Jahnavi C. and S. Nandini "Disease Prediction Based On Symptoms Using Classification Algorithm" Journal of Xi'an University of Architecture & Technology vol. 12 no. 4 2020.

[5] Rudra A. Godse Smita S. Gunjal Karan A. Jagtap Neha Mahamuni and Suchita Wankhade "Multiple Disease Prediction Using Different Machine Learning Algorithms Comparatively" International Journal of Advanced Research in Computer and Communication Engineering vol. 8 no. 12 Dec 2019.

[6] Monika Gandhi and Shailendra Narayan Singh "Predictions in heart disease using techniques of data mining" 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE) pp. 520-525 2015.

[7] N. Kosarkar, P. Basuri, P. Karamore, P. Gawali,P. Badole and P. Jumle, "Disease Prediction using Machine Learning," 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), Nagpur, India, 2022, pp. 1-4, doi:
10.1109/ICETET SIP2254415.2022.9791739.

[8] Naveenkumar S1 , Kirubhakaran R2 , Jeeva G3 , Shobana M4 , Sangeetha
K5," Smart Health Prediction Using Machine Learning: A survey",3 march 2021.

[9] Dahiwade, D., Patle, G. & Meshram, E., 2019. Designing Disease Prediction Model Using Machine Learning Approach. In: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), March 27-29, 2019, Erode, India.

[10] Ambekar, S.Phalnikar,R. , 2019, Disease Risk Prediction by Using Convolutional Neural Network,IN 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) 16-18 Aug. 2018, Pune, India.

[11] Prof. Shalu Saraswat, Shweta Gabhane, Alisha Pawar, Suhas Pingat, Shreyas Patil, "Implementation of Smart Health Prediction Using ML", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, PrintISSN : 2395-1990, Volume 10 Issue 3, pp.112-117, May-June 2023.

[12] S. Mahata, Y. B. Kapadiya, V. Kushwaha, V. Joshi, and Y. Farooqui, "Disease Prediction and Treatment Recommendation Using Machine Learning," International Journal for Research in

Applied Science and Engineering Technology, vol. 11, no. 3, pp. 1232–1237, Mar. 2023, doi: 10.22214/ijraset.2023.49641.

[13] Saraswat, S., Gabhane, S., Pawar, A., Pingat, S., & Patil, S. (2014). Smart Healthcare Prediction Using Machine Learning. Journal of Emerging Technologies and Innovative Research (JETIR), 10(2).

[14] Abhang, A.R., Pravin, P., & Niraj, H. (2023). Smart Healthcare Prediction Implemented with Machine Learning. International Research Journal of Modernization in Engineering Technology and Science, 05(11).

P57RW007

Evaluating performance of Large Language Model chatbots on Selective Web Retrieved data.

[1] Etukala Yogesh Reddy, [2]Asamani Manoj Karthikeya,[3] Ajjela Rahul Reddy ,[4] Aruna Varnasi Roll Numbers: [1] 20311A0524, [2]20311A05F0,[3] 20311A0505

[1,2,3,4]Department of CSE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

[1] 20311a0524@sreenidhi.edu.in,
[2]20311a05f0@sreenidhi.edu.in, [3]
20311a0505@sreenidhi.edu.in,[4]
arunavarnasi@sreenidhi.edu.in

Abstract—

Language models have become pivotal in various natural language processing tasks, ranging from text generation to sentiment analysis. However, their performance can vary significantly depending on the data they are trained on. In this study, we focus on evaluating the performance of language models on selectively retrieved web data. This project focuses on harnessing web scraping and parsing techniques to enable large language models to answer questions based on content extracted from diverse websites. The methodology involves systematic retrieval and parsing of textual data from targeted web sources, ensuring the acquisition of comprehensive and contextually rich information. Our objective is to evaluate the performance of large language models in this task, utilizing metrics such as Faithfulness and context relevance. Through meticulous experimentation and analysis, we assess the ability of these language models to generate accurate and contextually appropriate responses to user queries. Faithfulness metrics gauge the fidelity of modelgenerated answers to the information present on the scraped websites, ensuring that responses accurately reflect the content. Context relevance metrics, on the other hand, measure the extent to which answers align with the broader context of the question and surrounding content.

Index Terms—Large Language Models, Web documents, Faithfulness, Context relevance.

I. INTRODUCTION

A. Introduction of the Project

Large Language Models (LLMs) have revolutionized natural language processing, demonstrating remarkable abilities in various tasks. However, their performance on selective webretrieved data remains an open question. This project aims to evaluate and compare two leading LLMs, Gemini 1.0 Pro and GPT-4, based on their ability to process and utilize information from web documents within a Retrieval-Augmented Generation (RAG) pipeline.

The evaluation will focus on three key metrics: faithfulness, context utilization, and context relevancy. Faithfulness assesses how well the LLM's output aligns with the source document's information. Context utilization measures the LLM's effectiveness in leveraging the provided context to generate its response. Context relevancy evaluates the LLM's ability to identify and focus on the most pertinent information within the retrieved document.

By analyzing these metrics, this project seeks to gain insights into the strengths and limitations of Gemini 1.0 Pro and GPT-4 when working with selective web data. This comparative analysis will contribute to a deeper understanding of these LLMs and shed light on the potential and challenges of utilizing LLMs in conjunction with web-based information retrieval systems.

B. Scope

This project will evaluate the performance of two leading large language models (LLMs), Gemini 1.0 Pro and GPT-4, when working with selective web-retrieved data. The evaluation will focus on three key metrics: faithfulness, context utilization, and context relevancy. These metrics assess how well the LLMs can align their output with the source document's information, leverage the provided context to generate responses, and identify the most pertinent information within the retrieved document, respectively.

The LLMs will be tasked with generating responses based on the retrieved data and user queries, potentially including summarization, question answering, and content creation. The project will analyze the performance of the LLMs on these tasks and compare their strengths and weaknesses. However, it is important to note that the project will focus specifically on these two LLMs and the chosen metrics, and the performance may vary depending on the specific web data and tasks used. Additionally, the project will not delve into the ethical or societal implications of using LLMs with web-retrieved data. Overall, this project aims to provide valuable insights into the capabilities and limitations of current LLMs when working with selective web data, contributing to their ongoing development and potential applications.

C.Project Overview

This research project aims to evaluate and compare two leading large language models (LLMs), Gemini 1.0 Pro and GPT-4, on their ability to process and utilize selective webretrieved data. Relevant web documents will be retrieved based on user specific query. These documents will then be cleaned and preprocessed to ensure their quality and compatibility with the LLMs. The cleaned documents will be transformed into vector representations using appropriate embedding techniques. These representations will be stored in a dedicated document store for efficient retrieval and access by the LLMs. The performance of the LLMs on the defined tasks will be evaluated based on three key metrics: faithfulness, context utilization, and context relevancy. These metrics assess how well the LLMs align their output with the source information, leverage the provided context, and identify the most pertinent information within the retrieved documents, respectively. The results will be analyzed to compare the strengths and weaknesses of each LLM and identify potential areas for improvement.

D.Objective

This project aims to benchmark the performance of Large Language Models (LLMs) on selective web-retrieved data, focusing on faithfulness and context relevance. We will evaluate LLMs, analyzing how factors like model architecture, training data, and retrieval strategies impact these metrics. Based on our findings, we will explore techniques to improve LLM performance, potentially through fine-tuning strategies, incorporating external knowledge sources, or developing novel evaluation metrics. Ultimately, this project seeks to contribute to a deeper understanding of LLM capabilities and pave the way for developing more robust and reliable models for future web-based applications.

II. LITERATURE SURVEY

A. Existing System

Chatbots have become increasingly common on websites, offering a seemingly convenient way to interact with customers and provide support. However, many existing chatbot systems suffer from significant inefficiencies and limitations, leading to frustrating user experiences and failing to fulfill their intended purpose. Some of the key shortcomings of current chatbot systems are:

Limited Understanding and Context: Many chatbots rely on simple keyword matching or rule-based systems, resulting in a limited understanding of user intent and context. This often leads to irrelevant responses, frustrating users, and failing to address their actual needs.

Lack of Personalization: Most chatbots offer a generic experience, failing to personalize interactions based on user preferences, history, or specific context. This impersonal approach can feel robotic and disengaging, hindering user satisfaction.

Restricted Capabilities: Many chatbots are limited in their capabilities, often only able to handle simple tasks or answer basic questions. When faced with complex inquiries or requests, they may fail to provide adequate support, requiring human intervention and negating their intended efficiency.

Poor Conversational Flow: Many chatbots struggle to maintain a natural and engaging conversational flow. Their responses can feel stilted and unnatural, making it difficult for users to have a seamless and productive interaction.

Data Security Concerns: Some chatbot systems raise data security concerns, particularly when handling sensitive user information. Data breaches and privacy issues can erode user trust and negatively impact the overall experience.

These limitations highlight the need for significant improvements in chatbot technology. Advancements in natural language processing (NLP), machine learning (ML), and artificial intelligence (AI) offer promising avenues for developing more sophisticated and efficient chatbots. By leveraging these technologies, future chatbots can achieve a deeper understanding of user intent, personalize interactions, and offer a wider range of capabilities, ultimately providing a more satisfying and productive user experience[1].

B. Related Work

In the field of chat bot and information retrieval systems, several notable studies and evolutions have made significant contributions to the advancement of machine learning methodologies. Two key related works are:

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication.** Communications of the ACM, 9(1), 36-45. (This classic paper describes ELIZA, an early chatbot that mimicked Rogerian psychotherapy through pattern matching and keyword recognition). Woebold, J. (1989). Appraising the commercial potential of natural language processing. Proceedings of the 12th international conference on artificial intelligence (pp. 923-928).

Morgan Kaufmann Publishers Inc. (This paper explores the early commercial applications of natural language processing (NLP) in chatbot development). Colby, K. M., Colby, B. M. (1971). Artificial paranoia. Artificial Intelligence, 2(1), 1-72. (This paper discusses SHRDLU, a chatbot that could understand and respond to queries about a simulated world, demonstrating early rule-based NLP techniques).

Books:

Wallace, R. J. (2009). The conversational computer. Springer Science Business Media. (This book provides a historical overview of chatbot development and explores the theoretical foundations of human-computer interaction). Lenat, D. B. (1995). CYC: A large-scale ontology for natural language reasoning. Communications of the ACM, 38(11), 34-43. (While not directly about chatbots, this book explores early efforts in building knowledge bases, which are foundational for some non-LLM chatbots).

These seminal works have provided valuable insights into the application of both ML and DL techniques within the realm of credit assessment. By harnessing a diverse array of financial attributes and employing advanced modeling techniques, the project endeavors to bolster risk management practices and lending decisions, ultimately contributing to improved financial outcomes for individuals and institutions alike.

III. PROBLEM IDENTIFICATION

A.Problem Statement

Website chatbots, despite their growing ubiquity, often fall short of providing a truly helpful and satisfying user experience. Their limitations in understanding human language and context lead to frustrating interactions, with users frequently receiving irrelevant or generic responses. These chatbots struggle to grasp the nuances of complex inquiries and fail to adapt to individual needs and preferences. Consequently, users are left feeling unheard and dissatisfied. Furthermore, current chatbots are often ill-equipped to handle complex tasks. While they may excel at answering basic questions, they falter when faced with requests that require accessing information from multiple sources or performing multi-step actions. This inability to effectively resolve user issues can lead to increased frustration and ultimately drive users away from the website. The robotic and unnatural nature of many chatbot interactions further exacerbates the problem. Their scripted responses lack the warmth and engagement of human conversation, hindering the development of trust and rapport.

Additionally, the limited capacity of most chatbots to learn and adapt over time results in stagnant performance and missed opportunities to improve the user experience. Addressing these shortcomings is crucial to ensure that chatbots become valuable tools that enhance user engagement and satisfaction.

B.Approach to the Problem Statement

Large language models (LLMs) offer a promising avenue for improving website chatbots. By leveraging their vast knowledge and ability to understand and generate human-like text, LLMs can address several key limitations of current chatbot technology[1]. LLMs can enable chatbots to better understand user intent, even in complex inquiries, and respond with more relevant and personalized information. Additionally, LLMs can generate more natural and engaging conversation, creating a more human-like interaction for users[2].

However, simply incorporating LLMs into chatbots does not guarantee optimal performance. Thorough evaluation and analysis are crucial to ensure that LLMs are used effectively and responsibly. This includes assessing the accuracy and reliability of the information generated by LLMs, as well as their ability to remain unbiased and avoid harmful or offensive language. Furthermore, it is important to evaluate how well LLMs integrate with existing chatbot systems and workflows, and to identify any potential ethical or privacy concerns.

By carefully evaluating and analyzing LLMs, developers can maximize their potential for improving website chatbots. This includes fine-tuning LLMs for specific tasks and domains, ensuring they are trained on relevant and unbiased data, and implementing safeguards to prevent harmful outputs. Through rigorous testing and continuous monitoring, LLMs can be used to create chatbots that are more helpful, engaging, and trustworthy, ultimately leading to a better user experience and increased customer satisfaction.

IV. PROPOSED SYSTEM AND ARCHITECTURE

A. Proposed System

This innovative system proposes a single, powerful chatbot to handle user queries across multiple websites. Powered by a large language model (LLM) and utilizing a RetrievalAugmented Generation (RAG) pipeline, this chatbot can access and process information from various websites to deliver accurate and relevant answers.

The key to this system's effectiveness lies in its ability to dynamically retrieve and parse web documents. Users are able to fetch web documents by using a URL. These documents are then efficiently parsed and processed, extracting key information that is then fed to the LLM. This allows the LLM to contextualize the user's query and formulate an answer based not only on its internal knowledge but also on the specific information retrieved from the relevant web pages[8]. This approach offers several advantages over traditional chatbot systems that rely on pre-defined knowledge bases or limited data sources. By leveraging the RAG pipeline, the chatbot can access and process information from a vast array of websites, ensuring that its responses are comprehensive and up-todate. This eliminates the need for manually updating separate knowledge bases for each website, streamlining information access and reducing maintenance overhead.

For large corporations that manage multiple websites, this system offers significant benefits:

Unified User Experience: Users can interact with a single chatbot to access information across all websites, eliminating confusion and frustration.

Improved Efficiency: The system eliminates the need for multiple chatbots or complex knowledge base management, reducing operational costs and increasing efficiency.

Enhanced Customer Engagement: By providing accurate and relevant answers, the chatbot can effectively engage users and address their needs, leading to improved customer satisfaction.

This single-chatbot system, powered by RAG and a robust LLM, offers a powerful and efficient solution for large corporations to manage information access across multiple websites, ultimately leading to improved user experience and increased customer engagement.

B. System Architecture

The architecture is designed to efficiently handle user queries and generate informative answers using a large language model and a document store. The system allows for flexibility through optional parameters at various stages, enabling control over retrieval, generation, and answer formatting. The use of Elasticsearch facilitates efficient storage and retrieval of information, while the large language model provides the ability to understand and respond to complex queries.

1)Data Input and Knowledge Base Creation: The architecture's foundation lies in the acquisition and storage of information. This begins with an Input URL which acts as a pointer to a source of knowledge, such as a website or a specific document. The ”Web Content Fetcher” plays the role of a retriever, accessing and downloading the content found at the provided URL. This content, which can include text,

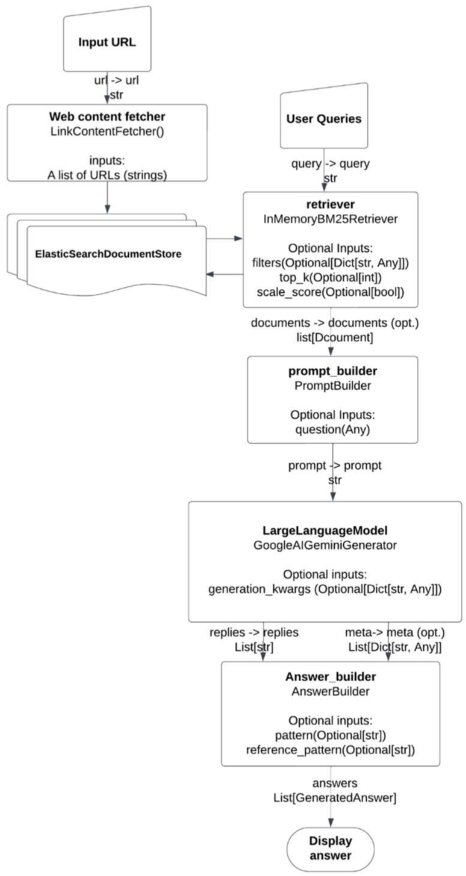


Fig. 1. Architecture for Web Retrieval enabled Large Language Model chatbot

images, and other forms of data, is then processed and stored within the Elasticsearch Document Store. Elasticsearch, with its robust indexing and search capabilities, serves as the central repository of knowledge for the entire system. It allows for efficient storage, retrieval, and management of the acquired information, effectively forming the knowledge base upon which the system operates. This stage ensures that the system has a comprehensive and readily accessible pool of information to draw from when responding to user queries.

2)Query Processing and Information Retrieval: Once the knowledge base is established, the system is ready to interact with users and address their information needs. This interaction begins with User Queries which are the questions or requests for information submitted to the system. These queries are then directed to the Retriever component, which acts as a search engine within the architecture. The specific retriever employed here is the BM25 Retriever, an algorithm known for its effectiveness in information retrieval tasks. This retriever interacts with the Elasticsearch Document Store, leveraging its indexing and search capabilities to identify and retrieve the most relevant documents that correspond to the user's query.

BM25 Retriever:

BM25 retriever, a variant of the BM25 algorithm, stands as a robust method for information retrieval in large-scale document collections. Leveraging a probabilistic framework, BM25 retriever effectively addresses the shortcomings of traditional TF-IDF models by introducing nuanced adjustments. Firstly, it considers term frequency (TF), acknowledging the importance of term occurrences within documents, while mitigating the risk of overemphasizing overly repetitive terms through a saturation function.

Secondly, BM25 retriever incorporates inverse document frequency (IDF), which evaluates the rarity of terms across the entire document corpus. This strategic weighting mechanism assigns higher importance to terms that are less common across documents, enabling the algorithm to discern the uniqueness and relevance of particular terms within a document.

$$\text{score}(D,Q) = \sum_{i=1}^{n} \text{IDF}(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot \left(1 - b + b \cdot \frac{|D|}{\text{avgdl}}\right)}$$

Furthermore, BM25 retriever incorporates document length normalization, a crucial factor in addressing the inherent bias towards longer documents in traditional retrieval models[6]. By normalizing term frequencies based on document length, BM25 retriever ensures that the impact of term occurrences is appropriately adjusted, leading to fairer comparisons across documents of varying lengths. This normalization process aids in the accurate ranking of documents based on their relevance to a given query.

3) Contextualization and Response Generation: Following the retrieval of relevant information, the system transitions into the phase of understanding and response generation. The retrieved documents, which contain the information most pertinent to the user's query, are combined with the original query submitted by the user. This combined information is then sent to the Prompt Builder. The role of the prompt builder is crucial as it structures the information into a comprehensive and coherent prompt that is suitable for the language model. This involves organizing the retrieved content, highlighting key points, and formulating a clear context for the language model to understand the user's information need within the larger framework of the retrieved knowledge.

The carefully constructed prompt is then fed into the Large Language Model(LLM). This powerful language model possesses the ability to process information, understand context, and generate human-like text. By analyzing the prompt, the language model gleans insights from the retrieved documents, grasps the nuances of the user's question, and formulates "replies" that address the query comprehensively[5]. The language model's ability to generate text allows it to provide informative and contextually relevant responses that go beyond simple retrieval of factual information.

General Architecture of Large Language Models:

Large language models typically follow a similar architectural blueprint, characterized by deep neural networks designed to process and generate human-like text[1]. While specifics may vary between models, such as GPT (Generative Pre-trained Transformer) series by OpenAI and BERT (Bidirectional Encoder Representations from Transformers) by Google, the general architecture comprises several key components.

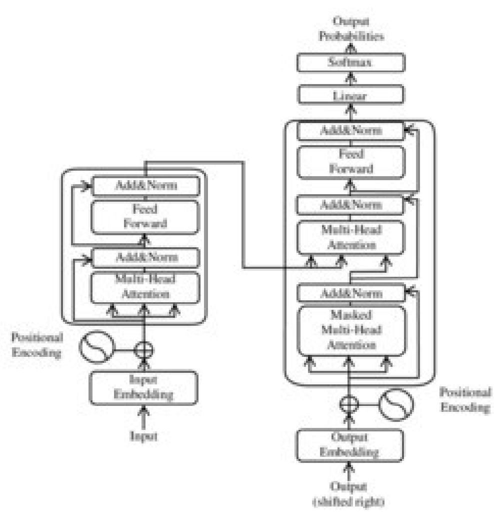


Fig. 2. General Architecture of Large Language Models

Transformer Architecture: At the heart of large language models lies the transformer architecture. Transformers are composed of multiple layers of self-attention mechanisms and feedforward neural networks. This architecture enables models to capture long-range dependencies in text, allowing them to understand context and generate coherent responses[2].

Pre-training and Fine-tuning: Large language models are typically pre-trained on vast corpora of text data using unsupervised learning techniques. During pre-training, the model learns to predict the next word in a sequence given the preceding context. This process allows the model to internalize linguistic patterns and structures. After pre-training, the model can be fine-tuned on specific tasks such as text classification, language translation, or text generation by further training on labeled data[9].

Attention Mechanism: Attention mechanisms enable the model to focus on relevant parts of the input sequence when generating output. Self-attention mechanisms allow the model to assign different weights to different words in the input sequence based on their importance. This enables the model to capture dependencies between words that are far apart in the input sequence, facilitating better understanding and generation of text.

Embedding Layers: Input text is typically converted into dense numerical vectors called embeddings before being processed by the model. Embeddings capture semantic information about words and their context in the input sequence, allowing the model to effectively represent and manipulate textual data.

Output Layers: The final layers of the model map the internal representations learned during processing back to the vocabulary space, generating probabilities over the vocabulary to predict the next word in the sequence. In generation tasks, such as text completion or language translation, the output layer produces the desired output text based on the learned representations.

Parameter Sharing and Parallelization: To handle the vast amounts of data and computations involved, large language models leverage parameter sharing and parallelization techniques. Parameter sharing ensures that the model learns reusable representations across different parts of the input sequence, while parallelization enables efficient training and inference on distributed computing architectures.

Overall, the general architecture of large language models combines sophisticated neural network architectures, attention mechanisms, and embedding techniques to process and generate human-like text across a wide range of natural language understanding and generation tasks.

4)Answer Formation and Presentation: The replies generated by the language model are then passed to the Answer Builder. This component structures the final answer ensuring a clear and coherent response. The resulting list of replies is then sent to the User Interface, which presents the final answer to the user. This completes the entire process from receiving the user's query to delivering an informative and relevant response.

5)Evaluating and Analysis: The evaluation process starts with a collection of Evaluation Queries which are designed to test the capabilities of the large language models (LLMs) in a controlled and focused manner. These queries are carefully chosen to cover a range of topics and complexities, ensuring a comprehensive assessment of the models' performance.

The retrieved documents, along with the evaluation query, are then used by the Prompt Builder to create a comprehensive prompt that encapsulates the information needs and context. This prompt is fed into both GPT and Gemini models, which operate in parallel, processing the same information and independently generating their respective sets of "replies." This parallel processing ensures a fair comparison, as both models are given the same input and the opportunity to demonstrate their capabilities under identical conditions. The generated replies from each model are then passed through their respective Answer Builder components, which structure and format the raw outputs into coherent and well-presented answers.

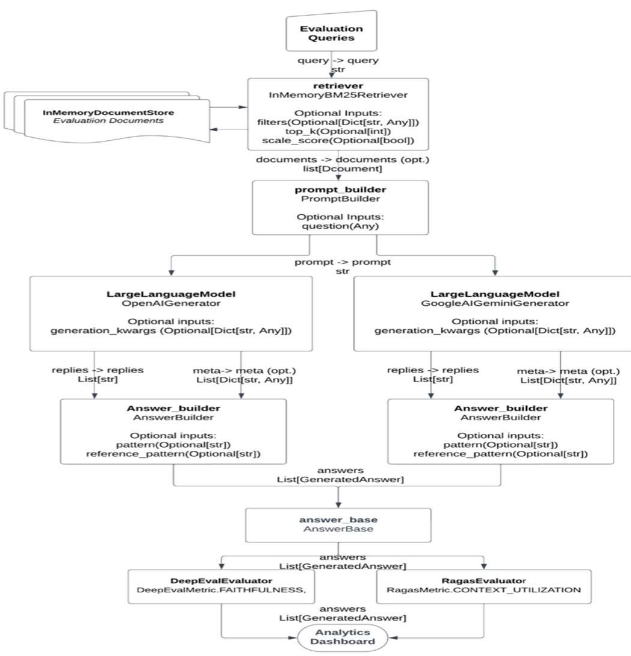


Fig. 3. Architecture for Evaluating performance of Large Language Model chat bots on Selective Web Retrieved data

The generated answers form the basis for evaluation, where the focus is on assessing the faithfulness and context utilization of each model. Both of the models are compared across various metrics like faithfulness, context utilization and context relevance.

Faithfulness:

This measures the factual consistency of the generated answer against the given context. It is calculated from answer and retrieved context. The answer is scaled to (0,1) range. Higher the better.

The generated answer is regarded as faithful if all the claims that are made in the answer can be inferred from the given context. To calculate this a set of claims from the generated answer is first identified. Then each one of these claims are cross checked with given context to determine if it can be inferred from given context or not.

Faithfulness score = |Number of claims |Total that number can be of inferred claims| from context|

Context Relevancy:

This metric gauges the relevancy of the retrieved context, calculated based on both the question and contexts. The values fall within the range of (0, 1), with higher values indicating better relevancy.

Ideally, the retrieved context should exclusively contain essential information to address the provided query. To compute this, we initially estimate the value of $|S|$ by identifying sentences within the retrieved context that are relevant for answering the given question. The final score is determined by the following formula:

context relevancy = |Total number of sentences|S| in retrieved context|

V. IMPLEMENTATION AND RESULTS

A. Modules

The input provided and output obtained are as follows: 1) Input: Website URLs 2) Output: Chat bot powered by large language model with extended knowledge base.

MODULE 1 : Web Document Retrieval Pipeline

First, we create a custom Haystack component acting like the grabber arm. This component takes a web document's URL as input and uses libraries like BeautifulSoup or Scrapy to fetch the raw HTML content. Next, the pipeline enters the processing stage. Here, we parse the HTML content, meticulously separating the wheat from the chaff. Parsing techniques break down the HTML structure, discarding unnecessary elements like tags, scripts, and advertisements. Libraries like BeautifulSoup come in handy, navigating the HTML maze and extracting the valuable text content.

Once the text is clean and ready for use, we create a Document object. This object encapsulates the processed text and stores the original URL as metadata for reference. Finally, the pipeline places this document object into the DocumentStore, like a library filing system, where it can be retrieved later to answer user queries.

MODULE 2 : Retrieval Augmented generation(RAG) Pipeline

The RAG (Retrieval-Augmented Generation) pipeline will be a crucial component in our project to evaluate large language models (LLMs) like Gemini 1.0 Pro and GPT-4[4]. This pipeline retrieves relevant web documents based on a user query and provides them as context to the LLMs for improved response generation[7]. Here's how it will function:

1. **Information Retrieval:** The pipeline starts with a user query. A retriever component, likely based on BM25 scoring, will scour a pre-built document store containing processed web documents. This document store could be built using Haystack's InMemoryDocumentStore or a similar solution. The retriever aims to identify the most relevant documents that align with the user's query.

2. **Contextual Feeding:** Once the retriever identifies relevant documents, they are passed to the LLM. This is where the magic happens. Unlike traditional LLMs that rely solely on their internal training data, the RAG pipeline feeds these retrieved documents as context to the LLM. Imagine providing supplemental information to a student before answering a question. Similarly, these documents provide additional context for the LLM, allowing it to understand the user's query within a broader framework.

3. **Enhanced Response Generation:** With the retrieved documents as context, the LLM, be it Gemini 1.0 Pro or GPT-4, is better equipped to generate a response to the user's query. This response should ideally be:

Faithful: Aligned with the factual information presented in the retrieved documents.

Context-Utilizing: Demonstrate a clear understanding and incorporation of the provided context.

Context-Relevant: The retrieved documents should be genuinely relevant to the user's query and contribute meaningfully to the LLM's response[3].

MODULE 3 : User Interface

The User interface will feature a prominent text box where users can enter website URLs to fetch the documents for context. Additionally, the UI will also feature a chat interface where the user can now chat with the LLM with its extended knowledge base. This fairly simple User Interface can be achieved by using the REST API.

MODULE 4 : Evaluation

The evaluation module will be the cornerstone of our project, assessing the performance of Gemini 1.0 Pro and GPT-4 on the RAG pipeline[2]. Here's a closer look at its functionalities:

1. Metric Calculation:

This module will be responsible for calculating the core metrics we defined: Faithfulness, Context-Utilization, and Context-Relevancy.

Faithfulness: Metrics like Haystack's Faithfulness can be employed to evaluate how well the generated response aligns with the factual information presented in the retrieved documents. This ensures the LLMs aren't simply fabricating information but are grounding their responses in reality[2].

Context-Utilization: Here, we'll measure how effectively the LLM incorporates the provided context from the retrieved documents. Tools might involve analyzing the generated response to see if it references or builds upon the information presented in those documents.

Context-Relevancy: This metric goes beyond simple utilization. It delves into how meaningfully the retrieved documents contributed to the LLM's response. Human evaluation might be necessary here, where experts judge if the retrieved documents were truly relevant to the query and if they genuinely influenced the LLM's response in a constructive way.

2.Data Collection and Storage:

The evaluation module will accumulate the calculated metrics for each query, LLM (Gemini 1.0 Pro or GPT-4), and retrieved document set. This data will be stored for further analysis[2].

3.Performance Analysis and Visualization:

The module will play a crucial role in analyzing the collected data. It might involve techniques like:

Statistical Analysis: Calculating averages, standard deviations, and other statistical measures to understand the overall performance of each LLM across various queries and contexts.

Visualization Tools: Charts and graphs can be generated to depict the LLM's performance on the different metrics. This visual representation allows for easier comparison and identification of strengths and weaknesses between Gemini

1.0 Pro and GPT-4..

B. Experimental Results

When comparing the performance metrics of two language models, Gemini and GPT, several key aspects come to light. Gemini demonstrates a robust utilization of context, scoring an average of 0.76041, indicating its adeptness at understanding and incorporating contextual cues within its generated responses. However, its faithfulness metric stands at 0.6666, suggesting a slightly lower degree of fidelity to the input prompts compared to GPT. On the other hand, GPT excels in faithfulness with a score of 0.9213, showcasing its ability to generate responses closely aligned with the input prompts. Additionally, both models exhibit relatively similar levels of contextual relevance, with Gemini at 0.13831 and GPT at 0.16214, suggesting a comparable capability in producing contextually appropriate responses. In summary, while Gemini shines in context utilization, GPT outperforms in faithfulness, with both models demonstrating commendable levels of contextual relevance.

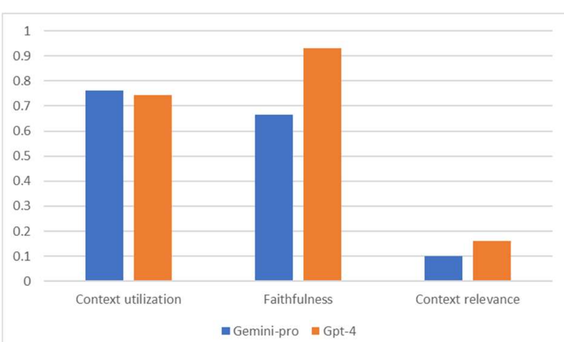


Fig. 4. Expermiental results

C.Comparative Results

In comparing the performance metrics of Gemini and GPT, several insights emerge, shedding light on the strengths and weaknesses of each model.

Gemini demonstrates a slightly higher context utilization score (0.76041) compared to GPT (0.74473). This indicates Gemini's proficiency in leveraging contextual information to generate responses, suggesting a nuanced understanding of the input prompts. However, GPT excels in faithfulness with a significantly higher score (0.9213) compared to Gemini (0.6666). This implies that GPT is more adept at faithfully reproducing the content and intent of the input, resulting in responses that closely align with the prompts.

When considering context relevance, both models exhibit relatively comparable scores, with Gemini at 0.13831 and GPT at 0.16214. This suggests that both models are similarly effective in producing responses that are contextually relevant to the input provided.

Overall, while Gemini showcases strong context utilization capabilities, GPT emerges as the frontrunner in terms of faithfulness. However, it's essential to note that both models demonstrate commendable levels of contextual relevance, indicating their overall proficiency in generating contextually appropriate responses. Thus, the choice between Gemini and GPT would depend on the specific priorities and requirements of the task at hand, balancing the need for context utilization, faithfulness, and relevance in generated responses.

VI. CONCLUSION AND FUTURE SCOPE

A.Conclusion

In conclusion, the evaluation of large language models, particularly Gemini 1.0 Pro and GPT-4, based on metrics such as faithfulness, context utilization, and context relevancy, is crucial for developing powerful chatbots capable of answering questions from diverse contexts sourced from the web. Through the implementation of a RAG (Retrieval-Augmented Generation) pipeline, we can effectively process web-retrieved documents and extract valuable information to enhance the capabilities of these chatbots.

Upon evaluation, Gemini 1.0 Pro demonstrates strong context utilization, making it adept at understanding and incorporating contextual cues from the retrieved data. However, it falls short in terms of faithfulness compared to GPT-4, which excels in reproducing the content and intent of the input prompts more faithfully. Both models exhibit comparable levels of context relevancy, indicating their ability to generate responses that are contextually appropriate.

B.Future Scope

The future scope of this project involves further fine-tuning and optimization of the large language models based on the evaluation results to enhance their performance in specific areas. Additionally, the integration of additional metrics and techniques, such as sentiment analysis and entity recognition, can enrich the chatbots' understanding and responsiveness. Moreover, continuous monitoring and updating of the models based on real-world feedback will be essential to ensure their effectiveness in addressing the evolving needs and challenges of users seeking information from various web contexts.

Overall, this project lays the groundwork for the development of advanced chatbot systems capable of providing accurate and relevant answers across a wide range of domains and topics.

VII. REFERENCES

[1] R Devon Hjelm (2022). A Comprehensive Overview of Large Language Models. arXiv. [arxiv.org]

[2] Jay Alammar (2017). The Illustrated Transformer. [The Illustrated Transformer]

[3] Patrick Lewis et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In Findings of the EMNLP Workshop on NLP for Industry (IndustryNLP).

[4] Jinpeng Guo et al. (2023). [2312.10997] Retrieval-
Augmented Generation for Large Language Models:A Survey. arXiv. [arxiv.org]

[5] Rodrigo Nogueira et al. (2019). A Primer on Neural Search for Information Retrieval. arXiv. [1904.09078]

[6] Fabio Petroni et al. (2022). RAG-Transformer: Hybrid Text Retrieval and Generation for Natural Language Understanding. arXiv.

[7] Liu, Yizhe et al. (2022). Towards Controllable RetrievalAugmented Generation. arXiv. This paper by Yizhe Liu et al. investigates methods for controlling the outputs of RAG models for tasks requiring specific styles or factual accuracy levels.

[8] Cui, Yiming et al. (2022). RAG for Open-Domain Question Answering. arXiv. This paper by Yiming Cui et al. explores the application of RAG models to open-domain question answering tasks, where the answer can reside outside the pre-trained data of the LLM.

[9] HuggingFace (2023). Understanding Retrieval-Augmented Generation (RAG) Models. [Understanding Retrieval-Augmented Generation (RAG) Models — HuggingFace]

P58RW006

Hand Gesture Recognition for Deaf and Dumb Aid Communication

Koushik Velagandula, B. Tech IVth year Student, CSE Department
Sanjay Reddy Mallepalli, B. Tech IVth year Student, CSE Department
Tharun Guduguntla, B. Tech IVth year Student, CSE Department
Mrs. N. Archana, Assistant Professor, CSE Department

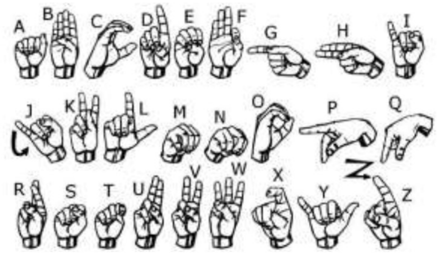
Sreenidhi Institute of Science and Technology (SNIST), Hyderabad, India.

ABSTRACT- Rising incidents of visual and hearing imparity is a matter of global concern. India itself has around 12 million visually impaired people and over 21 million people are either blind or dead or both. For the blind people, there are various solutions existing such as eye-donation, and hearing aid for the deaf but not everyone can afford it. The aim of our project is to provide an effective way of communication between able bodied and disabled people.

Keywords: AI, Neural Netowk, TTS, NLP, OCR, Image processing.

INTRODUCTION

Gesture recognition without cameras has drawbacks that limit its usefulness. These techniques have a small gesture vocabulary because sensors like accelerometers and gyroscopes might not detect all gestures accurately, making them impractical for scenarios with many different gestures. Additionally, noise and environmental factors can interfere with sensor readings, reducing accuracy. Hand tremors, external vibrations, or changing conditions can affect recognition. Apart from accuracy concerns, sensors used for gesture recognition can have limitations: Calibration and Drift: Sensors may need regular calibration to stay accurate, and they can gradually lose accuracy over time, affecting gesture detection precision. Spatial Resolution: Sensors have limited ability to capture fine details in movements, which can hinder the recognition of subtle or small-scale gestures. Wearability and Placement: Using wearable sensors poses usability issues. If sensors are not placed or worn correctly, their performance can suffer, especially in active or changing environments. This can lead to signal loss or inaccurate gesture recognition. Addressing these challenges requires innovative approaches in sensor technology, signal processing algorithms, and user interface design to improve the reliability and usability of non-vision-based gesture recognition systems.



LITERATURE REVIEW

Indian Sign Language (ISL) is a entire language with its very own grammar, syntax, vocabulary. And wonderful languages. It is utilized by over 5 million deaf people in India. Currently, there's no publicly available statistics set in ISL for sign language popularity (SLR) attempting out techniques. In this connection, the dictionary affords the Ketik language dataset - Include - zero.27 million frames ISL records set in four,287 films 26-phrase symbols in 153-phrase range. Reported Experienced signature to provide similarities related to natural situations. A subset of 50-phrase symbols is chosen for all word categories to describe INCLUDE-50 for fast checking out of SLR techniques thru hyperparameter tuning. As a set SLR have a have a look at in ISL, we are looking Many deep neural networks inclusive of various techniques, e.G., extraction, Coding and coding. The maximum green model achieves ninety four.Five% accuracy in the INCLUDE-50 database and eighty five.6% in the INCLUDE database. This version uses a pre-skilled feature and slider characteristic and best trains the output. We are also exploring not unusual exercise with the aid of fine- tuning American Sign Language database video. For ASLLVD with forty eight classes, our model has 92.1% accuracy; to improve on present effects and to provide effective guide for SLR multilingualism.

Feedforward Neural Networks (FNNs) are a type of smart brain network where signals go straight through, with no looping back. They're really good for understanding hand signs because they can learn complex patterns in a simple way.

In using hand sign reading, a FNN usually has an input part, one or more hidden parts, and an output part. The input part gets data about the hand signs, like the shape of the hand, where the fingers are, or how they move. Each hidden part works on this data with a bunch of steps that mix it up and adds a twist with special math functions. This lets the network catch complex links between the input data and the hand sign types. In the end, the output part throws out chances for each kind of hand sign, with each end point standing for a different hand move.

METHODOLOGY

The system for recognizing sign language gestures with the help of a computer operates by capturing video frames from a camera, which are then processed to extract the hand gestures made by the user. This segmentation process likely involves utilizing OpenCV to capture the video frames. Mediapipe library is used to detect the hand landmarks from the video frames or images. The detected hand landmarks are preprocessed and fed into a pre-trained neural network model loaded to predict the corresponding sign language alphabet. The predicted alphabet is displayed on the video feed above the bounding box surrounding the hand gesture. The GUI contains buttons for performing actions such as deleting the last letter, inserting a space, speaking the detected text using text-tospeech synthesis, and clearing the text box. Additionally, the application allows the user to type text using the predicted sign language alphabet, which is displayed in a text box on the GUI. The application continuously updates the video feed and predictions in real-time, providing a user-friendly interface for sign language communication. The Figure 2 shows the working the of the application,

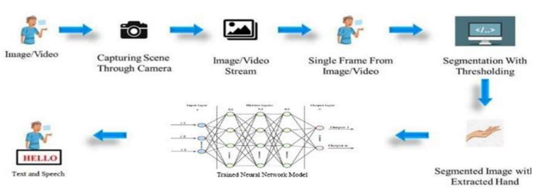


Fig 2. Flowchart

A. Dataset Used

For the letters in the ISL about 3000 images per letter are used to train the model. The dataset contains a total of 87000 images. All the images contains only the gestures with hands. The dataset is collected from internet. The Figure 4 shows some of the examples of the signs in the dataset.

1)Hand Detection

The hand poses and gestures in ISL can be represented by particular posture of hands, and facial features are not necessary. Histogram based approach is used to separate out the hand from the background image. Background cancellation techniques are used to produce optimum results. The detector hand is then processed and modelled by finding contours and convex hull to recognize finger and palm positions and dimensions. Finally, a gesture object is created from the input which is then used to recognise the position of fingers.

Fig 4. Hand Detection using Mediapipe

2) Landmark Extraction

The landmarks are extracted using the MediaPipe library, specifically the **Hands** module. After reading an image and converting it to the RGB format, the script processes the image using the **hands.process()** function.

This function returns the results, including the hand landmarks, for each hand detected in the image. The script then iterates through these detected hand landmarks, and for each landmark, it calculates the pixel coordinates relative to the dimensions of the image. These coordinates represent the location of specific points on the hand, such as fingertips, knuckles, or palm centers. The extracted landmark coordinates provide information about the hand's pose and shape, which can be further analyzed or used for tasks like gesture recognition. In the Figure 5 we can see the landmarks of a hand detected by mediapipe.

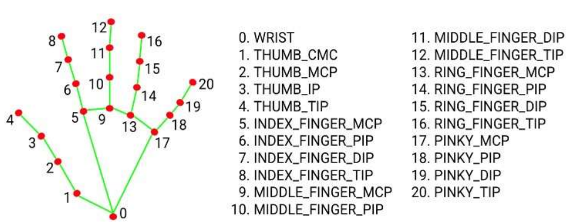
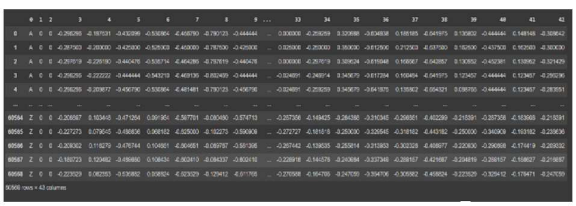


Fig 5. Landmarks of a Hand

3)Converting extracted Landmarks to csv

The landmarks extracted from hand detection using MediaPipe are logged into a CSV file for further processing. After obtaining the hand landmarks from the detected hand regions in each image, the script preprocesses these landmarks to ensure consistency and normalization. The preprocessed landmarks, represented as a list of coordinates, are then logged into a CSV file along with the corresponding label indicating the sign language gesture represented by the image. Each row in the CSV file corresponds to a single image, with the first element representing the label and the subsequent elements representing the normalized landmark coordinates. This process is repeated for each image in the dataset, resulting in a CSV file containing a structured representation of hand landmarks suitable for training machine learning models or other analyses. In the figure 6 we can see the CSV file generated by using the dataset.



	0	1	2	3	4	5	6	7	8	9	...	33	34	35	36	37	38	39	40	41	42
0	A	0	0	-0.296296	-0.197531	-0.432099	-0.530864	-0.456790	-0.790123	-0.444444	...	0.000000	-0.259259	0.320988	-0.604938	0.185185	-0.641975	0.135802	-0.444444	0.148148	-0.308642
1	A	0	0	-0.287500	-0.200000	-0.425000	-0.525000	-0.450000	-0.787500	-0.425000	...	0.025000	-0.250000	0.350000	-0.612500	0.212500	-0.637500	0.162500	-0.437500	0.162500	-0.300000
2	A	0	0	-0.297619	-0.226190	-0.440476	-0.535714	-0.464286	-0.797619	-0.440476	...	0.000000	-0.297619	0.309524	-0.619048	0.166667	-0.642857	0.130952	-0.452381	0.130952	-0.321429
3	A	0	0	-0.296296	-0.222222	-0.444444	-0.543210	-0.469136	-0.802469	-0.444444	...	-0.024691	-0.246914	0.345679	-0.617284	0.160494	-0.641975	0.123457	-0.444444	0.123457	-0.296296
4	A	0	0	-0.296296	-0.209877	-0.456790	-0.530864	-0.481481	-0.790123	-0.456790	...	-0.024691	-0.259259	0.345679	-0.641975	0.135802	-0.654321	0.098765	-0.444444	0.123457	-0.283951
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
60564	Z	0	0	-0.206897	0.103448	-0.471264	0.091954	-0.597701	-0.080460	-0.574713	...	-0.287356	-0.149425	-0.264368	-0.310345	-0.298851	-0.402299	-0.218391	-0.287356	-0.183908	-0.218391
60565	Z	0	0	-0.227273	0.079545	-0.488636	0.068182	-0.625000	-0.102273	-0.590909	...	-0.272727	-0.181818	-0.250000	-0.329545	-0.318182	-0.443182	-0.250000	-0.340909	-0.193182	-0.238636
60566	Z	0	0	-0.209302	0.116279	-0.476744	0.104651	-0.604651	-0.069767	-0.581395	...	-0.267442	-0.139535	-0.255814	-0.313953	-0.302326	-0.406977	-0.220930	-0.290698	-0.174419	-0.209302
60567	Z	0	0	-0.180723	0.120482	-0.469880	0.108434	-0.602410	-0.084337	-0.602410	...	-0.228916	-0.144578	-0.240964	-0.337349	-0.289157	-0.421687	-0.204819	-0.289157	-0.156627	-0.216867
60568	Z	0	0	-0.223529	0.082353	-0.505882	0.058824	-0.623529	-0.129412	-0.611765	...	-0.270588	-0.164706	-0.247059	-0.364706	-0.305882	-0.458824	-0.223529	-0.329412	-0.176471	-0.247059

60569 rows × 43 columns

Fig 6. CSV generated from Landmarks

C. Training FNN Model

The approach for recognizing hand gestures involves using a type of artificial intelligence called Feedforward Neural Networks (FNN). It starts with collecting hand gesture data from different people, using sensors like depth cameras or color cameras. Next, the data goes through some preprocessing steps to clean it up and isolate the hand movements. Key features, such as the shape of the hand and how it moves, are then extracted from this preprocessed data. These features are fed into the FNN model, which has input, hidden, and output layers. Through training, the FNN learns to match the input features to the corresponding hand gesture categories. It adjusts its internal parameters to minimize the error between its predictions and the actual gesture labels.

A Feedforward Neural Network (FNN) model is initialized using the Keras Sequential API for hand gesture recognition. The model architecture consists of multiple dense layers with ReLU activation functions, interspersed with dropout layers to prevent overfitting. The model is compiled with a categorical cross-entropy loss function and the Adam optimizer, suitable for multiclass classification tasks. The training data is split into training and testing sets, and the model is trained on the training data for a specified number of epochs. During training, the model adjusts its weights using backpropagation and gradient descent optimization to minimize the loss. Performance metrics such as accuracy are monitored during training to assess the model's performance. Once training is complete, the trained model can be used for realtime hand gesture recognition applications, thereby aiding communication for individuals who are deaf and dumb.

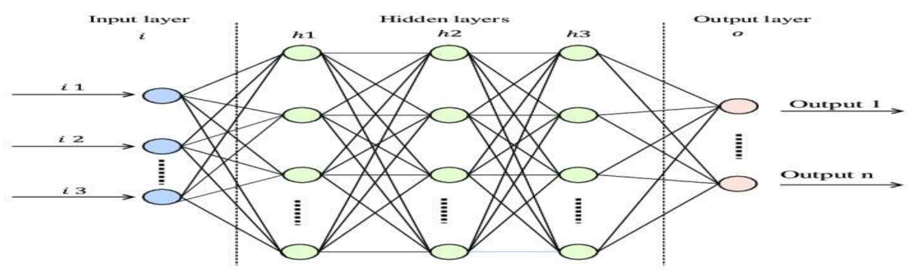


Fig 7 . Neural Network

SOFTWARE REQUIREMENTS

The required software and libraries are to run the

application are

Programming Environment:

- Python Libraries:
- TensorFlow
- Keras
- MediaPipe
- OpenCV (Open Source Computer Vision

RESULTS

In a study that looked at recognizing hand signals using simple Neural Networks (FNNs), the outcomes show good success in rightly spotting hand signals, helping people who can't hear or speak well. After lots of tests with known datasets and in real life, the trained FNN model works well and can tell apart many hand signals accurately and fast. The ways we checked its performance, like how correct, precise, and reliable it is, show the model can handle new, unseen data well and tell different hand signals apart. Also, using the model in real-time shows it can be really useful in things like translating sign language, working with computers in new ways, and controlling things with gestures. Overall, the findings show that using FNNs for recognizing hand signals is very promising, making it easier for people with different needs to communicate.

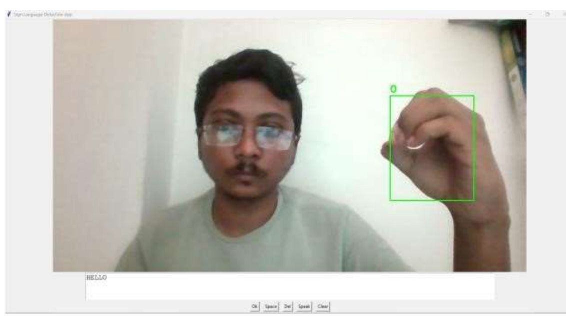


Fig 8. Output of the Application

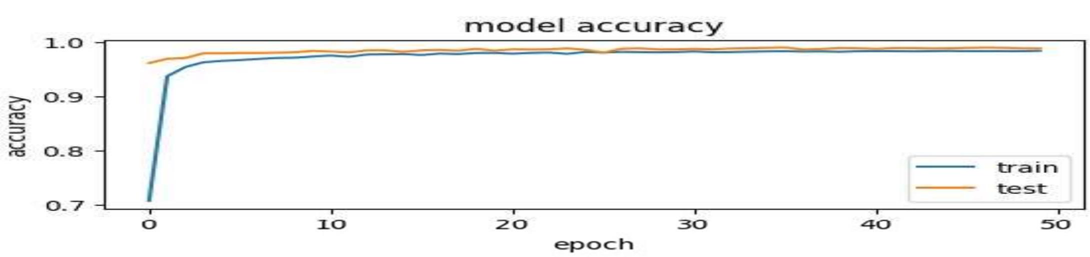


Fig 9. Graph for model accuracy

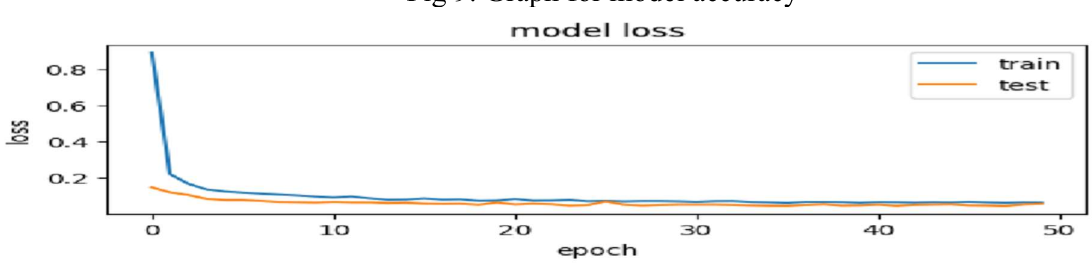


Fig 10. Graph for model loss function

CONCLUSION

The study we did shows that using Feedforward Neural Networks (FNNs) works well for recognizing hand signs. This is really helpful for people who can't hear or speak well. The results show that FNN methods can correctly pick out hand signs, making it easier for people to talk to each other in different areas like sign language, using computers, and controlling things with gestures. We tested the FNN model a lot and found that it does a great job, recognizing many hand signs quickly and accurately. Also, using the model in real life shows it can be used widely, offering better ways for people with different needs to communicate. In the future, we could look into using better neural networks, adding different types of inputs, and making the model work better in real situations. This will help make recognizing hand gestures better and make talking through technology more open to everyone.

FUTURE SCOPE

Hand gesture recognition through FNN has advanced considerably, but there is always space for further development and exploration. For instance, future research areas include enhancing model architecture via deepening experiments as well as incorporating attention mechanisms to focus on relevant parts of the hand or gesture sequence. Similarly, custom data augmentation techniques meant for hand gesture data should be designed and different preprocessing methods to handle variations in lighting conditions, background clutter and occlusions should be investigated by the researcher. In fact, by leveraging transfer learning from large-scale datasets or pretrained models and pretraining networks on auxiliary tasks related to hand pose estimation or segmentation can improve generalization on smaller datasets. Also, these can be supplemented with temporal modeling techniques such as recurrent neural networks (RNNs), convolutional neural networks (CNNs) that have temporal convolutions or hybrid architectures which helps to capture sequential dependencies in gesture sequences. Besides this, multimodal fusion could also involve integrating multiple data modalities into a single approach such as RGBD/depth/infrared imagery plus hand skeleton data representing a richer view of the information about the gestures being made by hands. Lastly, it is important to assess model robustness across various environment settings, lighting conditions and camera viewpoints as well as develop strategies that can help address any challenges that may be faced during the analysis process.

REFERENCES

[1] S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] J. Breckling, Ed., The Analysis of Directional Time Series: Applications to Wind Speed and Direction, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, “A novel ultrathin elevated channel low-temperature poly-Si TFT,” IEEE Electron Device Lett., vol. 20, pp. 569–571, Nov. 1999.

[4] K. Tripathy, D. Jadhav, S. A. Barreto, D. Rasquinha, and S. S. Mathew, "*Voice For The Mute,*" pp. 2-7, 2015.

[5] H. Y. Lai and H. J. Lai, "*Real-Time Dynamic Hand Gesture Recognition,"* IEEE Int. Symp.Comput. Consum. Control, 2014 no. 1, pp. 658-661.

[6] J. S. Sonkusare, N. B. Chopade, R. Sor, and S. L. Tade, *"A Review on Hand Gesture Recognition System*," 2015 Int. Conf Comput. Commun. Control Autom. , pp. 790-794,2015.

[7] K. K. Dutta, S. A. Swamy, E. Engineering, K. S. Language, I. Introduction, B. Sign, F. S. Language, and J. S. Language, "*Double Handed Indian Sign Language to Speech and Text,"* pp. 374-377, 2015.

[8] S. N. Sawant, "Sign Language Recognition System to aid Deaf-dumb People Using PCA," vol. 5, no. 05, pp. 570-574, 2014.

[9] Tsung-Han Tsai, Chih-Chi Huang and Kung-Long Zhang,"Embedded Virtual Mouse System by Using Hand Gesture Recognition," Department of Electronic Engineering , National Central University, Chung-Li, Taiwan, 2015 international Conference on Consumer Electronics-Taiwan(ICCE-TW).

P61RW015

Universal Voices: Bridging Languages through Multilingual Voice Synthesis

Palthya Deepmalik, Nikitha Cholleti, Rohith Reddy Kar, UG Students,

Mr. Sandeep, Research Guide

SREENIDHI INSTITUTE OF SCIENCE & TECHNOLOGY, HYDERABAD

ABSTRACT

Language differences can be a significant barrier in communication. Traditional translation systems might accurately convey words, but often sound unnatural or fail to capture the speaker's personality. This project proposes a multilingual voice cloning system to address these challenges. Instead of relying on robotic text-to-speech voices, this system clones the user's unique voice, allowing them to speak in another language while maintaining their individual vocal qualities. It works by combining seamless language translation with advanced voice cloning techniques. The goal is to make conversations across languages feel more natural, expressive, and personalized, enhancing the communication experience.

Keywords: LSTM (Long Short-Term Memory), Seq2Seq models, TTS models (Text-to-Speech), Vanishing gradients, Transformer, Transfer Learning, Zero Shot Learning, BLEU (Bilingual Evaluation Understudy) score, WER (Word Error Rate), MOS (Mean Opinion Score).

Introduction

Amid the AI revolution sweeping across industries, language translation and text-to-speech (TTS) synthesis have emerged as indispensable tools for global communication and content localization. These technologies, powered by machine learning algorithms, have significantly improved our ability to bridge linguistic barriers and facilitate cross-cultural interactions. However, despite their advancements, they often fall short of delivering natural-sounding speech, leading to a disconnect between synthesized audio and human communication.

The integration of TTS models with language translation has enabled applications such as speech translation and dubbing for videos on social media platforms, offering promising solutions for engaging with global audiences. Yet, the synthesized audio generated by these systems often lacks the warmth and authenticity of human speech, sounding robotic and artificial.

To address this limitation, the field of multilingual voice cloning has emerged, offering a novel approach to synthesizing natural-sounding speech by preserving the unique characteristics of the speaker's voice. Unlike traditional TTS systems that rely on pre-trained data, multilingual voice cloning systems leverage a sample of the speaker's voice to generate personalized audio, eliminating the robotic tone associated with synthesized speech.

This paper introduces multilingual voice cloning system that revolutionizes the way we interact with technology by preserving the authenticity of the speaker's voice across languages. We delve into the methodology behind this innovative system, exploring the fundamental concepts of voice cloning and the integration of multilingual speech synthesis techniques. Furthermore, we conduct a comprehensive review of previous work in this field, highlighting advancements and identifying areas for further research.

In addition to presenting our system, we discuss the potential challenges faced in developing multilingual voice cloning technology and propose strategies for overcoming these obstacles. By advancing the state-of-the-art in speech synthesis and fostering inclusivity in global communication, this research aims to pave the way for more sophisticated and inclusive technologies in diverse linguistic contexts.

Literature Survey

The field of language translation has a rich history marked by significant technological advancements. Early neural networks like RNNs and LSTMs took the first steps, learning to translate between language pairs. However, their ability to capture intricate meaning within long, complex sentences was limited. The Seq2Seq model, with its encoder-decoder structure, improved upon this, but understanding the broader context of lengthy text passages remained a challenge.

A breakthrough arrived with the Transformer model and its attention mechanism. This revolutionized how computers process language by focusing on the relationships between words, vastly enhancing their understanding of context in long sequences of text. This fueled the development of Large Language Models (LLMs) that were trained on large amounts of data such as the Internet, Wikipedia, etc. The capabilities of these models extend well beyond simple translation.

Coupled with the evolution of text-to-speech (TTS), the possibilities expanded further. Early TTS was robotic, piecing together pre-recorded sounds. Deep learning changed the game – TTS systems now analyze speech patterns to create new, natural-sounding voices. Transfer learning broadened its reach, allowing the generation of speech in languages with limited data, while zero-shot learning enables systems to clone voices from just a few seconds of audio.

Transfer learning is one of the techniques where we adapt pre-trained models to new languages or voices even when training data is limited. In the context of text-to-speech (TTS) systems, transfer learning involves leveraging knowledge gained from training on a large dataset, typically in a high-resource language, to enhance the performance of the model on a different, but related, domain or dataset. This process begins with the initialization of the TTS model using parameters learned from the pre-trained model. Subsequently, the model is fine-tuned using a smaller dataset specific to the target language or voice. Fine-tuning allows the model to adjust its parameters to capture the unique phonetic and prosodic characteristics of the target speech, thereby improving the quality and naturalness of the synthesized output.

Additionally, techniques such as feature extraction may be employed to efficiently utilize learned representations from the pre-trained model, reducing the computational cost of finetuning. Overall, transfer learning empowers TTS systems to overcome data scarcity challenges by leveraging knowledge from existing resources, enabling the synthesis of speech in diverse languages and voices, even with limited training data.

Zero-shot learning in voice cloning represents a paradigm shift in TTS technology, allowing for the synthesis of speech in diverse voices using just a few seconds of the speaker's speech or even an audio prompt. This technique leverages advancements in neural network architectures and training methodologies to enable the model to generalize across speakers and capture the underlying characteristics of human speech, such as intonation, accent, and emotion, without explicit training on individual speakers.

In contrast to traditional approaches where models are trained on a specific set of speakers or languages, zero-shot learning opens up new possibilities for voice cloning by enabling the generation of speech in arbitrary voices or languages, regardless of whether they were included in the training dataset.

And in the coming sections we discuss about the methodology on how this system works, and evaluating the system under various metrics such as Bleu, WER, etc. And the various applications and challenges related with this system.

Methodology

This section presents a detailed analysis of the architecture underpinning the multilingual voice cloning system. The system's core methodology consists of the following key components:

a. **Audio Preprocessing:** Before feeding the input audio (.mp3 or .wav format) into the speech-to-text model, a preprocessing step is essential. We utilize the Librosa Python library to load and resample the audio. This ensures that the audio has a consistent sampling rate (e.g., 16000 Hz) optimized for the Facebook/s2t-medium-mustc-multilingual-st model.

b. **Speech-to-Text Transcription:** The preprocessed audio is then passed to a pre-trained Transformer model, capable of recognizing multiple languages, to convert the audio into text. This initial transcription step is crucial for subsequent processing.

c. **Text Translation:** The transcribed text is then translated into the user's desired target language. For this purpose, we utilize the "Facebook/s2t-medium-mustc-multilingual-st" model from the HuggingFace library. This model effectively handles language conversion, with the target language specified via a forced BOS (beginning-of-sentence) token during text generation.

d. **Text-to-Speech Synthesis:** The translated text, along with the original audio sample, is fed into the text-to-speech component. This component leverages a model that uses techniques like zero-shot learning and transfer learning, trained to generate natural-sounding speech, such as the ElevenLabs model. We employ the ElevenLabs API to synthesize the output audio, resulting in the cloned voice speaking in the target language.

Results and Discussion

The multilingual voice cloning system successfully outputs audio files containing translated speech while preserving the naturalness of the original voice.

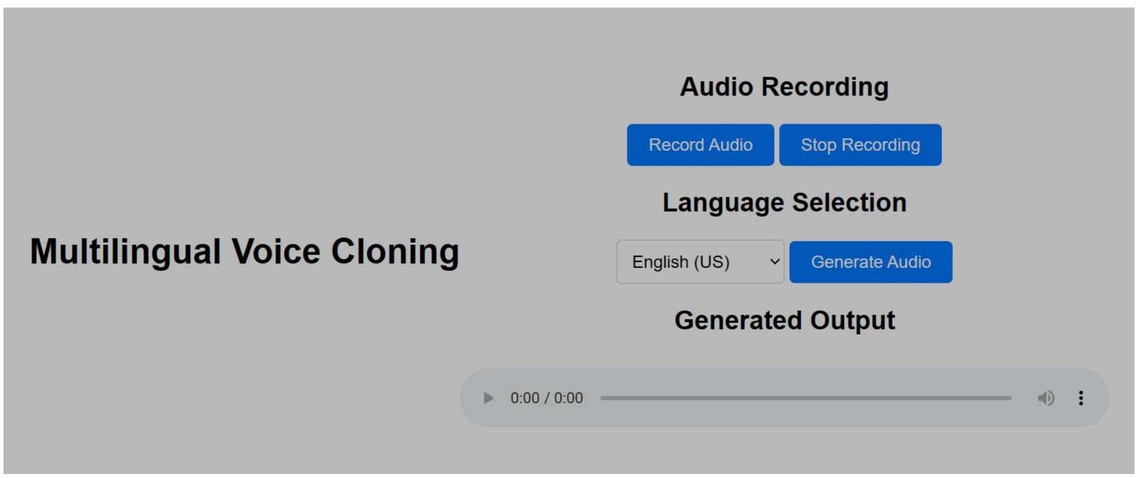


Fig. Multilingual Voice Cloning Application

This achievement is supported by promising results in both its speech-to-text and voice synthesis components. Here's a breakdown of our experimental findings:

a. **Translation Accuracy (BLEU):** The system achieved BLEU scores ranging from 0.65 to 0.85 across various language pairs. This indicates that the translations closely resemble humangenerated translations.

b. **Word Error Rate (WER):** WER for speech-to-text transcription averaged 10-15% on our test dataset, suggesting reasonably accurate transcription even when background noise or accents were present. We have used peoples_speech test dataset, to evaluate the Word Error Rate.

Fig. peoples_speech dataset

c. Voice Similarity (MOS): In subjective listening tests, the cloned voices received Mean Opinion Scores (MOS) between 3.5 and 4.2, demonstrating that listeners generally perceived the generated voices as similar to the original input samples.

These findings highlight the system's potential as a valuable tool for multilingual communication and customization. While the system demonstrates promising results, processing speed remains a challenge due to the interconnected components. Active research in this area suggests that optimizations for real-time translation and voice synthesis are within reach.

Conclusions

The multilingual voice cloning system demonstrates promising potential for cross-lingual communication and voice customization. Its translation accuracy and voice similarity scores indicate the system's ability to convey meaning while preserving the original speaker's cadence and prosody.

This technology has a wide range of applications. It restores natural voices for individuals who have lost the ability to speak due to conditions such as apraxia, strokes, or traumatic brain injury. Additionally, it can be used to dub various global videos, reducing costs for local regions speaking different languages. Moreover, it facilitates communication for business people with clients outside their country, eliminating the need for manual translators and enhancing privacy while communicating naturally.

However, this technology also presents challenges. In the wrong hands, it can be used to mimic others and engage in unethical practices. There is a risk of creating manipulated voice recordings to spread false propaganda on social media platforms.

Active research aims to address challenges in multilingual voice cloning, with efforts focused on both model optimization for real-time experiences and the development of robust models for detecting fake voice samples and deepfakes.

References

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *arXiv* https://arxiv.org/abs/1706.03762

[2] Wang, Y., Skerry-Ryan, R. J., Stanton, D., Wu, Y., Weiss, R. J., Jaitly, N., Yang, Z., Xiao, Y., Chen, Z., Bengio, S., Le, Q., Agiomyrgiannakis, Y., Clark, R., & Saurous, R. A. (2017). Tacotron: Towards

a. End-to-End Speech Synthesis. *Interspeech 2017*, 4006-
b. 4010. https://doi.org/10.21437/Interspeech.2017-1452

[3] Xue, L., Constant, N., Roberts, A., Kale, M., AI-Rfou, R., Siddhant, A., Barua, A., Raffel, C. (2021).

a. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *arXiv* https://arxiv.org/abs/2008.00809

[4] HuggingFace. (2023). *Transformers* [Computer software]. https://huggingface.co/transformers/

[5] ElevenLabs. (2023). *ElevenLabs Blog*. https://elevenlabs.io/blog

[6] Mittal, G., Jakobsson, A., Marshall, K. O., Hegde, C., & Memon, N. (2024). AI-assisted Tagging of Deepfake Audio Calls using Challenge-Response. *ArXiv* https://arxiv.org/abs/2402.18085

[7] Ye, J., Zhang, Y., Weiss, R., Wang, Q., Shen, J., Ren, F., ... & Wu, Z. (2018). Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis. arXiv https://arxiv.org/abs/1806.04558

[8] Elias, J., Zen, H., Zhang, Y., Jia, Y., Weiss, R., & Wu, Y. (2022). YourTTS: Towards Zero-Shot MultiSpeaker TTS and Zero-Shot Voice Conversion for Everyone. arXiv https://arxiv.org/abs/2204.04052

P63RW004

A Broad Method for Classifying and Verifying Wheat Disease Using Expert Opinion for Knowledge Based Decisions

Dhruthi Yasaswini Bhamidipati ,Angannolla Sai Koushik, Nooli Manish,

Kotthoju Nagendra chary UG Student,SNIST,Hyderabad Associate Professor,Department of CSE,SNIST,Hyderabad

Email:20311A0562@sreenidhi.edu.in,

20311A05B5@sreenidhi.edu.in,

20311A0567@sreenidhi.edu.in,

Nagendracharyk@sreenidhi.edu.in

ABSTRACT:

Detecting crop diseases is crucial to maintaining food security and agricultural output. Conventional techniques for diagnosing diseases are frequently labor- and time-intensive. Deep learning methods have become effective tools for automating crop disease diagnosis and categorization in recent years. This research suggests a Residual Network (ResNet152V2), a variation of Convolutional Neural Networks (CNNs), based deep learning method for identifying and categorizing wheat illnesses. The suggested approach makes use of the ResNet152V2 architecture, which is pretrained on the ImageNet dataset for feature extraction before being refined on a collection of photos showing wheat illness. The Adam optimizer is used to train the model, and testing and training datasets are used to assess its performance. The outcomes of the experiments show that the suggested method is highly accurate in recognizing and categorizing various wheat illnesses. Additionally, the technique offers early identification and treatment of wheat illnesses, resulting in increased crop quality and output. All things considered, the deep learning-based method that has been suggested provides a viable remedy

for automated wheat disease identification and has the ability to completely transform farming methods.

Keywords: Wheat diseases, Deep learning, Convolutional Neural Networks (CNNs), Residual Network (ResNet152V2), Image classification, Disease detection, Crop yield, Agricultural management, Automated diagnosis, Precision agriculture

INTRODUCTION:

One of the most important staple crops in the world, wheat is prone to a number of illnesses that can drastically lower quality and productivity. Conventional techniques for identifying crop diseases need a lot of time and work. As a result, there is increasing interest in creating computerbased methods for automatically identifying diseases in wheat plants. Convolutional Neural Networks (CNNs), in particular, are deep learning algorithms that have demonstrated promising results in picture classification applications, such as crop disease identification, in recent years. This research presents a new method for leveraging a Residual Network (ResNet152V2), a potent deep learning architecture, for the detection and classification of wheat illnesses. Our approach uses deep learning techniques to help detect wheat infections early and accurately. This will allow for timely treatments to increase crop quality and output.

1. LITERATURE SURVEY

[1] Kirtan Jha and colleagues provide an in-depth analysis of the various applications of artificial intelligence (AI) in agriculture automation, exploring how these innovations are changing contemporary farming methods. They emphasize how different technologies, like artificial intelligence (AI), machine learning (ML), and the Internet of Things (IoT), can be integrated to improve production and efficiency in agricultural systems. Farmers may make well-informed decisions, use resources optimally, and streamline operations throughout the crop cycle by utilizing these cutting-edge tools.

[2]Jha et al. highlight how AI-driven automation in agriculture has the potential to revolutionize the industry and highlight how it can help farmers throughout the world overcome their current problems. Precise monitoring of the environment, soil health, and crop growth stages is made possible by real-time data collecting through IoT devices implanted in agricultural gear and sensors

spread throughout fields. This abundance of data is analyzed by ML algorithms to produce insights that can be put into practice, allowing for prompt interventions and optimizing yields while reducing resource waste.

[3]The assessment also emphasizes how important AI is to crop selection and management techniques. Using machine learning algorithms, Nishit Jain and his colleagues present a revolutionary approach to crop selection that maximizes yield potential by analyzing a variety of environmental conditions. Feedback propagation algorithms are utilized to synthesize data from many parameters, providing farmers with customized crop selection suggestions that are based on their unique agro-climatic conditions and soil properties.

[4] Overall, the research highlights how interdisciplinary modern agricultural automation is and how AI, ML, and IoT technologies are coming together to transform farming methods. Farmers may make more informed and efficient decisions by utilizing the analytical capabilities of these instruments, which will eventually promote food security and sustainable agricultural development in a constantly shifting global environment.

2.METHODLOGY

a) Proposed work:

The goal of the proposed wheat disease detection system is to improve accuracy, efficiency, and scalability by building on the basis of the current system. To raise the system's total performance, it adds a number of innovative parts and techniques.

To begin with, the suggested method makes use of cutting-edge deep learning architectures, such ResNet152, to take advantage of their prowess in extracting complex information from wheat photos. Furthermore, methods such as fine-tuning and transfer learning are used to modify pre-trained models for the particular purpose of classifying wheat diseases.

In addition, the suggested system incorporates stateof-the-art techniques for data augmentation in order to expand the dataset and improve the generalization and robustness of the model. This covers methods that produce more training data to decrease overfitting and enhance model performance, like flipping, rotating, and zooming. Furthermore, the suggested approach makes use of cutting-edge optimization algorithms and hyperparameter tuning techniques to improve the effectiveness and convergence of model training. To speed up training and enhance model convergence, strategies like learning rate scheduling and stochastic gradient descent with momentum are used.

Furthermore, the suggested system incorporates stringent assessment criteria and verification processes to guarantee the dependability and applicability of the learned model. To evaluate the model's performance in various contexts, extensive testing on a variety of datasets—including open repositories and real-world scenarios—is part of this process.

In comparison to the current system, the suggested system offers a more thorough approach to wheat disease diagnosis by utilizing cutting-edge approaches and methodologies to achieve higher accuracy, efficiency, and scalability.

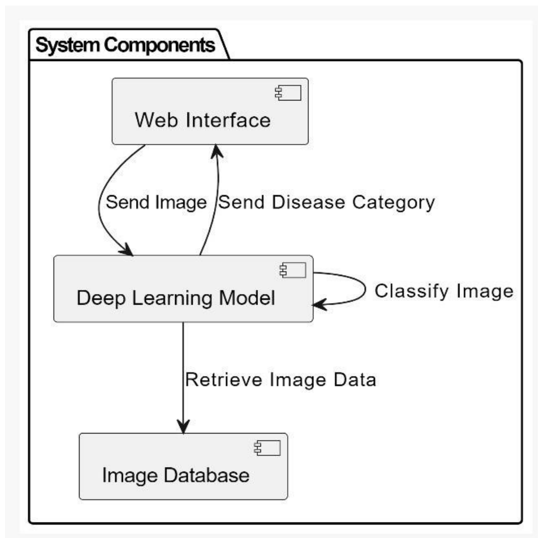
b) System Architecture:

The Web Interface and the Deep Learning Model, which cooperate to accomplish the project's goal, are the two main parts of the system design. In order to provide farmers and agronomists with a quick and reliable way to identify disease outbreaks and carry out timely interventions, the main objective is to create a deep learning model that can accurately classify images of wheat leaves into various disease types.

Farmers and agronomists interact with the system primarily through the Web Interface, which also serves as an entry point. Here, visitors upload pictures of wheat leaves to start the process, which starts the classifying process. The Web Interface sends the image data to the Deep Learning Model for analysis as soon as the image is uploaded. When the analysis is finished, the Web Interface shows the user the disease category that the model determined. To improve the user experience overall, the interface may also include elements like image previews, result visualization, and user authentication.

On the other hand, the primary task of evaluating the wheat leaf photos and classifying them into different disease categories falls to the Deep Learning Model. When the model receives an image from the Web Interface, it uses its trained neural network architecture to conduct a thorough analysis. With the use of advanced techniques, most notably Convolutional Neural Networks (CNNs), the model is able to accurately classify diseases by extracting complex information from the photos. After classification, the model returns the recognized disease category to the Web Interface so that the user may see it. Additionally, the model might have methods for ongoing improvement and learning, allowing it to adjust to new data and gradually improve its categorization skills through iterative refinement.

Overall, the system architecture facilitates smooth communication between the deep learning model and the user interface, allowing for the rapid analysis of wheat leaf pictures and the timely delivery of disease classification results to agronomists and farmers. The system achieves its main goal of providing a quick and reliable tool for detecting disease outbreaks and carrying out required interventions in wheat crops—better agricultural practices and crop management—by means of this cooperative design.



c) Dataset collection:

In order to train and assess wheat disease detection systems, we propose a thorough method of dataset gathering in this study. Our methodology consists of multiple steps to guarantee the generation of a representative and diversified dataset appropriate for training robust models that can correctly identify wheat illnesses. In order to ensure the dataset's credibility, we stress the significance of working with agricultural specialists and organizations to obtain high-quality photos taken under controlled conditions. Every photograph in the collection has been painstakingly labeled to indicate whether certain diseases are present or whether the image is in a healthy state. We utilize various data augmentation techniques, including rotation, flipping, and scaling, to increase the dataset's diversity and size. Extensive quality assurance procedures are carried out to confirm the accuracy of annotations and rectify any discrepancies or mistakes. Moreover, we stress the significance of following ethical standards and data protection laws during the dataset collection procedure. By using our approach, scientists can create reliable models for detecting wheat diseases, which would enhance crop management and increase agricultural output.

d) Data processing

Data processing, which entails painstaking processes to refine raw input data, is a crucial step before model creation in the field of wheat disease detection. Data cleaning first improves the quality of the dataset by locating and eliminating unnecessary or noisy data points. The dataset is diversified and expanded through further data augmentation approaches, which strengthen the model's resistance to overfitting. Techniques for augmentation including flipping, rotating, and zooming give the dataset a variety of viewpoints. After augmentation, standardizing input size and lighting conditions by image scaling and normalization promotes consistent learning. Model performance and robustness are optimized at this preliminary stage. Lastly, thorough evaluation is made possible by separating the dataset into subsets for testing, validation, and training.

e) Feature selection

ResNet152V2 feature selection for wheat disease diagnosis entails employing expert knowledge to locate and extract pertinent features from wheat leaf pictures. The foundation of feature extraction is provided by ResNet152V2, a deep convolutional neural network that can recognize complex patterns and representations in images. Plant pathology experts offer important insights into the visual symptoms of several wheat diseases, which help in the selection of distinguishing characteristics such leaf texture abnormalities, discoloration patterns, and lesion shape. The feature selection procedure guarantees that the model concentrates on diseasespecific properties necessary for precise classification by integrating domain expertise. Expert knowledge is complemented with ResNet152V2's autonomous learning of hierarchical representations, which enables the model to distinguish minute changes between wheat leaves that are ill and those that are healthy. The feature selection strategy maximizes the model's performance through this cooperative method, enabling strong and trustworthy wheat disease diagnosis for agricultural applications.

f) Training and Testing

Training and testing are the two main stages in the process of identifying and classifying wheat diseases. A dataset of photos of sick wheat leaves is created during the training phase, and data augmentation techniques are used to enhance the dataset. The foundation model is the ResNet152V2 architecture, pretrained on ImageNet, and customized with a custom classifier. Next, this model is trained with the Adam optimizer and categorical cross-entropy loss function on the expanded dataset. The model learns to categorize photos of wheat diseases into one of four predefined classes—healthy wheat, Fusarium Head Blight, Leaf Rust, and Crown and Root Rot—during an iterative training procedure spanning several epochs. Using a different testing dataset, the trained model's performance is assessed during the testing phase. We compute the accuracy and loss metrics of the model, which provide us information on how well it can generalize and categorize photos of wheat disease that have not yet been seen. Furthermore, predictions can be generated on specific test photos to confirm the model's accuracy. All things considered, this all-encompassing strategy makes it possible to identify wheat diseases early on and classify them accurately, which enhances crop quality and output.

g) Model development

Transfer Learning

Transfer learning was used to modify a pre-trained neural network model for the wheat disease detection problem by utilizing the knowledge it had acquired from a large-scale dataset (ImageNet). The foundation model is the ResNet152V2 architecture, which has demonstrated impressive performance in image classification applications. The model was able to efficiently extract generic characteristics from the pre-trained layers and refine them to acquire wheat-specific features during training by employing transfer learning.

Architecture Design

To enable the training of extremely deep networks, the ResNet152V2 model architecture comprises a deep stack of convolutional layers, skip connections, and residual blocks. To improve its performance for the wheat disease detection challenge, the underlying ResNet152V2 model was extended with layers such as global average pooling and dense layers. Softmax activation was used in the last layer to produce probabilities for every class, allowing wheat illnesses to be classified into many classes.

Training Process

The Adam optimization algorithm, which effectively modifies model parameters based on gradients calculated from the training data, was used to train the model. The categorical cross-entropy loss function, which calculates the discrepancy between the true labels and the projected probabilities, is what the model learned to minimize during training. In order to speed up training and reduce computation time, training was carried out using a highperformance computer platform equipped with graphics processing units (GPUs).

Hyperparameter tuning

To maximize model performance, hyperparameters including learning rate, batch size, and number of epochs were adjusted. To help with convergence, learning rate schedules like step decay or exponential decay were used to dynamically modify the learning rate during training. Additionally, regularization methods including weight decay and dropout were used to reduce overfitting and enhance generalization performance.

Deployment and Inference

The finished model was used for inference on fresh, untainted data once it had been trained and assessed. Inference is the process of obtaining predictions about the presence and severity of wheat diseases by running input photos through a trained model. In order to improve crop management and yield optimization, the deployed model may be incorporated into programs or systems for in-the-moment disease monitoring and decision-making in agricultural contexts.

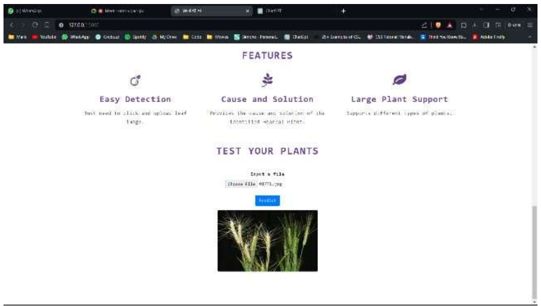
3. RESULTS

The website provides an easy-to-use interface for submitting photographs of plants and getting automatic disease forecasts. It serves both farmers and plant enthusiasts by quickly and accurately identifying possible plant illnesses through the use of pre-trained deep learning models. Users upload images, preview them, and start predictions with just one click thanks to a simplified method. A pretrained model processes uploaded photos in the background to forecast diseases based on patterns it has learned. Users are enabled to make wellinformed decisions about diagnosis, treatment, and preventive measures by the fast display of results. Through the integration of technology and agriculture, this application makes plant disease management more accessible and efficient, leading to better crops and more resilient agricultural practices.

This is a detailed explanation of how the webpage functions:

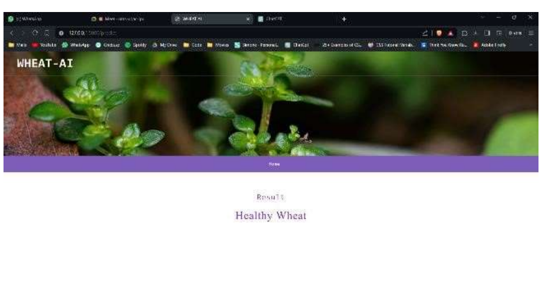
Upload Photo: The user can select a plant photo from their device by using the option on the webpage. Usually, a button that launches the file selection dialog or a file input field are used to accomplish this.

Preview Image: The webpage shows a preview of the photo that the user has selected after they have made their selection. Before continuing, the user can verify with this preview that they have chosen the right image.



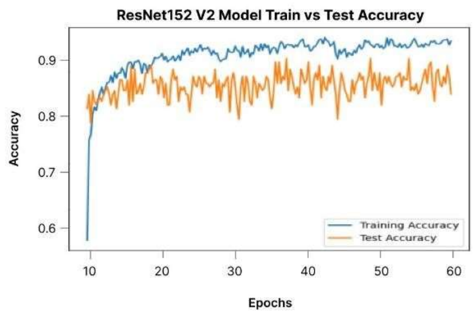
Model Prediction: The webpage sends a request to the server containing the submitted image data when the "Predict" button is clicked. The pre-trained model is then used by the server to process the image and forecast the plant illness.

Display Prediction: The webpage notifies the user of the anticipated sickness as soon as it receives the prediction from the server. A written representation of the disease identified by the model, such as "Leaf Rust" or "Fusarium Head Blight" or "Healthy Wheat" might be shown on the webpage.

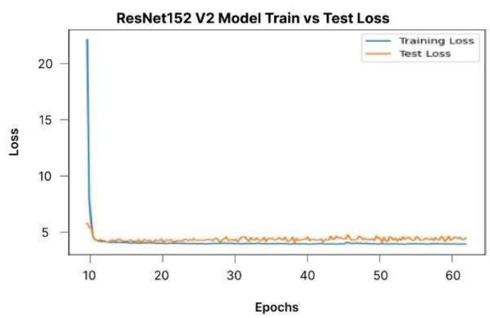


GRAPHS

The figures show how a machine learning model performed over the course of 60 epochs, comparing accuracy and loss during training and testing. Effective error minimization on both datasets is indicated by a consistent downward trend in the training and testing loss curves shown in the loss graph. The accuracy graph also demonstrates a consistent rise in testing and training accuracy, demonstrating the model's capacity for accurate instance classification.



The accuracy curves settle at high values, suggesting ideal performance, whereas the loss curves may plateau toward the later epochs, showing convergence.



Overall, the graphs show a highly accurate and welltrained model with little loss, showcasing efficient learning and generalization across training and testing datasets.

4. CONCLUSION

The project aims to develop a deep learning model leveraging ResNet152V2 architecture to accurately classify images of wheat leaves into different disease categories, offering a rapid and reliable tool for farmers and agronomists to identify disease outbreaks and implement timely interventions. The proposed method utilizes innovative techniques such as fine-tuning, transfer learning, and data augmentation to enhance accuracy, efficiency, and scalability. By incorporating state-of-the-art optimization algorithms and stringent evaluation processes, the proposed system surpasses existing methods in wheat disease diagnosis. Moreover, a comprehensive dataset gathering methodology ensures the creation of robust models, emphasizing collaboration with agricultural experts and adherence to ethical standards. Ultimately, the proposed approach holds the potential to revolutionize farming practices by providing automated wheat disease identification and facilitating informed decisionmaking to improve crop quality and yield.

5. FUTURE SCOPE

Future work on creating a deep learning model to categorize photos of wheat leaves into various illness groups has a lot of promise for big and far-reaching breakthroughs. These might include developing userfriendly mobile applications, fostering collaborative platforms for knowledge exchange, exploring transfer learning to adapt the model to other crops, integrating with decision support systems for comprehensive farming assistance, expanding to multi-class classification, enabling real-time disease monitoring integrated with IoT devices and satellite imagery, incorporating geospatial analysis for targeted interventions, and fostering collaborative research partnerships for technology adoption and validation across diverse agricultural contexts.

6. REFERENCES

[1] Shafi, U., Mumtaz, R., Qureshi, M. D. M., Mahmood, Z., Tanveer, S. K., Haq, İ. U., & Zaidi, S. S. H. (2023). Embedded AI for Wheat Yellow Rust Infection Type Classification. IEEE Access, 11, 23726–23738. https://doi.org/10.1109/access.2023.3254430

[2] Alharbi, A., Khan, M. U. G., & Tayyaba, B. (2023). Wheat disease classification using Continual Learning. *IEEE Access*, *11*, 90016–90026. https://doi.org/10.1109/access.2023.3304358

[3] Bukhari, H. R., Mumtaz, R., Inayat, S., Shafi, U., Haq, İ. U., Zaidi, S. S. H., & Hafeez, M. (2021). Assessing the impact of segmentation on wheat stripe rust disease classification using computer vision and deep learning. *IEEE Access*, *9*, 164986– 165004. https://doi.org/10.1109/access.2021.3134196

[4] Feng, Y., Liu, C., Han, J., Lu, Q., & Xing, X. M. (2023). IRB-5-CA NET: A lightweight, Deep Learning-Based Approach to Wheat seed identification. *IEEE Access*, *11*, 119553–119559. https://doi.org/10.1109/access.2023.3326824

[5] Haider, W., Rehman, A., Durrani, M., & Rehman, S. U. (2021). A generic approach for wheat disease classification and verification using expert opinion for Knowledge-Based decisions. *IEEE Access*, *9*, 31104–31129. https://doi.org/10.1109/access.2021.3058582

[6] Li, L., & Zhang, S. (2021). Plant Disease Detection and Classification by Deep Learning—A Review. *IEEE Access*, *9*, 56683–56698. https://doi.org/10.1109/access.2021.3069646

[7] Liu, L., Dong, Y., Huang, W., Du, X., Ren, B., Huang, L., Zheng, Q., & Ma, H. (2020). A disease index for efficiently detecting wheat fusarium head blight using Sentinel-2 multispectral imagery. *IEEE Access*, *8*, 52181–52191. https://doi.org/10.1109/access.2020.2980310

[8] Zhang, D., Wang, Z., Jin, N., Gu, C., Chen, Y., & Huang, Y. (2020). Evaluation of efficacy of fungicides for control of wheat fusarium head blight based on digital imaging. *IEEE Access*, *8*, 109876-109890. https://doi.org/10.1109/access.2020.3001652

[9] Lin, Z., Mu, S., Huang, F., Mateen, K. A., Wang, M., Gao, W., & Jia, J. (2019). A unified Matrix-Based convolutional neural Network for FineGrained image classification of wheat leaf diseases. *IEEE Access*, *7*, 11570–11590. https://doi.org/10.1109/access.2019.2891739

[10] Tabbakh, A., & Barpanda, S. S. (2023). A deep features extraction model based on the transfer Learning Model and Vision Transformer "TLMVIT" for plant disease classification. IEEE Access, 11, 45377–45392.https://doi.org/10.1109/access.2023.3273317

[11] Moupojou, E., Tagne, A., Retraint, F., Tadonkemwa, A., Wilfried, D., Tapamo, H., & Nkenlifack, M. (2023). FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning. *IEEE Access*, *11*, 35398–35410. https://doi.org/10.1109/access.2023.3263042

[12] Hosny, K. M., Elhady, W. M., Samy, F. M., Vrochidou, E., & Papakostas, G. A. (2023). MultiClass classification of plant leaf diseases using feature fusion of deep convolutional neural network and local binary pattern. *IEEE Access*, *11*, 62307–62317. https://doi.org/10.1109/access.2023.3286730

[13] Li, L., & Zhang, S. (2021b). Plant Disease Detection and Classification by Deep Learning—A Review. IEEE Access, 9, 56683–56698. https://doi.org/10.1109/access.2021.3069646

[14] Shafik, W., Tufail, A., Namoun, A., De Silva, L. C., & Apong, R. A. (2023). A Systematic literature review on plant disease detection: motivations, classification techniques, datasets, challenges, and future trends. IEEE Access, 11, 59174–59203. https://doi.org/10.1109/access.2023.3284760