iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VII Month of publication: July 2025

DOI: https://doi.org/10.22214/ijraset.2025.72888

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

Enhancing Healthcare through Multi Disease Prediction using Machine Learning

Anjali U
Information Technology, Jawaharlal Nehru Technological University Hyderabad, UCESTH

Abstract: This study presents a unified machine learning-based system for predicting multiple diseases diabetes, Parkinson's disease, and heart disease through a single interface. Support Vector Machine (SVM) is used for diabetes and Parkinson's prediction, while Logistic Regression handles heart disease classification. The models are trained on publicly available datasets and integrated into an interactive web app using Streamlit. This approach enhances diagnostic efficiency, supports early detection, and provides a scalable solution to assist clinical decision-making.
***Keywords**: Machine Learning, Disease Prediction, Support Vector Machine (SVM), Logistic Regression, Streamlit.*

I. INTRODUCTION

Modern healthcare systems are increasingly relying on intelligent technologies to support diagnosis and treatment. Among these, machine learning stands out for its ability to learn from medical data and assist in identifying health conditions. As the demand for quick and reliable health assessments grows, there is a clear need for automated solutions that can predict the risk of multiple diseases from basic input information. This project introduces a digital tool that applies machine learning techniques to forecast the possibility of having diabetes, Parkinson's disease, or heart disease, all through a single application interface. The solution aims to provide a supportive tool for early detection and improve overall healthcare accessibility.

The rise of machine learning in healthcare has revolutionized the way diseases are detected and predicted. As medical records and clinical data become more extensive and complex, conventional diagnostic approaches struggle to process and evaluate this information efficiently. Many existing systems are tailored to identify only one illness at a time, which can be limiting when patients present symptoms linked to multiple conditions. By applying machine learning, it's possible to analyze various health parameters at once, enabling faster and more comprehensive diagnosis that supports better patient outcomes.

Chronic illnesses such as diabetes, Parkinson's disease, and heart disorders are widespread and often challenging to manage without early detection. Early identification significantly increases the chances of effective treatment. However, most digital diagnostic tools are limited to detecting a single disease at a time. This constraint motivated the creation of a unified system capable of screening for multiple conditions simultaneously. By leveraging machine learning, this project aims to improve the speed, simplicity, and accessibility of disease prediction.

This project aims to develop a web-based machine learning application to predict the likelihood of three major diseases: diabetes, Parkinson's disease, and heart disease. Support Vector Machine (SVM) is used for predicting diabetes and Parkinson's disease, while Logistic Regression is applied for heart disease, based on the characteristics of each dataset. The system is deployed through a user-friendly interface built with the Streamlit framework, ensuring ease of access and usability.

II. LITERATURE SURVEY

[1] Swaroop Sana (2024)

Sana's work proposes a machine learning-based solution for predicting common chronic diseases such as diabetes, Parkinson's, and heart disease. The paper highlights how combining SVM and logistic regression models with a simple user interface can improve accessibility and diagnostic accuracy. The system's design supports non-expert users, aiming for broader usability in early disease detection.

[2] Xie et al. (2021)

This survey reviews various deep learning approaches used in health diagnostics. The authors explore how CNNs and RNNs manage diverse clinical data sources, identifying trends across multiple illnesses. Besides evaluating performance, they stress limitations related to privacy, interpretability, and the lack of standard datasets, noting these as key barriers in real-world deployment.

[3] Chen et al. (2020)
This study investigates how machine learning techniques can be leveraged to analyze large-scale health data effectively. The authors examine the challenges posed by inconsistent and unbalanced datasets commonly found in real-world medical environments. They explore the use of various supervised learning models, including decision trees, support vector machines, and ensemble methods like random forests. In addition, the research highlights the importance of efficient computation and secure data handling when building predictive systems for clinical applications.
[4] Choi et al. (2016)
In this paper, Choi and team introduce "Doctor AI," a recurrent neural network model that predicts a patient's future diagnoses based on historical health data. The innovation lies in its sequential learning method, which accounts for the temporal nature of clinical visits. It shows potential for enhancing proactive care by alerting clinicians to emerging risks earlier.

III. OBJECTIVE

The main objective of this project is to develop an integrated machine learning system capable of predicting multiple diseases specifically diabetes, Parkinson's disease, and heart disease through a single platform. The system aims to enhance early detection by applying suitable algorithms like Support Vector Machine (SVM) and Logistic Regression based on the nature of the datasets. Additionally, the goal is to make the tool accessible and user-friendly by deploying it as a web application using Streamlit, thereby supporting faster clinical decision-making and improving patient outcomes.

IV. SYSTEM ANALYSIS

A. Existing System

Most existing health prediction systems detect only one disease at a time, such as diabetes, Parkinson's, or heart disease. They often use basic algorithms like decision trees or logistic regression and lack multi-disease integration. Data preprocessing is minimal, focusing mainly on simple cleaning without handling missing or noisy data, which limits their accuracy and effectiveness.

B. Proposed System

The proposed system addresses these issues by integrating the prediction of diabetes, Parkinson's disease, and heart disease into a single, unified framework. It employs Support Vector Machine (SVM) algorithms for diabetes and Parkinson's disease predictions, and Logistic Regression for heart disease, chosen based on dataset characteristics and diagnostic effectiveness. The system incorporates advanced preprocessing steps, including missing data handling, feature scaling, and feature selection, to enhance model performance and reliability.

V. METHODOLOGY

1) Data Collection and Preprocessing
- Datasets for diabetes, Parkinson's, and heart disease were obtained from publicly available sources.
- Data preprocessing included handling missing values, feature scaling using StandardScaler, and splitting into training and testing sets using train_test_split.

2) Model Selection and Training
- Support Vector Machine (SVM) was used for diabetes and Parkinson's prediction due to its accuracy with complex data.
- Logistic Regression was applied to heart disease prediction as it suits binary classification problems well.
- Models were trained using scikit-learn.

3) Model Evaluation
- Accuracy was calculated on both training and testing datasets using accuracy_score.
- Predictions were tested with sample inputs to validate real-time use.

4) Model Saving
- Trained models were saved using the pickle library so they could be loaded later without retraining.

5) Web Interface Development
- Streamlit was used to create an interactive web interface for users to input their data.
- The interface connects user input to the corresponding model and displays the disease prediction.

6) Integrated Prediction System

- A single interface handles predictions for all three diseases. The system is designed to be scalable and user-friendly.

VI. RESULTS AND ANALYSIS

A. Results

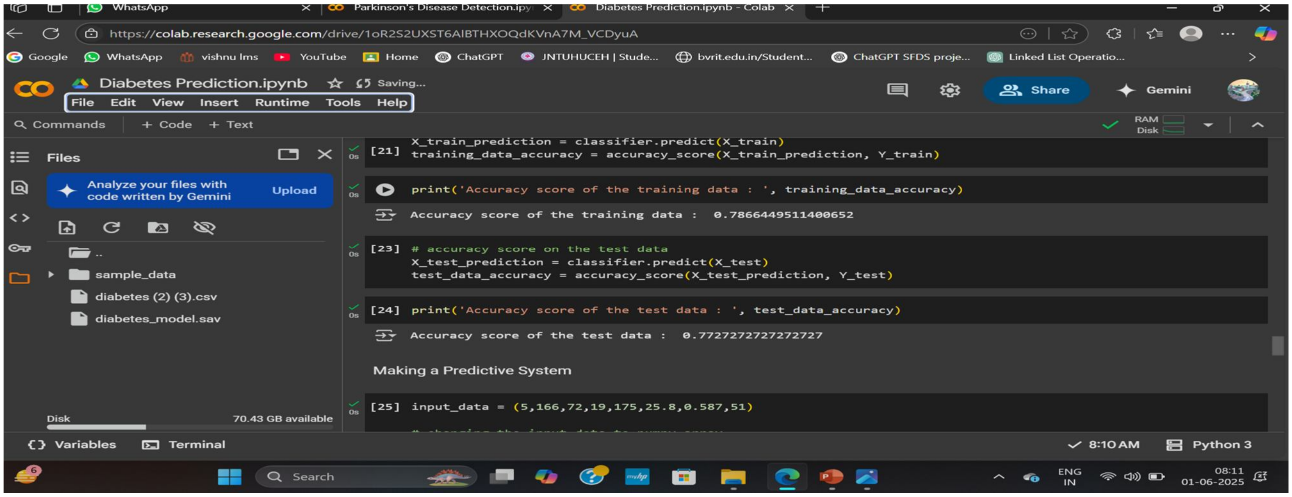


Fig.1 Accuracy Score of Diabetics Disease

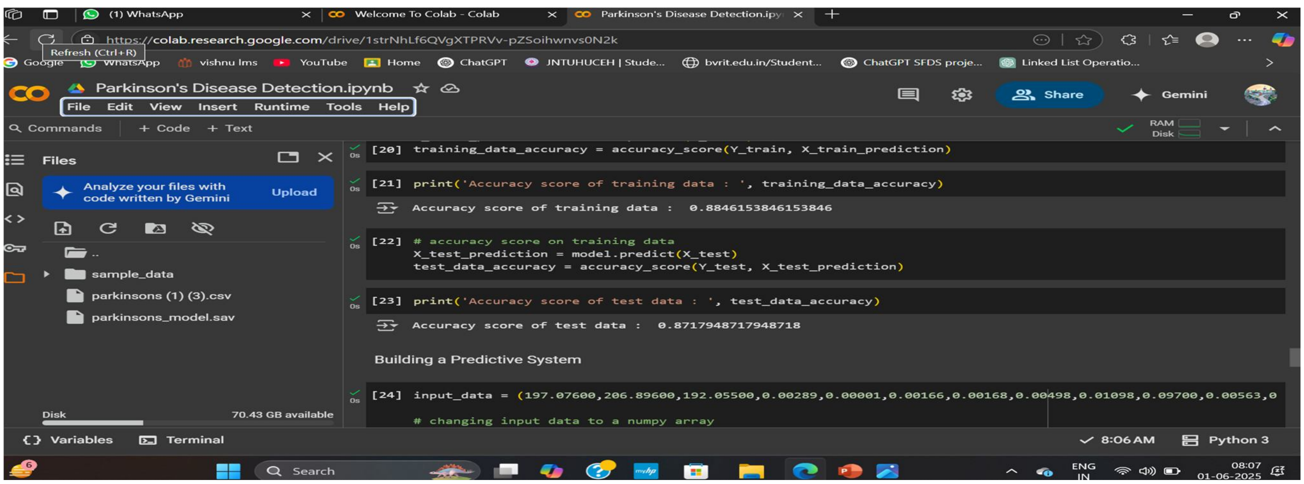


Fig.2 Accuracy Score of Parkinsons Disease

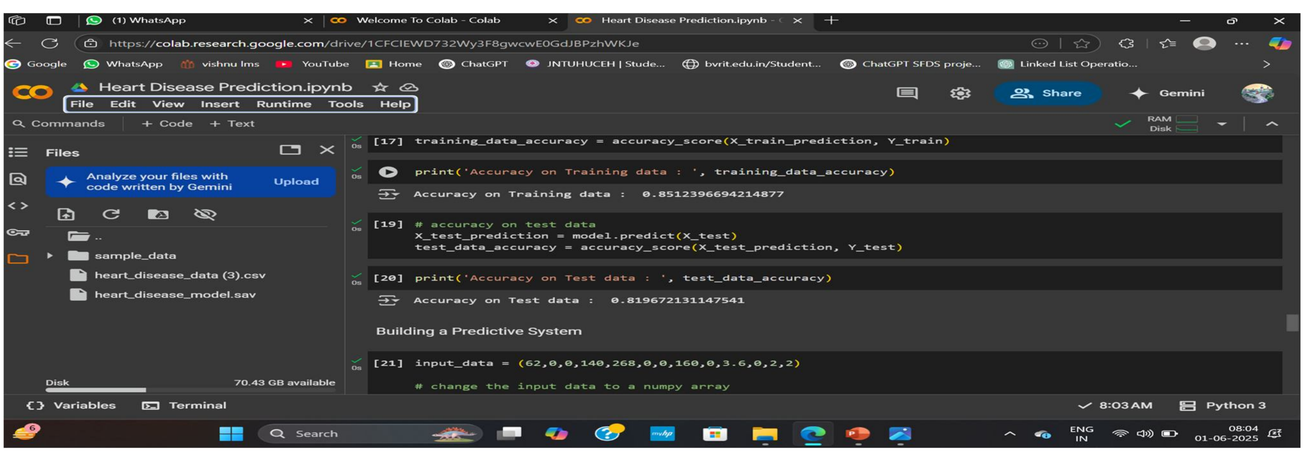


Fig.3 Accuracy Score of Heart Disease

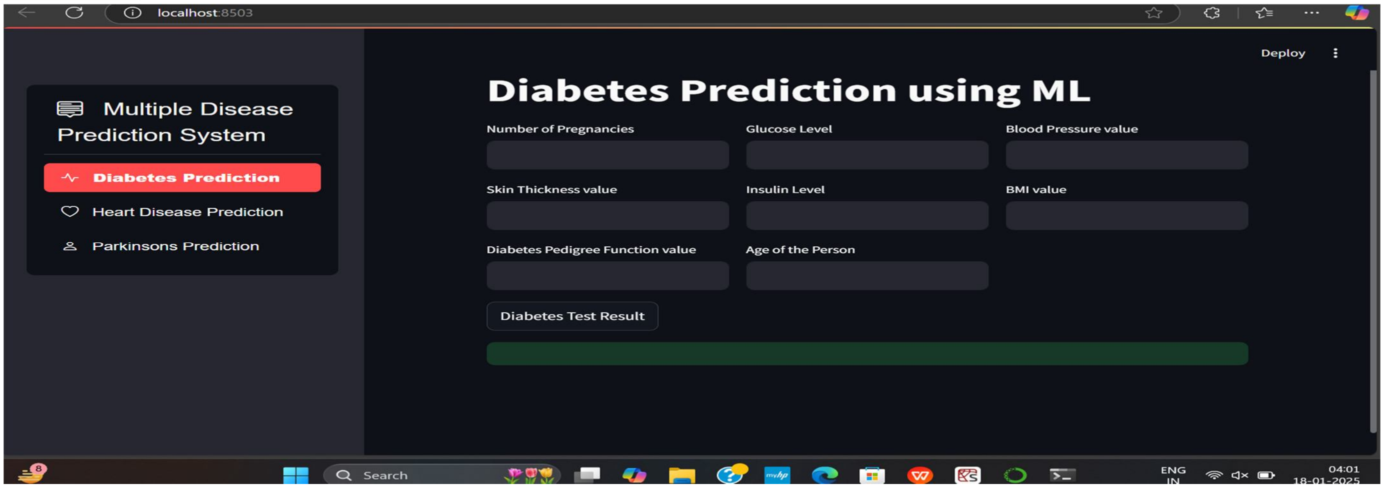


Fig.4 Diabetics Disease Prediction

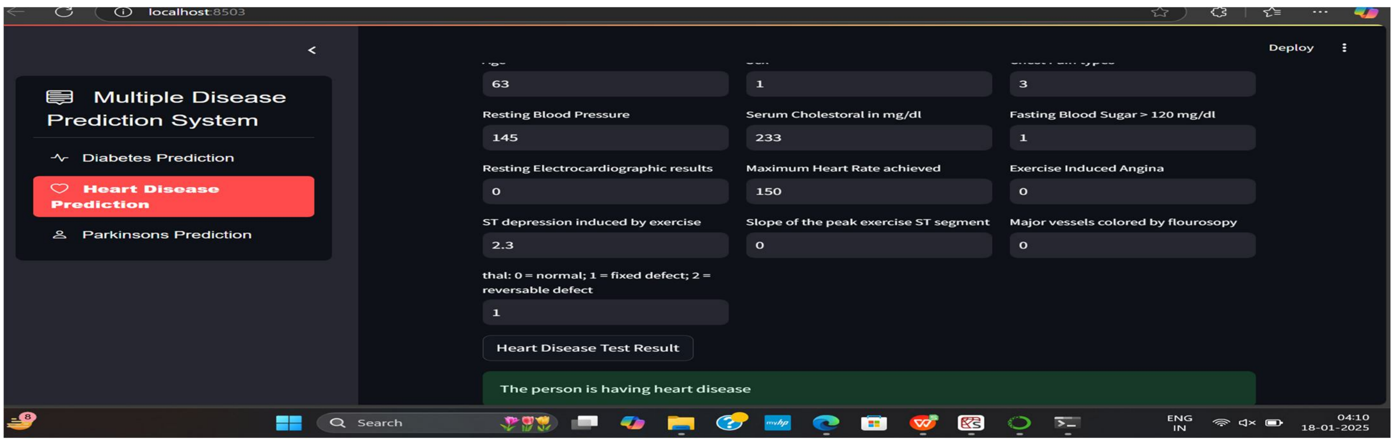


Fig.5 Heart Disease Prediction

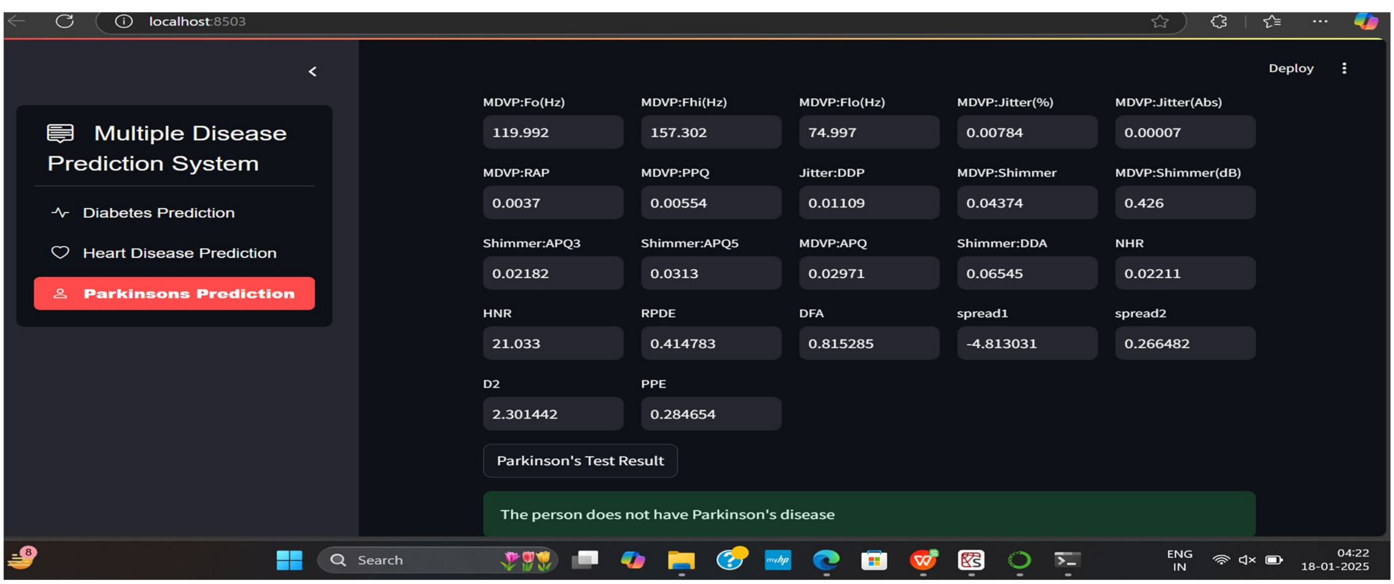


Fig.6 Parkinsons Disease Prediction

B. Analysis

SNo	Disease Name	Algorithm Name	Existing System Accuracy	Proposed System Accuracy
I	Diabetes	SVM Classifier	76%	77%
II	HeartDisease	Logistic Regression	80%	81%
III	ParkinsonsDisease	SVM Classifier	72%	87%

Table 1.Accuracy Score of Diseases

VII. CONCLUSION AND FUTURE SCOPE

This project successfully demonstrates the development of a machine learning-based system capable of predicting multiple diseases diabetes, Parkinson's disease, and heart disease using a unified platform. By applying Support Vector Machine (SVM) and Logistic Regression algorithms, the system offers reliable and accurate predictions. The use of a Streamlit-based interface allows for easy interaction, making the system accessible even to non-technical users. Overall, this approach improves early detection of diseases and supports quicker medical intervention, making it a valuable tool in modern healthcare.

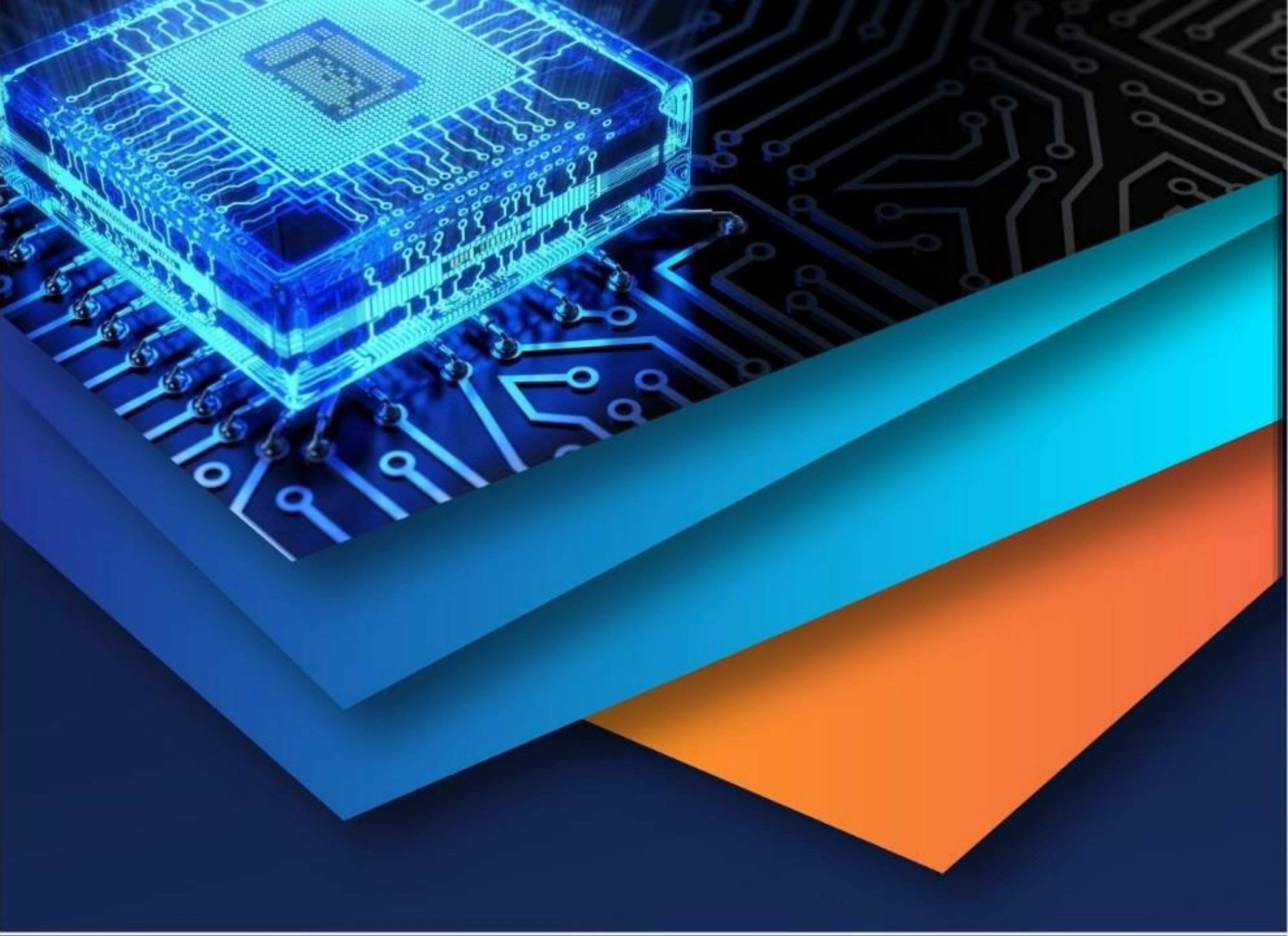
There are several opportunities to enhance this disease prediction system in future versions. Additional diseases can be integrated into the platform, making it more comprehensive. The system can also be improved by incorporating real-time patient data from wearable devices or medical records. Using advanced algorithms like deep learning could further improve accuracy. In addition, deploying the model on cloud platforms will allow for wider access and better scalability, making it usable in hospitals, clinics, and rural health centers.

VIII. ACKNOWLEDGEMENT

I sincerely thank Sri G. Narasimham sir for his continual support in helping me carry out this project. Also extend my heartfelt gratitude to the authors of the research papers reviewed in this study.

REFERENCES

[1] Sana, S., & Swaroop. (2024). Implementation of a Multi-Disease Prediction System Using Machine Learning Techniques. International Research Journal of Modernization in Engineering Technology and Science, ISSN: 2582-5208. https://doi.org/10.56726/IRJMETS49550
[2] Xie, S., Yu, Z., & Lv, Z. (2021). A Survey on Deep Learning Methods for Multi-Disease Prediction. Computer Modeling in Engineering & Sciences, 128(2), 489–522.
[3] Chen, M., Hao, Y., Cai, Y., Wang, Y., & Song, J. (2020). Utilizing Big Health Data for Disease Prediction Using Machine Learning Approaches. IEEE Access, 7, 44669–44678.
[4] Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., & Sun, J. (2016). Doctor AI: Applying Recurrent Neural Networks to Predict Clinical Events. Journal of Machine Learning Research, 17(1), 3011–3030.
[5] Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., ... & Ng, A. Y. (2018). Evaluating the CheXNeXt Algorithm for Chest Radiograph Diagnosis Using Deep Learning in Comparison to Radiologists. PLOS Medicine, 15(11), e1002686.
[6] Singh, K., Sharma, A., Verma, A., Maurya, R., & Perwej, Y. (2024). Development of a Machine Learning-Based System for Predicting Multiple Diseases. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 10, 673–684. https://doi.org/10.32628/CSEIT24103217



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY