International Journal of Innovative Research in Engineering
Volume 5, Issue 3 (May-June 2024), PP: 112-115.
www.theijire.com **ISSN No: 2582-8746**

Multi Disease Prediction System Using Machine Learning

Niyati Gaur[1], Simran Fatima[2], Raghvendra Pratap SIngh[3], Himanshu Patel[4]
[1]Assistant Professor, Computer Science and Engineering, School of Management Sciences, Lucknow, Uttar Pradesh, India.
[2,3,4] Final Year Students, Computer Science and Engineering, School of Management Sciences, Lucknow, Uttar Pradesh, India.

How to cite this paper:
Niyati Gaur[1], Simran Fatima[2], Raghvendra Pratap SIngh[3], Himanshu Patel[4], "Multi Disease Prediction System Using Machine Learning", IJIRE-V5I03-112-115.

Abstract: *The Multi Disease Prediction System (MDPS) utilizes machine learning techniques, particularly Logistic Regression and Support Vector Machines (SVM), to forecast various illnesses including diabetes, heart disease, and Parkinson's disease. It employs a user-friendly interface developed with Streamlit, enhancing accessibility. By analysing standard health indicators such as blood pressure, pulse rate, cholesterol, and heart rate, the system aims to improve early detection and enable personalized healthcare. The model demonstrates notable accuracy and precision, aiding in the identification of critical risk factors associated with these ailments. This paper highlights the essential role of machine learning in simultaneously predicting multiple diseases, addressing a current gap in technological capabilities. The project underscores the advancement of disease prediction through sophisticated machine learning methodologies.*

Key Word: *Streamlit, Machine Learning, Diabetes, Heart Disease, Parkinson's Disease, SVM, Logistic Regression.*

I.INTRODUCTION

Recent strides in machine learning have significantly impacted healthcare, introducing algorithms capable of predicting multiple diseases simultaneously, potentially revolutionizing medical diagnostics and patient outcomes. While existing models typically analyse individual diseases such as diabetes, cancer, or skin conditions, this research endeavours to develop a system utilizing Support Vector Machines (SVM) to predict three interrelated diseases: heart disease, diabetes, and Parkinson's disease, based on user-provided health parameters. Various machine learning algorithms, including Decision Tree, Random Forest, and SVM, are integrated into a user-friendly web application developed with the Streamlit library. This approach not only facilitates interactive application development but also enhances disease prediction accuracy. SVM's versatility in managing both linear and nonlinear data renders it particularly suitable for complex medical diagnostic tasks. By harnessing these technologies, this study establishes a multi-disease prediction framework evaluated for its effectiveness in diagnosing Parkinson's disease, diabetes, and heart disease. The utilization of machine learning in this context propels us towards more accurate, timely, and personalized healthcare interventions, thereby improving patient outcomes and optimizing healthcare systems.

II.MATERIAL AND METHODS

Predicting different diseases at once is complex and requires a detailed methodology. The approach for the Multi Disease Prediction project is outlined as follows:

1) **Data Collection:** Sources such as electronic health records (EHRs), medical literature, and public health databases are utilized to collect data specifically concerning diabetes, heart illness and Parkinson's illness.
2) **Data Preprocessing:** This step involves cleaning and transforming the raw data to make it suitable for analysis by machine learning algorithms.
3) **Model Selection:** In this stage, appropriate machine learning algorithms are identified and selected based on their suitability for predicting each specific disease. These algorithms are then trained with the pre processed data and evaluated using metrics like accuracy and precision to optimize performance.
4) **Data Splitting:** The data is divided into training and testing subsets. Machine learning models are trained on the training subset and their efficacy is tested on the testing subset.
5) **Deployment and Integration:** The final step involves deploying the trained models and integrating them into a cloud-based interactive web application. This application allows users to input specific health parameters and receive disease predictions, thereby improving both accessibility and user engagement.

Functional Requirement:

The system features an intuitive interface enabling users to input details about their symptoms, medical history, demographic information, and other relevant data. Subsequently, the system processes this information through built-in prediction models and presents forecasted diseases in a user-friendly format.

Non Functional Requirement:

The system must transparently indicate the range of values or confidence intervals associated with predicted disease outcomes. This provision of information empowers users to gauge the uncertainty of predictions, aiding them in making informed decisions regarding further medical advice or interventions. Additionally, the system should demonstrate reliability and consistency in its performance.

III.PROBLEM STATEMENT

Traditionally, disease prediction models operate in silos, focusing on individual conditions like heart disease, diabetes, or Parkinson's disease. This segmented approach often leads to fragmented analyses. The Multi Disease Prediction System aims to revolutionize this paradigm by harnessing machine learning to concurrently predict the likelihood of a patient developing multiple diseases. By leveraging sophisticated machine learning techniques, this system endeavours to accurately assess the risk of various diseases based on a patient's medical history and symptoms. The primary objective is to identify at-risk individuals early, particularly those vulnerable to conditions such as heart disease, kidney disease, and diabetes, thereby facilitating timely medical interventions to improve healthcare outcomes and patient management.

IV.EXISTING SYSTEM

The current system specializes in forecasting diabetes, heart disease, and Parkinson's disease through various machine learning techniques, including Naive Bayes, Decision Trees, Random Forest, Support Vector Machines (SVM), and Logistic Regression. Deployed using the Streamlit library, these models offer an accessible interface for users. Notably, SVM has exhibited significant efficacy, achieving 76% accuracy in predicting diabetes and 71% in Parkinson's disease. Complementary techniques like Logistic Regression and Decision Trees capitalize on specific data features, ensuring precise predictions. This setup underscores the potential of machine learning algorithms in disease prediction and hints at avenues for future enhancements to boost prediction accuracy.

V.PROPOSED SYSTEM

The proposed methodology entails employing multiple predictive models, rigorously evaluating their performances against each other. Key libraries such as pandas for data management, numpy for mathematical operations, scikit-learn for model training and evaluation, and pickle for model persistence are utilized. This approach facilitates the simultaneous prediction of diseases, enhancing user experience by streamlining the prediction process and potentially reducing mortality rates. With faster predictions and myriad other benefits over current systems, this method promises to elevate healthcare outcomes.

VI.RESULTS AND DISCUSSION

The realms of disease diagnosis and prediction stand on the brink of significant advancement through the integration of machine learning (ML) technologies. Accurate diagnoses are paramount for effective illness management and treatment. In our framework, we leverage the SVM algorithm for predictions. Patients input specific data into the system, which then assesses disease likelihood. The interface provides necessary value ranges and alerts for any incorrect or missing inputs. The predictive accuracy of our system hinges on the support vector machine algorithm's ability to yield precise outcomes, particularly with linear datasets.

Fig 7.1, Fig 7.2, Fig 7.3: Display the user interface (UI) alongside symptoms and predictions.

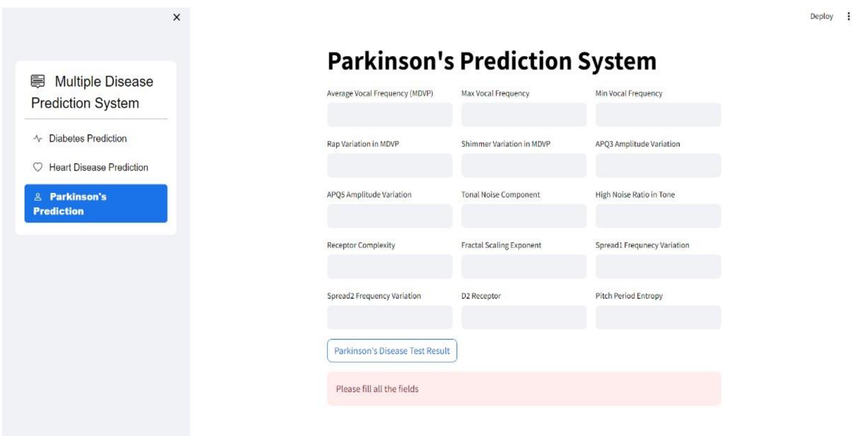

fig 7.1 Parkinson's Page

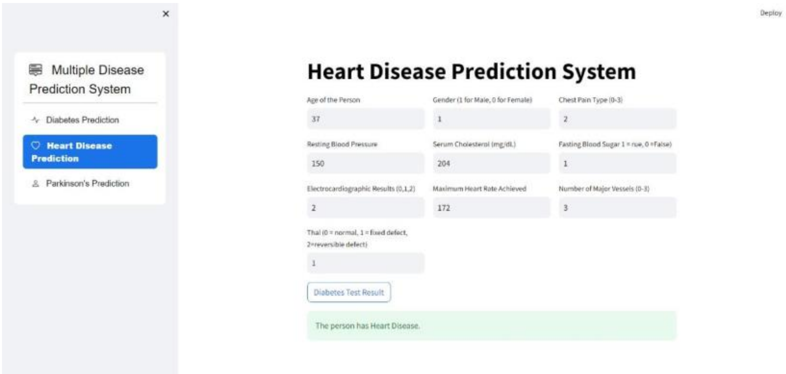

Fig 7.2 Heart Disease Page

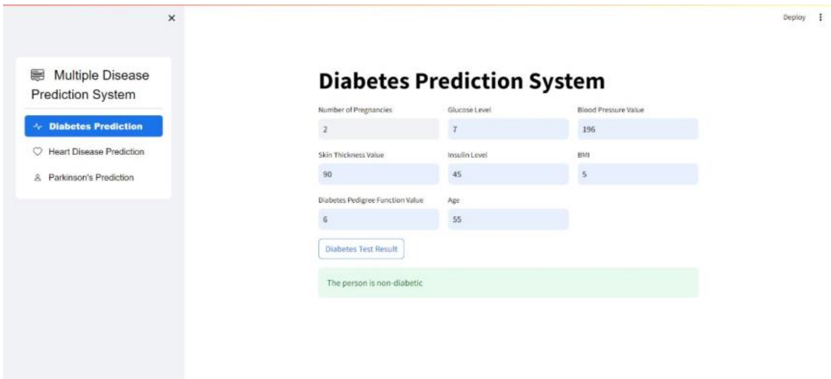

fig 7.3 Diabetes Page

VII.CONCLUSION

The objective of this study was to create a system that predicts multiple diseases accurately, removing the necessity for users to consult numerous websites. Detecting diseases early can increase life expectancy and reduce financial burdens. Various machine learning techniques, such as logistic regression and SVM, were employed to predict heart disease, diabetes, and Parkinson's disease. Through the SVM model, a framework for predicting multiple diseases was developed. Data was obtained from Kaggle.com and subjected to thorough preprocessing to ensure quality. The SVM algorithm achieved a 78% accuracy in predicting diabetes and 89% accuracy for Parkinson's disease, while logistic regression achieved 85% accuracy in predicting heart disease. These results highlight the potential of machine learning in disease prediction and improving patient outcomes. The implementation of the SVM model required careful handling and filtering. Integrating the trained model into an application allows for real-world disease prediction, benefiting researchers, healthcare providers, and individuals in making informed health decisions. Machine learning models enable targeted disease management, personalized treatment plans, and proactive interventions, enhancing patient care and optimizing resource allocation in healthcare. Additionally, they offer opportunities for population-level disease surveillance, potentially enabling early outbreak detection and preventive measures. In summary, this study showcases the effectiveness of SVM models in predicting multiple diseases and contributes to the advancement of machine learning in disease diagnosis, leading to more precise and timely healthcare solutions that enhance patient outcomes and healthcare efficiency.

VIII.ACKNOWLWDGEMENT

We extend our sincere thanks to Prof. Niyati Gaur, our mentor, for her valuable guidance and support throughout our project. Her assistance at every stage and her motivational support were crucial in enabling us to complete our work effectively. Additionally, we appreciate the Department of Computer Science and Engineering, School of Management Sciences, for approving our project in our chosen area of specialization.

References

[1] Laxmi Deepthi Gopisetti, Srinivas Karthik Lambavai Kummera, Sai Rohan Pattamsetti, Sneha Kuna, Niharika Parsi, Hari Priya Kodali, "Multiple Disease Prediction Model by using Machine Learning and Streamlit" 2023 IEEE, 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)

[2] Akkem Yaganteeswarudu, "Multi Disease Prediction Model by using Machine Learning" 2020 IEEE, 5th International Conference on Communication and Electronics Systems (ICCES)

[3] Elsevier B.V," Diabetes Prediction Using Machine Learning" 2019, International Conference on Recent Trends in Advanced Computing.

[4] KM Jyoti Rani, "Diabetes Prediction Using Machine Learning" July 2020, International Journal of Scientific Research in Computer Science Engineering, and Information Technology

[5] Firdous, Shimoo, Wagai, Gowher A, Sharma, Kalpana, "A survey on diabetes risk prediction using machine learning approaches", November 2022, Journal of Family Medicine, and Primary Car.

[6] Krittanawong, C. Virk, H. U., Bengaluru, S., Wang, Z., Johnson, K. W., Pinotti, R., Zhang, H., Kaplin, S., Narasimhan, B., Kitai, T., Baber, U., Halperin, J. L., & Tang, W. H. (2020). Machine learning prediction in cardiovascular diseases.

[7] Chaimaa Boukhatem, Heba Yahia Youssef, Ali Bou Nassif. February 2022 IEEE, Advances in Science and Engineering Technology International Conferences (ASET)

[8] Supriya Kamoji, Dipali Koshti, Valiant Vincent Dmello, Alrich Agnel Kudel, Nash Rajesh Vaz, Prediction of Parkinson's Disease using Machine Learning and Deep Transfer Learning from different Feature Sets, July 2021 IEEE, 6th International Conference on Communication and Electronics Systems (ICCES).

[9] Rohit Surya, A.T., Yaswanthram, P., Nair, P.R., Rajendra Prasath, S.S., Akella, S.V. (2022). Prediction of Parkinson's Disease Using Machine Learning Models—A Classifier Analysis. In: Bianchini, M., Piuri, V., Das, S., Shaw, R.N. (eds) Advanced Computing and Intelligent Technologies. Lecture Notes in Networks and Systems, vol 218. Springer, Singapore. https://doi.org/10.1007/978-981-16-2164-2_35.

[10] Makarious, M. B., Leonard, H. L., Vitale, D., Iwaki, H., Sargent, L., Dadu, A., Violich, I., Hutchins, E., Saffo, D., Kim, J. J., Song, Y., Maleknia, M., Bookman, M., Nojopranoto, W., Campbell, R. H., Hashemi, S. H., Botia, J. A., Carter, J. F., Craig, D. W., . . . Nalls, M. A. (2022). Multi-modality machine learning predicting Parkinson's disease. Npj Parkinson's Disease, 8(1), 1-13. https://doi.org/10.1038/s41531-022-00288-w.