iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62250

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Analysis and Prediction of Multiple Disease using Machine Learning

Jagadevi Puranikath[1], Fathima Farheen[2], Aruna Patil[3], Aishwarya K.M.[4]
Ballari Institute of Technology And Management Ballari

***Abstract:** This paper uses the machine learning methods for predicting diabetes, heart disease, and Parkinson's disease based on user-provided input data. Its primary goal is to enhance early disease detection and prompt medical intervention using personalized predictions. By training models on relevant datasets, accurate predictions are achieved across various diseases, expanding our understanding and predictive capabilities in healthcare. The research's broad scope contributes significantly to addressing multiple medical conditions, fostering a more holistic approach to healthcare delivery. Through data-driven insights, this tool aids in reducing medical costs by enabling early intervention and proactive management of medical conditions. Overall, this study highlights the potential of model in healthcare for disease pre-diction and underscores its role in advancing personalized medicine and improving clinical decision-making.*

I. INTRODUCTION

In today's digital era, data is incredibly valuable, especially in fields like healthcare where it includes patient information. Traditionally, most models focused on predicting one disease at a time, like diabetes, heart disease, or Parkinson's. There hasn't been a unique system capable of predicting multiple diseases simultaneously. The paper is proposing a new system using Streamlit, a model for building interactive web apps with Python. Our system will predict multiple diseases such as diabetes, heart disease, and Parkinson's. This model will also helpful in predicting various disease later on using machine learning algorithms to make these predictions based on user's input. A method for serializing and deserializing Python object performs as a middle step between getting test results and consulting with a doctor. Plus, not everyone will need to consult a doctor right away because user can get valuable insights from the system. Compared to existing systems, our model is more flexible and efficient. It uses just one set of inputs to predict various diseases, giving users the option to choose which features they want to focus on or rely on the system's recommendations. By predicting diseases in advance, our model can potentially lower mortality rates. Patient data is collected and utilized to predict diseases from the input given by the patients. This enables predictive models to offer valuable insights into healthcare decision-making and facilitates early disease prevention. Furthermore, the system's approach to collecting and utilizing patient data for predictive modeling aligns with the contemporary emphasis on data-driven healthcare decision-making. The insights derived from this model contribute to early disease prevention strategies, fostering a proactive approach to healthcare that goes beyond reactive responses to symptoms. Overall, our system addresses the limitations of existing models by offering flexibility, efficiency, and is able to detect various disease with just one set of input data. This simplifies the process for organizations analyzing patient health reports.

II. LITERATURE SURVEY

[1]. Prof. Shital Kolte "The literature survey underscores the significance of machine learning in software engineering, particularly its role in enabling computers to continuously improve performance on specific tasks through the utilization of statistical procedures and data-driven learning. Machine learning is closely associated with computational statistics and has ties to mathematical optimization. While sometimes confused with data mining, which primarily focuses on exploratory data, machine learning encompasses both supervised and unsupervised learning, with applications in the establishment of benchmark behavior profiles and the detection of anomalies."

[2].Dr.Anu Rathee "The literature survey indicates a significant stride in the field of machine learning, particularly in its applications within healthcare. The ability to predict multiple diseases simultaneously using machine learning models, such as Support Vector Machines (SVM), is highlighted as game-changer in medical diagnostics with the potential to enhance patient outcomes. The research specifically focuses on the application of SVMs for predicting three prevalent diseases: heart disease, diabetes, and Parkinson's disease."

[3]. Mr. Analp Pathak "The primary objective of the paper is to develop an efficient and accurate Machine Learning algorithm for

disease prediction.

Supervised Machine Learning, specifically, is employed for predicting diseases, and various algorithms such as Decision Tree, Random Forest, Naïve Bayes, and KNN are proposed. The intention is to achieve early and accurate disease prediction, contributing to better patient care. The cited reference underscores the significance of this approach in the context of the current technological landscape and required to proactive healthcare solutions. Overall, the literature survey sets the stage to address the issue of early disease prediction and patient care."

[4]. Dr. K Kishore Raju "The literature survey indicates a growing recognition of the challenges faced by medical professionals in diagnosing symptoms and identifying diseases early due to the vast volume of data. Current solutions involve the development of applications. However, existing ML models are often disease-specific, with separate applications for heart disease, diabetes, and Parkinson's disease, etc."

[5]**.** Divya Mandem "The literature survey highlights the current challenges in healthcare, where individuals experiencing illness face time-consuming and expensive processes to consult with doctors, especially when out of reach of medical facilities. The emergence of automated software for disease prediction is recognized as a potential solution to save time and money for patients, particularly in situations where immediate medical attention may not be readily available.

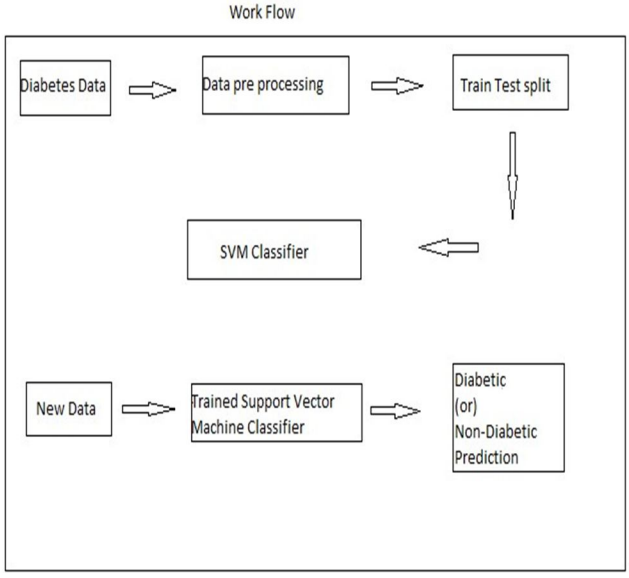
[6]. Dnyaneshwari Mahajan "A Survey on Prediction diseases using Machine Learning. It elaborated various machine learning algorithms and worked towards finding the best algorithm by analyzing their features. Every algorithm has given different result in different situations. Increase the accuracy of predicting the early heart disease with high accuracy and minimum cost and complexity."

[7] P. Shinde "The literature survey includes the challenges in healthcare data collection and processing, highlighting the vast and complex multidimensional data generated in the digital era.

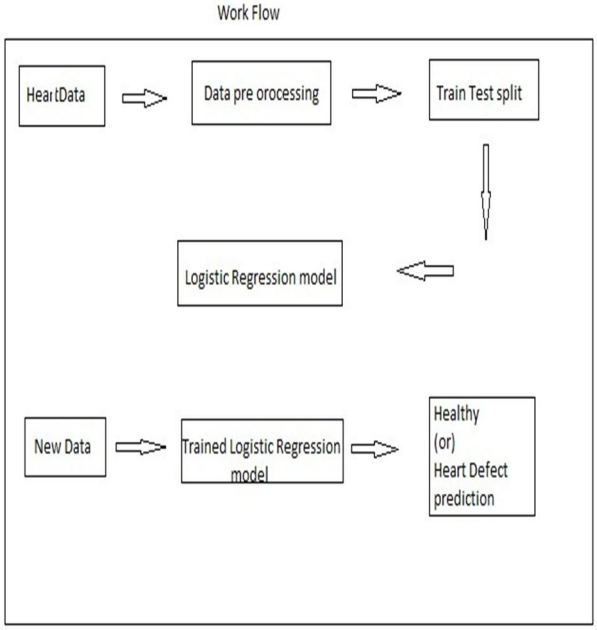
III. PROPOSED METHODOLOGY

The proposed methodology involves using Logistic Regression and SVM algorithms to provide user-friendly web-based system for predicting multiple diseases accurately. The process includes collecting datasets, preprocessing the input data, selecting similar features, training models, and developing a web interface with distinct modules for diabetes, diabetes, and heart diseases. Users input parameters through the interface, and the trained models process the information to provide accurate predictions. The system's performance is determined by using parameters like precision, accuracy, influx of information encompasses clinical factors, hospital resources, diagnostic information, patient records, and medical equipment. The requirement for effective decision-making in healthcare necessitates the processing this immense and intricate dataset. Medical data mining emerges as a crucial tool in uncovering hidden patterns within medical datasets, offering substantial potential for the healthcare industry."

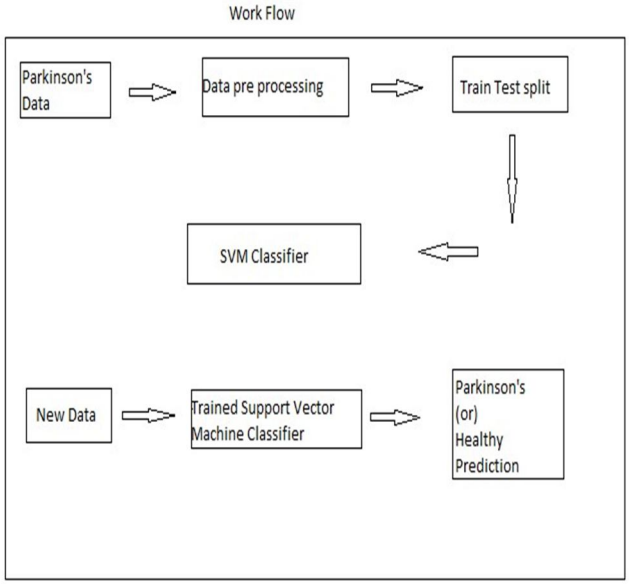
IV. SYSTEM ARCHITECTURE



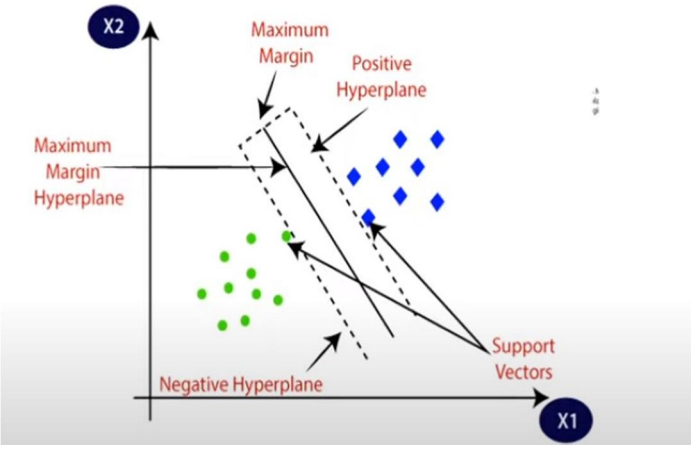
This system is designed to assist doctors in predicting diabetes. It starts by gathering patient information like blood sugar levels and weight. The data is then cleaned and organized so that the system understands it properly. Next, the data is divided into two parts: one for training and one for testing. Training data helps the system to learn patterns that differentiate diabetic and non-diabetic patients, using Support Vector Machine. Like a student studying for a test, the system is then evaluated on the unseen testing data to assess its learning. Finally, the trained system is used to analyze new patient data, predict diabetes and acting as a useful tool to detect early.



User required to enter patient heart data, this dataset include several health parameters which corresponds to a person's healthiness of the heart. First, make the dataset ready and once the dataset is ready then need to process this dataset, but cannot feed this raw data into machine learning algorithm. So, it is required to process this dataset to make fit and compatible for algorithm to learn. Once data is processed, it is splitted into training data and testing data because to train the tool using training data and evaluate performance of tool using testing data so this process is called strain test split. In this case by using logistics regression model to classify whether a person has heart disease or not.



When a user opts for Parkinson's disease prediction, the Parkinson Prediction module focuses on gathering specific parameters crucial for the assessment. Several parameters are used for disease prediction. These vocal characteristics plays role in analyzing and predicting the likelihood of Parkinson's disease. By assessing vocal patterns such as average, maximum, and minimum fundamental frequencies, the module aims to identify potential indicators associated with Parkinson's disease. This focused approach allows the system to offer predictions based on relevant vocal features, contributing to the early detection and monitoring of Parkinson's disease through user-provided data.



A Support Vector Machine (SVM) is like a smart boundary setter in machine learning. Imagine you have data points on a graph, and SVM helps draw a line (or plane in higher dimensions) that separates different groups. It aims to maximize the gap between these groups, making the decision boundary as clear as possible. The data points closest to this boundary are called support vectors. SVM is handy for tasks like classification, where to predict which group a new data point belongs on its features.

V. EXISTING SYSTEM

Machines are good at predicting diseases, but they struggle in predicting the specific types of diseases. They can't cover all the potential conditions people might have. The current systems only work with structured data and not accurate or specific. They're also expensive, making them inaccessible to users. Additionally, they often fail to predict the full range of conditions a person might experience. For example, if patient is having diabetes, then that patient might also have high risk of heart disease. Overall, existing prediction systems have limitations in accuracy, affordability, and comprehensiveness.

VI. PROPOSED SYSTEM

The Proposed system of analysis and prediction of multiple disease using machine learning is that algorithm like SVM and all other various tools to create a system which predicts the disease of the patient using the symptoms and by taking those symptoms comparison is done with the system's dataset that is previously available. By taking those datasets and comparing with the patient's disease predicting the accurate percentage disease of the patient. The dataset and symptoms given to the system that gives prediction where the data is preprocessed for the future. Then the classification of those data is done by utilizing the algorithm such as SVM. It predicts probable diseases by using datasets such as Diabetes, Heart Disease and Parkinson's Disease. Overall, current prediction systems excel in conditions of accuracy, affordability, and comprehensiveness.

VII. RESULTS AND DISCUSSIONS

System using the SVM algorithm to detect whether a patient has a disease based on input values. The system prompts users to input certain values within specified ranges. If the values are outside the ranges or invalid, a warning sign is displayed. SVM is good for linear data and original data is split into train and test datasets for accuracy assessment.

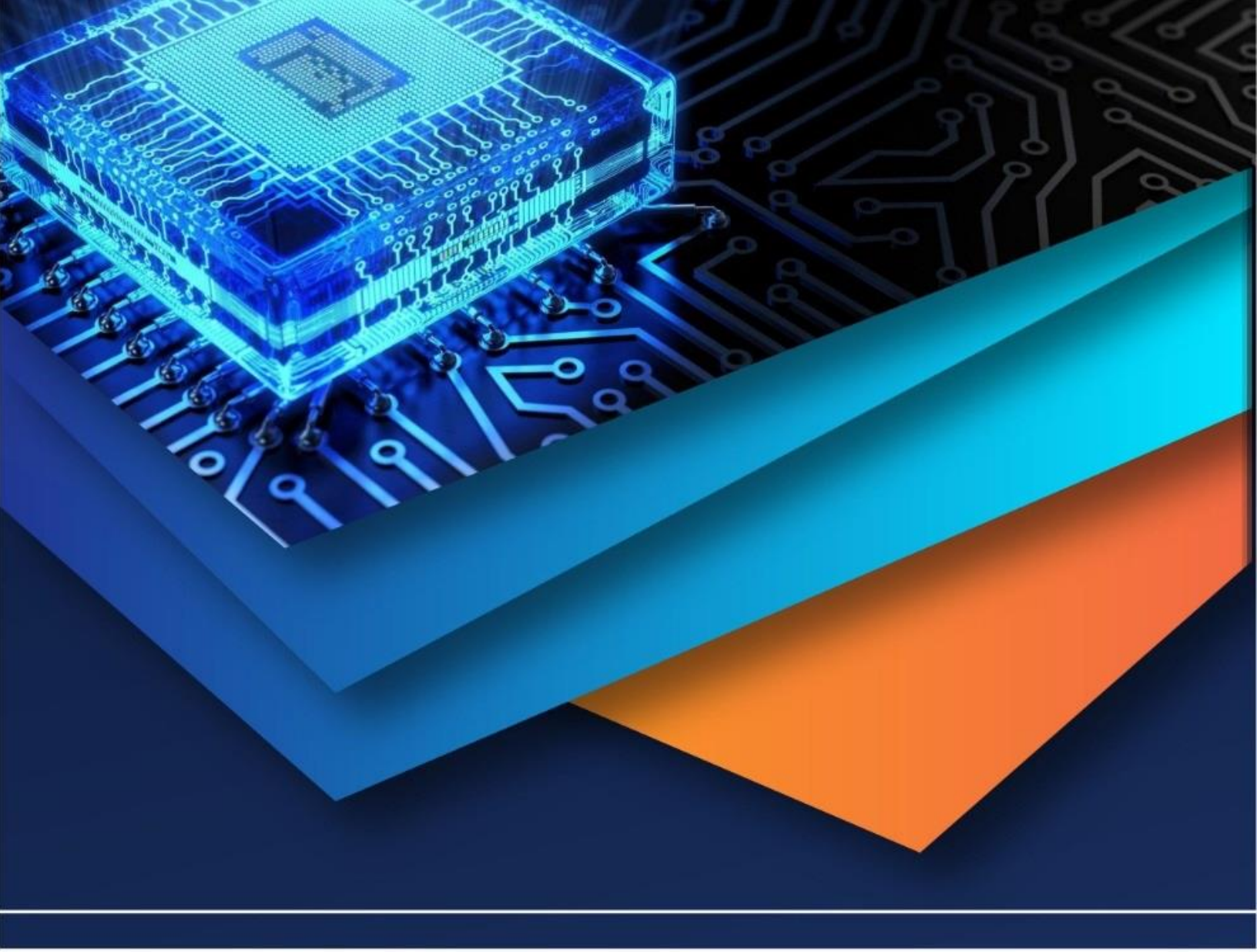
VIII. CONCLUSION

This project utilizes machine learning model to predict various medical conditions based on input data provided by the user. Gather comprehensive data on diabetes, heart disease, and Parkinson's disease, including patient details and relevant health metrics.

The model uses the collection of data to recognize some symptoms which will be helpful to detect diseases. Develop an easy-to-use interface that healthcare professionals can navigate effortlessly, allowing them to input patient data and to receive a clear outcome in accessible manner. Creating a tool that helpful for earlier detection of various diseases.

REFERENCES

[1] P. S. Kolte, "Disease Prediction by Machine Learning from Healthcare Communities", JSRST, vol. 6, no. 5, pp. 29-35, 2021.
[2] D. RATHEE, "Multiple Disease Prediction Using Machine Learning", IRE Journals, vol. 6, no. 12, pp. 411-415, 2023.
[3] M. Pathak, "a machine learning model for early prediction of multiple diseases", Turkish journal of computer and mathematics education, vol. 12, no. 6, pp. 4013-4023, 2021.
[4] D. K. K. Raju, "MULTIPLE DISEASE PREDICTION", JETIR, vol. 10, no. 4, pp. 246-253, 2023.
[5] D. Mandem, "Multi Disease Prediction System", IJIRT, vol. 8, no. 6, pp. 503-509, 2021.
[6] Dnyaneshwari Mahajan, "A Survey on prediction techniques of heart disease using Machine Learning", vol. 12, pp.15-18, 2023.
[7] Priyanka P. Shinde, "Multiple disease prediction using Machine learning algorithms", ELSEVIER, pp. 2214-7853,2021.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY