MULTIPLE DISEASE PREDICTION AND DRUG RECOMMENDATION SYSTEM USING MACHINE LEARNING

Dr.J.Amutharaj, Likhith Gowda L, Mohammed Taufique, Nanjunda K M

[1]Dr.J.Amutharaj, Department of ISE, Institute Rajarajeswari College of Engineering, Karnataka, India
[2] Likhith Gowda L, Department of ISE, Institute Rajarajeswari College of Engineering, Karnataka, India
[3] Mohammed Taufique, Department of ISE, Institute Rajarajeswari College of Engineering, Karnataka, India
[4] Nanjunda K M, Department of ISE, Institute Rajarajeswari College of Engineering, Karnataka, India

ABSTRACT

The increasing prevalence of chronic diseases and the exponential growth of healthcare data have created a strong need for intelligent, data-driven diagnostic solutions. Traditional medical diagnosis often requires expert interpretation, repeated clinical examinations, and significant time. These limitations encourage the use of machine learning-based decision-support systems that can efficiently analyze patient attributes and predict potential diseases in real-time. This research work proposes a unified ***Multiple Disease Prediction and Drug Recommendation System*** *capable of predicting three major diseases—****Diabetes****,* ***Heart Disease****, and* ***Parkinson's Disease****—using supervised machine learning algorithms. The system integrates essential stages such as data preprocessing, feature standardization, model training, evaluation, and deployment through an interactive web interface developed using Streamlit.*

The machine learning models used include Logistic Regression for diabetes prediction, Random Forest for heart disease prediction, and Support Vector Machine (SVM) for Parkinson's disease detection. A rule-based drug recommendation module is incorporated to suggest suitable medications based on commonly prescribed drugs for each predicted condition. Publicly available UCI datasets were used for training and evaluation, and results show high accuracy for all models, with SVM achieving 94% for Parkinson's, Random Forest achieving 88% for heart disease, and Logistic Regression yielding 82% accuracy for diabetes. The proposed system is designed to assist healthcare providers and individuals in early detection, reducing human error, and offering timely treatment guidance. Overall, this work demonstrates the effectiveness of machine learning in healthcare analytics and highlights its potential to enhance preventive medicine and personalized treatment.

Keyword : - *Machine Learning, Disease Prediction, Drug Recommendation, Healthcare Analytics, Classification Models.*

1. Introduction

Machine learning has emerged as a transformative technology in modern healthcare. With the availability of large medical datasets, it is now possible to apply supervised learning algorithms to predict diseases accurately. Chronic diseases such as diabetes, heart disease, and Parkinson's disease require early detection to prevent severe health complications. Traditional diagnosis heavily depends on the expertise of medical professionals and manual interpretation of laboratory results, which may sometimes lead to delays or inaccurate predictions. The effective use of machine learning reduces manual intervention and enhances predictive performance by recognizing complex patterns in patient data.

Predictive modeling also supports preventive healthcare by identifying risk factors before the onset of severe symptoms. The proposed Multiple Disease Prediction System aims to integrate three disease models into a single unified platform. This improves accessibility, reduces redundancy, and provides multi-disease screening in one interface. As per Kong, early-stage disease classification with ML significantly improves patient outcomes [3]. It is reported that automated prediction systems minimize diagnostic variability and provide consistent results across datasets [2]. This study addresses these gaps by combining disease prediction with drug recommendation to provide a comprehensive healthcare-assistive application.

1.1 Background Study

Disease prediction systems analyze clinical parameters such as glucose levels, blood pressure, vocal frequency, cholesterol levels, and other health indicators. These values help identify hidden correlations between symptoms and diseases. Early machine learning models focused on single diseases, but recent advancements have shown that multi-disease frameworks are more beneficial. The integration of multiple algorithms and datasets in one system improves prediction accuracy and reduces the dependence on individual diagnostic models.

.

1.2 Need for an Integrated System

Most existing diagnostic systems provide predictions for only one disease at a time, requiring patients to rely on separate applications or medical consultations. This creates inconvenience and increases the likelihood of delayed diagnosis. A multi-disease prediction system eliminates such limitations by offering consolidated insights. It also supports telemedicine, remote health consultation, and early risk assessment for individuals without direct access to healthcare professionals.

2. Methodology

The methodology includes data collection, preprocessing, model training, evaluation, and deployment. Each dataset undergoes outlier removal, handling of missing values, normalization, and label encoding to ensure model consistency.

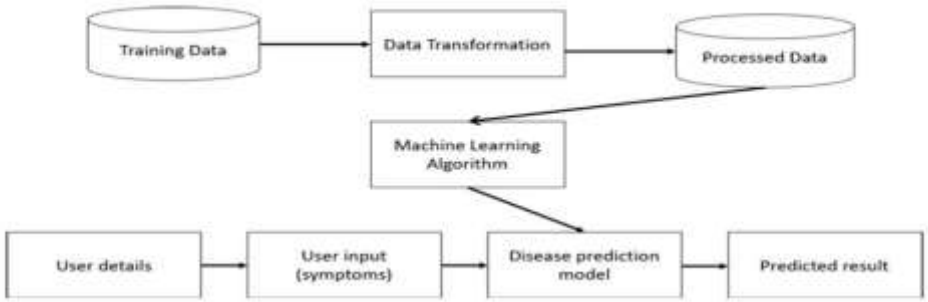

Chart -1 Workflow of the sysytem

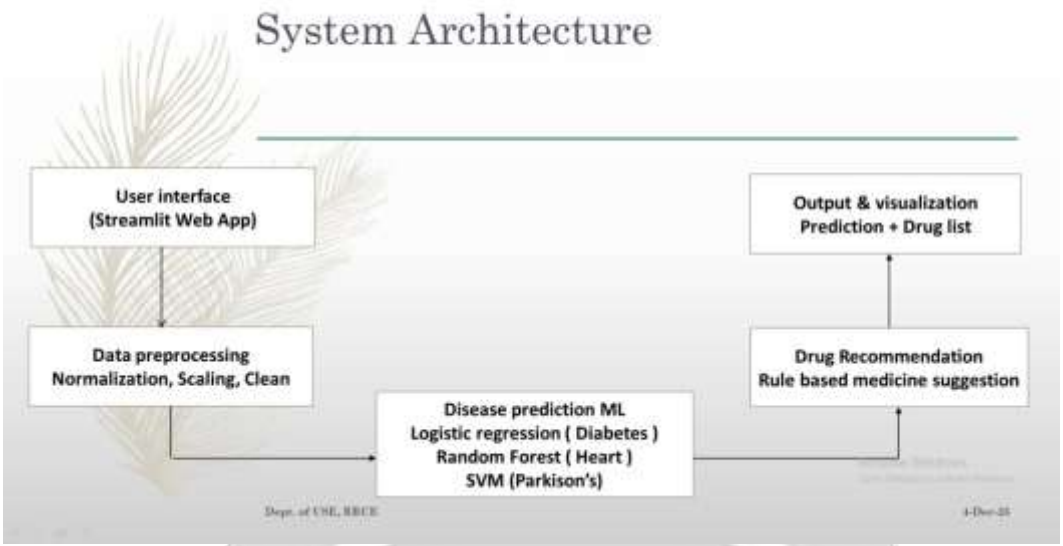

Fig -1 Block Diagram of the System

2.1 Machine Learning Algorithms Used:

Logistic Regression is used for diabetes prediction due to its simplicity and performance on linearly separable data. Random Forest, an ensemble model, is used for heart disease prediction due to its robustness and ability to handle multiple features. SVM is used for Parkinson's disease detection because it effectively manages high-dimensional biomedical attributes. Together, these algorithms form a hybrid system that ensures reliable predictions.

Table -1

Data Sets	Records	Features	Disease
PIMA	768	8	Diabetes
Cleveland	303	13	Heart Disease
Parkison's	195	24	Parkison's

2.2 Drug Recommendation Module

The drug recommendation module is rule-based, mapping each predicted disease to standard medications commonly prescribed in clinical practice.

- List Item – 1 Metformin for Diabetes
- List Item - 2 Aspirin/Statins for Heart Disease
- List Item – 3 Levodopa for Parkinson's
- List Item - 4 Drug suggestions displayed in UI

3. RESULTS AND DISCUSSION

The system was evaluated using accuracy, precision, recall, and error rate. Experimental results indicate that SVM performs best due to its ability to separate nonlinear and high-dimensional data. Graphs demonstrate the comparative performance of models across diseases.

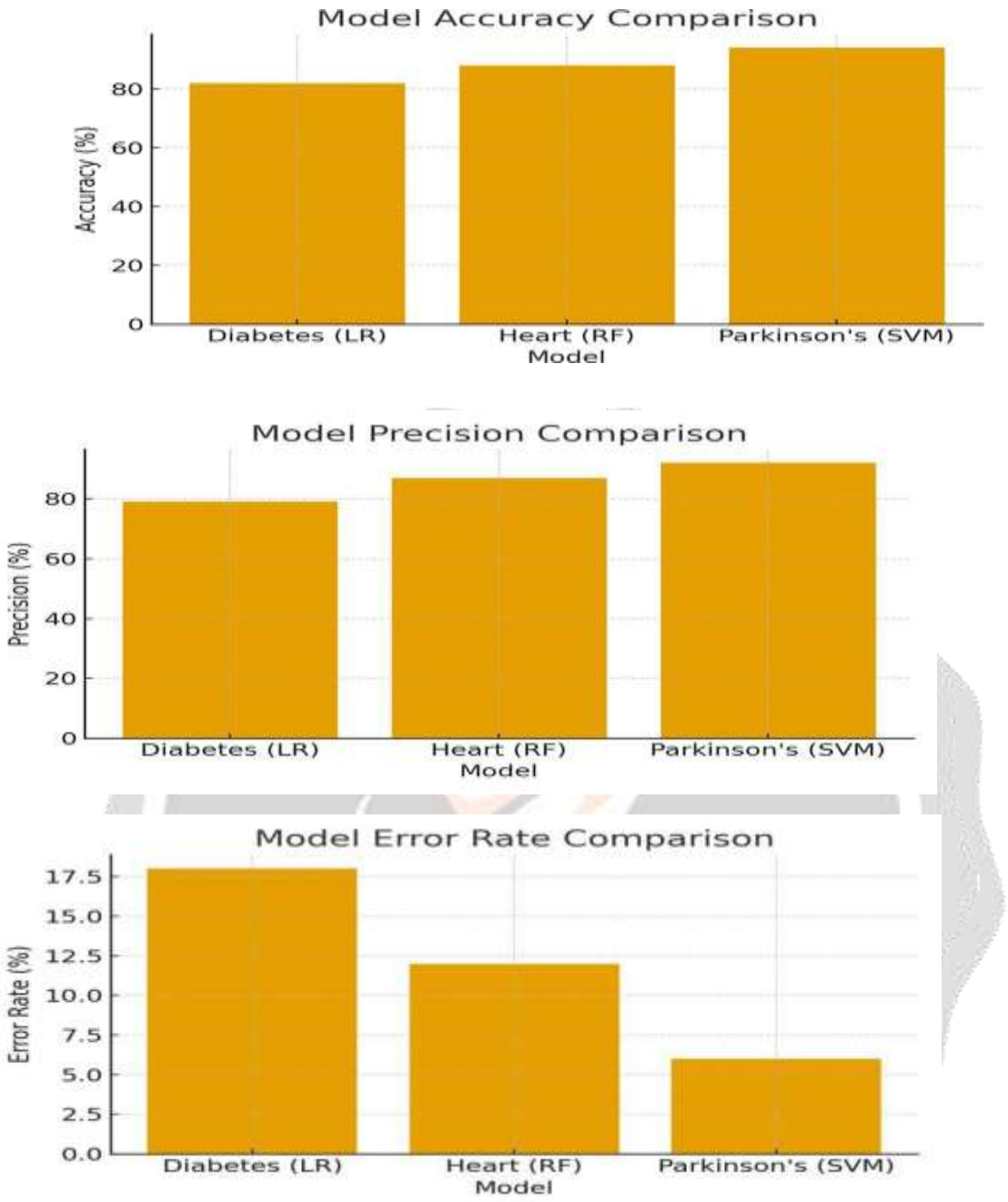

Chart -2 Model Accuracy, Precision and error rate Chart

3.1 User Interface Results

The Streamlit interface allows users to enter values such as glucose level, cholesterol, blood pressure, voice frequency metrics, and instantly view predicted results. The UI is simple, intuitive, and responsive.

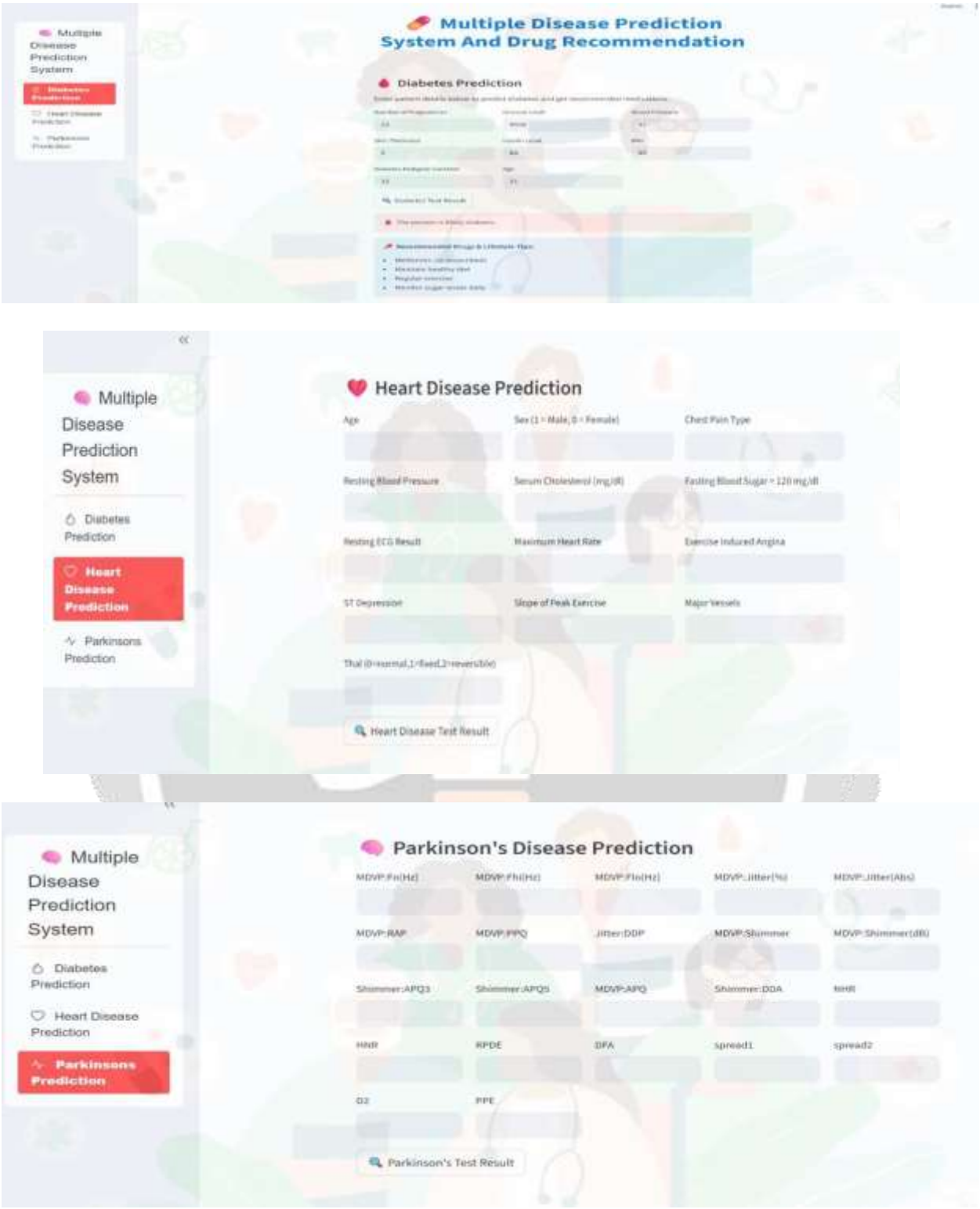

Fig -2 User Interface Output

4. CONCLUSIONS

The proposed system effectively predicts multiple diseases using machine learning and provides drug recommendations based on predicted outcomes. The integration of three disease models into one platform enhances usability and supports preventive healthcare. Future improvements include adding more diseases, increasing dataset size, integrating deep learning models, and supporting real-time patient data.

5. ACKNOWLEDGEMENT

The authors express gratitude to the Department of Information Science and Engineering at RRCE for providing resources and guidance throughout the project.

6. REFERENCES

[1].Reference on machine learning for healthcare applications. .
[2]. Study on multi-disease prediction frameworks. .
[3]. Kong S., Predictive Healthcare Analytics, 2024. .
[4]. Padhy et al., AI-Based Disease Prediction, 2024.