WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

HEART DISEASES PREDICTION SYSTEM

Chinmai Chouhan[1], Jharna Chopra[2]

[2] Guide Name:
Shri Shankaracharya Technical Campus

ABSTRACT :

Cardiovascular diseases (CVDs) are a major global health concern and represent the leading cause of death worldwide. Early prediction and intervention can significantly reduce morbidity and mortality rates. This paper presents the design, implementation, and evaluation of a Heart Diseases Prediction System using logistic regression, coupled with a graphical user interface (GUI) and an AI-driven chatbot. The system leverages patient health metrics such as age, cholesterol level, resting blood pressure, and other critical indicators to predict the likelihood of heart disease. The chatbot supports users by providing health advice and explaining medical terms, thereby enhancing the system's interactivity. Results indicate that the model achieves commendable accuracy on training and test data, and the interface offers ease of use for healthcare professionals and patients alike. This work demonstrates the feasibility and effectiveness of machine learning in medical diagnostics.

Keywords: Heart Disease, Machine Learning, Logistic Regression, Chatbot, GUI, Medical Diagnosis

1. Introduction

Heart disease remains the most prevalent cause of death globally. The ability to detect cardiovascular risks early has profound implications on public health. Traditional diagnostic procedures often require extensive tests and are time-consuming. With the rise of machine learning and big data analytics, predictive modeling in medicine has become a powerful tool to assist healthcare professionals.

This project focuses on creating a Heart Diseases Prediction System using logistic regression. The system allows the user to input medical parameters via a GUI, predicts the likelihood of heart disease, and provides visual insights. A supporting chatbot acts as a virtual health assistant, answering queries about the parameters and offering lifestyle advice. This paper outlines the design, development, and evaluation of the system.

2. Literature Review

Over the past decade, researchers have explored various machine learning techniques for medical diagnosis. Logistic regression, decision trees, support vector machines (SVM), and neural networks have been used for predicting diseases such as diabetes and heart conditions. The UCI Heart Disease dataset has frequently been used as a benchmark for such applications.

Previous works demonstrated accuracies ranging from 70% to 90%, depending on the features used and preprocessing techniques applied. However, the real-world applicability of these models is often limited due to the lack of user interfaces or decision support components like chatbots. Our project addresses this by offering both backend intelligence and frontend accessibility through integrated modules.

3. Methodology

This system is designed using the following major components: dataset preprocessing, model training, user interface (GUI), and chatbot integration.

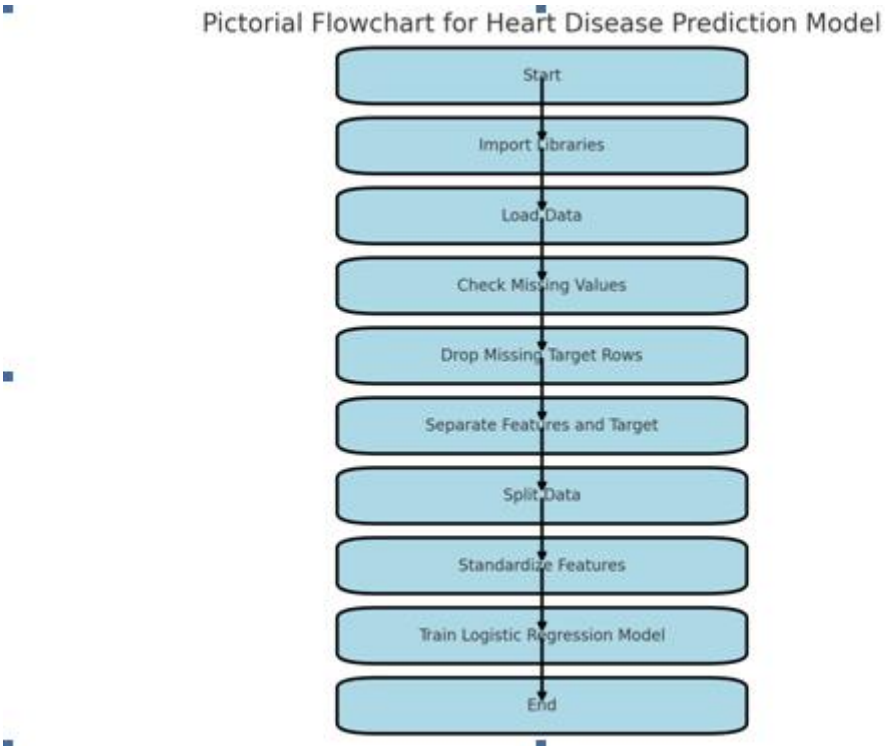
a) Dataset: We used the UCI Heart Disease dataset consisting of 303 entries and 14 features including age, sex, chest pain type, serum cholesterol, fasting blood sugar, maximum heart rate, and more. The 'target' variable indicates the presence of heart disease.

b) Preprocessing: Missing values were handled using SimpleImputer with a mean strategy. Features were standardized using StandardScaler to normalize input ranges.

c) Model: Logistic Regression was chosen for its interpretability and effectiveness in binary classification. The model was trained using an 80-20 train-test split. Evaluation metrics included training accuracy, test accuracy, and confusion matrices.

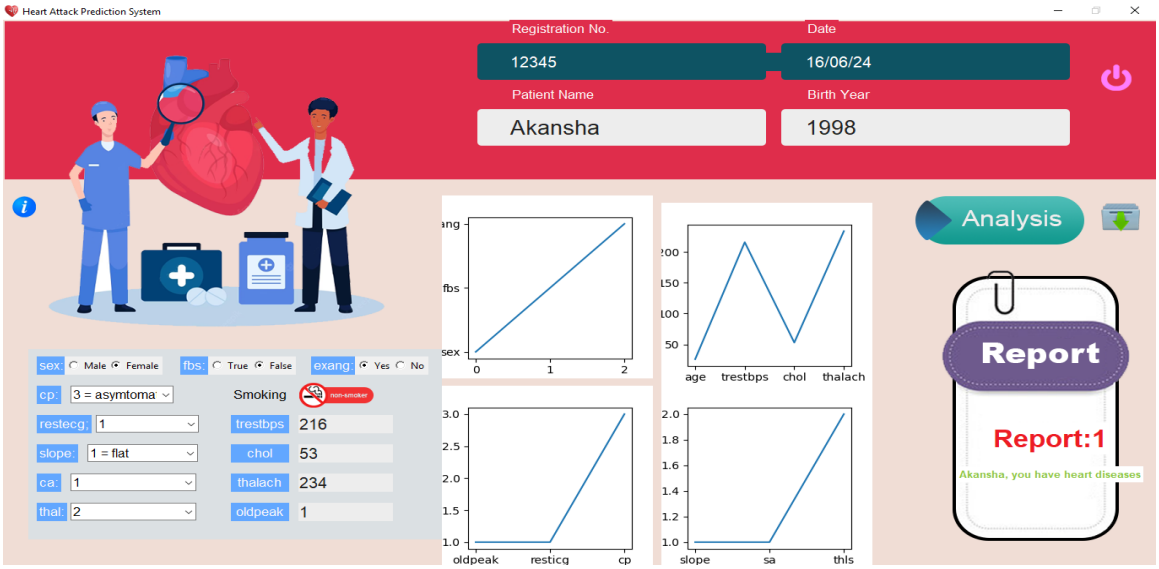
d) GUI: Developed using Python's Tkinter module, the GUI allows users to enter patient details, view visual graphs, and get instant prediction results.

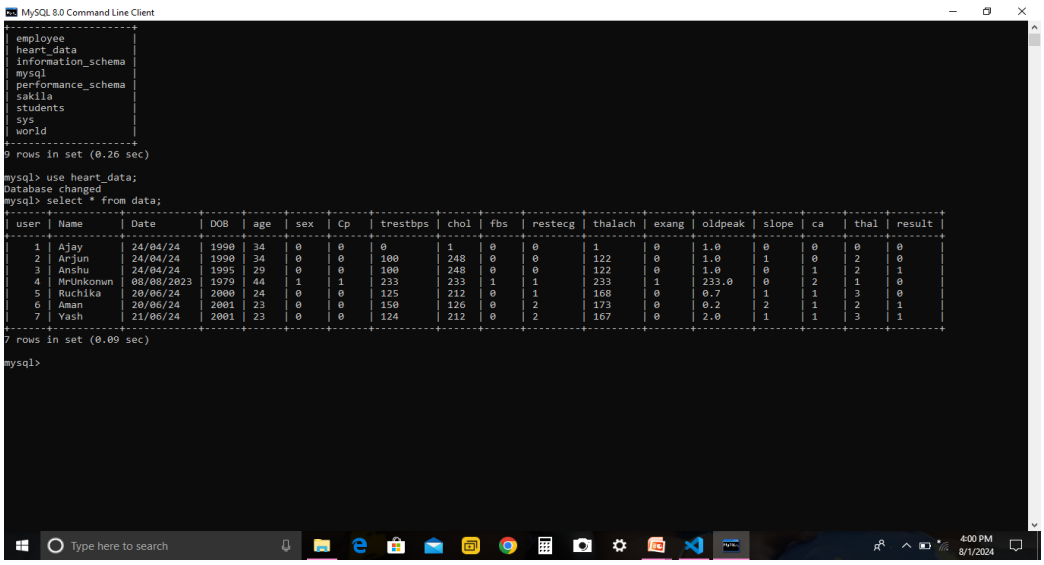
e) Chatbot: Implemented using Python, the rule-based chatbot responds to medical queries, explaining health indicators and offering general wellness tips.

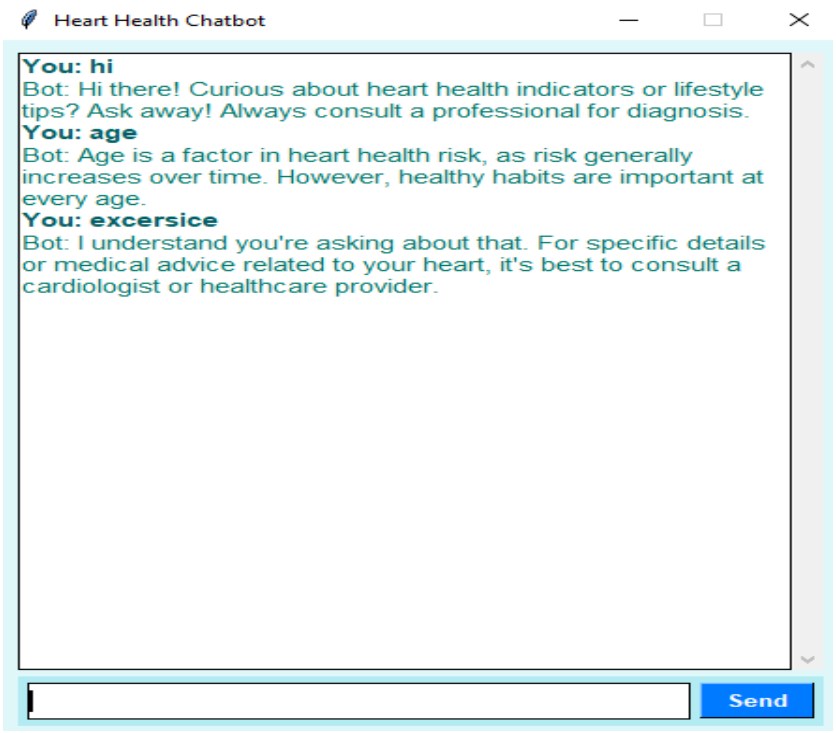


4. Results

The logistic regression model demonstrated reliable prediction capabilities. Accuracy on the training data was above 85%, and the test accuracy remained consistent around 83%, indicating minimal overfitting. Below are visual outputs used in the application to enhance user understanding:







5. Discussion

The predictive model achieves satisfactory performance, suggesting that machine learning is a viable tool for medical decision support. The graphical interface ensures that non-technical users can interact with the system effectively, while the chatbot adds value by assisting with terminology and wellness advice.

Limitations include reliance on a relatively small dataset and a rule-based chatbot rather than an NLP-based one. Future improvements could incorporate deep learning for improved accuracy and expand the chatbot with AI language models for more natural interaction.

6. Conclusion

The Heart Diseases Prediction System effectively combines machine learning, visual analytics, and human-computer interaction to offer an accessible, accurate, and educational tool for preliminary heart health assessments. This solution can assist clinics, healthcare professionals, and even individuals seeking health insights. The modular design ensures it can be further enhanced with more data and AI capabilities.

7. REFERENCES

[1] UCI Machine Learning Repository – Heart Disease Dataset
[2] Scikit-learn: Machine Learning in Python
[3] Tkinter GUI Library Documentation
[4] Logistic Regression: Theory and Applications in Healthcare
[5] Chatbot Design for Medical Use: A Review