International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Heart Disease Prediction System: A Definitive Approach

Shailendra W. Shende[1], Abhay Kumar Kashyap[2]

[1]*Department of Computer Science and Engineering*
Government College of Engineering, Chandrapur

[2]*Department of Computer Science and Engineering*
Government College of Engineering, Chandrapur

Abstract – Heart Disease Prediction System is a technological solution designed to predict the likelihood of a person developing heart disease based on various factors and health data. This system typically uses machine learning algorithm named Logistic Regression to analyse data such as age, sex, chest pain, resting blood pressure, cholesterol, fasting blood sugar, maximum heart rate achieved, exercise-induced angina, and more. By processing this data, the system can generate a prediction or recommendation about an individual's risk of heart disease, helping to identify at-risk individuals early and enable preventive actions.

Key Words: Heart Disease Prediction, Logistic Regression, Machine Learning, Graphical User Interface, Artificial Intelligence, Heart Disease Prediction System

1.INTRODUCTION

Heart disease prediction is a crucial step in early diagnosis proactive healthcare. With the increasing prevalence of heart related diseases worldwide, having an effective, data-driven approach to predict the likelihood of heart disease can significantly improve patient outcomes. In India around 27% of total heart related deaths are taking place. To counter the situation the heart disease prediction system with its user-friendly interface and data-driven approach plays a vital role in improving overall heart health outcomes and reducing the impact of heart related diseases on individuals and society.

2.1 LITERATURE SURVEY

Previous research has shown the potential of machine learning in diagnosing heart disease.

We referred D. Hemalatha et al. [1] which proposed that the models were trained and tested on standard cardiovascular datasets like the Cleveland Heart Disease dataset. The dataset constituted of 13 attributes and 1 target variable which has data 303 patients.

Rohit Bharti et al. [2] studied the effectiveness of combining traditional machine learning algorithms with deep learning techniques for heart disease prediction. The authors developed a hybrid approach that leverages the strengths of both methodologies to improve diagnostic accuracy. They concluded that machine learning algorithms performed better in this analysis.

Dr. T. Venu Gopal et al. [3] explored the application of various machine learning algorithms to predict heart disease based on patient clinical data. Heart disease remains one of the leading causes of mortality worldwide, and early prediction can significantly improve patient outcomes.

Nidhi Bhatia et al. [4] explored various data mining techniques for predicting heart disease, comparing their effectiveness and accuracy. The study evaluates multiple classification algorithms on heart disease datasets to determine which approaches yield the most reliable predictions.

V. Krishnaiah et al. [5] examined various data mining techniques and fuzzy approaches for heart disease prediction systems. The authors analyze multiple methodologies that combine machine learning algorithms with fuzzy logic to improve the accuracy of cardiovascular disease prediction.

Sellappan Palaniappan et al. [6] demonstrated how data mining can be applied to healthcare data to create predictive models that assist in early diagnosis and treatment planning.

Senthilkumar Mohan et al. [7] examined the use of hybrid machine learning approaches for improving heart disease prediction accuracy. The researchers combined multiple algorithms to overcome the limitations of

individual methods when applied to cardiovascular disease detection.

V. V. Ramalingam et al. [8] reviewed the application of various machine learning techniques for heart disease prediction. Heart disease remains one of the leading causes of mortality worldwide, and early detection through predictive models can significantly improve patient outcomes.

Baban U. Rindhe et al. [9] explored the application of machine learning algorithms for predicting heart disease based on various clinical parameters.

Pooja Rani et al. [10] presented a decision support system for predicting heart disease using machine learning techniques. The authors combine Apache Spark's ML library with a real-time processing framework to create an efficient solution for medical practitioners. The system analyzes patient data to predict heart disease risks. The research demonstrates how big data technologies can be integrated with machine learning to improve medical diagnostic processes.

Pothala Ramya et al. [11] stated that heart disease remains a leading cause of mortality worldwide, making early and accurate detection crucial for effective treatment and reduced mortality rates.

2.2 DATASET AND PREPROCESSING

2.2.1 DATASET DESCRIPTION

For this study, we use the Cleveland database of UCI repository of heart disease patients, which is widely used in heart disease prediction research. It contains the following attributes:

- **Age:** Age of the patient.
- **Sex:** Gender of the patient (1 = male, 0 = female).
- **Chest Pain:** Type of chest pain experienced by the patient (4 different categories).
- **Resting Blood Pressure:** Blood pressure measured at rest.
- **Serum Cholesterol:** Serum cholesterol in mg/dl.
- **Fasting Blood Sugar:** Blood sugar level measured after fasting (1 = > 120 mg/dl, 0 = <= 120 mg/dl).
- **Electrocardiographicresults:** Electrocardiogram results at rest.
- **Maximum Heart Rate:** Maximum heart rate achieved.
- **Exercise-Induced Angina:** Exercise-induced angina is developing or not (1 = Yes, 0 = No).
- **Oldpeak:** Depression induced by exercise relative to rest.
- **Slope of the peak exercise ST segment:** Type of slope of the peak exercise ST segment.
- **Number of major vessels colored by fluoroscopy:** Count of vessels colored by fluoroscopy.
- **Thalassemia:** A blood disorder.

The target variable is whether the patient has heart disease or not (1 = Yes, 0 = No).

2.2.2 DATA PREPROCESSING

Data preprocessing steps involved:

- **Handling missing values:** Missing values were imputed using the mean for numerical columns and mode for categorical columns.
- **Feature scaling:** Standardization (zero mean, unit variance) was applied to numerical features.
- **Encoding categorical variables:** One-hot encoding was used for categorical features, such as chest pain type.
- **Splitting the dataset:** The dataset was split into training (80%) and testing (20%) sets.

2.3 METHODOLOGY

2.3.1 MACHINE LEARNING ALGORITHM

Logistic Regression: Logistic regression is a fundamental statistical and machine learning algorithm used for binary classification tasks, where the goal is to predict one of two possible outcomes. It models the probability of a certain class based on one or more independent variables.

We have used logistic regression because it uses statistical approach and works well with already identified independent variables. It gives best result on the numerical dataset as compared to other contemporary and new algorithms.

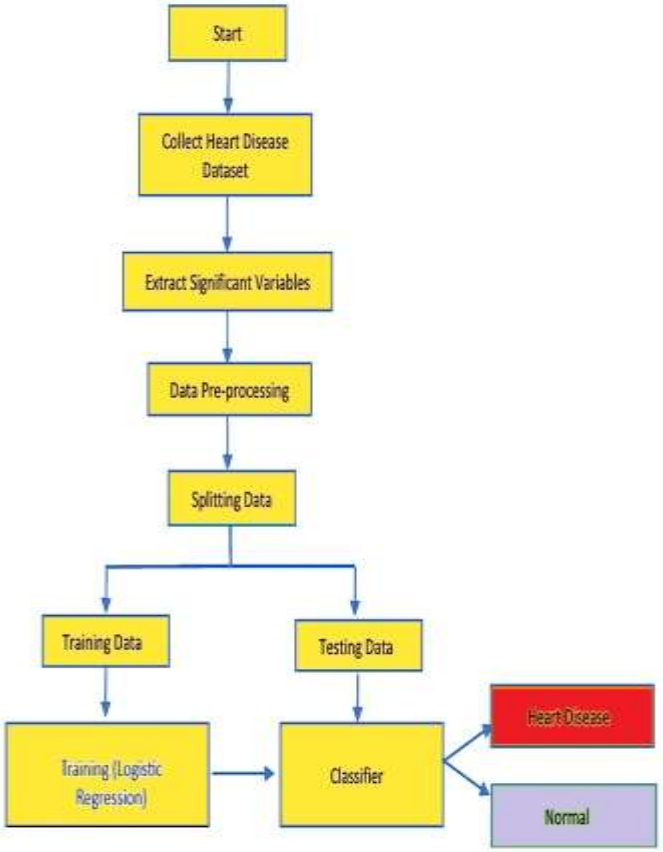


Fig.1Working of Model

2.3.2 FEATURE SELECTION

We utilized feature importance methods to identify the most influential features in predicting heart disease. We generated a heatmap which gave us information about features which were influential for determining whether the patient has heart disease or not.

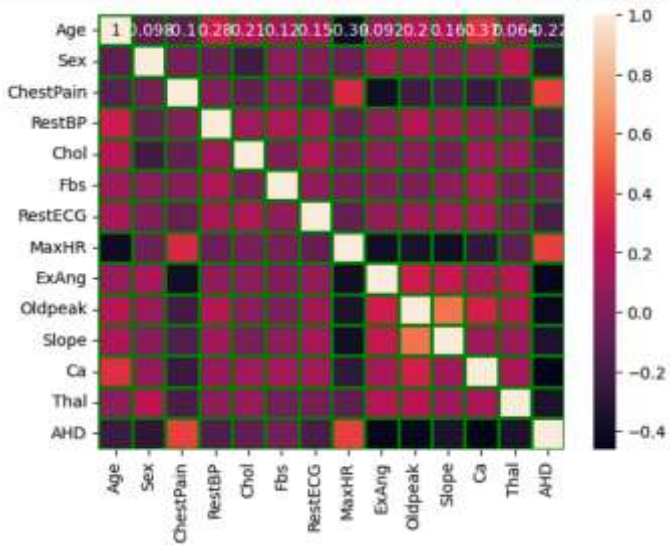


Fig. 2 Heatmap showing relations

It is clear from above heatmap that Age, Chest Pain Type, Resting Blood Pressure, Serum Cholesterol and Maximum Heart Rate are important features for determining whether the patient has heart disease or not.

2.3.3 UI DESIGNING

We developed a graphical user interface using Tkinter module available in Python programming language. The UI (User Interface) comprises of all the 13 features to be inputted by user and a predict button which predicts the target variable, i.e., whether the patient has heart disease or not.

The UI (User Interface) has been integrated with the model of Heart Disease Prediction to generate the prediction whether the patient has heart disease or not. For the integration of Heart Disease Prediction and UI (User Interface) we have used joblib module.

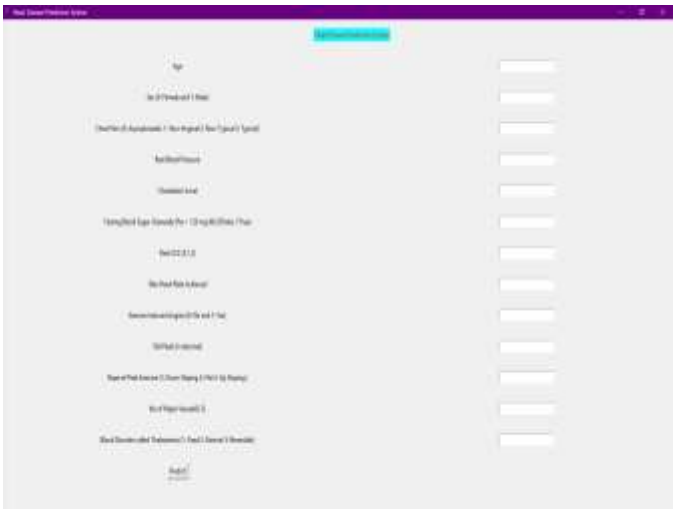
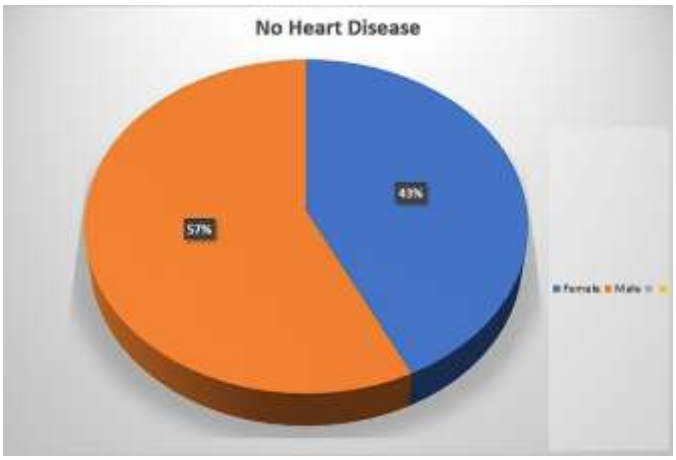


Fig. 3 GUI (Graphical User Interface)

2.4 RESULT AND DISCUSSION

2.4.1 RESULT

We have successfully completed the task of generating a GUI (Graphical User Interface) for predicting whether the patient has heart disease or not using a machine learning model which is trained and tested in the Cleveland dataset, which has 13 attributes and 1 target variable with 303 patients' data. The model uses Logistic Regression algorithm because this algorithm has best application on the data which is numerical in nature.



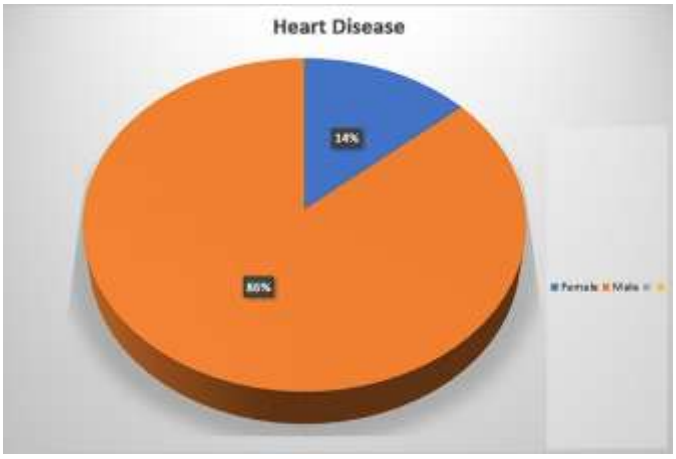
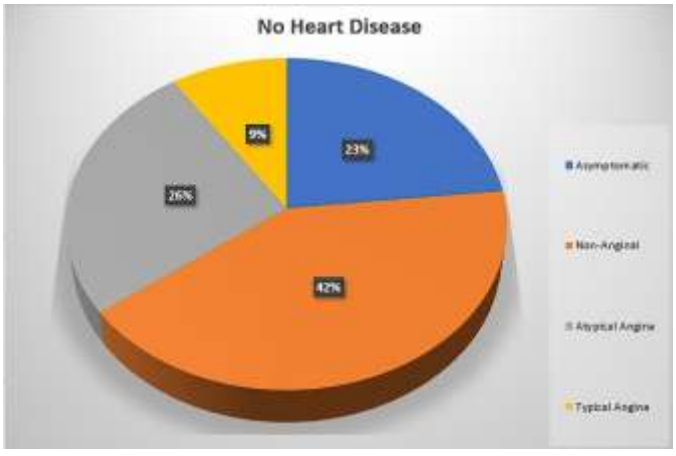


Fig. 4 Shows the patients having or not having heart disease based on sex.



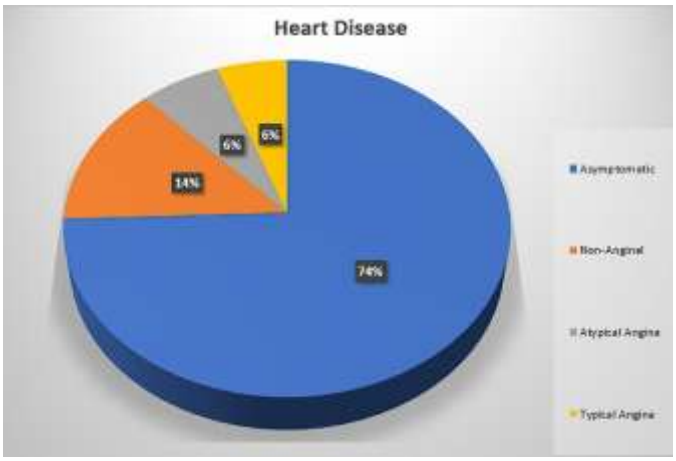


Fig. 5 Shows the patients having or not having heart disease based on type of chest pain.

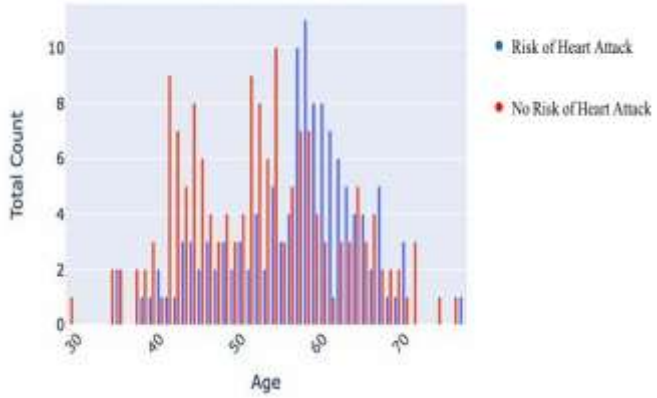


Fig. 6 Shows the risk of heart disease based on their age.

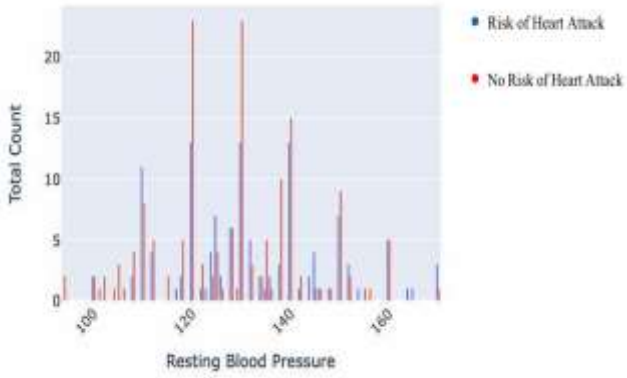


Fig. 7 Shows the risk of heart disease based on their resting blood pressure.

2.4.2 DISCUSSION

We would like to discuss and highlight the utility of the model for predicting whether the patient has heart disease or not. This model saves time and cost for predicting the result whether the patient has heart disease or not. This model paves the way for personalized healthcare where

the person would get the prediction based on the input of the health data needed for the result.

Pothala Ramya et al. [11] discussed limitations of current approaches and suggest future research directions, including integration with other data types and exploration of deep learning techniques.

Our future work could focus on integrating more complex model like Generative AI, as well as incorporating additional datasets for more generalized predictions.

3. CONCLUSIONS

This paper demonstrates that machine learning technique and UI designing modules in Python programming language offer new and robust methods for predicting heart disease. The heart disease prediction system offers a valuable tool for early detection, allowing healthcare professionals to make informed decisions and intervene promptly. The integration of real-time data processing ensures that patients can receive personalized care, ultimately reducing the risk of heart-related complications. As technology continues to evolve, further refinement of predictive models will lead to even more reliable and accessible solutions, contributing to better patient outcome and an overall improvement in cardiovascular health management.

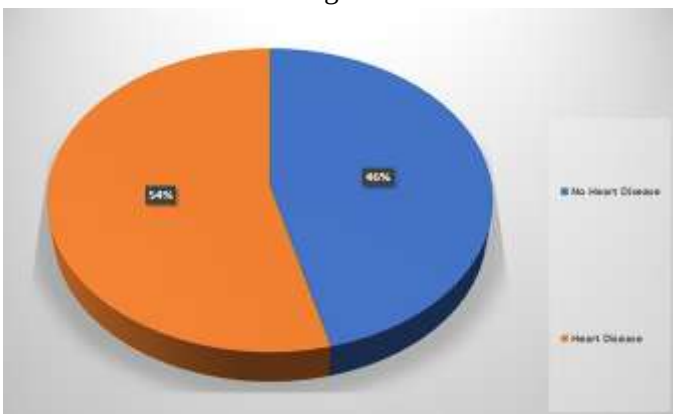


Fig. 8 Shows the total number of patients having or not having heart disease.

REFERENCES

1. D. Hemalatha, S. Poorani, Machine learning techniques for heart disease prediction, Journal of Cardiovascular Disease Research, 12 (1) ISSN: 0975-3583, 0976-2833 10.31838/jcdr.2021.12.01.05, 2021
2. Rohit Bharti, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande, Parneet Singh, Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning, Hindawi Computational Intelligence and Neuroscience, Volume 2021, Article ID 8387680, 2021
3. Dr. T. Venu Gopal, Avudurthi Sneha, HEART DISEASE PREDICTION USING MACHINE LEARNING TECHNIQUES, International Journal for Advanced Research in Science & Technology, Volume 13, Issue 07, Jul 2023, 2023
4. Nidhi Bhatia, Kiran Jyoti, An Analysis of Heart Disease using Different Data Mining Techniques, International Journal of Engineering Research & Technology (IJERT), Volume 1, Issue 8, October 2012, 2012
5. V. Krishnaiah, G. Narsimha, N. Subhash Chandra, Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review, International Journal of Computer Applications, Volume 136 – No.2, February 2016, 2016
6. Sellappan Palaniappan, Rafiah Awang, Intelligent Heart Disease Prediction System Using Data Mining Techniques, International Journal of Computer Science and Network Security (IJCSNS), Volume 8 No. 8, August 2008, 2008
7. Senthilkumar Mohan, Chandrasegar Thirumali, Gautam Srivastava, Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques, Institute of Electrical and Electronics Engineers (IEEE), Volume 7, 2019
8. V. V. Ramalingam, Ayantan Dandapath, M Karthik Raja, Heart disease prediction using machine learning techniques: a survey, International Journal of Engineering & Technology (IJET), 7 (2.8) (2018) pg. 684-687, 2018
9. Baban U. Rindhe, Nikita Ahire, Rupali Patil, Shweta Gagare, Manisha Darade, Heart Disease Prediction Using Machine Learning, International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), Volume 5, Issue 1, May 2021, 2021
10. Pooja Rani, Rajneesh Kumar, Nada M. O. Sid Ahmed, Anurag Jain, A decision support system for heart disease prediction based upon machine learning, Journal of Reliable Intelligent Environments, 25 January 2021, 2021
11. Pothala Ramya, Ashapu Bhavani, Sangeeta Viswanadham, Review: Heart Diseases Detection by Machine Learning Classification Algorithms, Journal of Harbin Institute of Technology (New Series), May 2022, 2022