



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 15, Issue 2, February 2026**

# Heart Disease Prediction using Logistic Regression

**Dr. A. N. Gedam<sup>1</sup>, Pratik Bhandari<sup>2</sup>, Srushti Chavre<sup>3</sup>, Revati Chougule<sup>4</sup>, Avanti Kadam<sup>5</sup>**

Professor, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>1</sup>

Student, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>2</sup>

Student, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>3</sup>

Student, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>4</sup>

Student, Department of Computer Engineering, AISSMS Polytechnic, Pune, India<sup>5</sup>

**ABSTRACT:** Cardiovascular diseases remain a leading global health concern, emphasizing the importance of early detection and continuous monitoring. This project aims to develop an integrated heart- health prediction system that combines machine learning with real-time health data collected through a mobile application. The backend utilizes a logistic regression model to analyze clinical parameters such as blood pressure, heart rate, RR interval, and user-recorded demographic information to predict the likelihood of heart-related abnormalities. The mobile application enables users to enter and track their vital signs, while the backend processes the data and returns a risk prediction. By providing an accessible, data-driven assessment tool, this system supports early identification of potential heart disease and enables timely medical intervention. This proposal outlines the methodology, system architecture, tools (React.js, CSS, React Router DOM, Axios, Fetch API, LocalStorage, Node.js, Express.js, Database – MongoDB) and machine-learning approach – Logistic Regression required to build the complete predictive platform.

**KEYWORDS:** cardiovascular diseases; early detection; machine learning; browser-based system; logistic regression; Heart Disease Risk Assessment ; heart monitoring.

## I. INTRODUCTION

Heart disease continues to be one of the leading causes of mortality worldwide, affecting millions of people each year and placing a significant strain on healthcare systems. Cardiovascular diseases often develop gradually, making early diagnosis extremely important for preventing severe complications and reducing death rates. However, conventional diagnostic techniques usually rely on advanced medical equipment, experienced healthcare professionals, and time-consuming procedures, which limits their availability in rural and resource-limited regions.

The growth of digital healthcare and the use of data-driven technologies have created new opportunities to overcome these challenges. The increasing collection of medical data and improvements in computational methods have enabled the development of intelligent systems capable of predicting diseases at an early stage.

In this project, a heart disease prediction system is developed using a logistic regression model to analyze important clinical and demographic parameters such as age, blood pressure, cholesterol levels, heart rate, and ECG-related features. Logistic regression is particularly suitable for this application due to its simplicity, computational efficiency, and high interpretability. The model not only predicts the presence or absence of heart disease but also helps in understanding the impact of each health parameter, supporting healthcare professionals in making accurate, reliable, and transparent decisions.



Fig 1: Heart Health Assessment

## II. PROBLEM STATEMENT

Heart disease often remains undetected during its early stages due to the absence of noticeable symptoms. The lack of affordable and easily accessible diagnostic tools further contributes to delayed detection and treatment, especially in resource-limited settings. As a result, many patients are diagnosed only when the condition becomes severe, increasing health risks and treatment costs.

Additionally, traditional manual diagnosis methods are time-consuming and highly dependent on medical expertise, making them prone to human error. To address these challenges, there is a growing need for an automated system that can accurately predict heart disease using easily available patient data. Such a system can support healthcare professionals in early detection, improve diagnostic accuracy, and enable timely medical intervention.

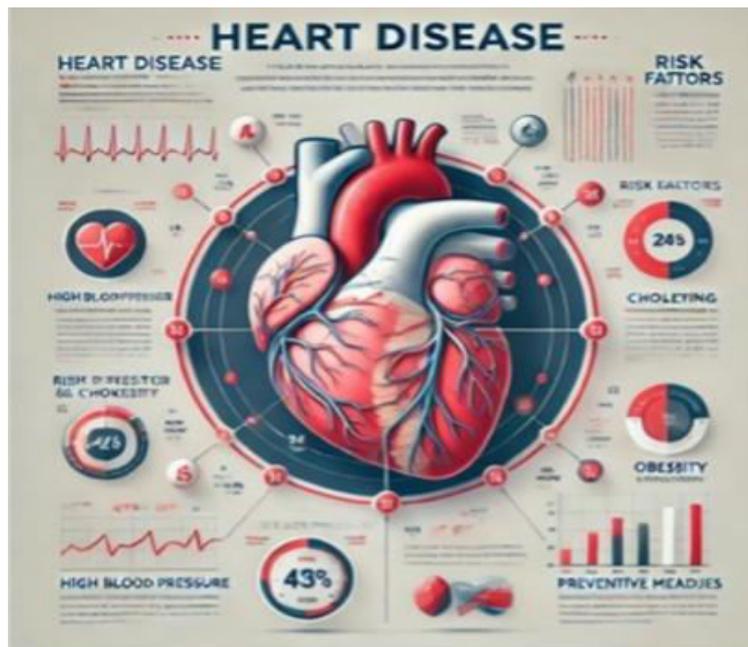


Fig 2: Problem Statement

III. LITERATURE SURVEY

Sr. No.	Publication	Inference
1	Mohammad Shafenoorr Amin, Yin Kia Chiam & Kasturi Dewi Varathan. <i>Identification of significant features and data mining techniques in predicting heart disease.</i>	This research focuses on improving heart disease prediction by identifying significant features and applying seven classification techniques. The hybrid Vote model, combining Naive Bayes and Logistic Regression, achieved the highest accuracy of 87.4%. Data mining enhances predictive accuracy by selecting key features from large healthcare datasets.
2	Mai Shouman, Tim Turner, Rob Stocker. <i>Using Data Mining Techniques in Heart Disease Diagnosis and Treatment.</i>	Data mining has been effective in diagnosing heart disease, the leading cause of death globally, with hybrid techniques showing improved accuracy. However, its application to identifying suitable treatments has received less attention. This paper addresses research gaps and proposes a model to explore the use of data mining for heart disease treatment prediction.
3	K. Srinivas, Dr. G. Raghavendra Rao, Dr. A. Govardhan. <i>Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal Mining Regions Using Data Mining Techniques.</i>	Heart disease is a leading cause of morbidity and mortality, with accurate medical diagnosis being crucial. This study analyzes self-reported cardiovascular disease rates in Singareni coal mining regions of Andhra Pradesh, India, using 15 attributes for prediction, including BMI, age, and income. Popular data mining techniques such as Decision Trees, Naive Bayes, and Neural Networks are applied for heart disease prediction. An automated diagnosis system could enhance healthcare and reduce costs.
4	Harshit Jindal, Sarthak Agarwal, Rishabh Khera, Rachna Jain and Preeti Nagrath. <i>Heart Disease Prediction using Machine Learning Algorithms.</i>	Heart disease cases are rising rapidly, making early prediction crucial. This research develops a heart disease prediction system using machine learning algorithms such as Logistic Regression and KNN to classify patients based on medical history. The proposed model shows better accuracy than Naive Bayes, effectively predicting heart disease, improving diagnosis precision, and reducing healthcare costs. The system is implemented in a .ipynb format to enhance medical care.
5	Jayshril S. Sonawane, D. R. Patil. <i>Prediction of Heart Disease using Learning Vector Quantization Algorithm.</i>	Heart disease diagnosis is complex and often relies on medical practitioners' expertise, which can lead to errors and delays. With advancements in machine learning, there is increasing interest in automated heart disease prediction systems. This paper presents a prediction system using the Learning Vector Quantization neural network algorithm, which analyzes 13 clinical features to predict the presence or absence of heart disease and evaluates its performance using various measures.
6.	Diaa Salama AbdElminaam, Nada Mohamed, Hady Wael, Abdelrahman Khaled, Adham Moataz. <i>MLHeartDisPrediction: Heart Disease Prediction using Machine Learning.</i>	Early prediction of heart disease is vital for saving lives, given its status as a leading cause of death. This study enhances prediction reliability using machine learning on three datasets with algorithms such as Logistic Regression and Random Forest. Logistic Regression achieved accuracies of 91.6% and 90.8%, while Random Forest reached 98.6%, demonstrating the effectiveness of these techniques in heart disease prediction. Further improvements can be achieved through collaboration with medical professionals.

#### IV. METHODOLOGY

The proposed methodology presents a structured framework for predicting heart disease using machine learning techniques. The system integrates user-provided health parameters with historical medical data to generate accurate and reliable predictions, supporting early diagnosis and informed decision-making.

##### **Data Collection and User Input**

The process begins with the collection of health-related data from two primary sources: user input and existing health datasets. User input includes key medical parameters such as heart rate, blood pressure, cholesterol levels, and other relevant clinical indicators. Simultaneously, historical health data stored in the database is utilized to enrich the prediction process and improve model learning.

##### **Data Processing and Feature Preparation**

The collected data undergoes preprocessing to ensure quality and consistency. This step includes handling missing values, normalization of numerical attributes, and selection of relevant features. Proper preprocessing ensures that the machine learning algorithms receive clean and structured data, which enhances prediction accuracy.

##### **Application of Machine Learning Algorithms**

Preprocessed data is fed into machine learning algorithms for analysis. Logistic Regression is primarily used due to its effectiveness in binary classification problems such as heart disease prediction. In parallel, a Decision Tree model is also applied to capture nonlinear relationships between health parameters and disease outcomes.

##### **Prediction Generation**

Both models are trained using the processed health data. After training, prediction results are generated based on the input parameters. These results indicate the likelihood of heart disease presence, enabling risk assessment at an early stage.

##### **Model Comparison and Selection**

The performance of Logistic Regression and Decision Tree models is evaluated using standard metrics such as accuracy, precision, and recall. Based on comparative analysis, the best-performing model is selected to ensure reliable and efficient prediction results.

##### **Final Prediction Output**

The selected model produces the final prediction, which is presented to the user in an understandable format. This output supports healthcare professionals and individuals in making timely medical decisions and adopting preventive measures for heart disease management.

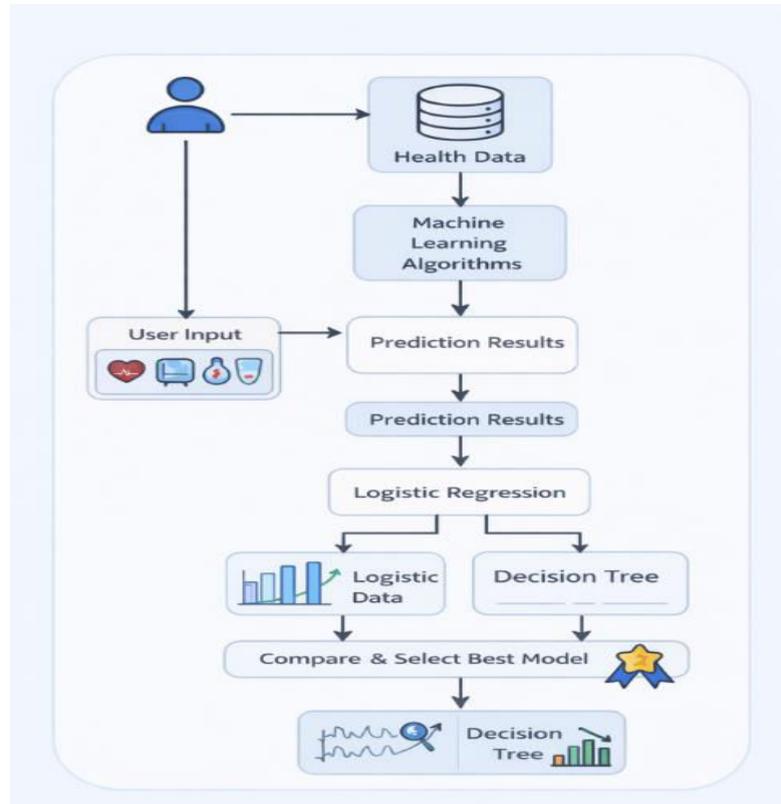


Fig 3: Proposed Methodology

## V. KEY FEATURES

The Heart Disease Prediction System incorporates several important features designed to support early diagnosis, improve prediction accuracy, and assist healthcare decision-making using machine learning techniques:

- **Early Disease Prediction:** The system predicts the risk of heart disease at an early stage using patient health parameters, helping in timely medical intervention and prevention.
- **Machine Learning–Based Analysis:** Logistic Regression is used as the primary algorithm to analyze medical data and classify patients as having heart disease or not with reliable accuracy.
- **Multiple Model Comparison:** The system compares different machine learning models such as Logistic Regression and Decision Tree to select the best-performing model for final prediction.
- **Use of Clinical Health Data:** The model utilizes key clinical attributes like age, blood pressure, cholesterol, heart rate, and other medical indicators to ensure accurate and meaningful predictions.
- **User-Friendly Interface:** The system provides an easy-to-use interface where users can input their health details and quickly receive prediction results.
- **Cost-Effective and Time-Efficient Diagnosis:** By automating the prediction process, the system reduces dependency on manual diagnosis, saving time, lowering healthcare costs, and minimizing human error.

## VI. CHALLENGES AND RESEARCH GAPS

Despite the growing use of machine learning in healthcare, several challenges and research gaps still exist in heart disease prediction systems. One major challenge is the **limited availability of high-quality and diverse medical datasets**, especially in developing regions, which affects the generalization and reliability of prediction models. Many existing studies rely on small or region-specific datasets, making it difficult to apply the models universally across different populations.

From a technical perspective, **model accuracy and interpretability** remain key concerns. While machine learning algorithms can achieve good accuracy, understanding how predictions are made is often difficult, which can reduce trust among healthcare professionals. Ethical issues such as **data privacy, security of patient information, and algorithmic bias** also pose significant challenges. Additionally, integrating these systems into real-world clinical environments requires further research, particularly in improving real-time performance, scalability, and collaboration between automated systems and medical experts to ensure safe and effective decision-making.

## VII. CONCLUSION

The logistic regression-based heart disease prediction model developed in this study demonstrates the significant potential of machine learning techniques in improving healthcare diagnostics and clinical decision-making. With an accuracy exceeding 80%, the model proves to be reliable and effective in identifying individuals at high risk of heart disease at an early stage. Early risk detection enables timely medical intervention, continuous monitoring, and preventive care, which can significantly improve patient outcomes. By leveraging clinical and demographic data, the proposed system serves as a supportive and interpretable diagnostic tool, especially in areas with limited access to specialized healthcare professionals, thereby helping to reduce the overall burden of heart disease.

## VIII. FUTURE SCOPE

Future enhancements will focus on improving the model's accuracy and real-world applicability by incorporating additional risk factors such as genetic information, lifestyle habits, and more detailed clinical parameters. Exploring advanced machine learning approaches, including ensemble methods and deep learning techniques, may further enhance predictive performance. Additionally, the development of a user-friendly interface will be prioritized to ensure ease of use for both healthcare providers and patients. These improvements will help evolve the system into a more robust, accessible, and impactful solution for the early prediction and prevention of heart disease.

## REFERENCES

1. Mohammad Shafenoor Amin, Yin Kia Chiam & Kasturi Dewi Varathan. Identification of significant features and data mining techniques in predicting heart disease.
2. Cincy Raju, Philipsey, E., Siji Chacko, L Padma Suresh, & Deepa Rajan
3. S. A Survey on Predicting Heart Disease using Data Mining Techniques.
4. Sana Bharti, Dr. Shaliendra Narayan Singh. Analytical Study of Heart Disease Prediction Comparing with Different Algorithms.
5. Montu Saw, Tarun Saxena, Sanjana Kaithwas, Rahul Yadav, Nidhi Lal. Estimation of Prediction for Getting Heart Disease Using Logistic Regression Model of Machine Learning.
6. Yanwei Xing, Jie Wang and Zhihong Zhao, Yonghong Gao. Combination data mining methods with new medical data to predicting outcome of coronary heart disease.
7. K. Srinivas, Dr. G. Raghavendra Rao, Dr. A. Govardhan. Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal Mining Regions Using Data Mining Techniques.
8. Mai Shouman, Tim Turner, Rob Stocker. Using Data Mining Techniques in Heart Disease Diagnosis and Treatment.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details