



HEART DISEASE PREDICTION USING MACHINE LEARNING

Amit Meshram¹, Abhishek Pawar², Pratiksha Tidke³, Tanu Rangarkar⁴, Komal Rewaskar⁵

Professor, Department of Computer Science and Engineering,

Nagarjuna Institute of Engineering Technology and Management, Nagpur, Maharashtra, India¹

UG Students, Department of Computer Science and Engineering,

Nagarjuna Institute of Engineering Technology and Management, Nagpur, Maharashtra, India²⁻⁵

Abstract: Heart disease is one of the leading causes of death worldwide. Early diagnosis and prediction can play a vital role in preventing life-threatening conditions. The traditional methods for predicting heart disease are often manual, time-consuming, and prone to errors. In this research, a machine learning-based model is proposed to predict the likelihood of heart disease based on clinical data such as age, gender, blood pressure, cholesterol, and other medical attributes. Various algorithms like Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM) were implemented and compared. The dataset used was the UCI Heart Disease Dataset. The results show that Random Forest Classifier achieved the highest accuracy of 88.5%, making it a reliable model for real-world applications. The study aims to assist medical practitioners in making better and faster diagnostic decisions.

Keywords: Heart disease prediction, Data mining, Risk factors, Feature selection, Real-world healthcare data, Neural network, Deep learning

I. INTRODUCTION

Cardiovascular diseases (CVDs) are among the most severe health problems globally. According to the World Health Organization (WHO), around 17.9 million people die each year from cardiovascular diseases, representing 32% of global deaths. Early prediction can significantly reduce the mortality rate by enabling preventive healthcare measures.

Traditional medical diagnosis relies on the expertise of physicians and manual interpretation of medical reports, which may lead to human errors. Machine learning (ML), a branch of Artificial Intelligence (AI), can automate and enhance the accuracy of disease prediction by analysing large sets of medical data.

This research focuses on developing a Heart Disease Prediction System using ML algorithms. The system analyses patient data and predicts whether a person is at risk of heart disease.

II. METHODOLOGY

The proposed system uses machine learning models trained on patient health data. The methodology consists of several phases:

2.1 Data Collection

The dataset used is the Cleveland Heart Disease Dataset from the UCI Machine Learning Repository.

It includes 14 medical attributes such as:

- Age, Sex, Chest Pain Type (cp)
- Resting Blood Pressure (trestbps)
- Serum Cholesterol (Chol)
- Fasting Blood Sugar (fibs)
- Resting ECG (restitch)
- Maximum Heart Rate (thelarche)
- Exercise Induced Angina (exchange)
- Old peak (ST depression)
- Slope of ST segment (slope)
- Major Vessels (ca)
- Thalassemia (that)
- Target (presence of heart disease)



2.2 Data Preprocessing

- Handling missing values
- Data normalization
- Encoding categorical variables
- Splitting dataset (80% training, 20% testing)

2.3 Model Training

Machine learning algorithms applied:

1. Logistic Regression
2. Decision Tree Classifier
3. Random Forest Classifier
4. Support Vector Machine (SVM)

III. FLOWCHART

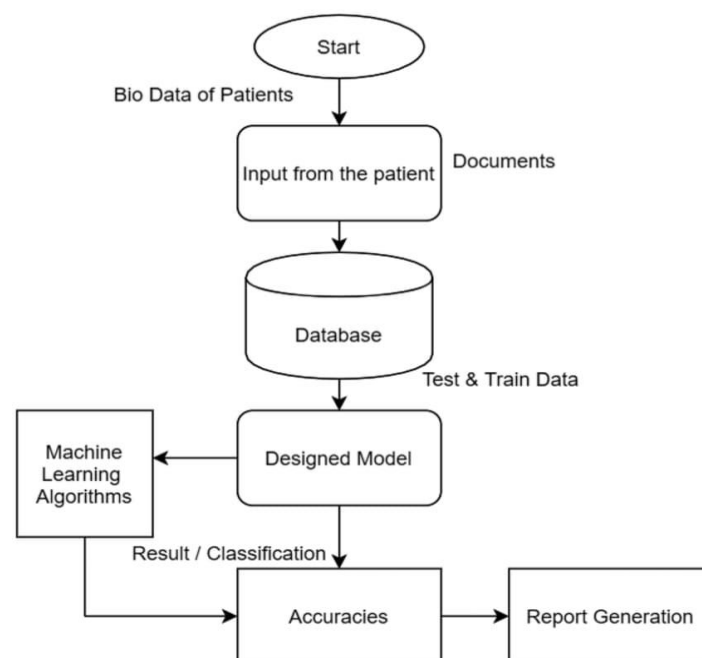


Fig. 1 Flow of Heart Disease Prediction

IV. RESULTS AND DISCUSSION

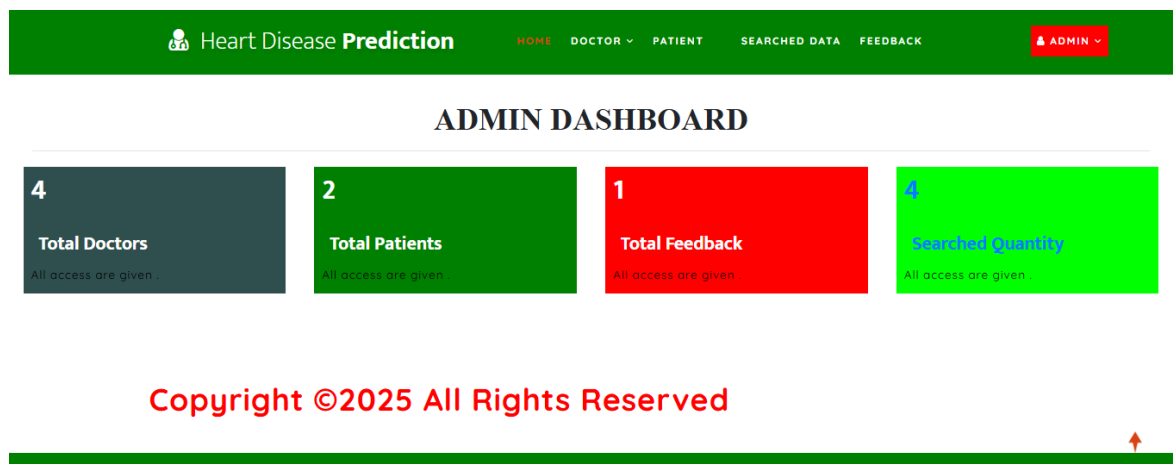


Fig.02 Admin Dashboard



REGISTER NOW

First Name First Name	Last Name Last Name
Username Username	Password Password
Email Enter Email	Contact Enter Contact
Date Of Birth mm/dd/yyyy	Image Choose File No file chosen
City Enter Address	User Type ☐ Patient ☐ Doctor

REGISTER

Fig.03 Register page

Heart Disease Prediction [HOME](#) [PREDICT](#) [MY DETAIL](#) [FEEDBACK](#) [HISTORY](#) **HELLO, PRATIKSHA**

ADD HEART DETAIL

Age 50	Sex F	CP 2	Trestbps 120	Cholestrol 219	Fbs 0	Restecg 1	Thalach 158	Exang 1	OldPeak 1.6	Slope 1	CA 0
--	---------------------------------------	--------------------------------------	--	--	---------------------------------------	---	---	---	---	---	--------------------------------------

Thal

SEND HEART DATA

©2025 Heart Disease Prediction [Facebook](#) [Twitter](#) [Google Plus](#)

Fig 04. Add Heart Details

Heart Disease Prediction [HOME](#) [PREDICT](#) [MY DETAIL](#) [FEEDBACK](#) [HISTORY](#) **HELLO, PRATIKSHA**

View Searched Data

Copy Excel CSV PDF Search:

#	Date	Accuracy	Result	HeartValue	Action
1	Nov. 2, 2025, 5:52 p.m.	88.52459016393442	Healty	['22', '1', '93', '83', '84', '123', '45', '32', '321', '35', '5', '3', '4']	Delete
2	Nov. 2, 2025, 5:49 p.m.	88.52459016393442	Healty	['23', '1', '45', '76', '87', '5', '55', '5', '67', '77', '4', '3', '4']	Delete
3	Oct. 31, 2025, 5:23 p.m.	88.52459016393442	Unhealty	['50', '1', '2', '120', '219', '0', '1', '158', '1', '1.6', '1', '0', '2']	Delete
4	Oct. 31, 2025, 8:49 a.m.	88.52459016393442	Unhealty	['41', '0', '1', '130', '204', '0', '0', '172', '0', '1.4', '2', '0', '2']	Delete

Showing 1 to 4 of 4 entries Previous Next

©2025 Heart Disease Prediction [Facebook](#) [Twitter](#) [Google Plus](#)

Fig 05. Data Store



Heart Disease Prediction

HOME

DOCTOR

PATIENT

SEARCHED DATA

FEEDBACK

ADMIN

View Doctor

Copy

Excel

CSV

PDF

Search:

#	Full Name	Image	Email	Contact	Address	Category	Status	Assign	Action
1	Ram Rahate			2435465768	Amravati	Cardiologist	Authorized	Cancel	
2	Dr.T.R Murlidhar			1234567890	Hydrabad	Cardiologist	Authorized	Cancel	
3	Dr.Sarita Rao			9518924507	Delhi	cardiologist	Authorized	Cancel	
4	khushi Singh			1234567890	Nashik	cardiologist	Authorized	Cancel	

Fig 06. All Doctors Member

The Random Forest Classifier outperformed other models in terms of accuracy and stability.

The model successfully predicted heart disease based on medical attributes.

Visualization of feature importance showed that cholesterol, age, and max heart rate were among the top contributing factors.

V. CONCLUSION

Heart disease prediction conclusions highlight the success of machine learning in identifying risk factors like age, gender, blood pressure, cholesterol, and lifestyle habits, which can lead to more accurate and timely diagnoses. Models using methods like Random Forest and multilayer perceptron can achieve high accuracy, enabling early intervention, personalized treatment, and reduced mortality rates. However, the accuracy of these models depends on data quality, the relevance of features, and the handling of complex, high-dimensional datasets. The heart disease prediction model demonstrates how data-driven techniques can play a vital role in healthcare decision-making. By analysing key health indicators such as age, blood pressure, cholesterol levels, and other clinical factors, the model can effectively estimate the likelihood of heart disease in individuals. The results show that machine learning algorithms—particularly [insert algorithm used, e.g., Logistic Regression, Random Forest, or Neural Networks]—can achieve high accuracy and provide reliable predictions.

REFERENCES

- [1]. UCI Machine Learning Repository: Heart Disease Dataset.
- [2]. Detrano, R. et al., International Application of a New Probability Algorithm for the Diagnosis of Coronary Artery Disease, American Journal of Cardiology, 1989.
- [3]. J. Han, M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2011.
- [4]. Scikit-learn Documentation, [https://scikit-learn.org/] (https://scikit-learn.org/)
- [5]. Zhang, Y. et al., Heart Disease Prediction using Machine Learning Algorithms, IJRTE, 2022.
- [6]. World Health Organization (WHO) – Cardiovascular Diseases Report, 2023.
- [7]. Cardiac disease risk prediction using machine learning algorithms, *PMC*.
- [8]. A Survey on Machine Learning Techniques for Heart Disease Prediction, *Springer*. [SpringerLink](#)
- [9]. Heart Disease Detection Using Machine Learning Methods (JMAI). [jmai.amegroups.org](#)
- [10]. Machine learning-based heart disease diagnosis, PubMed. [PubMed](#)
- [11]. Optimizing heart disease diagnosis with advanced machine learning, *BMC Cardiovascular Disorders*.