iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70225

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Predictive Modeling for Heart Diseases Detection

Ayush Vaidya[1], Hemangi Pustode[2], Janhavi Lede[3], Om Kolhe[4], Shweta Suryawanshi[5], Omkar Dudhbure[6]
[1, 2, 3, 4, 5]*Student,* [6]*Assistant Professor, Computer Engineering Department, RTM Nagpur University*

Abstract: ***In this Research paper focuses on the development and application of predictive modeling techniques for the early detection of heart disease. Heart disease remains a leading cause of death globally, making early diagnosis and prevention essential. This project seeks to develop a reliable system for predicting the risk of heart disease by utilizing modern machine learning and data analysis techniques, drawing on patient data such as demographics, lifestyle habits, medical background, and clinical test results. By applying various predictive algorithms, such as decision trees, support vector machines, and deep learning models, the system is trained to identify patterns and correlations within the dataset that are indicative of potential cardiovascular issues. The project also emphasizes the use of feature selection techniques to enhance model accuracy and efficiency while mitigating overfitting. The end goal is to create an automated, real-time decision support tool for healthcare providers, enabling them to diagnose heart disease risk more effectively and provide timely interventions.***
Keywords: Predictive modelling, Heart Disease Detection, Prediction, Real-time data, Machine learning, Wearable devices.

I. INTRODUCTION

Introduction to Predictive Modeling and Heart Disease Detection Mega Project Cardiovascular diseases (CVDs) are a leading cause of death worldwide, contributing to millions of lives lost each year. Early detection and intervention are critical to improving patient outcomes and reducing the burden of heart disease on global health systems. However, diagnosing heart disease often requires complex, time-consuming processes and the integration of multiple risk factors, which makes timely and accurate predictions challenging. This is where predictive modeling and advanced analytics can play a transformative role.

The goal of this project is to develop an intelligent, data-driven solution for the early detection of heart disease, leveraging predictive modeling techniques to enhance diagnosis accuracy and decision-making in healthcare. By analyzing a combination of demographic data, medical history, lifestyle habits, and clinical test results, the project aims to predict the likelihood of heart disease in individuals, even before symptoms become apparent.

Predictive modeling in healthcare uses historical patient data and advanced machine learning algorithms to uncover patterns and correlations that might otherwise go unnoticed. In the context of heart disease, various algorithms—such as regression models, support vector machines, decision trees, and neural networks—are employed to classify and predict patient risk levels. These tools can help healthcare providers make more informed decisions and offer personalized care plans that are tailored to each patient's unique risk profile.

As part of this project, we will also explore the integration of real-time data collected from wearable devices, providing continuous monitoring of patients' cardiovascular health. By combining these real-time metrics with historical data, the predictive model will improve its ability to track changes in health status and provide timely alerts for potential heart disease.

Ultimately, this mega project aspires to enhance the efficiency of heart disease detection and prevention, contributing to improved patient outcomes, reduced healthcare costs, and better public health. By utilizing the power of predictive analytics, the project envisions creating a system that is not only accurate and reliable but also scalable and accessible across different healthcare settings worldwide.

II. LITERATURE SURVEY

In this paper [1] detection and prediction of heart disease have long been subjects of intense research due to the increasing prevalence of cardiovascular diseases (CVDs) globally. Over the years, numerous studies and methodologies have emerged, harnessing predictive modeling.

Methods for improving early detection and enhancing patient health outcomes. This literature survey reviews the key research and developments in the field of predictive modeling for heart disease detection, with a focus on machine learning, feature selection techniques, and integration of real-time health data. [2]

Predictive Modeling Techniques in Heart Disease Detection: -

A variety of machine learning algorithms have been used to predict the likelihood of heart disease based on patient data. [3]

©IJRASET: All Rights are Reserved

Logistic Regression and Decision Trees: Traditional statistical methods like logistic regression have been commonly used for predicting heart disease risk, as highlighted by Ozdemir and Gunes (2017) These models typically use specific factors like age, cholesterol levels, blood pressure, and smoking status to estimate the risk of heart disease. Breiman (2001) introduced decision trees and random forests, which were later applied in heart disease prediction, demonstrating significant improvements in classification accuracy by capturing complex relationships between the features. [4]

Support Vector Machines (SVMs): Vimal et al. (2016) explored the use of support vector machines for heart disease prediction. SVMs are well-known for their ability to classify large datasets with high-dimensional features, achieving high accuracy rates in identifying heart disease risks by creating an optimal hyperplane that separates patients at risk from those who are not. [5]

Artificial Neural Network (ANNs): Deep learning models, such as artificial neural networks, have shown considerable promise in heart disease prediction. Chaurasia et al. (2018) used multilayer perceptron's (MLPs) to predict heart disease, demonstrating that ANNs outperform traditional algorithms when provided with large and diverse datasets. These networks can capture nonlinear relationships between input features, which is essential when analyzing complex medical data.[6] Feature Selection and Dimensionality Reduction: - A significant challenge in heart disease prediction is the high dimensionality of the dataset and the presence of irrelevant or redundant features. Several studies have focused on improving prediction accuracy by selecting the most informative features.

Correlation-based Feature Selection (CFS), introduced by Hall in 1999, is a method that assesses the relevance of features by considering their interrelationships. In heart disease detection, CFS has been widely used to reduce the number of features and improve model efficiency without compromising accuracy.[7]

Principal Component Analysis (PCA) is a widely used method for reducing data dimensions by highlighting the most important variables in a dataset. Jain et al. (2015) utilized PCA to transform the original feature space into a new space that maximizes variance, thus improving prediction performance. Genetic Algorithms for Feature Selection: Chavez et al. (2019) investigated the use of genetic algorithms (GA) for optimizing feature selection. GA helps find the optimal subset of features by mimicking the process of natural evolution, thus ensuring the highest predictive power for heart disease classification models.

Real-time Health Data and Integration with Wearables: -In recent years, the advent of wearable devices and continuous health monitoring systems has opened new avenues for heart disease prediction. These devices are capable of delivering instant readings of heart rate, blood pressure, oxygen levels, and other key indicators of cardiovascular health. Wearable Devices for Continuous Monitoring: Research by Zhao et al. (2020) explored the integration of wearable health devices with predictive models for heart disease detection. These devices provide continuous monitoring, enabling the prediction of heart disease risk in real-time by incorporating data such as daily activity levels, heart rate variability, and sleep patterns.[8]

Challenges and Limitations: -While significant progress has been made, there remain several challenges and limitations in predictive modeling for heart disease detection. Data Quality and Imbalance: One of the primary challenges in heart disease prediction is dealing with imbalanced datasets, where instances of heart disease are fewer than those without it. Techniques such as oversampling, under sampling, and synthetic data generation (e.g., SMOTE) have been applied to address this issue, but they remain an ongoing challenge in ensuring fairness and accuracy in predictions. Interpretability of Models: While machine learning models like deep learning and random forests offer high accuracy, they often act as "black boxes," making it difficult to interpret how predictions are made. Ribeiro et al. (2016) introduced the concept of Explainable AI (XAI), which provides interpretability to complex models. In heart disease detection, interpretability is crucial to ensure that healthcare professionals can trust and act on model recommendations. Generalization to Diverse Populations: Many predictive models are trained on data from specific populations, leading to potential biases when applied to other demographic groups. Ensuring that models generalize well to diverse populations remains a key improvement in predictive modeling for heart disease.[9]

Future Directions: -The future of predictive modeling for heart disease detection lies in the integration of multi-modal data, such as genetic information, lifestyle data, and real-time monitoring from wearables. The combination of these data sources will create a more comprehensive and personalized approach to heart disease prediction. Moreover, the continued advancement of deep learning models and the incorporation of Explainable AI techniques will make predictive models more reliable and understandable for clinicians. Predictive modeling for heart disease detection has evolved significantly, driven by advancements in machine learning, data analytics, and wearable technology. By combining these techniques with real-time health data, the potential to improve early detection and prevention of heart disease has never been greater. However, challenges such as data quality, model interpretability, and generalization to diverse populations remain important areas for future research and development. The findings from this literature survey will serve as the foundation for the mega project, guiding the design and implementation of an effective heart disease detection system that leverages cutting-edge predictive modeling techniques.[10]

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

III.EXSISTING SYSTEM

The current approach to heart disease detection largely depends on conventional diagnostic techniques that involve evaluating a patient's demographic information, lifestyle habits, medical history, and results from clinical tests such as ECG, cholesterol levels, and blood pressure. These assessments are manually reviewed by healthcare professionals to identify potential signs of cardiovascular conditions.

While these methods are well-established in medical practice, they face several limitations:

- Time-consuming and subject to human error.
- Limited ability to analyze complex, non-linear relationships between multiple risk factors.
- Dependent on symptomatic presentation, potentially missing early-stage or asymptomatic cases.
- Lack of personalized risk assessment, as most assessments use general clinical thresholds.
- Often not integrated with real-time data from modern health monitoring tools like wearable devices.

Due to these shortcomings, there is a growing need for intelligent, data-driven systems that can enhance diagnostic accuracy, provide early warnings, and support preventive care strategies.

IV.PROPOSED SYSTEM

Our predictive modeling in heart disease detection aims to leverage advanced machine learning algorithms and data analytics to predict the likelihood of heart diseases in patients. By utilizing a large dataset of clinical records, including patient demographics, medical history, lifestyle factors, and diagnostic test results, the system will apply techniques such as classification models (e.g., decision trees, random forests, and support vector machines) to identify patterns and risk factors associated with heart diseases. The model will be trained to classify individuals into high-risk and low-risk categories, enabling healthcare professionals to intervene early and prioritize patients who need immediate attention. Additionally, the system will incorporate real-time data, such as wearable devices that monitor heart rate and other vital signs, allowing for continuous assessment and dynamic risk predictions. By combining both historical and real-time data, the system aims to provide a comprehensive, data driven approach.

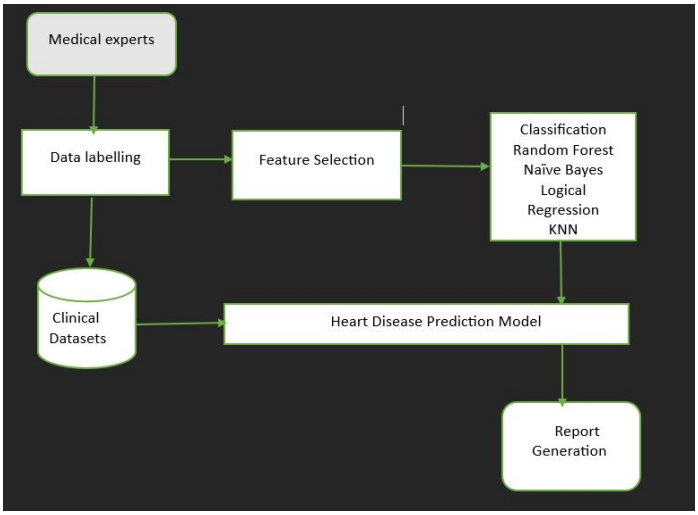


Fig-1-flow diagram of predictive modeling for heart diseases detection

The system architecture for a predictive modeling project focused on heart disease detection typically follows a layered and modular approach, integrating various components to handle data ingestion, preprocessing, model training, and prediction. At the core, it begins with data acquisition, which involves collecting structured and unstructured data from multiple sources such as electronic health records (EHR), medical imaging, patient history, and wearable devices. This data is then processed in the data preprocessing layer, where it undergoes cleaning, feature extraction, normalization, transformation to ensure high-quality input for modeling.

The model training layer involves applying machine learning algorithms (e.g., logistic regression, decision trees, support vector machines, or deep learning methods) to the processed data. The model may be fine-tuned using techniques such as cross-validation,

hyperparameter optimization, and feature selection to improve accuracy. Models are trained on historical data to identify patterns and predict potential heart disease occurrences based on patient characteristics, medical history, and lifestyle factors.

Once the model is trained, it is deployed in the prediction layer, where it can receive real-time or batch data and generate predictions. The prediction results are passed through the interpretation layer, which might include visualization tools and explainability algorithms to provide actionable insights for healthcare professionals. Additionally, the monitoring and evaluation layer tracks the model's performance over time, ensuring that it maintains high predictive accuracy and adapts to evolving trends in medical data. Finally, the architecture may include integration with healthcare systems like Electronic Health Records (EHR) and clinical decision support systems (CDSS) to provide seamless and timely alerts, recommendations, and patient management.

Overall, the architecture must be scalable, secure, and compliant with healthcare data regulations, ensuring patient privacy while delivering reliable, real-time predictions to support clinical decision-making.

V. CONCLUSIONS

Predictive modeling is transforming heart disease detection by using machine learning and data analytics to assess individual risk based on clinical factors like age, gender, cholesterol, blood pressure, ECG results, and more. These models, built with tools like Scikit-learn, TensorFlow, and PyTorch, utilize algorithms such as logistic regression, decision trees, SVMs, and deep learning for accurate predictions.

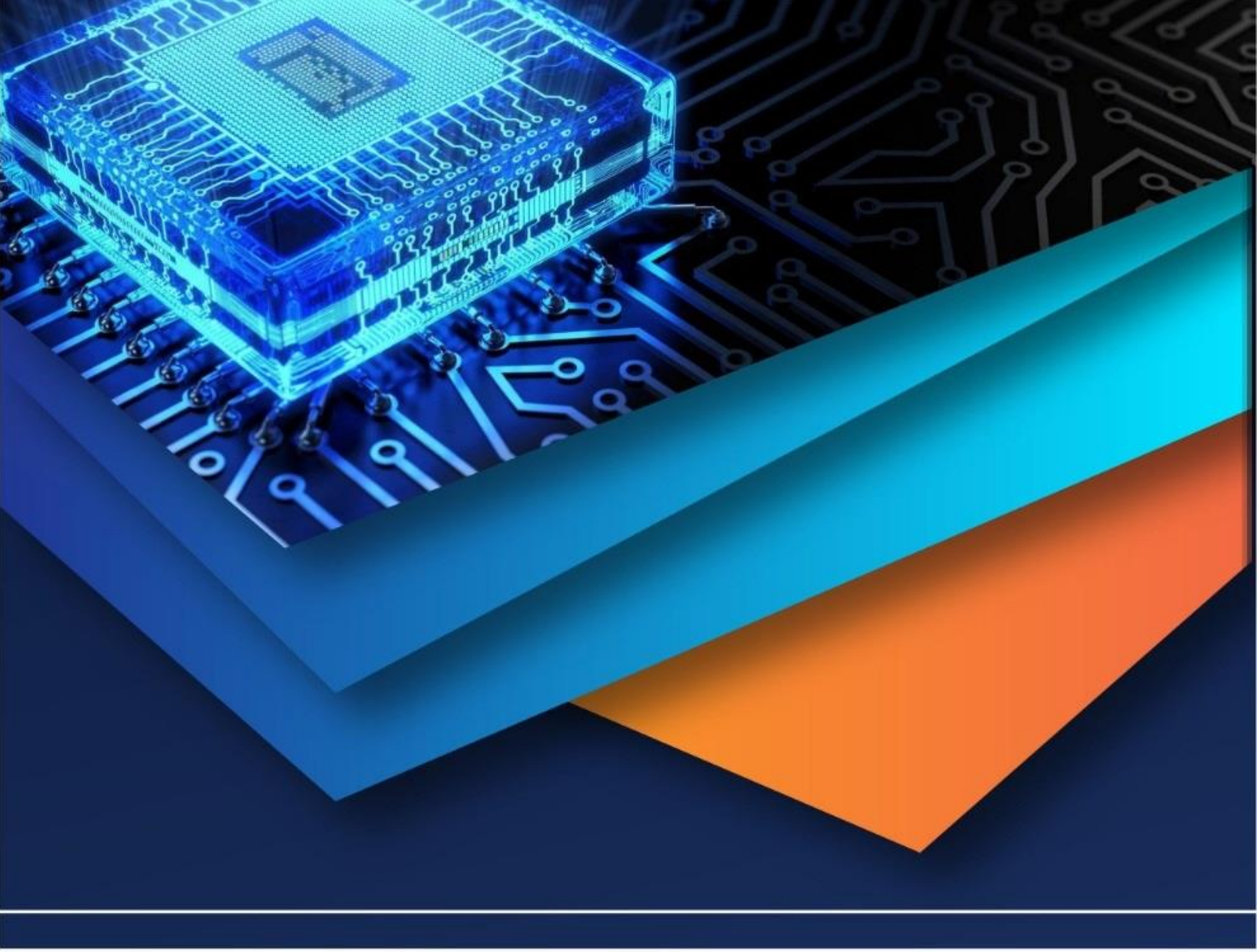
With user-friendly interfaces, healthcare providers can input patient data, which is then analyzed in real time—often integrated with cloud systems and electronic health records (EHRs). Beyond early diagnosis, these models support personalized treatment, risk stratification, and proactive care.

Future enhancements include integrating wearable device data for continuous monitoring, making predictive modeling a vital tool in preventive cardiology and improving long-term health outcomes.

REFERENCES

[1] M.P. Behera, A. Sarangi, D. Mishra, S.K. Sarangi A Hybrid Machine Learning algorithm for Heart and Liver Disease Prediction Using Modified Particle Swarm Optimization with Support Vector Machine Procedia Computer Science, 218 (2023), pp. 818-827.
[2] A. Ghasemieh, A. Lloyed, P. Bahrami, P. Vajar, R. Kashef A novel machine learning model with Stacking Ensemble Learner for predicting emergency readmission of heart-disease patients Decision Analytics Journal, 7 (2023), Article 100242.
[3] C.M. Bhatt, P. Patel, T. Ghetia, P.L. Mazzeo Effective heart disease prediction using machine learning techniques Algorithms, 16 (2) (2023), p. 88.
[4] Shore Wala, V. Early detection of coronary heart disease using ensemble techniques. Inform. Med. Unlocked2021,26, 100655. [CrossRef]
[5] Coffey, S., Roberts-Thomson, R., Brown, A., Carapetis, J., Chen, M., Enriquez-Sarano, M., ... & Prendergast, B. D. (2021). Global epidemiology og valvular heart disease. Nature Reviews Cardiology,18(12),853-864.
[6] De Hert, M., Detraux, J., & Vancampfort, D. (2022). The intriguing relationship between coronary heart disease and mental disorders. Dialogues in clinical neuroscience.
[7] Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Alonso, A., Beaton, A. Z., Bittencourt, M. S., ... & American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2022). Heart disease and stroke statistics—2022 update: a report from the American Heart Association. Circulation, 145(8), e153-e639.
[8] Katarya, R., & Srinivas, P. (2020, July). Predicting heart disease at early stages using machine learning: A survey. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 302-305). IEEE.
[9] Ahsan, M. M., & Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. Artificial Intelligence in Medicine, 128, 102289.
[10]] Priyanka S. Sangle; R. M. Goudar; A.N. Bhute2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)