IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

Development of a Heart Disease Prediction Model using Artificial Intelligence

Prof. Aditya Lawhale[1], Prof. Diksha Fulzele[2], Ayush Singh[3]
Assistant Professor, Computer Science & Engineering[1,2]
UG Student, Computer Science & Engineering[3]
Tulsiramji Gaikwad Patil College of Engineering and Technology, India
aditya.cse@tgpcet.com, diksha.cse@tgpcet.com, ayushsingh0968@gmail.com

Abstract: *Research on the development and implementation of a Heart Disease Prediction Model has been done. The research focuses on developing a heart disease prediction model based on AI, which helps to predict the probability of a patient getting a heart disease by analyzing various medical parameters. The system is designed to enhance accuracy and reliability by using appropriate machine learning algorithms and data preprocessing techniques. The model has been trained and tested using heart disease datasets, and the results have shown that AI can be used to predict heart disease effectively, which helps healthcare professionals to choose the appropriate diagnosis. The system is aimed at enhancing diagnostic efficiency, reducing the need for unnecessary tests, and improving health outcomes..*

Keywords: Heart Disease Prediction

I. INTRODUCTION

Heart disease is the most common and most dangerous of all health problems. It is responsible for a large number of deaths that occur each year. It encompasses different heart-related problems like coronary heart disease, heart failure, arrhythmias and more. According to the World Health Organization, cardiovascular disease is the number one cause of global deaths. They kill around 17.9 million people every year. The growing number of heart diseases suggests that early diagnosis and proper management are urgently needed.

The traditional ways of diagnosing heart diseases are associated with complex clinical procedures that are not only time consuming, but they are also costly and need specialized personnel. It is also worth noting that it might be too late before the symptoms can be identified due to the fact that most of them show up when the disease is already at a critical stage. Artificial Intelligence is now a powerful solution in the healthcare segment, offering data-driven, accurate, and scalable answers.

This Research Paper on Development of Heart Disease Prediction Model Using AI strives to investigate the possibility of predicting the presence of heart disease based on the specific parameters of the patient, such as age, blood pressure, cholesterol level, and other important clinical indicators. This model is proposed to increase the accuracy of diagnosis, reduce the dependence on manual interpretation and, in the end, help physicians to make an informed decision.

DOI: 10.48175/IJARSCT-27669

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

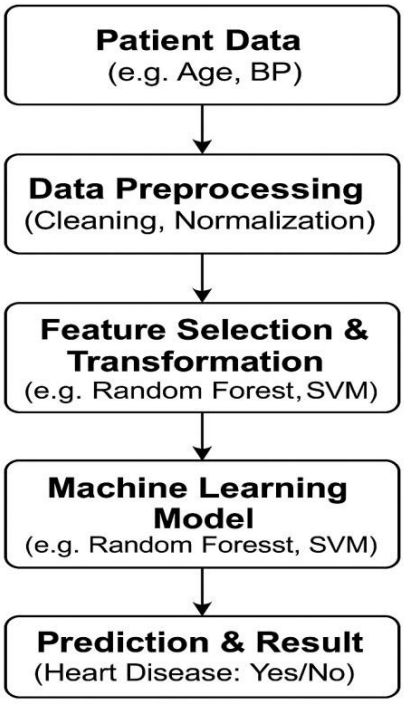


Figure 1: Architecture

II. LITERATURE REVIEW

In the medical field Artificial Intelligence (AI) and Machine Learning (ML) have received substantial attention in the context of medical diagnosis, thanks to the ability to analyse a large amount of data and identify patterns that are not immediately visible using traditional methods. Numerous studies have utilized AI techniques for the early prediction of heart disease with significant success.

Detrano et al. (1989) have introduced the Cleveland heart disease dataset that has become a benchmark for rating the prediction models of heart disease. It includes different medical parameters, like age, cholesterol level, and blood pressure, and which are usually used as inputs in machine learning-based diagnostic systems.

Gudadhe and others (2010) recommended utilizing Support Vector Machines (SVM), Decision Trees and Naïve Bayes classifiers in heart disease classification. Among them, SVM had the best results in identifying heart disease, underlining the model's adaptability with high-dimensional data.

Another study by Ayer et al. (2013) combined logistic regression and artificial neural networks (ANNs) for cardiovascular risk prediction based on clinical data. They showed that AI models significantly outperformed traditional risk scoring methods used by physicians.

Patel et al. (2016) conducted research on ensemble learning techniques like Random Forest and Gradient Boosting among others, which resulted in higher accuracy than individual algorithms.

Latest improvements include deep learning methods. In his research Mohan et al. (2019) has developed hybrid model, which is the combination of deep neural networks and optimization algorithms to enhance the accuracy and reliability of the heart disease detection.

Studies suggest that the progress of AI in healthcare diagnostics is dynamic and continuous. Thus, it does not already exclude the existing challenges and difficulties that AI faces in relation to the quality of the data used, interpretability of the AI models and their subsequent implementation in the medical practice.

This research builds upon these previous works by developing and evaluating multiple machine learning models on pre-processed clinical data, aiming to create a robust, interpretable, and accessible prediction system.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

III. METHODOLOGY

The research is based on the following steps: systematic, which is aimed at ensuring the effective building and testing of a predictive model for heart disease, and, ultimately, the final goal of creating an intelligent and reliable AI-based early diagnosis system.

3.1 Requirements Analysis

The first step is identifying the central issues for model development. This consists of comprehending the heart disease variables, using the right machine learning strategies, and choosing the performance proof metrics.

3.2 Dataset Collection

The dataset used in this study is acquired from publicly available sources like the UCI Machine Learning Repository. The dataset contains different characteristics such as age, gender, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, and maximum heart rate.

3.3 Data Preprocessing

Preprocessing is essential to ensure the dataset is clean and ready for model training. The following steps are performed:

- Handling missing values
- Encoding categorical variables
- Feature scaling using standardization or normalization
- Splitting the dataset into training and testing sets

3.4 Feature Selection

Techniques of choice feature selection are applied to identify the most relevant attributes in influencing disease. It helps to reduce noise in dataset and model leads to increased performance. Techniques such as recursive feature elimination can be taken into account.

3.5 Model Implementation

Various machine learning algorithms are implemented and compared, including:

- Logistic Regression
- Random Forest
- Support Vector Machine (SVM)

Each model is trained using the preprocessed dataset, with the training set used to learn patterns from input features such as age, cholesterol, resting blood pressure, and maximum heart rate. Hyperparameter tuning is performed using techniques like Grid Search or Randomized Search to optimize model performance.

3.6 Model Evaluation

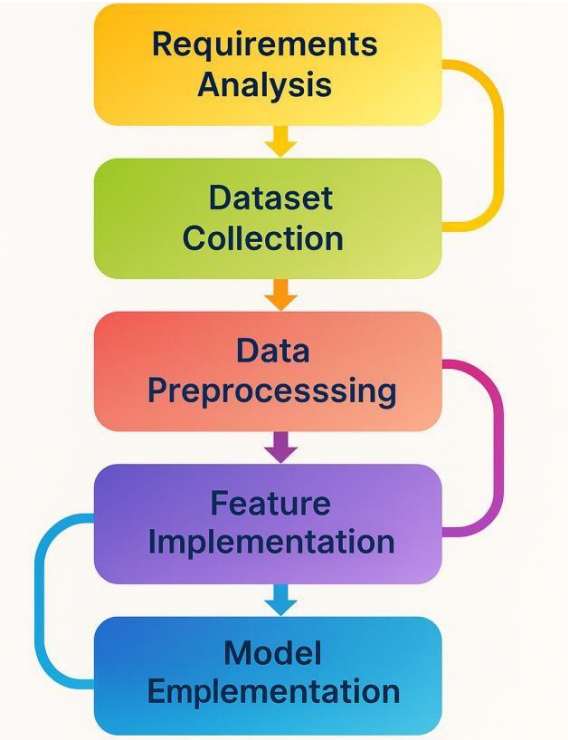
The models are evaluated using performance metrics such as:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix Cross-validation is also employed to ensure robustness and generalizability of the result.

3.7 Deployment and Integration

The final model is deployed using a user-friendly interface developed with tools like Streamlit, allowing users to input health data and receive instant predictions. This enhances accessibility for both healthcare professionals and patients.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67



IV. CHALLENGES

The heart disease prediction model using AI faces various obstacles even though it shows great potential during development. The research faced its primary hardship through issues involving available data quality and accessibility. The UCI Machine Learning Repository and other publicly accessible heart disease datasets provide limited data sets which lack comprehensive population representation according to size and diversity. Data preprocessing faced major hurdles because of missing values together with inconsistent data entries and the unavailability of real-time updates.

The main challenge emerged from handling data with imbalances which featured an excessive number of patients who lacked heart disease in comparison to those who had the condition. When the dataset contains imbalanced classes, models will prioritize the majority class excessively and perform poorly on the minority group which usually represents the most important medical diagnostic category.

The topic of model interpretability emerged as a significant source of concern. The complex algorithms Random Forest and Support Vector Machines demonstrate high accuracy yet their internal processes remain incomprehensible to users lacking technical expertise. In a clinical environment AI tool adoption faces resistance because healthcare practitioners need well-defined explanations regarding medical diagnostic outcomes.

The training process produced a problem called overfitting that occurred mainly with small datasets. The model demonstrated excellent performance during training but experienced reduced accuracy when tested on new data which demonstrates insufficient generalization ability. The situation required the adoption of both cross-validation methods and regularization techniques to achieve better model stability.

The management of patient data privacy and security brought an additional layer of complexity to the situation. The project required continuous attention to responsible and secure handling of sensitive health data as well as compliance with data protection guidelines throughout its duration. The introduction of this model to actual clinical processes creates practical obstacles. The system needs thorough evaluation to guarantee smooth integration with hospital systems and to guarantee user-friendly access for medical personnel together with sustainable model maintenance.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 4, June 2025
Impact Factor: 7.67

V. CONCLUSION

Created is an AI based heart disease prediction model in the research. The main goal was to help in early detection and diagnostic of cardiovascular conditions. The system uses machine learning methods and clinical data to build a model that can accurately and efficiently predict the probability of heart disease in people based on their health parameters.

Machine learning models like Logistic Regression, Random Forest, and Support Vector Machine can be effectively applied to classify patients with heart disease vs. no heart disease. The classification accuracy is promising. Data preprocessing, feature selection, and model evaluation enhanced the performance.

It has the potential to become a valuable decision-support tool for healthcare professionals. They could make rapid assessments and improve patient outcomes. This model also offers the opportunity for scalability and accessibility if it is deployed digitally. It could be integrated into current health systems and remote monitoring platforms.

Future research could concentrate on the incorporation of real- time data, expanding the dataset, implementing deep learning methods, and verifying clinical validation to amplify model accuracy and reliability. Additionally, ethical issues related to patient privacy and AI decision-making transparency will be important priorities in such systems deployment.

Despite its success, this study also acknowledges certain limitations. The accuracy of prediction heavily depends on the quality of data, and the model may not perform well with imbalanced or incomplete datasets. Furthermore, some machine learning algorithms behave as "black boxes," lacking transparency in decision-making, which can be critical in medical applications.

REFERENCES

[1]. (Detrano et al., 1989) provided an international application of a new probability algorithm for coronary artery disease diagnosis. The article was published in the American Journal of Cardiology, volume 64, number 5, pages 304-310.

[2]. S. Gudadhe K. Wankhade, and S. Dongre, Decision Support System for Heart Disease based on SVM and ANN, in Proc. Int. Conf. on Computer and Communication Technology, Sep. 2010, 741–745.

[3]. The research paper entitled Comparison of Logistic Regression and Artificial Neural Network Models in Predicting Clinical Outcomes is written by authors T. Ayer, J. Chhatwal, O. Alagoz, K. Kahn, and B. Ladabaum. The research paper was published in Medical Decision Making, volume 33, issue 5.

[4]. Heart disease prediction with machine learning and data mining techniques by M. Patel, B. Teja, and K. Prajapati has been published in International Journal of Computer Applications. The date of publication is Feb. 2017 and it is available in vol. 160, no.

[5]. B Mohan, C Thirumalai and G Srivastava Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques, IEEE Access, vol. 7, pp. 81542-81554, 2019.

[6]. The rise of artificial intelligence and the uncertain future for physicians by Krittanawong A. European Journal of Internal Medicine, vol. 48, e13–e14, Nov. 2017.

[7]. UCI Machine Learning Repository: Heart Disease Data Set.[Online]Available: https://archive.ics.uci.edu/ml/datasets/heart+disease

[8]. I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," Artificial Intelligence in Medicine, vol. 23, no. 1, pp. 89–109, 2001.