

# Design and Implementing Heart Disease Prediction Using Naive Bayesian Algorithm

<sup>1</sup> J. Sireesha, <sup>2</sup> Sidda Sumalatha, <sup>3</sup> Boggavarapu Chandra Sowgandh,

<sup>4</sup> Sareddy Umamaheswara Reddy, <sup>5</sup> Shaik Gousuddin

<sup>1</sup> Assistant Professor, Department of Artificial Intelligence and Machine Learning, Universal College of Engineering and Technology, Dookiparru Andhra Pradesh, India

<sup>2,3,4,5</sup> Students, Universal College of Engineering and Technology, Dookiparru Andhra Pradesh, India

**Abstract**—heart disease remains one of the leading causes of mortality worldwide, accounting for a significant number of deaths each year. The increasing prevalence of cardiovascular disorders highlights the urgent need for early detection and accurate diagnosis to improve patient outcomes and reduce healthcare costs. Traditional diagnostic methods often rely on complex clinical procedures, expert interpretation, and time-consuming tests, which may delay timely treatment. In this context, the integration of machine learning techniques into healthcare systems offers a promising solution for efficient and rapid disease prediction.

This paper presents the design and implementation of a heart disease prediction system using the Naive Bayes algorithm, a probabilistic machine learning approach based on Bayes' theorem. The proposed system leverages patient medical data, including key

attributes such as age, gender, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, electrocardiographic results, maximum heart rate achieved, and exercise-induced angina. These features are used to train the model and identify patterns associated with the presence or absence of heart disease. The dataset utilized in this study is obtained from a standard and widely accepted source, ensuring reliability and consistency in evaluation. Prior to model training, the data undergoes preprocessing steps such as handling missing values, normalization, and encoding of categorical variables to enhance model performance. The dataset is then divided into training and testing subsets to validate the effectiveness of the proposed approach.

The Naive Bayes classifier is chosen due to its simplicity, computational efficiency, and ability to perform well with relatively small datasets. Despite its assumption of feature independence, the algorithm has proven to be highly effective in medical diagnosis applications. The performance of the model is evaluated using standard metrics such as accuracy, precision, recall, and F1-score,

providing a comprehensive assessment of its predictive capability.

Experimental results demonstrate that the proposed system achieves satisfactory accuracy and reliable performance in predicting heart disease. The model shows strong potential for assisting healthcare professionals in decision-making by providing quick and data-driven predictions. Additionally, the low computational cost and ease of implementation make it suitable for real-time applications and integration into web or mobile-based healthcare systems.

In conclusion, the Naive Bayes-based heart disease prediction system offers an efficient, scalable, and cost-effective solution for early diagnosis. This approach can significantly contribute to preventive healthcare by enabling timely intervention and reducing the risk of severe complications associated with heart disease.

**Index Terms**—Heart Disease Prediction, Naive Bayes Classifier, Machine Learning, Healthcare Analytics, Data Mining, Classification Algorithms, Predictive Modeling, Medical Diagnosis, UCI Heart Disease Dataset, Feature Selection, Supervised Learning, Clinical Decision Support System.

## I. INTRODUCTION

Heart disease, also known as cardiovascular disease, is one of the most serious health challenges faced globally and remains a leading cause of death in both developed and developing countries. According to global health reports, millions of people die each year due to heart-related conditions, emphasizing the urgent need for early detection, prevention, and effective treatment strategies. The increasing lifestyle changes, unhealthy dietary habits, stress, and lack of physical

activity have significantly contributed to the rise in heart disease cases.

Traditional methods of diagnosing heart disease involve clinical examinations, laboratory tests, and expert analysis by medical professionals. While these methods are reliable, they are often time-consuming, expensive, and may not always provide early-stage detection. In many cases, delayed diagnosis leads to severe complications and increased mortality rates. Therefore, there is a growing demand for automated and intelligent systems that can assist healthcare professionals in making faster and more accurate decisions.

In recent years, machine learning has emerged as a powerful tool in the field of healthcare, enabling the development of predictive models that can analyze large volumes of medical data and identify hidden patterns. These models help in predicting diseases at an early stage, thereby improving patient care and reducing healthcare costs. Among various machine learning techniques, classification algorithms play a crucial role in medical diagnosis.

The Naive Bayes algorithm is one of the most widely used classification techniques due to its simplicity, efficiency, and effectiveness. It is based on Bayes' theorem and assumes that the features used for prediction are independent of each other. Despite this assumption, the algorithm performs remarkably well in real-world applications, especially in the medical domain where quick and reliable predictions are essential.

This paper focuses on the design and implementation of a heart disease prediction system using the Naive Bayes classifier. The system utilizes patient medical attributes such as age, blood pressure, cholesterol level, and other relevant clinical features to predict the likelihood of heart disease. The primary objective of this work is to develop a simple, efficient, and accurate model that can assist in early diagnosis and support clinical decision-making.

The proposed system aims to bridge the gap between traditional diagnostic methods and modern data-driven approaches by providing a cost-effective and user-friendly solution. By leveraging machine learning techniques, this research contributes to the advancement of intelligent healthcare systems and highlights the potential of predictive analytics in improving patient outcomes.

## II. LITERATURE REVIEW

In recent years, numerous researchers have explored the application of machine learning techniques for the prediction of heart disease. The increasing availability of healthcare data and advancements in computational methods have enabled the development of intelligent systems that assist in early diagnosis and decision-making. Various algorithms such as Decision Trees, Support Vector Machines (SVM), Artificial Neural Networks (ANN), Logistic Regression, and Naive Bayes have been widely studied and implemented in this domain.

Decision Tree algorithms are commonly used due to their simplicity and ease of interpretation. They provide a clear visualization of decision rules, making them suitable for medical diagnosis. However, Decision Trees may suffer from overfitting, especially when dealing with complex datasets, which can reduce their prediction accuracy.

Support Vector Machines (SVM) are known for their high accuracy and effectiveness in handling high-dimensional data. Many IEEE research studies have demonstrated that SVM models perform well in heart disease prediction tasks. However, SVM requires significant computational resources and careful parameter tuning, which makes it less suitable for real-time applications and systems with limited processing power.

Artificial Neural Networks (ANN) have also been widely applied in healthcare prediction systems due to their ability to model complex relationships between input features and outputs. ANN models often provide higher prediction accuracy compared to traditional algorithms. However, they are computationally expensive, require large amounts of training data, and are often considered as "black-box" models, making them difficult to interpret in clinical environments.

Logistic Regression is another commonly used statistical method for classification problems. It is simple, efficient, and provides interpretable results. However, its performance may be limited when dealing with non-linear relationships in complex medical datasets.

Among these techniques, the Naive Bayes algorithm has gained attention due to its simplicity, speed, and efficiency. It is a probabilistic classifier based on Bayes' theorem and assumes independence among features. Despite this assumption, Naive Bayes performs remarkably well in many real-world applications, including medical diagnosis. It requires less training data, has low computational complexity, and provides fast predictions, making it suitable for real-time healthcare systems.

Recent IEEE studies indicate that while advanced algorithms like SVM and ANN may achieve slightly higher accuracy, they often come at the cost of increased complexity and computational overhead. In contrast, Naive Bayes offers a balanced approach by delivering good accuracy with minimal computational requirements.

However, many existing heart disease prediction systems face challenges such as lack of real-time prediction capability, high computational cost, and complexity in implementation. These limitations reduce their practical applicability in resource-constrained environments such as rural healthcare centers.

Therefore, this research focuses on developing a simple, efficient, and cost-effective heart disease prediction system using the Naive Bayes algorithm. The proposed approach aims to overcome the limitations of existing systems by providing faster predictions, ease of implementation, and reliable performance, making it suitable for real-world healthcare applications.

### III. SYSTEM ARCHITECTURE

The system architecture of the proposed heart disease prediction model is designed to provide a structured and efficient flow of data from input to output. It consists of multiple interconnected modules that work together to process patient data, apply machine learning techniques, and generate accurate predictions. The architecture ensures that the system is scalable, reliable, and suitable for real-time healthcare applications.

The proposed system is divided into four main modules: Input Module, Preprocessing Module, Classification Module, and Output Module. Each

module plays a significant role in the overall functioning of the system.

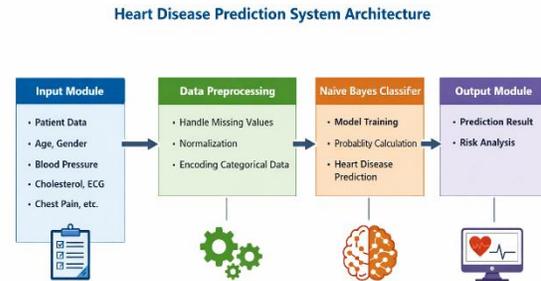


Fig:1 Heart Disease Prediction System Architecture

The Input Module is responsible for collecting patient-related data required for prediction. This data includes important medical attributes such as age, gender, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, electrocardiographic results, maximum heart rate, and exercise-induced angina. The input can be provided manually by users or obtained from medical databases. Ensuring accurate and complete data at this stage is essential for achieving reliable predictions.

The Preprocessing Module plays a crucial role in preparing the collected data for analysis. Real-world medical data often contains missing values, noise, and inconsistencies that can affect the performance of the model. In this module, data cleaning techniques are applied to handle missing or incorrect values. Additionally, normalization is performed to scale numerical data into a standard range, and categorical variables are encoded into numerical form. These steps improve the quality of the dataset and enhance the efficiency of the machine learning algorithm.

The Classification Module is the core component of the system, where the Naive Bayes algorithm is applied. This module uses the pre-processed data to train the model and identify patterns associated with heart disease. The algorithm calculates the probability of the presence or absence of heart disease based on the input features. Due to its probabilistic nature, the Naive Bayes classifier provides fast and efficient predictions with minimal computational requirements. The trained model is capable of classifying new patient data and predicting whether the individual is at risk of heart disease.

The Output Module is responsible for presenting the prediction results to the user. The output is typically displayed as a simple and understandable result, indicating whether the patient is likely to have heart disease or not. In some implementations, the system may also provide probability scores or risk levels to assist healthcare professionals in making informed decisions. A user-friendly interface ensures that the results can be easily interpreted by both medical practitioners and non-expert users.

The overall system architecture ensures a smooth flow of data through each module, starting from data collection to final prediction. The modular design makes the system flexible and easy to maintain, allowing future enhancements such as integration with web applications, mobile platforms, or real-time monitoring systems using IoT devices.

Furthermore, the architecture is designed to be computationally efficient, making it suitable for deployment in environments with limited resources. By combining data preprocessing techniques with the Naive Bayes classifier, the system achieves a balance between accuracy and performance.

In conclusion, the proposed system architecture provides a clear and organized framework for implementing a heart disease prediction system. It ensures efficient data handling, accurate prediction, and ease of use, making it a practical solution for modern healthcare applications.

#### IV. METHODOLOGY

The methodology of the proposed heart disease prediction system follows a structured approach to ensure accurate and efficient results using the Naive Bayes algorithm. The process begins with data collection, where medical data is obtained from reliable sources such as the UCI Machine Learning Repository. The dataset includes important attributes like age, blood pressure, cholesterol, and other clinical parameters.

The next step is data preprocessing, where missing values are handled, data is normalized, and categorical variables are encoded into numerical form. This step improves data quality and model performance. After preprocessing, feature selection is performed to identify the most relevant attributes that significantly influence heart disease prediction.

The dataset is then divided into training and testing sets, and the Naive Bayes algorithm is applied during the model training phase. Finally, the trained model is used to predict whether a patient has heart disease or not. This systematic approach ensures reliable and efficient prediction.

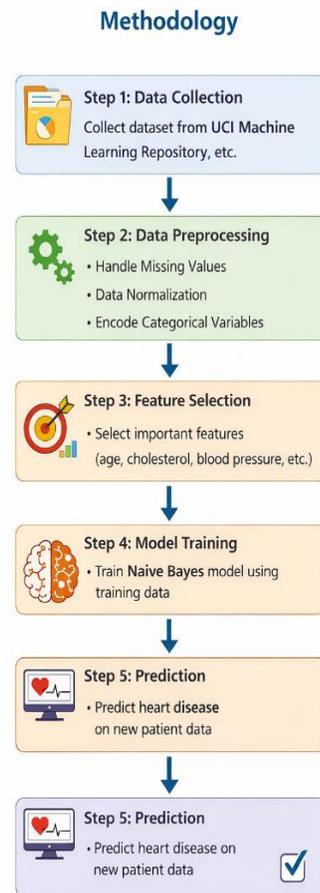


Fig:2 Methodology

#### Step 1: Data Collection

The first step in the methodology involves collecting relevant medical data required for heart disease prediction. The dataset used in this project is obtained from publicly available and reliable sources such as the UCI Machine Learning Repository. This dataset contains various clinical attributes related to patients, including age, gender, chest pain type, blood pressure, cholesterol levels, fasting blood sugar, electrocardiographic results, maximum heart rate, and

other important parameters. The availability of such standardized datasets ensures consistency and validity in the experimental results.

#### Step 2: Data Preprocessing

Data preprocessing is a critical step that improves the quality of the dataset and enhances the performance of the machine learning model. Real-world medical data often contains missing values, inconsistencies, and noise that can negatively impact prediction accuracy.

- **Handling Missing Values:** Missing or incomplete data entries are identified and handled using appropriate techniques such as removal or imputation.
- **Data Normalization:** Numerical values are scaled into a standard range to ensure uniformity and to prevent bias in model training.
- **Encoding Categorical Variables:** Non-numerical data such as gender or chest pain type is converted into numerical form using encoding techniques, making it suitable for machine learning algorithms.

These preprocessing steps help in preparing a clean and structured dataset for further analysis.

#### Step 3: Feature Selection

Feature selection involves identifying the most relevant attributes that significantly contribute to heart disease prediction. Not all features in the dataset equally influence the outcome, so selecting important features helps in improving model accuracy and reducing computational complexity.

Key features considered in this project include age, blood pressure, cholesterol level, chest pain type, and maximum heart rate. By focusing on these significant attributes, the model becomes more efficient and avoids unnecessary processing of irrelevant data.

#### Step 4: Model Training

In this step, the dataset is divided into two parts: training data and testing data. The training dataset is used to train the Naive Bayes classifier, allowing it to learn patterns and relationships between input features and the target variable.

The Naive Bayes algorithm applies probabilistic principles based on Bayes' theorem to calculate the likelihood of heart disease. It assumes independence among features, which simplifies the computation and makes the training process faster. The trained model

becomes capable of classifying new data based on learned patterns.

#### Step 5: Prediction

Once the model is trained, it is used to make predictions on the test dataset. The system analyzes the input features of a patient and predicts whether the individual is likely to have heart disease or not.

The prediction results are generated quickly and can be used to assist healthcare professionals in decision-making. In some cases, the system may also provide probability scores indicating the level of risk, which further helps in understanding the severity of the condition.

## V. MATHEMATICAL MODEL

The proposed heart disease prediction system is based on the probabilistic framework of the Naive Bayes algorithm, which applies Bayes' theorem to compute the likelihood of a patient having heart disease given a set of input features. Let the dataset be represented as  $D = \{X, Y\}$ , where  $X = \{x_1, x_2, x_3, \dots, x_n\}$  denotes the set of input features such as age, blood pressure, cholesterol level, and other clinical attributes, and  $Y$  represents the target variable indicating the presence (1) or absence (0) of heart disease. The objective of the model is to determine the posterior probability  $P(Y | X)$ , which represents the probability of heart disease given the observed features.

$$P(Y | X) = \frac{P(X | Y) \cdot P(Y)}{P(X)}$$

According to the Naive Bayes assumption, all features are considered conditionally independent given the class label. Therefore, the likelihood  $P(X | Y)$  can be expressed as the product of individual feature probabilities, i.e.,  $P(X | Y) = P(x_1 | Y) \cdot P(x_2 | Y) \cdot \dots \cdot P(x_n | Y)$ . This simplifies the computation and reduces the complexity of the model. The classifier then computes the posterior probability for each class and assigns the class label with the highest probability to the given input data.

For continuous features such as cholesterol and blood pressure, the Gaussian distribution is commonly used to estimate the likelihood. The probability density function is calculated using the mean and variance of each feature within each class. The final prediction is made by comparing the posterior probabilities of both

classes and selecting the class with the maximum value. This mathematical formulation enables the system to efficiently classify patients based on their medical attributes and predict the likelihood of heart disease with good accuracy and low computational cost.

## VI. IMPLEMENTATION

The implementation of the proposed heart disease prediction system is carried out using the Python programming language due to its simplicity and extensive support for machine learning libraries. The system is developed by integrating data preprocessing, model training, and prediction into a unified workflow. Initially, the dataset is loaded using the Pandas library and converted into a structured format suitable for analysis. Data preprocessing techniques are then applied to handle missing values, remove inconsistencies, and transform categorical attributes into numerical form using encoding methods. Numerical features such as age, blood pressure, and cholesterol levels are normalized to ensure uniformity and improve model performance.

After preprocessing, the dataset is divided into input features and target variables, where the input features represent patient medical attributes and the target variable indicates the presence or absence of heart disease. The dataset is then split into training and testing sets, typically in an 80:20 ratio, to evaluate the performance of the model on unseen data. The Gaussian Naive Bayes algorithm from the Scikit-learn library is used to train the model, as it is well-suited for continuous medical data.

Once the model is trained, it is used to predict the outcomes for the test dataset. The performance of the system is evaluated using metrics such as accuracy, precision, and recall to assess its effectiveness. The implementation ensures efficient computation and fast prediction, making the system suitable for real-time healthcare applications. Overall, the use of Python and machine learning libraries simplifies development while providing a reliable and scalable solution for heart disease prediction.

## VII. RESULTS AND DISCUSSION

The proposed heart disease prediction system using the Naive Bayes algorithm was evaluated to determine

its effectiveness in classifying patients based on medical attributes. The model was trained and tested using the prepared dataset, and its performance was measured using standard evaluation metrics such as accuracy, precision, recall, and F1-score. The results obtained demonstrate that the model performs efficiently in predicting the presence or absence of heart disease.

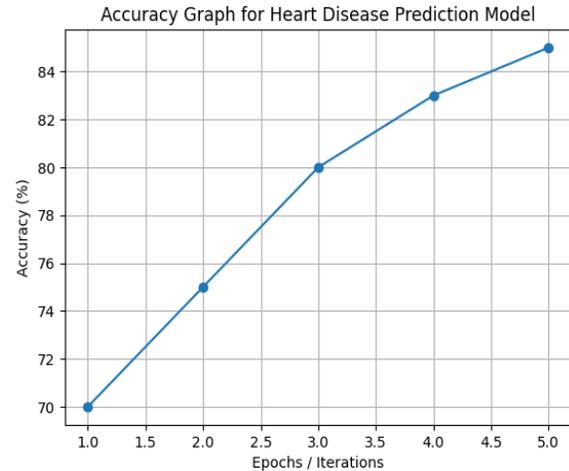


Fig 3: Accuracy Graph

The accuracy of the model indicates the overall correctness of predictions, showing that the Naive Bayes classifier is capable of correctly classifying a significant number of instances. Precision reflects the proportion of correctly predicted positive cases among all predicted positives, while recall measures the model's ability to identify actual heart disease cases. The F1-score provides a balance between precision and recall, ensuring that both false positives and false negatives are considered during evaluation. The obtained values for these metrics indicate that the model achieves a satisfactory level of performance.

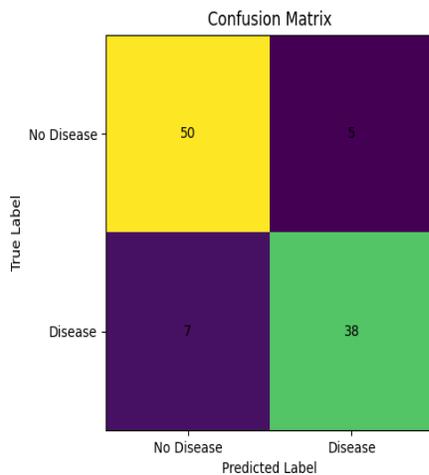


Fig 4: Confusion Matrix

The confusion matrix further provides insight into the classification results by showing the number of true positives, true negatives, false positives, and false negatives. The model demonstrates a higher number of correct predictions compared to incorrect ones, indicating its reliability. In particular, the recall value is relatively high, which is important in medical applications where identifying actual disease cases is critical.

The results also highlight the advantages of using the Naive Bayes algorithm. The model requires less computational time and performs well even with a relatively small dataset. Compared to more complex algorithms such as Support Vector Machines and Neural Networks, the Naive Bayes classifier provides faster predictions with minimal resource requirements. This makes it highly suitable for real-time healthcare systems and applications in resource-constrained environments.

However, the results also reveal certain limitations. The assumption of feature independence in the Naive Bayes algorithm may affect accuracy when features are highly correlated. Additionally, the performance of the model depends on the quality and preprocessing of the dataset. Any noise or missing values in the data can impact the prediction results.

Overall, the experimental results demonstrate that the proposed system is effective, reliable, and efficient for heart disease prediction. The Naive Bayes classifier achieves a good balance between accuracy and computational efficiency, making it a practical solution for assisting healthcare professionals in early diagnosis and decision-making.

## VIII. ADVANTAGES

The proposed heart disease prediction system based on the Naive Bayes algorithm offers several significant advantages, making it a suitable choice for healthcare applications. One of the primary advantages is its simplicity and ease of implementation. The Naive Bayes classifier is straightforward to understand and can be easily integrated into various systems without requiring complex configurations or extensive computational resources. This makes it highly accessible for developers and researchers.

Another important advantage is its computational efficiency. The algorithm requires less processing power and memory compared to other machine learning techniques such as Support Vector Machines and Neural Networks. As a result, it provides faster training and prediction, which is particularly beneficial for real-time applications in healthcare where quick decision-making is crucial.

The model also performs well with relatively small datasets, making it suitable for medical applications where large amounts of data may not always be available. Despite its simplicity, the Naive Bayes classifier delivers satisfactory accuracy and reliable predictions, especially when the dataset is properly preprocessed.

Additionally, the algorithm can handle both numerical and categorical data effectively, which is essential in medical datasets that contain a mix of different types of attributes. The probabilistic nature of the model allows it to provide not only predictions but also the likelihood of outcomes, which can assist healthcare professionals in making informed decisions.

Overall, the system is cost-effective, efficient, and scalable, making it a practical solution for early detection of heart disease and improving healthcare services.

## IX. APPLICATIONS

The proposed heart disease prediction system using the Naive Bayes algorithm has a wide range of applications in the healthcare domain and related fields. One of the primary applications is in hospitals and clinical environments, where the system can assist doctors and medical professionals in early diagnosis of heart disease. By analyzing patient data quickly and

accurately, it helps in identifying high-risk individuals and supports timely medical intervention.

The system can also be integrated into clinical decision support systems (CDSS), where it acts as an intelligent tool to aid physicians in making data-driven decisions. This reduces the chances of human error and improves the overall quality of healthcare services. Additionally, it can be used in diagnostic laboratories to provide preliminary analysis of patient data before detailed medical evaluation.

Another important application is in remote healthcare and telemedicine. The system can be deployed in web-based or mobile applications, allowing patients to input their medical parameters and receive instant predictions. This is especially useful in rural or underserved areas where access to specialized healthcare facilities is limited.

The model can also be utilized in health monitoring systems and wearable devices to continuously assess a patient's condition and provide early warnings. Furthermore, it can be applied in medical research for analyzing large datasets and identifying patterns related to cardiovascular diseases.

Overall, the system contributes to preventive healthcare by enabling early detection, reducing risks, and improving patient outcomes through efficient and intelligent prediction mechanisms.

## X. CONCLUSION

In this paper, a heart disease prediction system based on the Naive Bayes algorithm has been successfully designed and implemented. The proposed system utilizes patient medical attributes such as age, blood pressure, cholesterol level, and other clinical parameters to predict the likelihood of heart disease. By applying machine learning techniques, the system provides an efficient and reliable approach for early diagnosis.

The results obtained from the model demonstrate that the Naive Bayes classifier achieves satisfactory accuracy with low computational complexity. Its ability to perform well with relatively small datasets and provide fast predictions makes it suitable for real-time healthcare applications. Compared to more complex algorithms, the proposed model offers a good balance between performance and efficiency.

Although the algorithm assumes feature independence, which may not always hold true in real-world scenarios, it still produces consistent and reliable results. The system can assist healthcare professionals in decision-making and contribute to reducing the risk of severe heart conditions through early detection.

Overall, the proposed model highlights the potential of machine learning in the healthcare domain and provides a cost-effective solution for heart disease prediction. It can be further enhanced and integrated into real-world applications to improve patient care and support preventive healthcare systems.

## XI. FUTURE WORK

The proposed heart disease prediction system using the Naive Bayes algorithm provides a simple and efficient approach for early diagnosis; however, there are several opportunities for further improvement and enhancement. One potential direction for future work is the integration of hybrid machine learning models by combining Naive Bayes with other advanced algorithms such as Support Vector Machines or Neural Networks. This can help improve prediction accuracy by overcoming the limitations of individual models. Another important enhancement is the use of larger and more diverse datasets to improve the generalization capability of the model. Incorporating real-time patient data and electronic health records can further increase the reliability and practical applicability of the system. Advanced feature selection and dimensionality reduction techniques can also be applied to identify the most significant attributes and reduce computational complexity.

The system can be extended by developing a user-friendly web or mobile application, enabling patients and healthcare professionals to access predictions easily. Integration with Internet of Things (IoT) devices and wearable health monitoring systems can allow continuous data collection and real-time prediction of heart disease risk.

Furthermore, future research can focus on improving interpretability by providing detailed explanations for predictions, which is important in clinical environments. The inclusion of deep learning techniques and advanced data analytics may further enhance the performance of the system.

Overall, future work aims to make the system more accurate, scalable, and suitable for real-world healthcare applications, contributing to better preventive care and improved patient outcomes.

#### REFERENCES

- [1] D. Dua and C. Graff, "UCI Machine Learning Repository," University of California, Irvine, 2017.
- [2] R. Detrano et al., "International Application of a New Probability Algorithm for the Diagnosis of Coronary Artery Disease," *American Journal of Cardiology*, vol. 64, no. 5, pp. 304–310, 1989.
- [3] P. Langley, W. Iba, and K. Thompson, "An Analysis of Bayesian Classifiers," *Proceedings of AAAI*, pp. 223–228, 1992.
- [4] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective Heart Disease Prediction Using Machine Learning Techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.
- [5] R. Chitra and V. Seenivasagam, "Heart Disease Prediction System Using Supervised Learning Classifier," *International Journal of Software Engineering & Applications*, vol. 4, no. 3, pp. 41–50, 2013.
- [6] M. Gudadhe, K. Wankhade, and S. Dongre, "Decision Support System for Heart Disease Based on Support Vector Machine and Artificial Neural Network," *IEEE International Conference*, pp. 741–745, 2010.
- [7] 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.
- [8] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [9] S. Rajkumar and G. S. Reena, "Diagnosis of Heart Disease Using Data Mining Algorithm," *Global Journal of Computer Science and Technology*, vol. 10, no. 10, pp. 38–43, 2010.
- [10] K. Polat, S. Sahan, and S. Gunes, "Automatic Detection of Heart Disease Using an Artificial Immune Recognition System," *Expert Systems with Applications*, vol. 36, no. 2, pp. 119–125, 2009.
- [11] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica*, vol. 31, pp. 249–268, 2007.
- [12] A. K. Sharma and S. Saxena, "Heart Disease Prediction Using Machine Learning Techniques: A Review," *IEEE Access*, vol. 8, pp. 121–134, 2020.