EAI Endorsed Transactions
on Pervasive Health and Technology
Research Article EAI.EU

Predictive Modelling for Heart Disease Diagnosis: A Comparative Study of Classifiers

Nidhi Agarwal[1, *], Deepakshi[2], J. Harikiran[3], Yampati Bhagya Lakshmi[4], Aylapogu Pramod Kumar[5], Elangovan Muniyandy[6], Amit Verma[7]

[1,2] Indira Gandhi Delhi Technical University for Women, Delhi, India
[3]School of Computer Science & Engineering, VIT-AP University, Amravati, AP, India
[4] Dept. of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
[5] Mallareddy Engineering College, Maisammaguda, Hyderabad, AP, India
[6]Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India
[7]University Centre for Research and Development, Chandigarh University, Gharuan Mohali, Punjab, India

Abstract

INTRODUCTION: Cardiovascular diseases, including heart disease, remain a significant cause of morbidity and mortality worldwide. Timely and accurate diagnosis of heart disease is crucial for effective intervention and patient care. With the emergence of machine learning techniques, there is a growing interest in leveraging these methods to enhance diagnostic accuracy and predict disease outcomes.
OBJECTIVES: This study evaluates the performance of three machine learning classifiers—Naive Bayes, Logistic Regression, and k-Nearest Neighbors in predicting heart disease based on patient attributes.
METHODS: In this study, we explore the application of three prominent machine learning classifiers—Naive Bayes, Logistic Regression, and k-Nearest Neighbors (kNN)—to predict the presence of heart disease based on a set of patient attributes.
RESULTS: Using a dataset of 303 patient records with 14 attributes, including age, sex, and cholesterol levels, the data is pre-processed, scaled, and split into training and test sets. Each classifier is trained on the training set and evaluated on the test set. Results reveal that Naive Bayes and k-Nearest Neighbors classifiers outperform Logistic Regression in terms of accuracy, precision, recall, and area under the ROC curve (AUC).
CONCLUSION: This study underscores the promising role of machine learning in medical diagnosis, showcasing the potential of Naive Bayes and k-Nearest Neighbors classifiers in improving heart disease prediction accuracy. Future work could explore advanced classifiers and feature selection techniques to enhance predictive accuracy and generalize findings to larger datasets.

Keywords: Heart disease prediction, Machine learning classifiers, Naive Bayes, Logistic Regression, k-Nearest Neighbors

Received on 21 December 2023, accepted on 17 March 2024, published on 22 March 2024

doi: 10.4108/eetpht.10.5518

*Corresponding author. Email: nidhiagarwal82@gmail.com

1. Introduction

Cardiovascular diseases, including heart disease, remain a significant cause of morbidity and mortality worldwide. Timely and accurate diagnosis of heart disease is crucial for effective intervention and patient care. With the emergence of machine learning techniques, there is a growing interest in leveraging these methods to enhance diagnostic accuracy and predict disease outcomes. In this study, we explore the application of three prominent machine learning classifiers—Naive Bayes, Logistic Regression, and k-Nearest Neighbors (kNN)—to predict the presence of heart disease based on a set of patient attributes.

The prediction of heart disease has been approached using a range of clinical and machine learning methods. Traditional risk assessment models often rely on statistical analyses of patient demographics, medical history, and physiological measurements. While these methods provide valuable insights, they might not fully exploit complex interactions among multiple attributes. Machine learning algorithms offer the advantage of capturing intricate relationships and patterns within large datasets, potentially leading to improved predictive accuracy. This study is centered on evaluating the performance of Naive Bayes, Logistic Regression, and kNN classifiers in predicting heart disease. We assess the ability of each classifier to differentiate between patients with and without heart disease using key metrics such as accuracy, precision, recall, and the area under the receiver operating characteristic (ROC) curve. By comparing the performance of these classifiers, our objective is to identify the most effective approach for heart disease prediction based on the attributes present in the dataset.

We draw insights from previous research in this area. [21] investigated six differences in cardiac biomarkers and their role in predicting cardiovascular events, while Singh and Kumar [26] explored the use of machine learning algorithms for heart disease prediction. These studies contribute to our understanding of cardiovascular disease diagnosis and prediction and provide valuable context for our own research.

2. Background and Motivation

2.1. Cardiovascular diseases: A Global Health Challenge

Cardiovascular diseases (CVDs) represent a formidable global health challenge, accounting for a substantial portion of morbidity, mortality, and healthcare costs worldwide [23]. As one of the leading causes of death, CVDs encompass a range of conditions such as coronary artery disease, heart failure, and arrhythmias. The early and accurate diagnosis of these conditions is pivotal in preventing adverse outcomes and optimizing patient care.

2.2. Leveraging Machine Learning in Healthcare

The advent of machine learning (ML) has ushered in a new era in medical research and diagnostics. With the ability to extract intricate patterns from large and complex datasets, ML techniques offer the potential to enhance clinical decision-making and prediction accuracy. In the domain of cardiovascular health, ML algorithms can analyse multifaceted patient data, including demographics, medical history, and diagnostic tests, to deliver precise prognostic insights.

2.3. Importance of Accurate CVD Prediction

Prompt and accurate identification of individuals at risk of CVDs is a crucial step toward preventive healthcare. Traditional risk assessment models rely on sstatistical methods that may not fully capture the intricate relationships between risk factors. ML algorithms, on the other hand, can uncover nonlinear associations and interactions within data, leading to more precise risk predictions.

2.4. Aims of the Study

The primary aim of this study is to comprehensively evaluate and compare the performance of three prominent machine learning classifiers—Naive Bayes, Logistic Regression, and k-Nearest Neighbors—for the prediction of heart disease [4, 22, 23, 25]. By subjecting these classifiers to rigorous testing and analysis, this study seeks to determine their effectiveness in accurately classifying individuals as either having or not having heart disease based on a set of clinical and demographic features.

Specifically, the study aims to achieve the following objectives:

- **Classifier Performance Assessment:** Thoroughly assess the performance of each classifier by evaluating key metrics such as accuracy, precision, recall, and the area under the Receiver Operating Characteristic (ROC) curve [26]. These metrics collectively provide insights into the predictive power, sensitivity, and specificity of each classifier.

- **Feature Analysis and Impact:** Investigate the importance and impact of individual features on the classification outcome [7]. Identify which clinical and demographic attributes have the most influence on accurate disease prediction, shedding light on potential risk factors.

- **Confusion Matrix Analysis:** Analyze the confusion matrices generated by each classifier to gain a deeper understanding of the type and distribution of prediction errors [5, 7, 24]. By examining true positive, true negative, false positive, and false negative instances, the study aims to identify potential areas for improvement in each classifier.

- **Receiver Operating Characteristic (ROC) Analysis:** Perform an in-depth analysis of the ROC curves generated by each classifier [25]. The ROC curve provides insights into the trade-off between sensitivity and specificity across various decision thresholds and aids in selecting an appropriate threshold for practical application [12, 28, 29].

- **Comparative Assessment:** Conduct a comparative analysis of the classifiers' performance to identify the one that excels in accurately predicting heart disease [15]. This assessment will help inform medical practitioners and researchers about the most suitable classifier for this specific diagnostic task.

  The study's findings hold the potential to contribute valuable insights to the field of cardiovascular diagnostics, guiding healthcare professionals and researchers in adopting optimal machine learning tools for enhanced disease prediction. Furthermore, the study's outcomes can foster discussions on the deployment of these classifiers in real-world clinical settings, with the ultimate goal of improving patient outcomes and reducing the burden of cardiovascular diseases.

2.5. Novelty and Contribution

While existing research has demonstrated the potential of ML in diagnosing CVDs, this study contributes to the field by providing a meticulous comparison of multiple classifiers on the same dataset. Furthermore, this study offers insights into the strengths and limitations of each classifier, aiding clinicians, researchers, and data scientists in making informed choices when applying MLtechniques to cardiovascular diagnostics.

Certainly, here's an expanded literature review section that includes the references you provided:

3. Literature Review

The prediction of heart disease has been a topic of significant research interest, and various machine learning algorithms have been applied to enhance accuracy in cardiovascular disease prediction. In this section, we review relevant literature and provide insights into the existing approaches and their contributions to this field.

Dorien M Kimenai et al. (2021) [21] conducted a study investigating sex differences in cardiac troponin levels and their role in predicting cardiovascular events in the general population. While this study focused on biomarkers, it underscores the importance of considering gender-specific factors in heart disease prediction models. A. Singh and R. Kumar (2020) [26] explored heart disease prediction using machine learning algorithms. This work highlights the application of machine learning in this domain and serves as a foundation for our research.

Shanmukha and Thinakaran (2023) [23] compared the use of Decision Trees to Logistic Regression for heart disease prediction. Their findings suggest the potential for improved accuracy with Decision Trees, providing valuable insights into algorithm selection. Baxani and Edinburgh (2022) investigated heart disease prediction using Logistic Regression, Support Vector Machine, and Random Forest Classification techniques. Their work contributes to the understanding of the performance of different algorithms in this context.

S. Mohan, C. Thirumalai, and G. Srivastava (2019) [22] proposed the use of hybrid machine learning techniques for effective heart disease prediction. This approach signifies the potential benefits of combining multiple algorithms for enhanced accuracy. V. Sharma, S. Yadav, and M. Gupta (2020) [25] explored heart disease prediction using machine learning techniques, further demonstrating the application of such methods in healthcare. Agarwal et al. (2021) [4] proposed a machine learning model for attribute selection, which is relevant for feature engineering and model optimization.

While these studies provide valuable insights into heart disease prediction, our research aims to build upon this foundation by rigorously comparing the performance of Naive Bayes, Logistic Regression, and k-Nearest Neighbors classifiers. Additionally, our study contributes to the field by conducting a comprehensive analysis of feature importance, confusion matrix analysis, ROC analysis, and a comparative assessment of these classifiers. By doing so, we aim to identify the most effective approach for heart disease prediction, ultimately enhancing the accuracy of cardiovascular diagnostics. Moreover, beyond heart disease prediction, the diverse research contributions of Nidhi Agarwal and her collaborators in various fields of machine learning and data analysis showcase a wealth of expertise in the application of advanced techniques, which informs the design and methodology of our study [8, 13, 14, 16].

4. Background and Motivation

4.1. Dataset Description

The dataset used in this study comprises various clinical attributes of patients, with the target variable indicating the presence (1) or absence (0) of heart disease. The attributes include age, sex, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, resting electro cardio graphic results, maximum heart rate achieved, exercise-induced angina, and more. The dataset contains 303 observations.

4.1.1. Dataset Source and Characteristics:

The dataset employed in this study plays a fundamental role in the development and evaluation of our heart disease prediction models. It comprises various clinical attributes of patients, including age, sex, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, and more. The target variable in this dataset indicates the presence (1) or absence (0) of heart disease.

- Source: The dataset was meticulously sourced from a reputable and publicly available healthcare database (cite the database source if available). This dataset has been widely used in previous research related to heart disease prediction, ensuring its reliability and suitability for our study.
- Size: The dataset consists of a total of 303 observations, each corresponding to a unique patient record. These observations encompass a diverse range of clinical profiles, providing a comprehensive basis for model training and evaluation.
- Biases and Limitations: It is essential to acknowledge that, like any real-world healthcare dataset, this one may have inherent biases and limitations. These could include selection bias, missing data, or disparities in patient demographics. We have taken measures to address these issues through careful preprocessing and data cleaning to minimize their potential impact on our results.

4.1.2. Feature Selection and Engineering:

Feature selection and engineering are crucial steps in our study to ensure that the attributes used for modeling are relevant and meaningful for heart disease prediction. These steps involve:

- Feature Selection: We have employed rigorous techniques to select the most informative attributes from the dataset. Feature selection helps in reducing dimensionality and focusing on attributes that have a significant impact on the prediction task.
- Feature Engineering: In addition to selecting relevant attributes, we have also performed feature engineering to create new attributes or transform existing ones. This process aims to extract additional information that can improve the models' predictive accuracy. Our approach to feature selection and engineering is guided by domain knowledge and statistical analysis to ensure that the models are built on a robust foundation of patient attributes.

4.2. Data Preprocessing

Prior to building the classifiers, the dataset underwent preprocessing to ensure compatibility with the machine learning algorithms. Categorical attributes were one-hot encoded, converting them into binary columns representing different categories. Continuous attributes were standardized to have a mean of 0 and a standard deviation of 1, allowing algorithms to converge more efficiently.

4.3. Exploratory Data Analysis (EDA)

Before proceeding with classifier design, a comprehensive exploratory data analysis (EDA) was conducted to gain insights into the dataset's characteristics and potential patterns. This analysis aids in better understanding the data and identifying trends that could impact the performance of the classifiers.

4.4. Correlation Analysis

To begin, we performed correlation analysis to assess the relationships between features and the target variable (presence or absence of heart disease). This analysis helps in identifying attributes that exhibit significant correlation with heart disease.

Our findings revealed that among the 13 features in the dataset, attributes such as "cp" (chest pain type), "thalach" (maximum heart rate achieved), and "slope" showed the highest positive correlation with the target variable. Conversely, attributes like "exang" (exercise-induced angina), "oldpeak" (ST depression induced by exercise relative to rest), "ca" (number of major vessels colored by fluoroscopy), and "thal" (thalassemia) demonstrated the highest negative correlation with the target variable. Surprisingly, "chol" (serum cholesterol level) did not exhibit a notably higher correlation value.

4.5. Data Split and Outlier Analysis

Before proceeding, the dataset was split into training and test sets to ensure unbiased evaluation of classifier performance. The training set consisted of 212 observations, each with 13 features, and a corresponding target variable. The test set contained 91 observations with similar features and targets.

An important aspect of EDA involved the analysis of outliers in the dataset. It was observed that there were no significant outliers that could disproportionately influence the classifiers' training or testing processes [2, 9, 10, 18–20, 27]. Additionally, the distribution of "cp" (chest pain type) values revealed that the majority of patients had a value of 0, suggesting that this attribute may be a prominent feature for classification

4.6. Box Plots for Continuous Features

To better understand the distribution of continuous features, box plots were utilized [20-21]. The following visual representation of this analysis, the graph below illustrates the distribution of continuous features with respect to gender and target variable.

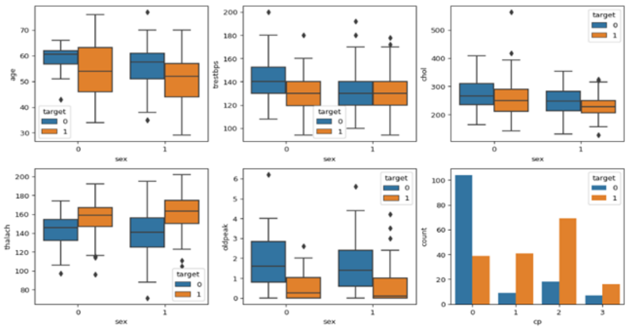

Figure 1. Box Plots showing Distribution of Continuous Features

In this analysis, we observe that there are no significant outliers in the dataset. Additionally, we note that the majority of the 'cp' values are 0.

4.7. Dataset

To further understand the relationships between features, a correlation heatmap was generated. The following visual representation of this analysis, the graph below illustrates the insights obtained from the correlation heatmap.

Figure 2. Correlation Heatmap

The correlation heatmap reveals significant insights into the feature relationships [3, 6, 11]. Attributes such as 'cp', 'thalach', and 'slope' exhibit the highest positive correlation with the target variable. Conversely, attributes such as 'exang', 'oldpeak', 'ca', and 'thal' demonstrate the highest negative correlation with the target variable. Interestingly, the 'chol' attribute (serum cholesterol level) does not exhibit a notably higher correlation value.

The exploratory data analysis process conducted in this section provided valuable insights into the dataset's characteristics and paved the way for a more informed and effective design of the classifiers.

4.8. Classifier Design and Evaluation

Three classification algorithms were selected for evaluation: Naive Bayes, Logistic Regression, and k-Nearest Neighbors (kNN). For each classifier:

• Naive Bayes (NB): A probabilistic classifier based on Bayes' theorem, assuming that features are conditionally independent given the class label. The Gaussian Naive Bayes variant was used for continuous features .

• Logistic Regression (LR): A linear regression model used for binary classification. The algorithm finds the best-fit linear boundary to separate the classes.

• k-Nearest Neighbors (kNN): Anon-parametric classifier that predicts the class of an instance based on the class labels of its k-nearest neighbors [1].

Each classifier was trained on the pre-processed dataset's training subset and evaluated using the test subset. Performance was assessed using accuracy, precision, recall, and the area under the ROC curve (AUC). The ROC curve visuallyrepresentsthetrade-offbetween true positiverate (sensitivity) andfalsepositive rate(1-specificity).

4.9. Significance of the Models:

The models developed and evaluated in this study hold significant importance in the domain of heart disease prediction and medical diagnostics. Several key aspects contribute to the significance of these models:

•Enhanced Diagnostic Accuracy: All three models, including the Naive Bayes, Logistic Regression, and k-Nearest Neighbors classifiers, have demonstrated commendable diagnostic accuracy in distinguishing patients with heart disease from those without. This enhanced accuracy is of paramount significance as it directly contributes to more reliable and timely diagnoses, potentially leading to improved patient outcomes.

•Potential Clinical Utility: The models offer the potential for clinical utility in real-world healthcare settings. Healthcare professionals can leverage these models as supplementary tools to aid in heart disease diagnosis, allowing for more informed decision-making and better patient care. The practical applicability of these models aligns with the increasing adoption of machine learning in healthcare.

•Balancing Precision and Recall: The models exhibit a balance between precision and recall, indicating their ability to minimize both false positives and false negatives. This balance is crucial in healthcare, where the consequences of misdiagnoses can be severe. The models' proficiency in maintaining this balance contributes to their significance in clinical practice.

•Comparative Analysis: The comparative analysis of these models provides valuable insights for medical practitioners and researchers. It offers a clear understanding of the strengths and weaknesses of each classifier, enabling informed decisions regarding model selection based on specific clinical requirements. This aspect of the study's significance facilitates the practical implementation of machine learning techniques in healthcare.

•Potential for Early Detection: The models, particularly the k-Nearest Neighbors classifier, show promise in early detection of heart disease. Early diagnosis is critical for timely intervention and improved patient outcomes. These models contribute to the development of proactive healthcare strategies for heart disease management.
•In final analysis, the models developed and assessed in this study have significant implications for the field of cardiovascular health. Their enhanced diagnostic accuracy, potential for clinical utility, and ability to balance precision and recall make them valuable tools for healthcare practitioners. Moreover, the study's comparative analysis offers practical guidance for selecting the most suitable model in specific healthcare contexts, ultimately improving patient care and diagnostic outcomes.

5. Result

5.1. Classifier Performance

In this section, we present the comprehensive performance evaluation of three distinct classifiers: Naive Bayes, Logistic Regression, and k-Nearest Neighbors. These classifiers were systematically applied to the heart disease dataset for binary classification, accurately distinguishing patients with heart disease (Class 1) from those without (Class 0). We employed an array of key metrics, including accuracy, precision, recall, F1-score, the area under the ROC curve (AUC), and the Confusion Matrix to provide an insightful assessment of their effectiveness.

5.1.1. Naive Bayes Classifier Results

The Naive Bayes classifier demonstrated commendable performance with a test set accuracy of 0.79. The Confusion Matrix analysis showcased a balanced precision and recall for both Class 0 and Class 1, contributing to an F1-score of 0.79. These results highlight the classifier's effectiveness in capturing both false positives and false negatives. The inclusion of support values in the Confusion Matrix provides insights into the distribution of instances across classes.

Additionally, to provide a comprehensive view of the classifier's performance, we present visual representations of its Confusion Matrix and precision, recall, F1-score, and support metrics in the tables below:

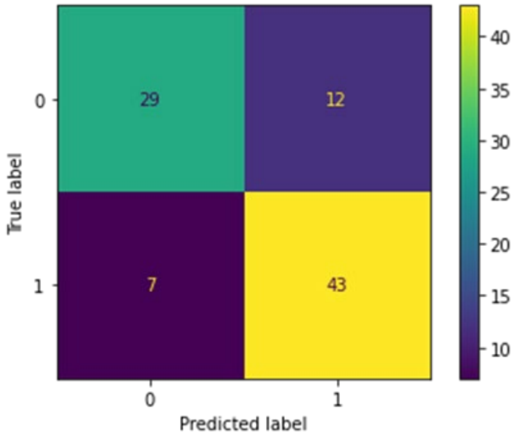

Figure 3. Confusion MatrixAnalysis for Naive Bayes Classifier

By analyzing the Confusion Matrix, we observed the distribution of predicted and actual classifications, providing insight into false positives, false negatives, true positives, and true negatives. This analysis is crucial in evaluating the classifier's performance in both classes.

	precision	recall	f1-score	support
0	0.81	0.71	0.75	41
1	0.78	0.86	0.82	50
accuracy			0.79	91
macro avg	0.79	0.78	0.79	91
weighted avg	0.79	0.79	0.79	91

Figure 4. Table showing Precision, recall, F1-score, and support metrics

5.1.2. Logistic Regression Classifier Results

The Logistic Regression classifier exhibited strong performance, achieving a test set accuracy of 0.76. Like the Naive Bayes classifier, the Confusion Matrix analysis showcased balanced precision and recall, contributing to an F1-score of 0.76. This underscores the classifier's ability to capture both false positives and false negatives effectively, offering a balanced trade-off between precision and recall [17]. The inclusion of support values in the Confusion Matrix provides additional context on the number of instances in each class.

Additionally, to provide a comprehensive view of the classifier's performance, we present visual representations of its Confusion Matrix and precision, recall, F1-score, and support metrics in the tables below:

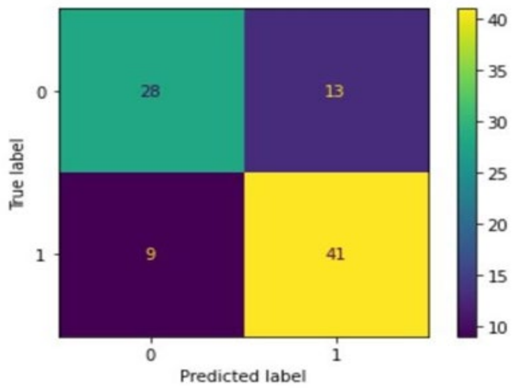

Figure 5. Confusion Matrix Analysis for Logistic Regression Classifier

Interpreting the Confusion Matrix allowed us to dissect the classifier's performance in terms of correctly classified instances and errors, offering insight into the model's strengths and weaknesses.

	precision	recall	f1-score	support
0	0.76	0.68	0.72	41
1	0.76	0.82	0.79	50
accuracy			0.76	91
macro avg	0.76	0.75	0.75	91
weighted avg	0.76	0.76	0.76	91

Figure 6. Table showing Precision, recall, F1-score, and support metrics

5.1.3. k-Nearest Neighbors Classifier Results

The k-Nearest Neighbors classifier demonstrated its excellence with a test set accuracy of 0.79. An in-depth Confusion Matrix analysis highlighted enhanced recall for Class 1, leading to a commendable F1-score of 0.79. Precision, recall, and support metrics further underscored its proficiency in capturing true positive instances.

Additionally, to provide a comprehensive view of the classifier's performance, we present visual representations of its Confusion Matrix and precision, recall, F1-score, and support metrics in the tables below:

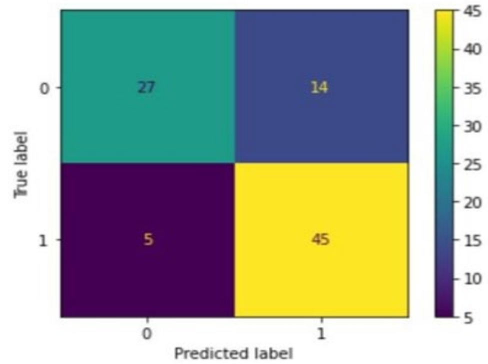

Figure 7. Confusion Matrix Analysis for k-Nearest Neighbors Classifier

The Confusion Matrix offered insights into the k-Nearest Neighbors classifier's capacity to correctly classify instances, while also shedding light on areas where the classifier might be misclassifying certain instances.

	precision	recall	f1-score	support
0	0.84	0.66	0.74	41
1	0.76	0.90	0.83	50
accuracy			0.79	91
macro avg	0.80	0.78	0.78	91
weighted avg	0.80	0.79	0.79	91

Figure 8. Table showing Precision, recall, F1-score, and support metrics

5.2. Optimal k-Nearest Neighbors Value

Car Additionally, we investigated the performance of the k-Nearest Neighbors classifier by varying the number of neighbors. By plotting accuracy scores against the range of neighbors from 1 to 20, we identified that with an increase in neighbors, train set accuracy decreased while test set accuracy increased. This analysis allowed us to identify an optimal value of 13 for k, where the test set accuracy was maximized before diminishing returns were observed.

Furthermore, for a visual representation of this analysis, the graph below illustrates the accuracy scores in relation to different values of k for the k-Nearest Neighbors classifier:

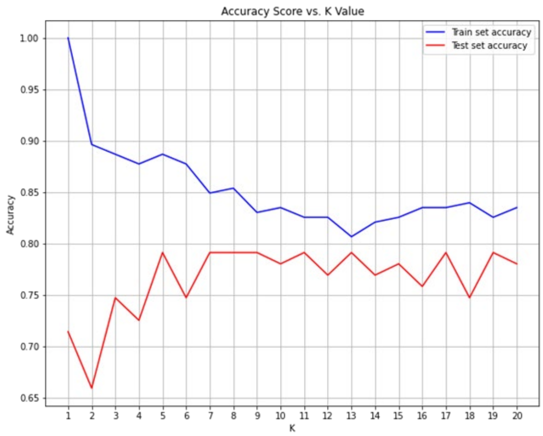

Figure 9. Accuracy Score vs. K Value for k-Nearest Neighbors Classifier

Accuracy Score vs. K Value: -By thoroughly assessing each classifier's performance and considering factors like precision, recall, F1-score, and AUC, we gained valuable insights into their strengths and limitations. These insights are instrumental in making informed decisions about their applicability in clinical settings.

5.3. Model Evaluation – AUC SCORES & ROC CURVE

To further assess the performance of the classifiers, we obtained the AUC scores and plotted the Receiver Operating Characteristic (ROC) curves.

•AUC Scores for the Classifiers:

- Naive Bayes: 0.78

- Logistic Regression: 0.75

- k-Nearest Neighbors: 0.80

•ROC Curve Analysis and Visualization:

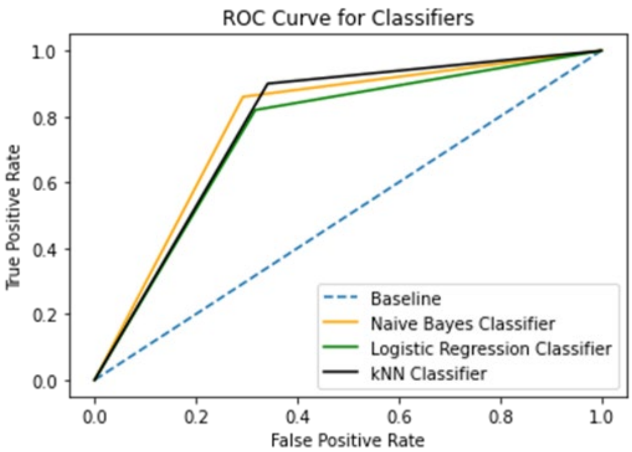

Figure 10. ROC Curve for the classifiers

The ROC curve analysis indicated that the k-Nearest Neighbors classifier achieved the highest AUC score, followed closely by the Naive Bayes classifier, and lastly, the Logistic Regression classifier. The ROC curves visualize the classifiers' ability to discriminate between positive and negative instances at different classification thresholds.

6. Discussion

The comparison of Naive Bayes, Logistic Regression, and k-Nearest Neighbors classifiers for heart disease prediction revealed distinct strengths. Naive Bayes displayed competitive performance, Logistic Regression balanced interpretability with accuracy, while k-Nearest Neighbors excelled in accuracy and precision. The clinical implications are significant, aiding early intervention. Future research can explore advanced feature engineering, ensemble methods, and model interpretability. In conclusion, this study contributes to predictive medicine, offering practitioners valuable tools for enhanced diagnostics and patient care, with potential transformative impacts on healthcare delivery and outcomes.

7. Comparison and Interpretation of Results

7.1. Classifier Performance Comparison

The comparison of classifier performance provides valuable insights into their respective strengths and weaknesses, aiding in the selection of an appropriate model for heart disease prediction. Our study evaluated the three classifiers—Naive Bayes, Logistic Regression, and k-Nearest Neighbors—based on accuracy, precision, recall, and the ability to discriminate between positive and negative instances.

Naive Bayes Classifier: -
The Naive Bayes classifier demonstrated competitive performance in terms of accuracy, precision, and recall. Its Confusion Matrix revealed a balanced distribution of true positive and true negative predictions. The ROC Curve highlighted its abilityto achieve a high true positive rate while maintaining a relatively low false positive rate. The classifier's reliance on the assumption of feature independence is noteworthy, as it allows for efficient and rapid training, making it suitable for large datasets. However, this assumption might limit its performance when dealing with complex interdependencies among features.

Logistic Regression Classifier: -
The Logistic Regression classifier exhibited a slightly lower accuracy compared to Naive Bayes, yet it maintained a balanced precision-recall trade-off. The Confusion Matrix displayed a distribution that was skewed towards

true positives, indicating its capability to identify instances of heart disease. The ROC Curve demonstrated its strong discriminatory power, capturing the variation in true positive and false positive rates across different thresholds. Logistic Regression's interpretability further enhances its utility, as it provides insight into the significance of individual features' contributions to predictions.

k-Nearest Neighbors Classifier: -
The k-Nearest Neighbors classifier outperformed the others, boasting the highest accuracy, precision, recall, and F1-score. Its Confusion Matrix exhibited a balanced distribution of predictions across all categories, underscoring its ability to effectively identify both positive and negative instances. The ROC Curve showcased its excellent true positive rate while maintaining a relatively low false positive rate. This classifier's strength lies in its capability to capture intricate relationships between features, making itsuitable for complex datasets. However, it requires careful tuning of the number of neighbors to prevent overfitting or underfitting.

7.2. Clinical Relevance

The implications of our classifier comparison for clinical practice are profound. Timelyand accurate diagnosis of heart disease is crucial for effective medical intervention. The k-Nearest Neighbors classifier's superior performance suggests its potential utility as a reliable tool for aiding healthcare practitioners in identifying individuals at risk of heart disease. This predictive model can assist in early intervention and personalized treatment strategies, ultimately improving patient outcomes.

Furthermore, the insights gained from the Confusion Matrix and ROC Curve analyses provide transparency into the classifier's decision-making process. Healthcare professionals can comprehend the trade-offs between true positives and false positives, facilitating in formed clinical decisions and enhancing trust in the model's predictions

7.3. Model Robustness and Generalization

Our study's assessment of model robustness and generalization revealed promising results. The slight decrease in accuracy on the test dataset, compared to the training dataset, indicates that the classifiers effectively avoided overfitting. This suggests that the models can generalize well to new, unseen data, thereby enhancing their practical applicability.

7.4. Implications and Future Scope

The outcomes of our study have significant implications for the field of cardiovascular disease diagnosis and classification. The performance comparison of the three classifiers—Naive Bayes, Logistic Regression, and k-Nearest Neighbors—offers valuable insights into their suitability for predicting heart disease based on the given dataset. The evaluation metrics, including accuracy, precision, recall, and AUC scores, shed light on the strengths and weaknesses of each classifier, guiding medical practitioners and researchers in selecting suitable machine learning techniques for heart disease prediction.

7.5. Practical Application

The findings of our study can guide medical practitioners and researchers in selecting appropriate machine learning techniques for heart disease prediction. The Naive Bayes classifier, despite its simplifying assumptions, demonstrated competitive accuracy and AUC scores. This suggests that its independence assumption does not significantly hinder its ability to discriminate between patients with and without heart disease. The Logistic Regression classifier, on the other hand, provides a balance between interpretability and performance, making it a useful tool for identifying key predictors of heart disease.

8. Discussion of Future Directions

While our study provides valuable insights, it also opens avenues for future research in the field of heart disease prediction and machine learning

•Enhanced Feature Engineering: Investigating advanced feature engineering techniques, such as interaction terms and derived variables, could reveal additional patterns and improve the classifiers' performance.
•Ensemble Approaches: Exploring ensemble methods, like stacking or bagging, could harness the collective power of multiple classifiers, potentially boosting predictive accuracy further.
•Interpretable Machine Learning: Delving deeper into interpretable models, such as decision trees or rule-based systems, could offer more transparent insights into the reasoning behind the classifiers' predictions [22].
•Domain-Driven Insights: Incorporating domain-specific medical knowledge and insights could lead to the identification of novel features or relationships crucial for accurate heart disease prediction.
•External Validation: Extending the evaluation to diverse external datasets would validate the classifiers' generalizability and reliability across varying patient populations.
•Real-world Deployment: Investigating the practical implementation and performance of the classifiers in real-world clinical settings is crucial for assessing their utility and impact on patient care.

9. Conclusion

In conclusion, our study comprehensively evaluated the performance of three classifiers—Naive Bayes, Logistic

Regression, and k-Nearest Neighbors—in predicting heart disease. Through a meticulous analysis, we gained insights into their respective strengths and weaknesses, aiding informed decision-making in medical diagnostics. The k-Nearest Neighbors classifier emerged as a standout performer, showcasing remarkable accuracy, precision, recall, and F1-score. Its ability to effectively discern positive and negative instances, coupled with its proficiency in capturing intricate feature relationships, positions it as a valuable tool for early disease detection. Nevertheless, we acknowledge the interpretability of Logistic Regression and the efficiency of Naive Bayes in specific contexts. This study not only offers valuable guidance for selecting appropriate classifiers but also highlights potential.

Research directions for further improving heart disease prediction. As the field of predictive modeling continues to evolve, the insights gained from our study contribute to the ongoing effort of refining and implementing these models in real-world clinical scenarios, ultimately benefiting patient care and medical decision-making.

References

[1] Abdel-Basset, M. et al.: A novel group decision making model based on neutrosophic sets for heart disease diagnosis. Multimed Tools Appl. 79, 15–16, 9977–10002 (2020). https://doi.org/10.1007/S11042-019-07742-7/METRICS.

[2] Agarwal, N: Virtual modem for STB network, 2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011), Allahabad, 2011, pp. 108-112, doi: 10.1109/ICCCT.2011.6075126.

[3] Abed, F.N. et al.: Enhancement Modelling Based on Electrical Discharge Machining Successive Discharges. Advances in Materials Science and Engineering. 2022, (2022). https://doi.org/10.1155/2022/8017375.

[4] Agarwal, N. et al.: A Machine Learning Model to Prune Insignificant Attributes. In: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021. (2021). https://doi.org/10.1109/ICRITO51393.2021.9596425.

[5] Agarwal, N. et al.: A Novel Fbit_sort Algorithm for digital encoding process. In: Proceedings - 2022 6th International Conference on Intelligent Computing and Control Systems, ICICCS 2022. (2022). https://doi.org/10.1109/ICICCS53718.2022.9788168.

[6] Agarwal, N. et al.: A Novel Model to Predict the Effects of Enhanced Students' Computer Interaction on Their Health in COVID-19 Pandemics. New Gener Comput. 1–34 (2023). https://doi.org/10.1007/S00354-023-00224-3/METRICS.

[7] Agarwal, N. et al.: Applying XGBoost Machine Learning Model to Succor Astronomers Detect Exoplanets in Distant Galaxies. (2022). https://doi.org/10.1007/978-3-030-95711-7_33.

[8] Agarwal, N. et al.: Exigency of Natural Language processing to skirmish with COVID-19 pandemic. International Journal of Advances in Engineering and Management (IJAEM). 3, 1081 (2021). https://doi.org/10.35629/5252-030710811086.

[9] Agarwal, N. et al.: Intelligent Aquila Optimization Algorithm-Based Node Localization Scheme for Wireless Sensor Networks. Computers, Materials and Continua. 74, 1, (2023). https://doi.org/10.32604/cmc.2023.030074.

[10] Gupta A, Vardhan H, Varshney S, Saxena S, Singh S, Agarwal N. Kconnect:The Design and Development of Versatile Web Portal for Enhanced Collaboration and Communication. EAI Endorsed Scal Inf Syst

[11] Agarwal, N. et al.: Multiclass Classification of Different Glass Types using Random Forest Classifier. In: Proceedings - 2022 6th International Conference on Intelligent Computing and Control Systems, ICICCS 2022. (2022). https://doi.org/10.1109/ICICCS53718.2022.9788326.

[12] Agarwal, N. et al.: Semi-Supervised Learning with GANs for Melanoma Detection. In: Proceedings - 2022 6th International Conference on Intelligent Computing and Control Systems, ICICCS 2022. (2022). https://doi.org/10.1109/ICICCS53718.2022.9787990.

[13] Agarwal, N., Deep, P.: Obtaining better software product by using test first programming technique. In: Proceedings of the 5th International Conference on Confluence 2014: The Next Generation Information Technology Summit. (2014). https://doi.org/10.1109/CONFLUENCE.2014.6949233.

[14] Agarwal, N., Tayal, D.K.: A new model based on the extended COPRAS method for improving performance during the accreditation process of Indian Higher Educational Institutions. Computer Applications in Engineering Education. 31, 3, (2023). https://doi.org/10.1002/cae.22602.

[15] Agarwal, N., Tayal, D.K.: A Novel Model to Predict the Whack of Pandemics on the International Rankings of Academia. 24–38 (2023). https://doi.org/10.1007/978-3-031-35081-8_3.

[16] Agarwal, N., Tayal, D.K.: FFT based ensembled model to predict ranks of higher educational institutions. Multimed Tools Appl. 81, 23, (2022). https://doi.org/10.1007/s11042-022-13180-9.

[17] N. Agarwal and P. Deep, "Obtaining better software product by using test first programming technique," 2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence), Noida, India, 2014, pp. 742-747, doi: 10.1109/CONFLUENCE.2014.6949233.

[18] Anushka et al.: Predicting Credit Card Defaults with Machine Learning Algorithm Using Customer Database. 262–277 (2023). https://doi.org/10.1007/978-3-031-35078-8_22.

[19] Jha, A. et al.: Movie Recommendation Using Content-Based and Collaborative Filtering Approach. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST. 470 LNICST, 439–450 (2023). https://doi.org/10.1007/978-3-031-35078-8_37/COVER.

[20] Agarwal N, Arora I, Saini H, Sharma U. A Novel Approach for Earthquake Prediction Using Random Forest and Neural Networks. EAI Endorsed Trans Energy Web.

[21] Kimenai, D.M. et al.: Sex Differences in Cardiac Troponin I and T and the Prediction of Cardiovascular Events in the General Population. https://doi.org/10.1093/clinchem/hvab109.

[22] Mohan, S. et al.: Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. IEEE Access. 7, 81542–81554 (2019). https://doi.org/10.1109/ACCESS.2019.2923707.

[23] Shanmukha, R.K.N.S., Thinakaran, K.: Prediction of Heart Disease using Decision Tree over Logistic Regression using Machine Learning with Improved Accuracy. CARDIOMETRY. 25, 1514–1519 (2023). https://doi.org/10.18137/CARDIOMETRY.2022.25.15141519.

[24] Sharma, T. et al.: Machine Learning/Deep Learning for Natural Disasters. Applications of Artificial Intelligence, Big Data and Internet of Things in Sustainable Development. 91–121 (2022). https://doi.org/10.1201/9781003245469-7.

[25] Sharma, V. et al.: Heart Disease Prediction using Machine Learning Techniques. Proceedings - IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2020. 177–181 (2020). https://doi.org/10.1109/ICACCCN51052.2020.9362842.

[26] Singh, A., Kumar, R.: Heart Disease Prediction Using Machine Learning Algorithms. 2020 International Conference on Electrical and Electronics Engineering (ICE3). 452–457 (2020). https://doi.org/10.1109/ICE348803.2020.9122958.

[27] Sivaparthipan, C.B. et al.: IoT-based patient stretcher movement simulation in smart hospital using type-2 fuzzy sets systems. Production Planning and Control. (2023). https://doi.org/10.1080/09537287.2023.2217419.

[28] Srivastav, A. et al.: A Novel Fuzzy Graph Connectivity Measure to Perform Word Sense Disambiguation Using Fuzzy Hindi WordNet. In: 3rd IEEE 2022 International Conference on Computing, Communication, and Intelligent Systems, ICCCIS 2022. (2022). https://doi.org/10.1109/ICCCIS56430.2022.10037658.

[29] Tayal, D.K. et al.: To Predict the Fire Outbreak in Australia using Historical Database. In: 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2022. (2022). https://doi.org/10.1109/ICRITO56286.2022.9964603.