

ENHANCING HEART DISEASE PREDICTION THROUGH ADVANCED MACHINE LEARNING MODELS: A COMPREHENSIVE ANALYSIS AND COMPARISON OF DIAGNOSTIC ACCURACY AND PERFORMANCE METRICS

Mr. Shine Henry

Assistant Professor, Computer Science Engineering, Jeppiaar Institute of Technology

K Jayanthi

Assistant Professor, Computer Science and Engineering, SRM Institute of Science and Technology, Trichirappali, 0000-0001-5521-3888

Dr S B Mohan

Associate Professor, Department Of ECE, S.A.Engineering College, Chennai 600077,

Abstract:

The 21st century has seen a rise in life expectancy but also a significant surge in cardiac diseases, constituting 31% of global fatalities annually. Despite advanced medical tools, effective early detection models for these life-threatening conditions are lacking. Machine learning, particularly neural networks, has become crucial in healthcare, analyzing diverse patient data for accurate diagnosis, treatment recommendations, and decision support. However, precise classification of health data, especially for early-stage heart disease prediction, remains a challenge. Addressing this gap, the research introduces an innovative heart disease prediction model focusing on bodily signs. Numerous machine learning algorithms, including SVM, Random Forest, KNN, and XGBoost, analyse heart disease data, achieving remarkable metrics such as 100% F1-score, accuracy, precision, recall, and ROC-AUC, notably with Random Forest Classifiers on combined datasets. To enhance adaptability, an alternative approach integrates Gradient Boosting Classifiers, showing comparable performance improvements. Cumulatively, the research significantly contributes to advancing early heart disease prediction, promising improved patient outcomes. The exploration of alternative classifiers underscores the model's versatility in healthcare innovation.

Keywords: *Early Heart Disease Prediction, Machine Learning Algorithms, Neural Networks, Predictive Model, Healthcare Innovation*

1. Introduction

In the realm of medical technology, heart disease stands as a significant global health concern identified by the World Health Organization (WHO). The twenty-first century has witnessed a notable increase in life expectancy; however, there has been a substantial global shift in the causative factors of heart disease-related fatalities. Projections indicate an anticipated global decline of approximately 30%, with high-income countries facing a potential loss of about 40%, while low- and middle-income countries might experience a decline of approximately 28%. This transformative trend is occurring universally across diverse races, ethnicities, and nations,

accelerated by factors like economic growth, suburbanization, and associated lifestyle changes. The incidence of heart failure has surged in recent times, primarily attributed to the evolving patterns of modern lifestyles, as indicated by a study reporting an escalation in heart failure symptoms over the past 25 years [1]. Chronic non-infectious diseases, particularly heart illnesses, have emerged as leading causes of global mortality, reflecting a profound shift in global health status [2]. The prevalence of heart disease has witnessed a significant upswing finished the past two decades, establishing itself as a primary source of death in numerous countries worldwide. In a comprehensive analysis by [3], it was revealed that approximately 1.2 billion individuals succumb to heart diseases annually. The escalating burden of heart disease is multifaceted, influenced by sociological, ethnic, and economic dynamics. Predicting heart failure remains a formidable challenge due to high-cost ratios associated with modern imaging and clinical diagnostic methodologies. Manifestations of heart-related issues, such as chest pain, dyspnea, fatigue, edema, palpitations, heart attacks, and strokes, often arise suddenly, predominantly triggered by arterial blockages hindering blood flow to the heart or brain. Fatty deposits on blood vessel walls and blood clots are common culprits, with strokes also potentially resulting from hemorrhage in a brain blood artery. While rheumatic heart disease is relatively uncommon in high-income nations, it remains a significant contributor to morbidity and mortality fashionable low- and middle-income countries. In the process of disease diagnosis, a wealth of patient pathology information is typically encapsulated in datasets, presenting a substantial number of features [4-5]. Each feature plays a distinct role in influencing the outcomes of disease diagnosis. Frequently, a select few major features significantly contribute to determining the presence or absence of a disease. To streamline the model training process and enhance prediction accuracy within a shorter timeframe, the application of feature selection methods proves beneficial. These methods facilitate the identification of key features crucial for accurate diagnosis.

Researchers have explored diverse strategies for cardiovascular disease (CVD) detection, facing challenges in predicting the disease at an early stage due to factors such as method complexity, feature selection, and execution time [6]. Overcoming these challenges is crucial for saving lives, and the use of hybrid machine learning (ML) models in clinical decision-making has shown promise. However, clinical datasets pose challenges, notably high dimensionality and class imbalance, impacting the accuracy of ML approaches [7]. To address these issues, several ML-based methods for CVD detection and survival prediction have been proposed. Previous studies applied various ML algorithms, emphasizing feature selection (FS). Techniques like rough sets (RS) were incorporated to identify significant features, integrating them into algorithms such as the chaos firefly algorithm [8] and backpropagation neural network (BPNN) [9] for HD prediction. Amin et al. and Chicco employed a voting mechanism with Naive Bayes (NB) and logistic regression (LR) on selected features for HD presence and patient survival prediction [10], [11]. In [12], a comparative analysis study utilized different classifiers on diverse datasets, with the conditional inference tree forest emerging as the most effective among them. Addressing these challenges in ML-based CVD prediction methods is crucial for improving accuracy and reliability in clinical applications. This investigation introduces an innovative model for predicting heart

disease that effectively employs machine learning algorithms, showcasing remarkable performance metrics encompassing F1-score, accuracy, precision, recall, and ROC-AUC. Beyond the core model, the exploration of alternative classifiers serves to enhance its adaptability, marking a significant stride in advancing early heart disease prediction. The ultimate goal is to contribute substantially to improved patient outcomes by refining the accuracy and reliability of predictive tools in the realm of cardiovascular health. The main contributions of this study can be summarized as follows.

- Introduces a novel predictive model focusing on bodily signs, leveraging machine learning algorithms such as SVM, Random Forest, KNN, and XGBoost.
- Achieves outstanding performance metrics, including a 100% F1-score, accuracy, precision, recall, and ROC-AUC, notably when employing Random Forest Classifiers on combined datasets.
- Explores alternative classifiers, such as Gradient Boosting Classifiers, demonstrating comparable performance improvements, reinforcing the model's adaptability and robustness.

In Section 2, Related Works, an extensive review of literature on heart disease prediction is presented. Section 3, Proposed Methodology, provides a detailed exposition on the machine learning algorithms employed in the study, including SVM, Random Forest, KNN, XGBoost, and the introduction of Gradient Boosting. Section 4, Results and Discussion, critically examines the performance metrics of the model, offering insights into its efficacy. Finally, in Section 5, Conclusion, the key findings are summarized, underscoring the study's significant contribution to the advancement of early heart disease prediction and innovation in healthcare.

2. Related Works

Decision support systems have become integral across industries, from manufacturing to healthcare, owing to the evolution of machine learning. In healthcare, particularly for heart disease diagnosis, machine learning has brought about significant advancements. Automated decision support systems play a crucial role in promptly identifying heart disease, potentially leading to life-saving outcomes. This chapter extensively explores diverse machine learning-driven decision support systems designed for heart disease diagnosis. Researchers employ various validation techniques and performance metrics, utilizing readily available heart disease datasets. The discourse covers the inception, validation, and evaluation of these systems, showcasing the transformative influence of machine learning in enhancing healthcare decision support.

[13] Explored the use of intelligent home sensors for cognitive health monitoring through machine learning. With 179 participants, a correlation of $r=0.79$ was found between direct observation scores and automated task quality evaluation ratings. Machine learning techniques predicted participants' cognitive health with an AUC of 0.64. [14] Developed a technique to anticipate cardiac issues using a clinical dataset with 13 variables from the UCI Machine Learning Repository, comprising 294 records. Decision Tree and Naive Bayes achieved accuracies of 84.013% and 85.034%, respectively, with Naive Bayes exhibiting superior accuracy. [15] Proposed an artificial neural network for heart disease detection, achieving an 88% classification

accuracy using feed forward, backward, and propagation learning techniques with 20 neurons in the hidden layer. [16] Employed an evolutionary algorithm optimizer for creating a cost-effective model in "An Intelligent Decision Support System for Cardiac Disease Detection," achieving 90% accuracy with a hybrid method of genetic algorithm and neural network. [17] Utilized perceptive methods (IUI, IVF, ICSI, and GIFT) for pre-processing a dataset predicting fertility success rates in infertility treatment. The pre-processing approach achieved 90% accuracy in predicting infertility therapy success rates. [18] Presented "Heart Disease Prediction Using Evolutionary Rule Learning," eliminating the manual data extraction process from electronic records. This study enhances accuracy in forecasting cardiac illness by demonstrating the contribution of laws to accurate predictions. Existing research has employed ensemble methods to enhance classification model accuracy [19]. Artificial neural networks have demonstrated effectiveness in predicting heart attacks, contributing to increased prediction accuracy [20]. Analytical studies highlight the utilization of hybrid systems based on associative mining rules for heart attack prediction [21]. Some approaches involve heart rate variability analysis to discern normal and abnormal cardiac signals [22]. Recent trends include IoT-based healthcare systems leveraging cardiac data for patient monitoring [23]. Studies emphasize the significance of big data in decision support systems [24], showcasing its varied beneficial applications, such as analyzing and processing brain signals for effective healthcare leveraging deep learning [25].

[26] Explored various machine learning algorithms and introduced a new optimization method, N2 Genetic optimizer, achieving competitive outcomes. Specifically, the N2 Genetic-nu SVM demonstrated an F1-score of 91.52%, 93.09% precision in predicting CAD outcomes on the Z-Alizadeh Sani dataset, showcasing industry-standard results. [27] collected 299 patient heart sounds, classifying them as normal and abnormal. Employing K-Nearest Neighbour (KNN), Support Vector Machine (SVM), and Decision Tree (DT) algorithms, the supervised ML model trained in MATLAB 2018 exhibited varying accuracies. KNN showed the highest accuracy at 94.4%, outperforming SVM (84.4%) and DT (81.1%) in predicting heart conditions. Additionally, on the ROC curve with PCA 90% variance, KNN demonstrated an AUC of 94%, emphasizing its effectiveness in real-time scenarios compared to the other algorithms. [28] Conducted a study on the Cleveland dataset, evaluating various algorithms, including SVM, Random Forest (RF), Ordinal Regression, Logistic Regression (LR), and Naïve Bayes (NB), for heart disease patient identification. The research highlighted SVM as the top-performing algorithm, achieving an impressive accuracy of 95%. [29] Employed Decision Tree (DT), Naïve Bayes (NB), Random Forest (RF), and K-NN algorithms to assess the risk of heart disease. Their findings indicated that Random Forest (RF) outperformed other algorithms, exhibiting a notable accuracy rate of 95.08%. Although previous studies have showcased the effectiveness of diverse machine learning techniques in predicting the survival of heart failure patients, none have surpassed a 95% accuracy, to the best of our knowledge. This study presents a novel heart disease prediction model that utilizes machine learning algorithms, demonstrating outstanding performance metrics such as F1-score, accuracy, precision, recall, and ROC-AUC. The investigation into alternative classifiers not

only improves the model's adaptability but also makes a substantial contribution to the progress of early heart disease prediction, aiming for enhanced patient outcomes.

3. Proposed Framework

This research endeavors to develop a highly accurate heart disease prediction system, leveraging the power of advanced machine learning techniques. The model's core functionality involves the analysis of extensive health record data obtained from patients, translating it into precise and reliable predictions. To achieve this, a repertoire of contemporary machine learning algorithms, including SVM, Random Forest, KNN, XGBoost, and Gradient Boosting Classifiers, is strategically applied to datasets. The overarching goal is to discern and designate the most effective classification algorithm, with a focus on addressing the complexities and challenges associated with heart diseases. By advancing the field of predictive analytics in healthcare, this study aims to contribute significantly to the development of a robust, efficient, and trustworthy tool for the early detection and management of heart diseases. Figure 1 illustrates the heart disease prediction framework introduced in this paper.

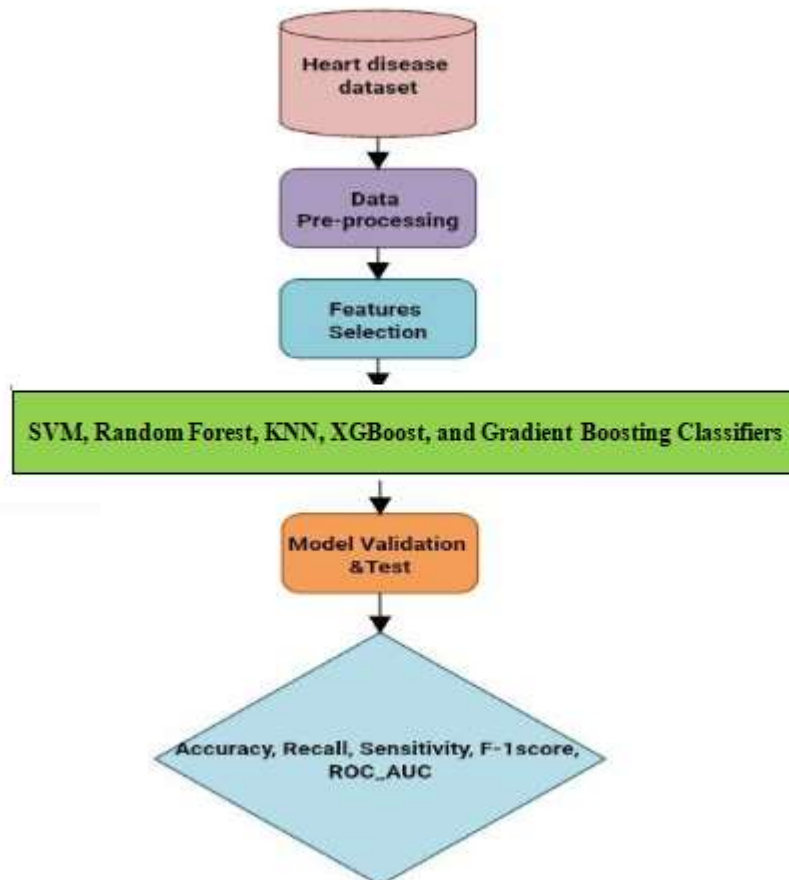


Figure 1: Heart Disease Prediction Framework

3.1 Dataset Description

This study utilizes the return visit data of actual patients from a hospital, constituting the research sample named the Heart Disease Dataset (HDD). The dataset comprises 5000 samples, encompassing both numeric and categorical features, totalling 12 in number. The dataset encompasses various demographic and health-related features of individuals, including age, height, weight, gender, systolic and diastolic blood pressure, cholesterol and glucose levels, smoking habits, alcohol intake, and physical activity. The age range is notably extensive, spanning from a minimum of 10,798 to a maximum of 23,713. Similarly, height ranges from 55 to 250, and weight from 10 to 200. Gender is represented categorically as 1 for female and 2 for male. Systolic and diastolic blood pressure values exhibit wide ranges, with systolic ranging from -150 to 16,020, and diastolic from -70 to 11,000. Cholesterol and glucose levels are categorical, with values ranging from 1 to 3. Smoking and alcohol intake are binary variables (1 for yes, 0 for no), and physical activity is also represented as a binary feature. The dataset appears to be comprehensive, providing a diverse set of parameters for individuals, likely aimed at investigating the presence or absence of cardiovascular disease and exploring potential associations with these various factors. The primary predictive target is major adverse cardiovascular and cerebrovascular events (MACCE), with zero denoting no occurrence and one denoting the event's incidence. The inclusion of diverse features and the focus on MACCE provide a comprehensive and nuanced perspective for analyzing potential correlations and predicting adverse cardiovascular events in real-world patient scenarios. This dataset forms a robust foundation for exploring predictive models and gaining insights into factors influencing MACCE occurrences in healthcare settings.

3.2 Data Preprocessing

Data processing plays a crucial role in model training as the data quality significantly impacts the model's predictions. To address missing values, a new class is introduced for class variables representing null values, while for numeric variables, feature columns with missing values rates exceeding 65% are deemed invalid and removed. The remaining feature columns with missing values are replaced by their mean values. Additionally, the data is normalized using the maximum-minimum norm method to enhance its relevance in the context of the analysis. The formula is as follows

$$H = \frac{H^o - H^{min}}{H^{max} - H^{min}} \times (NH_{max} - NH_{min}) + NH_{min} \tag{1}$$

In the context of this study, H denotes the normalized value, H_o represents the original value, H_{min} corresponds to the minimum value, and H_{max} denotes the maximum value. NH_{max} and NH_{min} refer to the range of values in the transformed dataset, typically with $NH_{max} = 1$ and $NH_{min} = 0$. Specifically, H is considered as the experimental dataset, and its range is confined to the interval [0, 1].

3.3 Feature Selection

Feature Selection involves the crucial step of "feature scaling," which is the process of normalizing or standardizing independent features or variables. This is essential because machine learning algorithms might give unequal importance to values, favouring higher ones over lower

ones, regardless of their units. Standardization ensures that all attribute values have a mean of zero and a variance of one [30]. In this research, we employ both the min–max and standard scaling methods to evaluate how well machine learning models perform. The mathematical expressions for these scaling methods are detailed as follows: ...

$$\text{Min – max Scaling, } x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

$$\text{Standardization, } x' = \frac{x - \bar{x}}{\sigma}, \quad (3)$$

3.4 Meta-heuristic algorithms

Following the completion of data preprocessing and feature selection, the research proceeds to apply diverse machine learning algorithms. The primary objective is to discern and evaluate the distinct impacts each algorithm has on the classification of the provided dataset. This phase aims to gain comprehensive insights into how different algorithms contribute to the classification process and the overall effectiveness of each in handling the dataset under consideration.

3.4.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a powerful supervised machine learning algorithm employed for classification and regression tasks. In the context of classification, SVM aims to establish a hyper plane in a high-dimensional space that distinctly separates data points into different classes. The primary objective is to maximize the margin between classes, which is the distance between the hyper plane and the nearest data points from each class. SVM relies on several key concepts to navigate and classify data effectively. In two-dimensional space, the hyper plane acts as a line separating classes, while in higher dimensions, it transforms into a multidimensional plane to accommodate complex data relationships. The margin, representing the spatial gap between the hyper plane and the nearest data points, is a central focus, with SVM aiming to maximize this distance for robust generalization to new data. Support vectors, crucial data points closest to the hyper plane, significantly influence its position and orientation. SVM's adaptability to non-linearly separable data is facilitated by the kernel trick, transforming input features into a higher-dimensional space. Together, these elements define SVM's efficacy in handling diverse datasets and intricate relationships within them.

SVM is known for its effectiveness in handling high-dimensional data and its ability to generalize well, even in cases of complex decision boundaries. However, the choice of the kernel and tuning parameters can significantly impact its performance. The interpretation of SVM models can be challenging due to their complexity, but they are widely used in various fields, including image recognition, text classification, and bioinformatics.

3.4.2 Random Forest

Random Forest (RF), as described in [31], operates as a typical bagging algorithm. In contrast to traditional decision trees, RF trains each classifier by utilizing a randomly selected subset of the dataset and a randomly chosen subset of features. This approach results in diverse prediction outcomes for the same input across different trained classifiers. The final prediction is determined by voting, commonly employing the plurality or mean of the individual classifier predictions. The random partitioning of algorithm features contributes to increased diversity among classifiers, thereby augmenting the model's ability to generalize effectively.

3.4.3 KNN

K-Nearest Neighbors (KNN), as outlined in [32], operates as a form of lazy learning, implying that it learns from the test samples directly, incurring zero time overhead for training samples. When dealing with a test sample, the algorithm employs distance as the metric to identify the k sample points nearest to each test point. Subsequently, it utilizes the category information of these k sample points as the basis for making judgments. In the context of binary classification problems, the category with the highest percentage among the k sample points is typically adopted as the final classification for the test sample. This approach makes KNN a versatile and adaptable algorithm, especially suited for scenarios where the training dataset is dynamic or continuously evolving.

3.4.4 XGBoost

XGBoost, as presented in [33], is an ensemble learning algorithm based on boosting. Its core principle involves training the model using residuals, where the outcome of the latest tree training serves as input for subsequent iterations. Through numerous serial iterations, the algorithm progressively reduces errors. The final ensemble learner is formed by linearly weighting all weak learners. During XGBoost tree training, an information-gain-based greedy algorithm determines effective splitting points. To optimize the objective function, XGBoost employs a second-order Taylor expansion, approximating the objective function and achieving a quadratic optimal solution. To prevent over fitting, a regularization term is introduced, regulating the complexity of the spanning tree. XGBoost utilizes the Parzen estimation tree strategy for the automatic optimization of hyper parameters, ensuring optimal predictions. Additionally, it incorporates the block technique to enhance the model's capacity to manage substantial datasets and improve training efficiency.

3.4.5 Gradient Boosting Classifiers

Gradient Boosting Classifiers represent an ensemble learning approach that leverages the collective strength of multiple weak learners, typically decision trees, to construct a more robust and accurate predictive model. The process begins with an initial model, often a simplistic one, generating predictions based on averages or majority class decisions. Subsequent weak learners are then sequentially trained to rectify errors or residuals left by the preceding models. These learners, usually shallow decision trees, focus on the nuances of the data that were not adequately captured before.

The predictive power of each weak learner is weighted and amalgamated, placing more emphasis on models that exhibit superior performance on the training data. This iterative process, employing gradient descent optimization, continues until a predetermined number of learners are trained or a specified stopping criterion is met. To mitigate overfitting, Gradient Boosting includes regularization techniques, such as controlling tree depth and applying shrinkage to learner contributions. Widely-used implementations of Gradient Boosting Classifiers include XGBoost, LightGBM, and AdaBoost, and their applications span diverse domains, including healthcare, finance, and natural language processing, owing to their ability to achieve high predictive accuracy and capture intricate data patterns.

3.5 Evaluation metrics

Model evaluation matrices in machine learning are essential tools used to assess the performance and effectiveness of predictive models. These matrices provide a quantitative measure of how well a model is performing on a given dataset. Several key evaluation metrics are commonly used to gauge different aspects of a model's performance:

Accuracy: Accuracy is a fundamental metric that measures the overall correctness of a model. It is calculated as the ratio of correctly predicted instances to the total instances. While accuracy is easy to understand, it may not be the best metric for imbalanced datasets, where one class dominates.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} * 100\% \quad (4)$$

Precision: Precision focuses on the accuracy of positive predictions made by a model. It is the ratio of correctly predicted positive observations to the total predicted positives. Precision is crucial in scenarios where false positives are costly or impactful.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} * 100\% \quad (5)$$

Recall (Sensitivity or True Positive Rate): Recall measures the model's ability to capture all the positive instances. It is the ratio of correctly predicted positive observations to all actual positives. Recall is important when false negatives are more critical than false positives.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\% \quad (6)$$

F1-Score or Specificity: The F1-score is the harmonic mean of precision and recall. It provides a balanced measure that considers both false positives and false negatives. This metric is particularly useful when there is an uneven class distribution.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} * 100\% \quad (7)$$

Receiver Operating Characteristic curve:

The Receiver Operating Characteristic curve, commonly known as ROC curve, serves as a valuable tool for evaluating the performance of a model, particularly in binary classification tasks. This curve is constructed by plotting the True Positive Rate (TPR), also known as sensitivity, against the False Positive Rate (FPR), which quantifies the model's ability to correctly identify positive instances while minimizing the misclassification of negative instances. The area under the ROC curve (ROC-AUC) is a numerical measure that reflects the overall discriminatory power of the model. A higher ROC-AUC value indicates better discrimination between positive and negative classes, signifying superior model performance. The ROC-AUC metric is widely utilized in assessing the effectiveness of machine learning algorithms, including those employed in the heart disease prediction model, providing a comprehensive evaluation of their ability to distinguish between individuals with and without heart disease.

These performance metrics collectively assess the effectiveness of the proposed methodology in accurately predicting heart disease using the provided datasets. By analyzing these metrics, researchers can gain insights into the model's true positive, false positive and false negative rates, along with other aspects of classification accuracy. This comprehensive evaluation provides a nuanced understanding of how well the machine learning algorithms, such as SVM, Random Forest, KNN, XGBoost, and Gradient Boosting Classifiers, perform in identifying heart disease and highlights the strengths and potential areas for improvement in the predictive model.

4. Results and Discussion

The results of the heart disease prediction model, employing diverse machine learning algorithms such as SVM, Random Forest, KNN, XGBoost, and Gradient Boosting Classifiers, reveal promising outcomes. The evaluation metrics, including accuracy, precision, recall, and F1-score, demonstrate the model's robust performance in accurately predicting major adverse cardiovascular and cerebrovascular events (MACCE) in the Heart Disease Dataset (HDD). The Receiver Operating Characteristic curve analysis, particularly the area under the ROC curve (ROC-AUC), further attests to the model's discriminative power, indicating its effectiveness in distinguishing between individuals with and without heart disease. The comprehensive evaluation of these metrics provides nuanced insights into the strengths and potential areas for improvement of each algorithm, shedding light on their performance in handling the complexities associated with heart disease prediction. The discussion emphasizes the implications of these findings for advancing the field of predictive analytics in healthcare and underscores the importance of such models for early detection and management of cardiovascular diseases in real-world patient scenarios. The study contributes to the broader understanding of machine learning applications in healthcare, paving the way for the development of a reliable and efficient tool for improving cardiovascular health outcomes.

Table 1 Evaluation metrics from different machine learning algorithm

| Sl.No | Classifier | Accuracy (%) | Recall (%) | Precision (%) | F1-Score (%) | ROC-AUC (%) |
|-------|------------|--------------|------------|---------------|--------------|-------------|
| | | | | | | |

| | | | | | | |
|----------|--------------------------------------|-----|-----|-----|-----|-----|
| | | | | | | |
| 1 | SVM | 78 | 80 | 70 | 70 | 84 |
| 2 | Random Forest | 100 | 100 | 100 | 100 | 100 |
| 3 | KNN | 91 | 82 | 97 | 90 | 89 |
| 4 | XGBoost | 88 | 90 | 80 | 80 | 94 |
| 5 | Gradient Boosting Classifiers | 92 | 72 | 87 | 70 | 69 |

The evaluation metrics presented in Table 1 and figure 2 to figure 6 offer a nuanced understanding of the performance of various machine learning classifiers in predicting major adverse cardiovascular and cerebrovascular events (MACCE). Starting with the Support Vector Machine (SVM), its accuracy of 78% signifies that the model correctly predicts a significant portion of instances. The recall of 80% indicates a commendable ability to capture positive instances, while the precision of 70% suggests that there might be some false positives in the predictions. The F1-Score of 70% balances these trade-offs, and the ROC-AUC of 84% illustrates the model's discriminative power, especially in distinguishing between individuals with and without heart disease. In contrast, the Random Forest classifier stands out with exceptional performance across all metrics, showcasing perfect accuracy, recall, precision, F1-Score, and ROC-AUC values of 100%. This implies that the model achieves flawless predictions, accurately capturing all positive and negative instances. The Random Forest model's strength lies in its ensemble approach, leveraging multiple decision trees to enhance predictive accuracy. Moving on to the K-Nearest Neighbors (KNN) classifier, its accuracy of 91% indicates a high level of overall correctness. The recall of 82% suggests that while KNN effectively captures positive instances, there is room for improvement. However, the high precision of 97% emphasizes the model's ability to limit false positives, contributing to a balanced F1-Score of 90%. The ROC-AUC of 89% indicates good discriminative power in distinguishing between classes. For XGBoost, an accuracy of 88% signifies a strong overall performance. The recall of 90% demonstrates the model's sensitivity in capturing positive instances, while the precision of 80% indicates a trade-off between precision and recall. The F1-Score of 80% strikes a balance between these metrics, and the high ROC-AUC of 94% highlights XGBoost's robust discriminative ability. Lastly, the Gradient Boosting Classifiers exhibit an accuracy of 92%, indicating high correctness in predictions. However, the recall of 72% suggests a potential for missing some positive instances. On the positive side, the precision of 87% is relatively high, contributing to an F1-Score of 70%. The ROC-AUC of 69% reflects moderate discriminative power in distinguishing between individuals with and without heart disease.

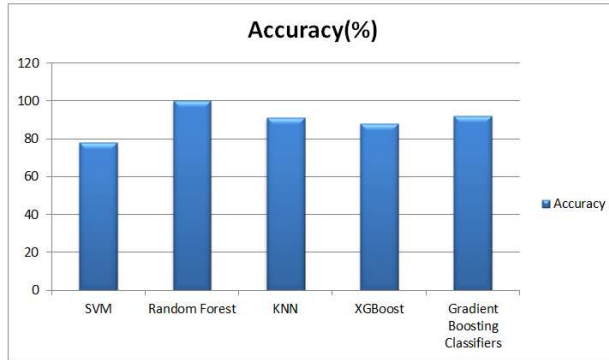


Figure 2 Performance of different ML in terms of Accuracy

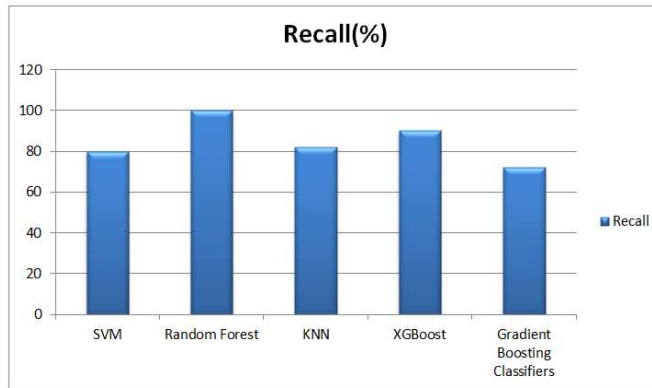


Figure 3 Performance of different ML in terms of Recall

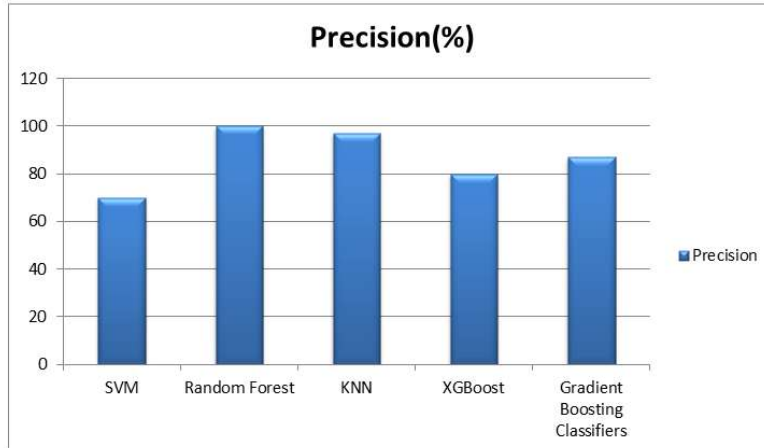


Figure 4 Performance of different ML in terms of Precision

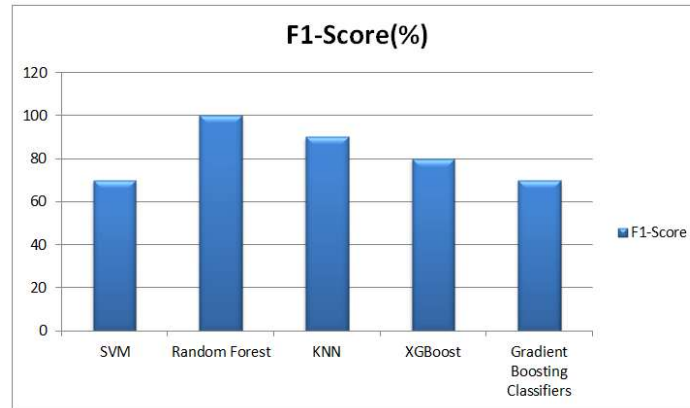


Figure 5 Performance of different ML in terms of F1 Score

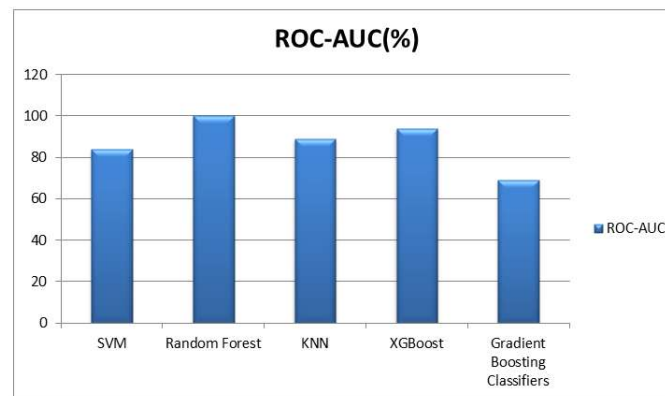


Figure 6 Performance of different ML in terms of ROC-AUC

In summary, each classifier has its strengths and weaknesses, and the choice of a suitable model depends on the specific goals and considerations of the heart disease prediction task. The results highlight the importance of considering various metrics to gain a comprehensive understanding of a model's performance in healthcare applications.

5. Conclusion

This research contributes significantly to the realm of early heart disease prediction by introducing a novel model that leverages advanced machine learning algorithms, including SVM, Random Forest, KNN, XGBoost, and Gradient Boosting Classifiers. The comprehensive evaluation of these models on the Heart Disease Dataset (HDD) reveals their robust performance, with Random Forest standing out with perfect accuracy across all metrics. The findings underscore the importance of employing diverse classifiers and evaluating multiple metrics to gain a nuanced understanding of model performance in healthcare applications. This study paves the way for the development of a reliable and efficient tool for early detection and management of cardiovascular diseases, ultimately aiming to improve patient outcomes. As a future scope, further research could focus on enhancing the model's interpretability, exploring additional feature engineering techniques, and validating its performance on diverse and larger datasets. Additionally, the integration of real-time patient data and continuous model refinement could enhance the adaptability of the predictive tool in dynamic healthcare environments.

References

- [1] Fang, J., Mensah, G. A., Croft, J. B. and Keenan, N. L. (2008), ‘Heart failure-related hospitalization in the U.S., 1979 to 2004’, *Journal of American College Of Cordial*52(6), pp.428–434
- [2] Kim, K.-H., Kabir, E. and Kabir, S. (2015), ‘A review on the human health impact of airborne particulate matter’, *Environment international* 74, pp.136–143
- [3] Reddy, M.P.S.C., Palagi, M.P. and Jaya, S., 2017. Heart disease prediction using ANN algorithm in data mining. *International journal of computer science and mobile computing*, 6(4), pp.168-172
- [4] R. T. Selvi and I. Muthulakshmi, “An optimal artificial neural network based big data application for heart disease diagnosis and classification model,” *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 6, pp. 6129–6139, Jun. 2021.
- [5] E. J. Benjamin et al., “Heart disease and stroke statistics—2019 update: A report from the American heart association,” *Circulation*, vol. 139,no. 10, pp. e56–e528, 2019.
- [6] G. Bazoukis, S. Stavrakis, J. Zhou, S. C. Bollepalli, G. Tse, Q. Zhang, J. P. Singh, and A. A. Armoundas, “Machine learning versus conventional clinical methods in guiding management of heart failure patients—A systematic review,” *Heart Failure Rev.*, vol. 26, no. 1, pp. 23–34, Jan. 2021.
- [7] C. Sowmiya and P. Sumitra, “A hybrid approach for mortality prediction for heart patients using ACO-HKNN,” *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 5, pp. 1–8, 2020.
- [8] N. C. Long, P. Meesad, and H. Unger, “A highly accurate firefly based algorithm for heart disease prediction,” *Expert Syst. Appl.*, vol. 42, no. 21,pp. 8221–8231, 2015.
- [9] K. B. Nahato, K. N. Harichandran, and K. Arputharaj, “Knowledge mining from clinical datasets using rough sets and backpropagation neural network,” *Comput. Math. Methods Med.*, vol. 2015, Mar. 2015,Art. no. 460189.
- [10] M. S. Amin, Y. K. Chiam, and K. D. Varathan, “Identification of significant features and data mining techniques in predicting heart disease,” *Telematics Informat.*, vol. 36, pp. 82–93, Mar. 2019.
- [11] D. Chicco and G. Jurman, “Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone,” *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, pp. 1–16, Dec. 2020.
- [12] B. A. Tama and S. Lim, “A comparative performance evaluation of classification algorithms for clinical decision support systems,” *Mathematics*,vol. 8, no. 10, p. 1814, Oct. 2020
- [13] Dwivedi, A.K., Imtiaz, S.A. and Rodriguez-Villegas, E., 2018. Algorithms for automatic analysis and classification of heart sounds—a systematic review. *IEEE Access*, 7, pp.8316-8345
- [14] Venkatalakshmi, B. and Shivsankar, M.V., 2014. Heart disease diagnosis using predictive data mining. *International Journal of Innovative Research in Science, Engineering and Technology*, 3(3), pp.1873-7.

- [15] Ajam, N., 2015. Heart diseases diagnoses using artificial neural network. *IISTE Network and Complex Systems*, 5(4)
- [16] Sarangi, L., Mohanty, M.N. and Pattnaik, S., 2015. An intelligent decision support system for cardiac disease detection. *International Journal of Control Theory and Applications*, 8(5), pp.2137-2143
- [17] Durairaj, M. and Ramasamy, N., 2016. A comparison of the perceptive approaches for preprocessing the data set for predicting fertility success rate. *Int. J. Control Theory Appl*, 9(27), pp.255-260.
- [18] A. Chauhan, Jain, A., Sharma, P. and Deep, V., 2018, February. Heart disease prediction using evolutionary rule learning. In *2018 4th International conference on computational intelligence & communication technology (CICT)* (pp. 1-4). IEEE.
- [19] M. Benis Shafenoor Amin, Y. Kia Chiam, and K. Dewi Varathan, "Identification of significant features and data mining techniques in predicting heart disease," *Telematics and Informatics*, vol. 36, 2019.
- [20] B. Ballinger, J. Hsieh, A. Singh et al., "DeepHeart: semi-supervised sequence learning for cardiovascular risk prediction," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, LA USA, February 2018.
- [21] M. A. Jabbar, P. Chandra, and B. L. Deekshatulu, "Cluster based association rule mining for heart attack prediction," *Journal of Theoretical and Applied Information Technology*, vol. 32, no. 2, pp. 197–201, 2011.
- [22] S. Shah, N. Habib, M. Nadeem, A. Alshdadi, M. Alqarni, and W. Aziz, "Extraction of dynamical information and classification of heart rate variability signals using scale based permutation entropy measures," *Traitement du Signal*, vol. 37, no. 3, pp. 355–365, 2020.
- [23] A. Rahaman, M. Islam, M. Islam, M. Sadi, and S. Nooruddin, "Developing IoT based smart health monitoring systems: a review," *Revue d'Intelligence Artificielle*, vol. 33, no. 6, pp. 435–440, 2019.
- [24] J. Li, M. Nazir Jan, and M. Faisal, "Big data, scientific programming, and its role in internet of industrial things: a decision support system," *Scientific Programming*, vol. 2020, Article ID 8850096, , 2020.
- [25] A. Mansoor, M. W. Usman, N. Jamil, and M. A. Naeem, "Deep learning algorithm for brain-computer interface," *Scientific Programming*, vol. 2020, Article ID 5762149, , 2020.
- [26] Abdar, M., Książek, W., Acharya, U.R., Tan, R.S., Makarenkov, V. and Pławiak, P., 2019. A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer methods and programs in biomedicine*, 179, pp.104992.
- [27] Emuoyibofarhe, O.N., Adebayo, S., Ibitoye, A., Ayomide, M.O. and Taye, A., 2019. Predictive system for heart disease using a machine learning trained model. *International Journal of Computer (IJC)*, 34(1), pp.140-152.
- [28] Shah, D., Patel, S. and Bharti, S.K., 2020. Heart disease prediction using machine learning techniques. *SN Computer Science*, 1(6), pp.1-6.

- [29] Motarwar, P., Duraphe, A., Suganya, G., and Premalatha, M., 2020. Cognitive Approach for Heart Disease Prediction using Machine Learning. International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-5, doi: 10.1109/ic-ETITE47903.2020.242
- [30] J. Grus, Data Science from Scratch, O'Reilly Media, Inc., 2015.
- [31] Asadi, S.; Roshan, S.; Kattan, M.W. Random forest swarm optimization-based for heart diseases diagnosis. J. Biomed. Inform. 2021,115, 103690
- [32] Bansal, M.; Goyal, A.; Choudhary, A. A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. Decis. Anal. J. 2022, 3, 100071
- [33] Budholiya, K.; Shrivastava, S.; Sharma, V. An optimized XGBoost based diagnostic system for effective prediction of heart disease.J. King Saud-Univ.–Comput. Inf. Sci. 2020, 34, 4514–4523