JES, Vol. 54, No. 1, Jan 2026 DOI: 10.21608/JESAUN.2025.369613.1457 Part B: Electrical Engineering

Journal of
Engineering Sciences
Faculty of Engineering
Assiut University

JES

مجلة
العلوم الهندسية
كلية الهندسة
جامعة أسيوط

journal homepage: http://jesaun.journals.ekb.eg

Early Cardiovascular Disease Detection Using Predictive Machine-Learning Models: Evaluation and Insights

Received 19 March 2025; Revised 1 September 2025; Accepted 1 September 2025

Nahed Tawfik[1]
Nourhan Zayed [2*]

Keywords
Cardiovascular disease; machine learning; heart illnesses; prediction model

Abstract: As the foremost global cause of mortality, cardiovascular diseases (CVDs) underscore the critical importance of early diagnosis. Machine learning (ML) has emerged as a transformative tool in healthcare, enabling precise vulnerability assessment through data-driven pattern recognition. Implementing this technology in cardiology is critical for assessing risks, detecting nascent problems, and customizing treatment plans. This study intended to cultivate a predictive machine-learning paradigm for the nascent detection of heart disease. Eight classifiers, namely, k-nearest-neighbors, support-vector machine, logistic regression, random forest, decision tree, artificial neural networks, gradient boosting, and CN2 rule induction, were utilized to calibrate cardiovascular disease diagnosis predictions in the field of ML. One of the contributions of the proposed method is its enhanced nascent detection of cardiovascular disease, matched up with existing models using the same dataset in terms of both robustness and complexity. This study appraised assorted classifiers and their efficacies, offering helpful information for the development of trustworthy prediction models for coronary diseases. These models were scrutinized for heart illness using the UCI dataset and achieved improved accuracy and performance metrics. From the perspective of performance indicators (AUC, accuracy, F1-score, precision, & recall), the conclusions of this study demonstrate the efficacy of the CN2 rule induction and random forest models for detecting cardiovascular disorders.

1. Introduction

The heart portrays a decisive contribution in nurturing life by efficiently channeling oxygen-rich blood and modulating imperative hormones to keep blood pressure within a healthy range. Variations from the typical functioning of the heart can lead to heart-related ailments, ordinally termed to as cardiovascular diseases (CVDs) [1]. CVDs encompass assorted ailments that impact the heart and blood vessels. These ailments include coronary artery disease, peripheral arterial problems, rheumatic heart illnesses, congenital malformations, pulmonary embolisms, arrhythmias, and cardiomyopathies affecting the heart muscle [2]. Coronary heart disease, an important subtype of CVD, is responsible for 64% of substantial cases of cardiovascular illnesses [3]. Although it mainly affects males, females

---

[1] Researcher, Computers and Systems Department, Electronics Research Institute (ERI), Cairo, Egypt. nahedtawfik@eri.sci.eg
[2*] Assoc. Professor, Computers and Systems Department, Electronics Research Institute (ERI)/ Mechatronics Engineering, British University in Egypt, Cairo, Egypt. nourhan@eri.sci.eg , nourhan.zayed@bue.edu.eg

https://doi.org/10.21608/JESAUN.2025.369613.1457

are equally susceptible. Among all cardiovascular disorders, coronary artery disease is of particular concern because of its strong correlation with the global death toll. In line with WHO (the World Health Organization), cardiovascular illnesses have severe consequences and cause approximately eighteen million deaths worldwide annually. These alarming statistics emphasize the need for scientific investigations and advancements in medicine directed at treating and mitigating the impact of CVDs on a global scale [4].

Multiple vulnerability factors, such as hypertension, obesity, abnormal lipid profiles, diabetes, tobacco use, physical inactivity, alcohol intake, and cholesterol levels, contribute to the development of CVDs as illustrated in Figure 1. According to the WHO, CVDs will continue to be a notable vulnerability to human life and a leading cause of death for a long time, perhaps even after 2030.

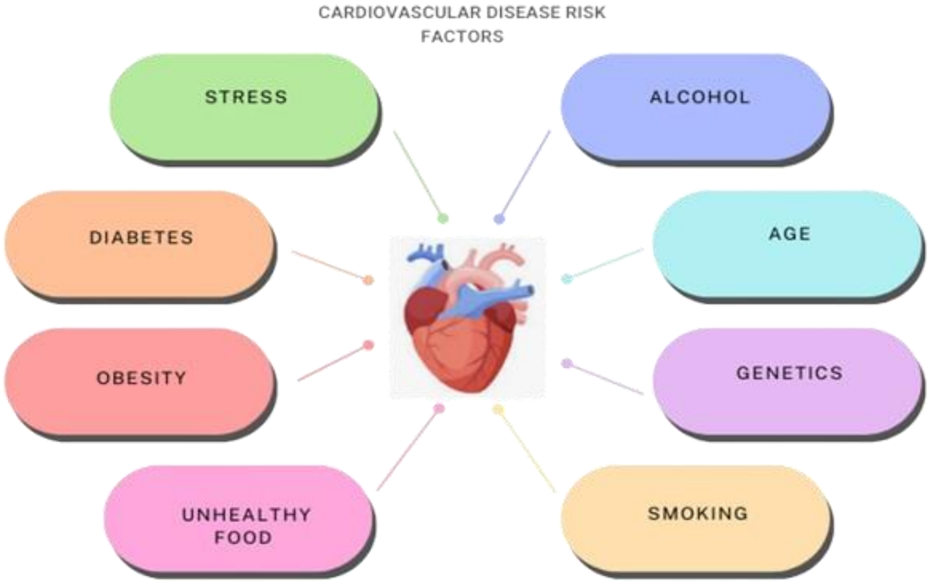


Fig. 1: Risk factors of CVDs.

The healthcare sector offers significant revolutionary potential through the application of machine-learning algorithms. The tremendous progress in this technology is due to its superior ability to process data analyzed side by side with humans. Consequently, the healthcare industry has witnessed an increase in several artificial intelligence (AI) applications that utilize the fast and precise capabilities of ML. These applications have led to the development of creative remedies to multifaceted healthcare problems. Although sundry machine-learning techniques have been employed to detect cardiovascular disorders, research gaps in present detection methods still require more resolution, and predictive models need enhancement.

The etiology of heart illness is an unresolved global issue characterized by high rates of disability, morbidity, and mortality. Consequently, the efficient and effective first-stage prediction of potential yield in individuals with CVD using AI is essential. AI can analyze extensive datasets, recognize signatures, and extract valuable insights, thereby presenting considerable promise for treating CVDs. AI exhibits competence in the nascent prediction of CVD incidence and progression, facilitating prompt preventive and intervention measures. Moreover, it improves venerability categorization by integrating patient-specific characteristics, facilitating customized monitoring, and focusing on interventions. These techniques optimize the diagnosis, treatment planning, and patient deliverables, heralding a new era in heart disease management.

The employment of observable data in clinical care, particularly for cardiovascular illnesses, presents considerable problems including data privacy and ethical implications. Protecting sensitive patient information while employing these data in machine-learning models is a substantial challenge. These models frequently require extensive personal health data for effective training, increasing the likelihood of unwanted access or misuse, and potentially resulting in considerable privacy violations and the erosion of patient trust. Ethical issues in utilizing patient data require comprehensive frameworks to regulate their collection, storage, and usage. Transparent communication on data management methods and compliance with regulations are decisive for alleviating concerns about data misuse and promoting collaboration among healthcare professionals. Moreover, given that data are frequently heterogeneous information sourced from diverse sources, guaranteeing data quality while upholding ethical standards introduces additional issues for real-world applications. Confronting these obstacles is essential for enhancing the application of observable clinical data to ameliorate CVD findings while protecting patient rights [5].

Machine learning has considerable transformational potential in healthcare. This extraordinary progress arises from its data processing capabilities, which notably exceed human capacities. Consequently, several AI applications in healthcare utilize the speed and accuracy of ML to foster novel countermeasures for multifaceted healthcare concerns. Diverse machine-learning methodologies have been employed to identify cardiovascular disorders. Numerous studies have explored multifaceted techniques, including neural networks (NNs) and ML, to enhance predictive accuracy. Although these studies provide meaningful insights, variations in the datasets, models, and yield underscore the intrinsic difficulty of prediction tasks. Nevertheless, despite these improvements, further studies are needed to enhance current models and improve the overall efficacy of CVD prediction. The ML models in this field demonstrate the importance of ongoing research to enhance the accuracy, reliability, and generalizability of predictive models. This will make clinical interventions and patient care more successful.

This study geared to analyze the influence of existing ML-based evaluation methods on early heart disease analysis. Subsequently, eight machine-learning-based predictive models were investigated and utilized to enhance the identification of cardiovascular disorders. The models included in this study were the k-nearest neighbor (k-NN), support vector machine (SVM), logistic regression (LR), decision tree (DT), random forest (RF), artificial NN (ANN), gradient boosting, and CN2 rule induction. By utilizing the UCI dataset, we deployed several preprocessing approaches, namely, feature scaling, normalization, and dimensionality reduction, to enhance the data for efficient ML-based evaluation. Finally, we appraised and weighted against the success rate of ML-based evaluation algorithms for assessing cardiac ailments.

The key advancements of this investigation are described as follows:

1. Proposing an accurate ML–based system for the early prediction of CVD using the publicly available UCI heart disease dataset.
2. Conducting a rigorous comparative analysis of eight ML classifiers on the same publicly available dataset to establish a consistent performance benchmark. This enables direct comparison with prior work while providing new insights into the relative strengths of these models, particularly the superior performance observed for Random Forest and CN2 Rule Induction.
3. Applying a comprehensive preprocessing pipeline—including handling missing values, normalization, and multi-method feature selection—to ensure a robust and reproducible evaluation framework.
4. Evaluating all models using a comprehensive set of five performance metrics (accuracy, precision, recall, F1-score, and AUC), offering a more complete view of model performance than many existing studies.

5. Presenting a detailed performance comparison with cutting-edge results reported in previous literature on the same dataset, demonstrating competitive or superior outcomes for the top-performing models identified in our analysis.

The rest of the paper is structured in the following manner: Section 2 introduces the relevant studies, Section 3 outlines the datasets, Section 4 details the proposed approach, Section 5 discusses the experimental observations and findings, and Section 6 concludes with future prospects.

2. Related Work

Recently, ML and deep-learning techniques have become increasingly popular disease prediction tools, particularly in the field of cardiovascular health. Substantial and continuous scientific efforts are currently being focused on predicting cardiac diseases. Machine-learning algorithms provide cost-effective and time-efficient methods for predicting CVDs. These strategies assume that computers can recognize data motifs and perform tasks either alone or with human assistance. Advancements in machine-learning algorithms have improved the accuracy and efficiency of human disease diagnosis [6]. Overall, these studies emphasize the ability of machine- and deep-learning models to promptly identify heart problems. In the following section, we offer a concise summary and explore important studies that adapted computational intelligence techniques to predict cardiac disease at a nascent stage.

An adaptive Hoeffding tree (AHT) algorithm for the detection of cardiac illness was demonstrated in [7]. The current analysis administered data from the Internet of Medical Things and machine-learning techniques. The study utilized a dataset of 3000 patient records, of which 33% were female. Sixteen important patient variables were selectively used for classification. The authors assessed the AHT algorithm using information from several medical institutions and juxtaposed it with other machine-learning techniques. The endpoints indicated a diagnostic accuracy of 95.67%. The authors in [8] exploited a deep-learning framework for identifying cardiac illnesses at an early stage using image categorization. Our examination appraised the effectiveness of a deep convolutional neural network (DCNN) on a publicly accessible dataset consisting of 1050 patients and 14 variables related to heart disease. The features extracted from the dataset were input into the DCNN to classify them into either normal or heart illness categories. The experimental endpoints demonstrated a validation accuracy of 91.7%.

In [9], the authors employed the UCI dataset to evaluate the execution of a machine-learning model for predicting cardiac disease. This method was appraised using aggregate-learning techniques. This study implemented sampling and feature selection techniques to determine the most relevant qualities. The authors implemented classifier models, such as DTs, LR, k-NN, SVM, naïve Bayes, and RF. The stacking classifiers achieved the accuracies of 99% with random oversampling, 93% with the synthetic minority oversampling technique (SMOTE), and 91% with the ADASYN sampling technique. A heart disease prognosis system that uses machine-learning classifiers, specifically LR and naïve Bayes, was introduced [10]. Classifiers were employed with datasets acquired from the UCI machine-learning repository. The dataset consisted of 13 characteristics and one target feature, whereas the Hungarian dataset contained 10 features. The precision of the model was increased using SMOTE. The LR accomplished an accuracy of 90% on the Cleveland dataset, whereas naïve Bayes achieved an accuracy of 85% on the Hungarian dataset. After implementing the SMOTE, the accuracy of the Cleveland dataset increased to 92%, whereas that of the Hungarian dataset reached 85%.

The authors of [11] addressed the premature detection of heart disease using machine-learning methods. The authors selected significant attributes from the dataset, considering complex patient attributes, such as allergy history, HDL and VDL levels, and irregular heart rate. The authors utilized several machine-learning techniques to enhance the accuracy of their predictions. The dataset obtained from the UCI ML repository was subjected to feature selection using regression and chi-square tests. The MLPPCA and chi-square approaches predicted heart disease patients with the accuracies of 98.08% and 98.25%, respectively. In another study [12], the authors utilized an ANN to predict heart illness. This algorithm combines the forward and backward approaches to provide a prediction framework. It exhibited an accuracy of 96%. The study in [13] exhibited hybrid classifiers that utilize an ensemble model and leverage a majority voting mechanism to enhance the prediction accuracy. Furthermore, the proposed work introduced a preprocessing and feature selection method that utilizes genetic algorithms. In addition to the use of 10-fold cross-validation to address potential overfitting issues. Experimental assessments were performed using the UCI dataset of patients with CVD, achieving a classification accuracy of 98.18%.

A model was cultivated in [14] to predict the survival of patients with heart failure using different machine-learning algorithms: DT regressor (DTR), DT, RF, XGBoost, and gradient boost. The RF obtained the highest accuracy of 97.78% out of all the approaches. In [15], a clustered genetic optimization method was exploited for the UCI dataset to predict heart disease. The data were first divided into segments by applying the k-means algorithm and then categorized using the proposed technique. The derivatives were benchmarked with those from basic classifiers, such as DTs, SVM, k-NN, and RF. The proposed model showed a precision of 94.56%.

Rashme et al. [16] proposed a CVD prediction model based on machine-learning algorithms to facilitate the first-stage prediction. The authors employed an RF for feature selection, whereas LR, k-NN, SVM, XGBoost, and DT were utilized for a performance comparison with the proposed model. According to the experimental findings, the XGBoost classifier showed a good performance, achieving a high accuracy of 75.10% in predicting CVD. The study in [17] emphasized the use of a machine-learning algorithm to identify cardiovascular illnesses. Four distinct machine-learning techniques were employed to train the model using patient datasets. The model was implemented using FLASK, a web-based framework that allows users to input 13 parameters for predictions. Upon a thorough evaluation of all four algorithms, the k-NN classifier was selected because of its notable prediction accuracy of 85.83%, which is acceptable for this model.

An ontology-based DT model was proposed in [18] to predict CVD. The authors employed machine-learning methods and ontologies in this approach. First, data preprocessing was performed to properly format the data for classification, and then the DT technique was utilized for classification and rule development. Ontology engineering was utilized to design and depict knowledge. The Semantic Web Rule Language was employed to build the DT rules in the ontology. The fusion of ML with semantic web technologies yielded a model with a classification accuracy of 75%. The authors of [19] disclosed an optimally configured long short-term memory technique to overcome the challenges encountered in CVD prediction. This technique leveraged the salp swarm algorithm (SSA) for feature elimination and the genetic algorithm (GA) to optimize the LSTM network configuration. The SSA eliminated irrelevant features and selected the optimal ones, whereas the GA fine-tuned the LSTM model by optimizing the network configuration and determining an appropriate time window size. This model demonstrated an accuracy of 97.11%.

Kurian et al. [20] performed a comparative assessment of sundry machine-learning classifiers geared at predicting heart disease, utilizing a limited set of attributes. Their evaluation encompassed k-NN, gradient boosting, SVM, and other techniques directed at discerning the attribute correlation and effectiveness. They tested the selected algorithms on the Cleveland heart disease dataset, wherein naïve Bayes achieved the highest performance with an accuracy of 94.20%. The study in [21] presented a model for predicting cardiovascular disorders based on grid-search cross-validation. This model used a dataset that included samples from both patients and healthy individuals. The model is based on combining feature selection techniques utilizing the Pearson correlation coefficient and permutation importance algorithms. Initially, the multilayer perceptron (MLP), LR, extra trees, DTs, RF, naïve Bayes, k-NN, and C5.0 were scrutinized using their default values. The RF model was determined to be the best performer based on an experimental study, with an accuracy of 94.63%. Subsequently, traditional machine-learning methods were implemented with the grid-search cross-validation model and the most optimal features, notably enhancing their accuracy. Consequently, the extra-trees classifier achieved an accuracy of 98.57%.

In [22], a predictive model for CVD was introduced. The model was based on data from patients with heart disease obtained through simple random sampling from two hospitals in Pakistan. Multifaceted machine-learning techniques were deployed for classification and prediction, including LR, RF, DTs, naïve Bayes, and SVM. The RF method attained the highest accuracy of 85.01%, sensitivity of 92.11%, and ROC curve of 87.73%. Additionally, it showed the lowest rates of misclassification errors and specificity for cardiovascular illnesses at 43.48% and 8.70%, respectively. Machine-learning algorithms are essential for precisely predicting cardiac diseases by revealing concealed signatures in data, providing predictions, and improving performance by analyzing past data. These methods facilitate a more accurate detection and prediction of cardiac diseases, whereas deep learning powered by ANNs is essential for handling intricate calculations on extensive datasets. Both methodologies are pivotal for discerning the major characteristics and motifs in structured and unstructured data, and for enhancing data analysis and processing efficiency.

3. Materials and Datasets

3.1. Dataset

In this current analysis, the UCI heart disease dataset was adopted [23]. Previously disparate datasets from Cleveland, Hungary, Switzerland, Long Beach, and Stalog were compiled in September 2021 to create the dataset. The final heart disease dataset contained 1025 samples, with equal numbers of cases in each class. Specifically, 499 samples belonged to the healthy class, and 526 samples belonged to the class with heart difficulties. Therefore, techniques to address the imbalanced classes need not be employed. A detailed description of the UCI heart disease dataset is showcased in Table 1.

3.2. Preprocessing

We implemented assorted preprocessing techniques to ensure the quality and suitability of the data for analysis. This section describes the specific methods employed for each preprocessing task, including handling missing values, addressing outliers, normalizing or standardizing numeric values, and encoding categorical variables. Data cleaning is pivotal in the preparation of data for analysis. This involves refining raw data, organizing it for visualization, and basing it on data prognostication. By eliminating erroneous, corrupted, inadequately structured, and redundant information from a dataset, data cleansing helps prevent inaccuracies, corruption, and incompleteness.

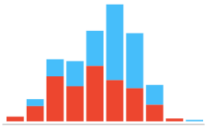
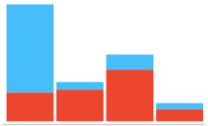
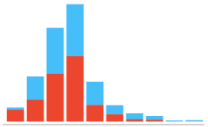
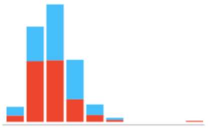
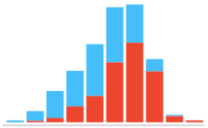
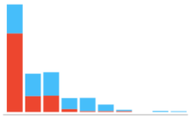
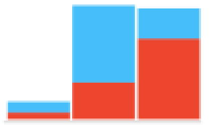
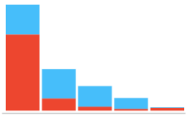
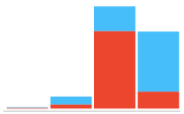
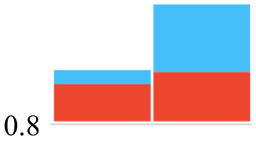
1. Imputation method: This approach involves substituting missing values (NaN) with either the average value (for continuous variables) or the most frequent value (for discrete variables). This technique maintains completeness and the integrity of the dataset.

2. Normalization: We implemented standard deviation scaling methods to normalize the values and bring them to a common scale. In these methods, the data can be made normal by dividing them by the standard deviation (σ) and subtracting the mean (μ) of each characteristic. This procedure guarantees that every feature has a standard deviation of one and an average value of zero. Consequently, this normalization technique facilitates a faster apex in data analysis or modeling processes. Table 2 presents the statistics of these features.
3. Feature selection: In this step, we targeted to identify the most informative clinical features that could be applied to determine heart disease prognosis. We leverage several feature selection techniques, such as information gain, chi-square, and correlation-based methods, to determine the relevant features, as listed in Table 3.

Table 1. Detailed description of the datasets

No.	Features	Description	Type
1	Age	Patient's age in years	Numeric
2	Sex	Male = 1; female = 0	Categorical
3	Chest pain (CP) type	Typical angina = 1, atypical angina = 2; nonanginal pain = 3, and asymptomatic = 4	Numeric
4	Resting blood pressure (trestbps)	The trestbps is assessed in mm Hg upon paying an entrance fee to the hospital.	Numeric
5	Serum cholesterol (chol)	The patient’s serum cholesterol level assessed in mg/dL	Numeric
6	Fasting blood sugar (fbs)	The fbs level if >120 mg/dL indicates 1 = true and 0 = no	Categorical
7	Resting electrocardiographic measures (restecg)	Rest: normal = 0; ST-T wave abnormality = 1; left ventricular hypertrophy as determined by the Estes criteria = 2	Numeric
8	Maximum heart rate (thalach)	The highest heart rate attained	Numeric
9	Exertional angina (exang)	The angina produced owing to exercise where yes = 1 and no = 0	Categorical
10	Exercise-induced ST depression (oldpeak)	The exercise-induced ST depression juxtaposed with the resting condition	Numeric
11	Slop indicating peak exercise (slop)	The slop ST segment where upsloping = 1, flat = 2, and downsloping = 3	Numeric
12	Major vessels number (ca)	The important vessels number, ranging from 0 to 3, colored via fluoroscopy	Numeric
13	Value of Thal	The value of thal is normal = 3, fixed defect = 6; reversible defect = 7	Numeric
14	Heart disease diagnosis (angiographic disease status)	The projected attribute is redemonstrated by a value of 0 if the narrowing diameter is less than 50%, and by a value of 1 if the narrowing diameter is greater than 50%.	Categorical

Table 2. Statistics of features

#	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.
1	age		54.43	58	56	0.17	29	77
2	cp		0.94	0	1	1.09	0	3
3	trestbps		131.61	120	130	0.13	94	200
4	chol		246	204	240	0.21	126	564
5	restecg		0.53	1	1	1.00	0	2
6	thalach		149.11	162	152	0.15	71	202
7	oldpeak		1.072	0.0	0.8	1.096	0.0	6.2
8	slope		1.39	1	1	0.45	0	2
9	ca		0.75	0	0	1.37	0	4
10	thal		2.32	2	2	0.27	0	3
11	sex			1		0.615		

#	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.
12	fbs			0		0.421		
13	exang			0		0.639		
14	target			1		0.693		

Table 3. Feature selection

Rank	Name	Info. gain	Gain ratio	Gini	χ^2	ReliefF
1	cp	0.208	0.121	0.137	217.824	0.095
2	thal	0.208	0.158	0.137	19.373	0.051
3	ca	0.169	0.104	0.112	228.148	0.120
4	oldpeak	0.159	0.080	0.103	193.305	0.061
5	exang	0.145	0.157	0.096	130.471	0.052

After obtaining the ranked lists from each feature selection method, we compared the top-ranked features across all methods to identify those consistently appearing in multiple rankings. Features selected by at least three different techniques were arranged for inclusion, as they demonstrated robustness across selection criteria. We then conducted iterative model training with cross-validation to evaluate the contribution of each candidate feature set to overall performance. The final set of features was chosen based on their ability to improve predictive accuracy while maintaining interpretability and alignment with established cardiovascular risk factors. In our study, this process led to the selection of cp, thal, ca, oldpeak, and exang as the most informative predictors for inclusion in the paradigms.

4. Methodology

In this research, multifaceted machine-learning techniques were leveraged to compare the endpoints with those of the proposed architecture, as shown in Figure 2. These methods are briefly introduced as they are widely recognized.

4.1. Decision Tree

A DT assigns an input feature to an internal or intermediary node. These intermediary nodes represent points at which a decision must be made based on several possibilities. Connections, re-exhibited as arcs, connect nodes, and originate from a node bearing a label as an input feature. Each arc corresponds to one of the potential merits of an output or earmarked feature. However, depending on the input feature, the arc leads to a subordinate decision node. The tree effectively classifies a dataset into a given class or specific probability distribution when the leaf nodes are marked with a class or probability distribution among the classes. Consequently, these nodes hold the final chosen class [24].

4.2. Random Forest

The Random Forest (RF) algorithm operates as an aggregate of decision trees, where each tree is trained on a bootstrap sample of the data. During prediction, RF aggregates outputs via majority voting, enhancing robustness against overfitting [25]. When operationalized to a sufficiently large dataset, RF has been acknowledged as one of the most accurate learning algorithms [25].

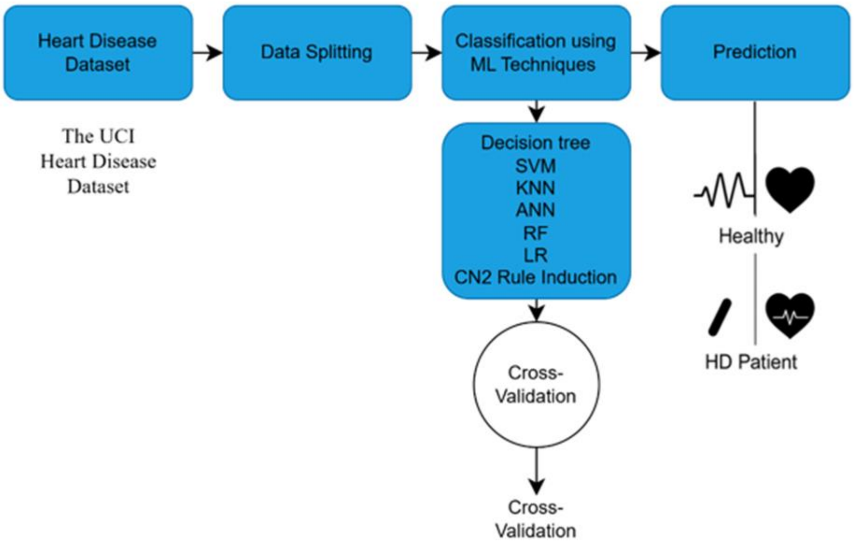


Fig. 2: Proposed cardiovascular illness prediction method

4.3. K-Nearest Neighbors (k-NN)

The k-NN method is a straightforward approach. In contrast to other machine-learning algorithms, k-NN **does not use training** data to create a model. Instead, it is a lazy learning strategy, because learning occurs only when the test data are examined. The vectors in the input data are p-dimensional. The vectors of the training examples and matching class labels are stored during the training phase. The distances between the new and stored vectors are computed during the classification phase. Subsequently, the k instances that are most similar to the updated input data are selected. The class that appears most frequently among the chosen vectors is adopted to categorize the new data. Although the distance can be computed using any metric, the Euclidean distance is frequently utilized [26].

4.4. Support Vector Machine (SVM)

An SVM constructs a model for classifying new instances into multiple categories. It can be categorized as a “nonprobabilistic classifier” as it is a spatial representation of instances [27]. The SVM model creates a hyperplane that separates the examples into different categories. It is a machine-learning approach that can be adapted to regression and classification problems. This simplifies the coordinates of the distinct observations, and the resulting SVM classifier acts as a decision boundary that effectively separates the two classes. Entering the parameters XTrain and YTrain into the fitting process allows the SVM model to be assessed for accuracy. Subsequently, the model performance is appraised using the score function with the XTest and YTest parameters. This approach computes the score of the SVM model and thus measures its performance [28].

4.5. Logistic Regression Model

Another ordinarily employed classification method is LR. In LR, the variables indicate a class, representing a categorical dependent variable. LR, a supervised learning classification approach, is adopted to calculate the probability of a target variable. The classification of the target variables is based on a specific categorical approach. The dependent variables in the LR analysis are dichotomous and classified into two distinct groups [29].

4.6. CN2 Rule Induction

CN2 is a learning approach specifically matured to function effectively when the available training data are insufficient. CN2 is derived from ID3 and the quasi-optimal algorithm (AQ), and it generates a set of rules similar to AQ and effectively addresses noisy data, such as ID3. The CN2 algorithm utilizes entropy as an assessment measure and has a beam width of 5. Rule filtering is employed with a minimum rule coverage of one and a maximum rule length of five [30].

4.7. Artificial Neural Networks

ANNs mimic the operation of biological NNs in living organisms, thereby enabling data processing through interconnected nodes [31]. ANNs can adapt and change their structures throughout the learning period. This examination utilized a hidden layer of one hundred neurons and employed a rectified linear unit activation function. The regularization parameter, represented by α, was assigned a value of 0.0001, and the maximum number of iterations was capped at 200 [32]. The optimal parameters for each classifier were established via grid search and cross-validation to guarantee reliable performance. The Random Forest model, exhibiting the highest accuracy, utilized 300 trees with an average depth of 20, resulting in nearly 60,000 parameters. The SVM utilized a radial basis function (RBF) kernel, with an ideal regularization parameter (C) of 1 and a gamma value of 0.01, employing about 150 support vectors. The KNN model exhibited optimal performance with k=5, encompassing about 22,000 parameters related to the training data. In the AdaBoost model, we employed 100 weak learners (decision stumps), yielding 100 parameters, whereas the Decision Tree model was trimmed to a depth of 10, comprising approximately 1,023 parameters. The optimal values were determined by comprehensive hyperparameter tweaking and validation, guaranteeing that each model attained the most favorable balance between accuracy and generalization.

5. Discussion and Results

5.1. Performance Measure

The evaluation of machine-learning performance was conducted through several essential metrics: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Using these measures, it is possible to calculate widely utilized performance parameters, such as accuracy, precision, recall, and F1-score. These metrics offer a thorough evaluation of how myriad machine-learning paradigms perform in making predictions [33].

- Accuracy

Accuracy is an ordinarily employed metric for assessing the overall quality of a classification model and represents the proportion of correctly classified cases.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

- Precision

Precision is assessed according to the total number of predictions achieved by the classifier. It can be defined as the ratio of TP to the total predictions (TP + FP) made by the model.

$$Percision = \frac{TP}{TP + FP}$$

- Recall

Recall denotes the efficiency of the classifier in accurately labeling TP cases, reflecting its capability to correctly identify positive instances.

$$Recall = \frac{TP}{TP + FN}$$

- F1-Score

F1-score is a widely implemented measure that integrates information from two common metrics: precision and recall. It is the harmonic mean of the precision and recall.

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Percision}$$

- AUC

The AUC is a comprehensive measure for assessing the efficiency of a model in classifying positive and negative classes. The AUC values range from 0 to 1. An AUC of 1 indicates a perfect model that perfectly distinguishes between classes, whereas an AUC of 0.5 signifies a performance equivalent to random guessing. A value below 0.5 indicates that the model performs worse than chance, thus highlighting the need for improvement. Higher AUC values indicate superior model performance [34]. Comprehensive Assessment of Machine-Learning architectures: Using the UCI heart disease dataset, this paper offers a thorough assessment of seven distinct machine-learning classifiers: k-NN, SVM, LR, DT, RF, ANNs, and gradient boosting. This comprehensive evaluation provides insightful information for future studies and valuable applications in healthcare by assisting in the identification of the most successful simulations for predicting cardiovascular illnesses.

This proposed work improved the performance of predictive paradigms by implementing diverse preprocessing methodologies, namely, feature scaling, normalization, and dimensionality reduction. This substantially enhances the precision and dependability of machine-learning models in the medical domain. Showcasing the potency of RF and CN2 Rule Induction paradigms: The conclusions of this study highlight the excellent performance and accuracy of the RF and CN2 rule induction representations in predicting cardiovascular illnesses. These frameworks are excellent candidates for practical implementation in early detection systems for cardiovascular health because they demonstrate improved performance metrics, namely, the AUC, F1-score, precision, and recall.

The authors investigated the traditional machine-learning approaches outlined in Section 4. In our study, commonly adopted evaluation metrics, namely, the AUC, accuracy, F1-score, precision, and recall, were leveraged for performance evaluation. The results are listed in Table 4, which displays the average of the 10-fold validation using the optimal configuration. The results indicated that RF and CN2 Rule Induction performed well, with the highest accuracy of 100%. The ANN method exhibited a comparable performance (97.5%), followed by Gradient boosting (97.2%). In this current analysis, the CN2 Rule Induction and RF models achieved 100% accuracy due to several factors, including the utilization of a dataset that contains highly discriminative features—such as cp, ca, and thal, which consistently ranked at the top across multiple feature selection methods, enabling strong categorization between classes. The preprocessing steps, including handling missing values, normalization, and rigorous feature selection, likely enhanced the quality of the input data. Moreover, the ability of CN2 Rule Induction and RF models to capture complex, non-linear relationships and feature interactions may have further boosted predictive performance. Despite these promising findings, a potential for applying dataset-specific motifs to influence performance, and therefore emphasize the need for external validation on larger and more diverse datasets to confirm the paradigm' generalizability.

The 10-fold cross-validation is an effective method for assessing model performance and mitigating the negative influence of overfitting. This method involves partitioning the dataset into 10 equalsized subgroups called “folds.” The model was trained nine times, with the remaining folds utilized for testing, guaranteeing that each fold served as a test set once. The procedure was executed 10 times, and the outcomes were averaged to yield a comprehensive performance assessment. Utilizing all the data points for training and testing across several iterations, 10-fold cross-validation offers a more dependable evaluation of the generalization capability of the model, juxtaposed with a singular train-test division. This method is particularly advantageous for smaller datasets because it optimizes the utilization of existing data. It is widely utilized in classification tasks to evaluate measures, such as accuracy, precision, recall, and F1-score, which facilitate the evaluation of the prediction efficacy of the model. The test set is a distinct segment of the dataset that remains unused during the training and validation stages. It functions as a conclusive evaluation instrument for examining the success rate of a model using novel data.

In this investigation, k-NN exhibited the lowest performance, achieving an accuracy of 84%, which was 7.7% lower than that of LR. Figure 3 displays the findings obtained using a stratified 10-fold cross-validation. Each result represents the average accuracy obtained during the 10-fold cross-validation. Although the proposed method yielded promising derivatives, a key limitation was the relatively small sample size. Therefore, this approach requires additional validation by using larger and more diverse datasets. Furthermore, collaboration with clinicians is essential to refine and optimize our frameworks for practical applications in real-world settings.

Table 4. Evaluation results of the different models

Model	AUC	Accuracy	F1-Score	Precision	Recall
k-NN	0.887	0.753	0.753	0.756	0.753
LR	0.917	0.842	0.841	0.845	0.842
SVM	0.967	0.923	0.923	0.924	0.923
ANN	0.996	0.975	0.975	0.975	0.975
Gradient Boosting	0.995	0.972	0.972	0.972	0.972
CN2 Rule Induction	1.000	1.000	1.000	1.000	1.000
RF	1.000	1.000	1.000	1.000	1.000

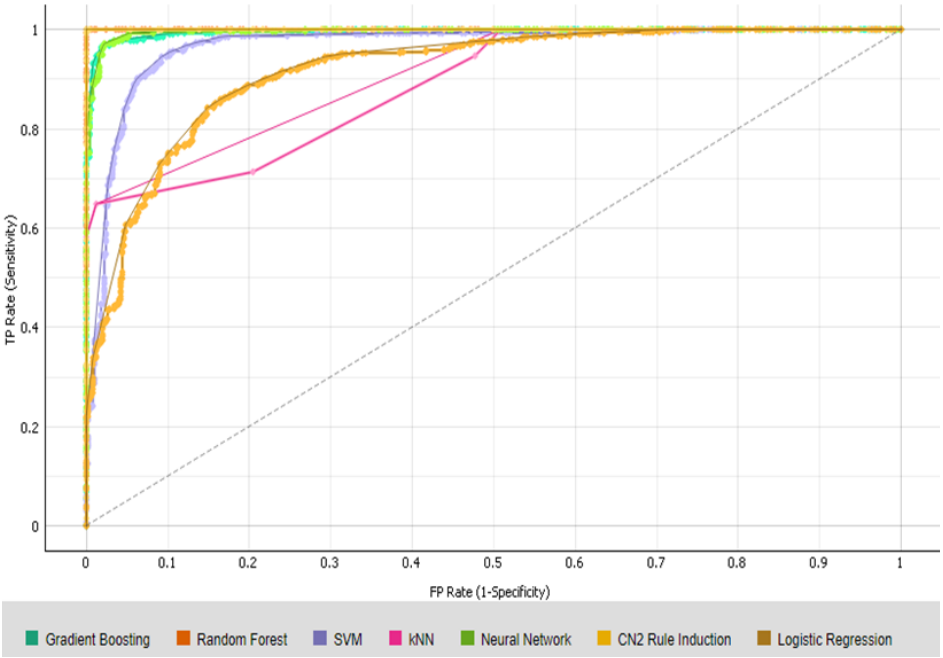


Fig. 3: Outcomes obtained using stratified 10-fold cross-validation

5.2. Comparison with Existing Methods

Table 5 presents a comparison of the existing methods with the proposed method. The proposed method outperformed all the existing methods.

Table 5. Comparison of the existing and proposed methods

Reference	Methods	Datasets	Results
[10]	LR Naïve Bayes and SMOTE	Cleveland dataset, 10 features of the Hungarian dataset from the UCI repository	Cleveland dataset: LR: 92%, NB: 90% UCI dataset NB: 85%, LR: 81%
[14]	DTR, XGBoost (XGB), DT, GB, and RF	Kaggle (299) instances	DTR: 85.24%, XGB: 94.44% DT: 96.77%, GB: 96.77% RF: 97.78%
[35]	SVM, DT, k-NN, LR, Naïve Bayes, and RF	Kaggle (1025) dataset	SVM: 68.83%, k-NN: 87.01%, LR: 75.97% DT: 98.05% , NB: 75% RF: 95.13%
[29]	Linear Regression	303 records of UCI repository	LR: 87.10%
[36]	ANN, SVM, RF, DT, and k-NN	Cleveland dataset	ANN: 84.4%, RF: 81.67% SVM: 83.33%, DT: 73.33% k-NN: 61.67%
[37]	Classification and regression tree (CART)	IEEE data port and 745 health records from heart disease dataset	CART: 87%
[38]	k-NN, LR, and RF	UCI dataset	k-NN: 73%, LR: 88% RF: 88%
[39]	SVM, AdaBoost, RF, MLP, GB, k-NN, DT, Gaussian Naïve Bayes (NB)	Cleveland dataset	AB: 87%, SVM: 92% RF: 91%, MLP: 97 GB: 98%, k-NN: 89% DT: 100%, Gaussian NB: 79%
[40]	LR, k-NN, RF, ET, XGBoost, composite classifier	Stat log, Hungary, Cleveland	k-NN: 81%, RF: 90% ET: 90%, XGB: 91% Composite: 89%
[3]	k-NN, RF, LR, Gradient Boost, SVM, CNN, and XGBoost	The Cardiovascular heart disease dataset and the Cleveland Dataset	k-NN: 96.50% and 91.80% RF: 98.60% and 91.09% LR: 95.50% and 88.52% GB: 98.00% and 86.89% SVM: 95.50% and 78.69% CNN: 97.50% and 86.89% XGBoost: 98.50% and 86.89%
[41]	GAN-1D-CNN GAN-Bi-LSTM	Cleveland	99.10% 99.30%
Proposed method	k-NN, LR, SVM, ANN, Gradient Boosting, DT, CN2 Rule Induction, and RF	The Heart Disease UCI dataset	k-NN: 75.3%, LR: 84.2% SVM: 92.3%, **ANN: 97.5%** GB: 97.2% CN2 Rule Induction: 100% RF: 100%

Our research underscores a significant trade-off between model complexity and interpretability in early identification of cardiovascular disease (CVD). Our analysis revealed that sophisticated models, including gradient boosting and random forests, typically demonstrated enhanced predictive accuracy by effectively capturing complicated, non-linear interrelations among diverse risk factors. Nonetheless, this improved throughput entails diminished interpretability, rendering these models as a "black box" in which the rationale for their predictions remains obscure. Conversely, simpler models such as logistic regression provide a more straightforward and comprehensible methodology, aligning more closely with conventional clinical risk-scoring methods. The distinct, measurable contribution of each aspect renders these models more appropriate for therapeutic environments where trust and transparent communication with patients are essential. The conflict between performance and transparency is a significant factor in clinical adoption. Consequently, we advocate for future research to investigate methodologies for explainable AI (XAI) to furnish transparent insights for physicians and address the opaque characteristics of certain models. This strategy would enhance the clinical applicability of our results by facilitating the implementation of robust, intricate models in an effective and reliable manner.

6. Conclusions

The proposed work refined a healthcare application precisely planned for detecting heart illnesses in patients, assimilating those exhibiting symptoms. By leveraging the RF algorithm, our approach yields improved accuracy weighted against with existing methods. Moreover, our solution is highly cost-effective in terms of development. Furthermore, the research endpoints obtained in our study can provide valuable data for subsequent analyses, facilitating the development of enhanced diagnostic techniques for heart disease detection, along with their associated findings and interpretations. This study will provide valuable support for evaluating third-party individuals using the proposed robust methodology. The enhanced combination of feature selection, paradigm selection, and accuracy highlights the unique contributions of our research to predicting heart disease. Our study enhances this discipline by offering a more accurate and resilient forecasting framework that exceeds the accuracy of previous studies. Overall, our study contributes both theoretically and practically to the healthcare field.

Future research directions on this topic may expand the scope by assimilating larger medical imaging datasets. Employing such data could improve image-based predictions of heart illnesses and potentially yield more accurate and dependable diagnostic tools in cardiovascular health. Furthermore, exploring aggregate paradigms that integrate the advantages of several algorithms may yield notable opportunities to enhance the accuracy of cardiac disease diagnosis. These methodologies emphasize the intricate domain of cardiac disease prediction research and highlight the necessity of ongoing refinement and innovation in this critical area. Future studies should prioritize framework enhancement and the augmentation of datasets. In addition, future studies can be directed toward integrating deep-learning techniques with traditional machine-learning methods to improve competency and interpretability. Our research offers a fundamental assessment of learning algorithms for the first-stage diagnosis of cardiovascular illness utilizing a publicly accessible dataset. Although the predicted efficacy of these models is encouraging, we acknowledge a significant constraint in their present scope. The models created in this study necessitate additional clinical validation and real-world testing prior to their potential implementation in a healthcare environment. Our findings, derived from retrospective data, require validation through prospective clinical trials with varied patient populations to appropriately evaluate the models' generalizability, robustness, and clinical value. Subsequent research must emphasize these essential measures to guarantee that the models can efficiently and securely transition from research settings to practical applications, thereby offering a definitive pathway for their advancement and deployment in a clinical context. Also, metrics such as calibration, decision curves, and cost-benefit analyses are vital for clinical decision support systems. Furthermore, future studies should also emphasize the need for clinical validation in more extensive and diverse patient populations. We will also discuss the importance of collaborating with clinicians to ensure the practical applicability of our models.

References

[1] Saghiv, M.S.; Sagiv, M.S. Cardiovascular Function. Basic Exercise Physiology 2020, 285–369, doi:10.1007/978-3-030-48806-2_6.
[2] Ho, K.-J. Cardiovascular Diseases. Nutritional Aspects of Aging 2018, 75–100, doi:10.1201/9781351075145-3.
[3] Ogunpola, A.; Saeed, F.; Basurra, S.; Albarrak, A.M.; Qasem, S.N. Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases. Diagnostics 2024, 14, doi:10.3390/diagnostics14020144.
[4] Cardiovascular Diseases Available online: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1 (accessed on 7 May 2024).
[5] Liu, F.; Panagiotakos, D. Real-World Data: A Brief Review of the Methods, Applications, Challenges and Opportunities. BMC Med Res Methodol 2022, 22, 287, doi:10.1186/s12874-022-01768-6.
[6] Krittanawong, C.; Virk, H.U.H.; Bangalore, S.; Wang, Z.; Johnson, K.W.; Pinotti, R.; Zhang, H.J.; Kaplin, S.; Narasimhan, B.; Kitai, T.; et al. Machine Learning Prediction in Cardiovascular Diseases: A Meta-Analysis. Scientific Reports 2020 10:1 2020, 10, 1–11, doi:10.1038/s41598-020-72685-1.
[7] Elbasi, E.; Zreikat, A.I. Heart Disease Classification for Early Diagnosis Based on Adaptive Hoeffding Tree Algorithm in IoMT Data. International Arab Journal of Information Technology 2023, 20, 38–48, doi:10.34028/IAJIT/20/1/5.
[8] Arooj, S.; Rehman, S. ur; Imran, A.; Almuhaimeed, A.; Alzahrani, A.K.; Alzahrani, A. A Deep Convolutional Neural Network for the Early Detection of Heart Disease. Biomedicines 2022, Vol. 10, Page 2796 2022, 10, 2796, doi:10.3390/BIOMEDICINES10112796.
[9] Rohit Chowdary, K.; Bhargav, P.; Nikhil, N.; Varun, K.; Jayanthi, D. Early Heart Disease Prediction Using Ensemble Learning Techniques. J Phys Conf Ser 2022, 2325, 012051, doi:10.1088/1742-6596/2325/1/012051.
[10] Bouqentar, M.A.; Terrada, O.; Lamrani, D.; Ouhmida, A.; Cherradi, B.; Raihani, A. Primary Prediction of Heart Disease Using Machine Learning Algorithms and SMOTE. In Proceedings of the 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET); IEEE, May 18 2023; pp. 1–7.
[11] Aggarwal, R.; Kumar, S. MLPPCA: Heart Disease Detection Using Machine Learning. In Proceedings of the 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC); IEEE, November 25 2022; pp. 457–461.
[12] J., A.; M., P. Forward and Backward Propagation Network Based Heart Disease Prediction. IJIRMPS - International Journal of Innovative Research in Engineering & Multidisciplinary Physical Sciences 2022, 10, doi:10.37082/IJIRMPS.
[13] Ashri, S.E.A.; El-Gayar, M.M.; El-Daydamony, E.M. HDPF: Heart Disease Prediction Framework Based on Hybrid Classifiers and Genetic Algorithm. IEEE Access 2021, 9, 146797–146809, doi:10.1109/ACCESS.2021.3122789.
[14] Ali, M.M.; Al-Doori, V.S.; Mirzah, N.; Afsari Hemu, A.; Mahmud, I.; Azam, S.; Al-Tabatabaie F, F.; Ahmed, K.; Bui, F.M.; Moni, M.A. A Machine Learning Approach for Risk Factors Analysis and Survival Prediction of Heart Failure Patients. Healthcare Analytics 2023, 3, 100182, doi:10.1016/j.health.2023.100182.
[15] Vijaya, J. Heart Disease Prediction Using Clustered Genetic Optimization Algorithm. Proceedings of the International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics, ICIITCEE 2023 2023, 1072–1077, doi:10.1109/IITCEE57236.2023.10091050.

[16] Rashme, T.Y.; Islam, L.; Jahan, S.; Prova, A.A. Early Prediction of Cardiovascular Diseases Using Feature Selection and Machine Learning Techniques. Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021 2021, 1554–1559, doi:10.1109/ICCES51350.2021.9489057.

[17] Bhardwaj, S. Employment of The Selective Machine Learning Algorithm for The Early and Effective Detection and Diagnosis of Cardiovascular Disease. International Journal of Research in Medical Sciences and Technology 2021, 11, 239–246, doi:10.37648/IJRMST.V11I01.024.

[18] El Massari, H.; Gherabi, N.; Mhammedi, S.; Sabouri, Z.; Ghandi, H. Ontology-Based Decision Tree Model for Prediction of Cardiovascular Disease. Indian Journal of Computer Science and Engineering 2022, 13, 851–859, doi:10.21817/INDJCSE/2022/V13I3/221303143.

[19] Revathi, T.K.; Balasubramaniam, S.; Sureshkumar, V.; Dhanasekaran, S. An Improved Long Short-Term Memory Algorithm for Cardiovascular Disease Prediction. Diagnostics 2024, 14, doi:10.3390/diagnostics14030239.

[20] Kurian, N.S.; Renji, K.S.; Sajithra, S.; Yuvasree, R.; Jenefer, F.A.; Swetha, G. Prediction of Risk in Cardiovascular Disease Using Machine Learning Algorithms. International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2022 - Proceedings 2022, 162–167, doi:10.1109/ICSCDS53736.2022.9760879.

[21] Moushi, O.M.; Ara, N.; Helaluddin, M.; Mondal, H.S. Enhancing the Accuracy and Explainability of Heart Disease Prediction Models through Interpretable Machine Learning Techniques. 2023 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2023 - Proceedings 2023, 1–6, doi:10.1109/ICICT4SD59951.2023.10303572.

[22] Khan, A.; Qureshi, M.; Daniyal, M.; Tawiah, K. A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction. Health Soc Care Community 2023, 2023, 1–10, doi:10.1155/2023/1406060.

[23] Heart Failure Prediction Dataset Available online: https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction (accessed on 10 May 2024).

[24] Song, Y.Y.; Lu, Y. Decision Tree Methods: Applications for Classification and Prediction. Shanghai Arch Psychiatry 2015, 27, 130–135, doi:10.11919/j.issn.1002-0829.215044.

[25] More, A.S.; Rana, D.P. Review of Random Forest Classification Techniques to Resolve Data Imbalance. In Proceedings of the 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM); IEEE, October 2017; pp. 72–78.

[26] Cunningham, P.; Delany, S.J. K-Nearest Neighbour Classifiers - A Tutorial. ACM Comput Surv 2022, 54, 1–25, doi:10.1145/3459665.

[27] Sherif, F.F.; Ahmed, K.S. A Machine Learning Approach for Stroke Differential Diagnosis by Blood Biomarkers., doi:10.12720/jait.15.1.1-9.

[28] Awad, M.; Khanna, R. Support Vector Machines for Classification. Efficient Learning Machines 2015, 39–66, doi:10.1007/978-1-4302-5990-9_3.

[29] G, A.; Ganesh, B.; Ganesh, A.; Srinivas, C.; Dhanraj; Mensinkal, K. Logistic Regression Technique for Prediction of Cardiovascular Disease. Global Transitions Proceedings 2022, 3, 127–130, doi:10.1016/J.GLTP.2022.04.008.

[30] Asiri, A.A.; Khan, B.; Muhammad, F.; Rahman, S.U.; Alshamrani, H.A.; Alshamrani, K.A.; Irfan, M.; Alqhtani, F.F. Machine Learning-Based Models for Magnetic Resonance Imaging (MRI)-Based Brain Tumor Classification. Intelligent Automation and Soft Computing 2023, 36, 299–312, doi:10.32604/iasc.2023.032426.

[31] Tawfik, N.; El Din, M.F.; Dessouky, M.I.; Abd Ei-Samie, F.E. Processing of Corneal Images with A Cepstral Approach. 23rd International Conference on Computer Theory and Applications, ICCTA 2013 - Proceedings 2013, 49–54, doi:10.1109/ICCTA32607.2013.9529605.

[32] Khan, Z.; Mishra, D.K.; Sharma, V.; Sharma, A. Empirical Study of Various Classification Techniques for Heart Disease Prediction. In Proceedings of the 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA); IEEE, October 30 2020; pp. 57–62.

[33] Foody, G.M. Challenges in the Real World Use of Classification Accuracy Metrics: From Recall and

Precision to the Matthews Correlation Coefficient. PLoS One 2023, 18, e0291908, doi:10.1371/journal.pone.0291908.

[34] Jin Huang; Ling, C.X. Using AUC and Accuracy in Evaluating Learning Algorithms. IEEE Trans Knowl Data Eng 2005, 17, 299–310, doi:10.1109/TKDE.2005.50.

[35] Tushar, A.M.; Wazed, A.; Shawon, E.; Rahman, M.; Hossen, M.I.; Jesmeen, M.Z.H. A Review of Commonly Used Machine Learning Classifiers in Heart Disease Prediction. 2022 IEEE 10th Conference on Systems, Process and Control, ICSPC 2022 - Proceedings 2022, 319–323, doi:10.1109/ICSPC55597.2022.10001742.

[36] Gangadhar, M.S.; Sai, K.V.S.; Kumar, S.H.S.; Kumar, K.A.; Kavitha, M.; Aravinth, S.S. Machine Learning and Deep Learning Techniques on Accurate Risk Prediction of Coronary Heart Disease. Proceedings - 7th International Conference on Computing Methodologies and Communication, ICCMC 2023 2023, 227–232, doi:10.1109/ICCMC56507.2023.10083756.

[37] Ozcan, M.; Peker, S. A Classification and Regression Tree Algorithm for Heart Disease Modeling and Prediction. Healthcare Analytics 2023, 3, 100130, doi:10.1016/j.health.2022.100130.

[38] Patidar, S.; Kumar, D.; Rukwal, D. Comparative Analysis of Machine Learning Algorithms for Heart Disease Prediction. Advances in Transdisciplinary Engineering 2022, 27, 64–69, doi:10.3233/ATDE220723.

[39] Venkatesh, V.; Rai, P.; Reddy, K.A.; Praba, S.; Anushiadevi, R. An Intelligent Framework for Heart Disease Prediction Deep Learning-Based Ensemble Method. In Proceedings of the 2022 International Conference on Computer, Power and Communications (ICCPC); IEEE, December 14 2022; pp. 274–280.

[40] S, N.; K, V.; B, I.; Kalshetty, Jagadevi.N. Heart Disease Prediction Using Artificial Intelligence Ensemble Network. In Proceedings of the 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon); IEEE, October 16 2022; pp. 1–6.

[41] Sarra, R.R.; Dinar, A.M.; Mohammed, M.A.; Ghani, M.K.A.; Albahar, M.A. A Robust Framework for Data Generative and Heart Disease Prediction Based on Efficient Deep Learning Models. Diagnostics 2022, 12, 2899, doi:10.3390/diagnostics12122899.