Digital Health Trends Journal. 2025;1(1):63-69.
doi:10.34172/dhtj.07
https://dhtjournal.ae

Digital Health Trends Journal

Original Article

Comparative Study of Classification-Based Data Mining Algorithms for Predicting Cardiovascular Diseases

Sepideh Seyedi-Sahebari[1], Ali Farzaneh[2*]

[1]Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran
[2]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

***Corresponding author:** Ali Farzaneh, Email: farzanehali78@gmail.com

Abstract

Background: Cardiovascular diseases (CVDs) remain a leading cause of mortality, demanding timely and accurate diagnosis. Traditional clinical assessments are often prone to errors, highlighting the need for predictive models that leverage large-scale clinical data, including data mining techniques that can extract data from complex medical datasets. This study comparatively analyzed classification-based data mining algorithms for predicting CVDs and evaluating their performance across multiple metrics to identify the most effective predictive model for clinical applications.

Methods: The UCI Heart Disease dataset (270 records with 14 clinical attributes) was used. Data preprocessing involved cleaning, normalization, discretization, and partitioning into training (70%) and testing (30%) sets. NB, ANN, kNN, SVM, and CART algorithms were implemented using Orange. Model performance was evaluated by accuracy, sensitivity, specificity, precision, recall, F-measure, and AUC using hold-out validation and 5-fold cross-validation. Feature importance and decision rules were extracted from tree-based models for interpretability.

Results: SVM and NB achieved the highest overall predictive performance (accuracy: 84.44%, sensitivity: 86.00%, specificity: 82.50%, AUC: 0.9136; accuracy: 84.07%, AUC: 0.9133). ANN and KNN demonstrated moderate predictive ability, while CART (accuracy: 78.52%) provided interpretable decision rules. Decision tree (DT) analysis identified thalassemia status, chest pain type, and number of major vessels colored as the most influential attributes. Several clinically interpretable rules were extracted, offering potential guidance for risk assessment. Statistical comparisons indicated no significant difference between SVM and NB performance, suggesting both models provide reliable predictions.

Conclusion: SVM and NB offer robust predictive capabilities for CVD, outperforming traditional statistical approaches. DT models provide additional interpretability, facilitating clinical understanding and application. These findings underscore the importance of evaluating multiple predictive models in context-specific datasets to identify optimal approaches for risk assessment, resource allocation, and quality of care improvement, thereby enhancing early detection and supporting evidence-based CVD management.

Keywords: Comparison, Classification, Data mining, Algorithms, Prediction, Cardiovascular diseases

Received: July 4, 2025, **Revised:** August 20, 2025, **Accepted:** September 15, 2025, **ePublished:** September 27, 2025

Background

Providing high-quality care while maintaining cost-effectiveness is one of the major challenges facing healthcare organizations, including hospitals and medical centers. High-quality care relies on accurate diagnosis and effective treatment, as inappropriate clinical decisions may lead to catastrophic and unacceptable consequences.[1] Cardiovascular diseases (CVDs) represent a critical domain where timely and precise diagnosis is essential, given their high mortality rates worldwide. The term CVD encompasses a wide spectrum of disorders that affect the heart and blood vessels, influencing the circulation and pumping function of the heart.[2-5]

According to the World Health Organization, in 2012, nearly 17 million deaths worldwide were attributed to heart attack and stroke, positioning CVDs as the leading cause of adult mortality.[6] Reports further indicate that CVDs account for almost half of all deaths in developed nations and remain a major cause of mortality in developing countries, including Iran. According to official statistics from the Iranian Ministry of Health, over 40% of deaths in the country are due to CVDs.[2,5,7] The diagnostic process for CVDs is highly complex and requires exceptional accuracy; however, traditional clinical assessments, typically based on physicians' expertise and experience, are prone to errors.[8] Despite technological advances in recent decades that have facilitated diagnosis, controlling CVDs continues to be a primary priority for healthcare systems worldwide.[9]

Data mining has emerged as a promising approach to extract valuable insights from large-scale medical databases that often remain underutilized. The applications of this process in healthcare include supporting decision-making in treatment selection, clinical guideline development, resource allocation, and evaluating therapeutic outcomes [10]. Broadly, data mining techniques can be categorized into descriptive and predictive types. While descriptive mining identifies patterns and associations from past activities,

predictive mining utilizes historical data to forecast future outcomes.[11,12] The efficiency of these algorithms depends on several factors, such as the availability of required variables, the size and quality of datasets, and the accuracy of data.

Although conventional diagnostic markers for CVDs include symptoms, electrocardiograms, and enzymatic indicators, incorporating additional patient attributes (e.g., age, gender, family history, and heart rate) can substantially enhance diagnostic accuracy. Numerous studies have applied data mining techniques to cardiovascular datasets, employing a number of models, such as decision trees (DTs), association rule mining, Bayesian networks, and artificial neural networks (ANNs). For instance, DT models have been used to identify risk factor relationships,[13] while DTs and association rules have evaluated critical risk predictors.[14, 15] Moreover, Bayesian networks have been applied to assess survival predictors in cardiac patients.[16]

However, most of these studies have focused on applying a single data mining algorithm or utilized limited evaluation criteria, making it difficult to determine the most reliable model for CVD prediction. Furthermore, there is limited evidence comparing the performance of multiple classification algorithms on large, real-world cardiovascular datasets using comprehensive performance metrics. This gap highlights the need for systematic evaluation to identify the most effective predictive models for clinical use.

Accurate prediction not only plays a vital role in guiding treatment strategies but also supports optimal resource allocation in hospitals, ultimately improving the quality of care delivery. Therefore, the present study aims to perform a comparative analysis of classification-based data mining algorithms for predicting CVDs by evaluating their performance across key metrics such as accuracy, sensitivity, specificity, precision, recall, and F-measure in order to identify the most effective model for CVD prediction.

Methods

Dataset

The method adopted in this study for predicting CVDs consists of several steps: dataset selection, data preparation and normalization, data mining and classification, model evaluation, and knowledge extraction.

The publicly available heart disease dataset from the University of California, Irvine (UCI) Machine Learning (ML) Repository was employed for this purpose. This dataset contains multiple clinical attributes that are highly relevant for identifying patterns associated with cardiovascular disorders. After handling missing values and removing outliers, a total of 270 complete records were retained for analysis. The dataset includes 14 attributes, where the 14th attribute (heartdisease) indicates the presence (positive cases) or absence (negative cases) of CVD, which was used as the target class.

The dataset is balanced, consisting of 120 negative cases (absence of the disease) and 150 positive cases (presence of the disease), ensuring that model training is not biased toward any class. The attributes contain both numerical (continuous) and categorical (discrete) variables. Five attributes were continuous, and the remaining attributes were discrete. To enhance classification performance and simplify analysis, all continuous variables were discretized using equal-width binning, converting them into discrete intervals. The discretization thresholds were determined based on standard clinical reference ranges and data distribution percentiles. For instance, age was grouped into four categories (29–39, 40–49, 50–59, and 60–77 years). In addition, resting blood pressure was classified as normal (< 120 mm Hg), elevated (120–129 mm Hg), stage 1 hypertension (130–139 mm Hg), and stage 2 hypertension (≥ 140 mm Hg). Further, cholesterol levels were classified according to established clinical cutoffs (< 200 mg/dL, 200–239 mg/dL, and ≥ 240 mg/dL). Other continuous variables, such as maximum heart rate and ST depression (old peak), were discretized using quartile-based thresholds derived from the study cohort.

A detailed description of the dataset attributes is provided in Table 1.

Data Preprocessing

The preprocessing phase involved three main steps:

- *Data cleaning:* Missing and inconsistent records were removed.
- *Normalization and discretization:* Continuous features were transformed into discrete intervals using equal-width binning to enhance model interpretability and improve classification performance.

Table 1. Description of the Attributes in the UCI Heart Disease Dataset

No.	Attribute (Dataset name)	Description
1	age	Age of the patient (in years)
2	sex	Gender of the patient
3	cp	Chest pain type
4	restbp	Resting blood pressure
5	chol	Serum cholesterol level
6	fbs	Fasting blood sugar
7	restecg	Resting electrocardiographic results
8	maxhrate	Maximum heart rate achieved
9	exang	Exercise-induced angina
10	oldpeak	ST depression induced by exercise relative to rest
11	slope	Slope of the peak exercise ST segment
12	ncolored	Number of major vessels (0–3) colored by fluoroscopy
13	thal	Thalassemia status: 3 = normal; 6 = fixed defect; 7 = reversible defect
14	heartdisease (class)	Diagnosis of heart disease (presence or absence)

Note. UCI: The University of California, Irvine.

- *Partitioning*: The dataset was divided into training and testing subsets with a 70:30 ratio, resulting in 189 samples for training and 81 samples for testing.

Classification Algorithms

The experiments were conducted using Orange (version 2.7.8), an open-source data mining and ML software that provides a wide range of tools for classification, regression, and clustering. The following classification-based data mining algorithms were implemented and compared, with their main parameters explicitly set for reproducibility:

1. Naïve Bayes (NB): Default prior probabilities and Gaussian likelihood for continuous features.
2. ANN: One hidden layer with 10 neurons, learning rate = 0.01, and activation function = sigmoid.
3. k-Nearest neighbors (kNNs): k = 5, Euclidean distance metric.
4. Support vector machine (SVM): Radial basis function kernel, C = 1, and gamma = 0.1.
5. Classification Tree (CART): Gini impurity criterion, ID3 Algorithm, and maximum depth = 5.

The primary objective of classification was to predict whether a patient has CVD (positive class) or not (negative class).

Model Evaluation

Two evaluation strategies were applied to ensure robustness:

- Hold-out validation: Using the 70:30 training–testing split described above.
- k-Fold cross-validation: The dataset was randomly partitioned into k = 5 equal subsets. In each iteration, one subset was used for testing, while the remaining four subsets were employed for training. The process was repeated 5 times until every subset had served as a test set once, and the final performance was obtained by averaging the results.

The following evaluation metrics were derived from the confusion matrix to comprehensively assess model performance:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

$$\text{Sensitivity (Recall)} = TP / (TP + FN)$$

$$\text{Specificity} = TN / (TN + FP)$$

$$\text{Precision} = TP / (TP + FP)$$

$$\text{F-measure} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

Knowledge Extraction

Following model evaluation, feature importance and decision rules were analyzed (particularly from tree-based models) to identify the most influential clinical attributes contributing to CVD prediction. This step provides insights for potential clinical interpretation and knowledge discovery.

Results

In total, 270 records were analyzed in the present study, comprising 120 negative cases (absence of CVDs) and 150 positive cases (presence of CVDs). The age of the participants ranged from 29 to 77 years, with a mean age of 54.43 years. Of the total sample, 87 individuals (32.2%) were female, and 183 individuals (67.8%) were male. The dataset included 14 clinical attributes relevant to CVD prediction, encompassing both continuous and categorical variables, which were all preprocessed and discretized to facilitate classification.

Classification Performance

Five widely used classification-based data mining algorithms—NB, ANN, kNN, SVM, and CART—were implemented to predict the presence of CVDs. The models were evaluated using multiple performance metrics, including accuracy, sensitivity, specificity, precision, recall, F-measure, and area under the receiver operating characteristic curve (AUC). Table 2 summarizes the performance of each algorithm.

The results indicated that SVM achieved the highest overall performance, slightly surpassing NB, with an accuracy of 84.44%, a sensitivity of 86.00%, a specificity of 82.50%, and an AUC of 0.9136. In addition, NB demonstrated comparable performance, with an accuracy of 84.07% and an AUC of 0.9133. Moreover, ANN and kNN exhibited moderate predictive ability, while the classification tree, though slightly lower in accuracy (78.52%), provided interpretable decision rules and a simplified model structure.

Although SVM achieved a slightly higher accuracy (84.44%) compared to NB (84.07%) and a marginally

Table 2. Performance Metrics of Classification Algorithms

Algorithm	Accuracy	Sensitivity	Specificity	AUC	F-measure	Precision	Recall
Classification tree	0.7852	0.8267	0.7333	0.8101	0.8105	0.7949	0.8267
Naïve bayes	0.8407	0.8467	0.8333	0.9133	0.8552	0.8639	0.8467
SVM	0.8444	0.8600	0.8250	0.9136	0.8600	0.8600	0.8600
Neural network	0.8074	0.8267	0.7833	0.9058	0.8267	0.8267	0.8267
kNN	0.7963	0.8400	0.7417	0.8831	0.8208	0.8025	0.8400

Note. AUC: The area under the receiver operating characteristic curve; SVM: Support vector machine; kNN: k-nearest neighbors.

higher AUC (0.9136 vs. 0.9133), the differences between the two classifiers were minimal. To assess whether these differences were statistically significant, a McNemar's test was conducted for accuracy, and a DeLong test was performed to compare AUC values. The results revealed no statistically significant difference between SVM and NB for either accuracy ($P>0.05$) or AUC ($P>0.05$), suggesting that both classifiers display comparable predictive performance for CVD in this dataset. These findings highlight that, despite the numerical superiority of SVM in some metrics, NB provides an equally effective predictive model while offering simpler interpretability and computational efficiency.

Decision Tree Analysis

To enhance interpretability, a DT was constructed using the ID3 algorithm. This algorithm selects the attribute with the highest information gain as the root node, iteratively splitting nodes to construct the tree. Information gain, gain ratio, and Gini index were calculated for each attribute (Table 3).

The attribute "thal" (thalassemia status) was identified as the root node due to its highest information gain, followed by chest pain type (cp) and number of major vessels colored (ncolored), among others. The final DT consisted of five levels, 41 nodes, and 27 leaves, providing a clear visual representation of attribute influence on CVD prediction (Figure 1). This tree illustrates the relative importance of clinical attributes in determining the likelihood of disease and offers interpretable rules for potential clinical application.

Based on the DT analysis, several clinical rules were extracted to predict the risk of CVDs. These rules provide clear guidance on how specific combinations of patient characteristics and clinical measurements influence the likelihood of disease. A summary of these extracted clinical rules is presented in Table 4.

Receiver Operating Characteristic Curve Analysis

ROC curves were generated to compare the discriminatory

Table 3. Attribute Evaluation for Decision Tree Construction

Attribute	Information gain	Gain ratio	Gini index
thal	0.2085	0.1712	0.0681
cp	0.1922	0.1115	0.0627
ncolored	0.1752	0.1120	0.0574
exang	0.1299	0.1420	0.0434
maxhrate	0.1151	0.0844	0.0359
oldpeak	0.1124	0.0748	0.0357
slope	0.1111	0.0864	0.0369
sex	0.0668	0.0737	0.0218
age	0.0564	0.0288	0.0190
restecg	0.0241	0.0228	0.0082
restbp	0.0160	0.0090	0.0054
chol	0.0152	0.0108	0.0051
fbs	0.0001	0.0003	0.0000

Table 4. Clinical Rules Extracted From the Decision Tree for Cardiovascular Disease Prediction

Rule	Path	Prediction	Interpretation
1	thal = negative → cp = typical angina → sex = female → maxhrate ≤ 159	Negative for CVDs	Female patients with negative thalassemia status, experiencing typical angina, and with maximum heart rate ≤ 159 are likely not to have CVDs.
2	thal = negative → cp = typical angina → sex = male → oldpeak > 1.5	Positive for CVDs	Male patients with negative thalassemia status, experiencing typical angina, and ST depression (oldpeak) greater than 1.5 are at high risk of CVDs.
3	thal = positive → cp = atypical angina → exang = yes	Positive for CVDs	Patients with positive thalassemia status, atypical angina, and exercise-induced angina are very likely to have CVD.
4	thal = negative → cp = non-anginal pain → cholesterol > 240	Positive for CVDs	Patients with negative thalassemia status, non-anginal chest pain, and very high cholesterol (>240 mg/dL) are at high risk of CVDs.

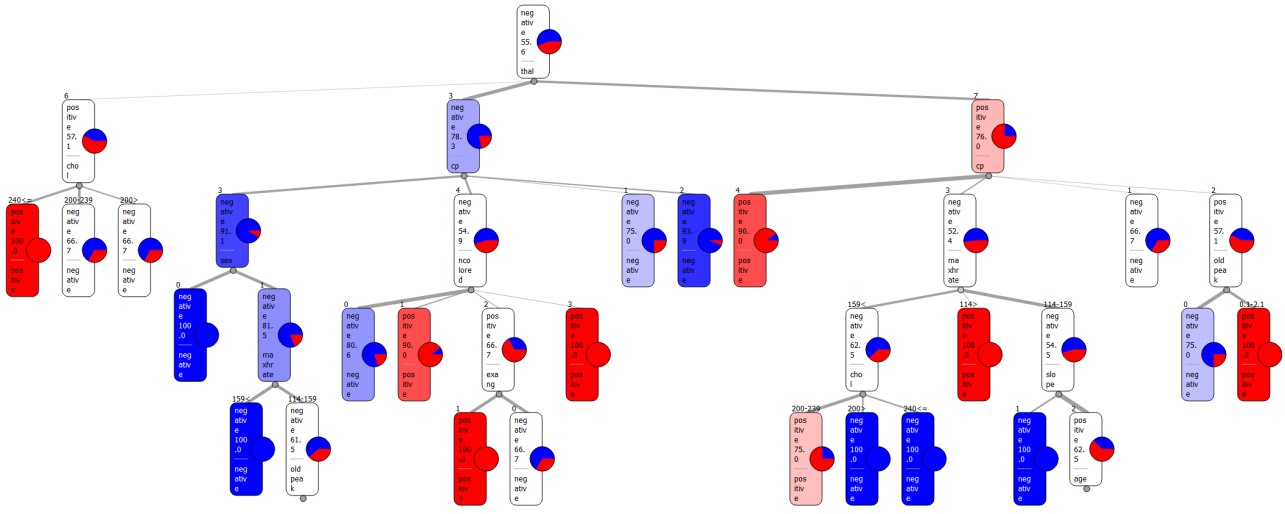

Figure 1. Decision Tree Constructed Using the ID3 Algorithm

power of each classifier. Consistent with performance metrics, SVM demonstrated the highest area under the curve (AUC = 0.9136), slightly outperforming NB (AUC = 0.9133), indicating superior predictive capability. The DT, while slightly lower in AUC, provided the advantage of interpretability, which is critical for clinical decision support (Figure 2).

Discussion

The present study comparatively evaluated five classification-based data mining algorithms—NB, ANN, kNN, SVM, and CART—for predicting CVDs using the UCI Heart Disease dataset. Our findings demonstrated that SVM and NB exhibited the strongest overall performance, achieving accuracies of 84.44% and 84.07%, respectively, with corresponding AUC values of 0.9136 and 0.9133. These models outperformed ANN (accuracy: 80.74%, AUC: 0.9058), kNN (accuracy: 79.63%, AUC: 0.8831), and CART (accuracy: 78.52%, AUC: 0.8101) in both hold-out and 5-fold cross-validation settings. Statistical analyses, including McNemar's test for accuracy and DeLong's test for AUC, revealed no significant differences between SVM and NB ($P > 0.05$), indicating comparable predictive efficacy. These results underscore the utility of ML approaches in enhancing CVD risk stratification, particularly through models that balance predictive power with computational efficiency.

The superior performance of SVM and NB aligns with their inherent strengths in handling high-dimensional, nonlinear data typical of clinical datasets. The effectiveness of SVM stems from its ability to maximize the margin between classes via kernel functions (e.g., radial basis function in this study), which facilitates robust separation of CVD-positive and -negative cases despite potential class imbalances or noise in attributes such as serum cholesterol or resting blood pressure.[17] Similarly, the probabilistic framework of NB, assuming conditional independence among features, enables efficient computation and high sensitivity (86.00% for SVM and 84.67% for NB),

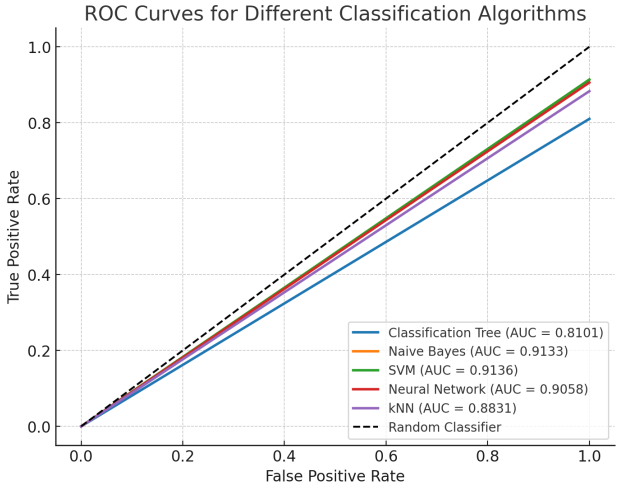

Figure 2. ROC Curves of All the Classification Algorithms. *Note*. ROC: Receiver operating characteristic; AUC: The area under the receiver operating characteristic curve; SVM: Support vector machine; kNN: k-nearest neighbors

making it particularly suitable for early detection where false negatives could have severe clinical consequences.[18] The moderate performance of ANN and kNN may be attributed to their sensitivity to hyperparameters and data scaling; for instance, the single hidden layer of ANN with sigmoid activation might not capture complex interactions as effectively as deeper architectures, while the reliance of kNN on the Euclidean distance could be influenced by outliers in continuous features, such as the maximum heart rate.[19] CART, although yielding the lowest accuracy, offers valuable interpretability through decision rules, which is crucial for clinical translation.[20] The DT identified thalassemia status (thal), chest pain type (cp), and number of major vessels colored by fluoroscopy (ncolored) as the most influential attributes, based on information gain (0.2085 for thal) and Gini index rankings. These features reflect pathophysiological mechanisms: Thalassemia defects impair myocardial perfusion, atypical chest pain signals ischemic events, and ncolored indicates coronary artery occlusion severity. The extracted rules, such as "thal = positive → cp = atypical angina → exang = yes," predicting positive CVD, provide actionable insights into risk assessment, potentially guiding targeted interventions, such as angiography or lifestyle modifications.

Comparisons with prior literature reveal both consistencies and variations in model performance, highlighting the context-specific nature of ML applications in CVD prediction. Using the same UCI dataset, our SVM accuracy (84.44%) is comparable to the findings of the study by Alsabhan and Alfadhly, demonstrating SVM accuracies exceeding 85% with similar preprocessing steps, including discretization and normalization.[12] However, more recent investigations employing advanced ensemble or optimization techniques have achieved higher metrics. For example, a hybrid framework by Teja and Rayalu using optimizable KNN attained 95.04% accuracy and 0.99 AUC on the UCI dataset through correlation-based feature selection and hyperparameter tuning, surpassing our results, possibly due to their integration of genetic algorithms for feature optimization.[13] Likewise, Kumar et al reported 99.7% accuracy with hybrid-optimized kNN, emphasizing the benefits of addressing data imbalance via oversampling, which was not explicitly applied here beyond initial partitioning.[16] Meta-analyses provide a broader context. Krittanawong et al pooled AUCs for SVM in coronary artery disease prediction at 0.92 (95% CI: 0.81–0.97), closely mirroring our 0.9136, while boosting algorithms yielded 0.88 (95% CI: 0.84–0.91).[6] A more recent meta-analysis by Liu et al on electronic health record-based models found random forest and deep learning pooled AUCs of 0.865 (95% confidence interval [CI]: 0.812–0.917) and 0.847 (95% CI: 0.766–0.927), respectively, outperforming conventional scores, such as QRISK3 (AUC: 0.765; 95% CI: 0.734–0.796).[4] These syntheses indicate that while our models' performance is robust within the constraints of traditional classifiers,

incorporating deep learning or ensemble methods—such as those in the study by Cai et al, achieving AUCs up to 0.99 via convolutional neural networks—could elevate accuracy, particularly in larger or multimodal datasets.[5]

Discrepancies in performance across studies may arise from methodological differences, including dataset variations (e.g., UCI's 270 records vs. larger cohorts like UK Biobank), feature engineering (e.g., our equal-width binning vs. advanced selection in hybrid models), and evaluation strategies.[2] For instance, Abdeldjouad et al reported KNN accuracies of 90.8% on UCI data, potentially due to more aggressive outlier removal or cross-validation folds.[21] Our focus on interpretability via CART complements these findings, as tree-based models facilitate clinical adoption by elucidating risk pathways, akin to the findings of Han et al, where DTs identified similar key predictors, including chest pain and vessel occlusion.[15] Nonetheless, the limitations of the UCI dataset (i.e., small size and potential selection bias) may underestimate real-world generalizability, as evidenced by meta-analytic heterogeneity ($I^2 > 99\%$ in the study by Liu et al[4]), underscoring the need for external validation.

The strengths of this study include its comprehensive metric evaluation (accuracy, sensitivity, specificity, precision, recall, F-measure, and AUC), dual validation approaches, and emphasis on interpretability, addressing gaps in prior single-algorithm studies. On the other hand, the limitations encompass reliance on a single dataset, absence of advanced ensembles or deep learning, and potential overfitting in tree models despite depth constraints. Future research should validate these models on diverse, prospective cohorts (e.g., integrating electronic health records or imaging data) and explore hybrid approaches (e.g., SVM-NB ensembles) in order to mitigate biases.[8] Additionally, incorporating explainable artificial intelligence techniques (e.g., SHapley Additive exPlanations) can enhance rule extraction, fostering integration into clinical decision support systems.[22]

In summary, this comparative analysis affirmed SVM and NB as reliable tools for CVD prediction, with CART adding interpretive value. By benchmarking against contemporary literature, our findings contribute to the evolving landscape of ML in cardiology, advocating for tailored, multi-model evaluations to optimize early detection and resource allocation in healthcare.

Conclusion

Our findings demonstrated that data mining algorithms generally yield more accurate predictive models for CVD than traditional statistical methods. Among the evaluated techniques, SVM and NB exhibited superior accuracy and reliability, whereas DT models, utilizing fewer variables, provided simpler and more interpretable models. Comparative analysis with previous studies suggests that the performance of predictive algorithms varies according to the type and volume of data, the nature of variables, and the characteristics of the target population. Therefore, it is recommended that the performance of multiple algorithms be evaluated in each specific context to identify the most suitable approach for developing predictive models. While it is not feasible to claim the creation of a model that can definitively predict CVD, appropriate data collection, rigorous preprocessing, and enhancement of data quality can lead to models with satisfactory predictive accuracy. Ultimately, this study highlights that employing predictive modeling and data mining techniques can offer a more precise understanding of CVD prevalence, thereby supporting evidence-based planning, optimal resource allocation, and the delivery of higher-quality healthcare services to the population.

Acknowledgements

The authors would like to express their sincere gratitude to all those who assisted in this study.

Authors' Contribution

Conceptualization: Sepideh Seyedi-Sahebari, Ali Farzaneh.
Data curation: Ali Farzaneh.
Formal analysis: Ali Farzaneh.
Funding acquisition: Sepideh Seyedi-Sahebari, Ali Farzaneh.
Investigation: Ali Farzaneh.
Methodology: Sepideh Seyedi-Sahebari.
Project administration: Sepideh Seyedi-Sahebari.
Resources: Ali Farzaneh.
Software: Sepideh Seyedi-Sahebari.
Supervision: Ali Farzaneh.
Validation: Sepideh Seyedi-Sahebari.
Visualization: Ali Farzaneh.
Writing–original draft: Sepideh Seyedi-Sahebari, Ali Farzaneh.
Writing–review & editing: Sepideh Seyedi-Sahebari, Ali Farzaneh.

Competing Interests

The authors declare that they have no competing interests.

Consent for Publication

Not applicable.

Data Availability Statement

The dataset supporting the conclusion of this article is included within the article.

Ethical Approval

Not applicable.

Funding

This study was self-funded by the authors and received no external financial support from any funding organization.

Intelligence Use Disclosure

The authors used Copilot for grammar correction and language editing to improve manuscript readability. All AI-generated language suggestions were reviewed and edited by the authors.

References

1. World Health Organization (WHO). Cardiovascular Diseases (CVDs): Key Facts. Geneva: WHO; 2021.
2. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol. 2020;76(25):2982-3021. doi: 10.1016/j.jacc.2020.11.010.
3. Syed MG, Trucco E, Mookiah MR, Lang CC, McCrimmon RJ,

Palmer CN, et al. Deep-learning prediction of cardiovascular outcomes from routine retinal images in individuals with type 2 diabetes. Cardiovasc Diabetol. 2025;24(1):3. doi: 10.1186/s12933-024-02564-w.
4. Liu T, Krentz A, Lu L, Curcin V. Machine learning based prediction models for cardiovascular disease risk using electronic health records data: systematic review and meta-analysis. Eur Heart J Digit Health. 2025;6(1):7-22. doi: 10.1093/ehjdh/ztae080.
5. Cai Y, Cai YQ, Tang LY, Wang YH, Gong M, Jing TC, et al. Artificial intelligence in the risk prediction models of cardiovascular disease and development of an independent validation screening tool: a systematic review. BMC Med. 2024;22(1):56. doi: 10.1186/s12916-024-03273-7.
6. Krittanawong C, Virk HU, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. Sci Rep. 2020;10(1):16057. doi: 10.1038/s41598-020-72685-1.
7. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One. 2017;12(4):e0174944. doi: 10.1371/journal.pone.0174944.
8. Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC Med Inform Decis Mak. 2020;20(1):16. doi: 10.1186/s12911-020-1023-5.
9. Kanwar MK, Lohmueller LC, Kormos RL, Teuteberg JJ, Rogers JG, Lindenfeld J, et al. A Bayesian model to predict survival after left ventricular assist device implantation. JACC Heart Fail. 2018;6(9):771-9. doi: 10.1016/j.jchf.2018.03.016.
10. Tylman W, Waszyrowski T, Napieralski A, Kamiński M, Trafidło T, Kulesza Z, et al. Real-time prediction of acute cardiovascular events using hardware-implemented Bayesian networks. Comput Biol Med. 2016;69:245-53. doi: 10.1016/j.compbiomed.2015.08.015.
11. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng. 2018;2(3):158-64. doi: 10.1038/s41551-018-0195-0.
12. Alsabhan W, Alfadhly A. Effectiveness of machine learning models in diagnosis of heart disease: a comparative study. Sci Rep. 2025;15(1):24568. doi: 10.1038/s41598-025-09423-y.
13. Teja MD, Rayalu GM. Optimizing heart disease diagnosis with advanced machine learning models: a comparison of predictive performance. BMC Cardiovasc Disord. 2025;25(1):212. doi: 10.1186/s12872-025-04627-6.
14. Li G, Wang H, Zhang M, Tupin S, Qiao A, Liu Y, et al. Prediction of 3D cardiovascular hemodynamics before and after coronary artery bypass surgery via deep learning. Commun Biol. 2021;4(1):99. doi: 10.1038/s42003-020-01638-1.
15. Han J, Pei J, Tong H. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers; 2022.
16. Kumar R, Garg S, Kaur R, Johar MG, Singh S, Menon SV, et al. A comprehensive review of machine learning for heart disease prediction: challenges, trends, ethical considerations, and future directions. Front Artif Intell. 2025;8:1583459. doi: 10.3389/frai.2025.1583459.
17. Vapnik V. The Nature of Statistical Learning Theory. Springer Science & Business Media; 2013.
18. Murty MN, Devi VS. Introduction to Pattern Recognition and Machine Learning. World Scientific; 2015.
19. Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep Learning. Cambridge: MIT Press; 2016.
20. Braiek HB, Khomh F. On testing machine learning programs. J Syst Softw. 2020;164:110542. doi: 10.1016/j.jss.2020.110542.
21. Abdeldjouad FZ, Brahami M, Matta N. A hybrid approach for heart disease diagnosis and prediction using machine learning techniques. In: International Conference on Smart Homes and Health Telematics. Cham: Springer International Publishing; 2020.
22. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:1-10.