JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.6738

EVALUATING LOGISTIC REGRESSION, SVM, KNN, AND ENSEMBLE MODELS FOR ACCURATE HEART DISEASE RISK PREDICTION

Amalia Shifa Aldila[1]; Lawrence Adi Supriyono[2*]

Data Science Study Program[1]
Software Engineering Study Program[2]
Universitas Jakarta Internasional, Jakarta, Indonesia[1,2]
www.uniji.ac.id[1,2]
amalia.shifa@uniji.ac.id, lawrence.supriyono@uniji.ac.id[*]

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *Cardiovascular disease remains the most significant contributor to global mortality, highlighting the importance of early and precise risk assessment within preventive healthcare frameworks. Alongside the rapid growth of clinical data availability, machine learning approaches have increasingly been adopted to assist medical decision-making, particularly for interpreting complex and high-dimensional health information. This research investigates the predictive capability of six supervised machine learning models in determining the likelihood of cardiovascular disease incidence: Logistic Regression, Support Vector Machine, k-Nearest Neighbors, Decision Tree, Random Forest, and Gradient Boosting. The Cleveland Heart Disease dataset from the UCI Machine Learning Repository served as the study's foundation. It includes 303 patient samples with a total of 76 recorded attributes. From this dataset, 14 clinically significant variables frequently reported in previous studies were selected for analysis. Considering the relatively small dataset size and the possibility of redundant or low-impact features, a feature selection approach was implemented to improve model robustness, minimize overfitting, and enhance interpretability. The data preparation process involved cleaning, normalization, feature selection, and division into datasets for testing and training. Metrics like accuracy, precision, recall, and F1-score were used to evaluate the model. The results of the experiment show that Random Forest and Logistic Regression models produced the highest predictive performance, followed by k-Nearest Neighbours and Support Vector Machine. These results indicate that supervised machine learning techniques, when supported by appropriate feature selection methods, are effective as decision-support tools for the early detection of cardiovascular disease.*

Keywords: *Classification Algorithms, Cardiovascular Disease, Prediction, Supervised Learning*

Intisari— *Penyakit kardiovaskular hingga saat ini masih menempati posisi teratas sebagai penyebab kematian global, sehingga identifikasi risiko secara dini menjadi aspek krusial dalam strategi pencegahan. Perkembangan dan ketersediaan data klinis yang semakin melimpah mendorong pemanfaatan pembelajaran mesin (machine learning) sebagai solusi pendukung dalam pengambilan keputusan medis, khususnya untuk mengolah data kesehatan yang bersifat kompleks dan multidimensional. Penelitian ini bertujuan untuk menganalisis kemampuan prediktif enam algoritma pembelajaran terawasi, yaitu Logistic Regression, Support Vector Machine, k-Nearest Neighbors, Decision Tree, Random Forest, dan Gradient Boosting, dalam memperkirakan kemungkinan terjadinya penyakit kardiovaskular. Data yang digunakan bersumber dari Cleveland Heart Disease dataset yang tersedia pada UCI Machine Learning Repository, dengan total 303 data pasien dan 76 variabel awal. Dari keseluruhan atribut tersebut, penelitian ini memfokuskan analisis pada 14 fitur klinis yang paling sering digunakan dan dianggap relevan berdasarkan penelitian terdahulu. Mengingat ukuran dataset yang relatif terbatas serta adanya potensi fitur yang saling beririsan atau kurang berpengaruh terhadap hasil prediksi, dilakukan tahapan seleksi fitur guna meningkatkan efektivitas model, menekan risiko overfitting, dan memperjelas interpretasi hasil. Proses praproses data mencakup pembersihan data, normalisasi, pemilihan fitur, serta entri data ke dalam rangkaian analisis dan penelitian. Kinerja setiap* *model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasilnya menunjukkan bahwa Random Forest dan Regresi Logistik menghasilkan kinerja prediksi terbaik, diikuti oleh k-Nearest Neighbors dan Support Vector Machine.Secara keseluruhan, hasil penelitian ini membuktikan bahwa kombinasi algoritma pembelajaran terawasi dengan teknik seleksi fitur yang tepat mampu menghasilkan sistem pendukung keputusan yang efektif dan andal dalam mendukung prediksi dini penyakit kardiovaskular.*

Kata Kunci*: Algoritma Klasifikasi, Penyakit Kardiovaskular, Prediksi, Pembelajaran Terawasi*

INTRODUCTION

Cardiovascular disease (CVD) continues to pose a significant global health burden, contributing to a large proportion of annual mortality and exerting sustained pressure on healthcare infrastructures worldwide [1]. Numerous well-established risk factors—such as hypertension, diabetes mellitus, tobacco use, physical inactivity, and unhealthy dietary habits—substantially elevate the risk of cardiovascular complications [2], [3]. Although clinical diagnostic techniques have advanced considerably, the early detection of individuals with a high risk of CVD remains challenging, particularly in healthcare environments with limited resources [4].

In recent years, growing emphasis has been placed on the early detection and prediction of heart disease, given its critical role in reducing mortality through timely clinical intervention [5]. Parallel to this trend, rapid advancements in machine learning (ML) have enabled new data-driven approaches for cardiovascular risk evaluation by effectively analyzing complex and heterogeneous clinical datasets. Supervised learning techniques—including Logistic Regression, Support Vector Machines (SVM), Decision Trees, and ensemble-based models—have demonstrated considerable potential in uncovering disease-related patterns from medical data [6]. Compared to conventional statistical approaches, these methods are better suited to modeling nonlinear relationships and interactions among clinical variables. Nevertheless, their broader application in real-world healthcare settings is often limited by challenges such as small dataset sizes, redundant features, limited interpretability, and concerns regarding model generalization [7] [8].

Several publicly available datasets have been widely employed in cardiovascular disease prediction research, among which the Cleveland heart disease dataset from the UCI Machine Learning Repository is one of the most extensively used [9]. This dataset comprises 14 clinically relevant attributes related to cardiovascular conditions, including demographic characteristics, cholesterol measurements, and electrocardiographic indicators, which are commonly utilized as input variables in predictive modeling tasks [10]. Prior studies have reported that ensemble learning techniques, particularly Random Forest and Gradient Boosting, tend to yield superior predictive performance and enhance the reliability of clinical decision-support systems [11], [12].

Despite these advancements, predictive accuracy remains constrained by issues such as data imbalance, incomplete records, and limited model transparency [13]. To mitigate these limitations, recent research has explored hybrid modeling strategies and deep learning techniques aimed at further improving prediction outcomes [14]. For example, combining ensemble learning approaches with artificial neural networks (ANNs) has been shown to improve the robustness and reliability of cardiovascular risk predictions [15], [16]. In addition, the development of more informative feature representations and the incorporation of medical domain knowledge are increasingly recognized as essential components for enhancing model performance [17], [18].

The development of advanced predictive systems for cardiovascular disease holds considerable promise for improving healthcare delivery by providing clinicians with timely and accurate decision-support insights. As artificial intelligence and machine learning technologies continue to gain traction in clinical environments, their integration into medical decision-support systems is expected to contribute to meaningful improvements in diagnostic accuracy and patient outcomes [19].

This research undertakes an extensive performance assessment of multiple supervised machine learning approaches for cardiovascular disease prediction, utilizing the Cleveland Heart Disease dataset as the experimental benchmark. Rather than focusing on a single modeling technique, the study compares six widely used algorithms—Logistic Regression, Support Vector Machine, k-Nearest Neighbors (KNN), Decision Tree, Random Forest, and Gradient Boosting—to examine their predictive behavior and robustness. The central aim of this investigation is to determine the most dependable model for estimating cardiovascular disease risk while demonstrating the practical relevance of machine learning methodologies in the analysis of medical datasets.

MATERIALS AND METHODS

This study applies machine learning techniques to the Cleveland UCI Heart Disease dataset using Python 3.8 to compare classification performance across models. The study follows a structured analysis pipeline that includes data preparation, feature refinement, model construction and assessment, and the determination of the most accurate predictive model. The dataset employed in this research is the Cleveland UCI Heart Disease dataset, comprising 303 patient entries and 14 variables, where 13 variables function as inputs and the remaining variable represents the output. Table 1 shows the attributes of the Cleveland UCI heart disease data:

Table1. The Cleveland Dataset

No	Parameter	Icon	Description
1.	Age	Age	Patients age, in years
2.	Sex	Sex	0 = female, 1 = male
3.	Chest pain	Cp	4 types of chest pain (1: typical angina, 2: atypical angina, 3: non aginal pain, 4: asymptotic
4.	Rest blood pressure	Trestbps	Resting systolic (mm/Hg)
5.	Serum cholesterol	Chol	Serum chol in mg/dL
6.	Fasting blood sugar	Fbs	Fasting blood sugar > 120 mg/dL (0:false, 1:true)
7.	Rest electrocardiograph	Restecg	0: normal, 1: having ST-T, 2: left-ventricular hypertrophy
8.	MaxHeart rate	Tahalach	Maximum heart rate achieved
9.	Exercise – induced angina	Exang	0: no, 1: yes
10	ST depression	Oldpeak	ST depression induced by exercise
11	Slope	Slope	1: unslopping, 2: flat 3: down slopping
12	Number of vessels	Ca	Number of major vessels
13	Thalassemia	Thal	Defect types; 3:normal, 6: fixed defect 7: reversible defect
14	Target	Class	Diagnosis of heart disease; 0:normal, 1: presence

Source: (UCI Repository, 2021).

In this research, the data cleaning process began with handling missing values. The fourteen attributes retained from the Cleveland dataset were systematically reviewed to assess whether any missing data were present. Observations containing missing data, denoted by the symbol "?", were handled through either removal or imputation, depending on the severity and characteristics of the missing values. Cases containing missing values in key clinical features, including *thal* and *ca*, were excluded from the analysis due to their low frequency and the potential distortion that imputation might cause.

After the preprocessing stage, categorical features such as *thal* and *cp* were encoded numerically using label encoding to enable their use in machine learning algorithms. Continuous variables, including *chol* and *trestbps*, were assessed for extreme values through boxplot visualization, but all observations were retained given their relevance in a clinical context.

The cleaning process is carried out namely, the process carried out to determine whether the data used has missing values or not, then the transformation process is carried out, at this stage smoothing, normalization, and aggregation is carried out [18]. The next stage is to train the data by classifying the data in two forms, namely train set, and test set. An 80:20 split was applied to the preprocessed dataset to create training and testing sets. Logistic Regression, SVM, and KNN classifiers were trained using the training data and evaluated on unseen test data. Model effectiveness was assessed using accuracy, precision, recall, and F1-score. To ensure feature comparability, Min–Max normalization was implemented to scale numerical variables to a (0,10) range. The final model selection was based on a comparison of accuracy outcomes across all algorithms.

1. Random Forest
   Random Forest is a supervised ensemble-based learning approach that is extensively utilized to address classification problems. The fundamental principle behind Random Forest is that increasing the number of trees generally enhances the model's overall accuracy and robustness. According to [20], the Random Forest model can be implemented using three main methodologies:

   a. Random Forest RI (Random Input Choice) – where randomness is introduced through feature selection;
   b. Forest RC (Random Combination) – which involves random blending of results from different trees; and
   c. A hybrid approach that combines both Forest RI and Forest RC to further improve model performance.

2. Decision Tree
   The method is capable of performing classification on datasets that include a mixture of categorical and numerical variables. Decision tree is the most frequently used algorithm in medical problems [18]. This algorithm is known as entropy, where the entropy of each data is calculated based on each attribute. The form is generally divided into two, namely predictions that give a maximum or minimum value of entropy, namely :

$$Entropy\ (S) = -\sum_{i=1}^{n} P_i\ log_2\ P_i \quad (1)$$

   Explanation:
   ***S**: the set of data samples*
   ***c**: the number of classes*
   ***Pi** : the proportion of samples in class i*

$$Gain(S,A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} . Entropy(S_v) \quad (2)$$

   Explanation:
   ***A**: the attribute used to split the data*
   ***Value A**: the set of all possible values of attribute A*
   ***Sv** : the subset of S where attribute A has value v*
   ***S** : the total number of samples in S*

3. k-Nearest Neighbor (KNN)
   An instance-based supervised learning method called k-Nearest Neighbors (KNN) uses local similarity analysis in the feature space to infer class labels. Instead of using an explicit training model, the technique aggregates the labels of the most comparable data instances to assign a class to an unseen sample. The dominant class then determines the final prediction. A distance function is used to calculate similarity between observations; in this study, the Euclidean distance is used to measure the proximity between feature vectors. The efficacy of KNN is highly sensitive to data quality, despite its conceptual simplicity and ease of implementation; noise and irrelevant features can significantly impair classification performance and predictive dependability [20].

4. Support Vector Machine (SVM)
   A type of supervised learning methods known as Support Vector Machines (SVM) can be used for both regression and classification tasks. SVMs can efficiently model complex relationships and nonlinear patterns in data by using kernel functions to convert input data into higher-dimensional feature spaces [21]. The SVM prediction function can be stated in the following general form:

$$f(x) = w^T\ x + b \quad (3)$$

   Explanation:
   w: weight vector
   x: input feature vector
   b: bias term

5. Logistic Regression (LR)
   A supervised learning technique called logistic regression (LR) is used to estimate the likelihood of a binary dependant variable in order to model dichotomous outcomes. Depending on the collection of predictor variables, this model is applied to situations where options are either indicated or not. The LR model for p as the independent variable looks like this:

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_{0+}\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)}} \quad (4)$$

   Here, P(Y = 1) denotes the probability that a patient is diagnosed with heart disease, while $\beta0$, $\beta1$, … , βp represent the regression coefficients of the model [22]. In the logistic model, there must be a linear regression model. The structure of the linear model is presented as follows:

$$g(x) = \ln\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_{0+}\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \quad (5)$$

   The function g(x) is defined to satisfy the properties of a linear regression framework

[22]. The independent variables are allowed to include a mixture of continuous and categorical attributes [23], [24]. In this research, the use of LR is to determine the accuracy value of the probability of patients being indicated for heart disease.

6. Gradient Boosting (GB)
   The Boosting algorithm follows a similar underlying principle to that of regression algorithms [25]. Let $D = \{xi,\ yi\}N$ denote the dataset. The learning objective is to estimate a function, $F(x)$, that closely represents the true mapping function $F*(x)$, from input x to output y, by minimizing the loss function, L (y, F(x)). The general form of the equation is as follows:

$$F_m(x) = F_{m-1}(x) + \rho_m h_m(x) \quad (6)$$

   Where ρm is the mth weight of the function $hm(x)$. This function is also called an ensemble model, such as a Decision tree. There are several attributes of this algorithm, namely:
   a. Learning rate (v) or shrinkage: This factor governs the extent to which individual trees affect the ensemble's final prediction.
   b. The *max_depth* setting limits the depth to which each decision tree is allowed to expand, similar to the parameter used in the Random Forest algorithm.
   c. Subsampling rate (subsample): defines the proportion of the dataset that is randomly selected for training each tree.
   d. Number of features (max_features): specifies how many features are considered when searching for the optimal split.
   e. Minimum samples for split (min_samples_split): specifies the smallest number of data instances that must be available at a node before a split is permitted.

7. Hyperparameters
   There are six supervised learning algorithms used. Before calculating the accuracy of each model, An initial hyperparameter optimization phase was implemented to select the most effective parameter settings for each model. Model stability was ensured through the use of six-fold cross-validation, while GridSearchCV was employed to identify optimal hyperparameter combinations.The following are the respective forms of the obtained hyperparameters:
   a. Hyperparameters Logistic Regression:
   b. {'C': 0.1, 'penalty': 'l2', 'random_state': 1}
   c. Hyperparameters SVM:
   d. {'n_neighbors': 7, 'n_jobs': 3}
   e. Hyperparameters KNN:
   f. {'C': 3, 'degree': 2, 'gamma': 0.65, 'kernel': 'rbf', 'probability': True, 'random_state': 0}
   g. Hyperparameters Random Forest:
   h. {'max_depth': 30, 'min_samples_split': 17, 'n_estimators': 30, 'random_state': 3}
   i. Hyperparameters Decision Tree:
   j. {'random_state': 5}
   k. Hyperparameters Gradient Boosting:
   l. {'random_state': 100}

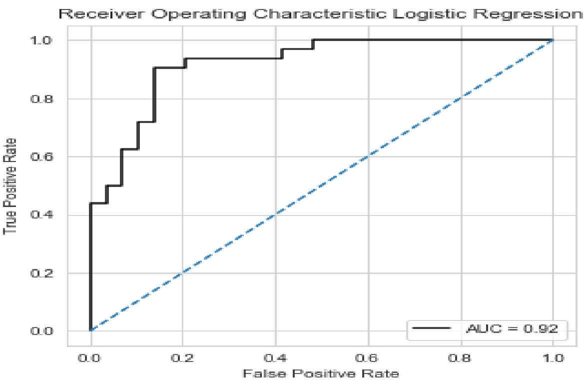
RESULTS AND DISCUSSION

Model accuracy served as the main criterion for assessing machine learning performance in this study, which used a total of 13 chosen characteristics to predict the existence of heart disease. The classification procedure was carried out to choose the best prediction model once the data preparation phase was finished. Six supervised learning algorithms—Logistic Regression, Support Vector Machine, k-Nearest Neighbors, Decision Tree, Random Forest, and Gradient Boosting—were used to develop the predictive models, which were based on the Cleveland Heart Disease dataset that was acquired from the UCI Machine Learning Repository.

Model discrimination ability was analyzed through ROC curve analysis and AUC measurements based on sensitivity and false positive rate calculations. As an alternative evaluation measure, the Area Under the Precision–Recall Curve (AUCPR) was also considered, reflecting mean precision over recall-based sampling. Model evaluation also included the F1 score or F_β. which reflects the harmonic balance between precision and recall. In addition, there are other methods to determine the performance of ML, namely the Matthews correlation coefficient or the cost curve [21]. However, in this paper the Matthews correlation coefficient method is not used, only ROC, AUC, recall, precision, and F1 score.

The evaluated models produced AUC scores between 0.88 and 0.93. Among them, the Logistic Regression model yielded an AUC of 0.92, accompanied by a recall of 0.84, precision of 0.90, and an F1-score of 0.87. Then, for the Support Vector Machine (SVM) model, the respective AUC, recall, precision, and F scores are 0.93, 0.84, 0.90, and 0.87. The AUC, recall, precision, and F scores of the Decision Tree model are 0.88, 0.78, 0.93, and 0.85. For the Gradient Boosting model, the respective AUC, recall, precision, and F scores are 0.90, 0.78, 0.93, and 0.85. The Random Forest and

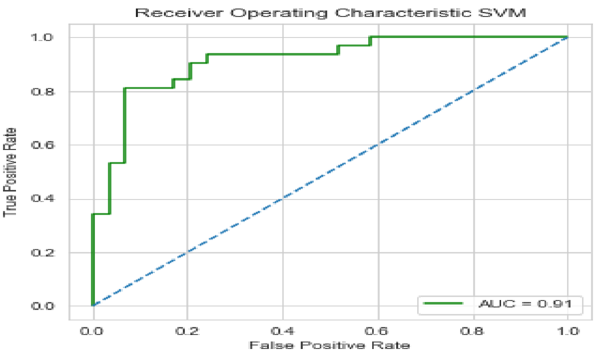
SVM models also have the same AUC, recall, precision, and F scores, namely 0.93, 0.84, 0.90, and 0.87. Then, for the KNN model, the respective values of AUC, recall, precision, and F scores are 0.91, 0.72, 0.92, and 0.81. The following is an image of the ROC graph of each model:



Source : (Research Result, 2025).

Figure 1. ROC of Logistic Regression

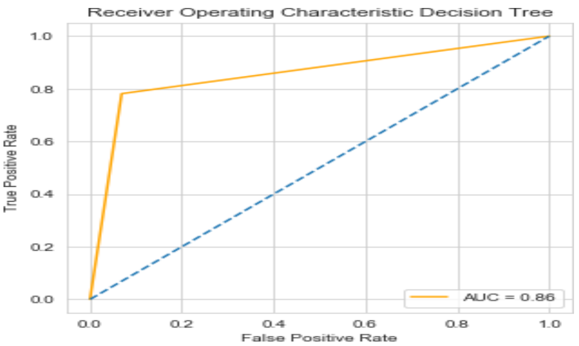
The Logistic Regression classifier demonstrated excellent discriminative performance, achieving an AUC of 0.92 through ROC curve analysis. AUC values closer to 1 are generally associated with improved model effectiveness. Therefore, it can be concluded that the Logistic Regression model indicating strong classification ability. It performed comparably to KNN and slightly below Random Forest and SVM, making it a reliable yet less complex alternative for risk prediction.



Source : (Research Result, 2025).

Figure 2. ROC of SVM

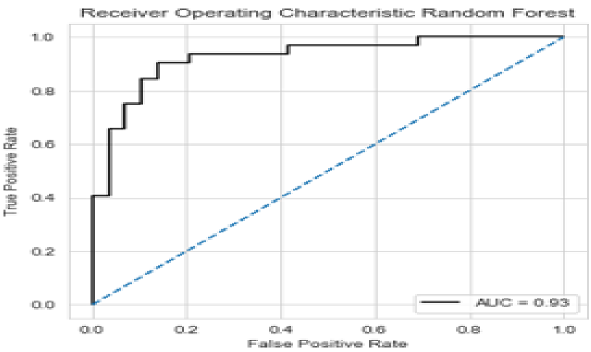
ROC curve analysis of the Support Vector Machine (SVM) classifier resulted in an AUC score of 0.91, reflecting its robust ability to differentiate between individuals with and without potential heart disease. The result confirms that the SVM classifier has a high capacity to accurately separate patients who are at risk of heart disease from those who are not. An AUC value nearing 1.0 signifies strong classification accuracy. Moreover, the SVM model performs on par with the Random Forest model in terms of class separation, highlighting its strong generalization capability, particularly when optimal hyperparameter tuning is applied.



Source : (Research Result, 2025).

Figure 3. ROC of Decision Tree

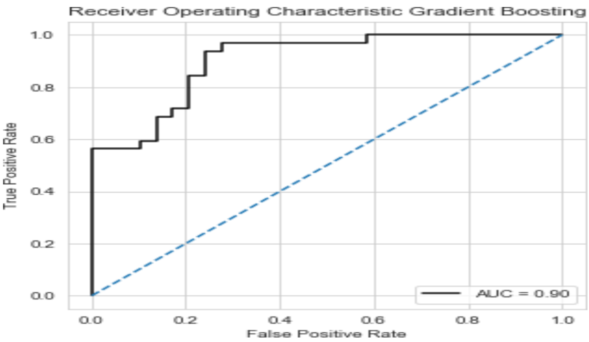
Based on ROC evaluation, the Decision Tree classifier achieved an AUC value of 0.86, reflecting moderate effectiveness in separating patients with and without heart disease. While the Decision Tree performs slightly below Logistic Regression and SVM, it nonetheless maintains a satisfactory level of classification effectiveness. Moreover, its inherent interpretability and structural simplicity make it a valuable choice for certain applications, particularly in contexts where model transparency and explainability are essential.



Source : (Research Result, 2025).

Figure 4. ROC of Random Forest

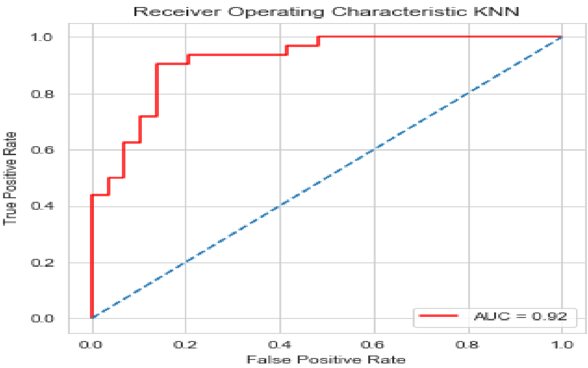
The Random Forest classifier recorded an AUC score of 0.93 based on ROC evaluation, highlighting its outstanding effectiveness in predicting heart disease and its robust ability to separate diseased and non-diseased cases. Among all tested models, Random Forest shows one of the highest predictive accuracies, making it a reliable choice for heart disease prediction due to its ensemble nature and robustness to overfitting. This model also showed excellent capability in distinguishing between patients with and without heart disease, supported by its ensemble structure and robustness to overfitting.



Source : (Research Result, 2025).

Figure 5. ROC of Gradient Boosting

The Gradient Boosting model recorded an AUC score of 0.90 based on ROC evaluation, highlighting its solid predictive capability and its ability to differentiate between individuals with and without heart disease risk. While slightly lower than Random Forest, the Gradient Boosting model remains a strong and reliable algorithm due to its ability to minimize error through iterative learning and handling complex patterns in data.



Source : (Research Result, 2025).

Figure 6. ROC of KNN

The K-Nearest Neighbors (KNN) algorithm's performance evaluation yielded an AUC value of 0.92, indicating its strong ability to accurately distinguish between patients with and without heart disease. The results show strong classification capability in separating heart disease cases from non-cases, reflecting KNN's strength in capturing local data relationships. However, the normalization of feature values and the number of neighbors selected can have a substantial impact on KNN performance.

In addition to these four performance metrics, this study also provided model correctness. 86.89% for Logistic Regression, 88.52% for KNN, 87.00% for Random Forest, 86.89% for SVM, 85.25% for Decision Tree, and 79.40% for Gradient Boosting were the achieved accuracy ratings. The KNN model outperformed all other models in terms of accuracy, followed by Random Forest and Logistic Regression. Accuracy, precision, recall, and F1-score were used to support the Area Under the ROC Curve (AUC) as the primary performance metric for a thorough assessment. The Random Forest model performed the best, achieving an AUC of 0.93 and demonstrating a robust capacity to distinguish heart disease cases from non-cases. Logistic Regression and KNN followed closely, each obtaining an AUC score of 0.92. In addition, the SVM model also recorded an AUC of 0.93, further indicating its strong predictive performance.

Despite KNN recording the highest accuracy of 88.52%, Random Forest remained superior when evaluated using AUC. This result underscores the importance of prioritizing AUC in medical prediction tasks, where imbalanced classes and false-negative errors can have serious clinical implications.

Table 2 presents a comparative summary of the evaluated classification models. The results demonstrate that model performance varies depending on the assessment metric used, indicating that no single algorithm consistently outperforms others across all criteria. Consequently, the choice of the most appropriate model should be guided by the evaluation metric that best supports the intended application, which in this study is effective risk discrimination as measured by the AUC. Additionally, a comparison between the findings of this study and those reported in previous works [2] [11], and [18] is provided in Table 2 below:

Table 2. Summary of AUC Scores from Evaluated Models in This Study

Model	AUC Score	Performance
Random Forest	0.93	Excellent (Best performer)
Logistic Regression	0.92	Excellent
K-Nearest Neighbors	0.92	Excellent
Support Vector Machine (SVM)	0.93	Excellent
Gradient Boosting	0.90	Excellent
Decision Tree	0.86	Good

Source: (Research Result, 2022).

Detail :

1. Random Forest has the highest AUC (0.93), demonstrating superior capability in classifying heart disease presence.
2. Logistic Regression and KNN also showed highly competitive results (AUC = 0.92).

3. SVM and Gradient Boosting maintained strong generalization performance.
4. Decision Tree had the lowest AUC (0.86) among the tested models, though still within acceptable predictive quality.

CONCLUSION

Using the Cleveland Heart Disease dataset from the UCI Machine Learning Repository, this study evaluated the predictive performance of six supervised machine learning techniques: Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree, Random Forest, and Gradient Boosting. Receiver Operating Characteristic (ROC) curves and the corresponding Area Under the Curve (AUC) measurements were used to evaluate the model. The Random Forest algorithm showed the best discriminative capacity among the assessed methods, obtaining an AUC value of 0.93. Logistic Regression and K-Nearest Neighbors followed closely, each attaining an AUC of 0.92. These outcomes suggest that ensemble-based techniques, such as Random Forest, are particularly effective for medical classification problems, as they enhance generalization performance while mitigating the risk of overfitting. The results suggest that ensemble-oriented machine learning approaches are highly effective tools for supporting early cardiovascular disease identification. These models can serve as dependable aids in clinical decision support, enabling earlier diagnosis and more informed interventions, which may ultimately lead to improved patient outcomes.

REFERENCE

[1] World Health Organization, “Cardiovascular Diseases (CVDs),” World Health Organization, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[2] S. Subramani *et al.*, “Cardiovascular diseases prediction by machine learning incorporation with deep learning,” *Front. Med. (Lausanne).*, vol. 10, 2023, doi: 10.3389/fmed.2023.1150933.

[3] A. A. Almazroi, E. A. Aldhahri, S. Bashir, and S. Ashfaq, “A Clinical Decision Support System for Heart Disease Prediction Using Deep Learning,” *IEEE Access*, vol. 11, pp. 61646–61659, 2023, doi: 10.1109/ACCESS.2023.3285247.

[4] M. A. Naser, A. A. Majeed, M. Alsabah, T. R. Al-Shaikhli, and K. M. Kaky, “A Review of Machine Learning’s Role in Cardiovascular Disease Prediction: Recent Advances and Future Challenges,” Feb. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/a17020078.

[5] D. Bertsimas, L. Mingardi, and B. Stellato, “Machine Learning for Real-Time Heart Disease Prediction,” *IEEE J. Biomed. Health Inform.*, vol. 25, no. 9, pp. 3627–3637, Sep. 2021, doi: 10.1109/JBHI.2021.3066347.

[6] H. Heidari, G. Hellstern, and M. Murugappan, “Heart Disease Detection using Quantum Computing and Partitioned Random Forest Methods,” 2022.

[7] C. D. Fernando, P. T. Weerasinghe, and C. K. Walgampaya, “Heart Disease Risk Identification using Machine Learning Techniques for a Highly Imbalanced Dataset: a Comparative Study,” *KDU Journal of Multidisciplinary Studies*, vol. 4, no. 2, pp. 43–55, Dec. 2022, doi: 10.4038/kjms.v4i2.50.

[8] A. Rahim, Y. Rasheed, F. Azam, M. W. Anwar, M. A. Rahim, and A. W. Muzaffar, “An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases,” *IEEE Access*, vol. 9, pp. 106575–106588, 2021, doi: 10.1109/ACCESS.2021.3098688.

[9] UCI, “Heart Disease Dataset”, 2021. [Online].Available: https://archive.ics.uci.edu/ml/datasets/heart+disease.

[10] KEEL Dataset Repository, “Cleveland Dataset”, 2021. [Online]. Available: https://sci2s.ugr.es/keel/dataset.php?cod=57

[11] A. Tiwari, A. Chugh, and A. Sharma, “Ensemble framework for cardiovascular disease prediction,” Comput. Biol. Med., vol. 146, Art. no. 105624, 2022, doi: 10.1016/j.compbiomed.2022.105624.

[12] J. Pirgazi, A. Ghanbari Sorkhi, and M. Iranpour Mobarkeh, “An Accurate Heart Disease Prognosis Using Machine Intelligence and IoMT,” *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/9060340.

[13] C. Zhang, J. Chang, Y. Guan, Q. Li, X. Wang, and X. Zhang, “A Low-Power ECG Processor ASIC Based on an Artificial Neural Network for Arrhythmia Detection,” *Applied Sciences (Switzerland)*,

vol. 13, no. 17, Sep. 2023, doi: 10.3390/app13179591.

[14] P. Shinde, M. Sanghavi, and T. A. Tran, "A Survey on Machine Learning Techniques for Heart Disease Prediction," *SN Comput. Sci.*, vol. 6, no. 4, Apr. 2025, doi: 10.1007/s42979-025-03860-2.

[15] A. M. Paciorek *et al.*, "Automated assessment of cardiac pathologies on cardiac MRI using T1-mapping and late gadolinium phase sensitive inversion recovery sequences with deep learning," *BMC Med. Imaging*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12880-024-01217-4.

[16] T. Polidori *et al.*, "Radiomics applications in cardiac imaging: a comprehensive review," *Radiologia Medica*, vol. 128, no. 8, pp. 922–933, Aug. 2023, doi: 10.1007/s11547-023-01658-x.

[17] A. Almulihi *et al.*, "Ensemble Learning Based on Hybrid Deep Learning Model for Heart Disease Early Prediction," *Diagnostics*, vol. 12, no. 12, Dec. 2022, doi: 10.3390/diagnostics12123215.

[18] C. A. ul Hassan *et al.*, "Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifiers," *Sensors*, vol. 22, no. 19, Oct. 2022, doi: 10.3390/s22197227.

[19] S. P. Knight *et al.*, "Cardiovascular Signal Entropy Predicts All-Cause Mortality: Evidence from The Irish Longitudinal Study on Ageing (TILDA)," *Entropy*, vol. 24, no. 5, May 2022, doi: 10.3390/e24050676.

[20] Y. Tian, I. Luthra, and X. Zhang, "Forecasting COVID-19 cases using Machine Learning models," Jul. 04, 2020. doi: 10.1101/2020.07.02.20145474.

[21] P. Shrivastava, S. Kashikar, P. H. Parihar, P. Kasat, P. Bhangale, and P. Shrivastava, "A systematic review on deep learning-enabled coronary CT angiography for plaque and stenosis quantification and cardiac risk prediction," *Eur. J. Radiol. Open*, vol. 14, Jun. 2025, doi: 10.1016/j.ejro.2025.100652.

[22] N. Kallus, B. Pennicooke, and M. Santacatterina, "More robust estimation of average treatment effects using kernel optimal matching in an observational study of spine surgical interventions," *Stat. Med.*, vol. 40, no. 10, pp. 2305–2320, May 2021, doi: 10.1002/sim.8904.

[23] T. Hočevar, B. Zupan, and J. Stålring, "Conformal prediction with orange," *J. Stat. Softw.*, vol. 98, 2021, doi: 10.18637/jss.v098.i07.

[24] V. A. Sangolgi, M. B. Patil, S. S. Vidap, S. S. Doijode, S. Y. Mulmane, and A. S. Vadaje, "Enhancing Cross-Linguistic Image Caption Generation with Indian Multilingual Voice Interfaces using Deep Learning Techniques," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 547–557.doi: 10.1016/j.procs.2024.03.244.

[25] C. Bentéjac, A. Csörgő, and G. Martínez-Muñoz, "A comparative analysis of gradient boosting algorithms," *Artif. Intell. Rev.*, vol. 54, no. 3, pp. 1937–1967, Mar. 2021, doi: 10.1007/s10462-020-09896-5.