BAREKENG: Journal of Mathematics and Its Applications
December 2025 Volume 19 Issue 4 Page 2739–2750
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss4pp2739-2750

COMPARISON OF SUPERVISED MACHINE LEARNING ALGORITHMS IN HEART FAILURE DISEASE

Jauhara Rana Budiani[1*], Nur Mahmudah[2]

[1,2]*Study Program of Statistics, Faculty of Sciences and Technology, Universitas Nahdlatul Ulama Sunan Giri*
Jln. Jenderal Ahmad Yani, No. 10, Bojonegoro 62115, Indonesia

*Corresponding author's e-mail: * **jbudiani@unugiri.ac.id***

Article History:

Received: 20th February 2025
Revised: 15th April 2025
Accepted: 23rd May 2025
Available online: 1st September 2025

Keywords:

Classification;
Disease prediction;
Heart Failure;
Machine Learning.

ABSTRACT

The heart is a vital organ in the human body that functions to pump blood throughout the body and to the lungs. The heart is located in the chest cavity. The heart is the main force that drives human life. Therefore, if there is a disturbance in heart function, this can cause a decrease in quality of life to death, one of which is heart failure. Heart failure, if not diagnosed and treated quickly, will result in death. Based on findings showing the high death rate due to heart failure, a classification is needed to predict heart failure using machine learning methods. Machine learning can help predict this disease to improve early detection and more accurate medical decision-making. This study focuses on predicting the likelihood of a patient experiencing heart failure. The machine learning algorithm method used is supervised machine learning classification, including decision trees, random forests, naïve bayes, SVM, and K-NN. The results showed that the best method for predicting heart failure was Random Forest with an accuracy of 74.35%, followed by SVM with an accuracy of 69.23%. Meanwhile, Naïve Bayes had the lowest accuracy of 51.28%. Based on these findings, Random Forest is recommended as the best method for heart failure prediction due to its ability to handle data complexity and provide more stable results. Once the best algorithm is obtained, the prediction results and early detection of heart failure will be more accurate.

How to cite this article:

J. R. Budiani and N. Mahmudah, "COMPARISON OF SUPERVISED MACHINE LEARNING ALGORITHMS IN HEART FAILURE DISEASE", *BAREKENG: J. Math. & App.*, vol. 19, iss. 4, pp. 2739-2750, December, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

The heart is an essential organ in the human body, responsible for circulating blood to the lungs and throughout the body. Positioned within the chest cavity, it acts as the primary force sustaining human life [1]. So that, if there is a problem with heart function, it can lead to a decreased quality of life and even death, one of which is due to heart failure. Heart failure is a disease with the potential to cause death. With advances in technology, machine learning algorithms provide an optimal solution for predicting heart disease. Thus, these algorithms are important for studies aimed at reducing the mortality rate of heart failure by earlier prediction classification [2].

One of the primary causes of death globally is heart disease, responsible for 16% of all fatalities. The death has increased from 2.7 million in 2000 to 9.1 million in 2021 [3]. Using data from the Global Burden of Disease and the Institute for Health Metrics and Evaluation (IHME) between 2014 and 2019, the leading cause of death in Indonesia is heart disease. According to the Basic Health Research Data (Riskesdas) from 2013 and 2018, the prevalence of heart disease rose from 0.5% in 2013 to 1.5% in 2018 [4]. Additionally, 50% of coronary heart disease patients have the potential to experience sudden cardiac arrest [5]. Based on this, which shows a high mortality rate in heart failure, a classification is needed to predict heart failure using machine learning technology methods [2]. This method is used in medical prediction systems because of its ability to handle various types of features and provide stable results. However, the performance of different machine learning algorithms in predicting heart failure still varies. Therefore, a comparison of the best-performing machine learning algorithms commonly used for heart failure prediction is necessary [6]. Early detection of the risk of heart failure is essential to prevent more serious complications and reduce mortality. Unfortunately, conventional methods of diagnosis often rely on manual examinations, which can be time-consuming and prone to the subjectivity of medical personnel. Therefore, a new approach is needed that is faster, more accurate, and able to handle large amounts of data to support the diagnosis process with machine learning [7].

RSUD Dr. R. Sosodoro Djatikoesoemo Bojonegoro is the main referral hospital in Bojonegoro Regency, one of which is heart failure cases [8]. Early detection of this disease is still a major challenge due to limitations in the screening methods used. Currently, the diagnosis of heart failure generally relies on laboratory tests [9]. However, this method can often only identify the disease after significant heart damage has occurred. Therefore, a technology-based predictive approach is needed that can help detect the risk of heart failure earlier [10]. Artificial Intelligence technology has shown great potential in the health sector, especially in disease prediction. But in reality, not all disease prediction tools can be represented as early detection in predicting chronic diseases, so this study develops an effective early detection model using the Decision Tree, Random Forest, Logistic Regression, Naïve Bayes, SVM, and K-NN approaches [11], [12].

Machine learning, as a part of artificial intelligence, offers promising solutions in the medical field, especially for disease prediction. With the ability to recognize complex patterns and relationships in health data, machine learning algorithms can help predict the risk of heart failure more accurately based on patient clinical data, such as blood pressure, cholesterol levels, medical history, and other parameters. Machine Learning is a method that requires learning data or training data as learning material, then evaluation is testing data [13]. In this study, Machine Learning algorithms used for classifying heart failure are Decision Tree, Random Forest, Logistic Regression, Naïve Bayes, SVM, and K-NN. Previous research on the classification of heart disease using machine learning has concluded that overall, machine learning techniques are effective in predicting heart disease. The best algorithms for predicting heart disease are K-NN, Naïve Bayes, and Random Forest [14].

This study focuses on evaluating the performance of several machine learning algorithms in predicting heart failure in the community in Bojonegoro Regency. By looking at the characteristics of local data and local demographic conditions, this study aims to determine which model is the most effective and accurate in the context of the region. The results of this study are expected to be the basis for developing an adaptive and local data-based medical decision support system to improve the quality of health services in Bojonegoro. Based on research comparing the K-Nearest Neighbors (K-NN) algorithm and Random Forest for heart failure disease prediction, the Random Forest method achieved the highest accuracy. Another study [6] examined how machine learning can be used to detect heart disease by applying the K-NN algorithm, Decision Tree, and Random Forest techniques. This study found that the Decision Tree method produced the best accuracy. Therefore, this research focuses on predicting the likelihood of a patient having heart failure. It utilizes supervised machine learning classification methods, including Decision Tree, Random Forest,

Naïve Bayes, SVM, and K-NN. A comparison of these methods is conducted by calculating accuracy, sensitivity, and specificity. The ultimate goal of this research is to identify the most effective machine learning method for heart failure classification.

2. RESEARCH METHODS

The algorithms used in this study are Decision Tree, Random Forest, Logistic Regression, Naïve Bayes, SVM, and K-NN. The following is a machine learning algorithm used to predict heart failure at Dr. Sosodoro Djatikoesoemo Hospital, Bojonegoro. The steps taken in this research are as follows in **Figure 1**.

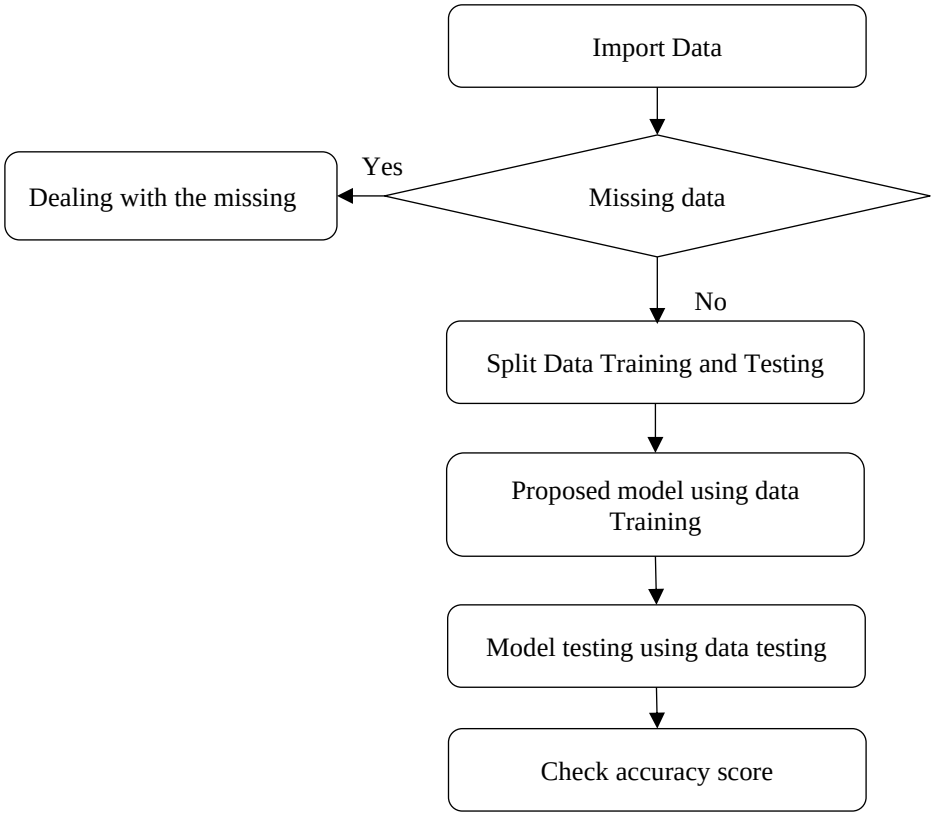

Figure 1. Flow Chart Comparison of Machine Learning Algorithms in Heart Failure Disease

The process begins by importing the data that will be used in the analysis or modeling. After that, a check is carried out to detect any missing data. If found, then the missing data needs to be filled in before proceeding to the next stage. Once the data is ready, the dataset is divided into two parts, which are training data and testing data. The proposed model is then built using the training data to recognize patterns in the dataset. After the training process is complete, the model is tested with the test data to evaluate its ability to make predictions on new data. The final stage is to assess the performance of the model by calculating an accuracy score to measure how well the model is at making predictions.

2.1 Data Collection

This research uses heart failure disease data from 2024 sourced from Dr. Sosodoro Djatikoesoemo Hospital in Bojonegoro Regency. A sample of 130 patient data was used. The variables used are heart failure, which is the target variable, with categories of no heart failure and heart failure. The age variable represents the patient's age in years. The gender variable indicates the patient's sex. The smoking status variable categorizes whether the patient smokes or does not smoke. The diet variable shows whether the patient follows a healthy eating pattern or not. The body mass index variable represents the ratio of the patient's weight to height, measured in kg/m^2. The blood pressure variable is measured in mmHg. The last variable is blood glucose level, measured in mg/dL. The variables used are also listed in **Table 1**.

Table 1. Research Variables

No	Variables	Description
1	Heart failure (Y)	0 : no heart failure 1 : heart failure
2	Age (X1)	Age of patient (year)
3	Gender (X2)	0 : female 1 : male
4	Smoking status (X3)	0 : do not smoke 1 : smoking
5	Diet (X4)	0 : diet 1 : do not diet
6	body mass index (X5)	body mass index of patient (kg/m²)
8	blood pressure (X6)	blood pressure of patient (mmHg)
9	Blood glucose level (X7)	Blood glucose level of patient (mg/dL)

2.2 Data Preprocessing

Data preprocessing is an important stage for machine learning methods. In this study, missing values and outlier data were identified to ensure data quality. After identifying the heart failure patient data, it was found that there was no missing data. The next step is to divide the data into training data and testing data. The training data is 70% and the testing data is 30%.

2.3 Decision Tree

The algorithm used in Decision Trees is the C4.5 algorithm. The C4.5 algorithm constructs a decision tree by creating branches until the rules are met, dividing a dataset into smaller groups based on a set of decision rules. The first step is to determine the root of the tree by calculating the highest gain for each variable or selecting the one with the lowest entropy index **[15]**.

$$Entropy(s) = \sum_{i=1}^{n} -p_i Log_2 P_i \quad (1)$$

where :

$Pi =$ the proportion of the class i in the dataset s

$$Gain(S, A) = Entropy(s) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} \cdot Entropy\ (s) \quad (2)$$

Model creation involves selecting important attributes as the initial node, dividing the data into two subsets based on the attribute value, and then repeating the division process until reaching the leaf node. The final step is model testing, which includes determining accuracy and interpreting the prediction results **[16]**.

2.4 Random Forest

Random forest is a classification technique that integrates multiple decision tree predictors, where each tree is influenced by a randomly selected vector, uniformly distributed across all trees in the forest. The random vector θ_k, is generated by the Random Forest algorithm independently of previous vectors and distributed across all trees. Each tree is then developed using the training dataset and the corresponding random vector θ_k,, resulting in a tree-structured set of classifiers $\{h(x, \theta_k), k = 1, \ldots\}$ applied to the input vector x. The generalization error in the Random Forest algorithm is given by **Equation (3)**.

$$PE^* = P_{X,Y}(mg(X, Y) < 0) \quad (3)$$

Here, the subscripts X and Y represent random vectors, indicating that the probabilities are computed in the X, Y space. The margin function, mg, measures how much the average number of votes for correct outputs on the random vectors exceeds the average number of votes for incorrect outputs. The margin function is defined as follows in **Equation (4)**.

$$mg\ (X, Y) = av_k I\ (h_k(X) = Y) - max_{j \neq Y} av_k I(h_k(X) = i) \quad (4)$$

where I is the indicator function.

Two key parameters for evaluating the performance of individual classifiers and their interdependencies are power and correlation. Random forests with random features are generated by selecting a small subset of input variables at each node. In this study, the random forest consisted of 20 trees, each constructed using 6 randomly selected features. The optimal number of trees and features was determined by testing various combinations and evaluating classification accuracy. As a result, 20 trees yielded the highest accuracy [17].

2.5 Logistic Regression

Binary logistic regression is a method used to predict binary dependent variables. To derive the logistic regression equation, the maximum likelihood ratio is used to assess the statistical significance of the variables. The logistic regression model for *k* independent variables is as follows in Equation (5).

$$P(Y=1) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k)}} \tag{5}$$

The probability of heart failure, $P(Y=1)$, is determined by the regression coefficients $\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k$. This linear model underlies the logistic regression framework. The natural logarithm of the odds ratio, $P(Y=1)$ to $(1-P(Y=1))$, forms a linear model in terms of X_i [18].

$$g(x) = ln\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k \tag{6}$$

2.6 Naïve Bayes

Naïve Bayes classification is a machine learning algorithm based on Bayes' theorem. It is a simple yet effective classification approach that assumes strong independence among features. Bayes' theorem provides a mathematical framework for calculating probabilities. It assumes that predictors are uncorrelated and function independently, with each attribute contributing individually to maximize the overall probability. Although it follows the Naïve Bayes model, it does not employ Bayesian methodology [19]. The probability function in Naïve Bayes is as follows in Equation (7).

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \tag{7}$$

where :

$P(X|Y)$ = posterior probability
$P(X)$ = prior probability
$P(Y)$ = prediction prior probability
$P(Y|X)$ = predictor probability

Naïve Bayes is a simple, easy-to-implement, and efficient classification algorithm that handles non-linear and complex data. However, it suffers from a lack of accuracy because it is based on the assumption of conditional independence between classes.

2.7 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a classification method commonly used in data mining. This method can be applied to both classification and regression cases. Although SVM originally operated on linear principles, this method has evolved to handle non-linear problems by incorporating the concept of kernels in high-dimensional spaces. In this space, SVM identifies separators, known as hyperplanes, which aim to maximize the distance or margin between different data classes. The optimal hyperplane is determined by calculating the margin and finding its maximum value. The process of identifying the most appropriate hyperplane as a class separator is the core of the SVM method [20]. A hyperplane in an n-dimensional space can be characterized by Equation (8) [21].

$$f(x) = w^t\phi(x) + b \tag{8}$$

where :

w = weighted parameter
x = input variable

$\phi(x)$ = feature transformation function
b = biased

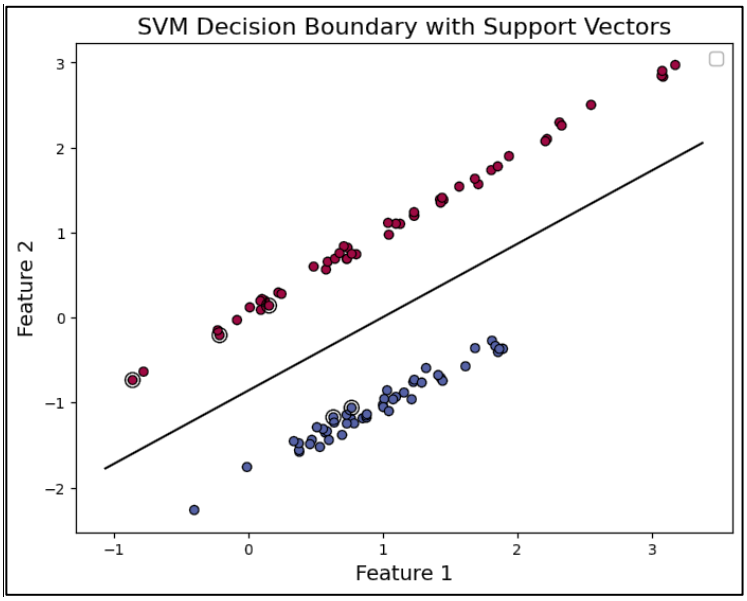

Figure 2. SVM Plot in Heart Failure Disease

Figure 2 is a plot of the SVM method, divided by the hyperplane line. The hyperplane is the separating line between classifications.

2.8 K-Nearest Neighbor (K-NN)

The K-Nearest Neighbors (K-NN) algorithm is a supervised classification method that assigns objects based on their proximity to their nearest neighbors. The distance between an attribute and its neighbors is typically measured using the Euclidean distance. The function of Euclidean distance is as follows in **Equation (9)**.

$$Eu = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \tag{9}$$

The algorithm relies on a set of labeled points to classify new data points. Data is clustered based on similarity, and K-NN can also handle missing values by estimating them based on neighboring data. Once missing values are addressed, various prediction techniques can be applied to the dataset, often improving accuracy when combined with other methods **[15]**. K-NN is easy to implement, as it requires no predefined models or assumptions. Additionally, it is versatile, capable of being used for classification, regression, and search tasks. However, despite being one of the simplest algorithms, its performance can be negatively affected by irrelevant variables.

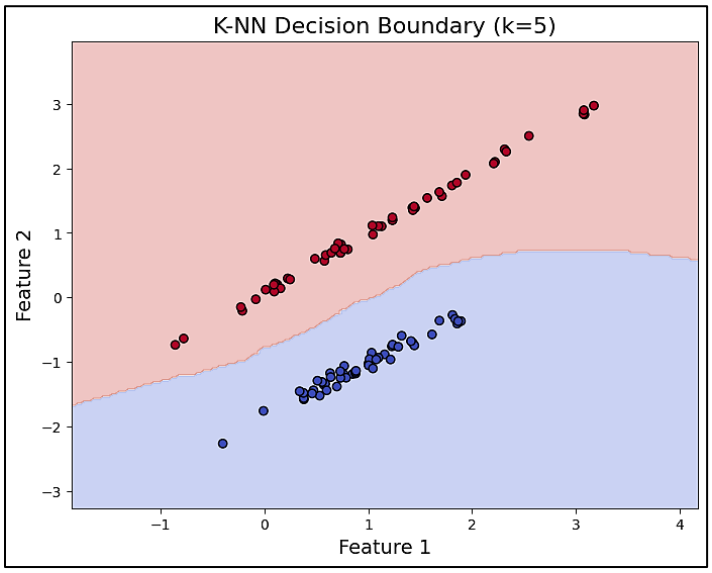

Figure 3. K-NN Plot in Heart Failure Disease

Figure 3 illustrates the classification results using the K-NN method. The red plot is classified as heart failure, and the blue plot is classified as not heart failure.

2.9 Performance Evaluation

The evaluation method for classification results used in this study includes accuracy, sensitivity, and specificity **[22]**. **Table 2** is a confusion matrix that shows the prediction results of heart failure data classification using a machine learning algorithm.

Table 2. Cofusion Matrix

Observation	Negative Predict	Positive Predict
Negative Actual	TN (*True Negative*)	FP (*False Positive*)
Positive Actual	FN (*False Negative*)	TN (*True Negative*)

The functions of accuracy, sensitivity, and specificity are as follows in **Equation (10)**, **Equation (11)**, and **Equation (12)**.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (10)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (11)$$

$$Specificity = \frac{TN}{TN + FP} \times 100\% \quad (12)$$

3. RESULTS AND DISCUSSION

3.1 Data Exploration

First, data exploration is carried out by making correlations between variables and outlier detection. **Figure 4** shows that the red color has a high positive correlation, while the dark blue color has a negative correlation.

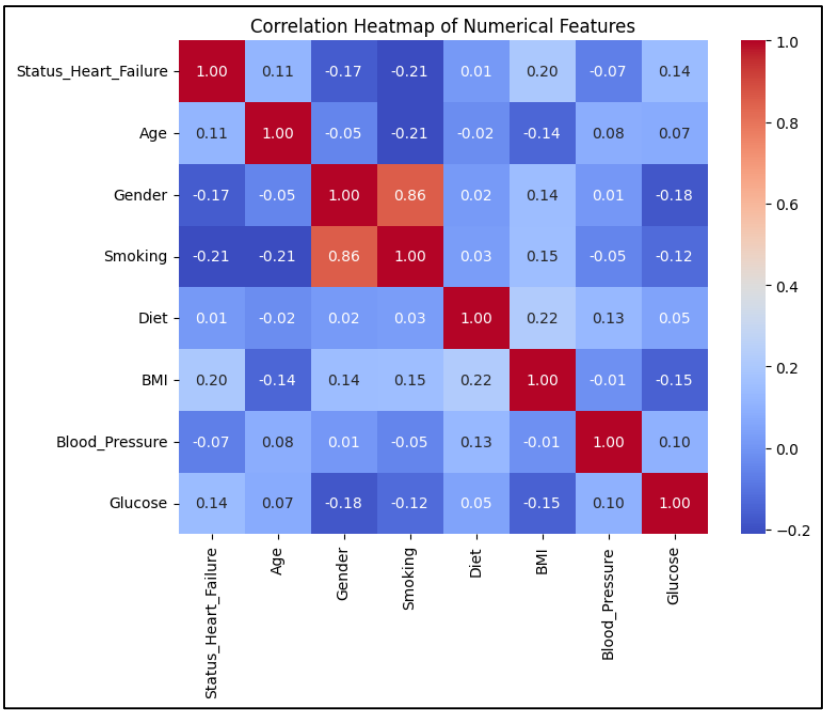

Figure 4. Heatmap Correlation All Variables

Figure 5 shows that there are outliers in Age, BMI, and Glucose. In this study, outlier data is not removed because the data is important and will affect the prediction results. Research **[23]** states that Random Forest is the least sensitive machine learning method to outlier data. Random Forest is an ensemble method that combines decision trees.

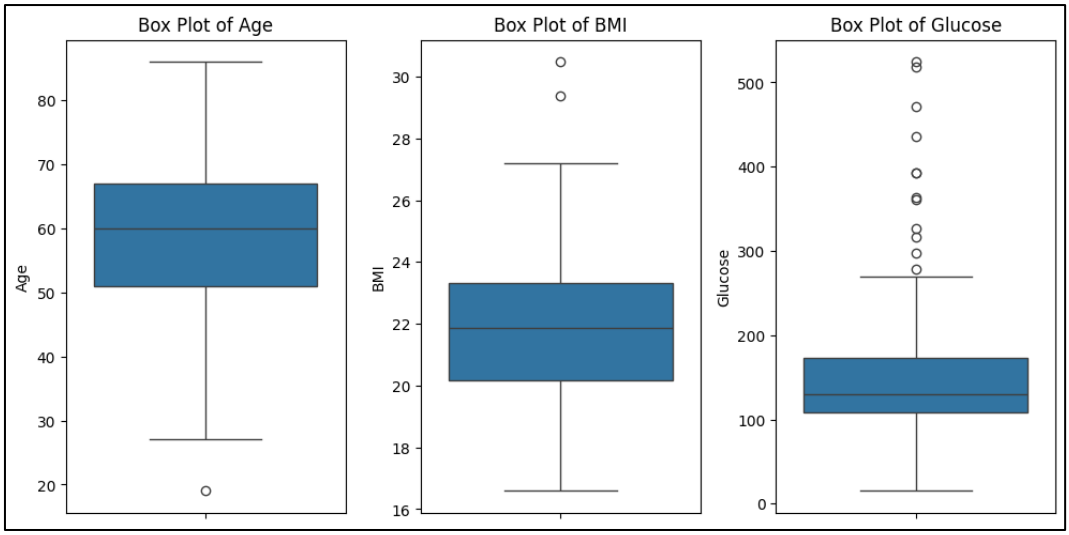

Figure 5. Outlier Data Detection

3.2 Result of Comparison Machine Learning Algorithm

The classification of heart failure patients using Decision Tree, Random Forest, Logistic Regression, Naïve Bayes, and K-NN methods produces a model capable of predicting heart failure. The evaluation method assesses accuracy, sensitivity, and specificity by generating a confusion matrix from the model's prediction results. The accuracy results of heart failure classification using machine learning are presented in **Figure 6**.

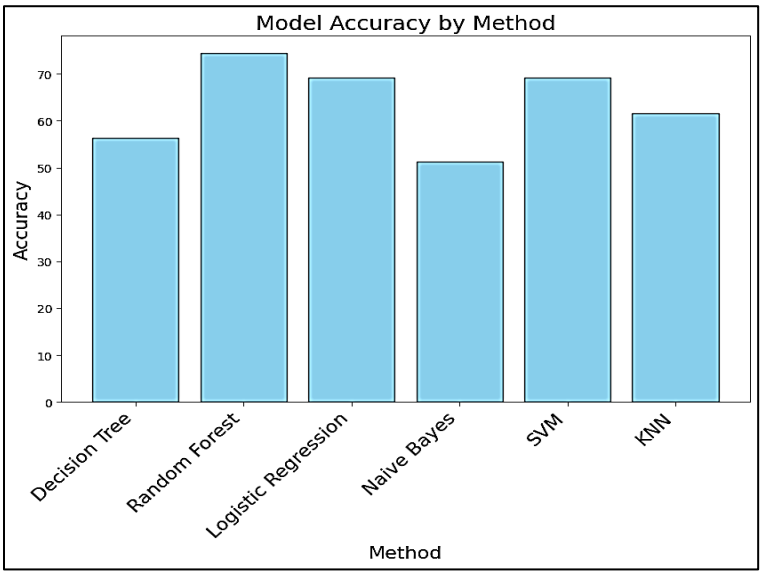

Figure 6. Model Accuracy

Based on the model evaluation results in **Figure 6**, the Random Forest method achieved the highest accuracy at 74.35%, while the Naïve Bayes method had the lowest accuracy. The machine learning model evaluation results for heart failure patients are presented in **Table 3.** The Random Forest method also obtained the highest sensitivity value at 89.28%, whereas the Naïve Bayes method had the lowest sensitivity. The highest specificity value (54.54%) was achieved by the Decision Tree and K-NN methods, while the lowest specificity was observed in the Logistic Regression and SVM methods. In machine learning-based classification, the Random Forest method demonstrated the best accuracy and sensitivity compared to other methods.

Table 3. Comparison of Performance Methods

Method	Accuracy	Sensitivity	Specificity
Decision Tree	56.41%	67.85%	27.27%
Random Forest	74.35%	89.28%	36.36%
Logistic Regression	69.23%	78.57%	45.45%
Naive Bayes	51.28%	50.00%	54.54%
SVM	69.23%	85.71%	27.27%
K-NN	61.53%	75.00%	27.27%

Based on **Table 3**, Random Forest (RF) demonstrates the best performance with an accuracy of 74.35%, followed by SVM with an accuracy of 69.23%. Other algorithms, such as Logistic Regression, KNN, Decision Tree, and Naïve Bayes, also show relatively good performance. Based on this, it is proven that Random Forest is more robust for outlier data than other algorithms. According to research conducted by **[23]**.

Table 4. Earlier Research of Heart Disease Prediction

Authors	Technique	Accuracy
F. Fredilio et al. [2]	KNN	86%
	Random Forest	96%
A. Hartono et al. [6]	K-NN	62%
	Decision Tree	90%
	Random Forest	87%
D. Shah et al. [14]	Naïve Bayes	88.157%
	K-NN	90.789%
	Decision tree	80.263%
	Random forest	86.84%

Authors	Technique	Accuracy
A. Damayunita et al. [24]	Naïve Bayes	88%
	K-NN	91%
	SVM	92%
F. S. Alotaibi [25]	Decision Tree	93.19%
	Logistic Regression	87.36%
	Random Forest	89.14%
	Naïve Bayes	87.27%
	SVM	92.30%
S. Xu et al. [26]	Random Forest	91.6%
	C4.5	89.6%
	SVM	89.2%
	Naïve Bayes	85.2%
	RBF Network	84.2%
	Adaboost	82.8%
A. Rahmah et al. [27]	SVM	81.51%
	Random Forest	83.33%
R. Baxani et al. [28]	Logistic Regression	91%
	SVM	90%
	Random Forest	99%
U. K. Lilhore [29]	SVM	90.25%
	MARS	88.12%
	Random Forest	90.12%
	Decision Tree	89.69%
	BGLM	86.99%

Research on the comparison of K-Nearest Neighbors (K-NN) and Random Forest algorithms for heart failure prediction [2] indicates that Random Forest achieves the highest accuracy in making early predictions for heart failure patients. The findings of this study are consistent with previous research. For instance, the study titled *"Cardiovascular Risk Prediction Method Based on CFS Subset Evaluation and Random Forest Classification Framework"* reported that the Random Forest method achieved a high accuracy of 97% [26]. According to research [27], the RF algorithm achieved the highest accuracy of 83.33%, indicating its effectiveness in classifying heart failure. As well as this research [28] using data from the Framingham Heart Study on heart attacks and heart failure, the random forest technique performed better than other methods with identical parameters. This study aligns with [29], which examined patients with a history of hepatitis C, showing that the Random Forest and SVM methods are effective tools for predicting the risk of patients developing Hepatitis C. This allows for early intervention and improves the quality of care in hospitals.

Compared to previous studies, this study achieves lower accuracy due to the absence of the ensemble method. To improve accuracy, various ensemble techniques or method combinations can be applied, such as tuning hyperparameters, using k-fold cross-validation, and addressing data imbalance.

4. CONCLUSION

After conducting research stages, starting from data pre-processing to model evaluation, it was found that the best method for predicting heart failure is Random Forest, with an accuracy of 74.34%. Random Forest is the most effective machine learning method for heart failure classification. This is because RF has proven to be less sensitive to outlier data and can handle data complexity and provide more stable results.

Based on the limitations of the research we conducted, it can be used as input for further research can be conducted to achieve higher accuracy. This research only includes 130 datasets, so it may not be able to generalize to other data. Future research requires more advanced methods to deal with outlier data and small data. Methods to overcome outliers include using bootstrapping and using SMOTE for imbalance.

AUTHOR CONTRIBUTIONS

Jauhara Rana Budiani: Conceptualization, Formal Analysis, Methodology, Writing - Original Draft. Nur Mahmudah: Data Curation, Investigation, Writing - Review and Editing. All authors discussed the results and contributed to the final manuscript.

FUNDING STATEMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. All research expenses were personally funded by the author.

ACKNOWLEDGMENT

The author gratefully acknowledges RSUD Sosodoro Djatikoesoemo Bojonegoro for granting permission and support in providing the data used in this research. Appreciation is also extended to the Institute for Research and Community Service (LPPM) of Universitas Nahdlatul Ulama Sunan Giri for its support. Such contributions were invaluable to the smooth execution and successful completion of this study.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest regarding the conduct, results, or interpretation of this research, whether financial or non-financial.

REFERENCES

[1] I. M. Adhikara, "MENGENAL JANTUNG," Direktorat Kajian dan Inovasi Akademik, 2024. [Online]. Available: https://mooc.ugm.ac.id/mengenal-jantung/#:~:text=Jantung%20adalah%20salah%20satu%20organ,seluruh%20tubuh%20dan%20paru%2Dparu.. [Accessed 19 November 2024].

[2] D. Sitompul, R. Hamonangan, D. J. Ziegel and E. Indra, "PERBANDINGAN ALGORITMA K-NEAREST NEIGHBORS (K-NN) DAN RANDOM FOREST TERHADAP PENYAKIT GAGAL JANTUNG," *Jurnal Teknlogi Informatika dan Komputer MH. Thamrin,* Vols. 9, no. 1, p. 471–486, 2023. Doi: https://doi.org/10.37012/jtik.v9i1.1432

[3] WHO, "THE TOP 10 CAUSES OF DEATH," 7 Agustus 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. [Accessed 18 November 2024].

[4] Kemenkes, "SEHAT NEGERIKU," 29 September 2022. [Online]. Available: https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20220929/0541166/penyakit-jantung-penyebab-utama-kematian-kemenkes-perkuat-layanan-primer/. [Accessed 18 November 2024].

[5] Kemenkes, "KEMENKES HEBAT, INDONESIA SEHAT," 23 September 2024. [Online]. Available: https://kemkes.go.id/id/rilis-kesehatan/kenali-gejala-jantung-sejak-dini. [Accessed 18 November 2024].

[6] A. Hartono, L. A. Dew, E. Yuniarti and S. T. H. Putri, "MACHINE LEARNING CLASSIFICATION FOR DETECTING HEART DISEASE WITH K-NN ALGORITHM, DECISION TREE AND RANDOM FOREST," *Eksakta: Berkala Ilmiah Bidang MIPA,* vol. 24 NO 04, pp. 513-522, 2023.

[7] T. A. E. Putri, T. Widiharih and R. Santoso, "PENERAPAN TUNING HYPERPARAMETER RANDOMSEARCHCV PADA ADAPTIVE BOOSTING UNTUK PREDIKSI KELANGSUNGAN HIDUP PASIEN GAGAL JANTUNG," *JURNAL GAUSSIAN,* vol. 11, no. 3, pp. 397-406, 2022. Doi: https://doi.org/10.14710/j.gauss.11.3.397-406

[8] M. R. Alfaqih, C. Indianti and A. Z. Abidin, "HUBUNGAN KLASIFIKASI HIPERTENSI DENGAN PENYAKIT JANTUNG KORONER DI POLI JANTUNG RSUD DR. R. SOSODORO DJATIKOESOEMO BOJONEGORO," *Jurnal Ilmiah Keperawatan (Scientific Journal of Nursing),* vol. 10, no. 3, pp. 514-521, 2024. Doi: https://doi.org/10.33023/jikep.v10i3.2289

[9] S. Y. Prasetyo, "PREDIKSI GAGAL JANTUNG MENGGUNAKAN ARTIFICIAL NEURAL NETWORK," *Jurnal Saintekom : Sains, Teknologi, Komputer dan Manajemen,* vol. 13, no. 1, pp. 79-88, 2023. Doi: https://doi.org/10.33020/saintekom.v13i1.379

[10] P. Purwono, P. Dewi, S. K. Wibisono and B. P. Dewa, "MODEL PREDIKSI OTOMATIS JENIS PENYAKIT HIPERTENSI DENGAN PEMANFAATAN ALGORITMA MACHINE LEARNING ARTIFICIAL NEURAL NETWORK," *INSECT,* vol. 7, no. 2, pp. 82-90, 2022. Doi: https://doi.org/10.33506/insect.v7i2.1828

[11] R. Pranandito and Hendry, "PERBANDINGAN PREDIKSI PENYAKIT SERANGAN JANTUNG MENGGUNAKAN MODEL MACHINE LEARNING," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)* , vol. 8, no. 4, pp. 1228-1237, 2023. Doi: https://doi.org/10.29100/jipi.v8i4.4165

[12] S. Adi and A. Wintarti, "KOMPARASI METODE SUPPORT VECTOR MACHINE (SVM), K-NEAREST NEIGHBORS (KNN), DAN RANDOM FOREST (RF) UNTUK PREDIKSI PENYAKIT GAGAL JANTUNG," *MATHunesa Jurnal Ilmiah Matematika* , vol. 10, no. 2, pp. 258-269, 2022. Doi: https://doi.org/10.26740/mathunesa.v10n2.p258-268

[13] I. Cholissodin, S. A. A. Soebroto, U. Hasanah and Y. I. Febiola, AI, MACHINE LEARNING & DEEP LEARNING (Teori & Implementasi), Malang: Fakultas Ilmu Komputer, 2020.

[14] D. Shah, S. Patel and S. K. Bharti, "HEART DISEASE PREDICTION USING MACHINE LEARNING TECHNIQUES," *SN Computer Science,* pp. 1-6, 2020.doi: https://doi.org/10.1007/s42979-020-00365-y

[15] R. N. Romadhon, A. Ogi and A. P. Agung, "IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI PELANGGAN," *Karimah Tauhid,* vol. 3, no. 2, pp. 1860-1874, 2024.doi: https://doi.org/10.30997/karimahtauhid.v3i2.11952

[16] C. Valsecchi, F. Grisoni, V. Consonni, D. Ballabio and R. Todeschini, MULTITASK LEARNING FOR QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIPS: A TUTORIAL. In: Hong, H. (eds) Machine Learning and Deep Learning in Computational Toxicology. Computational Methods in Engineering & the Sciences, Cham: Springer, 2023.doi: https://doi.org/10.1007/978-3-031-20730-3_8

[17] Z. Masetic and . A. Subasi, "CONGESTIVE HEART FAILURE DETECTION USING RANDOM FOREST CLASSIFIER," *Computer Methods and Programs in Biomedicine,* 2016.doi: https://doi.org/10.1016/j.cmpb.2016.03.020

[18] A. Das, "LOGISTIC REGRESSION." ENCYCLOPEDIA OF QUALITY OF LIFE AND WELL-BEING RESEARCH, CHAM: SPRINGER INTERNATIONAL PUBLISHING, 2024, pp. 3985-3986.doi: https://doi.org/10.1007/978-3-031-17299-1_1689

[19] P. M. Fatima M., "SURVEY OF MACHINE LEARNING ALGORITHMS FOR DISEASE DIAGNOSTIC," *J Intell Learn Syst Application,* vol. 9, pp. 1-16, 2017.doi: https://doi.org/10.4236/jilsa.2017.91001

[20] B. Kim, "COMPARISON OF HEART FAILURE PREDICTION PERFORMANCE USING VARIOUS MACHINE LEARNING TECHNIQUES," *International Journal of Internet, Broadcasting and Communication,* vol. 16(4), pp. 290-300, 2024.

[21] M. K. Biddinika, A. Masitha, H. Herman and V. A. N. Fatimah, "MACHINE LEARNING TECHNIQUES FOR HEART DISEASE PREDICTION USING A MULTI-ALGORITHM APPROACH," *JUITA: Jurnal Informatika* , vol. 12(2), pp. 149-158, 2024. Doi: https://doi.org/10.30595/juita.v12i2.24153

[22] L. R. Safitri, N. Chamidah, T. Saifudin, M. Firmansyah and G. T. Alpandi, "COMPARISON OF LOGISTIC REGRESSION AND SUPPORT VECTOR MACHINE IN PREDICTING STROKE RISK," *INFERENSI,* vol. 7, no. 2, pp. 123-128, July 2024.doi: https://doi.org/10.12962/j27213862.v7i2.20420

[23] M. Hamza and D. Larocque, "AN EMPIRICAL COMPARISON OF ENSEMBLE METHODS BASED ON CLASSIFICATION TREES," *Journal of Statistical Computation and Simulation,* vol. 75, no. 8, p. 629–643, August 2005.doi: https://doi.org/10.1080/00949650410001729472

[24] A. Damayunita, R. S. Fuadi and C. Juliane, "COMPARATIVE ANALYSIS OF NAIVE BAYES, K-NEAREST NEIGHBORS (KNN), AND SUPPORT VECTOR MACHINE (SVM) ALGORITHMS FOR CLASSIFICATION OF HEART DISEASE PATIENTS," *JOIN (Jurnal Online Informatika,* vol. 7, no. 2, pp. 219-225, 2022. doi: https://doi.org/10.15575/join.v7i2.919

[25] F. S. Alotaibi, "IMPLEMENTATION OF MACHINE LEARNING MODEL TO PREDICT HEART FAILURE DISEASE," *(IJACSA) International Journal of Advanced Computer Science and Applications,* Vols. Vol. 10, No. 6, pp. 261-268, 2019.doi: https://doi.org/10.14569/IJACSA.2019.0100637

[26] S. Xu, Z. Zhang, D. Wang, J. Hu, X. Duan and T. Zhu, "CARDIOVASCULAR RISK PREDICTION METHOD BASED ON CFS SUBSET EVALUATION AND RANDOM FOREST CLASSIFICATION FRAMEWORK," in *2017 IEEE 2nd International Conference on Big Data Analysis* , Beijing, 2017.doi: https://doi.org/10.1109/ICBDA.2017.8078813

[27] A. Rahmah, N. Sepriyanti, M. H. Zikri, I. Ambarani and M. Y. bin Shahar, "IMPLEMENTATION OF SUPPORT VECTOR MACHINE AND RANDOM FOREST FOR HEART FAILURE DISEASE CLASSIFICATION," *Public Research Journal of Engineering, Data Technology and Computer Science,* vol. 1, no. 1, pp. 34-40, 2023. Doi: https://doi.org/10.57152/predatecs.v1i1.816

[28] R. Baxani and M. Edinburgh, "Heart Disease Prediction Using Machine Learning Algorithms Logistic Regression, Support Vector Machine and Random Forest Classification Techniques," *SSRN,* July 1, 2022. Doi: https://doi.org/10.2139/ssrn.4151423

[29] U. K. Lilhore, P. Manoharan, J. K. Sandhu, S. Simaiya, S. Dalal, M. A. Baqasah, M. Alsafyani, R. Alroobaea, I. Keshta and K. Raahemifar, "HYBRID MODEL FOR PRECISE HEPATITIS-C CLASSIFICATION USING IMPROVED RANDOM FOREST AND SVM METHOD," *Scientifc Reports,* Vols. vol. 13, no. 1, p. 1–18, 2023.doi: https://doi.org/10.1038/s41598-023-36605-3