The Indonesian Journal of Computer Science
www.ijcs.net
Volume 14, Issue 2, April 2025
https://doi.org/10.33022/ijcs.v14i2.4765

Heart Failure Disease Classification Using Random Forest Algorithm With Grid Search Cross Validation Technique

Rapindra Septia[1], Junadhi[2], Susi Erlinda[3], Wirta Agustin[4]
septiarapindra@gmail.com[1], junadhi@usti.ac.id[2], susierlinda@usti.ac.id[3], wirtaagustin@usti.ac.id[4]
[1,2,3,4] Universitas Sains dan Teknologi Indonesia

Article Information

Received : 27 Feb 2025
Revised : 27 Apr 2025
Accepted : 30 Apr 2025

Keywords

Heart failure, classification, Random Forest, SMOTE, Winsorization

Abstract

Heart failure disease is one of the leading causes of death in the world, so early detection is very important to increase the patient's chances of survival. This study aims to develop a heart failure classification model using the Random Forest algorithm. The data used comes from the Heart Failure Prediction dataset on Kaggle, which consists of 299 samples with 13 relevant features. The pre-processing includes outlier handling using Winsorization method as well as class balancing with Synthetic Minority Over-sampling Technique (SMOTE). Models were developed and evaluated using accuracy, precision, recall, and F1-score metrics. The results showed that the optimized Random Forest model achieved an accuracy of 90%, higher than the other models tested. These findings indicate that the method used is effective in classifying patients at risk of heart failure, so it can be used as a tool in medical decision-making.

A. Introduction

Heart failure can be caused by a variety of disorders of the heart. This condition is one of the major contributing factors to the world's mortality rate. According to the World Health Organization (WHO), about 17 million people die each year from cardiovascular diseases, including heart failure, accounting for 32% of total deaths globally[1]. Heart failure is a condition where the heart is unable to distribute oxygen optimally due to impaired structure or function. Causes include abnormalities in the heart, blood vessels, or metabolism. The body attempts to compensate for these disorders through compensatory mechanisms, but often fails, leading to a decline in heart function. As the end stage of many cardiovascular diseases, heart failure has an increasing incidence and is a leading cause of global morbidity and mortality[2]. Heart failure is triggered by risk factors such as hypertension, diabetes, heart valve disorders, obesity, smoking, alcohol, family history, and aging. Controlling these factors with a healthy lifestyle and medication can reduce the risk [3]. Early detection of heart failure is important to prevent complications and reduce mortality. By recognizing risk factors sooner, treatment and lifestyle changes can be optimally applied, slowing down the disease and improving the patient's quality of life[4]. Early screening for heart disease is important for early detection and prevention of complications [5]. Heart failure can be prevented with a healthy lifestyle, such as maintaining an ideal weight, eating nutritious foods, limiting salt and sugar, avoiding cigarettes and alcohol, and exercising regularly[6].

In the midst of rapid technological advancements, machine learning has become one of the most important approaches in various sectors, including healthcare. This technology provides solutions to automatically analyze data without the need for direct supervision, thus supporting the process of disease diagnosis and clinical decision-making more efficiently [7]. General hospitals face challenges in managing heart failure, such as limited medical personnel and educational facilities [8]. This research develops a machine learning-based application with the Random Forest algorithm to detect heart failure and classify patient risk[9]. Random Forest is a machine learning algorithm for classification that builds multiple decision trees and combines them for more stable and accurate predictions. Using the bagging method, it trains a decision tree on a dataset by randomly selecting a subset of features, thus increasing the diversity and effectiveness of the model, especially on high-dimensional data [10].

Various studies have evaluated machine learning algorithms in heart failure prediction. Wahyu Nurgraha et al. (2024) compared several algorithms, with the result that Random Forest achieved the highest accuracy of 95%, making it an effective method for clinical decision support[11]. Aryo Sasi Kirono et al. (2024) compared Decision Tree, Random Forest, and XGBoost algorithms, where Random Forest showed the best performance with accuracy 0.86, precision 0.87, recall 0.86, and F1-score 0.86, so it was considered reliable for detecting and analyzing the causes of heart failure[12]. Socayo Adi et al. (2022) found that SVM and Random Forest achieved the highest accuracy of 97%, with Random Forest using n_estimators=30, making it optimal in heart failure prediction [13]. Shafira Amanda Putri et al. (2024) also showed that Random Forest has the best accuracy, which is 87.7% for test data and 92.6% for validation, making it the optimal model for heart disease diagnosis [14]. Filbert Duran Putri et al. (2024) compared Random Forest Classifier and LightGBM Classifier, with the result that Random Forest excels with 95.37% accuracy, making it more suitable for heart disease analysis [15]. Although Random Forest has been proven effective in heart failure classification, optimal hyperparameter selection is still a challenge. The use of Grid Search Cross-Validation (GridSearchCV) allows finding the best combination of n_estimators and max_depth to improve model accuracy and stability. However, the grid search process often requires high computational time. Therefore, this research focuses on Random Forest optimization using GridSearchCV to obtain the best performing model for early heart failure detection.

B. Research Method

This research was conducted systematically through the stages of data processing, model development, and evaluation, designed to produce an optimal prediction model. A detailed description is given in figure 1.

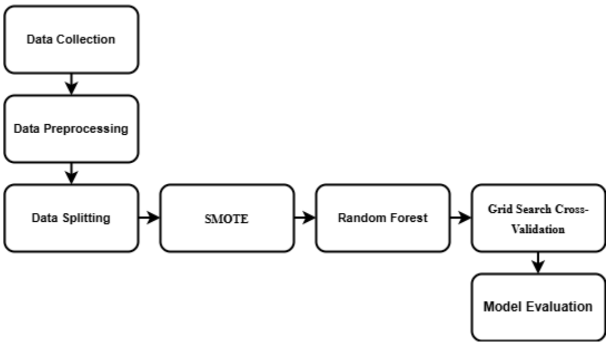

Figure 1. Research Method

This research started with data collection from Kaggle, followed by pre-processing using the IQR method to handle outliers. The data was then split with a 60:40 ratio for training and testing. The Random Forest model was developed and optimized through Grid Search Cross-Validation. Evaluation was performed using accuracy, precision, recall, and F1-score to assess the performance of the model in heart failure classification.

1. Data Collection

   The initial stage in this research is the data retrieval process. This research uses a dataset obtained from the Kaggle platform, which consists of 13 columns, with 12 input feature columns and 1 output (target) column. This dataset includes 300 patient data used for heart failure prediction analysis [16].

1	age	anaemia	creatinine	diabetes	ejection_f	high_bloo	platelets	serum_cre	serum_so	sex	smoking	time	DEATH_EVENT
2	75	0	582	0	20	1	265000	1.9	130	1	0	4	1
3	55	0	7861	0	38	0	263358	1.1	136	1	0	6	1
4	65	0	146	0	20	0	162000	1.3	129	1	1	7	1
5	50	1	111	0	20	0	210000	1.9	137	1	0	7	1
6	65	1	160	1	20	0	327000	2.7	116	0	0	8	1
7	90	1	47	0	40	1	204000	2.1	132	1	1	8	1
8	75	1	246	0	15	0	127000	1.2	137	1	0	10	1
9	60	1	315	1	60	0	454000	1.1	131	1	1	10	1
10	65	0	157	0	65	0	263358	1.5	138	0	0	10	1

Figure 2. Heart Failure Prediction Dataset

2. Data Preprocessing

   Data pre-processing is done to ensure the dataset is in optimal condition before the modeling stage. In this study, no missing values or duplicate data were found, so no further handling was required. Next, outlier detection was performed using boxplots and Interquartile Range (IQR) to identify extreme values that could affect the analysis results. To handle outliers, the Winsorization method was used by replacing the extreme values to the 10th and 90th percentile limits. This approach aims to maintain the stability of the predictive model and reduce the risk of bias due to outliers.
3. Data Splitting
   In heart failure classification using Random Forest algorithm, the dataset is divided into two main parts: 80% for training data used to train the model, and 20% for test data used to measure the performance of the model after the training process.
4. SMOTE
   This study uses SMOTE to handle class imbalance in heart failure classification with Random Forest. The dataset is divided into 80% training data and 20% test data with stratification. SMOTE is applied to the training data to increase the minority class samples, so that the model can learn better without bias. The Random Forest model was trained with the balanced data, and hyperparameter optimization was performed using Grid Search Cross-Validation. Model evaluation uses accuracy, precision, recall, F1-score, and confusion matrix to assess the performance of heart failure prediction.
5. Random Forest
   Random Forest models are trained by building multiple decision trees using random samples and features from the training data. The predictions of each tree are combined through majority voting for classification or averaging for regression. Hyperparameters such as number of trees (n_estimators) and tree depth (max_depth) are tuned to optimize model performance. Evaluation is done using accuracy, precision, recall, and F1-score. Random Forest is generally more accurate and resistant to overfitting than single decision trees.
6. Grid Search Cross Validation
   Tuning hyperparameter dengan Grid Search Cross-Validation digunakan untuk menemukan kombinasi parameter terbaik dalam model Random Forest. Proses ini menguji berbagai nilai, seperti n_estimators (100, 120, 150), max_depth (2, 5, 6, 8, 10), min_samples_split (4, 5, 6), min_samples_leaf (2, 3), random_state (42), dan criterion (entropy). Dengan validasi silang, Grid Search CV memastikan model bekerja optimal berdasarkan metrik evaluasi yang ditentukan.
7. Model Evaluation
   Once the model is trained, evaluation is performed using metrics such as Accuracy, Precision, Recall, and F1-Score. These metrics provide a more in-depth look at the effectiveness of the model in solving binary

classification problems. The following is an explanation of each metric and its formula.

a) Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b) Precision

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

c) *Recall*

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

d) F1-Score

$$F1 - Score = 2x\,\frac{Precision\ x\ Recall}{Precision + Recall} \quad (4)$$

C. Result and Discussion

The scatter plot shows the relationship between patient age and creatinine phosphokinase (CPK) levels with heart failure mortality status. The majority of patients had CPK levels below 1000, and there was no clear pattern of correlation between age and CPK levels. However, older patients appeared to have a higher risk of death. Interestingly, patients with very high CPK levels (>2000) mostly survived, suggesting that CPK levels are not the sole determinant of mortality. Therefore, further analysis with other variables such as blood pressure, ejection fraction, and serum creatinine levels is required.

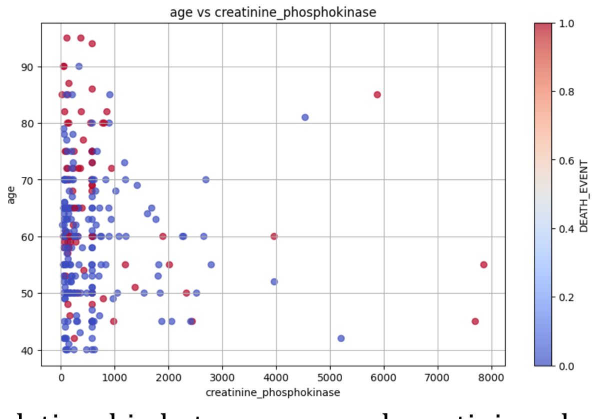

Figure 3. Relationship between age and creatinine phosphokinase

The scatter plot shows the relationship between patient age and ejection fraction, with colors indicating death status due to heart failure. Ejection fraction is an important indicator of heart function, with patients with values below 30% having a higher proportion of deaths. This is in line with the medical fact that low ejection fraction increases the risk of heart failure. In contrast, patients with ejection fraction above 40% survived, although there were still some cases of death. This finding suggests that other factors, such as age or other health conditions, also play a role in survival rates. Overall, this scatter plot confirms that low ejection fraction is associated with a higher risk of death from heart failure.

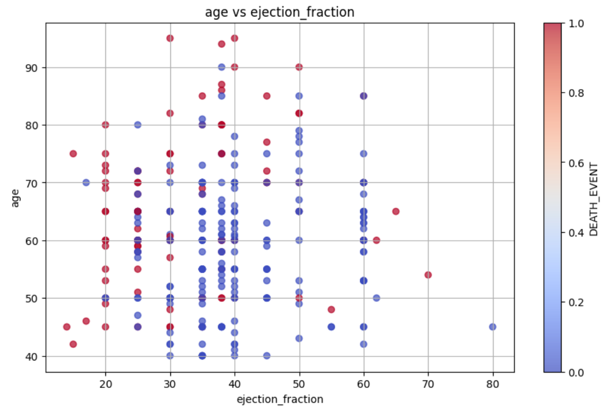

Figure 4. Relationship between age and ejection fraction

The scatter plot shows the relationship between patient age and platelet count, with the color indicating the status of death from heart failure. The distribution of platelet counts was evenly spread across ages with no clear pattern to mortality risk. The majority of patients had platelet counts within the normal range (150,000-350,000), both survivors and deceased, so platelet count is likely not a major factor in heart failure. Patients with high platelet counts (>600,000) were less likely to die, while patients with low platelets did not show a significant mortality trend. Therefore, further analysis is needed by considering other factors such as blood pressure, ejection fraction, and serum creatinine levels.

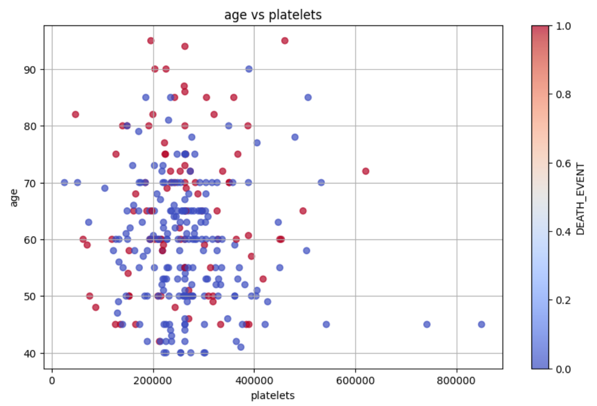

Figure 5. Age relationship of platelets

The scatter plot shows the relationship between patient age and serum creatinine levels, with colors indicating death status from heart failure. Serum creatinine is an indicator of kidney function, with high levels often associated with heart failure risk. Most patients have serum creatinine levels in the range of 0.5-2.5, but patients with higher levels (>2.5) are more likely to die, especially in the elderly. However, some patients with low levels also died, suggesting that other factors such as age and other health conditions played a role. Overall, high serum creatinine levels may increase the risk of death from heart failure, but further analysis considering other health factors is needed.

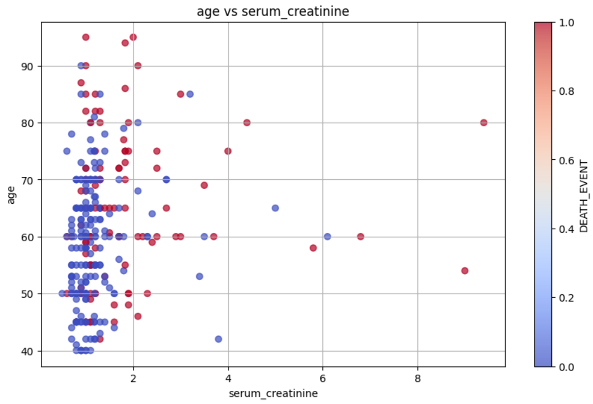

Figure 6. Age relationship of serum creatinine

Scatter plot analysis showed that ejection fraction and serum creatinine were the main factors influencing the risk of death from heart failure. Patients with low ejection fraction (<30%) and high serum creatinine levels (>2.5) had a greater chance of death, especially in the elderly. In contrast, creatinine phosphokinase and platelet count showed no significant pattern in influencing mortality. Therefore, ejection fraction and serum creatinine should be the main focus in heart failure prediction analysis and medical decision-making.

Table 1. *Confusion Matrix*

	Positive Prediction	**Negative Prediction**
Actual Positive	*9 (TP)*	*5 (FN)*
Actual Negative	*3 (FP)*	*32 (TN)*

Confusion Matrix shows the model correctly identified 9 patients with heart failure (TP), but misclassified 5 patients (FN) and 3 patients (FP). A total of 32 patients without heart failure were correctly classified (TN). These results help evaluate metrics such as accuracy, precision, recall, and F1-score to assess the model's performance in detecting heart failure.

Table 2. Model Evaluation Results

Class	**Precision**	**Recall**	**F1-Score**
0	0.97	0.91	0.94
1	0.81	0.93	0.87
Accuracy	-	-	**0.92**

The model achieved 92% accuracy, with a precision of 0.97 for class 0 and 0.81 for class 1, indicating reliability in reducing false positives. Recall of 0.91 (class 0) and 0.93 (class 1) indicated good positive case detection capability. F1-score of 0.94 (class 0) and 0.87 (class 1) shows the balance of precision and recall. These results confirm that the model performs well in heart failure classification.

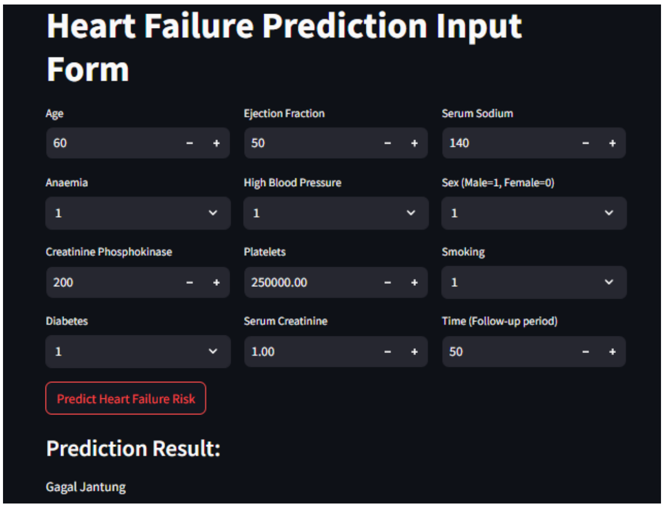

Figure 7. Streamlit

Figure 7 shows the implementation of Streamlit as the web interface for the heart failure disease classification system using the Random Forest algorithm with the Grid Search Cross Validation technique. This application allows users to enter patient data directly through the web interface. The entered data is then processed by the trained model to generate a prediction of the patient's condition.

D. Conclusion

This study successfully developed a heart failure disease classification model using the Random Forest algorithm optimized with Grid Search Cross Validation. The pre-processing process, including outlier handling and data balancing using SMOTE, proved to improve the performance of the model in detecting heart failure risk. Model evaluation showed that Random Forest achieved 92% accuracy, with a good balance between precision and recall. From the analysis, factors such as ejection fraction and serum creatinine have a significant influence on the risk of death from heart failure. In addition, the implementation of the model in a web-based application using Streamlit makes it easier for users to predict patient conditions quickly and efficiently. In the future, this research can be further developed by exploring additional features and using other optimization methods to further improve the accuracy of the model.

E. References

[1] World Health Organization, "Cardiovascular diseases."
[2] Yumi Thesiana, S.Kep.Ners, M.Kep, "Gagal Jantung." [Online]. Available: https://ayosehat.kemkes.go.id/gagal-jantung
[3] dr. Airindya Bella, "Ini Penyebab Gagal Jantung dan Faktor Risikonya."
[4] F. Z. Sudirman, W. Wisudawan, and J. Dase, "Karakteristik Kejadian Mati Mendadak Akibat Penyakit Kardiovaskular," *J. Kesehat. Tambusai*, vol. 5, no. 4, pp. 11099–11109, Nov. 2024, doi: 10.31004/jkt.v5i4.36895.
[5] U. Nagavelli, D. Samanta, and P. Chakraborty, "Machine Learning Technology-Based Heart Disease Detection Models," *J. Healthc. Eng.*, vol. 2022, pp. 1–9, Feb. 2022, doi: 10.1155/2022/7351061.

[6] F. Febby, A. Arjuna, and M. Maryana, “Dukungan Keluarga Berhubungan dengan Kualitas Hidup Pasien Gagal Jantung,” *J. Penelit. Perawat Prof.*, vol. 5, no. 2, pp. 691–702, Mar. 2023, doi: 10.37287/jppp.v5i2.1537.
[7] I. Akbar, F. Supriadi, and D. Indra Junaedi, “PEMANFAATAN MACHINE LEARNING DI BIDANG KESEHATAN,” *JATI J. Mhs. Tek. Inform.*, vol. 9, no. 1, pp. 1744–1749, Jan. 2025, doi: 10.36040/jati.v9i1.12663.
[8] S. Sutrisno, “Sistem Pakar Mendiagnosa Penyakit Gagal Jantung Pada Manusia Dengan Menggunakan Metode Certainty Factor Berbasis Web,” *J. Ilmu Komput. Dan Sist. Inf. JIKOMSI*, vol. 5, no. 1, pp. 20–27, Mar. 2022, doi: 10.55338/jikomsi.v5i1.207.
[9] U. Sunarya and T. Haryanti, “Perbandingan Kinerja Algoritma Optimasi pada Metode Random Forest untuk Deteksi Kegagalan Jantung,” *J. Rekayasa Elektr.*, vol. 18, no. 4, Dec. 2022, doi: 10.17529/jre.v18i4.26981.
[10] Siti Nurmaini, Annisa Darmawahyuni, AdeIriani Sapitri, Muhammad Naufal Rachmatullah, Firdaus, and Bambang Tutuko, *Pengenalan Deep Learning dan Implementasinya*. 2021.
[11] W. Nugraha and M. Syarif, “Evaluasi Performa Algoritma Klasifikasi dalam Prediksi Gagal Jantung: Studi Kasus Dataset Heart Failure Prediction,” *Techno.Com*, vol. 23, no. 4, pp. 897–908, Nov. 2024, doi: 10.62411/tc.v23i4.11685.
[12] A. S. Kirono and Y. Nataliani, “Perbandingan Algoritma Machine Learning dalam Analisis Penyebab Penyakit Gagal Jantung,” *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 10, no. 2, p. 296, Aug. 2024, doi: 10.26418/jp.v10i2.78369.
[13] 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 J. Ilm. Mat.*, vol. 10, no. 2, pp. 258–268, Jul. 2022, doi: 10.26740/mathunesa.v10n2.p258-268.
[14] N. Selayanti, S. A. Putri, M. Kristanaya, M. P. Azzahra, M. G. Navsih, and K. M. Hindrayani, “Penerapan Machine Learning Algoritma Random Forest Untuk Prediksi Penyakit Jantung,” *Pros. Semin. Nas. SAINS DATA*, vol. 4, no. 1, pp. 895–906, Oct. 2024, doi: 10.33005/senada.v4i1.376.
[15] F. Duran, F. Wijaya, Y. R. Hulu, M. Harahap, and A. Prabowo, “PERBANDINGAN KINERJA ALGORITMA RANDOM FOREST CLASSIFIER DAN LIGHTGBM CLASSIFIER UNTUK PREDIKSI PENYAKIT JANTUNG,” *Data Sci. Indones. DSI*, vol. 3, no. 2, pp. 98–103, May 2024, doi: 10.47709/dsi.v3i2.3831.
[16] Bhavik Jikadara, [Online]. Available: https://www.kaggle.com/datasets/bhavikjikadara/heart-failure-prediction