BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

Analysis of the Effect of SMOTE on the Performance of Random Forest in Classifying the Risk of Death in Heart Failure Patients

Alyiza Dwi Ningtyas[1,*], Rama Prameswara Ritonga[1], Rizka Salsabila Riadi[1], Arif Hamied Nababan[2], Surizar Rahmi Danur[3]

[1]Department of Computer and Informatics Engineering, Multimedia Graphic Engineering Technology Study Program, Politeknik Negeri Medan, Medan, Indonesia
[2]Department of Computer and Informatics Engineering, Software Engineering Technology, Politeknik Negeri Medan, Medan, Indonesia
[3]Faculty of Science and Technology, Information Systems, Institut Teknologi dan Bisnis Indonesia, Medan, Indonesia
Email: [1,*]alyizadwi@polmed.ac.id, [2]ramaprameswara@polmed.ac.id, [3]rizkasalsabila@polmed.ac.id, [4]arifhamied@polmed.ac.id, [5]surizar.rdanur@gmail.com
Correspondence Author Email: alyizadwi@polmed.ac.id
Received: **14/08/2026**, Revised: **30/10/2026**, Accepted: **31/10/2026**, Published**: 31/10/2026**

Abstract−Class distribution imbalance in the heart failure patient dataset has the potential to affect the ability of the classification model to identify patients experiencing death events. This study aims to analyze the effect of the Synthetic Minority Over-sampling Technique (SMOTE) application on the performance of the Random Forest algorithm in classifying death events in heart failure patients. The dataset used is the Heart Failure Clinical Records Dataset consisting of 299 patient data, with 12 predictor attributes and one target attribute, namely DEATH EVENT. Data processing was carried out using the Python programming language through the Google Colab platform. The research stages include data preparation, dataset division with a ratio of 80:20, application of SMOTE to training data, formation of a Random Forest model, and performance evaluation using accuracy, precision, recall, and F1-score metrics. In addition, Stratified 5-Fold Cross-Validation was applied as an additional evaluation to analyze model performance on several data divisions. The test results with an 80:20 data split showed that the model before and after applying SMOTE produced an accuracy value of 83.33%, a positive class precision of 80.00%, a recall of 63.16%, and an F1-score of 70.59%. Meanwhile, the results of Stratified 5-Fold Cross-Validation showed that the application of SMOTE increased the average recall value from 69.74% to 76.00% and the F1-score from 73.34% to 74.75%. However, the average accuracy value decreased from 83.93% to 83.60%, while the precision decreased from 77.93% to 73.74%. These findings indicate that the application of SMOTE has the potential to improve the ability of Random Forest in identifying patients with the label DEATH EVENT = 1, although it does not improve all evaluation metrics. This study is expected to provide an empirical contribution in evaluating the effectiveness of SMOTE in classifying imbalanced medical data, especially in improving the detection ability of minority classes.

Keywords: Heart Failure; Random Forest; SMOTE; Classification; Stratified 5-Fold Cross-Validation

1. INTRODUCTION

Heart failure is a clinical condition in which the heart is unable to pump blood optimally to meet the body's needs. This condition can lead to serious complications and increase the risk of death, making identifying factors related to patient outcomes an important part of health research. Clinical characteristics, such as age, ejection fraction , serum creatinine level, serum sodium level, high blood pressure, and history of comorbidities, can be used to analyze patterns related to patient mortality. Developments in information technology allow for the processing of clinical data using machine learning approaches to build classification models capable of recognizing patterns based on historical data. Therefore, the classification of mortality risk in heart failure patients is a relevant topic to be studied using a data mining approach, while still noting that the results of computational models require validation before being used in clinical decision-making [1].

A critical issue in the classification of mortality risk in heart failure patients is class imbalance in the dataset. This imbalance occurs when the number of observations in the non-death class is greater than the number of observations in the death class. This condition can affect the model's learning process, making it more likely to recognize the majority class than the minority class. Consequently, a high accuracy value does not necessarily indicate that the model is optimally capable of identifying patients who experience death. In the context of this study, the ability to recognize the death class is important because misclassification of this class can reduce the model's analytical utility. One approach that can be used to address this issue is the Synthetic Minority Oversampling Technique (SMOTE), a data balancing technique that generates synthetic samples for the minority class. However, the application of SMOTE needs to be evaluated to determine whether the addition of synthetic samples actually improves classification performance.

In data mining, classification algorithms are used to learn the relationship between predictor attributes and target classes based on historical data. One algorithm that can be used is Random Forest, an ensemble learning method that combines predictions from a number of decision trees to produce a final classification. Random Forest is relevant for clinical data because it can learn complex patterns of relationships between various patient attributes and does not require a linear relationship between predictors and targets. However, Random Forest's ability to classify data does not automatically eliminate the impact of class imbalance [2]. Therefore, this study examines the application of SMOTE as a preprocessing stage before Random Forest training. The focus of the study is not only to produce a classification model, but also to evaluate the change in Random Forest performance when trained using original data compared to training data that has been balanced using SMOTE [3],[4],[5].

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

SMOTE was chosen because it generates synthetic samples through interpolation between minority class samples and their nearest neighbors, thus differing from methods that simply duplicate existing observations [6]. This approach is expected to help the model learn the characteristics of the minority class more adequately. Meanwhile, Random Forest was chosen because it combines multiple decision trees, enabling it to learn diverse data patterns and reduce reliance on a single decision tree [7],[8],[9],[10],[11]. The combination of SMOTE and Random Forest allows the effect of data balancing to be observed while maintaining the same classification algorithm in both experimental scenarios. To avoid conclusions based solely on a single indicator, model performance will be assessed using accuracy , precision , recall , and F1-score , along with a confusion matrix and ROC-AUC if appropriate for the evaluation design. This comparison is important because SMOTE does not always improve all metrics and has the potential to generate synthetic samples that overlap with other classes [12],[13].

This study aims to analyze the effect of SMOTE implementation on the performance of the Random Forest algorithm in classifying the risk of death in heart failure patients. The dataset used is the Heart Failure Clinical Records Dataset consisting of 299 patient records with 13 predictor attributes and one target attribute, namely DEATH EVENT . The target variable represents the presence or absence of death during the observation period in the dataset, not the certainty of the patient's risk of death in the future. The class distribution in the dataset consists of 203 observations with a DEATH EVENT value of 0 and 96 observations with a DEATH EVENT value of 1, resulting in a class imbalance that serves as the basis for SMOTE testing. The study was conducted through two scenarios, namely Random Forest without SMOTE and Random Forest with SMOTE. Both scenarios will use consistent data sharing and model configurations so that performance changes can be compared more objectively. Through this design, the study is expected to demonstrate whether SMOTE provides performance improvements, especially in recognizing patient classes that experience death events.

Data processing and experiments were conducted using the Python programming language because it provides libraries that support dataset management, class balancing, modeling, and evaluation of classification results. Pandas and NumPy were used to read and process the data, Scikit-learn was used to split the dataset, build the Random Forest model, and calculate evaluation metrics, while imbalanced-learn was used to apply SMOTE. The dataset was divided into training and test data using a stratified split approach to maintain the proportion of classes represented in each part. SMOTE was applied only to the training data, while the test data retained its original distribution. This procedure was implemented to prevent data leakage and ensure that evaluations were carried out on data not used in the synthetic sample generation process. Next, the results of the two scenarios were compared through metric tables and visualizations so that performance changes could be systematically analyzed and reproduced through program code.

A study conducted by Fatah et al in 2025 entitled “Application of Random Forest Algorithm and SMOTE Technique for Predicting Death Due to Heart Failure Using RapidMiner” discusses the application of the Random Forest algorithm combined with the SMOTE technique to overcome class imbalance in a dataset of heart failure patients [14]. Furthermore, a study by Scrutinio et al published in 2025 compared a number of machine learning models, including Random Forest, to predict mortality in heart failure patients. The study showed that machine learning models can provide good discrimination ability, but the results also emphasized the importance of comparing models with statistical approaches using comparable variables [15]. A study conducted by Faradeya et al in 2025 discussed the Classification of Heart Failure Disease Using the Naive Bayes Algorithm. The results of the study discuss the application of the Naive Bayes algorithm to predict the risk of heart failure based on patient clinical data. This research was motivated by the increasing prevalence of heart failure and the complexity of its clinical impact [16].

Based on previous research, the use of Random Forest for classifying deaths in heart failure patients and the application of SMOTE to address class imbalance have become relevant approaches in data-driven healthcare research. However, this study specifically focuses on analyzing the effect of SMOTE by comparing Random Forest performance across two experimental scenarios using consistent datasets, data splits, and model configurations. The study's contribution lies in the empirical evaluation of changes in classification performance after class balancing, particularly in terms of the ability to recognize death event classes. Furthermore, the use of multiple evaluation metrics allows the study to identify whether improvements in one metric are accompanied by changes in the other. The results are expected to serve as a methodological reference for data mining research in healthcare, particularly in determining the need for data imbalance management before building a Random Forest model. The resulting model remains positioned as an experimental analysis tool, not a substitute for diagnosis or clinical decision-making.

2. RESEARCH METHODOLOGY

2.1 Random Forest Algorithm

Random Forest is an ensemble learning-based machine learning algorithm that combines multiple decision trees to produce a final prediction. This algorithm builds each tree using a randomly selected data sample and subset of attributes, then combines the prediction results of all trees to determine the final class [9],[5]. In classification problems, class determination is generally done through a majority voting mechanism. Combining multiple decision trees aims to improve the model's generalization ability and reduce the risk of overfitting compared to using a single decision tree. In this study, Random Forest was used to classify death events in heart failure patients based on clinical characteristics contained in the dataset [17]. Attributes such as age, ejection fraction, serum creatinine level, serum sodium level, and other clinical

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2661
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

characteristics were used as predictor variables, while DEATH EVENT was the target variable. The performance of Random Forest was evaluated in two scenarios, namely without the application of SMOTE and with the application of SMOTE on the training data. This comparison is used to determine changes in classification ability after class imbalance correction [4],[18],[19],[20],[21].

2.2 Synthetic Minority Over-sampling Technique (SMOTE)

Synthetic Minority Over-sampling Technique (SMOTE) is a data balancing technique used to address class imbalance by generating synthetic samples in the minority class. Unlike random oversampling methods that duplicate existing samples, SMOTE generates new samples through interpolation between a minority sample and one of its nearest neighbors, also from the minority class. This technique is an approach to improving the representation of the minority class in the classification learning process [22],[23],[24], [25].

2.3 Research Stages

The research completion process can be seen in Figure 1. The research stages begin with loading the Heart Failure Clinical Records Dataset, then conducting data exploration to examine the structure, missing values, and class distribution. Next, data preprocessing, attribute selection, and the division of training data and test data with a ratio of 80:20 are carried out. Modeling is carried out through two scenarios, namely the Random Forest algorithm without SMOTE and Random Forest with SMOTE applied to the training data to address class imbalance. Both models are then evaluated using the metrics of accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC where appropriate. The evaluation results are compared to analyze the effect of applying SMOTE on the performance of Random Forest in classifying death events in heart failure patients.

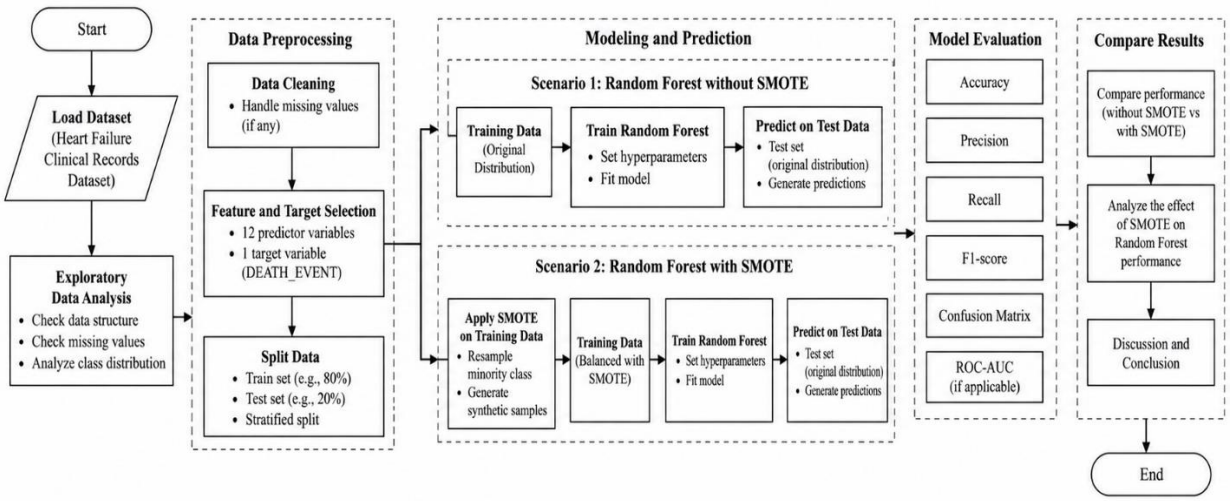


Figure 1. Research Stages

2.4 Research Data Attribute Description

This study uses the Heart Failure Clinical Records Dataset containing clinical records of heart failure patients. The dataset is used to analyze the effect of the Synthetic Minority Over-sampling Technique (SMOTE) on the performance of the Random Forest algorithm in classifying death events during the observation period. The dataset contains clinical attributes as predictor variables and the DEATH EVENT attribute as the target variable. Based on file inspection using python programming, the dataset consists of 299 observations and 13 attributes. The number of detected empty values is 0. The target variable DEATH EVENT has two classes, namely 0 (no death occurred during the observation period) and 1 (death occurred during the observation period). The following is shown in Table 1 which is a description of the attributes used in this study.

Table 1. Description of Research Dataset Attributes

No.	Attribute	Data Types	Information
1	Age	Numeric	Patient Age (Years).
2	Anemia	Binary	Anemia Status (0 = No, 1 = Yes).
3	Creatinine Phosphokinase	Numeric	Creatine Phosphokinase Levels in Blood (Mcg/L).
4	Diabetes	Binary	Diabetes Status (0 = No, 1 = Yes).
5	Ejection Fraction	Numeric	Cardiac Ejection Fraction Percentage (%).
6	High Blood Pressure	Binary	High Blood Pressure Status (0 = No, 1 = Yes).
7	Platelets	Numeric	Number of Platelets in Blood (Kiloplatelets/Ml).
8	Serum Creatinine	Numeric	Serum Creatinine Levels (Mg/Dl).

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2662
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

No.	Attribute	Data Types	Information
9	Sodium Serum	Numeric	Serum Sodium Level (Meq/L).
10	Sex	Binary	Gender (0 = Female, 1 = Male).
11	Smoking	Binary	Smoking Status (0 = No, 1 = Yes).
12	Time	Numeric	Patient Follow-up Period (Days).
13	DEATH EVENT	Binary	Target Variable: 0 = No Death Occurred During Observation; 1 = Death Occurred During Observation.

Source: Heart Failure Clinical Records Dataset, processed by researchers.

Based on Table 1, the dataset has 12 predictor attributes and one target attribute, namely DEATH EVENT. The predictor attributes describe the patient's demographic and clinical characteristics, while DEATH EVENT is used as a class label predicted by the Random Forest model.

3. RESULT AND DISCUSSION

3.1 Dataset

The raw data used in this study comes from the Heart Failure Clinical Records Dataset, which contains clinical attributes of heart failure patients and the target variable DEATH EVENT. To provide an overview of the structure and characteristics of the data before the preprocessing and modeling stages, Table 2 presents several examples of the raw data used in this study.

Table 2. Data

No	Age	Anemia	Creatinine Phosphokinase	Diabetes	Ejection Fraction	High Blood Pressure	...	...	Serum Sodium	Sex	Smoking	Time	Death Event
1	75	0	582	0	20	1	...	...	130	1	0	4	1
2	55	0	7861	0	38	0	...	...	136	1	0	6	1
3	65	0	146	0	20	0	...	...	129	1	1	7	1
4	50	1	111	0	20	0	...	...	137	1	0	7	1
5	65	1	160	1	20	0	...	...	116	0	0	8	1
6	90	1	47	0	40	1	...	...	132	1	1	8	1
7	75	1	246	0	15	0	...	...	137	1	0	10	1
8	60	1	315	1	60	0	...	...	131	1	1	10	1
9	65	0	157	0	65	0	...	...	138	0	0	10	1
10	80	1	123	0	35	1	...	...	133	1	1	10	1
11	75	1	81	0	38	1	...	...	131	1	1	10	1
12	62	0	231	0	25	1	...	...	140	1	1	10	1
13	45	1	981	0	30	0	...	...	137	1	0	11	1
14	50	1	168	0	38	1	...	...	137	1	0	11	1
…	…	…	…	…	…	…	…	…	…	…	…	…	…
299	50	0	196	0	45	0	...	...	136	1	1	285	0

3.2 Class Distribution Analysis and Data Imbalance

Class distribution analysis was performed to determine the proportion of each class in the target variable DEATH_EVENT and to identify any data imbalances. The results of the class distribution analysis on the heart failure patient dataset are presented in Table 3.

Table 3. Class Distribution

Class DEATH EVENT	Information	Amount of Data	Percentage (%)
0	Did Not Experience Any Deaths	203	67.89
1	Experiencing a Death	96	32.11

Based on the analysis results, the dataset consists of 299 patient data with 13 attributes. The class distribution shows that 203 data (67.89%) belong to the DEATH EVENT = 0 class, namely patients who did not experience a death event during the follow-up period, while 96 data (32.11%) belong to the DEATH EVENT = 1 class, namely patients who experienced a death event during that period. With a majority to minority class ratio of 2.11:1, the dataset shows a class imbalance that needs attention. Therefore, SMOTE was applied to the training data to test whether class balancing can

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2663
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

improve the performance of the Random Forest algorithm. The following can be seen in Figure 2 Death Event Class Distribution Diagram.

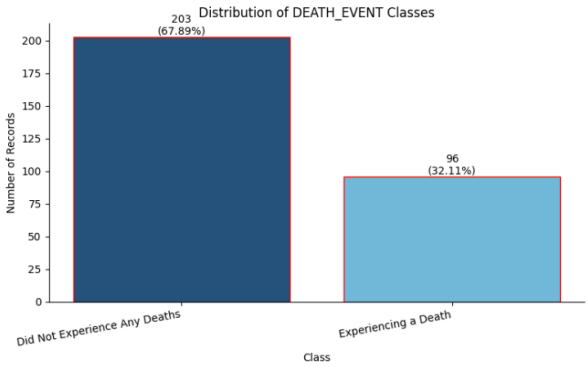


Figure 2. Death Event Class Distribution Diagram

3.3 Preprocessing and Dataset Sharing

The data preprocessing stage was carried out to ensure that the dataset used had the appropriate quality and structure for the classification process. At this stage, checks were made for missing values , duplicate data, data types, and the validity of the DEATH EVENT target attribute . Next, the dataset was divided into 80% training data and 20% testing data using a *stratified sampling technique* to maintain the class proportions in both parts of the data. This dataset division served as the basis for testing the performance of the Random Forest algorithm before and after the application of SMOTE. The results of the preprocessing and dataset division are presented in Table 4 and Figure 3 below.

Table 4. Dataset Distribution

Subset	DEATH EVENT	Information	Amount Data	Percentage (%)
All Data	0	No deaths occurred during follow-up.	203	67.89%
All Data	1	Death occurred during follow-up	96	32.11%
Training Data	0	No deaths occurred during follow-up.	162	67.78%
Training Data	1	Death occurred during follow-up	77	32.22%
Test Data	0	No deaths occurred during follow-up.	41	68.33%
Test Data	1	Death occurred during follow-up	19	31.67%

To determine the class distribution in the heart failure patient dataset, an analysis was performed on the total data in the entire dataset, training data, and test data based on the target attribute DEATH_EVENT . This analysis aims to identify the imbalance in the number of data between class 0 (no deaths during the observation period) and class 1 (deaths occurred during the observation period). The class distribution before applying SMOTE is shown in Figure 3.

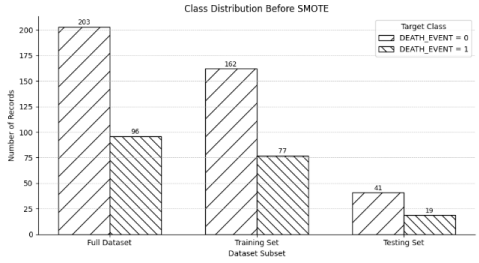


Figure 3. Dataset Class Distribution Before SMOTE Application

3.4 Implementation of SMOTE on Training Data

The Synthetic Minority Oversampling Technique (SMOTE) was applied to address class imbalance in the training data by generating synthetic samples in the minority class. This process was performed after the dataset was divided into 80% training data and 20% test data. SMOTE was applied only to the training data to maintain the distribution of the test data and allow for objective evaluation of the Random Forest model's performance. A comparison of the class distribution before and after SMOTE application is presented in Table 5 and Figure 4 below.

Table 5. Comparison of Training Data Class Distribution Before and After SMOTE Implementation

Target Class	Before SMOTE	After SMOTE	Synthetic Samples Added
DEATH EVENT = 0	162	162	0
DEATH EVENT = 1	77	162	85

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2664
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

To display the distribution image of the training data classes before and after applying SMOTE, see Figure 4 below.

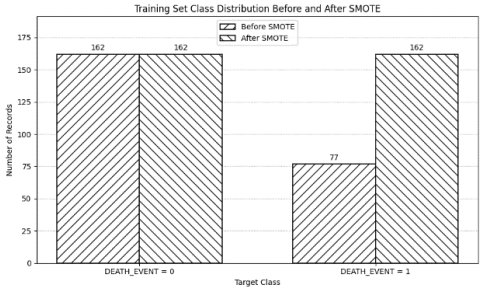


Figure 4. Distribution of Training Data Classes Before and After Implementing SMOTE

Based on the results of the SMOTE application, changes in the number of samples in each class can be observed by comparing the class distribution before and after the oversampling process. The addition of synthetic samples to the minority class aims to balance the class distribution in the training data so that the Random Forest model has a better opportunity to learn the characteristics of both classes. The effect of the SMOTE application on classification performance is then analyzed by comparing the evaluation results of the Random Forest model before and after SMOTE.

3.5 Random Forest Classification Results with SMOTE

The Random Forest algorithm was tested using the Synthetic Minority Over-sampling Technique (SMOTE) to analyze the model's performance in classifying the risk of death in heart failure patients on a dataset with an unbalanced class distribution. SMOTE was applied to the training data to balance the number of samples in the majority and minority classes, while the test data retained its original distribution. Model performance was evaluated using a confusion matrix, accuracy, precision, recall, and F1-score. The test results were then analyzed to determine the model's ability to classify each class and to serve as a basis for evaluating the effect of SMOTE on Random Forest performance. The results of the Random Forest classification evaluation with SMOTE are presented in Table 7.

Table 7. Random Forest Classification Using SMOTE

	precision	recall	f1-score	support
No Death Event (0)	0.8444	0.9268	0.8837	41
Death Event (1)	0.8000	0.6315	0.7058	19
accuracy	0.8333	0.8333	0.8333	60
macro avg	0.8222	0.7792	0.7948	60
weighted average	0.8304	0.8333	0.8274	60

Based on Table 7, the Random Forest model with SMOTE obtained an accuracy value of 83.33%. In the death class (1), the model produced a precision of 80.00%, a recall of 63.16%, and an F1-score of 70.59%. These results indicate that the model is able to identify most patients who did not experience a death event, but still does not optimally detect all patients in the death class. These evaluation values are the same as the results of the Random Forest test without SMOTE on an 80:20 data split, so the application of SMOTE has not shown any changes in predictions in this test scenario. To determine the number of correct and incorrect predictions in each class, the evaluation results of the Random Forest model with SMOTE are also presented in the form of a confusion matrix, as shown in Figure 5.

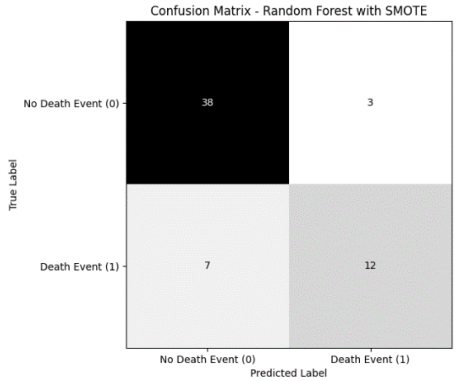


Figure 5. Confusion Matrix

3.6 Discussion of the Effect of SMOTE on Random Forest Performance

Based on the test results using an 80:20 data split, the Random Forest model before and after SMOTE application produced the same evaluation scores. Both models achieved an accuracy of 83.33%, a class 1 precision of 80.00%, a class

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2665
This work is licensed under a Creative Commons Attribution 4.0 International License.

BULLETIN OF COMPUTER SCIENCE RESEARCH
ISSN 2774-3659 (Media Online)
Vol 6, No 6, October 2026 | Hal 2660-2668
https://hostjournals.com/bulletincsr
DOI: 10.47065/bulletincsr.v6i6.1341

1 recall of 63.16%, and a class 1 F1-score of 70.59%. These similar values indicate that the application of SMOTE has not changed the model's prediction results on the used data split. This condition may occur because the addition of synthetic samples to the minority class in the training data does not always result in changes in classification decisions on the test data. Thus, the success of SMOTE in balancing the class distribution in the training data does not automatically increase the model's evaluation score.

Based on the confusion matrix results, the model successfully classified 38 class 0 data and 12 class 1 data correctly. Meanwhile, there were 3 class 0 data that were incorrectly predicted as class 1 and 7 class 1 data that were incorrectly predicted as class 0. These results indicate that the model still has difficulty in identifying some patient data that experienced a death event. This is also reflected in the class 1 recall value of 63.16%, which indicates that the model successfully identified 12 of the total 19 actual data in that class. Therefore, the relatively high accuracy value does not fully describe the model's ability to detect minority classes.

To obtain a more comprehensive picture of the effect of SMOTE, an additional evaluation was conducted using Stratified 5-Fold Cross-Validation. In this test, the data was divided into five parts while maintaining class proportions, so that each part was used alternately as validation data. SMOTE was applied only to the training data in each fold to avoid data leakage. The evaluation results using the mean and standard deviation values of each performance metric are presented in Table 10.

Table 10. Comparison of Random Forest Performance Using 5-Fold Cross-Validation

Evaluation Metrics	Without SMOTE (Mean)	Without SMOTE (Std)	With SMOTE (Mean)	With SMOTE (Std)
Accuracy	83.93%	3.96%	83.60%	2.32%
Precision	77.93%	6.63%	73.74%	3.00%
Recall	69.74%	9.78%	76.00%	6.43%
F1-score	73.34%	7.53%	74.75%	4.24%

Based on Table 10, the application of SMOTE resulted in different performance changes for each evaluation metric. The average accuracy value decreased from 83.93% to 83.60%, or by 0.33 percentage points. Precision also decreased from 77.93% to 73.74%, or by 4.19 percentage points. These results indicate that the application of SMOTE did not improve all aspects of Random Forest performance. The decrease in precision indicates that the proportion of correct positive predictions decreased in the average evaluation results.

In contrast, the average recall value increased from 69.74% to 76.00%, or by 6.26 percentage points. This increase indicates that the model with SMOTE has a better ability to recognize data from the positive class, namely patients with DEATH EVENT = 1. In addition, the average F1-score value increased from 73.34% to 74.75%, or by 1.41 percentage points. This increase in the F1-score indicates an improvement in the balance between precision and recall for the positive class. Thus, although the application of SMOTE slightly decreases accuracy and precision, this technique has the potential to improve the model's ability to detect minority classes.

The difference in results between the 80:20 and 5-Fold Cross-Validation tests needs to be considered in interpreting the research. In the 80:20 test, both models produced identical predictions on the test data, leaving all evaluation metrics unchanged. Meanwhile, in the 5-Fold Cross-Validation test, the average evaluation scores showed a change after applying SMOTE. This difference may occur because the two scenarios use different data allocation and evaluation processes. Therefore, the results of the two scenarios cannot be considered contradictory but should be understood based on the characteristics of each testing process.

In addition to the mean values, the standard deviations in Table 10 also show changes in performance variation between folds. The model with SMOTE has a lower standard deviation across all metrics than the model without SMOTE. For example, the standard deviation of recall decreased from 9.78% to 6.43%, while the standard deviation of F1-score decreased from 7.53% to 4.24%. These results indicate that the evaluation values of the model with SMOTE have smaller variations across the five folds used. However, the lower variation does not yet indicate a statistically significant difference in performance.

In conclusion, the study shows that the effect of SMOTE on Random Forest performance depends on the evaluation scenario and the metrics used. At an 80:20 data split, SMOTE did not significantly change the model's classification results. However, based on Stratified 5-Fold Cross-Validation, SMOTE resulted in an increase in the average recall and F1-score of the positive class, although accompanied by a decrease in accuracy and precision. Therefore, the effectiveness of SMOTE in this study was more evident in the model's ability to detect minority classes than in its overall accuracy. This finding emphasizes that evaluation of class imbalance management should not only focus on

The findings of this study are consistent with Fatah et al, which demonstrated the relevance of applying Random Forest and SMOTE to heart failure data with class imbalance. However, the present study specifically evaluates the effect of SMOTE by comparing model performance before and after its application. The results show that SMOTE improved recall from 69.74% to 76.00% and F1-score from 73.34% to 74.75% based on 5-Fold Cross-Validation [14].

Furthermore, in line with Scrutinio et al, the results indicate the importance of evaluating classification models using multiple performance metrics. Although SMOTE slightly decreased accuracy, it improved the model's ability to

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2666
This work is licensed under a Creative Commons Attribution 4.0 International License.

identify the death event class. Therefore, this study emphasizes that the effectiveness of SMOTE should not be assessed based solely on accuracy but also on recall and F1-score, especially for imbalanced medical data [15].

4. CONCLUSION

Based on the research results, the application of Synthetic Minority Over-sampling Technique (SMOTE) to the Random Forest algorithm showed different effects based on the test scenario used. In the test with an 80:20 data split, the model before and after the application of SMOTE produced an accuracy value of 83.33%, a precision of 80.00%, a recall of 63.16%, and an F1-score of 70.59%, so there was no change in performance based on these evaluation metrics. Therefore, Stratified 5-Fold Cross-Validation was added as a further evaluation to determine the effect of SMOTE on several different data splits and obtain a more comprehensive picture of the model's performance and consistency. The evaluation results show that the application of SMOTE increases the average recall value from 69.74% to 76.00% and the F1-score from 73.34% to 74.75%, but decreases the accuracy from 83.93% to 83.60% and the precision from 77.93% to 73.74%. These results indicate that SMOTE has the potential to improve the model's ability to identify patients with the label DEATH EVENT = 1 , although it has not improved all aspects of Random Forest performance. Thus, the effect of SMOTE needs to be evaluated based on various metrics and testing scenarios to obtain a more comprehensive picture of its effectiveness in handling class imbalance in the heart failure patient dataset.

5. DECLARATIONS

5.1 Author Contributions

Conceptualization: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Methodology: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Software: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Validation: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Formal Analysis: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Investigation: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Resources: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Data Curation: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Writing Original Draft Preparation: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Writing Review and Editing: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
Visualization: A. D. N., R. P. R, R. S. R., A. H. N., and S. R. D.;
All authors have read and agreed to the published version of the manuscript.

5.2 Data Availability

The data presented in this study are available on request from the corresponding author.

5.3 Funding

The research was funded by the authors' personal resources.

5.4 Institutional Review Board Statement

Not applicable.

5.5 Informed Consent Statement

Not applicable.

5.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

5.7 Generative AI and AI-assisted Technologies in The Writing Process

The authors used Generative AI to improve the writing clarity of this paper. They reviewed and edited the AI-assisted content and took full responsibility for the final publication.

REFERENCES

[1] J. P. Anggraini and A. Desiani, "Comparison of Random Forest and Extreme Gradient Boosting (XGBoost) Algorithms in the Classification of Heart Failure Diseases," *Computing J. Sist. Computing.*, vol. 14, no. 2, pp. 149–157, 2025.
[2] S. K. N. A. Putri, I. Jumiatin, I. Sulistia, N. A. B. Saputra, and N. Wiranda, "Implementation of Hyperparameter Tuning for Classification Models in Heart Disease Risk Prediction," *MALCOM Indonesia. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 4, pp.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*.
This work is licensed under a Creative Commons Attribution 4.0 International License.

1181–1189, 2025.
[3] K. Wirastuti, N. S. Riasari, D. Djannah, and M. Silviana, "Stroke Prevention Efforts through Stroke Risk Score Screening with Counseling Intervention and Examination of Stroke Risk Factors in Bojong Salaman Village, South Pusponjolo District, West Semarang," *J. ABDIMAS-KU J. Servant. Masy. Doctor.*, vol. 2, no. 1, pp. 23–29, Jan. 2023, doi: 10.30659/abdimasku.2.1.23-29.
[4] M. N. Sidqi, "Comparison of Performance of Random Forest and Decision Tree Algorithms in Classification of Heart Diseases," *J. Art. Intell. Multimed. Informatics*, vol. 3, no. 1, pp. 1–14, 2026.
[5] A. Guardi, Z. Saputra, and F. Budiman, "Comparative Analysis of Random Forest and XGBoost Models for Classification of Heart Disease Based on Clinical Data," *JURIKOM (Journal of Ris. Computer)*, vol. 13, no. 1, pp. 254–264, 2026, doi: 10.30865/jurikom.v13i1.9427.
[6] A. Rohman and S. Mujiyono, "Performance Analysis of Machine Learning Algorithms for Heart Disease Prediction Using Integrated Data Preprocessing Methods," *Multimatrix J. Computational Science.*, vol. 5, no. 1, pp. 18–22, 2023.
[7] C. Arafat *et al.*, "Comparison of Random Forest and Xgboost Algorithms for Classification of Heart Disease Based on Medical Data," *J. Fasilcom*, vol. 15, no. 2, pp. 430–435, 2025.
[8] P. Dewandra, I. G. Susrama, M. Diyasa, A. D. Rahajoe, and A. P. Previously, "Comparison of Random Forest and XGBoost Algorithms in the Classification of Stunting Status in Toddlers Based on Anthropometric Data," *Semin. Nas. Inform. Defending the State ISSN*, vol. 6, pp. 108–112, 2026.
[9] A. Setyawan, N. Sulistianingsih, and R. Rismayati, "Comparison of Machine Learning Algorithms for Heart Failure Detection Based on RFECV Feature Selection and SMOTE Data Balancing," *Semin. Nas. Corisindo*, vol. 3, no. September 2025, pp. 622–628, 2025.
[10] N. L. Desmita, S. Lonang, and D. T. Kumoro, "Comparative Analysis Of Decision Tree And Random Forest Algorithms For Predicting Diabetes Mellitus," *SainsTech Innov. J.*, vol. 8, 2025.
[11] E. M. Z. Darmawan and A. Fauzan Dianta, "Implementation of GridSearchCV Hyperparameter Optimization in a Heart Attack Prediction System Using SVM," *Technological. J. Ilm. Sist. Inf.*, vol. 13, no. 1, pp. 8–15, 2023.
[12] R. Ridwan, H. H. Handayani, S. A. P. Lestari, and Y. Cahyana, "Performance Evaluation of Random Forest and Gradient Boosting Algorithms for Classification of Heart Disease," *J. Komtika (Computing and Inform.*, vol. 9, no. 1, pp. 112–124, 2025.
[13] U. Hasanah, A. M. Soleh, and K. Sadik, "Effect of Random Under sampling, Oversampling, and SMOTE on the Performance of Cardiovascular Disease Prediction Models," *J. Matt. Stat. and Computing*, vol. 21, no. 1, pp. 88–102, 2024.
[14] Z. Fatah and I. Addewiyah, "Application of Random Forest Algorithm and SMOTE Technique for Prediction of Death Due to Heart Failure Using RapidMiner," *J. Jamastika*, vol. 4, 2025.
[15] D. Scrutiny *et al.*, "Prediction of mortality in heart failure by machine learning. Comparison with statistical modeling," *Eur. J. Intern. Med.*, vol. 133, pp. 106–112, Mar. 2025, doi: 10.1016/j.ejim.2025.01.020.
[16] M. A. Faradeya and E. R. Subhiyakto, "Classification of Heart Failure Disease Using a Naive Bayes Algorithm," *J. Algorithm. Technology Engineer. Garut*, pp. 115–127, 2025, doi: 10.33364/algorithm/v.22-1.2178.
[17] M. Armin, N. Arifin, and S. A. Rachmini, "Application of Random Forest Algorithm with Recursive Feature Elimination (RFE) Selection for Classification of Heart Failure," *J. Sist. Computing. and Artificial Intelligence*, vol. IX, pp. 303–311, 2026.
[18] M. R. A. Pramudya, C. Arafat, M. C. Ramadhan, F. A. Jafar, and E. Ismanto, "Comparison of Random Forest and Xgboost Algorithms for Classification of Heart Disease Based on Medical Data," *J. Fasilkom*, vol. 15, no. 2, pp. 430–435, 2025.
[19] F. Ferdianto, H. R. Ardiyanto, R. Alkazim, and F. Fathoni, "Prediction Of Heart Disease With Variables Of Metabolic Disease History And Smoking Habits Using Random Forest And Xgboost Algorithms," *JATI (Journal of Mhs. Tek. Inform.*, vol. 10, no. 4, pp. 7255–7260, 2026.
[20] S. A. S. Mola, Y. C. Luttu, and D. N. Rumlaklak, "Comparison of Machine Learning Methods in the Sentiment Analysis of InDriver Application User Comments on Unbalanced Datasets," *J. Sist. Business Inf.*, vol. 14, no. 3, pp. 247–255, 2024, doi: 10.21456/vol14iss3pp247-255.
[21] M. Metode, R. Forest, C. Putri, W. Kase, S. Yulianto, and J. Prasetyo, "Analysis of Stunting Risk Factors in Toddlers in Kesetnana Village," *J. Indonesian. Management. Inform. and Commun.*, vol. 6, no. 3, pp. 1556–1566, 2025.
[22] S. Lestanti *et al.*, "Stunting Classification Using Random Forest Classifier," *JATI (Journal of Student Tech. Inform.*, vol. 7, no. 6, pp. 3846–3851, 2023.
[23] A. Fitri, I. Afrianty, E. Budianita, and S. K. Gusti, "Implementation of Feature Selection Information Gain in Support Vector Machine Method for Stroke Disease Classification," *Bull. Informatics Data Sci.*, vol. 4, no. 1, pp. 22–33, May 2025, doi: 10.61944/bids.v4i1.116.
[24] A. Syaadah, A. A. Fadhilah, and I. Daulay, "Prediction of Heart Disease with Random Forest through Imputation and Feature Scaling Techniques in Data Preprocessing," *J. Sist. Inf. Computer.*, vol. 1, no. 1, pp. 73–92, 2024.
[25] T. Z. Jasman, M. A. Fadhlullah, A. L. Pratama, and R. Rismayani, "Analysis of Gradient Boosting, Adaboost and Catboost Algorithms in Water Quality Classification," *J. Tek. Inform. and Sist. Inf.*, vol. 8, no. 2, pp. 392–402, 2022, doi: 10.28932/jutisi.v8i2.4906.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 2668
This work is licensed under a Creative Commons Attribution 4.0 International License.