Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 4, October 2025, pp: 964-976 **e-ISSN: 2656-8632**

RESEARCH ARTICLE OPEN ACCESS

Heart Disease Classification Using Random Forest and Fox Algorithm as Hyperparameter Tuning

Afidatul Masbakhah, Umu Sa'adah, and Mohamad Muslikh

Department of Mathematics, Brawijaya University, Malang, East Java, Indonesia

Corresponding author: Umu Sa'adah. (e-mail: u.saadah@ub.ac.id), **Author(s) Email**: Afidatul Masbakhah (e-mail: afidah29600@gmail.com), Mohamad Muslikh (e-mail: mslk@ub.ac.id)

Abstract **Heart disease remains the leading cause of death worldwide, making early and accurate diagnosis crucial for reducing mortality and improving patient outcomes. Traditional diagnostic approaches often suffer from subjectivity, delay, and high costs. Therefore, an effective and automated classification system is necessary to assist medical professionals in making more accurate and timely decisions. This study aims to develop a heart disease classification model using Random Forest, optimized through the FOX algorithm for hyperparameter tuning, to improve predictive performance and reliability. The main contribution of this research lies in the integration of the FOX metaheuristic optimization algorithm with the RF classifier. FOX, inspired by fox hunting behavior, balances exploration and exploitation in searching for the optimal hyperparameters. The proposed RF-FOX model is evaluated on the UCI Heart Disease dataset consisting of 303 instances and 13 features. Several preprocessing steps were conducted, including label encoding, outlier removal, missing value imputation, normalization, and class balancing using SMOTE-NC. FOX was used to optimize six RF hyperparameters across a defined search space. The experimental results demonstrate that the RF-FOX model achieved superior performance compared to standard RF and other hybrid optimization methods. With a training accuracy of 100% and testing accuracy of 97.83%, the model also attained precision (97.83%), recall (97.88%), and F1-score (97.89%). It significantly outperformed RF-GS, RF-RS, RF-PSO, RF-BA, and RF-FA models in all evaluation metrics. In conclusion, the RF-FOX model proves highly effective for heart disease classification, providing enhanced accuracy, reduced misclassification, and clinical applicability. This approach not only optimizes classifier performance but also supports medical decision-making with interpretable and reliable outcomes. Future work may involve validating the model on more diverse datasets to further ensure its generalizability and robustness.**

Keywords **Classification; Random Forest; Hyperparameter Tuning; FOX Algorithm; Heart Disease**.

I. Introduction

Accurate detection and classification of heart disease remains one of the most significant challenges in healthcare systems around the world. Cardiovascular disease (CVC) continues to be the leading cause of death globally, with approximately 17.9 million deaths each year, representing 31% of total deaths worldwide [1], [2]. Early and accurate detection of heart disease is essential for effective treatment and reduction of mortality rates [2]. However, traditional diagnostic methods often rely on a variety of clinical tests and physician expertise, which can lead to delayed diagnosis, increased healthcare costs, and variations in subjective judgment [3]. Therefore, the development of an automated and reliable heart disease classification system is critical to assist healthcare professionals in making timely and accurate diagnostic decisions.

In this study, one of the medical diagnostic efforts, refers to the reliability of machine learning. Machine Learning (ML) techniques have emerged as a powerful tool for medical diagnosis, particularly in the classification of heart diseases, due to their ability to analyze complex medical data and identify patterns that may not be immediately visible to human observers [4], [5]. Among various ML algorithms, ensemble methods such as Random Forest (RF) have shown promising results in medical diagnostics due to their resistance to overfitting and their ability to handle high-dimensional data [6]. RF has been successfully applied to a wide range of medical classification tasks, achieving high accuracy in heart disease prediction [7], [8]. However, the performance of RF is highly

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 4, October 2025, pp: 964-976 e-ISSN: 2656-8632

dependent on its hyperparameters, such as the number of decision trees, the maximum depth, and the minimum number of samples for node separation [9]. Conventional approaches to hyperparameter tuning often involve grid search or random search methods, which consume a lot of computing resources and may not always find the optimal combination of hyperparameters [10].

Several studies have attempted to optimize the hyperparameter RF for the classification of heart disease. Zhang et. al. [11] used Bayesian optimization to tune the RF hyperparameter and reported an accuracy of 88.7%. Similarly, Barry et al. [12] used Particle Swarm Optimization (PSO) for hyperparameter tuning and achieves an accuracy of 92.3%. Recently, Genetic Algorithms (GA) have been used by Rahman et al. [13] to optimize the RF parameters, resulting in an accuracy of 93.5%. Although these metaheuristic approaches have shown improvement compared to manual tuning, they still experience slow convergence rates and the potential to get stuck in local optima when dealing with complex hyperparameter spaces of ensemble classifiers [14], [15]. Therefore, there is a significant research gap in the development of efficient and effective hyperparameter optimization methods that can overcome these limitations and further improve the performance of RF for the classification of heart diseases.

To address this gap, the study proposed the integration of the Fox Optimization Algorithm (FOX), a metaheuristic algorithm inspired by Red Fox Optimization (RFO) and the Fox Hunting Algorithm (FHA) [16] newly developed, to set the RF hyperparameter in the classification of heart diseases. The FOX algorithm, inspired by the fox's hunting behavior and social intelligence, has shown superior performance in solving complex optimization problems compared to traditional optimization methods [17]. The algorithm leverages exploration and exploitation strategies that mimic fox hunting tactics, allowing efficient navigation on hyperparameter space to find optimal or near-optimal solutions [18]. Unlike other metaheuristic algorithms that may get caught up in local optimization, FOX maintains a balance between exploration and exploitation through dynamic adjustment mechanisms, making it particularly suitable for complex hyperparameter optimization tasks in RF [19], [20].

This study aims to develop an efficient heart disease classification system using RF optimized with the FOX algorithm for hyperparameter tuning, thereby improving classification accuracy, reducing computational costs, and improving model interpretability for clinical applications. The main contributions of this study are: 1) the development of a new approach that integrates the FOX algorithm with Random Forest for hyperparameter optimization in the classification of heart diseases, improving accuracy and efficiency compared to state of the art methods; 2) a comprehensive comparative analysis of FOX-optimized RF against other optimization algorithms, including GridSearch (GS), RandomSearch(RS), PSO, Bat Algorithm (BA), and Firefly Algorithm (FA); 3) investigation of the impact of different hyperparameter settings on model performance, providing insight into the relationship between hyperparameters and classification accuracy for heart disease prediction; and 4) the development of interpretable models that identify the most significant features for the classification of heart disease, increasing the reliability and clinical applicability of the system.

The structure of this paper is as follows: Part II describes the datasets used and methodologies, including the RF algorithm and the FOX optimization technique. Part III presents the results of the experiment and comparisons with other methods. Part IV discusses the findings, clinical implications, and limitations of the study. Finally, Part V concludes the paper and suggests directions for future research.

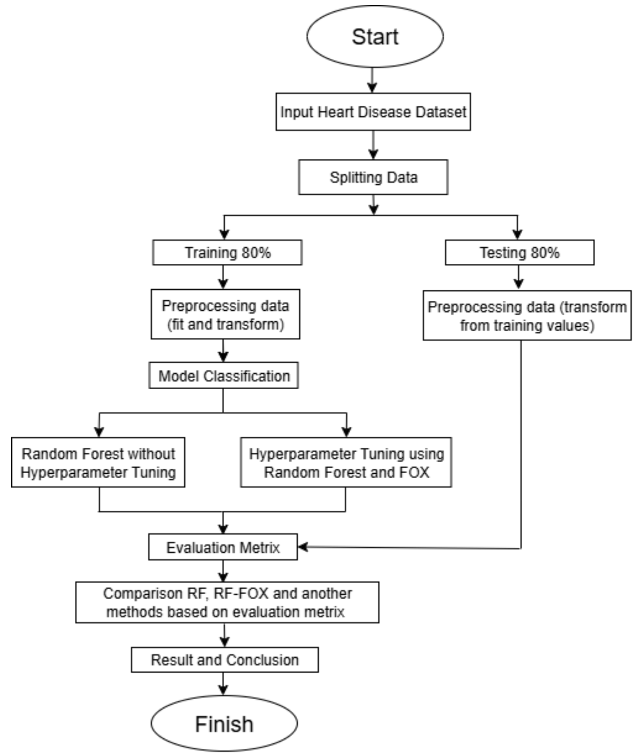

Fig. 1. The research methods workflow of heart disease classification and hyperparameter tuning Random Forest using FOX algorithm.

II. Method

This study examined the classification of heart disease using data obtained from the UCI Machine Learning

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (DOI): https://doi.org/10.35882/jeeemi.v7i4.932

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 4, October 2025, pp: 964-976 e-ISSN: 2656-8632

repository accessible through https://archive.ics.uci.edu/dataset/45/heart+disease. The research process starts from data collection, pre-processing of data that prepares data for training and testing using RF Classifier. This study focuses on hyperparameter optimization RF performed with the FOX algorithm and evaluates the performance of the classification model using specific metrics. The research flow can be seen in Fig. 1. This research was conducted using Python in Google Collaboration Notebook. The dataset analysis was carried out using the Sklearn, Matplotlib, Pandas, and Nuppy heritages. In addition, model validation was carried out using evaluation metrics from the sklearn package, including classification_report.

A. Data Collection

This study used heart disease data consisting of 303 data with 13 features and 1 target. The target class in this classification is a response variable consisting of a negative class (0) and a positive class (1). The dataset was analyzed descriptively and statistically before being classified, to identify the necessary data preprocessing steps. The variables of the research heart disease dataset consist of description and data type as shown in Table 1. Each of these variables contributes to the formation of predictive models that are able to accurately and efficiently identify individuals at high risk of heart disease.

Table 1. Dataset variable of heart disease consist of description and type of data.

Variable	Description	Data Type
Age	Patient age	Integer
Sex	Patient gender	Categorical
CP	Type of chest pain	Categorical
BP	Resting blood pressure	Integer
Chol	Serum cholesterol	Integer
Fbs	Fasting blood sugar	Categorical
Restecg	Rest Electrocardiogram	Categorical
Thalach	Max heart rate achieved	Integer
Exang	Exercise induced angina	Categorical
Oldpeak	ST depression	Integer
Slope	Slope	Categorical
Ca	Number of major vessels	Integer
Thal	Thalassemia	Integer
Target	0 : Negative 1 : Positive	Categorical

The workflow represents a comparative evaluation-not a parallel process. The baseline RF and optimized RF-FOX models were evaluated using the same preprocessed dataset. After separate evaluations, their performances were compared to assess the impact of FOX-based hyperparameter tuning.

B. Data Preprocessing

Data preprocessing in this study began with identifying categorical features on the data and splitting them to be training and testing data. After that, they were converted into integer using label encoding to ensure compatibility with ML algorithm [21]. This approach was preferred over one-hot encoding to reduce dimensionality, considering the dataset size. Next, outlier detection was performed using the interquartile range (IQR) method, which identified abnormal values based on data spread around the median. IQR was chosen for its robustness to non-normal distributions and its ability to handle data variability effectively [22]. Missing values in the “ca” and “thal” features were handled using mean imputation from training dataset for fit, next transform at training and testing data, a simple and widely used method that maintains dataset consistency and is resistant to outlier influence [23]. To standardize feature scales, min-max normalization was applied to numerical variables, scaling them to a [0,1] range. Max and min values using fit from training then were transformed at training and testing, so normalization of the testing data used min-max values from training data. This prevents features with large values from dominating others and improves model training consistency [24],[25],[26]. The normalization formula used is shown in Eq. (1) [27], [28]:

$$x_{inorm} = \frac{x_i - x_{imin}}{x_{imax} - x_{imin}}, \quad (1)$$

with x_{inorm} is the new value of the data sample x, x_{imin} is the smallest value and x_{imax} is the largest value in the feature column.

Normalization is essential before resampling, as class imbalance can bias model learning. In this dataset, a slight imbalance is presented between positive and negative classes. To address this, SMOTE-NC was used, generating synthetic samples for the minority class while handling both nominal and continuous features appropriately [29],[30],[31]. For continuous features, new samples were generated using interpolation between neighbors, as shown in Eq. (2) [32]:

$$x_{m,syn} = x_{m,i} + \lambda(x_{m,i} - x_{m,j}), \quad (2)$$

where $x_{m,i}$ is the value of the m continuous feature of the sample x, $x_{m,i}$, $x_{m,j}$ represent the same features of the sample x, and λ is a number in the range [0,1]. while nominal values were assigned based on the most frequent category among neighbors. Although SMOTE-NC may produce less representative samples if the original data distribution is suboptimal, this can be addressed through data visualization and distribution checks [33], [34]. Overall, these preprocessing steps ensure clean, balanced, and well-scaled input for model training.

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

C. Random Forest

Random Forest (RF) is an ensemble method that combines many decision trees to form a more robust and stable model [35]. Each decision tree was constructed from a random subset of training data, and the final result was obtained by voting a majority of the results of each tree. RF's advantage lies in its ability to handle datasets with complex features, as well as reducing the risk of overfitting that often occurs in single decision tree models [15], [36]. The use of RF learning ensembles, capable of producing models with low variability and higher accuracy [37]. Eq. (3) represent RF models,

$$R(a) = \sum_{k=1}^{N} Y_k \sum_{j \epsilon C_k} \theta(j; P_k) = \sum_{k=1}^{N} \sum_{j \epsilon C_k} Y_k \cdot \theta(j; P_k) \quad (3)$$

with $R(a)$ is the final decision of the classifier, Y_k average response value ke-k, C_k, P_k is a divisive parameter to divide the region of the decision and $\theta(j; P_k)$ is a function of the information limit based on the P_k and j-index. This equation shows the aggregation of the response values in each region k, taking into account the data subses and relevant divisor parameters [38].

D. Hyperparameter Tuning

RF performance can be maximized with precise hyperparameter tuning processes. Hyperparameter tuning is the process of finding the best combination of hyperparameters that can improve model performance [11]. In RF, some hyperparameters that need to be tuned include the number of trees (NEstimators), the maximum tree depth (MaxDepth), the number of features selected in each split (MaxFeatures), the minimum number of samples (MinSamplesSplit), the minimum number of leaves (MinSamplesLeaf) and the criteria (Criterion).

Table 2. Parameter default of Random Forest and range hyperparameter Random Forest.

No	Parameter	Default	Range value
1	NEstimators	100	[1, 500]
2	MaxFeatures	-	[1, 13]
3	MinSamplesSplit	2	[1, 20]
4	MinSamplesLeaf	2	[1, 10]
5	MaxDepth	-	[1, 50]
6	Criterion	0	[0, 1]

The advantage of hyperparameter tuning is its ability to improve model performance by adjusting important parameters to match the characteristics of the dataset being used [39],[40]. Frequently used tuning techniques include GS and RS, but these methods can take significant computational time [41], [42]. Hyperparameter tuning was done within a specific range of values such as in Table 2.

E. FOX Algorithm

The FOX algorithm is a nature inspired optimization method based on fox hunting behavior, which combines exploration and exploitation to locate optimal solutions [16]. During the exploration, it performs random walks guided by simulated ultrasound detection; once prey is detected, the algorithm enters an exploitation phase by estimating the time required to reach the target and executing a calculated jump [17]. FOX requires two main components: an objective function to evaluate fitness and boundary constraints to define the search spacem [18], [43].

FOX implements a static compromise between exploration and exploitation (50% each). In exploration, the algorithm uses random walks to find red fox prey. Meanwhile, in the exploitation phase, the algorithm calculates the distance to the prey, jump height, and new position as in Eq. (4) and Eq. (5) [16].

$$X_{i+1} = X_i(t) + (d_i \cdot DFP_i \cdot Jump_i \cdot c_1), \quad (4)$$

$$X_{i+1} = X_i(t) + (d_i \cdot DFP_i \cdot Jump_i \cdot c_2), \quad (5)$$

with DFP_i is distance fox from prey, d_i is unit direction vector between fox and prey, c_1 and c_2 is a constant that has been set at 0.18 and 0.82, respectively. This constant value comes from the observation of the jumping behavior of the red fox. It is known that the jump of a red fox usually points to the northeast or the opposite direction. FOX explores the surrounding environment using Eq. (6) following to calculate his new position (this is considered exploration) [16], [44].

$$X_{i+1} = BestX \cdot rand(1, dim) \cdot Min(T) \cdot a, \quad (6)$$

with

$$tt = \frac{\sum(Time_{st_{it}})}{dimension},$$
$$Min(T) = Min(tt),$$
$$a = 2\left(it - \frac{1}{(it)}\right).$$

where tt is the average of time calculated from the sum of the time variables divided by the dimension of the problem, it is the current iteration, and is the maximum number of iterations. $Max(it)$ is the calculation of variables and $MinT$ a has an important impact in the search phase to approach the best solution. Using a random function $rand(1, dimension)$, the fox can walk stochastically to explore prey [16], [43], [45]. In the context of RF hyperparameter tuning, the FOX algorithm looks for the optimal combination of hyperparameters by evaluating the performance of the model in each iteration. Compared to GS or RS, the FOX algorithm is able to balance the broad global solution search with the use of the best solutions that have already been found, thereby reducing the risk of being stuck on local solutions and accelerating convergence towards optimal solutions [43].

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

This study applied the FOX algorithm for tuning Random Forest hyperparameters, not a GS or Bayesian Optimization because FOX uses a metaheuristic, nature inspired strategy that balances exploration and exploitation through dynamic movements modeled after fox hunting. This allows FOX to search complex hyperparameter spaces efficiently and avoid local optima. FOX's jump based update and time based prey detection provide more adaptive and flexible tuning than conventional methods, making it highly effective for ensemble models like RF. This theoretical advantage was confirmed by the significant performance improvements shown in the experimental results.

F. Metrics Evaluation

Evaluation of optimized model performance can use several evaluation metrics based on confusion matrix, including the following [46], [47]. Confusion matrix is shown on Table 3. True Positive (TP) means the actual value and the predicted value are both positive, False Positive (FP) means the actual value is negative and the predicted value is positive, False Negative (FN) means the actual value is positive and the predicted value is negative, and True Negative (TN) means the actual value and the predicted value are both negative.

Table 3. Confusion matrix of classification heart disease using Random Forest model.

Actual	Prediction	
	Negative	Positive
Negative	TN	FP
Positive	FN	TP

Some of the metric values used are accuracy, precision, recall and F1 score, with the formula as follows.

a. Accuracy as in Eq. (7) is used to measure the correct prediction percentage of the overall data [48].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}. \quad (7)$$

b. Precision is used to measure the proportion of correct positive predictions to all positive predictions [49], like as Eq. (8).

$$Precision = TP/(TP + FP). \quad (8)$$

c. Recall (Sensitivity) is used to measure the model's ability to detect all true positive instances [47], like as Eq. (9).

$$Recall = TP/(TP + FN). \quad (9)$$

d. F1 Score is a harmonious average between precision and recall, which gives an idea of the balance between the two, especially if the dataset has an unbalanced class [50], The formula is like Eq. (10).

$$F1score = \frac{2(presision \cdot recall)}{presision + recall}. \quad (10)$$

G. Proposed Research Method

The proposed method hybrid RF classification with the FOX optimization algorithm aims to enhance hyperparameter tuning for heart disease diagnosis. The workflow began by inputting the dataset, followed by splitting the data into training and testing subsets with presentation of 80:20. Preprocessing steps were then applied to prepare the data for analysis. The heart disease classification is first conducted using the baseline RF model without any tuning and performance was evaluated. Subsequently, the FOX algorithm was used to search for the best hyperparameter combinations, given its ability to efficiently explore the parameter search space. The optimized RF model was then evaluated and compared with the baseline model using the same metrics. This comparison measures the level of improvement achieved through the hyperparameter optimization process. Result show a significant impact of using the FOX algorithm in improving heart disease classification performance. Furthermore, the depth analysis provides valuable insights into the effectiveness of the proposed approach and serves as foundation for further research in machine learning based medical diagnosis.

III. Result

This section contains the results of the preprocessing of the applied data, the evaluation of the proposed model, and its comparison with other models. In addition, this section also highlights the features that are most influential in classifying heart disease in individuals. The result of the data analysis consists of preprocessing data, hyperparameter RF-FOX and the result, evaluation matrix of RF-FOX performance, and comparison methods.

A. Data Analysis

The dataset contains 303 instances with 13 features and a binary target. Before the preprocessing, the data was split into training data and testing data with a ratio of 80:20 (in percentage) using the function called train_test_split of Sklearn, which was determined after testing with several other ratio ratios such as 90:10, 85:15, 80:20, 75:25, and 70:30 on basic models [51], [52], [53]. Each model, run 25 times to assess stability based on the average and standard deviation ratio 80:20 ratio was chosen because it yields the best performance. Ratio selection based on the base model was used to ensure that the selected comparison is fair and consistent.

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

Categorical variables such as "sex", "cp", "fbs", and others were encoded using label encoding to allow processing by machine learning algorithms. Prior to resampling, the class distribution was slightly imbalanced, with 165 positive and 138 negative samples. Descriptive analysis showed wide value ranges in features like "chol" and "thalach", justifying the use of min-max normalization. Outlier detection using the IQR method revealed no significant outliers. Two features, "ca" and "thal", contained missing values, which were handled using mean imputation. Furthermore, an examination was carried out on the target data that shows that the data class was unbalanced, so it was necessary to carry out data balancing. Before that, feature scaling was carried out by normalizing data using min-max scaling. After normalization, SMOTE-NC was applied to balance the classes. Correlation analysis among numerical features revealed no strong multicollinearity. Additionally, a Random Forest based feature importance analysis indicated that "cp", "thalach", and "oldpeak" were the most influential features, supporting their clinical relevance in heart disease prediction. The results of data balancing with SMOTE-NC are shown by Table 4.

Table 4. The result of resampling class data heart disease using SMOTE-NC.

Class	Amount of data		Persentase(%)	
	Original	Resampling	Original	Resampling
0	138	165	45.54	50
1	165	165	54.46	50

After prepocessing the data, the data can be used in the application of machine learning algorithms to classify heart disease diagnoses and optimize hyperparameters with optimization algorithms to obtain optimal optimization in global searches.

B. Hyperparameter Random Forest using FOX

RF was optimized using FOX algorithm by defining the accuracy as the objective function. The tuning process was run with a maximum of 25 iterations. Several population sizes were tested across 25 replications to determine the most stable and effective configuration. As shown in Table 5, increasing the number of fox agents leads to higher accuracy and lower standard deviation, indicating better convergence. The best balance was achieved with a population size of 40, which provided high accuracy (99.67%) and stable results. However, the results also show that larger populations increase computational time significantly from 180 seconds (10 agents) to 2751 seconds (50 agents). Thus, a population size of 40 was chosen as an optimal trade-off between accuracy and computational efficiency.

Table 5. The value of Random Forest-FOX average accuracy with various fox population.

Fox	Avg. Accuracy(%)	Std. Accuracy	Avg. Times/ s
10	0.9577	0.0061	180 s
20	0.9585	0.0067	425 s
30	0.9653	0.0017	1278 s
40	0.9667	0.0023	1818 s
50	0.9783	0.0006	2751 s

C. The Result of Random Forest-FOX Optimization

Each best fitness score of the FOX algorithm yields a different combination of hyperparameters for RF. The best and optimal hyperparameter RF using FOX algorithm performance was obtained from the hyperparameter configuration shown in the Table 6. From these combinations, Table 7 shows the performance of RF-FOX performance based on several evaluation metrics.

Table 6. The result of hyperparameter tuning Random Forest using FOX algorithm.

Hyperparameter	Value
Nestimators	200
MaxFeatures	8
MinSamplesSplit	2
MinSamplesLeaf	1
MaxDepth	5
Criterion	1

Based on Table 7, The performance results of the RF-FOX model in the training phase resulted in a score of 100%, this shows that the optimization of RF hyperparameters using the FOX algorithm not only improves the prediction accuracy (test phase) but also improves the reliability and consistency of classification. The RF-FOX model also achieves superior performance with a lower risk of overfitting, which means that the difference in training and testing phases is not too significant. Although the RF-FOX model achieved 100% training accuracy, we employed stratified 5-fold cross validation during tuning to prevent overfitting and assess robustness. The small gap between training and testing accuracy suggests minimal overfitting.

Table 7. Random Forest-FOX performance model of classification heart disease based on evaluation metrics.

Split Data	Accuracy	Precision	Recall	F1-Score
Training	1.0000	1.0000	1.0000	1.0000
Testing	0.9783	0.9783	0.9788	0.9780

D. Comparison Methods

Analysis of model performance compared the proposed RF-FOX model and several other superior

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 7, No. 4, October 2025, pp: 964-976 **e-ISSN: 2656-8632**

models that have been performed in previous studies after 25 experiments with the same treatment. This comparison was based on the results of the model's performance according to the calculation of evaluation metrics. The results of the comparison are shown in Table 8.

This study compared standard RF and several other hybrid methods based on the advantages of previous research. Among them are RF-GS, RF-RS, RF-PSO, RF-BA, and RF-FA. These models have been trained on datasets that have undergone the same prepocessing phase and optimized with other swarm intelligence algorithms that are considered superior and optimal. Hyperparameter optimization using PSO, BA, and FA algorithms is used as a benchmark in terms of accuracy and stability. PSO depends on speed updates [54], BA mimics echolocation [55], and FA resilient against local optimum [56]. Meanwhile, FOX has a balanced advantage between exploitation and exploration [57].

Table 8. Comparison RF-FOX with another hybrid methods of heart disease classification.

Methods	Avg. Accuracy	Avg. Precision	Avg. Recall	Avg. F1-Score
RF	0.8753	0.8516	0.8771	0.8777
RF-GS	0.9034	0.8919	0.9112	0.9006
RF-RS	0.9155	0.9167	0.9212	0.9159
RF-PSO	0.9545	0.9551	0.9549	0.9544
RF-BA	0.9310	0.9319	0.9320	0.9318
RF-FA	0.9636	0.9838	0.9839	0.9833
RF-FOX	**0.9783**	**0.9783**	**0.9788**	**0.9780**

Based on Table 8, it can be seen that RF-FOX improves the performance of the standard RF applied with the default parameters on the Table 2. RF-FOX compared to standard RF has achieved an increase of 0.1030 at accuracy values, 0.1267 at precision values, 0.1017 at recalls, and 0.1003 at F1-scores. The proposed RF-FOX model also outperformed other hybrid models in all four evaluation metrics. Higher accuracy indicates that many samples were correctly classified. The increase in precision values proves the model's ability to prioritize true positive classes and reduce false positives. The relatively superior recall value indicates better sensitivity in detecting all samples from each class and minimizing missed cases. Likewise, the high F1-score value on the proposed method, confirms that the model has a balanced performance between precision and recall in all classes. Table 8 overall shows that the RF-FOX model performs better than standard RF and other hybrid methods.

Table 9. validates the observed performance improvements, we conducted statistical analysis based on 25 experimental runs for each model. Table X summarizes the Wilcoxon signed-rank test comparing RF-FOX against other methods. RF-FOX consistently outperformed all compared models with p-values < 0.01, indicating that the improvements are statistically significant. In addition, RF-FOX achieved the highest mean accuracy (0.9989) with the lowest standard deviation (0.0006). A 95% confidence interval for the RF-FOX testing accuracy was [0.9783, 0.9989], suggesting both high precision and robustness. These results confirm that the superior performance of the RF-FOX model is not due to random variation but reflects a meaningful and reproducible improvement.

Table 9. Wilcoxon Signed-Rank test results comparing between RF-FOX and other models.

Methods	W-statistic	Z-value	p-value	significance
RF	0	-4.6152	0.0000	p <0.01
RF-GS	2	-4.3710	0.0001	p <0.01
RF-RS	3	-4.1620	0.0002	p <0.01
RF-PSO	4	-4.0143	0.0003	p <0.01
RF-BA	1	-4.5210	0.0000	p <0.01
RF-FA	3	-4.0821	0.0002	p <0.01

IV. Discussion

The study aims to develop an early diagnosis model for a person's heart condition, with the hope of reducing the high mortality rate caused by heart disease. This model was built using one of the Machine Learning algorithms, Random Forest, with a focus on optimizing the algorithm to improve the accuracy and efficiency of classification. Optimization was carried out in the hyperparameter tuning process using the FOX algorithm. This study used a heart disease dataset, obtained from the UCI repository. This dataset consists of 303 observations, 13 features and one target class. Before the classification process, the dataset was obtained through several steps, including coding labels for categorical features. Data normalization was then applied to maintain its distribution, followed by resampling with SMOTE-NC to handle unbalanced data.

Based on the results presented on Table 8, [58], [59], this study provides a comparison with previous studies with similar datasets using various different models. The comparison is shown in Table 10. Santh et. al. [53] using a hybrid RF method with a standard tuning hyperparameter i.e. GridSearch which shows significantly improved results from standard RF. Valarmathi et. al. [60] proposed standard hyperparameter tuning methods, namely GS, RS and the Tree-Based Pipeline Optimization Tool (TPOT)-genetic programming algorithm. Valarmati's research yields more promising results of 4.52% of the best performance of RF-TPOT compared to research [53]. Moreover, Torthi et al., [61] has also proposed a hybrid RF method with two swarm intelligence algorithms,

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

namely BA and PSO. The RF-BA-PSO method has outperformed other hybrid methods compared by Torthi such as GAPSO-RF and GA-RBF. Another previous research project by Parikh et al. [62] also conducted research on heart disease using several methods, such as Tree-based models like Random Forest, XGBoost, and Decision Trees and various hyper-parameter optimization methods including Grid Search, Random Search, Swarm and Evolutionary Algorithms which excel in accuracy and robustness. However, they are computationally inefficient and less effective in dynamic settings. A novel Hybrid Swarm Evolution Optimization (HySEOpt) was introduced, adjusting mutation rates based on performance curves and utilizes parallel processing for faster optimization achieving 98.01% accuracy. Parikh's research was able to outperform previous studies

As shown in Table 10, the proposed model in this study, RF-FOX achieved the highest accuracy of 99.89%, surpassing previous studies in this dataset. An increase of 1.88%-6.89% indicates that FOX effectively optimizes hyperparameters to improve model generalization and minimize misclassification. Overall, RF-FOX proved to be one of the superior solutions for improving the performance of heart disease classification with the dataset used. From a clinical perspective, this intervention is of great importance because its higher accuracy ensures a more precise classification of liver disease stages, allows timely intervention for severe cases and reduces complications up to death.

Table 10. Some previous studies on hyperparameter tuning Random Forest of heart disease dataset.

Author	Year	Methods	Accuracy
Santh et. al. [53]	2020	RF-GS	93.00%
Valarmathi et. al. [60]	2021	RF-RS RF-TPOT	95.04% 97.52%
Torthi et. al. [61]	2023	RF-BA-PSO	95.57%
Parikh et. al. [62]	2024	RF-HySEOpt	96.01%
Purposed Method	-	RF-FOX	97.83%

In addition to its strong predictive performance, the RF-FOX model provides clinical interpretability through feature importance analysis. The top features chest pain type (cp), maximum heart rate (thalach), and ST depression (oldpeak) are all established indicators of heart disease. Their alignment with clinical knowledge enhances confidence in the model's decisions, making it not only accurate but also medically meaningful for supporting diagnosis and treatment planning. The proposed model shows the potential for more reliable early diagnosis, leading to better and patient-ready treatment planning and outcomes. In a longitudinal context, regular retraining with the latest trends in data can undermine the model's accuracy and adaptability. Monitoring model performance metrics can help optimize retraining for long-term effectiveness.

Despite achieving high performance, the proposed RF-FOX model has limitations that may impact its generalizability in real-world clinical settings. The dataset used in this study, while widely adopted, is relatively small and lacks demographic diversity, which may limit the model's applicability to broader populations. Overfitting is also a concern, especially given the near-perfect performance observed; while cross-validation was applied, further external validation is necessary to confirm the model's robustness. In real-world environments, patient data often contain more noise, imbalance, and missing values than benchmark datasets. Therefore, future work should include testing on larger, multi-center datasets with varied patient profiles. Additionally, implementing model calibration, regularization techniques, and domain adaptation strategies could help improve generalization and reduce the risk of overfitting when deployed in diverse clinical environments.

Algorithm-based models can more quickly analyze larger data sets, reduce subjectivity, and detect complex patterns that humans may miss, thereby improving the efficiency of clinical assessment. However, external validation and clinical testing are essential for widespread adoption. This is necessary because the model's interpretation capabilities must be improved for clinical use. This proposed model can potentially assist healthcare workers in decision-making, improve diagnosis, and support the Clinical Decision Support System (CDS). To enhance clinical relevance, the RF-FOX model should be integrated into existing workflows, such as electronic health records and decision support systems. Deployment may face challenges including clinician trust, data privacy, and model transparency. Ethical concerns such as bias, informed consent, and over-reliance must also be addressed. Future efforts should focus on interpretability, collaborative validation, and ethical safeguards to ensure safe integration into healthcare environments.

V. Conclusion

This study aimed to enhance heart disease classification by integrating the FOX optimization algorithm with the Random Forest (RF) classifier for effective hyperparameter tuning. The proposed RF-FOX model achieved superior performance on the UCI

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

Heart Disease dataset, with a testing accuracy of 97.83, precision of 97.83%, recall of 97.88%, and F1-score of 97.89%. These results were statistically validated using the Wilcoxon signed-rank test ($p < 0.01$), confirming that the improvements were significant and reliable. Beyond its strong predictive performance, the FOX algorithm demonstrated efficient convergence and better generalization compared to other methods such as PSO, BA, FA, GS, and RS. The feature importance analysis identified ‘cp’, ‘thalach’, and ‘oldpeak’ as the most influential variables, reinforcing the model’s clinical relevance. For future work, the model should be validated on larger and more diverse datasets. Enhancements such as interpretability tools, regularization strategies, and integration into clinical decision support systems (CDSS) are also recommended to ensure practical deployment in healthcare settings.

Acknowledgment

Thank you for LPDP Kemenkeu Republik Indonesia for research cost assistance and support in completing this research.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The data analyzed in this study is heart disease data obtained from the UCI repository.

Author Contribution

Afidatul Masbakhah conducted data collection, data analysis, data processing, and interpretation of the research findings, and was responsible for manuscript writing. Umu Sa'adah served as the corresponding author, contributed to data analysis, and provided critical feedback on the manuscript and revisions. Mohamad Muslikh contributed to the writing of the mathematical theory. All authors reviewed and approved the final version of the manuscript and agreed to be responsible for all aspects of the work, ensuring its integrity and accuracy.

Declarations

Ethical Approval

This study utilized secondary data related to the classification of heart disease data. The dataset was obtained from the UCI Machine Learning Repository, which is an open-access and publicly available source. Therefore, this research did not involve direct interaction with human subjects and did not require additional ethical approval.

Consent for Publication Participants.

Consent for publication was given by all participants

Competing Interests

The authors declare no competing interests.

References

[1] S. Emmons-Bell, C. Johnson, and G. Roth, “Prevalence, incidence and survival of heart failure: a systematic review,” 2022, *BMJ Publishing Group*. doi: 10.1136/heartjnl-2021-320131.

[2] G. Savarese, P. M. Becher, L. H. Lund, P. Seferovic, G. M. C. Rosano, and A. J. S. Coats, “Global burden of heart failure: a comprehensive and updated review of epidemiology,” Dec. 01, 2022, *Oxford University Press*. doi: 10.1093/cvr/cvac013.

[3] A. Norhammar *et al.*, “Prevalence, outcomes and costs of a contemporary, multinational population with heart failure,” *Heart*, vol. 109, no. 7, pp. 548–556, Apr. 2023, doi: 10.1136/heartjnl-2022-321702.

[4] V. L. Roger, “Epidemiology of Heart Failure: A Contemporary Perspective,” *Circ Res*, vol. 128, no. 10, pp. 1421–1434, May 2021, doi: 10.1161/CIRCRESAHA.121.318172.

[5] J. G. F. Cleland, “The struggle towards a Universal Definition of Heart Failure—how to proceed?,” *Eur Heart J*, vol. 42, no. 24, pp. 2331–2332, Jun. 2021, doi: 10.1093/eurheartj/ehab082.

[6] A. Ward *et al.*, “Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population,” *NPJ Digit Med*, vol. 3, no. 1, Dec. 2020, doi: 10.1038/s41746-020-00331-1.

[7] R. Nakanishi *et al.*, “Machine Learning Adds to Clinical and CAC Assessments in Predicting 10-Year CHD and CVD Deaths,” *JACC Cardiovasc Imaging*, vol. 14, no. 3, pp. 615–625, Mar. 2021, doi: 10.1016/j.jcmg.2020.08.024.

[8] B. K. Tamarappoo *et al.*, “Machine learning integration of circulating and imaging biomarkers for explainable patient-specific prediction of cardiac events: A prospective study,” *Atherosclerosis*, vol. 318, pp. 76–82, Feb. 2021, doi: 10.1016/j.atherosclerosis.2020.11.008.

[9] X. Liu, J. Lu, H. Dai, D. Zhou, S. Cheng, and J. Wang, “Prevention and Health Promotion MACHINE LEARNING DRIVEN CORONARY HEART DISEASE RISK ASSESSMENT: ANALYSES OF NHANES 1999-2018 DATA.”

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

[10] D. P. Mishra, H. K. Gupta, G. Saajith, and R. Bag, "Optimizing Heart Disease Prediction Model with GridsearchCV for Hyperparameter Tuning," in *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing, IC-CGU 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/IC-CGU58078.2024.10530772.

[11] P. K. P, M. A. B. V, and G. G. Nair, "An efficient classification framework for breast cancer using hyper parameter tuned Random Decision Forest Classifier and Bayesian Optimization," *Biomed Signal Process Control*, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102682.

[12] K. A. Barry, Y. Manzali, R. Flouchi, and M. Elfar, "Heart disease approach using modified random forest and particle swarm optimization," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 14, no. 2, p. 1242, Apr. 2025, doi: 10.11591/ijai.v14.i2.pp1242-1251.

[13] C. M. Rahman and T. A. Rashid, "A new evolutionary algorithm: Learner performance based behavior algorithm," *Egyptian Informatics Journal*, vol. 22, no. 2, pp. 213–223, Jul. 2021, doi: 10.1016/j.eij.2020.08.003.

[14] N. Bacanin, T. Bezdan, E. Tuba, I. Strumberger, and M. Tuba, "Optimizing convolutional neural network hyperparameters by enhanced swarm intelligence metaheuristics," *Algorithms*, vol. 13, no. 3, Mar. 2020, doi: 10.3390/a13030067.

[15] M. Daviran, M. Shamekhi, R. Ghezelbash, and A. Maghsoudi, "Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm," *International Journal of Environmental Science and Technology*, vol. 20, no. 1, pp. 259–276, Jan. 2023, doi: 10.1007/s13762-022-04491-3.

[16] H. Mohammed and T. Rashid, "FOX: a FOX-inspired optimization algorithm," *Applied Intelligence*, vol. 53, pp. 1030–1050, 2023, doi: 10.1007/s10489-022-03533-0/Published.

[17] D. Połap and M. Woźniak, "Red fox optimization algorithm," *Expert Syst Appl*, vol. 166, Mar. 2021, doi: 10.1016/j.eswa.2020.114107.

[18] A. Masbakhah, U. Sa'adah, and M. Muslikh, "Feature Selection Risk Factors Cervical Cancer Using Hybrid Methods Random Forest and FOX-Inspired Optimization Algorithm," *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, vol. 9, no. 2, pp. 352–367, Nov. 2024, doi: 10.18860/ca.v9i2.29582.

[19] Z. Zhang, X. Wang, and L. Cao, "FOX Optimization Algorithm Based on Adaptive Spiral Flight and Multi-Strategy Fusion," *Biomimetics*, vol. 9, no. 9, p. 524, Aug. 2024, doi: 10.3390/biomimetics9090524.

[20] R. Sharma *et al.*, "Comparative performance analysis of binary variants of FOX optimization algorithm with half-quadratic ensemble ranking method for thyroid cancer detection," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-46865-8.

[21] X. Zhu, J. Li, J. Ren, J. Wang, and G. Wang, "Dynamic ensemble learning for multi-label classification," *Inf Sci (N Y)*, vol. 623, pp. 94–111, Apr. 2023, doi: 10.1016/j.ins.2022.12.022.

[22] K. Sumwiza, C. Twizere, G. Rushingabigwi, P. Bakunzibake, and P. Bamurigire, "Enhanced cardiovascular disease prediction model using random forest algorithm," *Inform Med Unlocked*, vol. 41, Jan. 2023, doi: 10.1016/j.imu.2023.101316.

[23] A. Desiani, N. R. Dewi, A. N. Fauza, N. Rachmatullah, M. Arhami, and M. Nawawi, "Handling Missing Data Using Combination of Deletion Technique, Mean, Mode and Artificial Neural Network Imputation for Heart Disease Dataset," 2021. [Online]. Available: https://doi.org/11.26554/sti.2221.6.4.333-312

[24] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Appl Soft Comput*, vol. 97, Dec. 2020, doi: 10.1016/j.asoc.2019.105524.

[25] M. Ahmed Ouameur, M. Caza-Szoka, and D. Massicotte, "Machine learning enabled tools and methods for indoor localization using low power wireless network ☆," *Internet of Things*, vol. 12, p. 0, 2020, doi: 10.1016/j.iot.2020.10.

[26] R. AŞLIYAN, "Examining Variants of Learning Vector Quantizations According to Normalization and Initialization of Vector Positions," *European Journal of Science and Technology*, Dec. 2022, doi: 10.31590/ejosat.1222296.

[27] B. Paul and B. Karn, "Heart disease prediction using scaled conjugate gradient backpropagation of artificial neural network," *Soft comput*, vol. 27, no. 10, pp. 6687–6702, May 2023, doi: 10.1007/s00500-022-07649-w.

[28] H. Benhar, A. Idri, and J. L Fernández-Alemán, "Data preprocessing for heart disease classification: A systematic literature review.," Oct. 01, 2020, *Elsevier Ireland Ltd*. doi: 10.1016/j.cmpb.2020.105635.

[29] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Inf Sci (N Y)*, vol. 513, pp. 429–441, Mar. 2020, doi: 10.1016/j.ins.2019.11.004.

[30] P. Zhang, Y. Jia, and Y. Shang, "Research and application of XGBoost in imbalanced data," *Int J Distrib Sens Netw*, vol. 18, no. 6, Jun. 2022, doi: 10.1177/15501329221106935.

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

[31] C. Vairetti, J. L. Assadi, and S. Maldonado, “Efficient hybrid oversampling and intelligent undersampling for imbalanced big data classification,” *Expert Syst Appl*, vol. 246, Jul. 2024, doi: 10.1016/j.eswa.2024.123149.
[32] E. C. Gök and M. O. Olgun, “SMOTE-NC and gradient boosting imputation based random forest classifier for predicting severity level of covid-19 patients with blood samples,” *Neural Comput Appl*, vol. 33, no. 22, pp. 15693–15707, Nov. 2021, doi: 10.1007/s00521-021-06189-y.
[33] C. H. Bhavani and A. Govardhan, “Cervical cancer prediction using stacked ensemble algorithm with SMOTE and RFERF,” *Mater Today Proc*, vol. 80, pp. 3451–3457, Jan. 2023, doi: 10.1016/j.matpr.2021.07.269.
[34] I. Priyana, N. Alamsyah, Budiman, A. P. Sarifiyono, and E. Rusnendar, “Predictive Boosting for Employee Retention with SMOTE and XGBoost Hyperparameter Tuning,” in *2024 International Conference on Smart Computing, IoT and Machine Learning, SIML 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 92–97. doi: 10.1109/SIML61815.2024.10578116.
[35] L. Breiman, “Random Forests,” 2001.
[36] S. Dhanka and S. Maini, “Random Forest for Heart Disease Detection: A Classification Approach,” in *2021 IEEE 2nd International Conference on Electrical Power and Energy Systems, ICEPES 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICEPES52894.2021.9699506.
[37] M. Pal and S. Parija, “Prediction of Heart Diseases using Random Forest,” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1742-6596/1817/1/012009.
[38] K. M. M. Uddin, A. Al Mamun, A. Chakrabarti, R. Mostafiz, and S. K. Dey, “An ensemble machine learning-based approach to predict cervical cancer using hybrid feature selection,” *Neuroscience Informatics*, vol. 4, no. 3, p. 100169, Sep. 2024, doi: 10.1016/j.neuri.2024.100169.
[39] P. Probst, “Hyperparameters, Tuning and Meta-Learning for Random Forest and Other Machine Learning Algorithms,” 2019.
[40] P. Probst, M. N. Wright, and A. L. Boulesteix, “Hyperparameters and tuning strategies for random forest,” May 01, 2019, *Wiley-Blackwell*. doi: 10.1002/widm.1301.
[41] J. Ivan and S. Y. Prasetyo, “Heart Disease Prediction Using Ensemble Model and Hyperparameter Optimization,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, pp. 290–295, 2023, doi: 10.17762/ijritcc.v11i8s.7208.
[42] A. Baita, I. A. Prasetyo, and N. Cahyono, “HYPERPARAMETER TUNING ON RANDOM FOREST FOR DIAGNOSE COVID-19,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 6, no. 2, Aug. 2023, doi: 10.33387/jiko.v6i2.6389.
[43] M. A. Jumaah, Y. H. Ali, T. A. Rashid, and S. Vimal, “FOXANN: A Method for Boosting Neural Network Performance,” 2024. doi: https://doi.org/10.48550/arXiv.2407.03369.
[44] M. H. Nadimi-Shahraki, H. Zamani, Z. Asghari Varzaneh, and S. Mirjalili, “A Systematic Review of the Whale Optimization Algorithm: Theoretical Foundation, Improvements, and Hybridizations,” *Archives of Computational Methods in Engineering*, vol. 30, no. 7, pp. 4113–4159, Sep. 2023, doi: 10.1007/s11831-023-09928-7.
[45] O. O. Akinola, A. E. Ezugwu, J. O. Agushaka, R. A. Zitar, and L. Abualigah, “Multiclass feature selection with metaheuristic optimization algorithms: a review,” Nov. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00521-022-07705-4.
[46] J. J. Tanimu, M. Hamada, M. Hassan, H. A. Kakudi, and J. O. Abiodun, “A Machine Learning Method for Classification of Cervical Cancer,” *Electronics (Switzerland)*, vol. 11, no. 3, Feb. 2022, doi: 10.3390/electronics11030463.
[47] J. Lu, E. Song, A. Ghoneim, and M. Alrashoud, “Machine learning for assisting cervical cancer diagnosis: An ensemble approach,” *Future Generation Computer Systems*, vol. 106, pp. 199–205, May 2020, doi: 10.1016/j.future.2019.12.033.
[48] Y. Rimal and N. Sharma, “Hyperparameter optimization: a comparative machine learning model analysis for enhanced heart disease prediction accuracy,” *Multimed Tools Appl*, vol. 83, no. 18, pp. 55091–55107, May 2024, doi: 10.1007/s11042-023-17273-x.
[49] Q. H. Doan, S. H. Mai, Q. T. Do, and D. K. Thai, “A cluster-based data splitting method for small sample and class imbalance problems in impact damage classification[Formula presented],” *Appl Soft Comput*, vol. 120, May 2022, doi: 10.1016/j.asoc.2022.108628.
[50] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, no. 1, Jan. 2020, doi: 10.1186/s12864-019-6413-7.
[51] A. P. Kumar, Y. Macha, and A. S. Kumar, “HEART FAILURE DETECTION USING OPTIMIZATION ALGORITHMS,” *J Theor Appl Inf Technol*, vol. 15, no. 7, 2025, [Online]. Available: www.jatit.org

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

[52] N. K. Chauhan and K. Singh, "Performance Assessment of Machine Learning Classifiers Using Selective Feature Approaches for Cervical Cancer Detection," *Wirel Pers Commun*, vol. 124, no. 3, pp. 2335–2366, Jun. 2022, doi: 10.1007/s11277-022-09467-7.

[53] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," *SN Comput Sci*, vol. 1, no. 6, Nov. 2020, doi: 10.1007/s42979-020-00365-y.

[54] M. G. El-Shafiey, A. Hagag, E. S. A. El-Dahshan, and M. A. Ismail, "A hybrid GA and PSO optimized approach for heart-disease prediction based on random forest," *Multimed Tools Appl*, vol. 81, no. 13, pp. 18155–18179, May 2022, doi: 10.1007/s11042-022-12425-x.

[55] S. A. Ardiyansa, N. C. Maharani, S. Anam, and E. Julianto, "OPTIMIZING HEART ATTACK DIAGNOSIS USING RANDOM FOREST WITH BAT ALGORITHM AND GREEDY CROSSOVER TECHNIQUE," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 2, pp. 1053–1066, May 2024, doi: 10.30598/barekengvol18iss2pp1053-1066.

[56] J. Huang, M. M. S. Sabri, D. V. Ulrikh, M. Ahmad, and K. A. M. Alsaffar, "Predicting the Compressive Strength of the Cement-Fly Ash–Slag Ternary Concrete Using the Firefly Algorithm (FA) and Random Forest (RF) Hybrid Machine-Learning Method," *Materials*, vol. 15, no. 12, Jun. 2022, doi: 10.3390/ma15124193.

[57] M. Youssef, M. A. Deif, R. Elgohary, H. Attar, M. Hafez, and S. M. Sharfo, "Fox Optimizer and Logistic Regression for Liver Diseases Classification," in *2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence, EICEEAI 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/EICEEAI60672.2023.10590225.

[58] V. Shalamov, V. Efimova, and A. Filchenkov, "Faster Hyperparameter Optimization via Finding Minimal Regions in Random Forest Regressor," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 378–386. doi: 10.1016/j.procs.2022.11.022.

[59] M. K. Suryadi, R. Herteno, S. W. Saputro, M. R. Faisal, and R. A. Nugroho, "A Comparative Study of Various Hyperparameter Tuning on Random Forest Classification with SMOTE and Feature Selection Using Genetic Algorithm in Software Defect Prediction," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 6, no. 2, pp. 137–147, Apr. 2024, doi: 10.35882/jeeemi.v6i2.375.

[60] R. Valarmathi and T. Sheela, "Heart disease prediction using hyper parameter optimization (HPO) tuning," *Biomed Signal Process Control*, vol. 70, Sep. 2021, doi: 10.1016/j.bspc.2021.103033.

[61] R. Torthi, A. D. K. Marapatla, S. Mande, H. K. V. Gadiraju, and C. Kanumuri, "Heart Disease Prediction Using Random Forest Based Hybrid Optimization Algorithms," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 2, pp. 134–144, 2024, doi: 10.22266/ijies2024.0430.12.

[62] V. Parikh, B. Sharma, A. Byotra, and A. Malhotra, "Optimizing Heart Disease Prediction Using a Hybrid Dynamic Swarm Evolution Approach," *SN Comput Sci*, vol. 5, no. 8, Dec. 2024, doi: 10.1007/s42979-024-03484-y.

Author Biography

Afidatul Masbakhah, S. Mat received her Bachelor's degree (S. Mat) in Mathematics from the Mathematics Study Program at Maulana Malik Ibrahim State Islamic University of Malang, with a focus on Statistics from 2018 to 2022. She has experience in fieldwork, serving in the data collection division and as an assistant financial manager at Pos Indonesia, Blitar at 2020. She conducted research on the implementation of the Bagging CART method for admission data as part of her undergraduate thesis at 2022. Currently, she is a Master's student in Mathematics at the Graduate School of Universitas Brawijaya started from 2023. She has also conducted research on hybrid machine learning methods combined with swarm intelligence for feature selection. Her research interests include intelligent computing, machine learning, and data science. She can be contacted via email at afidah29600@gmail.com.

Dr. Dra. Umu Sa'adah M.Si., the corresponding author, received the bachelor, master, and doctoral degree in Mathematics from Universitas Gadjah Mada, Indonesia in 1993, 2002, and 2015, respectively. She is currently an associate professor at the Department of Mathematics in the Faculty of Mathematics and Natural Sciences, Brawijaya University, Indonesia. She began her research career in 2002, focusing on bootstrap application. Her research interests are in artificial neural networks, bootstrap, data science, data mining, machine learning, statistics, and risk theory. She also has authored several books, including "Kupas Tuntas Algoritma Data Mining dan Implementasinya Menggunakan R" (2021), "Pengantar Algoritma dan

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932

Pemrograman dengan Python" (2023), and "Teori Risiko Aktuaria" (2023). She can be contacted at email: u.saadah@ub.ac.id.

Prof. Drs. Mohamad Muslikh, M.Si. Ph.D. earned his bachelor's degree in Mathematics from Universitas Padjadjaran in 1987, his master's degree in Mathematics from Universitas Gadjah Mada in 1996, and his Ph.D. in Mathematics from Universiti Putra Malaysia in 2019. He is currently a Professor at the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, and has served as the Head of the Research and Community Service Unit at the same faculty since 2020. He began his research career in 1990, focusing on applied and theoretical mathematics. He has supervised 20 undergraduate students, 3 master's students, and 1 doctoral student. He has authored several books, including "Analisis Real" (2012), "Ukuran dan Integral Lebesgue" (2013), "Fungsi Bernilai Himpunan" (2022) and has contributed extensively to journals and conferences in mathematical analysis, fixed-point theory, and set-valued functions. His research interests include real analysis, functional analysis, integral theory, set-valued theory, and fixed-point theory. He can be contacted at email: mslk@ub.ac.id.

Manuscript received 18 May 2025; Revised 10 July 2025; Accepted 30 July 2025; Available online 2 August 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i4.932