Research Article

# Heart Attack Classification with a Machine Learning Approach Based on the Random Forest Algorithm

Suleyman Dal and Necmettin Sezgin

***Abstract*— Heart attack diagnosis delays constitute a critical health problem that increases the risk of mortality. Timely and accurate identification of cardiac events is therefore essential to improve patient outcomes and reduce preventable deaths. This study aims to develop a random forest based classification model using the Heart Disease Classification dataset published on the Kaggle platform to support early diagnosis. This dataset consists of 1319 samples and 8 demographic, clinical and biochemical features for the diagnosis of heart disease. To evaluate the model's reliability and generalizability, a 10-fold cross-validation technique was employed. Through this method, each data instance contributed to both training and testing phases, enabling a more stable and robust performance assessment. This approach also reduced the risk of overfitting and ensured more representative evaluation metrics. The performance of the model was evaluated with ROC curve, training-validation curves, confusion matrix. In the evaluation process, especially in Fold 6, 100% accuracy, precision, recall and F1 score were obtained and it was revealed that the model showed superior performance in the classification task. In addition, as a result of the feature importance analysis, it was determined that troponin, potassium (kcm) and age variables came to the forefront in the decision process. This study aims to fill an important gap in the literature in terms of both strong classification performance and interpretability in the field of machine learning models for heart attack diagnosis.**

*Index Terms*— **Heart Attack Classification, Machine Learning, Random Forest Algorithm, Clinical Decision Support Systems**

## I. INTRODUCTION

THE HEART is a vital organ that systematically pumps blood to maintain the body's life functions. The heart is the most basic component of the cardiovascular system, together with arteries, veins and capillaries, which are involved in the efficient transport of oxygen and nutrients to the tissues [1]. Heart diseases are among the common health problems worldwide with high mortality risk. Among these diseases, heart attacks are responsible for more than 80% of all cardiovascular disease (CVD)-related deaths [1, 2]. Risks that trigger the occurrence of CVD include factors such as high cholesterol and blood pressure, sedentary lifestyle, age, genetic predisposition, obesity, diabetes, stress, excessive alcohol and smoking [3]. Some risk factors can be limited by lifestyle interventions such as smoking cessation, body weight control, regular physical activity and stress management. Diagnostic and imaging techniques such as medical history, physical evaluation, electrocardiography, echocardiography, cardiac magnetic resonance imaging and various blood analyses are widely used in the diagnosis of heart diseases. In the treatment of these diseases, methods such as lifestyle modifications, pharmacological treatment methods, angioplasty, coronary artery bypass surgery and pacemakers are applied by specialist physicians [4, 5].

The risk of death can be significantly prevented by early diagnosis of heart diseases and effective treatment options [6, 7]. In this context, the integration of developing technology into health systems is of vital importance. The use of data analysis-based methods effectively supports the medical decision-making process of specialised physicians in common diseases with high mortality rates such as cardiovascular diseases. In this context, machine learning (ML) methods have been widely embraced by researchers. In recent years, the development of ML methods has become an important auxiliary method by being actively used in the health sector as in almost every field [8, 9]. With the effective analysis of large data sets in the field of health, it can make significant contributions to disease prediction and treatment processes. These methods enable the development of clinical decision support models, especially by performing beyond human intuition. In this context, ML methods strongly support medical professionals with high accuracy in critical processes such as early diagnosis of heart diseases, patient risk classification and treatment response prediction. In this respect, random forest, which is one of the ML algorithms, is an effective method widely preferred in the field of health, especially in disease classification and risk prediction studies. This algorithm, which works on the principle of multiple decision trees, produces successful results in

**Süleyman Dal**, is with the Energy Coordination Office, Rectorate of Batman University, Batman, Türkiye States,(e-mail: suleyman.dal@batman.edu).
https://orcid.org/0000-0002-4564-8076

**Necmettin Sezgin**, is with the Department of Electrical and Electronics Engineering, Batman University, Batman, Türkiye, (e-mail: necmettin.sezgin@batman.edu).
https://orcid.org/ 0000-0002-4893-6014

Manuscript received May 5, 2025; accepted May 15, 2025.
DOI: 10.17694/bajece.1691905

classification problems with medical data thanks to its low error tolerance [10, 11].

In recent years, studies based on ML for heart attack diagnosis have been increasing and the performances of the models developed in this field have been extensively examined in the literature. In their study, Natarajan et al. use Firefly algorithm-assisted feature selection and ensemble learning methods such as Stacking and Voting to identify important attributes related to heart disease and improve prediction accuracy. In the applications on the Z-Alizadeh Sani dataset, the Stacking method performed successfully with an accuracy rate of 86.79% [12]. In another study, Jabbar et al. propose an effective classification method by combining the K-nearest neighbour (KNN) algorithm and genetic algorithm to improve the accuracy of heart disease diagnosis. Experimental results show that the proposed method significantly improves the classification accuracy in heart disease diagnosis [13]. In Enad and Mohammed's study, a comprehensive analysis is performed on the Cleveland dataset using quantum machine learning (QML) methods to support early and accurate diagnosis of heart diseases. In the study, quantum-based approaches (QNN, QSVM, Bagging-QSVM) were compared with traditional classifiers (SVM, ANN) after preprocessing and feature selection; in particular, the Bagging-QSVM model achieved the highest accuracy with 100% success in all key performance measures [2]. In another study, El-Sofany compared ten different machine learning algorithms on feature sets generated by three different feature selection methods (Chi-square, ANOVA and Mutual Information) aiming to accurately predict heart diseases at an early stage. The unbalanced data problem was overcome with the SMOTE method, and the XGBoost algorithm achieved the most successful results with the SF-2 feature set with superior performance values such as 97.57% accuracy, 96.61% sensitivity and 95.00% precision [1].

The main objective of this study is to develop a machine learning model that provides highly accurate results for early prediction of heart attack risk. In this context, using the Heart Disease Classification Dataset published on the Kaggle platform, a clinical decision support mechanism that can classify individuals' susceptibility levels to heart disease has been created with the Random Forest algorithm. The main objective of the model is to provide a reliable prediction system that can contribute to early diagnosis processes in clinical settings by learning meaningful patterns from patient data. In this context, it both increases the speed and accuracy of clinical decision-making processes and strengthens effective intervention opportunities by providing data-based decision support to specialist physicians in the early detection and management of high-risk individuals. The main contributions of the study can be listed as follows.

- A reliable and generalisable machine learning model with high classification performance has been developed for early diagnosis of heart attack risk using Random Forest algorithm.
- The performance of the model was evaluated through a systematic cross-validation process, resulting in a stable structure that can be integrated into clinical decision support systems.
- This study contributes to data-driven clinical interpretation by identifying key biomedical variables that influence classification decisions.

## II. MATERIAL AND METHODS

### A. *Material*

The dataset used in this study is the Heart Disease Classification Dataset, which was created and published on the Kaggle platform for the analysis of heart attack risk factors, one of the most common causes of death worldwide [37]. This dataset includes demographic, clinical and biochemical parameters that may be associated with heart attack. In the dataset recorded from 1319 individuals, a total of eight input parameters (age, gender, heart rate, systolic blood pressure, diastolic blood pressure, blood glucose level, CK-MB isoenzyme and troponin level) contain an output as an indicator of heart attack. Here, the input variables indicate clinical data on the individual's heart health, while the output variables indicate whether the individual has had a heart attack or not in a binary classification (0 = no heart attack, 1 = heart attack). These variables offer applicability for both clinical studies and artificial intelligence-based prediction models, as the effects of factors such as gender, hypertension, hyperglycaemia and cardiac enzyme levels, which reflect the main causative factors in heart diseases, are tested on heart attacks.

### B. *Methods*

#### 1) *Data Preparation and Preprocessing*

The dataset is a comprehensive dataset containing features for predicting the risk of heart attack. The data used in the study was loaded from a file named Heart_Attack.csv and includes demographic information, clinical measurements and biochemical parameters of individuals. In this context, the dataset is structured to analyse and classify the factors that may contribute to the occurrence of heart attack. In this process, the dataset was systmetically organised and missing or inconsistent data were removed.

**Label Encoding:** Label Encoding is a method that enables categorical variables to be represented in numerical format, making it possible to be processed by machine learning algorithms [14, 15]. In this direction, the class column (Positive--1, negative--0), which is used as the target variable within the scope of the classification problem in this study, was converted into numerical values. This process enables the model to make numerical distinction between categorical classes and enables the algorithm to carry out the classification process effectively.

**Data Balancing (SMOTE - Synthetic Minority Over-sampling Technique):** SMOTE is an oversampling technique that eliminates data imbalance between classes. In imbalanced data sets, when one class has fewer samples than the other, machine learning algorithms usually prioritise the class with the highest number of samples, leading to a decrease in the prediction success of the model in the minority class [16].

SMOTE increases the learning capability of the model by eliminating the data imbalance for the minority class through synthetic examples. In this way, the overall accuracy of the model improves by balancing the data distribution between classes [17, 18]. Equations 1 and 2 present the mathematical representation of the synthetic data generated by the SMOTE method [16, 18].

$$d(x_i, x_{neighbor}) = \sqrt{\sum_{k=1}^{\Pi} (x_{ihi} - x_{neighbor,k_i})^2} \quad (1)$$

$$x_{synthetic} = x_i + \lambda \times (x_{neighbor} - x_i) \quad (2)$$

In these equations, for each minority class sample, a new synthetic data point is generated using the distance information determined in the previous step. Here $x_i$ represents an instance of the minority class. $x_{neighbor}$ represents the selected neighbours. The parameter $\lambda$ is a randomly chosen coefficient between 0 and 1. $x_{synthetic}$ represents the new synthetic data sample generated. In this study, the distribution of the number of samples in the dataset before and after the SMOTE application is presented in Table 1. The SMOTE method eliminated the imbalance between classes by increasing the number of samples in the minority class.

TABLE I
NUMBER OF DATA BEFORE AND AFTER CORRECTION OF UNBALANCED CLASS DISTRIBUTION USING SMOTE METHOD

| Operation | Total sample count | Number of class | Class(0) sample count | Class(1) sample count |
|---|---|---|---|---|
| Before SMOTE | 1319 | 2 | 509 | 810 |
| After SMOTE | 1620 | 2 | 810 | 810 |

**Standard Scaler:** This method is a preprocessing and scaling method that equates the data mean to zero and the standard deviation to one. This technique is frequently used in machine learning applications to prevent the model from being stuck on a single feature. In this way, the use of the standard scaler ensures that all attributes are treated with similar weights during model training, which significantly prevents the model from developing bias towards certain features. This scaling method, expressed mathematically in Equation 3, is critical for improving the accuracy of machine learning models, optimising the training process, and reducing the imbalances that can be introduced by large-scale variables [16].

$$X_{scaled} = \frac{X - mean(X)}{std(X)} \quad (3)$$

$X_{scaled}$ represents scaled data with mean 0 and standard deviation 1, where X is the data value, $mean(X)$ is the mean of the dataset and $std(X)$ is the standard deviation of the dataset.

## III. MODELLING AND EVALUATION

### A. *Random Forest Classifier*

Random Forest [10] is a classification algorithm that utilises multiple decision trees in the training phase and stands out with high accuracy rates among supervised learning methods [19]. Although each tree used is effective in the prediction process independently of each other, the final decision is made according to the weighted preference of all trees [20]. In this way, it improves the generalisation performance of the model by reducing the high variance that each individual tree may show. Furthermore, since the feature selection is a random process, the correlation between the trees is minimised and the overlearning of the model is avoided [21]. As shown in Figure 1, the random forest algorithm can minimise errors by providing highly accurate analyses of complex data sets. In addition, it can work in harmony with effective methods to eliminate data imbalances between classes [22]. This method can be effectively applied in many fields such as biomedical diagnostic systems, financial risk analyses, and development of educational systems [16]. The random forest hyper parameters used in this study are presented in Table 2.

TABLE II
RANDOM FOREST CLASSIFIER HYPER PARAMETERS

| Parameters | Value | Description |
|---|---|---|
| n_estimators | 100 | The total number of trees to be created in the model. |
| max_depth | 10 | Determines the maximum depth of each decision tree. |
| max_samples | 0.8 | The proportion of samples to be used for training each tree (with bootstrap). |
| max_features | 'sqrt' | Determines the maximum number of features to be used per tree; square root is taken. |
| class_weight | 'balanced' | Used to automatically balance class imbalance in the dataset. |
| bootstrap | True | Ensures creation of sub-sample datasets using bootstrap sampling. |
| random_state | 42 | Fixes randomness to ensure reproducibility of results. |

*Evaluation Metrics*

**Cross-Validation:** In order to evaluate the overall performance of the model more reliably, a 10-fold stratified cross-validation method was applied on the dataset. In this method, the dataset is divided into 10 equal parts (folds) and each part is used once as test data and the remaining parts are used as training data. Thanks to the stratified fold technique, the distribution between classes in each layer is preserved similar to the original dataset, and the generalisation ability of the model is measured more accurately even in imbalanced data situations. This approach makes the accuracy assessment of the model more fair and stable, especially in areas such as health data where class imbalance is significant [23].

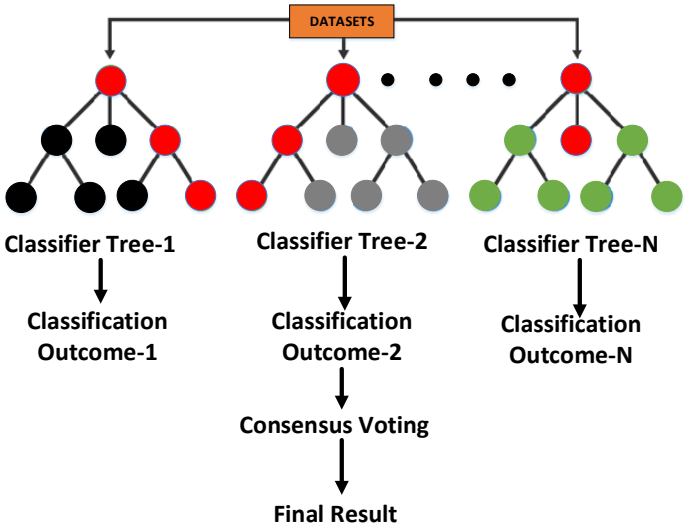


Fig.1. General operation of the random forest classification algorithm

**Performance Metrics Used:** The performance of the model was analysed with the following basic classification metrics:

**Accuracy:** Accuracy is defined as the ratio of the number of samples correctly classified by the model to the total number of samples. This metric is a fundamental measure of the overall success of the model. In this regard, it is calculated as the sum of true positive (TP) and true negative (TN) predictions divided by the sum of all predictions. By considering both positive and negative classes, it reflects the overall recognition ability of the model over all classes. The accuracy metric is expressed mathematically in equation 4 [24, 25].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

**Precision:** Precision refers to the proportion of samples that the model predicts as positive that are actually positive. This metric gains importance when the number of false positives (FP) is high. Precision metric is expressed mathematically in equation 5 [24, 26].

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

**Recall:** Sensitivity indicates how many true positive samples the model can accurately predict. In scenarios where false negatives (FN) need to be minimised, for example in disease detection, it is an important criterion in determining the success of the model. The Recall metric is expressed mathematically in equation 6 [24, 26].

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

**F1 Score:** The F1 score is the harmonic mean of precision and sensitivity and is used to balance both metrics. It is an ideal indicator to evaluate the overall performance of the model, especially in cases where there are imbalances between classes. The F1 Score metric is expressed mathematically in the following equation [16, 27].

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

F1 score is a metric that is critical in evaluating the performance of classification models. This metric, which takes a value between 0 and 1, reflects how a model balances between precision and recall. Especially in cases where false positive and false negative results are important, the F1 score provides a more sensitive indicator of the overall reliability of the model [24].

**Confusion Matrix :** Confusion Matrix analyses the classification success by comparing the model's predicted classes with the actual class labels through four basic components: True Positive (TP), false positive (FP), true negative (TN) and false negative (FN). These components categorically reveal both the model's ability to classify correctly and its tendency to make errors. The Confusion Matrix metric is expressed mathematically in the following equation [28, 29].

$$\begin{aligned} Precision &= TP/(TP + FP) \\ Recall &= TP/(TP + FN) \end{aligned} \quad (8)$$

This 2 × 2 complexity matrix allows to observe not only the overall accuracy rate, but also how successful the model is for each class. In this respect, it is very valuable in evaluating the reliability of the model, especially in data sets with unbalanced class distribution or in applications where positive classes are critical (e.g. in clinical scenarios such as disease detection, readmission prediction). It also forms the basis for various performance metrics such as precision, recall and F1 score. Thus, not only the correct prediction rate of the model, but also the types of errors it makes and the possible effects of these errors in the application context can be analysed more systematically [24, 29].

## IV. FEATURE IMPORTANCE

In machine learning applications, attribute importance ranking is a basic approach that reveals which attributes the model bases its predictions on the target variable [30]. This method both facilitates the understanding of decision-making processes and increases the transparency of the model by providing attributes related to the inner workings of the model [31]. Moreover, knowing the relative importance of attributes helps to eliminate unnecessary variables, especially in high-dimensional data structures, thus reducing the complexity of the model and minimising the risk of overfitting [32]. In this context, the main reasons for using attribute importance analysis are to determine which attributes contribute to the prediction process of the model, to highlight the critical variables required to improve the model performance and to strengthen the interpretability of the model. Especially in classification problems in the field of healthcare, understanding

which clinical indicators the decisions are based on is an important factor that enables expert physicians to use the model more confidently. In this way, artificial intelligence-based models not only make accurate predictions, but also clearly present the rationale for these predictions [33, 34].

In this study, attribute importance rating was performed using a model-based approach. Model-based attribute importance analysis is a method that aims to determine the contribution of each attribute to the decision process based on the internal structure of a trained machine learning model, derived directly from the prediction process [35, 36]. The multiple decision tree structure of the Random Forest algorithm allows the contribution of each attribute to the classification process to be statistically evaluated and a relative importance ranking is obtained according to these contributions. This model-based evaluation allows strong inferences to be made by revealing the attributes that interact with the data more effectively, especially in complex data structures where classical statistical methods are insufficient. This process also increases the accuracy and reliability levels of clinical decision support systems and strengthens both the performance and the acceptability of the model in the eyes of the user.

## V. RESULT AND DISCUSSION

### A. *Result*

In this study, the binary classification problem for heart attack diagnosis is addressed using the random forest algorithm. The dataset consists of 1319 samples and 8 demographic, clinical, and biochemical features, including age, gender, heart rate, systolic and diastolic blood pressure, blood glucose level, CK-MB isoenzyme, and troponin level. To ensure a robust and generalizable evaluation of model performance, a 10-fold cross-validation method was employed, allowing each data instance to contribute equally to model evaluation across different folds.

The experimental analyses show that the model works consistently and with high accuracy rates. The accuracy, precision, recall and F1 score values obtained in each layer are presented in detail in Table 3. The average accuracy value is 98.95%, the average precision is 99.38%, the average recall is 98.52% and the average F1 score is 98.94%. These high success rates indicate that the model has a strong discriminative ability between classes. In this context, especially Fold 6 stands out as the layer that best reflects the performance of the model.

TABLE III
RANDOM FOREST CLASSIFICATION PERFORMANCE (10-FOLD CROSS VALIDATION RESULTS)

| Fold | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| 1 | 0.9877 | 0.9877 | 0.9877 | 0.9877 |
| 2 | 0.9938 | 1.0000 | 0.9877 | 0.9938 |
| 3 | 0.9877 | 1.0000 | 0.9753 | 0.9875 |
| 4 | 0.9815 | 0.9875 | 0.9753 | 0.9814 |
| 5 | 0.9753 | 0.9753 | 0.9753 | 0.9753 |
| 6 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 7 | 0.9938 | 0.9878 | 1.0000 | 0.9939 |
| 8 | 0.9938 | 1.0000 | 0.9877 | 0.9938 |
| 9 | 0.9815 | 1.0000 | 0.9630 | 0.9811 |
| 10 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

The confusion matrix of Fold 6 presented in Figure 2 shows the classification performance of the random forest algorithm on the test data within the scope of cross-validation. In this particular fold, the model correctly classified all 162 samples with 100% accuracy. All true positive (TP = 81) and true negative (TN = 81) samples were predicted without error, and there were no false positive (FP) and false negative (FN) samples. This flawless classification within a single fold underscores the model's robustness under cross-validation settings. These findings suggest that the model may have the potential to maintain consistent performance in similar classification tasks. This result demonstrates that the model has high discriminative power for both classes and exhibits a strong generalisation performance without showing any signs of overfitting.

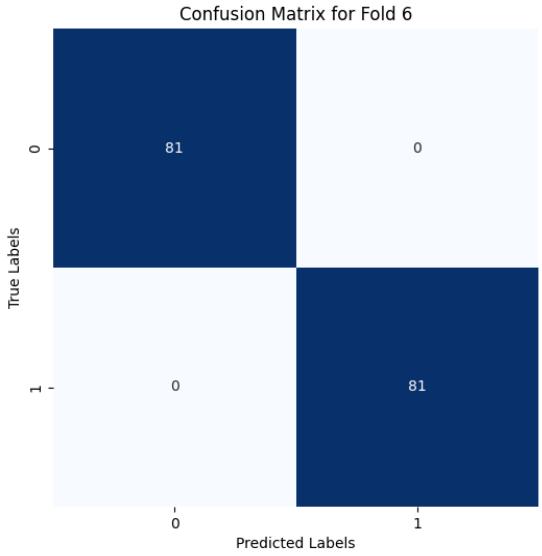


Fig.2. Confusion Matrix for Fold 6

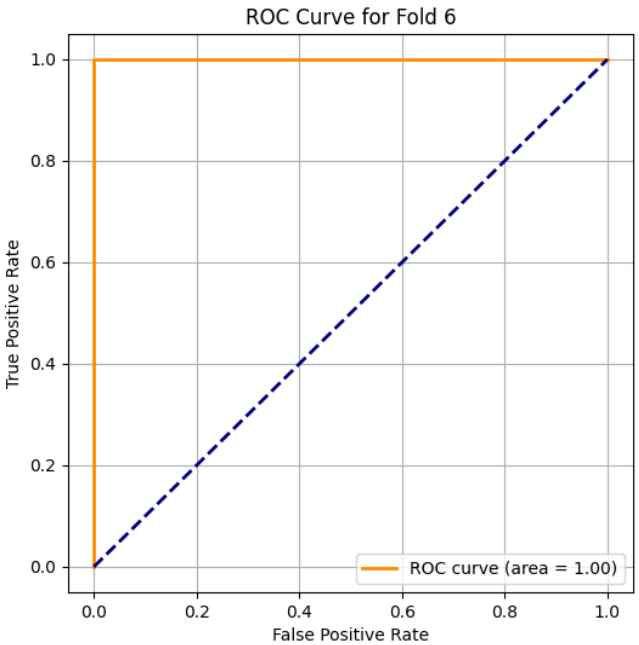


Fig.3. ROC Curve and AUC Value for Fold 6

The classification performance of the model is also supported by the ROC curve. As presented in Figure 3, the ROC curve of Fold 6 shows that the model provides 100% discrimination ability by maintaining both sensitivity and specificity at a high level. The AUC (Area Under Curve) value obtained was 1.00. This shows that the model can discriminate between classes with maximum accuracy.

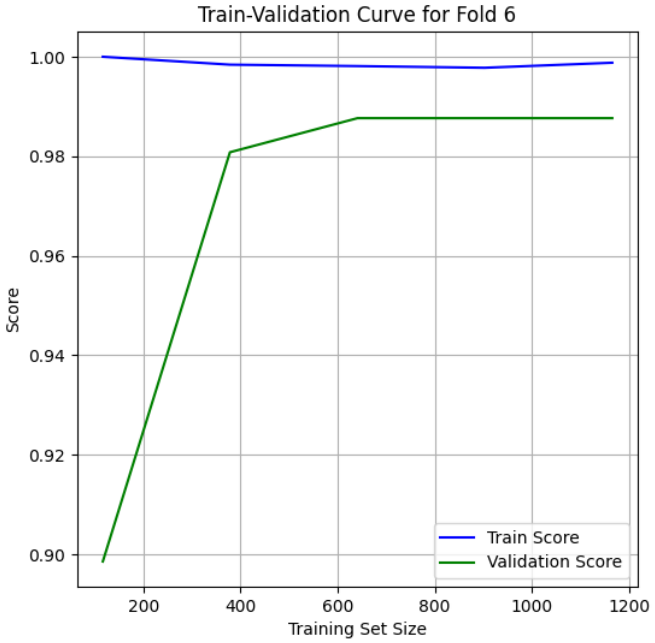


Fig.4. Training-Validation Curve for Fold 6

The training and validation curves generated to obtain insights into the learning process of the model are presented in Figure 4. When the graph is analysed, it is seen that as the training set grows, the validation success remains at a constant and high level (around 98.8%), while the training success is close to 100%. These results show that the model performs consistently throughout the learning process and gains a strong generalisation capability. Therefore, the model achieved high success not only on the training data but also on the validation data, providing a reliable and stable classification performance.

In this study, attribute importance rating is performed with a model-based approach. In model-based attribute importance analysis, the contribution of each attribute to the model is directly calculated by using the internal structure of the classification algorithm trained in the prediction process and the relative effect of the variables is revealed. Thanks to the multiple decision tree structure of the Random Forest algorithm, the contribution of each attribute to the classification process is statistically evaluated and the relative importance ranking is obtained according to these contributions. As presented in Table 4, according to the relative importance values calculated by the random forest algorithm, the troponin variable is the most dominant decision maker of the model with 58.13%. This variable is followed by kcm and age with 25.24% and 5.97%, respectively. The contributions of other attributes seem to be limited, indicating that certain variables play a dominant role in the decision process of the model. The attribute importance distribution graph presented in Figure 5 visually supports this situation and clearly demonstrates the determinant effect of the troponin variable in the classification. These findings are consistent with the literature stating that troponin levels are one of the main biomarkers in determining the risk of heart attack and prove that the model has both an explainable and reliable structure.

TABLE IV
RELATIVE IMPORTANCE VALUES OF FEATURES ACCORDING TO RANDOM FOREST ALGORITHM

| Feature | Importance |
|---|---|
| troponin | 0.5813 |
| kcm | 0.2524 |
| age | 0.0597 |
| pressurehight | 0.0255 |
| glucose | 0.0238 |
| pressurelow | 0.0213 |
| impluse | 0.0199 |
| gender | 0.0160 |

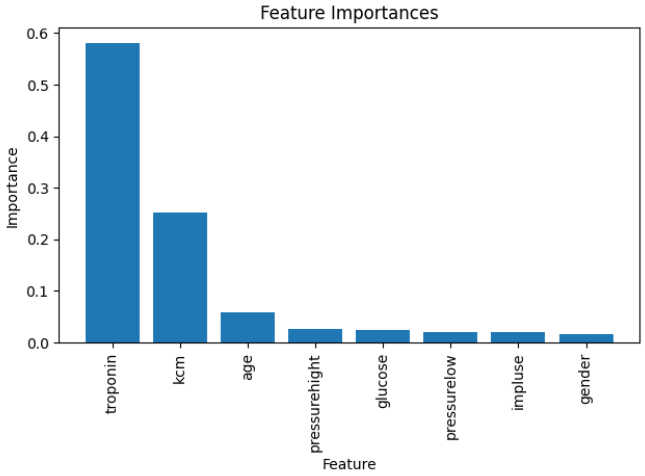


Fig.5. Importance distribution showing effects of attributes on the model

## B. *Discussion*

In this study, a classification model using the random forest method developed for the diagnosis of heart attack has attracted attention with its high accuracy and stable performance results. The dataset consisting of 1319 individuals was evaluated using 10-fold cross-validation to assess the model's performance in a systematic and reliable manner. The metrics obtained show that the classification success is strong; especially the correct classification of all examples in the test data in Fold 6 clearly reveals the discriminative power of the model. These findings suggest that machine learning approaches can be an effective support tool for early diagnosis of cardiovascular diseases.

However, the study has some limitations. The dataset used consists of a relatively limited number of individuals and reflects the characteristics of a specific group, which may limit the generalisability of the model to different populations. Furthermore, the analysis is based on only eight basic clinical parameters. Therefore, it is important to re-evaluate the developed model with more diverse and comprehensive data sets in order to increase its generalisation capacity.

## VI. CONLUSION

This study aims to develop a classification model using the random forest algorithm in order to diagnose the risk of heart attack at an early stage. In order to evaluate the performance of the model, a dataset consisting of a total of 1319 individuals was used, and performance measures such as accuracy, precision, recall and F1 score were analysed through cross-validation-based evaluation. In the evaluation process, 10-fold cross-validation was applied and 100% classification success was achieved in Fold 6, demonstrating the strong discrimination capacity of the model. In addition, troponin and kcm parameters stood out as the most effective variables in the model-based feature importance analysis; these findings are consistent with clinical evaluations in the literature. Furthermore, a significant disadvantage is that the dataset contains only eight clinical parameters and represents a specific population. This may limit the generalisability of the model to different demographic groups. In future studies, retraining and evaluating the model with more diverse and comprehensive datasets containing more samples and attributes is considered. Such approaches can be extended not only to heart diseases, but also to early diagnosis of other diseases such as diabetes and cancer, and can contribute to decision support systems in the field of health.

## REFERENCES


[1] H. F. El-Sofany, "Predicting heart diseases using machine learning and different data classification techniques," *IEEE Access,* 2024.

[2] H. G. Enad and M. A. Mohammed, "Cloud computing-based framework for heart disease classification using quantum machine learning approach," *Journal of Intelligent Systems,* vol. 33, no. 1, p. 20230261, 2024.

[3] T. A. Gaziano, A. Bitton, S. Anand, S. Abrahams-Gessel, and A. Murphy, "Growing epidemic of coronary heart disease in low-and middle-income countries," *Current problems in cardiology,* vol. 35, no. 2, pp. 72-115, 2010.

[4] C. Gupta, A. Saha, N. S. Reddy, and U. D. Acharya, "Cardiac Disease Prediction using Supervised Machine Learning Techniques," in *Journal of physics: conference series*, 2022, vol. 2161, no. 1: IOP Publishing, p. 012013.

[5] A. K. Dubey, A. K. Sinhal, and R. Sharma, "Heart disease classification through crow intelligence optimization-based deep learning approach," *International Journal of Information Technology,* vol. 16, no. 3, pp. 1815-1830, 2024.

[6] R. Rajkumar, K. Anandakumar, and A. Bharathi, "Coronary artery disease (CAD) prediction and classification-a survey," *Breast Cancer,* vol. 90, p. 94.35, 2006.

[7] P. Rani *et al.*, "An extensive review of machine learning and deep learning techniques on heart disease classification and prediction," *Archives of Computational Methods in Engineering,* vol. 31, no. 6, pp. 3331-3349, 2024.

[8] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN computer science,* vol. 2, no. 3, p. 160, 2021.

[9] Ö. F. Ertuğrul, S. Dal, Y. Hazar, and E. Aldemir, "Determining relevant features in activity recognition via wearable sensors on the MYO Armband," *Arabian Journal for Science and Engineering,* vol. 45, pp. 10097-10113, 2020.

[10] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," *Computer methods and programs in biomedicine,* vol. 108, no. 1, pp. 10-19, 2012.

[11] D. R. Edla, K. Mangalorekar, G. Dhavalikar, and S. Dodia, "Classification of EEG data for human mental state analysis using Random Forest Classifier," *Procedia computer science,* vol. 132, pp. 1523-1532, 2018.

[12] K. Natarajan *et al.*, "Efficient heart disease classification through stacked ensemble with optimized firefly feature selection," *International Journal of Computational Intelligence Systems,* vol. 17, no. 1, p. 174, 2024.

[13] B. Deekshatulu and P. Chandra, "Classification of heart disease using k-nearest neighbor and genetic algorithm," *Procedia technology,* vol. 10, pp. 85-94, 2013.

[14] N. Kosaraju, S. R. Sankepally, and K. Mallikharjuna Rao, "Categorical data: Need, encoding, selection of encoding method and its emergence in machine learning models—a practical review study on heart disease prediction dataset using pearson correlation," in *Proceedings of International Conference on Data Science and Applications: ICDSA 2022, Volume 1*, 2023: Springer, pp. 369-382.

[15] T. Amarbayasgalan, V.-H. Pham, N. Theera-Umpon, Y. Piao, and K. H. Ryu, "An efficient prediction method for coronary heart disease risk based on two deep neural networks trained on well-ordered training datasets," *IEEE Access,* vol. 9, pp. 135210-135223, 2021.

[16] Y. Hazar and Ö. F. Ertuğrul, "Process management in diabetes treatment by blending technique," *Computers in Biology and Medicine,* vol. 190, p. 110034, 2025.

[17] P. Soltanzadeh and M. Hashemzadeh, "RCSMOTE: Range-Controlled synthetic minority over-sampling technique for handling the class imbalance problem," *Information Sciences,* vol. 542, pp. 92-111, 2021.

[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research,* vol. 16, pp. 321-357, 2002.

[19] S. Hegelich, "Decision trees and random forests: Machine learning techniques to classify rare events," *European policy analysis,* vol. 2, no. 1, pp. 98-120, 2016.

[20] G. A. B. Suryanegara and M. D. Purbolaksono, "Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi," *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi),* vol. 5, no. 1, pp. 114-122, 2021.

[21] S. Suparyati, E. Utami, and A. H. Muhammad, "Applying different resampling strategies in random forest algorithm to predict lumpy skin disease," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi),* vol. 6, no. 4, pp. 555-562, 2022.

[22] R. Oktafiani, A. Hermawan, and D. Avianto, "Max Depth Impact on Heart Disease Classification: Decision Tree and Random Forest," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi),* vol. 8, no. 1, pp. 160-168, 2024.

[23] I. Tougui, A. Jilbab, and J. El Mhamdi, "Impact of the choice of cross-validation techniques on the results of machine learning-based diagnostic applications," *Healthcare informatics research,* vol. 27, no. 3, pp. 189-199, 2021.

[24] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.", 2022.

[25] M. Yin, J. Wortman Vaughan, and H. Wallach, "Understanding the effect of accuracy on trust in machine learning models," in *Proceedings of the 2019 chi conference on human factors in computing systems*, 2019, pp. 1-12.

[26] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 233-240.

[27] A. Humphrey *et al.*, "Machine-learning classification of astronomical sources: estimating F1-score in the absence of ground truth," *Monthly Notices of the Royal Astronomical Society: Letters,* vol. 517, no. 1, pp. L116-L120, 2022.

[28] J. Liang, "Confusion matrix: Machine learning," *POGIL Activity Clearinghouse,* vol. 3, no. 4, 2022.

[29] I. Markoulidakis and G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies,* vol. 12, no. 7, p. 113, 2024.

[30] V. A. Huynh-Thu, Y. Saeys, L. Wehenkel, and P. Geurts, "Statistical interpretation of machine learning-based feature importance scores for biomarker discovery," *Bioinformatics,* vol. 28, no. 13, pp. 1766-1774, 2012.

[31] F. Pan, T. Converse, D. Ahn, F. Salvetti, and G. Donato, "Feature selection for ranking using boosted trees," in *Proceedings of the 18th ACM conference on Information and knowledge management*, 2009, pp. 2025-2028.

[32] A. A. Megantara and T. Ahmad, "Feature importance ranking for increasing performance of intrusion detection system," in *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)*, 2020: IEEE, pp. 37-42.

[33] M. A. Jamil and S. Khanam, "Influence of one-way ANOVA and Kruskal–Wallis based feature ranking on the performance of ML classifiers for bearing fault diagnosis," *Journal of Vibration Engineering & Technologies,* vol. 12, no. 3, pp. 3101-3132, 2024.

[34] N. Silpa, V. M. Rao, M. V. Subbarao, R. R. Kurada, S. S. Reddy, and P. J. Uppalapati, "An enriched employee retention analysis system with a combination strategy of feature selection and machine learning techniques," in *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2023: IEEE, pp. 142-149.

[35] C. Strobl, A.-L. Boulesteix, A. Zeileis, and T. Hothorn, "Bias in random forest variable importance measures: Illustrations, sources and a solution," *BMC bioinformatics,* vol. 8, pp. 1-21, 2007.

[36] B. M. Greenwell, B. C. Boehmke, and A. J. McCarthy, "A simple and effective model-based variable importance measure," *arXiv preprint arXiv:1805.04755,* 2018.

[37] Bharath011, Heart Disease Classification Dataset, Kaggle, 2022. [Çevrimiçi]. Erişim adresi:https://www.kaggle.com/datasets/bharath011/heart-disease-classification dataset

## BIOGRAPHIES

**Süleyman Dal** Dr. Lecturer. Assist. Süleyman Dal is an academic specialised in the field of electrical and electronics engineering. He completed his undergraduate education at Çukurova University, Department of Electrical and Electronics Engineering in 2018, completed his master's degree at Batman University, Institute of Science and Technology in 2020, and completed his doctorate education at the Institute of Graduate Studies of the same university in 2021.

He is currently working as a Dr. Lecturer in the Energy Coordination Department within the Batman University Rectorate. His research interests include machine learning, signal processing and optimisation algorithms.

**Necmettin Sezgin** Prof. Dr. Necmettin Sezgin is an academic specialised in electronics. He received his bachelor's degree from Hacettepe University, his master's degree from Dicle University and his doctorate from İnönü University.

Since 2011, Prof. Sezgin has been working at Batman University, where he is currently a faculty member of the Department of Electrical and Electronics Engineering and Vice Rector of Batman University. His research interests include signal processing, biomedical signal analysis and electronic systems.