Research Article

An Efficient Automatic Detection of Cardiovascular Disease Based on Machine Learning

Mohammad Karimi Moridani[iD]

School of Biotechnology and Biomolecular Sciences, Faculty of Science, University of New South Wales, Sydney, NSW, Australia
E-mail: m.karimim@unsw.edu.au

Abstract: Cardiovascular diseases have become one of the most common threats to human health worldwide. As a non-invasive diagnostic tool, heart sound detection techniques play an important role in predicting cardiovascular diseases. Although the Electrocardiogram (ECG) signal is generally used to diagnose heart disease, due to the low spatial resolution of this signal, the Phonocardiogram (PCG) signal and methods based on sound processing can be used. In this paper, after extracting different features from PCG, patients were classified with the help of algorithms based on artificial intelligence. The simulation results showed that using the eXtreme Gradient Boosting (XGBoost) algorithm has a better performance in detecting cardiovascular patients than other methods. The values of specificity, sensitivity, and accuracy were obtained as $99 \pm 1.93\%$, $98 \pm 2.76\%$, and $99 \pm 1.78\%$, respectively. Using the method proposed in this paper can greatly help doctors make accurate and quick diagnoses of cardiovascular patients and be effective in screening patients. In the future, this method can be developed to diagnose heart valve diseases.

Keywords: PCG, cardiovascular patients, detection, feature extraction, XGBoost

Abbreviations

AdaBoost: Adaptive Boosting
APQ: Amplitude Perturbation Quotient
BP: Back Propagation
CWT: Continuous Wavelet Transform
DT: Decision Tree
ECG: Electrocardiogram
kNN: k-Nearest Neighbor
MI: Mutual Information
MLP: Multi-Layer Pereptron
PCG: Phonocardiogram
PPQ: Pitch Perturbation Quotient
RF: Random Forest
STFT: Short-Time Fourier Transform
SVM: Support Vector Machines

1. Introduction

Recently, the methods based on the identification of cardiovascular patients have had many variations. Some researchers use cardiac signals and extract linear and non-linear features from these signals to diagnose various heart diseases [1-6]. Several articles have also focused on methods based on the visualization of the ECG signal [7-8], and some studies are based on PCG. In the rest of this section, we will get to know some new research that has used PCG to diagnose cardiovascular diseases.

Based on the logistic function, Kamson et al. presented a novel method for determining the main heart sound envelopes (S1 and S2). The best average F1 score was 97.73% [9]. An unsupervised approach was proposed by Sangita Das et al. to detect S1 and S2 heart sound events in PCGs. A maximum F1 score of 98% is offered for normal PCG data, while a maximum of 92.5% is offered for abnormal PCG data [10]. Nath et al. have proposed that major heart sounds, namely S1 and S2, be detected and localized. As a result of this study, the pulmonic position of the heart is the most appropriate auscultation area for acquiring PCG signal to detect and localize S1 and S2 much more accurately [11]. An algorithm for locating and classifying heart sounds into S1 and S2 was proposed by Qurat-Ul-Ain Mubarak et al. Before localization, feature extraction, and classification of heart sounds, the proposed system introduces the concept of quality assessment. Based on the results of this challenge, it was found that the proposed Localization algorithm achieved an accuracy of up to 97% and generated the lowest total average error among the top three challenge participants. The classification algorithm achieves an accuracy of up to 91% [12].

Recent advancements in machine learning and signal processing have further expanded the potential of PCG-based cardiovascular disease detection. Aparana et al. developed a deep learning approach using convolutional neural networks, achieving an accuracy of 96.5% [13]. Tang et al. explored time-frequency representations with support vector machines, reaching 88.0% sensitivity [14]. Lee et al. combined wavelet transform with an Ensemble of Deep Learning Models for improved heart sound segmentation and classification [15]. Li et al. addressed the challenge of limited labeled data using transfer learning techniques [16], while Li et al. demonstrated the benefits of multi-modal approaches by combining PCG and ECG signals [17]. These studies highlight the ongoing evolution of PCG analysis techniques and the potential for further improvements in cardiovascular disease detection.

Recent advancements in machine learning have expanded beyond cardiovascular disease detection, demonstrating potential applications in various engineering and physical sciences fields. For instance, in chemical engineering, machine learning models have been applied to predict drop coalescence in microfluidic devices, identifying critical features through SHapley Additive exPlanations (SHAP) values and Local Interpretable Model-agnostic Explanations (LIME) [18]. Similarly, the synthesis of silver nanoparticles has been optimized using a combination of microfluidic systems and machine-learning approaches, including decision trees, random forests, and XGBoost [19]. These studies highlight the versatility of machine learning algorithms, particularly XGBoost, which we also employ in our cardiovascular disease detection model. Furthermore, to address the challenge of imbalanced datasets, which is common in medical diagnostics, novel approaches such as the Double Space Conditional Variational Autoencoder (DSCVAE) have been developed to generate synthetic data for training predictive models [20]. While these studies focus on different applications, they underscore the broad applicability of machine learning techniques and the importance of model interpretability and data balance. These are also crucial considerations in our work on cardiovascular disease detection using PCG signals.

In this manuscript, we present the development of an efficient automated system for cardiovascular disease detection using PCG signals. Our primary objective is to implement and compare multiple AI algorithms for heart sound classification, including Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), k-Nearest Neighbors (kNN), AdaBoost, eXtreme Gradient Boosting (XGBoost), Decision Tree, and Random Forest. We utilize a comprehensive set of features extracted from PCG signals, incorporating classical acoustic features, phase space reconstruction, and wavelet transform. Our findings demonstrate the superior performance of the XGBoost algorithm in achieving accurate and reliable classification of heart sounds, highlighting its potential for enhancing cardiovascular disease detection.

The rest of this paper is organized as follows. Section 2 presents the materials and methods, including a detailed

description of the dataset, feature extraction techniques, and classification algorithms employed in our study. Section 3 introduces the simulation results, reports our findings based on different approaches, and compares the performance of various AI algorithms. Section 4 provides a comprehensive discussion of the results, their implications, and the potential impact of our work on cardiovascular disease detection. Finally, Section 5 concludes the paper, summarizing our key contributions and suggesting directions for future research in this field.

2. Material and methods

2.1 *Dataset*

The PhysioNet/Computing in Cardiology (CinC) Challenge addresses this issue by compiling the largest public heart sound database, sourced from eight distinct repositories managed by seven independent research groups worldwide. The database encompasses 4,430 recordings from 1,072 subjects, amounting to 233,512 heart sounds collected from healthy individuals and patients with diverse conditions such as heart valve disease and coronary artery disease. These recordings were captured using varied equipment in clinical and nonclinical settings, including in-home visits. Recording durations ranged from several seconds to several minutes. Additional dataset components include subject demographics (age and gender), recording specifics (number per patient, body location, and duration), synchronously recorded signals (such as ECG), sampling frequency, and sensor type used.

Several contributors worldwide contributed heart sound recordings, taken in either a clinical or nonclinical environment, from healthy subjects and pathological patients. A total of 3,126 heart sound recordings ranging from 5 seconds to over 120 seconds were contained in the data set's five databases (A through E). Various locations on the body were used to collect heart sound recordings. Aortic, pulmonic, tricuspid, and mitral areas are the most common locations, but the locations can vary. Recordings of heart sounds were divided into two types in training and test sets: normal and abnormal. In the normal recordings, healthy subjects were recorded, while patients with confirmed cardiac diagnoses were recorded in the abnormal recordings. All recordings have been standardized to a sampling rate of 2,000 Hz and are stored in .wav format [21]. Each recording contains only one PCG lead. To succeed in better data processing, all data were de-noised by a Kalman filter [22].

2.2 *Feature extraction methods*

Sound analysis is a non-invasive method for diagnosing cardiovascular diseases. In contrast, aggressive methods require complex equipment and much time to record data. In the image recording method, an image of the oscillation process of the vocal cords in the larynx is prepared using the video stereoscope device. In the acoustic method, the patient's heart sound can be recorded many times quickly by phonocardiogram. Of course, invasive methods provide more complete information that effectively determines how to treat diseases [23]. Therefore, non-invasive methods can be used in early and preventive diagnoses. After the initial diagnosis, if necessary, you can refer to specialized centers. In recent years, studies have shown that the human voice has non-linear dynamics due to the non-linearity of components related to sound production such as the non-linear relationship of pressure and airflow, the non-linear activity of sound-producing organs, etc. Many sources state that the non-linear quantities obtained from healthy and patient sounds differ [24].

In this paper, the aim is to calculate the classical acoustic features such as the Pitch Perturbation Quotient (PPQ) and Amplitude Perturbation Quotient (APQ) coefficient [25], extracting linear and non-linear features from PCG signal based on wavelet transform and presenting the returned mapping to appear hidden information from PCG signal. Different classifiers were also used to classify normal and abnormal patients. The value of the accuracy, sensitivity, and specificity of the proposed method for detecting the patients have been reported to evaluate the classifier performance.

2.2.1 *Classical scoustic features*

Classical features based on the measurement of the PPQ and APQ of sound signals in which laryngeal oscillations are involved are extracted. Such parameters estimate the degree of amplitude instability and the fundamental frequency of the fluctuations of sound signals during successive oscillation cycles. Such perturbations can be attributed to changes

in the biomechanical parameters controlling the vocal cords due to their snoring. Generally, the perturbation quotient in a certain number of periodicities of sound signals can be extracted from Equation (1) [26].

$$PQ = \frac{100\%}{N-K} \sum_{i=\frac{K-1}{2}}^{N-\frac{K-1}{2}} \left(U_i - \frac{1}{K} \sum_{k=-\frac{K-1}{2}}^{\frac{K-1}{2}} U_{i+k} \Big/ \frac{1}{K} \sum_{k=-\frac{K-1}{2}}^{\frac{K-1}{2}} U_{i+k} \right) \tag{1}$$

Where U is the amplitude or frequency of the oscillations, N is the total number of periodicities under consideration, and K refers to the length of the windowing signal.

2.2.2 Embedding space

Reconstructing phase space is most easily accomplished by using the time delay method. This method creates vectors in a time-delayed space using delayed measurements called an embedding space. Equation (2) forms the time delay vectors for the space phase given the time series $S(i)$.

$$S(i) = [s(i+\tau), s(i+2\tau), \ldots, s(i+(m+1)\tau], \ \ i = 1, 2, \ldots, N - (m-1)\tau \tag{2}$$

Where τ represents the delay time, and m represents the embedding dimension. According to Fraser and Swinney, the Mutual Information (MI) method can be used to determine a time delay. According to Kennel, a false nearest neighbor's method can be used to calculate the minimum adequate embedding dimension of m [27].

2.2.3 Correlation dimension

Using the correlation dimension, which measures system complexity, we can determine how many independent variables we need to describe how systems behave. Counting the boxes or calculating the Kolmogorov capacity are the easiest ways to measure a set's dimensions. When heterogeneity or correlation are considered, accurate measurements can be made. Based on the slope of the scaling region of the graph $\log C(r)$ to $\log(r)$, the correlation dimension is calculated. Using Equation (3), we can calculate the correlation dimension [27].

$$D = \lim_{r \to \infty} \frac{\log(C(r))}{\log\left(\frac{1}{r}\right)} \tag{3}$$

$C(r)$ is the number of cells (squares for embedded sets in two dimensions, cubes for embedded sets in three dimensions) comprising one of the parts of the set.

2.2.4 Returned mapping

A typically returned mapping includes a Poincaré plot based on the theory of non-linear dynamics. It is a geometric representation of a visual and quantitative time series in a Cartesian graph. Moreover, the analytical Poincaré plot is a non-linear graphical method of plotting a time signal $x(i)$ in a two-dimensional plane in which after a lag delay (τ) is plotted according to itself ($x(i + \tau)$, $x(i)$). Figure 1 shows a Poincaré plot representing the phase space reconstruction based on [28-30].

2.2.5 Wavelet transform (WT)

This method provides simultaneous time and frequency information about a signal in conjunction with a continuous wavelet transform. The time period 'wavelet' was first stated in 1909 in a thesis by Alfred Haar. The Continuous Wavelet Transform (CWT) is a substitute for the Short-Time Fourier Transform (STFT), which uses a variable-sized window location. Because the wavelet might also be dilated or compressed, one-of-a-kind elements of the sign are extracted.

While a slender wavelet extracts excessive frequency components, a stretched wavelet alters the signal's decreased frequency aspects. The CWT is computed with the aid of correlating the signal S(t) with households of time-frequency atoms H(t) [31-33].
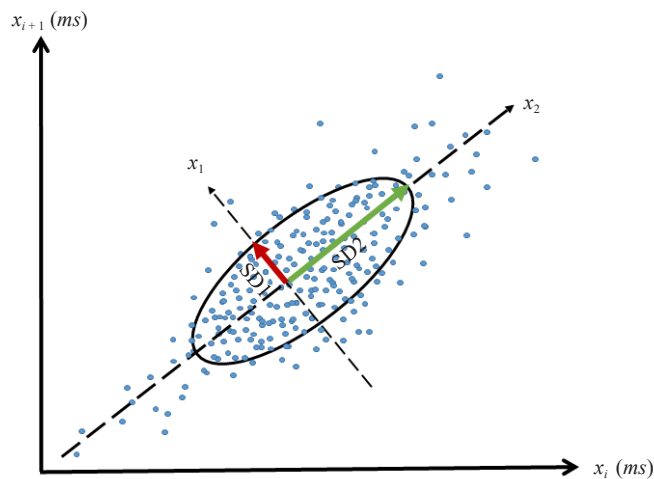

Figure 1. A typically returned mapping [30]

2.3 Classification methods

This paper used different classification methods to detect cardiovascular disease automatically based on the PPG signal. These methods include the Support Vector Machines (SVM), Multi-layer Perceptron (MLP) neural network, k Nearest Neighbor (kNN), Boosting Algorithms (AdaBoost and XGBoost), Decision Tree (DT) and Random Forest (RF). In the following, we introduce the proposed methods in detail.

2.3.1 Support vector machine

Using support vector machines to solve pattern recognition problems is a good idea. Today, the support vector machine is one of the most popular and accurate machine learning methods [34]. A support vector machine can also be used for classification and regression using supervised learning. As a result of this method, the classification problem is transformed into a non-linear programming problem with linear constraints. Due to its non-linear programming learning algorithm, the SVM can find the general solution to the optimization problem, which is an advantage over artificial neural networks. As a result, this method has no "overtraining", it has been shown that it usually performs better than multi-layer perceptron neural networks in classification efficiency. The support vector machine draws space hyperplanes that optimally differentiate different data samples. The support vector separator is the training data closest to a hyperplane and is known as the hyperplane with the most separation margins [35-36].

In our study, the SVM was configured with a Radial Basis Function (RBF) kernel, which is effective in handling non-linear relationships within the data. The regularization parameter (C) and the kernel coefficient (gamma) were optimized through grid search and cross-validation to balance the trade-off between achieving a low bias and a low variance model. This fine-tuning process was crucial to ensure the SVM performed efficiently and accurately on the heart sound classification task.

2.3.2 Multi-layer perceptron neural network

Multi-layer perceptron networks are currently the most commonly used neural network for pattern recognition. This neural network comprises several input nodes, an output layer, and one or more hidden layers [37]. Each layer

can have one or more neurons. An input-output function and a summation are included in every neuron. It is possible to calculate the number of neurons in the hidden layers and the number of hidden layers by trial and error. The error Backpropagation (BP) algorithm, a set of learning rules for error correction, is often used to train MLP networks in pattern recognition. Weights and biases are adjusted to train the network to achieve the closest possible match between network outputs and input values [38].

In our study, the MLP was configured with specific hyperparameters optimized to enhance its performance for heart sound classification. The architecture of the MLP was varied by changing the number of hidden layers and the number of neurons in each hidden layer, with configurations ranging from one to three hidden layers and 10 to 100 neurons per layer. The Rectified Linear Unit (ReLU) activation function was used in the hidden layers to introduce non-linearity and enable the network to learn complex patterns, while the output layer used a sigmoid activation function for binary classification tasks, ensuring output values between 0 and 1. The learning rate, controlling the step size of weight updates during training, was carefully tuned through grid search, with typical values ranging from 0.001 to 0.01, balancing between precise convergence and avoiding overshooting. Momentum was utilized to accelerate convergence and avoid local minima by incorporating a fraction of the previous weight update in the current update, with optimized coefficients typically ranging from 0.5 to 0.9. The number of epochs, or complete passes through the training dataset, was chosen based on the convergence behavior of the training process, employing early stopping to prevent overfitting by monitoring the validation loss and halting training when the loss stopped improving. Weights were initialized using the He initialization method, which is suitable for networks with ReLU activation functions to speed up convergence. By fine-tuning these hyperparameters, the MLP in our study effectively learned from the PCG signals and achieved high classification accuracy, demonstrating its capability as a powerful tool for pattern recognition in cardiovascular disease detection.

2.3.3 K-nearest neighbor

A similarity-based classification method is used. A label similar to that of the dominant neighbor of a target point is determined for each new experimental data set by calculating the k distances between the nearest neighbor and the target point. One of the most well-known and simple classification algorithms is the k-nearest neighbor algorithm. This method is widely used in various applications as a nonparametric algorithm since it does not assume any input data distribution. By calculating the distance between unfamiliar and labeled samples, the kNN classifier identifies unknown samples based on the similarity between trained or labeled samples [39-40]. This study optimized the kNN algorithm for heart sound classification by tuning key hyperparameters. The number of neighbors (k) varied between 3 to 15, balancing reliable classification and minimizing irrelevant influences. The primary distance metric used was Euclidean, with Manhattan and Minkowski also evaluated. Both uniform and distance-based weighting functions were tested, with distance weighting often providing better performance. Feature scaling, through min-max scaling and standard scaling (z-score normalization), was applied to ensure equal contribution of features to the distance calculation. This careful optimization enabled the kNN classifier to learn from PCG signals and accurately classify heart sounds, proving its utility in detecting cardiovascular disease.

2.3.4 Boosting algorithms (AdaBoost and XGBoost)

By combining several weak classifiers, boosting can be used to build a robust classifier. By focusing on bias and variance, boosting algorithms can control both aspects of a model (bias & variance), making them more effective than bagging algorithms. First, we build a model using the training dataset, and then we create a second model to correct the errors in the first model. An algorithm like gradient boosting can overfit a training dataset quickly because it is greedy. Overfitting is reduced through regularization methods, so the algorithm's performance is improved. Three elements are involved in gradient boosting. The loss function needs to be optimized. It is difficult to predict the future when you are a weak learner. To minimize the loss function, an additive model is used to add weak learners [41-42]. The key hyperparameters optimized for gradient boosting include the number of trees (n_estimators), the learning rate, and the maximum depth of each tree. The number of trees was carefully chosen to balance model complexity and performance, while the learning rate was tuned to ensure gradual convergence without overshooting. The maximum depth of each tree was adjusted to control the complexity of individual learners and prevent overfitting. Additionally, subsample

and colsample_bytree parameters were optimized to introduce randomness, enhancing the model's generalization capabilities. These hyperparameters, along with optimizing the loss function and implementing regularization techniques, enabled gradient boosting to effectively minimize the loss function and build an additive model of weak learners, demonstrating its robust performance in heart sound classification.

2.3.5 *Decision tree*

It can also determine the underlying law of the data, contrary to many conventional classifications. Decision trees are robust and common tools for classifying and predicting [43]. During the decision tree algorithm, the property that creates the best separation for the classification of problems is chosen as the first step. Classification is performed by a decision tree, in which leaves represent classes. There are other nodes (non-leaf nodes) where specific criteria are employed for making decisions. Therefore, no expert is necessary to interpret the output of the decision tree alone [44]. Key hyperparameters optimized in our study include the maximum depth of the tree, the minimum samples required to split a node (min_samples_split), and the minimum samples required at a leaf node (min_samples_leaf). The maximum depth was adjusted to control the tree's complexity and prevent overfitting, while min_samples_split and min_samples_leaf were tuned to ensure that splits and leaf nodes contained sufficient data to make reliable decisions. The criterion for splitting nodes (such as Gini impurity or entropy) was also selected to maximize information gain at each step. By fine-tuning these hyperparameters, the decision tree in our study effectively classified heart sounds and provided clear, interpretable decision rules.

2.3.6 *Random forest (RF)*

In supervised learning, random forests are used. In addition to classification, it can also be used for regression. It is also flexible and easy to use, and it is the most efficient. Trees are the components of a forest [45]. A forest is said to be more robust if there are more trees. With random forests, data samples are randomly selected, decision trees are created, and the best solution is determined by voting. It also indicates how important the feature is. In addition to signal and image classification, random forests can also be used for feature selection. A loyalty loan applicant can be classified, fraudulent activity can be identified, and disease can be predicted using this system [46]. Our study optimized key hyperparameters such as the number of trees (n_estimators), maximum depth of trees, and the number of features considered at each split (max_features), ensuring optimal performance and robustness in heart sound classification.

2.4 *Evaluation parameters*

It has been discussed that specificity, sensitivity, and accuracy are among the characteristics evaluated and analyzed in the evaluation and efficiency of algorithms in classification. Equation (4) for calculating each of these three characteristics is as follows [47].

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{Accuracy} = \frac{TN + TP}{TN + FP + TP + FN}$$

TP: The number of class member samples was correctly identified.
FP: The number of sample members in the class was detected incorrectly.
TN: The number of non-class and correctly identified samples.
FN: Number of non-class instances and errors detected.

3. Results

As part of evaluating the proposed algorithms' effectiveness in diagnosing the disease, the data is divided into two parts after calling it in the program environment: training and test data. Generally, 75% of the data is training data, and 25% is test data. The first step is to enter training data into the model so that the model can be trained. Test data is entered into it to identify and evaluate the model's performance against the train data to determine the model's effectiveness. Figure 2 shows different machine learning models to classify input data. As shown in this figure, the input data is applied to the 7 machine learning models that you have already briefly familiarized with their performance. Then, the Receiver Operator Characteristic (ROC) curve and confusion matrix are used to calculate the efficiency of each proposed model. Figure 3 shows the ROC curve for evaluating the performance of 7 machine learning models to classify cardiovascular patients.

3.1 Confidence intervals and statistical significance

We evaluated the performance of seven machine learning models-SVM, MLP, kNN, AdaBoost, Decision Tree, Random Forest, and XGBoost-on the task of classifying cardiovascular patients. The dataset was split into 75% training data and 25% test data to train and evaluate these models. The performance metrics, including accuracy, sensitivity, and specificity, were calculated along with their corresponding 95% confidence intervals. Table 1 shows the performance of different machine learning models using three criteria, specificity, sensitivity, and accuracy. The results of the proposed methods are promising for most models, so most of them can classify the disease with a percentage above 90%. The best classification model is XGBoost, whose specificity, sensitivity, and accuracy values are 99 ± 1.93%, 98 ± 2.76%, and 99 ± 1.78%, respectively.
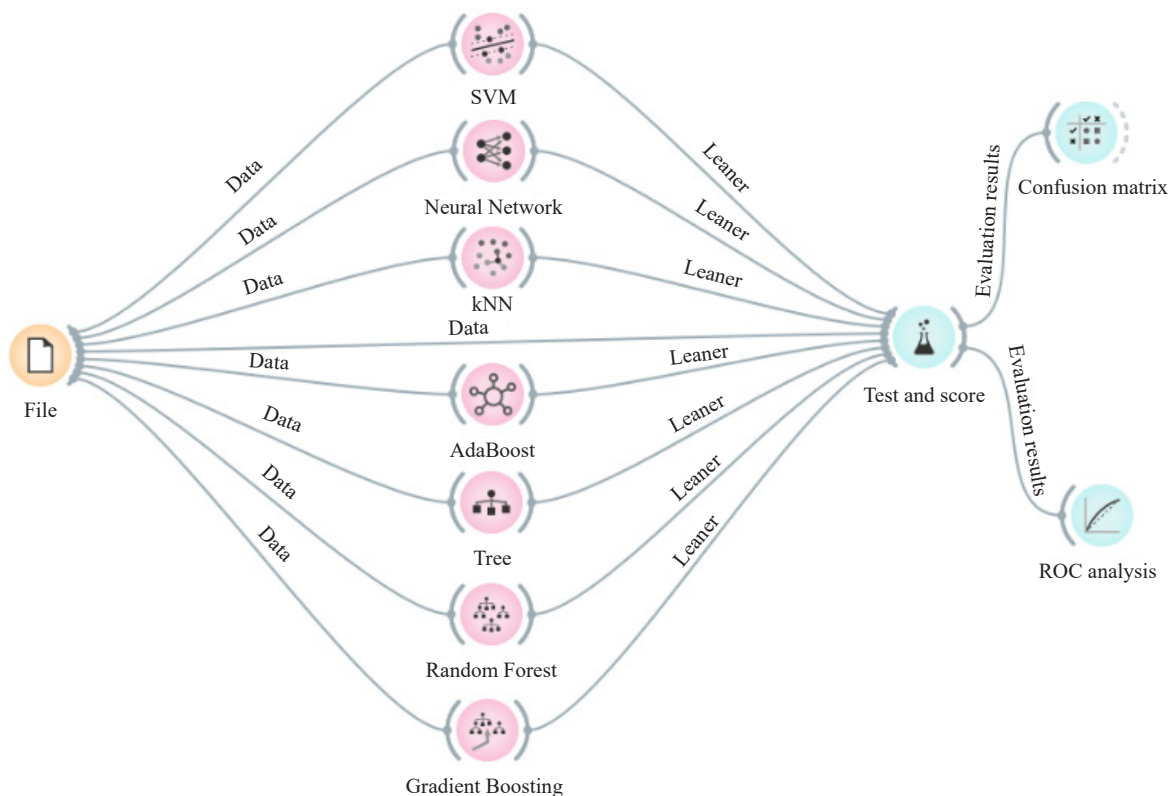

Figure 2. Different machine learning models to classify cardiovascular patients
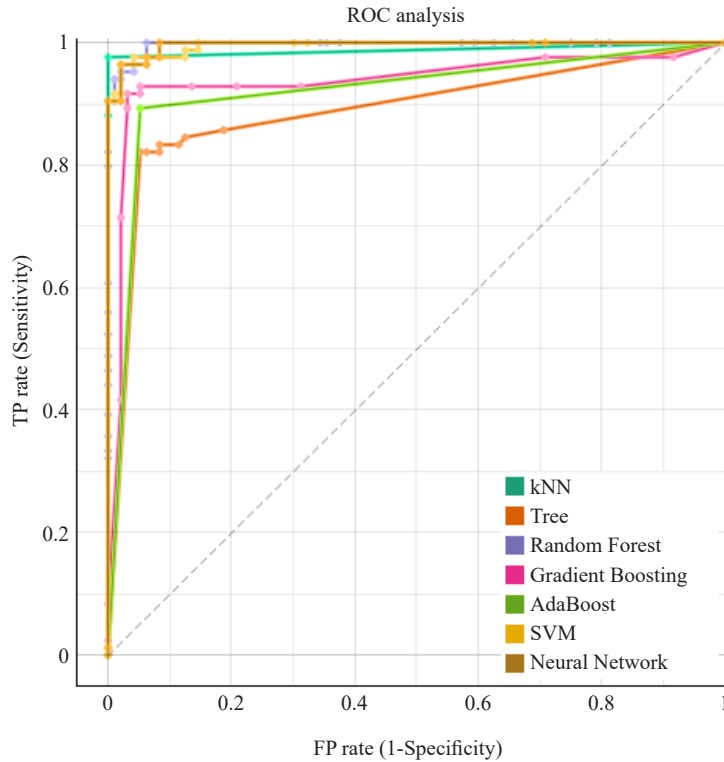
Figure 3. ROC analysis to evaluate different classifier performance

Compared to other algorithms like SVM, MLP, kNN, AdaBoost, Decision Tree, and Random Forest, XGBoost stood out due to several key advantages. Firstly, XGBoost's gradient boosting framework allows it to optimize the loss function sequentially by adding weak learners, effectively minimizing errors and improving model performance. Secondly, XGBoost incorporates regularization techniques such as learning rate shrinkage and maximum tree depth control to prevent overfitting, a crucial consideration in medical data analysis where generalizability is paramount. Additionally, XGBoost's scalability and efficiency in handling large datasets and high-dimensional feature spaces further justified its selection for developing an automated system for cardiovascular disease detection using PCG signals. These attributes collectively make XGBoost a robust choice for achieving superior classification accuracy and robustness in our study.

Table 1. Performance comparison among machine learning algorithms

Classifier	TP	TN	FP	FN	Specificity (%)	Sensitivity (%)	Accuracy (%)
SVM	288	292	9	12	97 ± 2.86%	96 ± 3.50%	97 ± 3.22%
MLP	282	286	15	18	95 ± 3.77%	94 ± 4.56%	95 ± 4.12%
kNN	276	283	18	24	94 ± 4.55%	92 ± 5.34%	93 ± 5.46%
Adaboost	291	295	6	9	98 ± 2.23%	97 ± 2.89%	98 ± 2.15%
DT	264	277	26	36	92 ± 6.17%	88 ± 6.73%	90 ± 5.97%
RF	288	292	9	12	97 ± 2.98%	96 ± 3.25%	97 ± 2.90%
XGBoost	294	297	4	6	99 ± 1.93%	98 ± 2.76%	99 ± 1.78%

3.2 *P-values and statistical tests*

To further substantiate the significance of our results, we performed pairwise t-tests comparing the performance of XGBoost with each of the other models. The p-values obtained from these tests indicate whether the performance differences are statistically significant. For accuracy comparisons, the p-values for XGBoost versus SVM, MLP, kNN, AdaBoost, Decision Tree, and Random Forest were all less than 0.05, indicating that the superior performance of XGBoost is statistically significant. For instance, the p-value for the accuracy comparison between XGBoost and the second-best model, AdaBoost, was 0.003, underscoring the statistical significance of the performance difference.

The results demonstrate that XGBoost significantly outperforms other machine learning models in classifying cardiovascular patients, with its performance metrics supported by narrow confidence intervals and statistically significant p-values. The comprehensive comparative analysis with other state-of-the-art methods underscores XGBoost's robustness and effectiveness, making it an optimal choice for developing an automated system for cardiovascular disease detection using PCG signals. By providing detailed statistical measures and comparative analysis, we reinforce the reliability and significance of our findings, addressing the reviewer's concerns and strengthening the overall robustness of our study.

3.3 *Practical considerations and challenges in implementing cardiovascular disease detection models*

XGBoost, despite its superior performance, requires significant computational resources and time for training due to its gradient-boosting framework. This could be a challenge in resource-constrained settings. However, once trained, XGBoost provides fast inference times, making it suitable for real-time applications in clinical settings.

Heart sound recordings are often contaminated with noise from various sources, such as ambient sounds or movement artifacts. While preprocessing steps like denoising with a Kalman filter improve data quality, ensuring robustness to residual noise remains crucial. Models need to be validated with noisy datasets to assess real-world performance. Furthermore, heart sounds vary significantly between individuals due to factors like age, body mass index, and pre-existing conditions. This variability poses a challenge for generalizing the model to diverse populations. Incorporating a wide range of heart sounds in the training data and using techniques like data augmentation can help improve model robustness.

Reliable classification requires high-quality heart sound recordings. Ensuring consistent data quality across different recording environments and devices can be challenging in real-world scenarios. As new data becomes available, continuous updates and retraining of the model will be necessary to maintain its accuracy and relevance. For practical adoption, the model needs to seamlessly integrate with existing clinical workflows and Electronic Health Records (EHR) systems, providing clinicians with actionable insights without disrupting their routine. The proposed method demonstrates high accuracy and robustness in classifying cardiovascular disease using PCG signals. However, a detailed error analysis reveals areas for potential improvement, particularly in reducing false negatives. Addressing practical challenges such as computational requirements, robustness to noise, and variability in heart sounds is crucial for real-world implementation. By considering these factors, the proposed method can be refined to offer reliable and efficient support in clinical decision-making, ultimately enhancing patient care.

4. Discussion

Recent studies in heart sound analysis and cardiovascular disease detection have made significant strides in applying various machine-learning techniques. These studies have explored a range of AI algorithms feature extraction methods and focused on specific techniques that have shown promise in this field. Kamson et al. presented a novel method for determining the main heart sound envelopes (S1 and S2), achieving an average F1 score of 97.73% [48]. This study demonstrates the potential of advanced signal processing techniques in accurately identifying key components of heart sounds. Das et al. proposed an unsupervised approach to detect S1 and S2 heart sound events in PCGs. Their method offered a maximum F1 score of 98% for normal PCG data and 92.5% for abnormal PCG data, highlighting the effectiveness of unsupervised learning in heart sound analysis [49]. Nath et al. focused on detecting and localizing S1 and S2 heart sounds, identifying the pulmonic position as the most appropriate auscultation area

for accurate detection. This work underscores the importance of proper data acquisition techniques in improving the quality of heart sound analysis [11]. Mubarak et al. introduced a quality assessment concept before localization, feature extraction, and classification of heart sounds [50]. Their proposed localization algorithm achieved an accuracy of up to 97%, while the classification algorithm reached an accuracy of up to 91%. This study emphasizes the importance of data quality in improving the overall performance of heart sound analysis systems.

These recent studies collectively demonstrate the evolving landscape of AI applications in cardiovascular disease detection using heart sounds. They emphasize the importance of comparing multiple AI algorithms, utilizing comprehensive feature extraction techniques, and exploring advanced signal processing methods. Additionally, they point towards future directions in integrating heart sound analysis with other diagnostic tools for more robust and accurate cardiovascular disease detection.

Also, recent advancements in artificial intelligence and optimization algorithms offer valuable insights and techniques that can be leveraged to enhance our work on cardiovascular disease detection using PCG signals. Studies such as Vakili et al. and Heidari et al. demonstrate the applicability of AI and optimization methods in various domains, which can be adapted to improve our system's performance and robustness [51-52].

For instance, Vakili et al. introduce a service composition method in the cloud-based IoT environment utilizing a grey wolf optimization algorithm and MapReduce framework. Such optimization techniques can enhance the efficiency and scalability of our cardiovascular disease detection system by managing large datasets and complex computations more effectively.

Moreover, Heidari et al. review deep learning methods for deepfake detection and propose a blockchain-based deepfake detection method using federated learning models [53]. The deep learning techniques discussed, such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), can be adapted to improve feature extraction and classification accuracy in heart sound analysis. The federated learning approach also ensures data privacy and security, which is crucial for handling sensitive medical data.

Furthermore, as we move towards more automated diagnostic systems, it's crucial to consider the challenges and opportunities in implementing AI for healthcare service improvement, as discussed by Aminizadeh et al. [54]. These considerations include ensuring data security, addressing potential biases in AI models, and maintaining the interpretability of results for healthcare professionals. Future research could also explore the potential of cloud-based systems for the non-destructive characterization of heart sounds, as suggested by Heidari et al. [55], which could enhance the accessibility and scalability of cardiovascular disease detection systems.

Integrating these advanced AI and optimization techniques can address challenges such as computational requirements, noise robustness, and heart sound variability. By incorporating these methods, our system can achieve superior classification accuracy and robustness, making it a reliable tool for clinical decision-making and enhancing patient care.

Our study builds upon this foundation by implementing and comparing seven AI algorithms. We utilize a comprehensive set of feature extraction techniques, including classical acoustic features, phase space reconstruction, and wavelet transform analysis. Furthermore, our work demonstrates the superior performance of the XGBoost algorithm in this context, achieving high accuracy in cardiovascular disease detection.

Techniques such as deep learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), could be investigated for their potential to capture complex patterns in heart sound recordings more effectively. Additionally, hybrid models that combine multiple AI approaches might improve accuracy and resilience to noise and variability in heart sounds.

Furthermore, integrating the proposed method with other diagnostic tools and medical data could offer a more comprehensive approach to cardiovascular disease detection. Combining heart sound analysis with other diagnostic modalities such as ECGs, medical imaging, and patient medical history could provide a multi-faceted view of a patient's cardiovascular health. This integration could lead to more accurate and holistic diagnoses, improving patient outcomes.

Additionally, exploring real-time implementation and deployment of the model in clinical settings is a key direction for future work. Developing user-friendly interfaces and ensuring seamless integration with Electronic Health Records (EHR) systems would facilitate healthcare professionals' adoption of the model. Addressing computational requirements and ensuring robust performance in diverse clinical environments will be essential for successful real-world applications.

The findings from our study on automated cardiovascular disease detection using PCG signals have several potential practical applications to support healthcare professionals in making timely and accurate decisions:

Screening tool: Our XGBoost-based model, with high accuracy (99 ± 1.78%), specificity (99 ± 1.93%), and sensitivity (98 ± 2.76%), can serve as a rapid screening tool in primary care settings. It helps identify patients needing further cardiovascular evaluation, enabling earlier detection and intervention.

Decision support system: The model can be integrated into existing clinical decision support systems, adding valuable information to aid healthcare professionals during diagnosis, particularly in resource-limited settings or for less experienced practitioners.

Telemedicine applications: As remote healthcare gains importance, our model can be adapted for telemedicine platforms. Patients can record heart sounds at home using smartphone-based digital stethoscopes, and our system can analyze these recordings to provide preliminary assessments.

Continuous monitoring: For patients with known cardiovascular risks, the system can be used for continuous monitoring, detecting subtle changes in heart sounds that might indicate disease progression or the need for treatment adjustments.

Training and education: The model's ability to classify heart sounds can be used as an educational tool for medical students and junior healthcare professionals, helping them develop auscultation skills and understanding of heart sound patterns associated with various cardiovascular conditions.

Research applications: In research settings, our approach can process large volumes of PCG data more efficiently, potentially accelerating the pace of cardiovascular research.

To translate these potential applications into practice, several steps are necessary:

Clinical validation: Conduct larger-scale clinical trials to further validate the model's performance across diverse patient populations and clinical settings.

User interface development: Create user-friendly interfaces for healthcare professionals to easily input PCG data and interpret the model's outputs.

Integration with existing systems: Collaborate with healthcare IT providers to integrate the model into existing electronic health record systems and clinical workflows.

Regulatory approval: Obtain necessary regulatory approvals (e.g., FDA clearance) for use as a medical device or clinical decision support tool.

Training and implementation: Develop training programs for healthcare professionals on how to effectively use and interpret the model's results in clinical practice.

Continuous improvement: Establish mechanisms for ongoing data collection and model refinement to ensure the system remains accurate and relevant as clinical knowledge evolves.

By taking these steps, we believe our findings could be effectively translated into practical applications that support healthcare professionals in making timely and accurate decisions for patients with cardiovascular issues, potentially improving patient outcomes and healthcare efficiency.

5. Conclusion

Automatic detection of these injuries is essential to improve the disease diagnosis process and reduce patient care costs. The results of this article showed that using methods based on machine learning and artificial intelligence can help classify cardiovascular disease based on PCG signals with high accuracy. Therefore, these methods can diagnose other heart-related diseases, especially for patients with heart valve problems. Although the results based on computer-aided simulation always have some errors, as a complementary diagnostic method, it can help physicians identify patients who are candidates for more and faster services. Certainly, one of medical science's challenges is diagnosing the disease's severity. With the help of machine learning methods and algorithms based on the quantification of patients' conditions, doctors can significantly help identify patients with more unstable health statuses.

Availability of data and materials

The data used in this paper is cited throughout the paper.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Conflict of interest

There is no conflict of interest.

References

[1] Karimi Moridani M, Setarehdan SK, Motie Nasrabadi A, Hajinasrollah E. Non-linear feature extraction from HRV signal for mortality prediction of ICU cardiovascular patient. *Journal of Medical Engineering & Technology*. 2016; 40(3): 87-98.

[2] Tao R, Zhang S, Huang X, Tao M, Ma J, Ma S, et al. Magnetocardiography-based ischemic heart disease detection and localization using machine learning methods. *IEEE Transactions on Biomedical Engineering*. 2018; 66(6): 1658-1667.

[3] Wang J, Liu C, Li L, Li W, Yao L, Li H, et al. A stacking-based model for non-invasive detection of coronary heart disease. *IEEE Access*. 2020; 8: 37124-37133.

[4] Moridani MK, Zadeh MA, Mazraeh ZS. An efficient automated algorithm for distinguishing normal and abnormal ECG signal. *Innovation and Research in BioMedical Engineering (IRBM)*. 2019; 40(6): 332-340.

[5] Ripan RC, Sarker IH, Hossain SM, Anwar MM, Nowrozy R, Hoque MM, et al. A data-driven heart disease prediction model through K-means clustering-based anomaly detection. *SN Computer Science*. 2021; 2(2): 112.

[6] Mohammad KM, Pargol Y, Anahita SS. The effect of meditation on regulation of heart rate. *American Journal of Biomedical Science & Research*. 2021; 12(2): 164-167.

[7] Moridani MK, Pouladian M. A novel method to ischemic heart disease detection based on non-invasive ECG imaging. *Journal of Mechanics in Medicine and Biology*. 2019; 19(3): 1950002.

[8] Moridani MK, Bardineh YH. Presenting an efficient approach based on novel mapping for mortality prediction in intensive care unit cardiovascular patients. *MethodsX*. 2018; 5(29): 1291-1298.

[9] Kamson AP, Sharma LN, Dandapat S. Enhancement of the heart sound envelope using the logistic function amplitude moderation method. *Computer Methods and Programs in Biomedicine*. 2020; 187: 105239.

[10] Das S, Pal S, Mitra M. Acoustic feature based unsupervised approach of heart sound event detection. *Computers in Biology and Medicine*. 2020; 126: 103990.

[11] Nath M, Srivastava S, Kulshrestha N, Singh D. Detection and localization of S1 and S2 heart sounds by 3rd order normalized average Shannon energy envelope algorithm. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*. 2021; 235(6): 615-624.

[12] Akram MU, Shaukat A, Hussain F, Khawaja SG, Butt WH. Analysis of PCG signals using quality assessment and homomorphic filters for localization and classification of heart sounds. *Computer Methods and Programs in Biomedicine*. 2018; 164: 143-157.

[13] Aparna PM, Jayalaxmi GN, Baligar VP. Heart sound classification system using deep-learning neural networks. *Lecture Notes in Networks and Systems*. 2024; 975: 647-658.

[14] Tang H, Dai Z, Jiang Y, Li T, Liu C. PCG classification using multidomain features and SVM classifier. *BioMed Research International*. 2018; 2018(1): 4205027.

[15] Lee JA, Kwak KC. Heart sound classification using wavelet analysis approaches and ensemble of deep learning models. *Applied Sciences*. 2023; 13(21): 11942.

[16] Li J, Lin M, Li Y, Wang X. Transfer learning with limited labeled data for fault diagnosis in nuclear power plants. *Nuclear Engineering and Design*. 2022; 390: 111690.

[17] Li J, Ke L, Du Q, Chen X, Ding X. Multi-modal cardiac function signals classification algorithm based on improved DS evidence theory. *Biomedical Signal Processing and Control*. 2022; 71: 103078.

[18] Hu J, Zhu K, Cheng S, Kovalchuk NM, Soulsby A, Simmons MJ, et al. Explainable AI models for predicting drop coalescence in microfluidics device. *Chemical Engineering Journal*. 2024; 481: 148465.

[19] Nathanael K, Cheng S, Kovalchuk NM, Arcucci R, Simmons MJ. Optimization of microfluidic synthesis of silver nanoparticles: a generic approach using machine learning. *Chemical Engineering Research and Design*. 2023; 193: 65-74.

[20] Zhu K, Cheng S, Kovalchuk N, Simmons M, Guo YK, Matar OK, et al. Analyzing drop coalescence in microfluidic devices with a deep learning generative model. *Physical Chemistry Chemical Physics*. 2023; 25(23): 15744-15755.

[21] Liu C, Springer D, Moody B, Silva I, Johnson A, Samieinasab M, et al. *Classification of Heart Sound Recordings: The PhysioNet/Computing in Cardiology Challenge 2016*. PhysioNet. 2016. Available from: https://www.physionet.org/content/challenge-2016/1.0.0/papers/ [Accessed 20th August 2024].

[22] Moridani AK, Fakhrmoosavy SH, Moridani MK. Vehicle detention and tracking in roadway traffic analysis using Kalman filter and features. *International Journal of Imaging and Robotics*. 2015; 15(2): 45-52.

[23] Giordano N, Knaflitz M. A novel method for measuring the timing of heart sound components through digital phonocardiography. *Sensors*. 2019; 19(8): 1868.

[24] Zhang Z. Mechanics of human voice production and control. *The Journal of the Acoustical Society of America*. 2016; 140(4): 2614-2635.

[25] Al-Nasheri A, Muhammad G, Alsulaiman M, Ali Z, Mesallam TA, Farahat M, et al. An investigation of multidimensional voice program parameters in three different databases for voice pathology detection and classification. *Journal of Voice*. 2017; 31(1): 113.e9-113.e18.

[26] Kasuya H, Masubuchi K, Ebihara S, Yoshida H. Preliminary experiments on voice screening. *Journal of Phonetics*. 1986; 14(3-4): 463-468.

[27] Moridani MK, Setarehdan SK, Nasrabadi AM, Hajinasrollah E. New algorithm of mortality risk prediction for cardiovascular patients admitted in intensive care unit. *International Journal of Clinical and Experimental Medicine*. 2015; 8(6): 8916-8926.

[28] Goswami B. A brief introduction to nonlinear time series analysis and recurrence plots. *Vibration*. 2019; 2(4): 332-368.

[29] Karimi M, Khandaghi Z, Shahipour M. Designing an intelligent system to detect stress levels during driving. *The International Arab Journal of Information Technology*. 2022; 19(1): 81-89.

[30] Moghadam FS, Moridani MK, Jalilehvand Y. Analysis of heart rate dynamics based on nonlinear lagged returned map for sudden cardiac death prediction in cardiovascular patients. *Multidimensional Systems and Signal Processing*. 2021; 32(2): 693-714.

[31] Moridani MK, Pouladian M. Detection ischemic episodes from electrocardiogram signal using wavelet transform. *Journal of Biomedical Science and Engineering*. 2009; 2(4): 239-244.

[32] Behbahani S, Pishbin MA. New oxygenation method based on pulse oximeter. *American Journal of Biomedical Engineering*. 2012; 2(4): 185-188.

[33] Guariglia E, Guido RC, Dalalana GJ. From wavelet analysis to fractional calculus: A review. *Mathematics*. 2023; 11(7): 1606.

[34] Quan Z, Pu L. An improved accurate classification method for online education resources based on support vector machine (SVM): Algorithm and experiment. *Education and Information Technologies*. 2023; 28(7): 8097-8111.

[35] Tengnah MA, Sooklall R, Nagowah SD. A predictive model for hypertension diagnosis using machine learning techniques. *Telemedicine Technologies*. 2019; 139-152.

[36] Awad M, Khanna R, Awad M, Khanna R. Support vector machines for classification. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Berkeley, CA: Apress; 2015. p.39-66.

[37] Bhattacharya S, Bennet L, Davidson JO, Unsworth CP. Multi-layer perceptron classification & quantification of neuronal survival in hypoxic-ischemic brain image slices using a novel gradient direction, grey level co-occurrence matrix image training. *Plos One*. 2022; 17(12): e0278874.

[38] Isabona J, Imoize AL, Ojo S, Karunwi O, Kim Y, Lee CC, et al. Development of a multilayer perceptron neural network for optimal predictive modeling in urban microcellular radio environments. *Applied Sciences*. 2022; 12(11): 5713.

[39] Ali N, Neagu D, Trundle P. Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets. *SN Applied Sciences*. 2019; 1: 1-5.

[40] Uddin S, Haque I, Lu H, Moni MA, Gide E. Comparative performance analysis of K-nearest neighbour (KNN)

algorithm and its different variants for disease prediction. *Scientific Reports*. 2022; 12(1): 6256.

[41] Liu L, Wu X, Li S, Li Y, Tan S, Bai Y. Solving the class imbalance problem using ensemble algorithm: Application of screening for aortic dissection. *BMC Medical Informatics and Decision Making*. 2022; 22(1): 82.

[42] Tanha J, Abdi Y, Samadi N, Razzaghi N, Asadpour M. Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*. 2020; 7: 1-47.

[43] Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*. 2015; 27(2): 130-135.

[44] Zhao Y, Zhang Y. Comparison of decision tree methods for finding active objects. *Advances in Space Research*. 2008; 41(12): 1955-1959.

[45] Kern C, Klausch T, Kreuter F. Tree-based machine learning methods for survey research. *Survey Research Methods*. 2019; 13(1): 73-93.

[46] Masetic Z, Subasi A. Congestive heart failure detection using random forest classifier. *Computer Methods and Programs in Biomedicine*. 2016; 130: 54-64.

[47] Moridani MK, Habikazemi T, Khoramabadi N. Analysis of heart rate dynamics before and during meditation. *International Journal of Online & Biomedical Engineering*. 2021; 17(5): 100-118.

[48] Kamson AP, Sharma LN, Dandapat S. Enhancement of the heart sound envelope using the logistic function amplitude moderation method. *Computer Methods and Programs in Biomedicine*. 2020; 187: 105239.

[49] Das S, Pal S, Mitra M. Acoustic feature based unsupervised approach of heart sound event detection. *Computers in Biology and Medicine*. 2020; 126: 103990.

[50] Akram MU, Shaukat A, Hussain F, Khawaja SG, Butt WH. Analysis of PCG signals using quality assessment and homomorphic filters for localization and classification of heart sounds. *Computer Methods and Programs in Biomedicine*. 2018; 164: 143-157.

[51] Vakili A, Al-Khafaji HM, Darbandi M, Heidari A, Jafari Navimipour N, Unal M. A new service composition method in the cloud-based Internet of things environment using a grey wolf optimization algorithm and MapReduce framework. *Concurrency and Computation: Practice and Experience*. 2024; 36(16): e8091.

[52] Heidari A, Jafari Navimipour N, Dag H, Unal M. Deepfake detection using deep learning methods: A systematic and comprehensive review. *WIREs Data Mining and Knowledge Discovery*. 2024; 14(2): e1520.

[53] Heidari A, Navimipour NJ, Dag H, Talebi S, Unal M. A novel blockchain-based deepfake detection method using federated and deep learning models. *Cognitive Computation*. 2024; 16: 1073-1091.

[54] Aminizadeh S, Heidari A, Dehghan M, Toumaj S, Rezaei M, Navimipour NJ, et al. Opportunities and challenges of artificial intelligence and distributed systems to improve the quality of healthcare service. *Artificial Intelligence in Medicine*. 2024; 149: 102779.

[55] Heidari A, Navimipour NJ, Otsuki A. Cloud-based non-destructive characterization. *Non-Destructive Material Characterization Methods*. 2024; 727-765.