IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 3, June 2026, pp. 2398~2409
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i3.pp2398-2409

Adaptive synthetic-based arrhythmia classification using machine learning techniques

Md. Rabiul Islam[1], Tapan Kumar Godder[2], Ahsan-Ul-Ambia[3], Ferdib-Al-Islam[4], Bulbul Ahamed[5], Anindya Nag[4], Ariful Islam[1]

[1]Department of Biomedical Engineering, Faculty of Engineering and Technology, Islamic University, Kushtia, Bangladesh
[2]Department of Information and Communication Technology, Faculty of Engineering and Technology, Islamic University, Kushtia, Bangladesh
[3]Department of Computer Science and Engineering, Faculty of Engineering and Technology, Islamic University, Kushtia, Bangladesh
[4]Department of Computer Science and Engineering, Faculty of Science and Engineering, Northern University of Business and Technology Khulna, Khulna, Bangladesh
[5]Department of Computer Science and Engineering, Faculty of Science and Engineering, Sonargaon University, Dhaka, Bangladesh

Article Info

Article history:

Received Feb 14, 2025
Revised Mar 31, 2026
Accepted Apr 22, 2026

Keywords:

Adaptive synthetic
Arrhythmia classification
Class imbalance
GridSearchCV
Machine learning
Paired t-test
Statistical significance test

ABSTRACT

Cardiac arrhythmias, characterized by irregular heart rhythms, pose a significant challenge for timely and accurate diagnosis. This paper presents an advanced framework for arrhythmia classification that addresses the class imbalance issue using adaptive synthetic (ADASYN) sampling combined with a wide-ranging set of machine learning (ML) techniques. Numerous classifiers are implemented, including logistic regression (LR), naive Bayes (NB), decision tree (DT), random forest (RF), k-nearest neighbors (KNN), multi-layer perceptron (MLP), gradient boosting (GB), adaptive boosting (AB), light gradient boosting machine (LGBM), and extreme gradient boosting (XGB). The experimental results demonstrate that RF, GB, LGBM, and XGB achieved a remarkable accuracy of 97%. Other models, such as DT and LR, also accomplished well, achieving 95% and 94% accuracy, respectively. KNN and NB yielded 93% and 81% accuracy, respectively, while AB underperformed with an accuracy of 24%. Precision scores across the models remained high, except for NB and AB. All models, except AB, demonstrated excellent recall. The proposed methodology outperforms previous works, setting a new benchmark for arrhythmia classification. These outcomes underscore the effectiveness of integrating ADASYN with ML techniques to enhance arrhythmia detection, with significant potential for improving clinical diagnostic processes and patient outcomes.

Corresponding Author:

Ferdib-Al-Islam
Department of Computer Science and Engineering, Faculty of Science and Engineering
Northern University of Business and Technology Khulna
Khulna–9100, Bangladesh
Email: ferdib.bsmrstu@gmail.com

1. INTRODUCTION

Arrhythmia is a common cardiac disorder that affects millions of people worldwide. It is characterized by an abnormal heart rhythm, which can be life-threatening if undiagnosed and untreated. Early detection and classification of arrhythmia are critical for active treatment and controlling of the condition. The traditional arrhythmia classification method electrocardiogram (ECG) is a non-invasive technique widely used to diagnose and classify arrhythmia [1]. These types of manual approaches are time-consuming and require specialized medical knowledge. However, the demands are accurate and efficient diagnosis and classification of arrhythmias. Arrhythmia is a cardiac complaint considered by irregular heartbeats, which can lead to severe impediments such as stroke, heart failure, or sudden cardiac arrest. Arrhythmias must be identified and categorized early to be effectively treated and managed.

Machine learning (ML) algorithms show great potential for arrhythmia classification, but their performance depends on the assortment of informative features. Recursive feature elimination (RFE) is a technique that can be applied to select appropriate features for classification, and this study aims to examine the performance of several ML models using this technique. The development of reliable and accurate automated arrhythmia classification systems can aid in the diagnosis and management of arrhythmias, ultimately improving patient outcomes. ML models have shown great potential in the classification of arrhythmia [2]. However, the performance of these algorithms depends on the selection of informative features. RFE is a procedure that can be used to select relevant features for classification [3], [4].

The researchers propose an arrhythmia classification procedure based on synthetic minority over-sampling technique (SMOTE) and feature selection [5]. The study aims to address the concern of imbalanced data in ECG arrhythmia classification, which can lead to low classification accuracy for minority classes. The proposed procedure applies SMOTE to produce synthetic samples to balance the data, and feature selection methods to select the most informative features for classification. The study demonstrates the potential of using SMOTE and feature selection techniques for accurate and efficient ECG arrhythmia classification [5]. Arrhythmia classification using data from the UC Irvine Machine Learning (UCI ML) repository has been extensively studied with various ML techniques to enhance diagnostic accuracy and efficiency. The classification of ECG signals, which are crucial for diagnosing arrhythmia, has been significantly enhanced using convolutional neural networks (CNNs), achieving a high accuracy of 99.88% with an ensemble of depth-wise separable convolutional (DSC) neural networks [6]. Random forest (RF) classifiers have also shown promising results, particularly in classifying atrial fibrillation (AF) and other heart rhythms, with an accuracy of 89% [7].

Comparative studies indicate that RF performs best for arrhythmia datasets, outperforming other classifiers like naive Bayes (NB), multi-layer perceptron (MLP), and support vector machine (SVM) [8]. Additionally, ensemble methods combining classifiers such as decision tree (DT), SVM, and RF have demonstrated superior performance on certain datasets [9]. Advanced techniques like the rider chaotic biography optimization (RCBO) algorithm integrated with deep stacked auto-encoders have been proposed to handle large datasets effectively, achieving high accuracy and specificity [10]. Moreover, the use of principal component analysis (PCA) for reducing features and wavelet packet decomposition (WPD) for feature extraction has been beneficial in improving classification accuracy [11]. The integration of ML with hardware-software complexes, such as mobile electronic cardiographs, has enabled real-time ECG analysis and arrhythmia detection, further enhancing diagnostic capabilities [12]. Techniques like genetic algorithm and artificial flora optimization have also been employed to optimize neural network topologies, reducing overfitting and improving generalization [13]. Lastly, teaching-learning-based optimization (TLBO) combined with k-means and fuzzy c-means clustering has shown high accuracy in classifying heart disease datasets, indicating the robustness of these methods in handling non-clustered data [14]. These advancements collectively contribute to the effective classification of arrhythmia using UCI ML repository data, highlighting the prospective of ML in healthcare diagnostics.

Deep learning now dominates ECG arrhythmia classification. CNNs learn hierarchical temporal features directly from raw beats or fixed-length windows and set early state-of-the-art (SOTA) on Massachusetts Institute of Technology–Beth Israel Hospital (MIT-BIH) and related corpora [15], [16]. Recurrent models (e.g., bidirectional long short-term memory (BiLSTM)) capture longer-range dependencies across beats and often boost minority-class sensitivity when the input is framed as a sequence [17]. More recently, transformer architectures model global context via self-attention on continuous ECG segments, with strong results for rhythm detection (including AF) without hand-crafted segmentation [18], [19]. At clinical scale, deep networks have matched or exceeded cardiologist performance for single-lead rhythm classification on ambulatory ECGs [20] and generalized to 12-lead clinical ECGs trained on millions of labeled exams, achieving high F1 and >99% specificity on multiple abnormalities [21].

Regarding real-world integration, large pragmatic studies show feasibility and utility. The Apple Heart Study (n ≈419k) validated smartwatch irregular-pulse notifications with a positive predictive value ~84% for AF when confirmed simultaneously by ECG patch monitoring [22]. In routine primary care, an AI-ECG decision-support tool embedded in the electronic health record (EHR) increased new diagnoses of low ejection fraction in a cluster-randomized trial (22,641 adults) [23]. Earlier work also demonstrated AI-ECG screening for systolic dysfunction using standard 12-lead ECGs in clinical workflows [24]. Together, these trends motivate the focus on robust imbalance handling (adaptive synthetic (ADASYN)) and ensemble learners as a complementary, cost-efficient avenue where deep models are not yet deployed or where dataset size, labeling cost, or compute are constrained.

The precise and prompt detection of arrhythmias is fundamental for effective clinical interference, yet it remains a stimulating task due to the complex and varied nature of arrhythmic patterns in ECG signals. Traditional diagnostic methods are not only time-consuming and subject to human error but also struggle with the class imbalance prevalent in arrhythmia datasets, where certain types of arrhythmias are significantly underrepresented. This imbalance can severely impair the performance of ML models, leading to biased predictions and poor generalization. Therefore, there is a pressing need for advanced computational approaches that can effectively handle imbalanced data and leverage the power of ML for robust arrhythmia classification. By integrating ADASYN and a comprehensive suite of ML models, this work aims to overcome these challenges, providing a more precise, reliable, and efficient solution for arrhythmia detection. This not only boosts indicative accuracy but also has the potential to considerably improve patient conclusions by facilitating timely and appropriate medical intervention. The main objective of the research is to deliver some effective techniques for the classification of arrhythmia using several ML models, in terms of accuracy and speed.

The main contributions of this system are—introduce a novel method for generating synthetic data tailored specifically for arrhythmia classification tasks. By leveraging adaptive techniques, the approach ensures that the synthetic data closely resembles real-world arrhythmia patterns, thereby enhancing the robustness and generalization proficiency of the classification model. Evaluate a range of SOTA ML models. Through rigorous experimentation and comparative analysis, the most suitable algorithm(s) is identified for achieving optimal classification accuracy and performance. Conduct comprehensive experiments using benchmark arrhythmia datasets to empirically validate the effectiveness of the proposed approach. By comparing the performance metrics such as accuracy, precision, and recall with existing methods, the superiority of the approach is demonstrated in accurately classifying different types of arrhythmias.

The continuing portions of paper are structured in following style. Section 2 outlines the approach used in this paper. Section 3 presents the findings of this research. Section 4 delivers the conclusion.

2. PROPOSED METHOD

The proposed system identifies whether a person's heart is healthy or suffering from arrhythmia disease. The implementation of this system has been done by using NumPy, Pandas, ADASYN, GridSearchCV, and Scikit-learn library. From Figure 1, it can be seen that, a robust framework is developed for arrhythmia classification that incorporates several key stages, including data preprocessing, class imbalance handling, and ML model application. Initially, the dataset underwent comprehensive preprocessing to handle missing values, confirming the integrity and comprehensiveness of the data. Following this, the data was split into training and testing sets to simplify model evaluation.

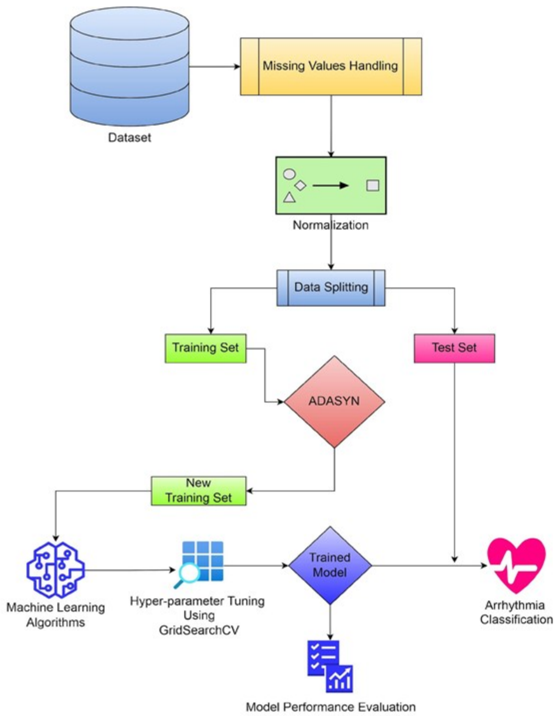

Figure 1. Workflow of the proposed method

To address the inherent class imbalance in arrhythmia datasets, ADASYN is employed, which generated synthetic samples for underrepresented classes within the training set, thus creating a more balanced dataset. Subsequently, a diverse array of ML models was applied to the balanced dataset, including logistic regression (LR), NB, DT, RF, k-nearest neighbors (KNN), MLP, gradient boosting (GB), adaptive boosting (AB), light gradient boosting machine (LGBM), and extreme gradient boosting (XGB). Each model's performance was rigorously assessed using accuracy, precision, recall, F1-score, Cohen's Kappa score, and Mathew's correlation coefficient (MCC) to determine their effectiveness in classifying arrhythmias. This comprehensive methodology ensured a thorough evaluation of different ML techniques in conjunction with advanced data preprocessing and balancing strategies.

2.1. Dataset description

This work uses the arrhythmia dataset, is found at the UCI ML repository [25]. There are 279 features and 452 records that make up the dataset. Each record is classified into one of the sixteen classes; a class label of 1 designates regular ECG patterns, while a class label ranging from 2 to 16 designates various arrhythmia kinds.

In this research, the normal class indicates a regular ECG pattern and the other 15 classes indicate cardiac arrhythmia. It should be noted that the UCI arrhythmia dataset does not provide raw ECG waveforms, but instead 279 pre-extracted numerical features that represent various clinical and signal-derived characteristics. Thus, the study is limited to working with these already processed features, and did not perform raw-signal–level preprocessing such as filtering or beat segmentation. This ensures reproducibility and comparability with prior works using the same dataset. Figure 2 demonstrates the imbalanced class distribution of the dataset.

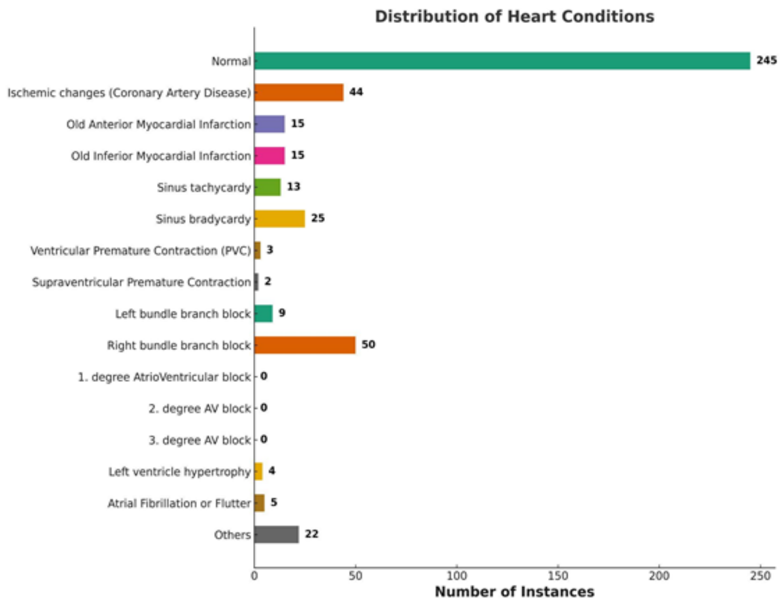

Figure 2. Imbalanced class distribution of the dataset

2.2. Data preprocessing

Since the dataset comprises extracted features rather than raw ECG signals, the preprocessing focused on ensuring data quality at the feature level. Missing values were identified and managed, while columns with constant values (all 0s or all 1s) were removed as non-informative. All remaining features were normalized using StandardScaler to achieve a comparable range.

2.2.1. Normalization

Normalizing or scaling feature variables in statistical evaluation and ML is crucial to ensure that they possess a comparable range and distribution. Utilizing this technique may enhance the efficiency of certain algorithms, while also facilitating the comparison and interpretation of coefficients associated with various characteristics. The process of data normalization is carried out on the feature values using the StandardScaler technique [26].

2.2.2. ADASYN

Data augmentation is a ML-based approach used to tackle the issue of class imbalance in datasets [27], [28]. Class imbalance refers to a situation where one class in a classification task is greatly unbalanced in the dataset. ADASYN utilizes a density distribution function to create artificial representations of the minority class. The value of the density distribution equation is adjusted according to the distribution of the data points, enabling more accurate production of synthetic samples that resemble the minority class. Synthetic samples are created for the minority class data points that provide challenges for the ML system to learn. By increasing the number of artificial samples for the underrepresented class, the disparity within the classes is reduced, enabling the ML algorithm to more effectively categorize the underrepresented class.

The arrhythmia dataset showed severe class imbalance, with the largest class (normal) having 245 samples while some minority classes contained fewer than 10. This imbalance (≈120:1) motivated the use of ADASYN. ADASYN is applied with k =5 nearest neighbors, as recommended in the original paper [27]. Synthetic samples were generated only within the training set, and the oversampling ratio was adaptively controlled so that harder-to-learn minority samples contributed more new instances.

2.2.3. GridSearchCV

In this study, GridSearchCV played a pivotal role in optimizing the performance of the ML models used for arrhythmia classification. GridSearchCV is a powerful hyperparameter tuning technique that systematically searches through a specified grid of hyperparameters to identify the optimal combination that maximizes model performance. By evaluating each combination of hyperparameters through cross-validation (CV), GridSearchCV ensures that the selected model configuration achieves the best possible performance on the given dataset [29], [30]. For this work, 5-fold GridSearchCV if applied to fine-tune the hyperparameters of various ML classifiers. Each model has its own set of hyperparameters that significantly influence its performance. The application of GridSearchCV involved defining a parameter grid for each classifier, which included a range of possible values for each hyperparameter. The tool then conducted an exhaustive search over the grid by fitting and evaluating the model for each combination of hyperparameters using CV. The CV process involved splitting the training data into several folds, training the model on some folds while validating it on the remaining fold, and then averaging the results to ensure robust and unbiased performance estimates.

2.3. Use of machine learning algorithms

Arrhythmia classification has been explored using various ML models, such as LR [31], NB [32], DT [33], RF [34], KNN [35], MLP [36], GB [37], AB [38], LGBM [39], and XGB [40]. Every method has distinct advantages and disadvantages, and the selection of an algorithm [41] is contingent upon the particular dataset and classification issue. To optimize model performance, GridSearchCV is applied with 5-fold CV across all classifiers. For each model, a grid of hyperparameter values was systematically explored, and the configuration yielding the best cross-validated accuracy was selected. Table 1 summarizes both the tested parameter ranges and the final chosen values.

Table 1. Hyper-parameters and tested ranges for ML models

Model	Hyper-parameter(s)	Tested range	Selected value
LR	max_iter, C	max_iter {500, 1000}, C {0.1, 1, 10}	max_iter =1000, C =1
NB	var_smoothing	{1e-9, 1e-8, 1e-7}	1e-9
DT	max_depth	{5, 10, 15}	10
RF	n_estimators, max_depth	n_estimators {100, 200, 300}, max_depth {None, 10, 20}	n_estimators =200, max_depth =None
KNN	n_neighbors	{3, 5, 7, 9}	5
MLP	hidden_layer_sizes, max_iter	hidden_layer_sizes {(50,), (100,)}, max_iter {500, 1000}	hidden_layer_sizes =(50,), max_iter =1000
GB	n_estimators, learning_rate	n_estimators {50, 100, 200}, learning_rate {0.05, 0.1, 0.2}	n_estimators =200, learning_rate =0.1
AB	n_estimators, learning_rate	n_estimators {50, 100, 200}, learning_rate {0.05, 0.1, 0.2}	n_estimators =200, learning_rate =0.1
LGBM	n_estimators, num_leaves	n_estimators {50, 100, 200}, num_leaves {31, 63}	n_estimators =50, num_leaves =31
XGB	n_estimators, max_depth, learning_rate	n_estimators {50, 100, 200}, max_depth {3, 6}, learning_rate {0.05, 0.1}	n_estimators =50, max_depth =3, learning_rate =0.1

3. RESULTS AND DISCUSSION

During the training process, the models were trained on the training data using CV to prevent overfitting. The training of the ML models has been done using CV, and evaluating the performance of the models using testing data. In this study, the dataset was split into 80-20% for training and testing, respectively, and the trained models were evaluated using the test data.

Figure 3 summarizes the comparative performance of the ML models across six evaluation metrics. Ensemble classifiers such as XGB, GB, LGBM, and RF achieved the highest scores, each recording 97% for accuracy, precision, recall, and F1-score, with strong agreement indices (MCC =0.90-0.96 and Cohen's Kappa =0.91-0.95), confirming their robustness. DT followed closely with 95% across all four primary metrics and solid agreement (MCC =0.94, Kappa =0.92). LR and MLP maintained balanced performance at 94% with MCC values of 0.93 and 0.88, and Kappa scores of 0.90 and 0.86, respectively. KNN also performed reliably with 93-94% and high agreement (MCC =0.90, Kappa =0.92). In contrast, NB produced only moderate results (81-82%, MCC =0.78, Kappa =0.75), while AB underperformed significantly, reaching just 24% accuracy and an MCC of 0.02 with a Kappa of 0.08. These findings clearly highlight the superiority of ensemble-based models in arrhythmia classification compared to traditional classifiers.

To validate the significance of model performance differences, paired t-tests were performed comparing AB with the top-performing models. The results confirmed that AB's accuracy was significantly lower ($p < 0.01$ in all cases). Specifically, comparisons yielded: RF vs AB ($t = -148.60$, $p = 1.23\times10^{-8}$), GB vs AB ($t = -148.60$, $p = 1.23\times10^{-8}$), LGBM vs AB ($t = -141.99$, $p = 1.48\times10^{-8}$), and XGB vs AB ($t = -194.57$, $p = 4.19\times10^{-9}$). The tests revealed statistically significant differences ($p < 0.01$), approving that AB's markedly lower accuracy (24%) is not due to random variation but represents a genuine performance gap. Table 2 shows the analysis by reporting 95% confidence intervals (CI) for the accuracy differences between AB and the top-performing models (RF, GB, LGBM, and XGB). All CI exclude zero. This reinforces our conclusion that AB's poor performance is inherent to its sensitivity under ADASYN-based oversampling, while ensemble methods maintain stable and superior accuracy.

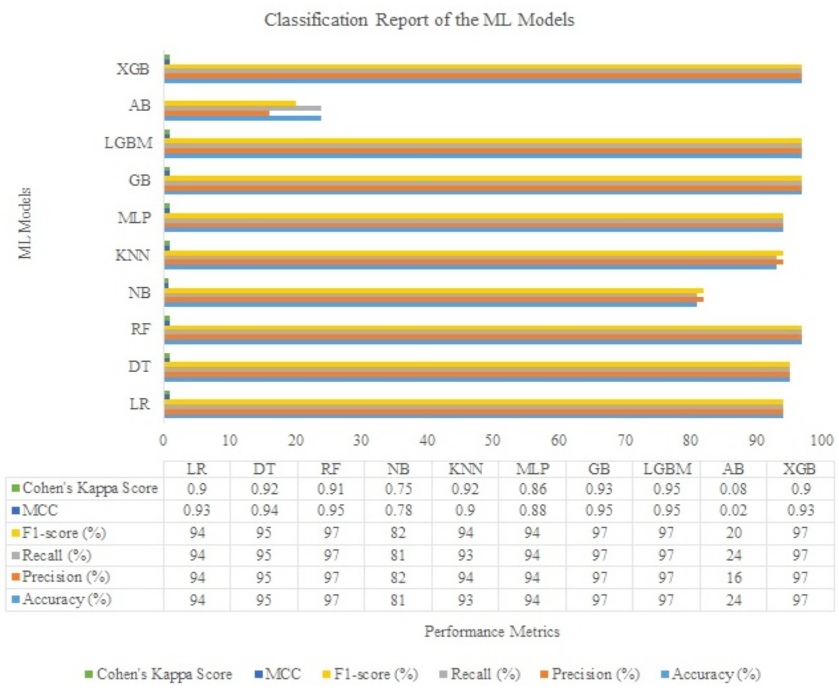

Figure 3. Performances of ML models

Table 2. Paired t-test results with 95% CI (AB vs. top models)

Comparison	t-value	p-value	95% CI of accuracy difference
RF vs AB	-148.60	1.23×10^{-8}	[0.71, 0.75]
GB vs AB	-148.60	1.23×10^{-8}	[0.71, 0.75]
LGBM vs AB	-141.99	1.48×10^{-8}	[0.70, 0.74]
XGB vs AB	-194.57	4.19×10^{-9}	[0.72, 0.76]

Figure 4 presents the area under the curve (AUC)-receiver operating characteristic (ROC) of the ML models, where ensemble methods achieved the best performance. XGB attained the highest AUC (0.99), followed closely by LGBM and GB (0.98 each), while RF also performed strongly with 0.97. Moderate results were observed for DT (0.94), KNN (0.92), MLP (0.91), and LR (0.88), whereas NB showed weaker discrimination (0.82). In contrast, AB performed poorly with an AUC of only 0.30, confirming its instability with ADASYN balancing. These findings emphasize the robustness of ensemble-based classifiers for arrhythmia classification. Figure 5 illustrates the confusion matrix for the RF model which performed best in the aforementioned metrics in classifying various heart conditions. The matrix demonstrates the true positives, false positives, false negatives, and true negatives for each heart condition, with color intensity reflecting the number of instances for each classification. The diagonal elements, such as 67 for “normal” and 82 for “right bundle branch block,” highlight the RF model’s strong performance in correctly classifying the most common conditions. However, some misclassifications are visible, particularly for “left bundle branch block” (56 misclassified instances) and “AF or flutter” (3 misclassifications). Despite these misclassifications, the RF model achieves high accuracy in predicting the majority of conditions, with certain less frequent categories showing room for further improvement.

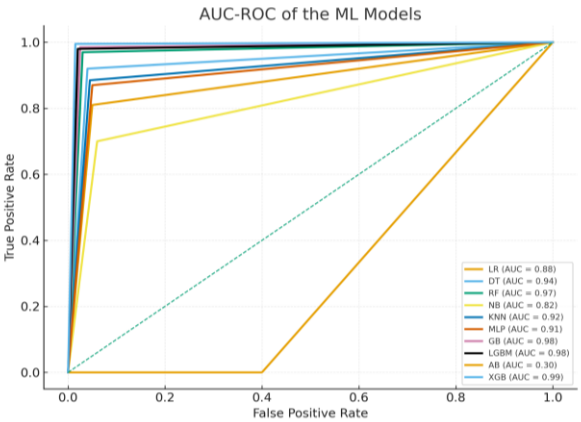

Figure 4. AUC-ROC curve of the ML models

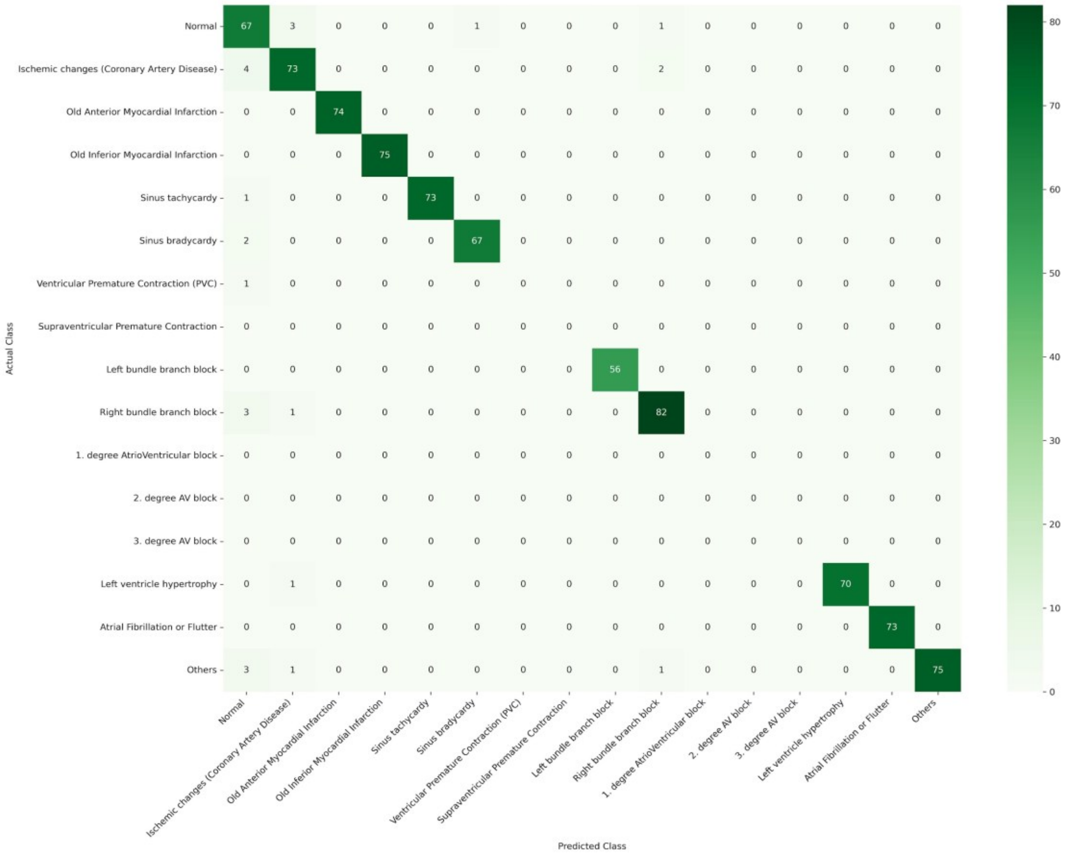

Figure 5. Confusion matrix for the best (RF) model

Figure 6 shows the prediction probabilities for a specific instance, highlighting the model's output for the predicted class (4) with a probability of 0.99. The table on the right lists the values of key features that most influence the prediction. For instance, feature 124 (with a value of 36.00) has the most significant impact on the prediction, while features such as 276 and 196 also contribute substantially. This visualization demonstrates the interpretability of the model, offering insights into which features drive the classification decision for this specific case.

Table 3 demonstrates the comparative analysis of the proposed work with the previous works. The results of this study demonstrate a significant improvement in the field of arrhythmia classification compared to previous works. By addressing the class imbalance problem using ADASYN and leveraging a comprehensive suite of ML models, superior classification performance was achieved. In previous works, classifiers such as PCA with SVM and RF achieved accuracies of 91.2% [2] and 89% [7], respectively, while NB reported a lower accuracy of 85% [32]. GB models like XGB and AB yielded higher accuracies of 95.65% [37] and 94.15% [38], respectively, demonstrating their strong predictive capability. However, RF alone in previous work performed at a suboptimal level with 84% [42] accuracy. In contrast, the proposed methodology using ADASYN significantly boosts the performance of several models. For instance, RF, GB, LGBM, and XGB all achieved an impressive accuracy of 97%. This marks a substantial improvement over their counterparts in previous works, particularly for RF, which improved from 84% to 97%, highlighting the impact of addressing the class imbalance issue.

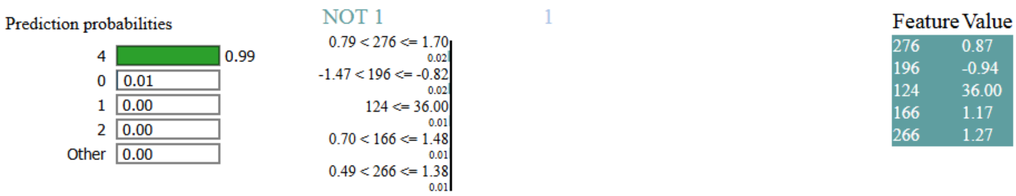

Figure 6. Prediction analysis using local interpretable model-agnostic explanations (LIME) for RF model

Table 3. Performance comparison of the proposed work with existing works

Reference	Method used	Accuracy (%)
[2]	PCA+SVM	91.2
[7]	PCA+RF	89
[32]	NB	85
[37]	XGB	95.65
[38]	AB	94.15
[42]	RF	84
Proposed work	ADASYN+RF	97
	ADASYN+GB	97
	ADASYN+LBGM	97
	ADASYN+XGB	97

Additionally, DT and LR models in the proposed framework also performed well, achieving 95% and 94% accuracy, respectively, which is superior to most previous approaches. Interestingly, while most classifiers in the proposed work performed exceptionally, NB experienced a drop in performance, with an accuracy of 81%. This is likely due to NB's assumption of feature independence, which may not hold true for the complex patterns present in arrhythmia data. Furthermore, KNN performed reasonably well, with 93% accuracy, although it did not reach the same high level as other models like RF or GB. One notable observation is the poor performance of AB in the proposed framework, which resulted in a significantly low accuracy of 24%. This sharp decline compared to prior work (94.15%) suggests that AB may be particularly sensitive to the synthetic samples generated by ADASYN, leading to overfitting or ineffective learning.

A drawback of this work is the restricted variability of the UCI arrhythmia dataset. Real-world ECG recordings are influenced by noise, baseline drift, electrode placement, and device-specific artifacts. In addition, patient populations differ by age, comorbidities, and ethnicity, which may affect ECG morphology and arrhythmia expression. Because this dataset lacks such diversity, the results may not fully generalize to heterogeneous clinical populations. Overall, the proposed approach has demonstrated that the integration of ADASYN with ML models, particularly ensemble methods like RF, GB, and XGB, is highly effective for arrhythmia classification. The ability to handle imbalanced datasets with ADASYN contributed to superior classification performance, making this methodology more reliable for real-world medical applications compared to earlier approaches.

Beyond algorithmic evaluation, an important consideration is clinical applicability. While this study employed the UCI arrhythmia dataset of pre-extracted features, the proposed framework can be readily adapted for real-time use. ECG devices such as wearable patches, Holter monitors, and smartwatches can continuously record raw ECG signals, which are then transformed into numerical features through embedded signal processing modules. These features could be directly fed into the ADASYN+ML pipeline, enabling real-time classification of arrhythmias at the point of care. In hospital settings, integration with EHR systems could provide clinicians with automated alerts during patient monitoring, improving early detection and diagnosis. Furthermore, this framework is computationally lightweight compared to deep learning methods, making it fit for deployment on portable or low-power devices. Challenges for clinical translation remain, including the need for prospective validation on raw ECG signals, ensuring interpretability of model outputs, and addressing regulatory and ethical issues related to data privacy and safety. Nevertheless, these scenarios highlight the real-world potential of this approach to complement existing clinical workflows.

An additional consideration is the feasibility of deployment on edge devices. Because the models are based on lightweight ML algorithms, they can be integrated into mobile ECG monitors, wearable devices, or bedside hospital monitors without requiring high-performance graphics processing units (GPUs). Embedded processors (e.g., advanced RISC machines (ARM)-based chips in smartwatches or Holter monitors) can execute such models in real time, enabling continuous monitoring and early arrhythmia detection at the patient's side. In hospital workflows, integration with monitoring systems or EHR-linked alerts could support clinicians by providing automated risk stratification alongside routine ECG recording. Moreover, while the experiments used the UCI arrhythmia dataset, future work should extend to external validation on widely used benchmark datasets such as MIT-BIH, pulmonary tuberculosis (PTB) diagnostic ECG, PhysioNet challenge datasets or on prospectively collected hospital ECG data. Such external testing will be critical to ensure generalizability across populations, device settings, and noise environments.

4. CONCLUSION

This paper proposed an enhanced framework for arrhythmia classification that addresses the critical issue of class imbalance using ADASYN combined with a comprehensive set of ML algorithms. Through rigorous experimentation, the effectiveness of this approach was demonstrated across multiple classifiers, achieving substantial improvements in accuracy compared to previous works. Notably, ensemble models such as RF, GB, LGBM, and XGB achieved an outstanding accuracy of 97%, marking a significant enhancement over past results. Other models, including DT (95%), LR (94%), and MLP (94%), also showed strong performance, underscoring the robustness of the methodology. Despite the overall success, some models like NB and KNN underperformed relative to others, achieving 81% and 93% accuracy, respectively. Moreover, AB exhibited an unexpectedly low accuracy of 24%, suggesting that it may not be well-suited for datasets with synthetic samples generated by ADASYN. The results of this study demonstrate the importance of addressing class imbalance in arrhythmia classification tasks, and the integration of ADASYN with advanced ML techniques proved to be a highly effective solution. This work outperforms prior studies and sets a new benchmark for arrhythmia detection and classification. Future research will focus on validating the framework using raw ECG signals from wearable and clinical monitoring devices, exploring additional data resampling techniques; further refining model optimization, and testing the proposed framework in real-time clinical settings to enhance its applicability in cardiac care diagnostics. Future research will also focus on deploying the framework on edge devices such as wearable ECG monitors and hospital bedside units. Given the relatively low computational footprint of this approach, real-time inference on portable devices is feasible. In addition, validation on external benchmark datasets and unseen hospital data will be prioritized to confirm robustness and clinical utility. It is also acknowledged that the UCI dataset used here lacks the variability and patient diversity encountered in real-world practice. Extending the framework to larger, heterogeneous datasets will be essential for ensuring clinical reliability.

FUNDING INFORMATION

The authors received no funding for this research.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Md. Rabiul Islam	✓	✓	✓	✓	✓			✓	✓				✓	✓
Tapan Kumar Godder		✓		✓		✓	✓			✓		✓	✓	
Ahsan-Ul-Ambia		✓		✓		✓	✓			✓		✓		
Ferdib-Al-Islam	✓	✓	✓	✓	✓				✓	✓	✓		✓	
Bulbul Ahamed	✓		✓			✓	✓		✓					✓
Anindya Nag	✓	✓	✓					✓	✓		✓			
Ariful Islam					✓	✓	✓			✓		✓		✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

The data used in this study are available in UCI ML repository at https://doi.org/10.24432/C5BS32, reference [25].

REFERENCES

[1] S. K. Berkaya, A. K. Uysal, E. S. Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, "A survey on ECG analysis," *Biomedical Signal Processing and Control*, vol. 43, pp. 216–235, May 2018, doi: 10.1016/j.bspc.2018.03.003.

[2] P. Shimpi, S. Shah, M. Shroff, and A. Godbole, "A machine learning approach for the classification of cardiac arrhythmia," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, Jul. 2017, pp. 603–607, doi: 10.1109/ICCMC.2017.8282537.

[3] A. Bahl *et al.*, "Recursive feature elimination in random forest classification supports nanomaterial grouping," *NanoImpact*, vol. 15, Mar. 2019, doi: 10.1016/j.impact.2019.100179.

[4] A. Jovic, K. Brkic, and N. Bogunovic, "A review of feature selection methods with applications," in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, May 2015, pp. 1200–1205, doi: 10.1109/MIPRO.2015.7160458.

[5] T. Wang, P. Chen, T. Bao, J. Li, and X. Yu, "Arrhythmia classification algorithm based on SMOTE and feature selection," *International Journal of Performability Engineering*, vol. 17, no. 3, pp. 263–275, 2021, doi: 10.23940/ijpe.21.03.p2.263275.

[6] P. Giudici, A. Gramegna, and E. Raffinetti, "Machine learning classification model comparison," *Socio-Economic Planning Sciences*, vol. 87, Jun. 2023, doi: 10.1016/j.seps.2023.101560.

[7] S. G. Chickaramanna *et al.*, "Classification of arrhythmia using machine learning algorithm," *Revue d'Intelligence Artificielle*, vol. 36, no. 4, pp. 529–534, Aug. 2022, doi: 10.18280/ria.360403.

[8] E. Haerani, F. Syafria, F. Lestari, N. Novriyanto, and I. Marzuki, "Classification academic data using machine learning for decision making process," *Journal of Applied Engineering and Technological Science*, vol. 4, no. 2, pp. 955–968, Jun. 2023, doi: 10.37385/jaets.v4i2.1983.

[9] S. S., K. H. Nehemiah, and K. A., "A framework to classify clinical data using a genetic algorithm and artificial flora-optimized neural network," *International Journal of Swarm Intelligence Research*, vol. 13, no. 1, pp. 1–22, Jul. 2022, doi: 10.4018/IJSIR.304719.

[10] J.-G. Choi, I. Ko, and S. Han, "Depression level classification using machine learning classifiers based on actigraphy data," *IEEE Access*, vol. 9, pp. 116622–116646, 2021, doi: 10.1109/ACCESS.2021.3105393.

[11] H. Javaid, E. Kumarnsit, and S. Chatpun, "Classification techniques in machine learning for age-related electroencephalography data analysis," *PSU Medical Journal*, vol. 3, no. 2, pp. 103–113, Jul. 2023, doi: 10.31584/psumj.2023255208.

[12] A. Farooq, M. Seyedmahmoudian, and A. Stojcevski, "A wearable wireless sensor system using machine learning classification to detect arrhythmia," *IEEE Sensors Journal*, vol. 21, no. 9, pp. 11109–11116, May 2021, doi: 10.1109/JSEN.2021.3062395.

[13] P. Gautam, M. Singh, and B. R. Mukindrao, "Machine learning-based arrhythmia classification," in *Integrating Digital Health Strategies for Effective Administration*, IGI Global Scientific Publishing, 2023, pp. 345–377, doi: 10.4018/978-1-6684-8337-4.ch016.

[14] K. Perumalla, A. Soni, R. Dey, and S. Rich, "ZeroIn: characterizing the data distributions of commits in software repositories," Apr. 2022, *arXiv: 2204.07863*.

[15] U. R. Acharya *et al.*, "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, pp. 389–396, Oct. 2017, doi: 10.1016/j.compbiomed.2017.08.022.

[16] T. F. Romdhane, H. Alhichri, R. Ouni, and M. Atri, "Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss," *Computers in Biology and Medicine*, vol. 123, Aug. 2020, doi: 10.1016/j.compbiomed.2020.103866.

[17] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Computers in Biology and Medicine*, vol. 96, pp. 189–202, May 2018, doi: 10.1016/j.compbiomed.2018.03.016.

[18] R. Hu, J. Chen, and L. Zhou, "A transformer-based deep neural network for arrhythmia detection using continuous ECG signals," *Computers in Biology and Medicine*, vol. 144, May 2022, doi: 10.1016/j.compbiomed.2022.105325.

[19] M.-U. Yang, D.-I. Lee, and S. Park, "Automated diagnosis of atrial fibrillation using ECG component-aware transformer," *Computers in Biology and Medicine*, vol. 150, Nov. 2022, doi: 10.1016/j.compbiomed.2022.106115.

[20] A. Y. Hannun *et al.*, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, no. 1, pp. 65–69, Jan. 2019, doi: 10.1038/s41591-018-0268-3.

[21] A. H. Ribeiro *et al.*, "Automatic diagnosis of the 12-lead ECG using a deep neural network," *Nature Communications*, vol. 11, no. 1, Apr. 2020, doi: 10.1038/s41467-020-15432-4.

[22] M. V. Perez *et al.*, "Large-scale assessment of a smartwatch to identify atrial fibrillation," *New England Journal of Medicine*, vol. 381, no. 20, pp. 1909–1917, Nov. 2019, doi: 10.1056/NEJMoa1901183.

[23] X. Yao *et al.*, "Artificial intelligence–enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial," *Nature Medicine*, vol. 27, no. 5, pp. 815–819, May 2021, doi: 10.1038/s41591-021-01335-4.

[24] Z. I. Attia *et al.*, "Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram," *Nature Medicine*, vol. 25, no. 1, pp. 70–74, Jan. 2019, doi: 10.1038/s41591-018-0240-2.

[25] H. Guvenir, B. Acar, H. Muderrisoglu, and R. Quinlan, "Arrhythmia [Dataset]," *UCI Machine Learning Repository*, 1997, doi: 10.24432/C5BS32.

[26] A. Apicella, F. Isgrò, A. Pollastro, and R. Prevete, "On the effects of data normalization for domain adaptation on EEG data," *Engineering Applications of Artificial Intelligence*, vol. 123, Aug. 2023, doi: 10.1016/j.engappai.2023.106205.

[27] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Jun. 2008, pp. 1322–1328, doi: 10.1109/IJCNN.2008.4633969.

[28] Y. Pristyanto, A. F. Nugraha, A. Dahlan, L. A. Wirasakti, A. A. Zein, and I. Pratama, "Multiclass imbalanced handling using ADASYN oversampling and stacking algorithm," in *2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, Jan. 2022, pp. 1–5, doi: 10.1109/IMCOM53663.2022.9721632.

[29] G. N. Ahmad, H. Fatima, S. Ullah, A. S. Saidi, and Imdadullah, "Efficient medical diagnosis of human heart diseases using machine learning techniques with and without GridSearchCV," *IEEE Access*, vol. 10, pp. 80151–80173, 2022, doi: 10.1109/ACCESS.2022.3165792.

[30] E. M. Senan *et al.*, "Diagnosis of chronic kidney disease using effective classification algorithms and recursive feature elimination techniques," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–10, Jun. 2021, doi: 10.1155/2021/1004767.

[31] X. Zou, Y. Hu, Z. Tian, and K. Shen, "Logistic regression model optimization and case analysis," in *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, Dalian, China, Oct. 2019, pp. 135–139, doi: 10.1109/ICCSNT47585.2019.8962457.

[32] A. N. Putra, Kemalasari, and H. Oktavianto, "A portable and mobile systems for identifying cardiac arrhythmia using naïve Bayes," in *2023 International Electronics Symposium (IES)*, Aug. 2023, pp. 173–178, doi: 10.1109/IES59143.2023.10242509.

[33] S. Lee, C. Lee, K. G. Mun, and D. Kim, "Decision tree algorithm considering distances between classes," *IEEE Access*, vol. 10, pp. 69750–69756, 2022, doi: 10.1109/ACCESS.2022.3187172.

[34] F.-Al-Islam, M. S. Sanim, K. M. Hasan, M. M. Alam, M. A. A. Walid, and M. R. Islam, "Crop-RecFIS: machine learning classifiers for crop recommendation and feature importance scores calculation," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, Jul. 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.10307833.

[35] F.-Al-Islam and L. Akter, "Early identification of Parkinson's disease from hand-drawn images using histogram of oriented gradients and machine learning techniques," in *2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*, Dec. 2020, pp. 1–6, doi: 10.1109/ETCCE51779.2020.9350870.

[36] S. Savalia and V. Emamian, "Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks," *Bioengineering*, vol. 5, no. 2, May 2018, doi: 10.3390/bioengineering5020035.

[37] R. Linghu and K. Zhang, "Real-time automatic arrhythmia detection system based on extreme gradient boosting and neural network algorithm," *Journal of Physics: Conference Series*, vol. 2449, no. 1, Mar. 2023, doi: 10.1088/1742-6596/2449/1/012033.

[38] B. Zhang, J. Wen, and H. Ren, "A classification method of arrhythmia based on AdaBoost algorithm," *Journal of Physics: Conference Series*, vol. 1682, no. 1, Nov. 2020, doi: 10.1088/1742-6596/1682/1/012060.

[39] Q. H. Hoang, X. B. Tran, L. H. Nguyen, A. V. Tran, and V. D. Nguyen, "Electrocardiography anomaly detection," in *2023 International Conference on Advanced Technologies for Communications (ATC)*, Oct. 2023, pp. 404–408, doi: 10.1109/ATC58710.2023.10318912.

[40] L. Akter and F.-Al-Islam, "Dementia identification for diagnosing Alzheimer's disease using XGBoost algorithm," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, Feb. 2021, pp. 205–209, doi: 10.1109/ICICT4SD50815.2021.9396777.

[41] S. Dhyani, A. Kumar, and S. Choudhury, "Analysis of ECG-based arrhythmia detection system using machine learning," *MethodsX*, vol. 10, 2023, doi: 10.1016/j.mex.2023.102195.

[42] J. Chittoria, S. S. Kamath, and V. Mayya, "Detection of cardiac arrhythmia using machine learning approaches," in *2022 IEEE Region 10 Symposium (TENSYMP)*, Jul. 2022, pp. 1–6, doi: 10.1109/TENSYMP54529.2022.9864533.

BIOGRAPHIES OF AUTHORS

Md. Rabiul Islam is pursuing his Ph.D. from the Department of Biomedical Engineering, Islamic University. He is currently working as an assistant professor in the same department. He has more than 15 publications both in national and international journals. He has also been supervising enormous number of students both in graduate and postgraduate level. His research interests include ML, bio-signal processing, medical image processing, prosthetics, and orthotics. He can be contacted at email: rabinub@gmail.com.

Dr. Tapan Kumar Godder graduated from the Department of Applied Physics and Electronics, Rajshahi University, Bangladesh and obtained his post-graduate degree from the same in 1995 respectively. He was awarded Master of Philosophy and Doctor of Philosophy from the Faculty of Applied Science and Technology, Islamic University, Bangladesh in 2007 and 2012 respectively. He has more than 40 publications in national and international journals and conferences. He can be contacted at email: tkict@yahoo.com.

Dr. Ahsan-Ul-Ambia graduated from the Department of Applied Physics and Electronics, Rajshahi University, Bangladesh and did his post-graduate degree from the same in 1994. He obtained Doctor of Philosophy degree from the Graduate School of Science and Technology, Shizuoka University Japan in 2009. He has more than 35 publications in national and international journals and conferences. He is also the author of some renowned books. He was the chair of the Department of Computer Science and Engineering and the Director of Information and Communication Technology Cell in Islamic University. He can be contacted at email: aambia@cse.iu.ac.bd.

Ferdib-Al-Islam completed B.Sc. Engineering in Computer Science and Engineering from Gopalganj Science and Technology University, Bangladesh and M.Sc. Engineering in the same subject from Khulna University of Engineering and Technology. He is currently working as an assistant professor in the Department of Computer Science and Engineering at Northern University of Business and Technology Khulna, Bangladesh. His research interests include ML, deep learning, explainable artificial intelligence, and computer vision. He acted as a reviewer and a member of technical program committee in reputed journals and conferences. He can be contacted at email: ferdib.bsmrstu@gmail.com.

Bulbul Ahamed is serving as a professor in the Department of Computer Science and Engineering at Sonargaon University, Bangladesh. His research interests encompass speech recognition, ML, artificial intelligence, and cybersecurity. He obtained his bachelor's degree in Computer Science and Engineering from Northern University Bangladesh. Additionally, he earned his M.Sc. degree in Computer Science and Engineering from United International University, Bangladesh. He received several prestigious awards, including the chancellor's gold medal and summa cum laude, in recognition of his outstanding academic achievements. He can be contacted at email: bulbul_cse@su.edu.bd.

Anindya Nag obtained B.Tech. degree from Adamas University in Kolkata, India, and an M.Sc. degree in Computer Science and Engineering from Khulna University in Bangladesh. His research focuses on wireless resource management in 6G and future generations, healthcare informatics, ML, deep learning, artificial intelligence, blockchain, cloud computing, and networking systems. He has written and collaborated on more than 30 articles, which include peer-reviewed journals and conference papers published by IEEE/ACM. He can be contacted at email: anindyanag@ieee.org.

Ariful Islam is working as an assistant professor in the Department of Biomedical Engineering at Islamic University, Bangladesh. His research interests include biosensors, biomedical instrumentation and diagnostics, medical device fabrication, nuclear medicine, bio-signal processing, bio-system modeling and simulation, biomechanics, biomaterials, and bioinformatics. He can be contacted at email: arif@bme.iu.ac.bd.