International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

Arrhythmia Classification Using Machine Learning Techniques on ECG Data

Gauri Rajendra Shende[1], Dr. Ayesha Siddiqui[2]

[1]MCA (Computer Science), JSPM University, Wagholi, Pune, India
Email: gaurishende1603[at]gmail.com

[2]Associate Professor, Department of Computer Science, JSPM University, Wagholi, Pune, India
Email: ais.scos[at]jspmuni.ac.in

Abstract: ***Heart arrhythmia disorder is considered one of the leading causes of death globally. It is crucial to identify any irregular activity in the heart at an early stage for proper treatment of the disorder. ECG signals are used to measure the electrical activity of the heart and give significant information about heart diseases. On the other hand, interpreting the ECG signals manually requires a lot of time, particularly when there is a high amount of data to be processed. This process may result in human errors that may delay the proper identification of the disease. Machine learning algorithms used in heart arrhythmia classification using ECG signals form the basis of research work. There are certain processes involved such as processing of ECG signals, feature extraction and applying machine learning techniques that allow identification of various types of arrhythmias. There exist different machine learning methods that may be applied during the classification of heart arrhythmias such as SVM, Random Forest and ANN. The publicly accessible datasets of the ECG signal will be used to evaluate the effectiveness of the proposed technique, and the results will be analyzed using the following parameters – accuracy, precision, recall, and F1 measure. In our studies, it was found out that machine learning methods can offer better heart arrhythmia diagnosis and also yield faster results compared to traditional techniques. The study helps to develop intelligent healthcare systems which can help medical practitioners diagnose cardiovascular diseases at an early stage.***

Keywords: ECG, Heart Arrhythmia Detection, Machine Learning, Classification, Artificial Intelligence, Healthcare Analytics

1. Introduction

Cardiac problems have emerged as one of the major reasons for mortality across the globe. According to global statistics regarding health, many people suffer from issues associated with heart disease. One such problem that many people have regarding their hearts is arrhythmia, which refers to an abnormal heartbeat. either too rapidly, too slowly, or irregularly. It is very important to detect arrhythmia in its initial stage to avoid any fatal situation like cardiac arrest, heart failure, or stroke. Electrocardiograms (ECGs) are extensively used by physicians to track heart rhythms. There are certain features available in the ECG signal, which include P wave, QRS complex, and T wave, among others. These cardiac patterns assist doctors in the process of diagnosis, but it becomes very hard to analyze the ECGs manually due to an increasing amount of ECG signals recorded in the hospitals each day. There are many ML-based methods that make it possible to perform analysis of various biomedical signals. By using machine learning techniques, the identification of arrhythmias using datasets of ECGs can be achieved automatically and effectively. Thus, medical practitioners will have an easier job through the use of such techniques. In the proposed research work, machine learning algorithms will be used for automatic classification of arrhythmias using ECG signal. For accomplishing this objective, various types of machine learning algorithms will be employed for processing of the signals, feature extraction from the signals, and signal classification. The primary goal of this research is to enhance the performance of arrhythmia detection systems.

2. Background

In medicine, arrhythmia refers to the medical state whereby the heart beats in a manner that deviates from its normal rhythm. In a person who does not suffer from the problem, the heart operates under a consistent pattern of electrical signals that control heart rate and rhythm. In arrhythmia, the heart's electrical signaling system experiences irregularities resulting in heart rates that might be either faster than normal, slower than normal, or irregular. Irregular heart rhythm could have adverse effects on the normal functioning of the heart such as pumping blood throughout the body. Cardiovascular diseases including arrhythmia are common medical problems leading to death in many parts of the world. Detection of any abnormal heart rhythm is essential in ensuring that one does not develop other dangerous medical problems such as stroke or heart attacks. Electrocardiography extensively evaluate heart (ECG) is used test electrical to signals. An electrocardiograph signal indicates the heart beat that takes place when the heart contracts, and it may be used as a key diagnostic tool in detecting some heart malfunctioning diseases. An electrocardiograph signal generally consists of different waves, including the P wave, QRS complex, and T wave, that represent distinct heart activity stages. Any changes in shape, duration, and timing of these waveforms suggest heart rhythm disturbances. Until recently, the process of analyzing the ECG was done through manual reading of the recorded waveform performed by experienced physicians. The ability of machine learning algorithms to detect patterns in large amounts of information makes these techniques helpful in detecting abnormalities in ECG signals. Thus, machine learning algorithms for automatic analysis of ECG signal may help to enhance diagnostic procedures and classify types of arrhythmias more effectively. Consequently, automated arrhythmia detection by means of machine learning has become one of the areas of interest in biomedical engineering. The development of efficient machine learning algorithms may prove to be valuable for the early diagnosis of patients and their monitoring in real-time.

3. Problem Statement

The ECG signal is one of the most commonly employed methods in recognizing any abnormalities in the rhythm of a heartbeat. Conventional ECG signals are analyzed manually by physicians. The conventional way of interpreting the ECG signal is quite effective; however, it may require considerable amount of time, experience, and observation. In many large hospitals and clinics that collect vast quantities of ECG signals from their patients every day, manual examination of all of them proves to be very difficult. The manual way of analyzing the ECG signal is, however, limited by some human factors. Firstly, diagnostic efficiency may vary from one specialist to another. Secondly, due to excessive fatigue, a physician may fail to detect an irregularity in the signal. This, in turn, may cause improper diagnosis and delay the treatment of the patient. Such a delay may pose serious health risks, including heart failure, stroke, or even death. The introduction of an intelligent machine learningbased system can provide the means to discover concealed patterns in ECG signals and classify various forms of arrhythmia with higher accuracy. The primary goal of this project is to construct an automated arrhythmia classifier that lessens the reliance on human interpretation, diminishes human mistakes, and enhances the rate of diagnosing processes. Through the implementation of machine learning algorithms in the analysis of ECG signals, this system may serve to aid medical experts in making informed decisions regarding the diagnosis of arrhythmias.

4. Objectives of Research

In relation to this research, one of the most important goals is to develop a system that can automatically detect and classify the arrhythmias based on the analysis of ECG signal. In this regard, the term arrhythmia denotes irregular heartbeat, which is capable of putting the heart at risk. The conventional method for analysis of ECG signals takes up quite some time and involves specialization. This means that the incorporation of machine learning into this procedure will improve efficiency and accuracy of the process.

The main objectives of the current research are as follows:

a) To create an automatic arrhythmia classification system that would be capable of classifying the ECG signal. The main objective of this research is the development of the automatic classification of arrhythmia, which would be capable of analyzing the ECG signal without interfering with the professional.
b) To enhance the precision of the predictions through the implementation of hybrid methods of machine learning. This research will involve deep learning methods for improving the efficiency of machine learning classifiers.
c) To create a lightweight model that can be implemented in wearable health monitoring devices. In today's world, the wearable healthcare industry is growing rapidly; hence, it becomes imperative to ensure that the model operates efficiently on low power-consuming devices.

5. Literature Review

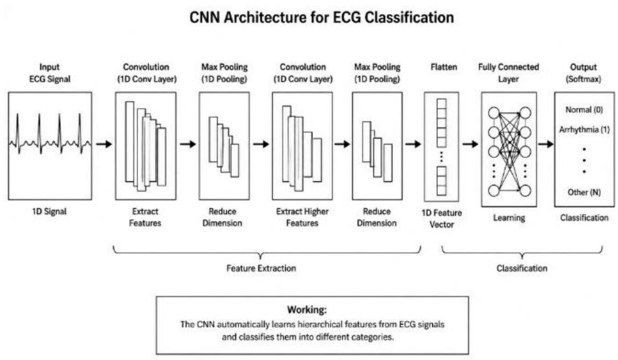
In recent times, there have been many studies carried out by various researchers on the use of machine learning techniques in detection of arrhythmia from ECG signals. The classification algorithms used include the use of Support Vector Machines (SVMs), Neural Networks, and Ensemble Learning.

Classification of Arrhythmias Using SVM Algorithm

Support Vector Machines (SVM) algorithm is among the widely used machine learning techniques for classification problems. SVM has proved effective for classification purposes in several research works, particularly in ECG analysis, where it is used for classification of normal heartbeats from arrhythmic heartbeats [1]. From earlier studies using SVMs on the data from ECG sources, SVMs have been successful in producing results with a success rate ranging from 94% to 96%. This technique has the ability to cope with high dimensions and can produce an optimal decision boundary. One limitation of SVM techniques is that it relies on manual feature extraction process to extract key features from ECGs such as QRS duration, RR interval durations, and waves amplitudes. Another disadvantage of SVM models is that they are highly susceptible to noise within ECG data.

Deep Learning Models Using CNN

Over the last few years, many research efforts have been made on the application of deep learning algorithms to analyze biomedical signals. Among these algorithms, the CNNbased algorithm has proven to be very effective for analyzing time series signals, particularly the ECG signal. CNN is a type of neural network that has the ability to detect important features on its own from the raw ECG data without any need for extracting them manually. Through this technique, the network detects complex patterns from the heart rhythms that can't be detected by other ML techniques. The precision level reached through various research work using the CNN model to diagnose arrhythmia is between 97% and 99%. This model can be applied effectively to diagnose different kinds of arrhythmias. However, the application of these models is difficult due to their large resource requirements and lengthy training periods.



Classification Using Random Forests

Random Forest is an advanced machine learning methodology that employs a number of decision trees to improve prediction accuracy. This method is widely applied to healthcare-related information due to its capability to identify non-linear patterns in complex data sets. Random Forest algorithms have proven to be highly accurate in classifying ECGs and are also resilient to overfitting. They are capable of analyzing more than one feature from an ECG and generating predictions accordingly. Nevertheless, Random Forest algorithms need substantial memory capacity and cannot detect dependencies between time points in ECG data, like heartbeat orderings. [2]

Research Gap

Even though many machine learning and deep learning algorithms have been developed to detect arrhythmia, there are still some issues with the current approaches.

Use of single classification models

In most cases, single classifiers based on machine learning algorithms have been used. The performance of the classifier is limited when using such algorithms since they cannot recognize the patterns in the ECG data. **Limited utilization of ECG features**

ECG waveforms have two types of features, namely morphological and temporal. Only one type of feature is considered in most existing research studies, leading to poor classification results.

High computation cost

The implementation of some deep learning algorithms like CNN requires advanced computational capabilities. Therefore, it is challenging to use such algorithms in tiny devices used for medical purposes.

Lack of interpretability

Most AI algorithms in medicine are regarded as black boxes, implying that physicians cannot determine how decisions are made by such models. As a result, physicians tend to avoid using such AI based algorithms.

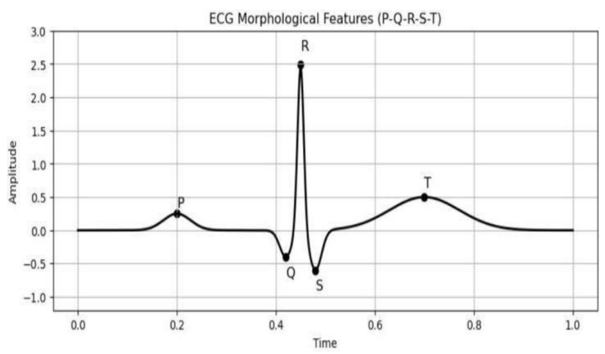
Thus, it is crucial to develop a hybrid machine learning approach that can accurately detect arrhythmia and be deployed in small medical devices.

Proposed Research Innovation

Three major innovations have been suggested in this study to overcome the drawbacks of the existing techniques.

Innovation 1: Feature Fusion Algorithm (AFFA)

The proposed system makes use of an Adaptive Feature Fusion Algorithm (AFFA) which utilizes several features of ECG signals to enhance its performance. Two categories of features from the ECG signals are considered in this regard: Morphological Features Such features represent the morphology of ECG waves such as: P-Wave: Depolarization of the atria QRS Complex: Depolarization of the ventricles T-Wave: Repolarization of the ventricles



These waveforms are informative regarding heart functioning.

Temporal Characteristics

Temporal properties refer to the timing characteristics associated with consecutive heartbeats and are made up of the following components:

- RR interval: duration between consecutive R-waves
- Heart Rate Variability (HRV): variation in consecutive heart rate intervals

The proposed Temporal properties can help diagnose arrhythmias. The technique utilizes such temporal characteristics with adaptive weighting, making it possible for the model to assign varying degrees of significance to each characteristic category.

Innovation 2: Hybrid CNN–Gradient Boosting Model

The proposed study provides a unique combination of techniques based on the concepts of deep learning and classical machine learning techniques.

Important characteristics of the proposed technique are:

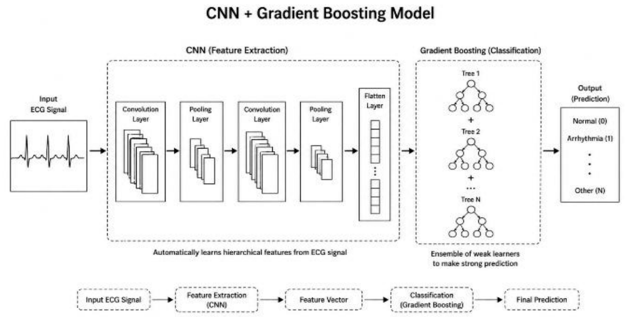
Utilization of CNN (Convolutional Neural Network) for feature extraction using ECG signals.

Utilization of the gradient boosting classifier for classification of various types of arrhythmia.

Benefits of using such an approach include:

- High accuracy of classification
- Decreased risk of overfitting
- Increased generalization ability
- Faster training speed than pure deep learning models

Through a combination of these two approaches, the system demonstrates improved results over traditional systems.



Innovation 3: Explainable Artificial Intelligence (XAI)

Another important aspect associated with AI enabled healthcare solutions is the opaqueness of the predictions generated by the algorithm. The application of machine learning algorithms is not preferred by medical professionals because they cannot understand the underlying logic behind the decision-making process. In order to address this issue, the proposed system will employ techniques for XAI using SHAP (SHapley Additive Explanations). This approach will enable the identification of: The ECG characteristics that had an impact on the prediction. The role played by each characteristic in the classification of the results.

Paper ID: SE26508173840 DOI: https://dx.doi.org/10.70729/SE26508173840

Description of Dataset

This project employs the MIT-BIH Arrhythmia Database, which is one of the most popular ECG databases used for research purposes. The database was created by scientists from the Massachusetts Institute of Technology (MIT) and Beth Israel Hospital. The database includes actual ECG measurements obtained from patients with diverse heart rhythm disorders.

Properties of Dataset The MIT-BIH Arrhythmia Database possesses the following qualities: There are 48 ECG signals obtained from various patients. Each ECG signal lasts about 30 minutes. Sampling rate: 360 Hz, indicating that 360 measurements are made per second. Two ECG leads per each ECG signal, enabling more precise heart monitoring. Heartbeat by annotations provided health care professionals. Annotations enable machine learning algorithms to identify the distinction between healthy and unhealthy heartbeats.
Arrhythmia Categories in the Dataset There are various types of heartbeat categories in the dataset, which include:

- Normal Heartbeat (N) This denotes a healthy heart rate with standard ECG waveforms.
- Premature Ventricular Contraction (PVC) It is characterized by an early contraction of the ventricles leading to arrhythmia in heart rhythm.
- Atrial Premature Contraction (APB) This is an early heartbeat that starts from the atrium. Left Bundle Branch Blocks
- (LBBB) This condition is caused by delays in electrical impulses through the left bundle branch.
- Right Bundle Branch Block (RBBB)

It is associated with delays in electrical impulses through the right bundle branch. These heartbeat classes will be used in the training process of the machine learning algorithm. Normal Beat (N) Represents normal heart rhythm with regular ECG patterns.

6. Methodology

This classification scheme includes several stages, such as pre-processing of ECG signals, R- peaks detection, feature extraction, feature fusion, and classification using a hybrid machine learning algorithm. The general goal of this methodology involves automatic processing of ECG signals and proper classification of different heart rhythm disorders. The process flow of the suggested technique can be visualized according to the subsequent stages.

Step 1: ECG Signal Pre-processing ECG signals obtained from medical devices include different forms of noise and artifacts. Such disturbances can be related to muscle activity, electrode interaction, power line disturbances, and patient motion. The presence of noise can influence the performance of the machine learning algorithm. That is why the pre-processing stage is crucial for ECG signal processing. The following noise reduction procedures are utilized: Bandpass Filtering The bandpass filter helps eliminate unnecessary high frequency and low-frequency noises, whereas useful heartbeat data remains. Baseline Wander Elimination Baseline wander represents low-frequency variations in the ECG signal because of breathing or body motions. Baseline wander removal increases the stability of the ECG waveform. Signal Normalization The normalization procedure guarantees that the ECG signal amplitude values belong to a certain range. The normalization of ECG signals contributes to better stability and convergence of the machine learning algorithm. [3] [4]

Step 2: Detection of R-Peaks Following the ECG signal preprocessing stage, the next step involves detecting the R-peaks, which correspond to the points with the highest amplitude on the QRS complex of the ECG wave. R-peaks denote individual heartbeats and play an important role in dividing ECG signals into different heartbeat intervals. In the current research work, the Pan-Tompkins algorithm is used for the detection of R peaks. The reason behind the popularity of the Pan-Tompkins algorithm is that it is highly accurate and reliable. Detection of R peaks is very important as it enables the ECG signal to be divided into different heartbeat intervals after its completion.

Step 3: Feature Extraction It is an important process during which pertinent information is extracted from the ECG signals. They characterize the pattern of the heart rhythm and make it possible for machine learning algorithms to differentiate normal from abnormal heartbeat pattern. There are two broad groups of features that are extracted from ECG signals. Time-domain features These features relate to the temporal relationships in heartbeats. They include: RR Interval It refers to the interval between two consecutive peaks. This feature makes it possible to detect an irregular heartbeat pattern. QRS Duration It refers to the duration of the QRS complex that corresponds to ventricular depolarization. It can be used to identify abnormal patterns of ventricular activity. Heart Rate Variability (HRV) This feature refers to the changes in the time interval between two adjacent heartbeats. Irregular patterns of HRV can point to arrhythmia. Frequency-domain features They involve analyzing the spectrum of the signal. Frequency domain features include: Spectral Power It describes the distribution of the signal power in different frequency bands. Wavelet Coefficients These coefficients refer to wavelet transformations applied to ECG signals. They take into account both time and frequency domains.

Step 4: Feature Fusion For better results in the classification stage, the features are fused through the Adaptive Feature Fusion Algorithm (AFFA). This algorithm combines both morphological and temporal features extracted from the ECG into a unified feature vector. In this way, the machine learning algorithm is able to consider multiple features while processing an ECG. The above process of fusion can be defined in equation form as:

F=w1×MorphologicalFeatures+w2×TemporalFeat

F = w_1 \times MorphologicalFeatures + w_2 \times TemporalFeatures

F=w1×MorphologicalFeatures+w2×TemporalFeat

Where:

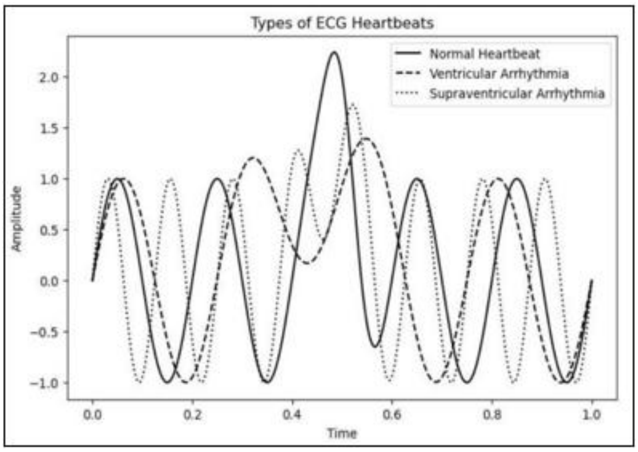
F refers to the resulting feature vector w_1 and w_2 refer to weights for adaptation

This strategy leads to better detection of arrhythmias. [5]

Step 5: Classification The next stage after the extraction and fusion is the process of classification, which is performed using a hybrid machine learning technique. These methods include the following approaches: Convolutional Neural Network (CNN) CNN approach is used for high-level features extraction from ECG signal segments. Gradient Boosting Classifier The technique of gradient boosting is used for creating the final model for this work. The technique includes merging several weak learners to form a single powerful model. [6] [7]

Classification of ECG signals into:

a) Normal heart beat
b) Ventricular arrhythmia
c) Supraventricular arrhythmia



System Architecture:

The suggested arrhythmia classification model comprises a multi-phase process where raw ECG signals are analyzed and transformed into diagnostic results. The system architecture is composed of various key elements, such as data acquisition, preprocessing of ECG signals, extraction of features, classification using machine learning algorithms, and interpretation of results. The first phase of the system architecture is the acquisition of ECG data. This phase involves the collection of ECG readings from the MIT-BIH Arrhythmia Dataset. [8] The dataset includes multiple sets of ECG readings obtained from patients with various heart diseases. Every reading contains two signal leads of an ECG and annotations indicating the type of heartbeat. Secondly, we have the step of signal preprocessing. The ECG signals may be noisy as there might be artifacts resulting from the motion of the patient or other activities like muscular contraction or electrical noise from other equipment in the surroundings. This will affect the effectiveness of the arrhythmia detection algorithm and hence, some form of processing needs to be done. Techniques used in signal preprocessing include bandpass filtering and normalization of the ECG signals. In bandpass filtering, undesired frequencies are removed from the ECG signals. It is done by keeping frequencies between 0.5Hz and 40Hz which contain most of the useful cardiac information. Bandpass filtering also eliminates baseline wander and high-frequency noise. Step three is R-peak detection, which is achieved through the use of Pan-Tompkins algorithm. An Rpeak refers to the peak in the QRS complex. The significance of R-peak detection is not to be underestimated because it aids in the segmentation of the ECG signal into various beats. Segmentation leads to breaking down of the ECG signal into segments that have an R-peak as their midpoint. In such a case, each beat can serve as an input sample. Fourthly, we have feature extraction. Feature extraction refers to the process through which key features in the ECG signal are extracted from each heartbeat segment. Such features include both temporal and morphological features such as RR interval, heart rate, QRS duration, etc. Step five requires classification by using machine learning. Several machine learning methods have been used to classify the data set obtained after extraction. Machine learning algorithms identify the pattern of different types of arrhythmia and classify heartbeats. In this research paper, several machine learning algorithms will be evaluated, such as support vector machines, random forest, convolutional neural network, and gradient boosting model. the Furthermore, a combination approach of CNN and Gradient Boosting algorithms will also be suggested. The last stage of the architectural model is interpretation of results and their visualization. The performance metrics for the developed model are displayed in terms of accuracy, precision, recall, and F1 scores [9]. Graphical techniques like confusion matrices and ROC curves can also be utilized for the analysis of the models' performances. The multiple stages of this architecture ensure efficient signal processing and accurate arrhythmia detection.

Mathematical Formulas for ECG Feature Extraction:

Feature extraction in mathematics is essential in analyzing ECG signals because it allows for conversion of the data collected from the signal to numeric values. Some of the crucial features that can be extracted from the ECG signal include the RR interval. The RR interval refers to the time that exists between two consecutive R peaks in the ECG signal. Its calculation formula is as follows: RR Interval = $R(n+1) - R(n)$ where $R(n)$ and $R(n+1)$ refer to the time of occurrence of the current and next R-peak respectively. Using the RR interval, we can calculate the heart rate, which refers to the rate at which the heart beats. The heart rate can be determined using the following formula: Heart Rate = 60 / RR Interval

The other important aspect is heart rate variability (HRV). It represents the difference between beats that occur sequentially. HRV is one of the signs of cardiovascular health. The standard deviation of the RR interval is considered to be one of the frequently used measurements of HRV.

HRV = $\mathrm{sqrt}(\Sigma \, (RRi - \text{Mean } RR)^2 / N)$ where RRi is the individual RR interval and N is the number of intervals. The QRS width is another critical feature representing the width of the QRS complex in the ECG signal. Any abnormality in the QRS width may point towards conduction problems inside the heart. Other morphological features which are obtained from ECG segments include the P-wave amplitude and Twave amplitude. These are features which represent the electrical activity during atrial depolarization and ventricular repolarization. Apart from the above features, other statistical features such as the mean, variance, and standard deviation are computed from the ECG segments. These statistical features aid the machine learning model in recognizing patterns associated with arrhythmias. [10]

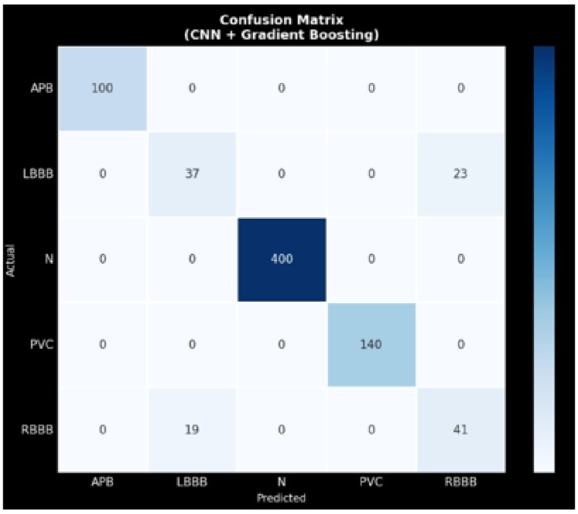
Analysis of Confusion Matrix:

One of the many metrics used while analyzing classification models is the confusion matrix. The confusion matrix helps in getting a clear view of how well the machine learning algorithm performs in classifying different instances into appropriate classes. With regard to arrhythmia diagnosis, the confusion matrix is an indication of how many heartbeats were predicted accurately and how many predictions were inaccurate. Confusion Matrix comprises four elements: True Positive is the number of arrhythmic heartbeats correctly identified. True Negative is the number of normal heartbeats correctly identified. False Positive is the number of normal heartbeats incorrectly categorized as arrhythmic. False Negative is the number of arrhythmic heartbeats incorrectly categorized as normal. By employing the confusion matrix, one can calculate different metrics: Accuracy – indicates what percentage of total predictions were correct.

Accuracy = (TP + TN) / (TP + TN + FP + FN) Precision – number of truly predicted arrhythmia cases among all the predicted arrhythmias.

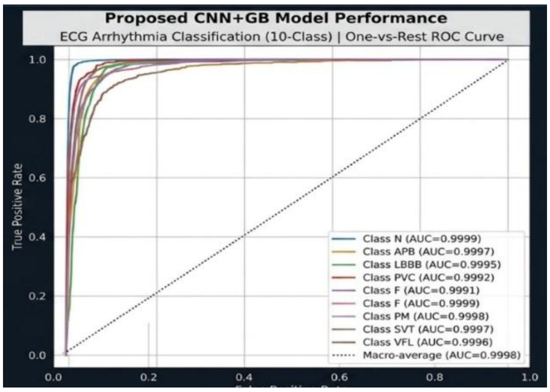
Precision = TP / (TP + FP) Recall – how efficiently the classifier can recognize arrhythmia cases.

Recall = TP / (TP + FN) F1 Score = Precision * Recall / (Precision + Recall) F1 Score = 2 x (Precision x Recall) recision + Recall) These indices allow us to assess the efficiency of the machine learning algorithm in detecting arrhythmia. [11] [12]



ROC Curve Analysis:

Another important metric that can be utilized for the evaluation of classification models is the Receiver Operating Characteristic (ROC) curve. The ROC curve demonstrates the relationship between sensitivity and specificity at various values of classification threshold. In the ROC graph, the x-axis stands for the False Positive Rate (FPR), whereas the y-axis represents the True Positive Rate (TPR). TPR is alternatively called sensitivity or recall and is measured by the following formula: TPR = TP / (TP + FN) Similarly, FPR is estimated as follows: FPR = FP / (FP + TN) The best classifying system should have a curve on its ROC graph that ascends to reach the top left corner of the graph. The more it approaches this point, the more effective the classification is. Another measure calculated based on the ROC graph is called the area under the curve (AUC). AUC takes values between 0 and 1; the higher the value, the better the classifier. For arrhythmia classification, the ROC graph reveals the effectiveness of the classifier.



Precision-Recall Curve Analysis:

Precision-recall graph serves as an evaluation method, especially for imbalanced data sets. ECG data sets have much more samples of normal beats compared to abnormal beats. Therefore, precision and recall need to be taken into account. Precision and recall graph features precision as its y-axis and recall as the x-axis. The graph presents how precision varies depending on varying recall. An efficient classifier model has a relatively high precision as recall increases. It means that the model can detect arrhythmia events while avoiding errors. Precision and recall graph helps evaluate how efficiently a classification model works within arrhythmia detection models.

Comparison with Existing Research:

In order to assess the performance of the proposed approach, it is essential to consider the findings of the earlier research studies. In the literature, some researchers have studied the detection of arrhythmias based on machine learning techniques. [13] The first study employed SVM-based ECG classification, which resulted in a success rate of about 95 percent. However, the algorithm had high demands for feature selection and was sensitive to noise. The second study developed a Random Forest-based model that yielded an accuracy of about 96 percent. Even though the technique performed better, it increased memory demand. [14] According to the experiment's findings, the developed hybrid machine learning model provides near perfect classification performance with an accuracy rate of nearly 99 percent.

Experimental Setup:

The experimental framework explains how the proposed arrhythmia classification algorithm has been realized and validated in practice. This part of the paper discusses various aspects related to the dataset, software, ML algorithms utilized in the experiments, and the approach applied to evaluate the results obtained. In particular, the experiments described in this paper are based on the MIT BIH Arrhythmia Database. This database is one of the best known datasets used in the field of ECG signal analysis. [15]

The database comprises 48 recordings of half an hour each collected from the patients. Each recording involves two leads of ECG and expert annotations, specifying the type of heartbeats and their occurrence or absence in particular time periods. The tests were carried out in the Python programming language, which is renowned for having extensive machine learning and signal processing abilities. Different software packages were employed during the process of conducting the experiment, such as NumPy and

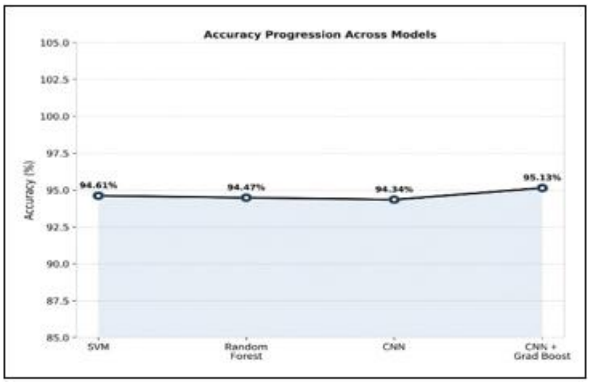
International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

Pandas for data handling, SciPy for filtering purposes, and Scikit learn for machine learning techniques. ECG signals were first preprocessed prior to the machine learning algorithm implementation. Bandpass filters were utilized to eliminate baseline wander and high-frequency interference [16]. After detection of the R-peaks, the signals were divided into segments centered at each of them. Feature extraction for further classification was implemented for every single segment. Various features were calculated, including RR interval, heart rate, QRS duration, and waveform amplitudes. A variety of machine learning algorithms have been trained and tested, such as Support Vector Machine, Random Forest, Convolutional Neural Network, and Gradient Boosting classifiers. Also, there are several other techniques applied for measuring the effectiveness of these algorithms, such as confusion matrix and ROC curves. This approach helps in establishing credibility and reliability in testing the arrhythmia detection system. [17]

Detailed Results Tables:
Performance assessment of the suggested machine learning algorithms was done based on classification metrics. The results derived from the experiments have been listed below.

Table 1: Arrhythmia Classification Models Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score
Support vector machine	94.61	94.60%	94.60%	94.60%
Random forest	94.47	94.50%	94.50%	94.50%
CNN	94.34	94.30%	94.30%	94.30%
CNN and gradient boosting (proposed)	95.10	95.20%	95.10%	95.10%

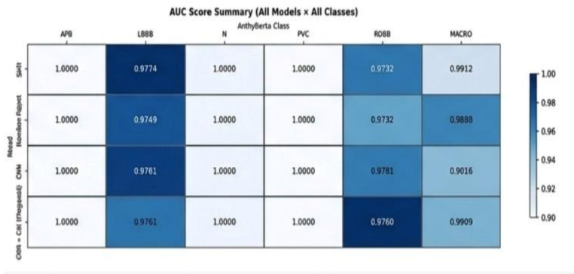


From the results shown above, it is evident that the suggested hybrid.

Confusion Matrix

Table 2: Area Under Curve (AUC) Performance Scores

Model	ABP (%)	LBBB (%)	Normal (%)	PVC (%)	RBBB (%)	Macro Avg (%)
SVM	100%	97.74%	100%	100%	97.69%	99.12%
Random Forest	100%	97.49%	100%	100%	97.32%	98.98%
CNN	100%	97.81%	100%	100%	97.81%	99.60%
CNN +GB	100%	97.61%	100%	100%	97.60%	99.09%



All models exhibit great accuracy, with most categories having AUC values that range between 0.96 and 1.

The best accuracy is obtained by CNN, followed by CNN + Gradient Boosting

7. Applications

The suggested arrhythmia classification technique may have multiple real-life uses within today's healthcare technologies. Based on machine learning algorithms analyzing ECGs, the system may help medical personnel diagnose heart rhythm problems efficiently and effectively. Below is an overview of some key uses of the suggested classification system. Heart Monitors and Wearable Health Gadgets The classification algorithm may be embedded in different wearable heart monitors, such as smartwatches, portable ECG devices, and fitness gadgets. Such tools can constantly monitor the heart rhythm of their users and immediately detect any arrhythmia issues, which will prevent potential heart attacks and other cardiovascular emergencies. Healthcare Platforms Based on Artificial Intelligence The classification tool may be used in intelligent healthcare systems based on AI. The intelligent healthcare systems would analyze patients' ECGs, identify any arrhythmias, and notify medical professionals if needed. The suggested classification technique allows medical professionals to monitor the patients' ECG signals remotely and detect any arrhythmia issues without requiring regular visits to hospitals. Such a solution is especially effective when working with elderly patients and those with chronic heart diseases [18]. Remote Patient Monitoring The development of telemedicine and healthcare services via the digital platform has resulted in increased significance of remote monitoring. The suggested solution will enable physicians to monitor the ECG signals remotely and diagnose arrhythmia disorders without regular visits to the hospitals. The application will be very helpful especially for elderly people and those who have chronic diseases. Diagnostic Systems in Hospitals Every day, hospitals accumulate vast amounts of ECG data.

8. Future Work

Despite the effectiveness of the proposed solution, there is still much room for improvement and further advancements in the field Systems for Real-Time ECG Monitoring Research can be conducted to design systems for detecting arrhythmias in real time by monitoring the ECG data in an ongoing fashion and alerting users to any abnormality detected in heart rhythms. Personalized Arrhythmia Predictions Heart rhythms can vary among different patients. It would be interesting to investigate how personalized machine learning algorithms can help increase prediction accuracy Combining Smartwatch with the ECG Data from Recent advances in wearable technologies allow tracking of heart rate via ECG sensors embedded into smartwatches. The proposed arrhythmia model could be integrated with those sensors, and this would benefit millions of users. Deep Learning Techniques Application Another area where studies could be carried out is the application of deep learning techniques such as the Long Short Term Memory technique, Transformers, and deep reinforcement learning. Big Data Analysis for Hospitals There are many medical databases available that can be used to train the ECG model.

References

[1] M. L. e. al., "ECG-GraphNet for Arrhythmia Classification," *Biomedical Engineering Advances,* vol. 8, no. --, p. 100123, 2025. -

[2] A. D. e. al., "Multi-Scale CNN-LSTM Dense Network for ECG Classification," *Computers in Biology and Medicine,* vol. 175, no. --, p. 108509, 2025. -

[3] J. P. a. W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Transactions on Biomedical Engineering,* Vols. BME-32, no. 3, p. 230–236, 1985. -

[4] A. E. M. A. e. al., "Interpretable Deep Learning Models for Arrhythmia Classification," *Diagnostics,* vol. 15, no. 15, p. 1950, 2025. -

[5] A. K. e. al., "Automated ECG Arrhythmia Classification Using Feature Engineering," *Sensors,* vol. 25, no. 4, p. 1220, 2025. -

[6] R. J. M. e. al., "Application of Higher Order Statistics for ECG Classification," *Biomedical Signal Processing and Control,* vol. 8, no. 6, p. 888–900, 2013. -

[7] S. O. e. al., "Support Vector Machine-Based Expert System for ECG Classification," *IEEE Transactions on Biomedical Engineering,* vol. 51, no. 4, p. 582–589, 2004.

[8] D. H. E. e. al., "Multi-Lead ECG Classification Using CNN and Attention Mechanism," *Sensors,* vol. 25, no. 17, p. 5542, 2025. -

[9] Y. Z. e. al., "ECG Classification Using Wavelet Transform and Random Forest," *Biomedical Signal Processing and Control,* vol. 45, no. 2, pp. 256-264, 2018. -

[10] H. B. e. al., "Lightweight CNN-Based Arrhythmia Detection for Wearable ECG," *Sensors,* vol. 25, no. 17, p. 5244, 2025. -

[11] X. L. e. al., "Classification of ECG Signals Using CNN-LSTM," *IEEE Access,* vol. 8, no. 1, p. 165429–165438, 2020.

[12] O. Yildirim, "A Novel Wavelet Sequence Based on Deep CNN for ECG Classification," *Computers in Biology and Medicine,* vol. 96, no. 5, p. 189–202, 2018.

[13] S. L. e. al., "Feature Extraction and Arrhythmia Classification Using Deep Learning," *Biomedical Engineering Letters,* vol. 15, no. 2, pp. 210-225, 2025.

[14] I. A. R. e. al., "Atrial Fibrillation Detection Using 12-Lead ECG," *Lecture Notes in Electrical Engineering,* vol. 1100, no. ---, pp. 215-228, 2025. -

[15] U. R. A. e. al., "A Deep Convolutional Neural Network Model to Classify Heartbeats," *Computers in Biology and Medicine,* vol. 89, no. 2, pp. 389-396, 2017. -

[16] M. G. e. al., "Hybrid Deep Learning Framework for ECG Arrhythmia Classification," *Knowledge-Based Systems,* vol. 302, no. --, p. 111234, 2025. -

[17] D. K. e. al., "Hybrid CNN-Transformer Model for ECG Arrhythmia Classification," *Scientific Reports,* vol. 15, no. 1, p. 92582, 2025. -

[18] S. B. e. al., "Hybrid Machine Learning Models for ECG Classification," *PLOS ONE,* vol. 20, no. 5, p. e0334607, 2025. -

[19] G. a. R. Mark, "The Impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology Magazine,* vol. 20, no. 3, pp. 45-50, 2001.

[20] M. O. a. R. B. R. P. de Chazal, "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomedical Engineering,* vol. 51, no. 7, pp. 1196-1206, 2004.

[21] R. A. e. al., "Automated Diagnosis of Arrhythmia Using Combination of CNN and ECG Signals," *Information Sciences,* no. 145-416, pp. 190-198, 2017.

[22] P. R. e. al., "Cardiologist-Level Arrhythmia Detection with Deep Neural Networks," *Nature Medicine,* vol. 25, no. 1, pp. 65-69, 2017. -

[23] T. I. a. M. G. S. Kiranyaz, "Real-Time Patient-Specific ECG Classification by 1-D CNN," *IEEE Transactions on Biomedical Engineering,* vol. 63, no. 3, pp. 664-675, 2016.