Comparative Analysis of Supervised Machine Learning Algorithms for ECG Arrhythmia Detection using Small Dataset

Igiri , C.G.[1], Anireh, V.I.E.[2], Nwiabu, N.D.[3], Matthias, Daniel[4]
[1,2,3,4]Department of Computer Science, Rivers State University, Nkpolu, Rivers State, Nigeria.
Corresponding Author: igiri.chima@ust.edu.ng, Tel.: +2348063483474

DOI: 10.56201/ijcsmt.v9.no4.2023.pg23.44

Abstract

Automatic detection and analysis of arrhythmias from Electrocardiogram (ECG) signal is beginning to take a center stage in recent times, due to delay the nature of ECG signals, coupled with the subjective interpretation of these signals by cardiologist. Arrhythmia detection plays a vital role in diagnosing and managing cardiovascular diseases. With the advancements in machine learning techniques, various supervised algorithms have been employed to improve the accuracy of arrhythmia detection. However, each supervised ML algorithm has its strength and weakness in predicting the various classes of arrhythmia. This research study presents a comparative analysis of five popular supervised machine learning algorithms: support vector machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), Naïve Baye (NB) and Decision Tree (DT) when applied to ECG arrhythmia detection with down sampled dataset. The goal is to evaluate and compare the performance of these algorithms in terms of accuracy, precision, recall, and F1 score. The study utilizes the MIT-BIH benchmark dataset, and experimental results provide insights into the strengths and limitations of each algorithm, aiding in the selection of the most suitable algorithm for accurate ECG arrhythmia detection. The Random Forest algorithm outperformed other algorithms in terms of accuracy, achieving an accuracy of 89.9% with RR interval based feature set.

Keywords: *Electrocardiogram, Cardiovascular, Arrhythmia, Machine Learning*

1.0 Introduction

According to World Health Organisation (WHO), Cardiovascular Diseases (CVDs) are a leading cause of mortality worldwide. Also, WHO, (2018) projected an increase of 23.4 million deaths by 2030 globally, due to heart abnormality.Arrhythmia which refers to abnormal heart rhythms that can range from harmless irregularities to life-threatening conditions, is a major CVD that requires early detection (Murat *et al.*, 2020; Carrillo-Alarcón *et al.*, 2020; Pandey *et al.*, 2020; Singh & Singh, 2019). Early and accurate detection of arrhythmias is crucial for effective diagnosis, and mitigating the side effect (Murat *et al.*, 2020). For clinical diagnosis of heart diseases, the Electrocardiogram (ECG) signal is mostly used by cardiologist, since it gives a complete

representation of the electrical activity of the heart on the surface of the skin, and it can be used to monitor the functionality of the cardiovascular system (Qin *et al.,* 2017; Sraitith *et al.,* 2020). It is made up of cycles of depolarization and repolarization, where the normal beats have a rhythm called sinus, which includes the QRS complex, the P, Q, and T waves (Murat *et al.,* 2020). Thus, heart rhythm disorders or alterations in the ECG waveform are evidences of underlying cardiovascular problems, such as arrhythmias. Figure 1 shows waves and intervals in two circles of an ECG circle. In common practice the standard 12-lead electrocardiogram, which measures electric potentials from 10 electrodes placed at different parts of the body surface, six in the chest and four in the limbs are used for taking ECG readings (Canon- *et al.,* 2022). These readings are then annoted and interpreted by the cardiologist to predict any possible abnormality in the heart. However, human interpretation of ECG signals can be unreliable and varies between several cardiologists, depending on the experience of the cardiologist (Sraitith *et al.,* 2020).

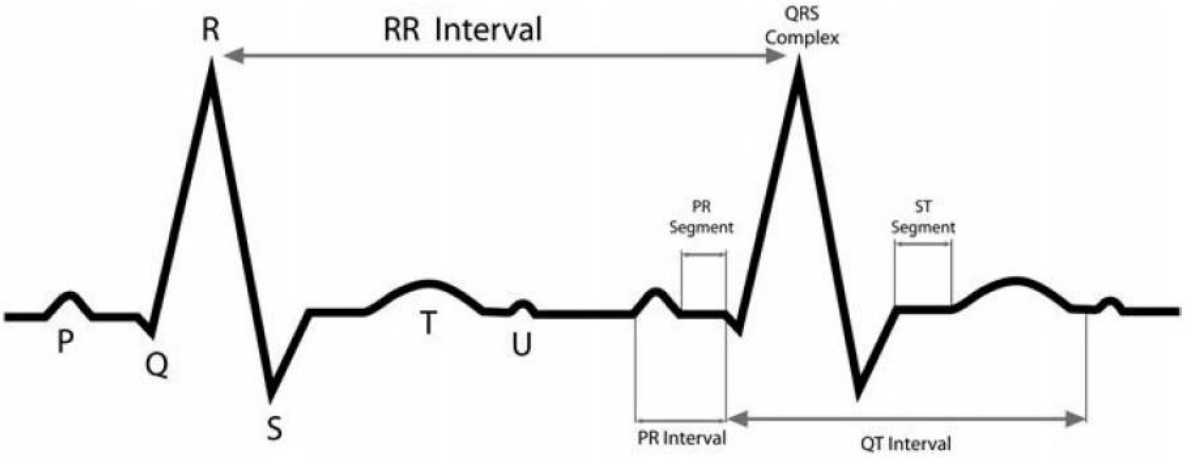

Figure 1. The ECG Waves, Segments, and Intervals (Source: Carrillo-Alarcon *et al.*, 2020).

Also, it is challenging to visually interpret the Electro Cardiogram (ECG) signals, due to its small amplitude and duration, and it can be time–consuming, less accurate, and prone to mistakes due to stress, exhaustion, fatigue. Hence, the need for real time automation of the process is necessary, since it is faster, more dependable, more efficient, more standardized, more accurate, and objective in nature (Vanajakshi *et. al.,* 2023; Taloba *et al.,* 2021). Also, in traditional methods monitoring is only done during visit of the patient to the hospital and for limited time, mainly 24 to 48 hours. However, early detection of certain types of transient, short-term or infrequent arrhythmias requires long-term monitoring (more than 24 h) of the electrical activity of the heart (Montenegro et al., 2022; Cañon-Clavijo *et al.,* 2023). The adaptation of IoT wearables, which are unobstructive to the users, have given rise to continuous remote heart monitoring, where the patient can live independently in areas of their choice and still be monitored remotely.

With the advent of machine learning techniques, there has been a growing interest in leveraging machine learning algorithms to automate ECG arrhythmia detection. These algorithms have the potential to improve the predictive accuracy for arrhythmia diagnosis systems. Supervised ML learning algorithms uses supervised learning, which is the types of machine learning in which

machines are trained using well "labelled" training data, and on basis of that data, machines predict the output ((Sraitith *et al.*, 2020)). The labelled data means some input data is already tagged with the correct output (Ahamed, *et al.*, 2021). The size of the dataset has always affected the ML algorithm performance. However, due to the difficulties in taking ECG readings, small amount of data was used in this study to demonstrate the performance of the ML algorithm with small available dataset. While several supervised machine learning algorithms have been applied to ECG arrhythmia detection (Cañon-Clavijo *et al.*, 2023) there is a need to comprehensively analyze and compare their performance. This study investigates the performance of five supervised machine learning models: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), Naive Bayes(NB) and Decision Tree (DT). The choice of algorithm can significantly impact the accuracy and reliability of the detection system. Therefore, it is important to evaluate and compare various algorithms to identify the most suitable approach for accurate ECG arrhythmia detection. The experiment is performed on the publicly available MIT-BIH dataset based on the AAAI recommendation, to assist healthcare professionals and researchers in selecting the most effective algorithm for automated arrhythmia detection. The rest of this paper is organized as follows: Section II introduces related literature reviews. In Section III, the materials and methods used are discussed. In section V, results are exposed, leading to the discussion in Section VI, as well as conclusion drawn.

2.0 Review of Related work

Arrhythmia detection is a critical task in healthcare, aiming to identify abnormal heart rhythms that can be indicative of various cardiac conditions. Several studies have been conducted to develop and evaluate different approaches for arrhythmia detection. Here is a review of some related work in this field: Han, (2017) developed a fuzzy neural networks model for classification of heart beat. The Genetic Algorithm (GA) was used for optimization of the input subspace, while the Gradient Decent (GD) method was used for training the model. The model achieved an accuracy of 82.5%.

Ramalingam *et al.*, (2018), provides an overview of numerous models that focus upon methods and methodologies for arrhythmia detection. Naive Bayes, Random Forest (RF), K-nearest Neighbor (KNN), Decision Trees (DT), and Support Vector machines (SVM), and Ensemble models are popular models, particularly among researchers.

A hybrid approach for automatic heartbeat classification into normal and abnormal using Elephant Herding Optimization (EHO) and Support Vector Machine (SVM) was proposed by (Hassanien *et al.*, 2018). The EHO was used to select the optimal feature set and parameter optimization, while the SVM is used for binary classification of heartbeat. The hybrid of EHO-SVM produced an accuracy of 93.31% on the MIT-BIH dataset. However, the use of manual feature extraction and selection, tends to affect the accuracy and only binary classification that was implemented. Also, conventional machine learning does not handle high dimensional-dataset.
Kumari *et al.*, (2021), used SVM classifier with Discrete Wavelet Transform (DWT) to classify arrhythmias, using the MIT-BIH and BIDMC databases. DWT was used to extract a total of 190 features, and an accuracy of 95.92% was achieved.
Singh & Singh (2019), developed a system for cardiac arrhythmia detection using linear SVM, random forest, and JRip ML algorithms. The highest accuracy of 85.58% was obtained with RF

classifier with gain ratio feature selection method. A subset of 30 features was used and this number tends to increase computational complexity of the model.

Qin *et al.*, (2017) develop a model for heartbeat classification using wavelet multi-resolution analysis to extract time-frequency domain feature sets, while PCA is used to select the most discriminant feature sets in order to reduce the dimension of the feature sets. Using the MIT-BIH arrhythmia datasets, an accuracy of 99.70% was achieved with the SVM and 10 fold cross validation. The use of handcrafted feature selection affects the system.

Kumar, *et al.,* (2020) used various machine learning techniques for the prediction of cardio vascular disease. The proposed model showed that random forests having highest accuracy achieved of 85.71% as compared to other classifier techniques.

Batra & Jawa, (2016), proposed arrhythmia detection model using Neural Networks(NN),DT ,RF, Gradient Boosting, and SVM.. The Combination of SVM and Gradient Boosting achieved the highest accuracy of 84.82% .

The study by Devadas, (2021) compared the performance of SVM, KNN, and Naive Bayes (NB) for multiclass classification of 16 classes of cardiac arrhythmia. Using Boruta Algorithm (BA) for feature selection and using the wrapper class around the RF, a total of 88 features generated yielded an accuracy of 71.4% with SVM. The use of handcrafted feature is a problem and leads to the problem of curse of dimensionality.

Marjia *et al.,* (2016) used several ML algorithm KStar, J48, Sequential Minimal Optimization (SMO) and Bayes Net and Multilayer perceptron for classification of heart diseases. For heart disease prediction the performances of Bayes Net and SMO classifiers are the optimum among the investigated five classifiers, with SMO giving highest accuracy of 89%.

Using augmentation to increase dataset and parameter turning to improve accuracy and training time, Shanmugavadivel *et al.,* (2022) proposed a CNN-hyper parameter tuned model for classification of arrhythmia using ECG image. Also, SVM, NB and Logistic Regression (LR) were used for the classification, and the tuned hyper parameter CNN model outperformed all other ML model with an accuracy of 94%. Keras turner was used for the hyper parameter tuning. However, the accuracy obtained needs improvement considering the sensitivity of the area of application.

In 2017, Sultan-Qurraie proposed a cardiac arrhythmias system using frequency domain feature sets. The use of frequency domain feature sets will add more computation to the system.

The advantages of ensemble was demonstrated by Pandey *et al.,(*2020), where ensemble of SVM was used for heartbeat classification of ECG signal into 4 classes. Wavelets, high order statistics, R-R intervals and morphological features were extracted and an overall accuracy of 94.4% was attained for patient specific classification. However, the use of ensemble increases computational intensity.

The wavelet packet entropy and RR intervals heartbeat features were used with RF algorithm to classifier heartbeat by (Li *et al.,* 2016). An accuracy of 94%

Sraitih *et al.,* (2021), proposed inter-patient classification of ECG for arrhythmia classification without feature extraction using SVM, KNN, RF and ensemble of the five class of arrhythmia. The SVM outperforms others with accuracy of 83% in terms of computational cost. Also, the use of

ensemble increases computational complexity, and the use of grid search for hyper parameter optimization may not guarantee optimal model parameters.
The need for developing smart real-time health monitoring system for cardiovascular disease is becoming trending, and Bazi *et al.,* (2020) implemented such system capable of mitigating the effect of cardiovascular abnormalities. Four temporal features extracted are pre-RR, post-RR, local RR interval and global average RR interval, while SVM was used as classifier. The result demonstrated the possibility of using the system for smart remote health monitoring.
In a multiclass classification, Alarsan & Younes, (2019) implemented an ECG classification model using RF and Gradient-Boosted Trees (GDB) implemented in ML-Libs and scala language on Apache spark framework. The ECG signal is considered as a big data and the need to process big data at the edge of the network was demonstrated. DWT is used for feature extraction, to extract summits feature such as temporal features, and morphological features, however the use of all feature set extracted increases computational intensity. The study used the RF classifier, and obtained an accuracy of 98.03% for multi class classification.

Ahamed *et al.*, (2021) propose a cardiovascular disease prediction system applying fve machine-learning algorithms (random forest, decision tree, Naive based, k-nearest neighbors, and support vector machine).

3.0 Materials and Methods

This section introduces the methods used in this research as well as the materials.

3.1 Proposed Method

The methods used in this research are illustrated in figure 2. The ECG signals are collected and annotated as described in the dataset, after which pre processing consisting of noise removal and normalization is performed. In our method, the data set are down sampled to ensure that total number of evaluable data set is small as well as to balance the various classes in the dataset. After which the dataset are segmented, then feature sets are generated using the RR intervals and passed to the classifiers for prediction. Based on the feature sets various classes of arrhythmia are then predicted.

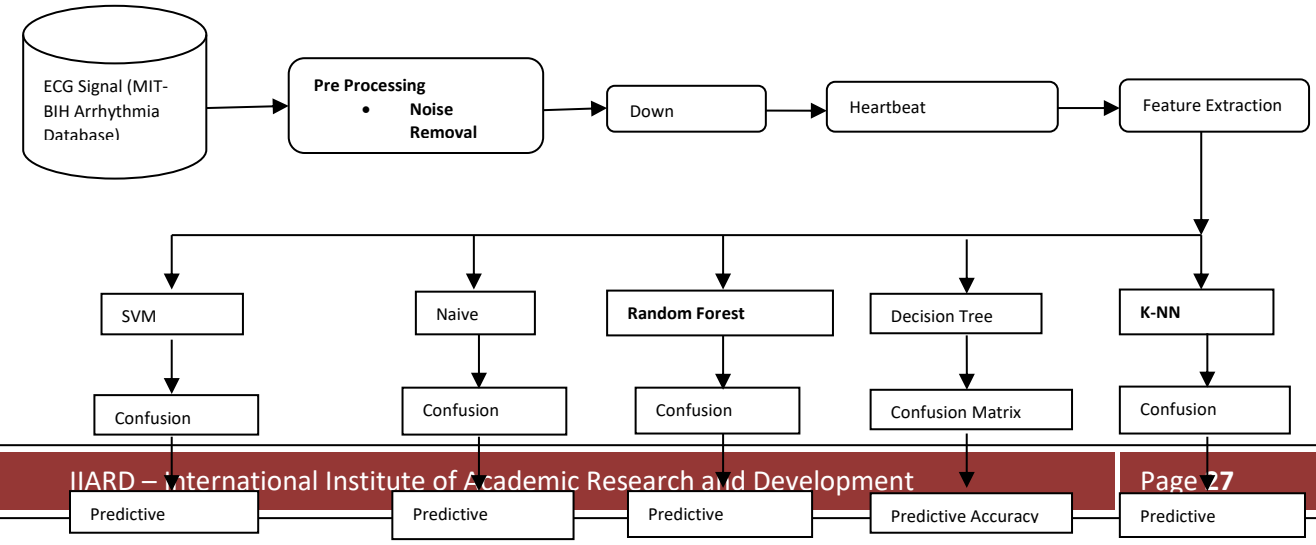

Figure 2. Proposed arrhythmia classification system

3.1.1 Dataset

The performance of the proposed heartbeat classification method has been evaluated with the MIT-BIH arrhythmia dataset (Moody *et al.*, 2001), which is a golden standard to evaluate ML classifiers on arrhythmia classification. It consists of 48 half-hour ECG records sampled at 360 Hz. The sampling rate determines the granularity and fidelity of the ECG waveform and it is important to consider this when designing the detection algorithms. Each ECG record contains two leads: lead II (modified limb lead II, obtained from electrodes on the chest) and lead V1' (modified lead V1, and in some records V2, V4, or V5). Based on recommendation by the Association for the Advancement of Medical Instrumentation (AAMI) for the evaluation of ECG classifiers the dataset consist of five (5) major heartbeat classes: N (normal beats), S (supraventricular ectopic beats), V (ventricular ectopic beats), F (fusion beats), and Q (unclassifiable beats). According to AAAI specification the Q beats and paced beats are removed. The dataset contains recorded ECG signals along with accompanying annotations or labels indicating the presence of various arrhythmias. The annotations are crucial for supervised learning approaches and performance evaluation. The dataset is then divided into 70% for training the model and 30% for testing of the model.

3.1.2 Preprocessing

Preprocessing is a crucial step in ECG arrhythmia detection as it helps to enhance the quality of the ECG signal and extract meaningful information. It involves several techniques and steps to remove noise, artifacts, baseline wander, and to normalize the signal. Commonly used denoising methods include: Filtering Techniques (which involves use of Low-Pass Filter, High-Pass Filter, and Bandpass Filter) , Wavelet Transform(which includes various wavelet families like Daubechies, Symlet, and Haar), and Ensemble Averaging. This study adopts the use of low pass and high pass filters . The High-pass filter removes the baseline wander and low-frequency noise, emphasizing the higher-frequency components of the ECG signal. The Bandpass Filter which combines the properties of low-pass and high-pass filters, to allow a specific frequency band relevant to the ECG signal of interest was used for noise removal. It effectively removes noise and artifacts outside the desired frequency range, enhancing the detection of arrhythmias.

3.1.3 Down Sampling

The study conducted an exploratory data analysis on the dataset in other to have insight into the dataset. Figure 3 shows a countplot of the dataset. The visualized countplot, shows that the dataset is highly imbalanced, as the number of instances of the classes tends to vary by far margin. Figure 4 shows a visualized plot of the balanced data. The data imbalance was resolved using RandomUnderSampling technique in python. This was used to reduce the instances of the majority classes being of the same size as the minority class.

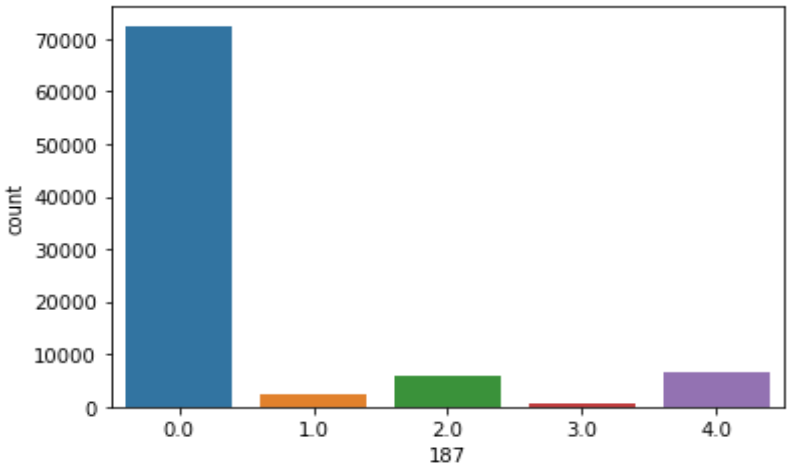

Figure 3: Countplot of the data imbalanced.

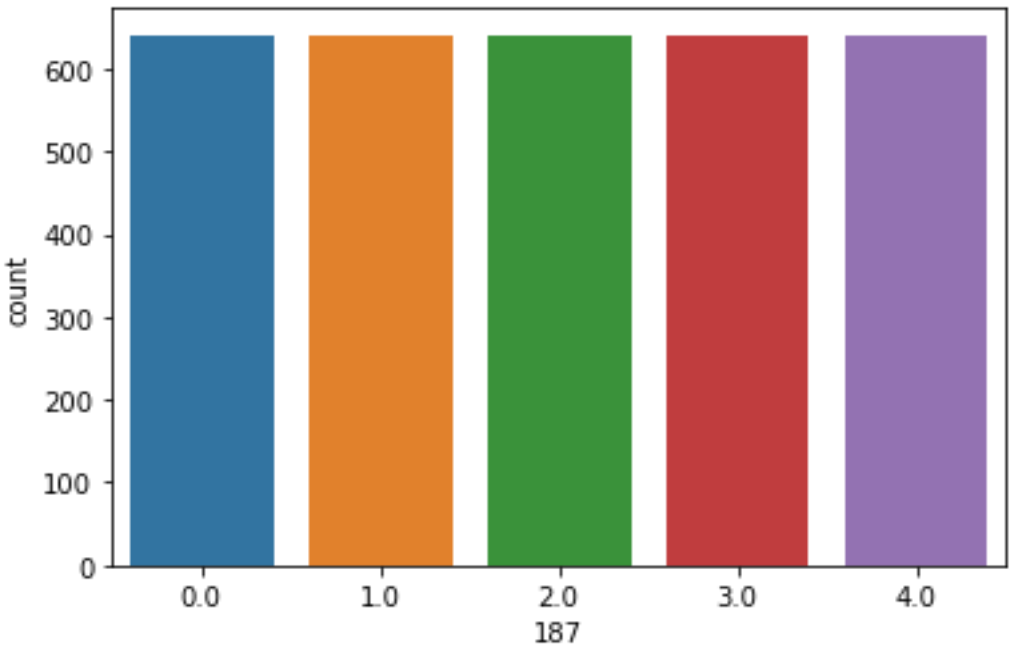

Figure 4: Countplot of the balanced data

3.1.4 Normalization

Normalization of the datasets was implemented using the Z-score statistic where the mean heart beat (μ) is subtracted from the heartbeat point $x(i)$, and then divided by the standard deviation(σ) of the waveform as shown in equation(1):

$$Z = \frac{(x(i) - \mu)}{\sigma} \quad (1)$$

3.1.5. Heartbeat segmentation

In the segmentation process the R-peak position is detected using the pan-tompkins algorithm within a window size of 180, centered at R-peak. Each heartbeat signal consists of 90 points on the right of R peaks and 90 points on the left hand side. The R-peak locations and RR-intervals were extracted using the BioSPPy (v.0.7.2)1 Python module (Asl et al., 2008).

3.1.6 Feature Engineering for the ECG Arrhythmia Detection

Feature engineering plays a crucial role in ECG arrhythmia detection as it involves extracting informative and discriminative features from the raw ECG signal. These features are then used as inputs for machine learning algorithms or other classification techniques. To train a ML classifier, one must first pre-process the ECG signals and then extract features from the signal employing known signal processing methods. Feature engineering which consist of feature extraction and selection ensures that most discriminant features are used, hence reducing computational complexities. Some popular choices for the input features are morphological features extracted from the time domain (such as inter-beat intervals, amplitudes, areas) frequency-domain features wavelet transforms and higher order statistics (HOS) (Montenegro et al., 2022). The RR-interval features that are computed from the time between consequent beats includes Pre-RR: the distance between the actual R-peak and the previous one, Post-RR: distance between the actual R-peak and the next one, Local-RR: average of the ten previous Pre-RR intervals, Global- RR: average of the Pre-RR values produced in the signal and the normalized values from each of the four intervals (Montenegro *et al.,* 2022)**.** Series of experiment were performed while selecting different combination of the features set, and finally a combination of Pre-RR and Local-RR produced the best accuracy.

3.2.7 Machine Learning Model

A machine learning model is an algorithm that learns patterns and relationships from data to make predictions or take actions without explicit programming. It is trained by optimizing its parameters based on a specified objective, such as minimizing error or maximizing accuracy. Once trained, the model can be used to make predictions on new, unseen data by applying the learned patterns and relationships. In this study, five supervised ML algorithms are explored, and their operations are presented in this section.

3.2.6.1 K-Nearest Neighbor

The K-Nearest Neighbors (KNN), also called as a non-parametric lazy algorithm because it does not use any model to fit, it is only based on memory, and it classifies feature vectors based on labels of the closest training samples in the feature space (Sraitih *et al*., 2021). The two parameters required by the algorithm are the value K and the threshold value, where. K value shows the number of nearby neighbors, and the threshold value is used for the evaluation of unusual neighbors. The k-nearest neighbors are collected by calculating the distance between an unknown feature vector and all the vectors in the training set using mainly the Euclidean distance given by

$$D(X,Y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{\frac{1}{p}} \quad (2)$$

where p is the order and $X = (x_1, x_2, ..., x_n)$, $Y = (y_1, y_2, ...y_n) \in R^n$.
The unknown feature vector is assigned then to the class to which the closest k samples mostly belong with the help of the votes got from the neighbors (Sraitih *et al*., 2021). The class with the most votes is considered as the prediction. As shown in figure 5, two feature vectors are represented with sky color rectangle and black color circle, while the new feature point is represented as star]. Using the Euclidean distance, the new data point is represented by the green star point. For K = 7, the star is close to the rectangle, hence the KNN algorithm classified it as a class A (Ahamed *et al.*, 2022). It then estimates the conditional probability for each class, that is, the fraction of points in A with that given class label. (Note I(x) is the indicator function which evaluates to 1 when the argument x is true and 0 otherwise).

$$P(y = j|X = x) = \frac{1}{K} \sum_{i\varepsilon A} I(y^{(i)} = j)$$
(4)

Finally, the users input x gets assigned to the class with the largest probability.

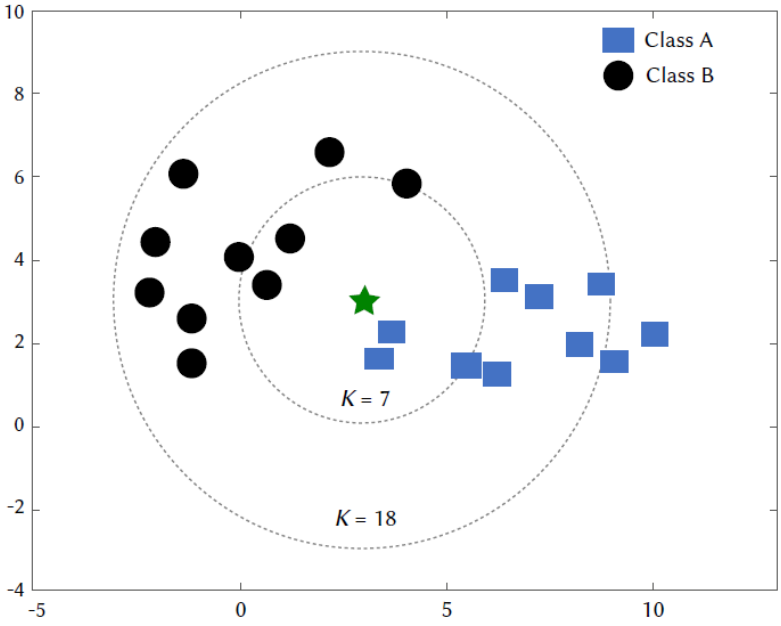

Figure 5: K-Nearest Neighbors (Source: Ahamed *et al.*, 2022).

However, it requires careful selection of K and appropriate feature representation for accurate arrhythmia classification. K-NN offers advantages such as simplicity, interpretability, and ability to handle nonlinear relationships in the data.

3.3.2 Random Forest

Random forest (RF) is another famous supervised machine learning technique, which is primarily an ensemble of decision trees to train and predict outcomes. Random forest constructs multiple decision trees during the training phase, and takes a final decision to choose the trees based on the

majority of voting (Ahamed *et al.*, 2022). The use of multiple trees helps reduce model overfitting. Figure 6 portrays the working of the Random Forest algorithm. Three different Decision trees were constructed from the dataset, but the final classification algorithm takes decisions based upon the majority of voters.

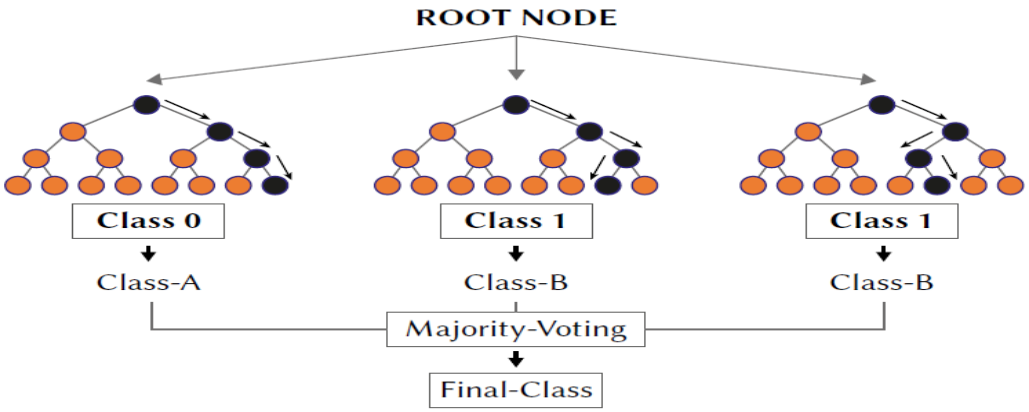

Figure. 2. Random Forest Algorithm (Source: Ahamed *et al.*, 2022).

The RF is a parametric algorithm regarding the number of trees in the forest and is also a stochastic method because of its two origins of randomness: random attribute sub-set selection and bootstrap data sampling. This randomness helps to avoid overfitting during the training process. Essentially, the constructed model depends on several parameters. The most important ones are the number of trees, the maximum depth, and the maximum split. The decision trees pick their splitting properties from a random subset of k characteristics at each internal node. The best split is taken within these randomly chosen attributes, and it builds the trees without trimming (Sraitih *et al.*, 2021). Random Forest can be effective in arrhythmia detection, due to its ability to handle complex and non-linear relationships between input features and output classes. It's important to note that the success of using Random Forest for ECG arrhythmia detection depends on the quality and diversity of the dataset, the choice of features, and the tuning of hyperparameters. Experimentation and fine-tuning may be required to achieve the best results.

3.2.6.3 Support Vector Machine (SVM)

SVM is a supervised learning algorithm used for binary classification, which can also be extended to handle multi-class problems as well. In the context of arrhythmia detection, Practically, the SVM technique performs a classification task by constructing a separating hyper-plane in n-dimensional space (*n* is the number of features used as inputs) that separates different class labels by maximizing the geometric margin between the input data classes mapped in a higher-dimensional space and minimizing the empirical classification error (Sraitih *et al.*, 2021). SVM is algorithm uses the hyperplane to classify the data with the main objective of finding the best and optimal hyper-plane in n-dimension space because it requires such type of hyper-plane which has a maximum margin. Hence, the thumb rule is choosing the hyper-plane, which separates the two classes better. In Figure 7, the circle and rectangle are two classes and separated from each other by

hyper-plane. The optimal hyper-plane will be chosen based upon the maximum margin between the classes **Ahamed *et al.*, 2022)**.

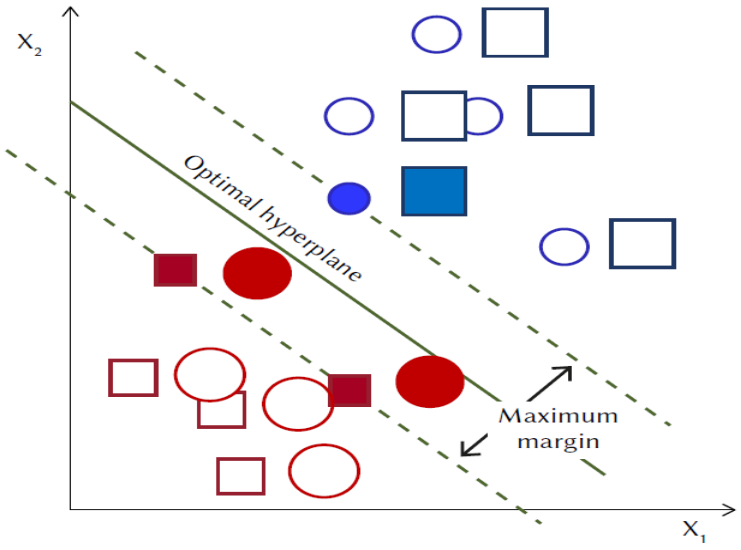

Figure 7. Support Vector Machine Algorithm (Source: Ahamed *et al.*, 2022).

SVM depends in all this classification process on the kernel functions (Sraitih *et al*., 2021), either as a linear or nonlinear classifier according to the type of its kernel function. A linear kernel function makes the SVM a linear classifier. On the other hand, the polynomial and sigmoid kernels make the SVM a non-linear classifier. We suppose a training set that consists of N samples (y_j,x_j), $j = 1, ..., N$, where x $\mathbb{R}^n$ indicates the n-dimensional feature vector of the jth example and y_j $\in$ signifies the matching class label and $y_j \in \pm 1$. A decision function $g(x)$ learned from the training set makes a representation of the optimal hyper-plane that predicts the class label in the subsequent tests. By using the kernel, the decision function is formulated as follow (Sraitih *et al*., 2021):

$$g(x) = sign\left(\sum_{i \in SV_s} \propto_i y_i K(x, x_i) + b\right) \quad (4)$$

where α is the Lagrange multiplier for each training data set and $K(x_j, x)$ is the kernel function that maps the data into higher dimensional space and is defined as in the case of the polynomial (Sraitih *et al*., 2021):

$$K(x, x_i) = \sum(x, x_i)^d \quad (5)$$

where d is the degree of the polynomial function.

3.2.6.4 Naive Bayes

Naive Bayes is a probabilistic machine learning algorithm that can be used in the detection of arrhythmia from ECG signals. Naive Bayes algorithm utilizes Bayes' theorem to calculate the conditional probabilities of each arrhythmia class given the observed features. The probability of the samples is calculated using following formula:

$$\mathbf{Prob(X \mid Y)=(Prob(Y \mid X)*Prob(X))/Prob(Y)} \quad (6)$$

The class (c, target) given predictor's posterior probability is as (X|Y). The prior probability of the class is probability(X). The probability of a given class's predictor is represented b y probability (Y|X). The predictor's prior probability is probability(Y) (Vanajakshi et al., 2023).

The algorithm assumes that the features are conditionally independent of each other, which is known as the "naive" assumption. Despite this simplifying assumption, Naive Bayes can still provide effective results in many cases.

3.2.6.5 Decision Tree

A decision tree is a non-parametric supervised learning method. It is a tree-structured classifier where the characteristic of a dataset is defined by internal nodes and branches of the tree represent the decision rules and the outcome is defined by each leaf node **Ahamed *et al.*, 2022** .

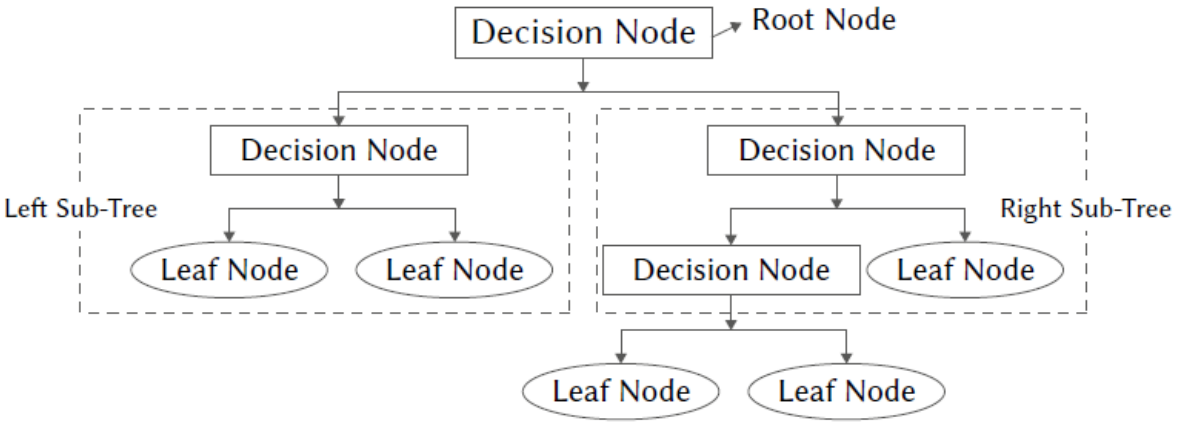

Figure. 8. Decision Tree Algorithm (Source: Ahamed *et al.*, 2022).

There are two nodes in the Decision tree, which are the leaf node and the decision node. Decision nodes have several branches and are used to make some decisions while leaf nodes are the outcome of such decisions and do not have any further branches. A test or decision is carried out based on features of the dataset. A decision tree asks a question, and it further partitions the tree into sub-trees based on the appropriate response (Yes / No), as portrayed in Figure 8 (**Ahamed *et al.*, 2022).**

4.0 Experimental Results and Discussion

Several experiments were performed in extracting and selecting the best feature sets that were later fed to the machine learning algorithms for classification.

4.1. Performance Measurements

The following metrics were used to evaluate the performance of the classification accuracy (Acc), sensitivity or recall (Sen), specificity (Spe), precision (Pre), and F1 score (F1), where TP is true positive, FN is false negative, TN is true negative, and FP is false positive. The equations for each metric are defined as follows:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} . 100 \quad (7)$$

$$\text{Sen} = \frac{\text{TP}}{(\text{TP} + \text{FN})} . \ 100 \quad (8)$$

$$\text{Spe} = \frac{\text{TN}}{(\text{TN} + \text{FP})} . 100 \quad (9)$$

$$\text{Pre} = \frac{\text{TP}}{(\text{TP} + \text{FP})} . \ 100 \quad (10)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

4.2 Results

4.2.1 Result from RF

Table 1 and Figure 9 show the classification report and the confusion matrix of the random forest model.

Table 1: Classification Report of the Random Forest Model

	Precision	Recall	F1-score	Support
n	0.75	0.84	0.79	135
q	0.91	0.85	0.88	135
v	0.87	0.91	0.89	135
s	0.97	0.85	0.91	135
f	0.97	0.99	0.98	135
accuracy			0.89	675
macro avg.	0.89	0.89	0.89	675
weighted avg.	0.89	0.89	0.89	675

The classification report shows the accuracy, precision, and F1 score. The result of the classification report shows that the random forest model is more efficient.

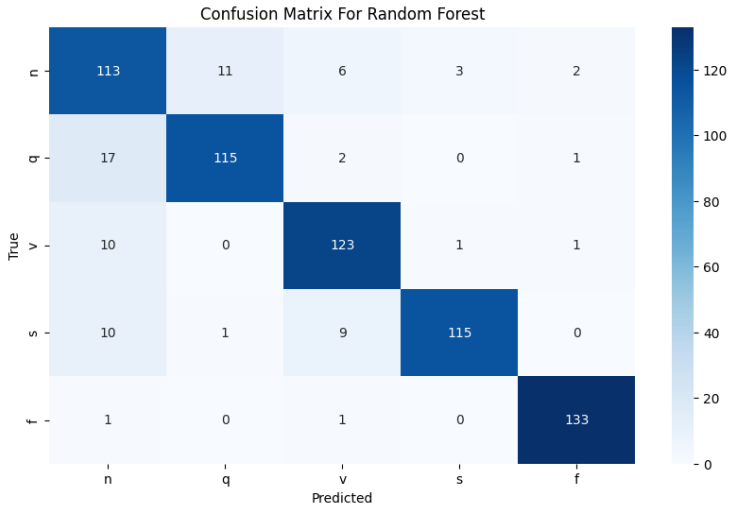

Figure 9: Confusion matrix of the Random Forest Model

The result of the confusion matrix shows that there are some levels of false positive and negative rates. The percentage of true and true negative rates is 89%, and the percentage of false positive and negative rates is 11%.

4.4.2 Decision Tree Results

Table 2 and Figure 10 show the classification report and the confusion matrix of the random forest model.

Table 2: Classification Report for Decision Tree Algorithm

	Precision	Recall	F1-score	Support
n	0.64	0.61	0.63	135
q	0.71	0.79	0.75	135
v	0.79	0.84	0.81	135
s	0.91	0.77	0.84	135
f	0.91	0.92	0.92	135
accuracy			0.79	675
macro avg.	0.79	0.79	0.79	675
weighted avg.	0.79	0.79	0.79	675

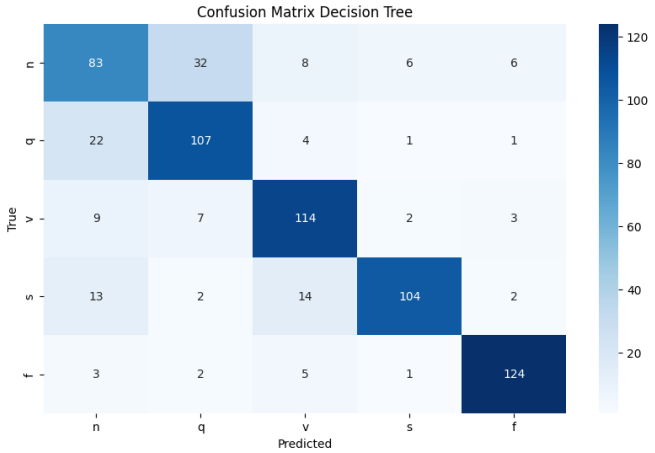

Figure 10: Confusion matrix of the Decision Tree Model

The result of the confusion matrix shows that there are some levels of false positive and negative rates. The percentage of true and true negative rates is 79%, and the percentage of false positive and negative rates is 21%.

4.2.3 SVM

Table 3: and Figure 11 show the classification report and the confusion matrix of the support vector machine.

Table 3: Classification Report for Support Vector Machine

	Precision	Recall	F1-score	Support
n	0.63	0.66	0.64	135
q	0.82	0.71	0.76	135
v	0.73	0.84	0.78	135
s	0.90	0.87	0.88	135
f	0.95	0.93	0.94	135
accuracy			0.80	675
macro avg.	0.81	0.80	0.80	675
weighted avg.	0.81	0.80	0.80	675

The classification report shows the accuracy, precision, and F1-score. The result of the classification report shows that the support vector machine shows an accuracy of 80% on the test data.

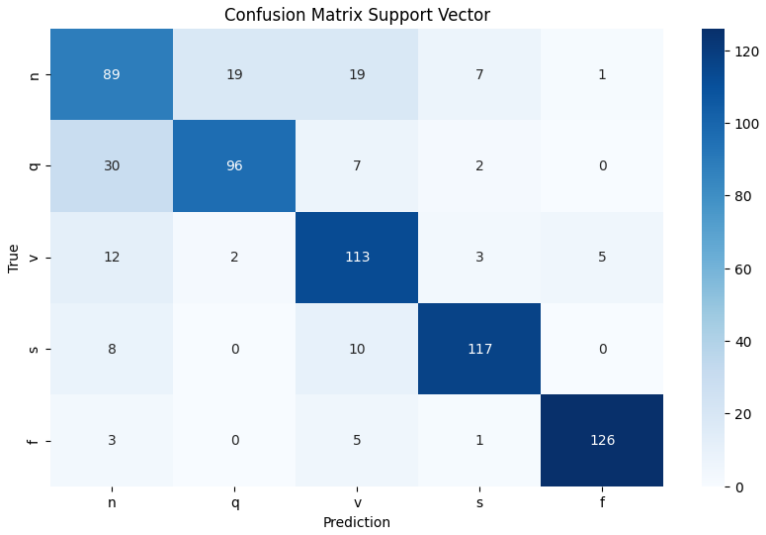

Figure 11: Confusion matrix of the Support Vector Model

The result of the confusion matrix shows that there are some levels of false positive and negative rates. The percentage of true and true negative rates is 80%, and the percentage of false positive and negative rates is 20%.

4.2.4 KNN

Table 4 and Figure 12 show the classification report and the confusion matrix of the support vector machine.

Table 4: Classification Report for K-Nearest Neighbor

	Precision	Recall	F1-score	Support
n	0.71	0.76	0.74	135
q	0.83	0.78	0.80	135
v	0.87	0.87	0.87	135
s	0.88	0.90	0.89	135
f	0.98	0.96	0.97	135
accuracy			0.85	675
macro avg.	0.86	0.85	0.85	675
weighted avg.	0.86	0.85	0.85	675

The classification report shows the accuracy, precision, and F1-score. The result of the classification report shows that the K-nearest neighbor shows an accuracy of 85% on the test data.

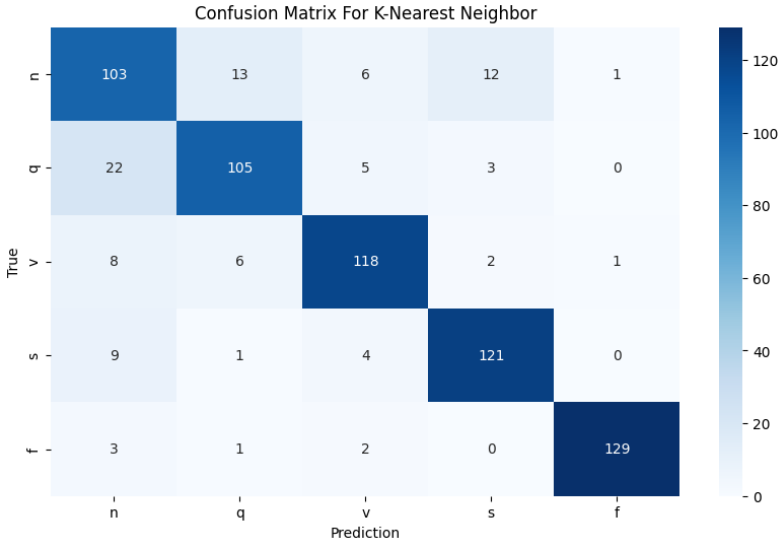

Figure 12: Confusion matrix of the K-Nearest Neighbor

The result of the confusion matrix shows that there are some levels of false positive and negative rates. The percentage of true and true negative rates is 85%, and the percentage of false positive and negative rates is 15%.

4.2.5 Naïve Baye

Table 5 and Figure 13 show the classification report and the confusion matrix of the support vector machine.

Table: Classification Report for Naïve Baye

	Precision	Recall	F1-score	Support
n	0.62	0.59	0.61	135
q	0.78	0.76	0.77	135
v	0.72	0.79	0.76	135
s	0.87	0.86	0.87	135
f	0.94	0.93	0.94	135
accuracy			0.79	675
macro avg.	0.79	0.79	0.79	675
weighted avg.	0.79	0.79	0.79	675

The classification report shows the accuracy, precision, and F1-score. The result of the classification report shows that the logistic regression shows an accuracy of 79% on the test data.

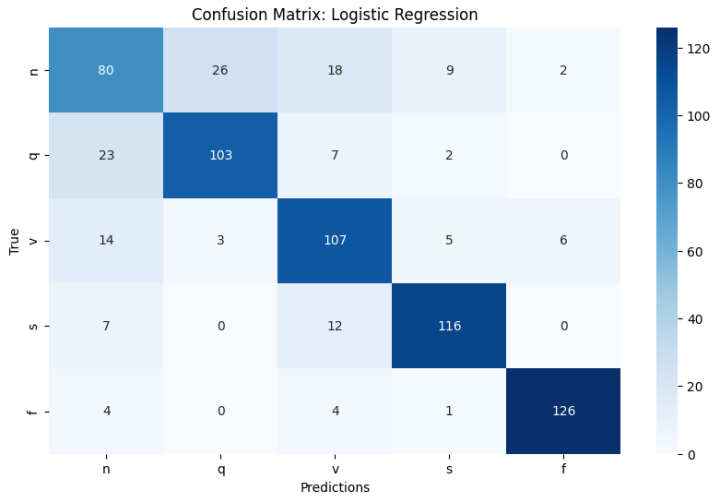

Figure 13: Confusion matrix of the Logistic Regression

The result of the confusion matrix shows that there are some levels of false positive and negative rates. The percentage of true and true negative rates is 79%, and the percentage of false positive and negative rates is 2%.

The comparison between the five models is presented in table 6 and figure 14.

Table 6: Comparison of the results obtained by the ML Models

Models	Sensitivity	Specificity	Precision	F1-score	Accuracy
RF	85	75	89	89	89
DT	76	74	79	79	79
SVM	80	75	81	80	80
KNN	85	75	86	85	85
NB	78	75	79	79	79

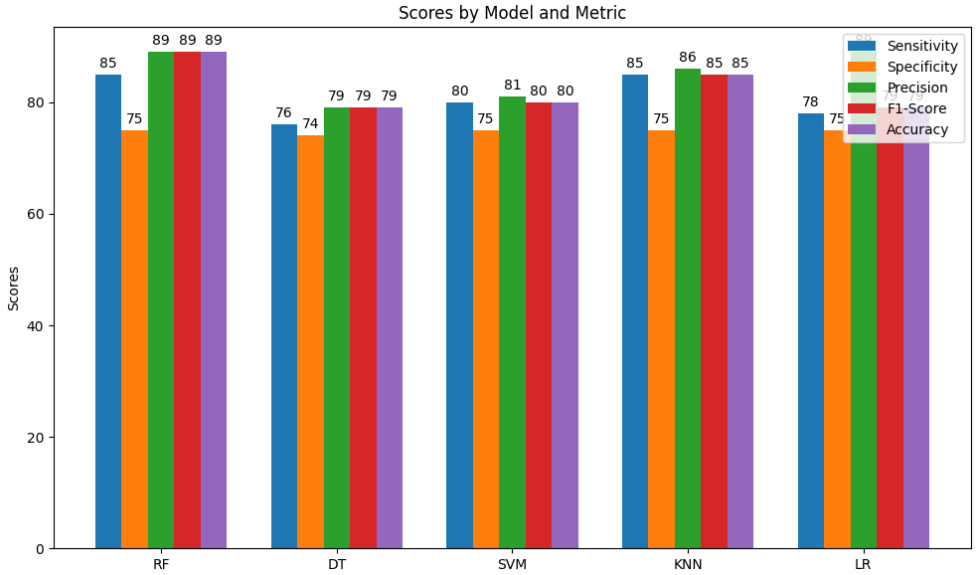

Figure 14: Comparison of the ML models

4.3 Discussion

In this study, several supervised ML algorithms to predict cardiac arrhythmia are investigated and found that RF presented the best performance. The excellent prediction performance of the model was validated with the MIT-BIH benchmark dataset. The prediction model may facilitate more responsive clinical care that is beneficial for patients through early identification. The Study have shown that ML models can effectively predict the risk of arrhythmias and can overcome several limitations in traditional arrhythmia detection by cardiologist. The RF algorithm had the highest accuracy of 89% and hence the lowest miss classification rate. DT and NB both hence the highest miss classification rate of 21% each.

4.4 Conclusion

The Study presents a comparison of supervised machine learning classification techniques in automatic arrhythmia classifier based on simple RR interval of the ECG signal, with small amount of ECG dataset. Five supervised machine learning models: Support Vector Machine (SVM),K-Nearest Neighbors (KNN), Random Forest (RF), and Decision Tree. The study tested the performance of these techniques in classifying: Normal Beat (NOR), Left Bundle Branch Block Beat (LBBB), Right Bundle Branch Block Beat (RBBB), Premature Atrial Contraction (PAC), and Premature Ventricular Contraction (PVC), on the MIT-BIH dataset. The study shows that the RF algorithm had the best performance of 89.00% in comparison with the other methods using the proposed method. Future study will tend to use deep learning to support automatic feature extraction as well as increase the size of the dataset.

REFERENCES

Ahamed, J., Koli, A.M., Ahmad, K., Jamal, A.M., & Gupta, B.B (2021) CDPS-IoT: Cardiovascular Disease Prediction System Based on IoT Using Machine Learning. *International Journal of Interactive Multimedia and Artificial Intelligence, 7(4).* DOI: 10.9781/ijimai.2021.09.002

Acharya, U. R., Fujita, H., Adam, M., Lih, O. S., Sudarshan, V. K., & Tan, J. H., et al. (2016). Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ecg signals: a comparative study. Information Sciences, 377, (pp.17-29).

Ahamed, J., Koli, A.M., Ahmad, K., Jamal, M.A & Gupta, B.B. (2021). CDPS-IoT: cardiovascular disease prediction system based on IoT using machine learning. *International Journal of Interactive Multimedia and Artifcial Intelligence*, 7(4), 78–86.

Alarsan, F.I., & Younes, M. (2019). Analysis and Classification of Heart Diseases using Heartbeat Features and Machine Learning Algorithms . *Journal of Big Data 6(81).*

Asl, B.M., Setarehdan, S.K., & Mohebbi, M. (2008). Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artificial . Intelligence. Medical*. 44, 51–64.

Batra, A., & Jawa, V. (2016). Classification of arrhythmia using conjunction of machine learning algorithms and ECG diagnostic criteria. *International Journal of Biology. Biomed.* 1, 1–7.

Bazi, Y., AlRahhal, M.M., AlHichri, H., Ammour, N., Alajlan, N., & Zuair, M. (2020). Real-Time Mobile-Based Electrocardiogram System for Remote Monitoring of Patients with Cardiac Arrhythmias. *International Journal of Pattern Recognition and Artificial Intelligence, 34(*10), 2058013-205801324. DOI: 10.1142/S0218001420580136.

Cañon-Clavijo,R.E., Montenegro-Marin, C.E., Gaona-Garcia P.A., & Ortiz-Guzm´an, J. (2023). IoT Based System for Heart Monitoring and Arrhythmia Detection Using Machine Learning. *Journal of Healthcare Engineering* 6401673, https://doi.org/10.1155/2023/6401673

Carrillo-Alarcón, J.C., Morales-Rosales, L.A., Rodríguez-Rángel, H., Lobato-Báez, M., Muñoz A., & Algredo-Badillo, I. (2020). A Metaheuristic Optimization Approach for Parameter Estimation in Arrhythmia Classification from Unbalanced Data. *Sensors* 2020, 20, 3139..

Devadas, R.M. (2021). Cardiac Arrhythmia Classification using svm, knn and naive bayes Algorithms. *International Research Journal of Engineering and Technology (IRJET), 8*(5), 3939-3941.

Han C-W. (2017). Detecting an ECG Arrhythmia using cascade Architectures of fuzzy neural networks. *Adv Sci Technol Lett* .143:272–5.

Hassanien, A.E., Kilany, M., & Houssein, E. H. (2020). Combining Support Vector Machine and Elephant Herding Optimization for Cardiac Arrhythmias. *arXiv:1806:08242v1*.

Kumar, N. K., Sindhu, G. S., Prashanthi, D. K., & Sulthana, A. S. (2020). Analysis and prediction of cardio vascular disease using machine learning classifiers. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* **.** 15-21**.** *IEEE.*

Kumari, C.U., Murthy, A.S.D., Prasanna, B.L., Reddy, M.P.P., & Panigrahy, A.K. (2021). An automated detection of heart arrhythmias using machine learning technique: SVM. *Materials Today Proceedings*, 45,1393–1398.

Li, T., & Zhou, M. (2016). Ecg classification using wavelet packet entropy and random forests. Entropy, 18(8), (pp.1-16).

Marjia S., Haider, A., & Uddin, M. S. (2016). Analysis of Data Mining Techniques for Heart Disease Prediction . 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT) DOI:10.1109/CEEICT.2016.7873142

Montenegro, L.; Abreu, M.; Fred, A.; Machado, J.M. (2022). Human- Assisted vs. Deep Learning Feature Extraction: An Evaluation of ECG Features Extraction Methods for Arrhythmia Classification Using Machine Learning. *Appl. Sci. 12*, 7404. https://doi.org/10.3390/app12157404

Moody, G.B., Mark R.G., & Goldberger A. L. (2001). PhysioNet: a Web-based Resource for the study of Physiologic Signals. *Journal of Engineering in Medicine & Biology Magazine IEEE* 2001; 20(3): 70-75.

Pandey, S.K., Janghel, R.R., & Vani, V. (2020).Patient Specific Machine Learning Models for ECG Signal Classification. *Procedia Computer Science, 167, 2181–2190* .

Qin, Q., Li, J., Zhang, L., Yue, Y., & Liu, C. (2017). Combining Low-dimensional Wavelet Features and Support Vector Machine for Arrhythmia Beat Classification. *Scientific ReportS.* DOI:10.1038/s41598-017-06596-z.

Ramalingam, V., Dandapath, V.A., & Raja, M.K. (2018). Heart disease prediction using machine learning techniques : a survey, *International Journal of Engineering & Technology,* 7, 684.

Shanmugavadive, K., Sathishkumar, V.E., & Kumar, M.S. (2022). Investigation of Applying Machine Learning and Hyperparameter Tuned Deep Learning Approaches for Arrhythmia

Detection in ECG Images. *Hindawi:Computational and Mathematical Methods in Medicine*. https://doi.org/10.1155/2022/8571970.

Sultan-Qurraie, S., & Afkhami, R.G. (2017). ECG arrhythmia classification using time frequency distribution techniques. *Biomedical Engineering Letters*, 7(4), 325–332.

Sraitih, M., Jabrane, Y., & Hajjam , E.H, A. (2021). An Automated System for ECG Arrhythmia Detection Using Machine Learning Techniques. *Journal of Clinical Medicine , 10*, 5450. https://doi.org/10.3390/jcm10225450

Taloba, A.I., Alanazi R., ,Shahin, O.R., Elhadad, A., Abozeid, A., & El-Aziz, R.M.A. (2021)Machine Algorithm for Heartbeat Monitoring and Arrhythmia Detection Based on ECG Systems. *Hindawi: Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2021/7677568.

Vanajakshi, P., Karthik. D.R., & Chanda, S. (2023). Prediction of Cardiac Arrhythmia using Machine Learning . *International Journal of Engineering Research & Technology (IJERT).* 12(2).

WHO. Cardiovascular Diseases. Available online: https://www.who.int/health-topics/cardiovasculardiseases/(accessed on 13 December 2019).