JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ARRYTHMIA CLASSIFICATION USING ECG DATA VALUES

[1]P.Gayatri,2Nanduri Naveen
[1]Assistant Professor,[2]MCA 2nd year,
[1]Master Of Computer Applications,
[1]Sanketika Vidya Parishad Engineering College,Visakhapatnam, India

ABSTRACT
A condition known as cardiac arrhythmia causes an irregular heartbeat that can be either too slow or too fast. Faulty electrical impulses that control the heartbeats cause it to occur. Certain severe arrhythmia diseases can lead to sudden cardiac death. Because of this, the main objective of electrocardiogram (ECG) examination is to accurately identify arrhythmias as life-threatening in order to offer an appropriate therapy and save lives. ECG signals are waveforms (P, QRS, and T) that represent the electrical activity of the human heart. Each waveform's duration, organisation, and spacing between different peaks are used to determine whether there are any heart issues. The parameters of the AR signal model are then determined by doing an autoregressive (AR) analysis on the signals. The training dataset neatly separates the groups of recovered AR features for three different types of ECG, giving each ECG signal in the training dataset good connection classification and heart condition diagnosis. To more accurately assess ECG signals, a novel method based on fractional Fourier transform (FFT) algorithms and two-event-related moving averages (TERMAs) is proposed. This study may aid in the examination of the most recent cutting-edge techniques used in arrhythmia situation detection. Our proposed machine learning method features cross-database training and testing with enhanced properties.
Keywords: Cardiac arrhythmia, irregular heartbeat, electrical impulses, electrocardiogram (ECG), P wave, QRS complex, T wave, autoregressive (AR) analysis, AR signal model, heart condition diagnosis, fractional Fourier transform (FFT), two-event-related moving averages (TERMAs), machine learning, cross-database training, arrhythmia situation detection.

I.INTRODUCTION

Arrhythmias come in a variety of forms, and because each type is linked to a certain pattern, it is easy to recognise and categorise each type[1]. Two broad categories can be used to group the arrhythmias[2]. The first classification includes arrhythmias that are created by a single irregular heartbeat, or morphological arrhythmia[3]. The second group of arrhythmias, referred to as rhythmic arrhythmias here, are those caused by a series of erroneous heartbeats[4]. The classification of normal heartbeats and the heartbeats that make up the first group are the main subjects of this study[5]. The ECG exam can detect all of these variations since these heartbeats modify the wave frequency or shape. A human being may have a difficult time diagnosing and categorising arrhythmias because it occasionally becomes essential to examine each heartbeat in the ECG data recorded by a Holter monitor, for example, over the course of hours or even days[6]. Additionally, human mistake owing to weariness during the processing of the ECG records is a potential[7]. Utilising computational methods for automatic classification is an alternative. Four phases can be used to create a fully autonomous system for classifying arrhythmias from signals obtained from an ECG device: ECG signal preprocessing, heartbeat segmentation, feature extraction, and learning/classification[8]. Every one of the four processes involves doing something, and the determination of the type of heartbeat is the ultimate goal[9]. ECG data preprocessing and heartbeat segmentation, the first two elements of such a classification system, have received extensive study in the literature. Precaution should be taken when selecting the preprocessing approaches because they directly affect the outcomes[10]. In the case of QRS detection, the results of the heartbeat segmentation stage are extremely near to ideal[11]. However, the classification-related phases (feature extraction and learning algorithms) still have potential for investigation and advancement. Even though the issue of ECG demarcation is still up for debate, the approaches described in the literature we have reviewed do not really benefit from it[12]. This study reviews previous research on ECG-based arrhythmia classification methods that has been published in the literature[13]. It also examines the key approaches for building these automatic systems and the two main paradigms for evaluation: intra-patient and inter-patient[14]. Additionally, the most used databases and issues with evaluating the most recent techniques described in the literature are reviewed. A workflow is suggested as a result of this debate to direct the assessment of future works. Please take note that this survey work's major contribution is the procedure for the evaluation process. A review of knowledge-based ECG interpretation techniques put forth in the 20th century is found in the literature[15]. Clifford et al. have out a thorough examination of the techniques for ECG signal analysis[16]. Their research concentrated on the signal's physiology as well as its processing methods, particularly feature extraction and classification. Clifford et al., in particular, did not concentrate on the topic of evaluating methodologies, which is unique to our work in addition to a more recent literature analysis on the subject. Additionally, a special analysis of feature selection is included in our feature extraction survey. The subsequent sections of this essay are structured as follows. Section 2 introduces the basic elements of ECG signals; Sections 3, 4, 5, and 6 describe the state-of-the-art; Sections 7, 8, and 9 discuss the evaluation standards developed by the Association for the Advancement of Medical Instrumentation (AAMI) and the databases recommended for these standards, along with criticisms of the systems currently in development and upcoming challenges. More precisely, Section 3 discusses the preprocessing methods used most often on ECG signals, and Section 4 introduces the idea of separating heartbeats from ECG signals and the methods used to do so. The depiction of a heartbeat or the feature extraction procedure, which is essential for the success of arrhythmia classification, is covered in Section 5. The most well-liked learning techniques for categorising arrhythmias are covered in Section 6 of the literature. In Section 7, the recommended assessment standard put out by AAMI is presented, along with information on the most popular databases used to evaluate classifying arrhythmia algorithms as indicated by the standard. Section 8 contains some comments on the topic of data selection for learning/evaluating models for the classification of arrhythmias and its effects on the outcome. The field's shortcomings and issues are finally covered in Section 9, which also highlights upcoming difficulties for the research community.

II.EXISTING SYSTEM

There are various existing systems and approaches related to arrhythmia diagnosis, treatment, and management. Here are a few examples:

1.Electrocardiogram (ECG): The electrical activity of the heart is measured by an ECG, a common diagnostic procedure for arrhythmia[17]. Electrodes are positioned on the chest, arms, and legs for this non-invasive test in order to look for electrical signals. Atrial fibrillation, ventricular tachycardia, and other arrhythmias can be diagnosed using an ECG.

2.Implantable cardioverter-defibrillator (ICD): A tiny device known as an ICD is inserted beneath the skin of the chest and connects to the heart[18]. It is used to monitor cardiac rhythm and shock the heart with electricity if a risky arrhythmia is found. ICDs can assist prevent sudden cardiac death and are frequently used to treat ventricular arrhythmias[19].

3.Cardiac ablation: Catheters are used to deliver radiofrequency radiation to the heart tissue during a process known as cardiac ablation, which results in the formation of scar tissue[20]. The aberrant electrical signals that lead to arrhythmias can be stopped by this scar tissue. Atrial fibrillation, ventricular tachycardia, and other arrhythmias are all treated with cardiac ablation[21].

4.Anti-arrhythmic medications: Anti-arrhythmic drugs are used to treat different forms of arrhythmias by regulating the electrical activity of the heart[22]. These drugs function by preventing the heart's aberrant electrical impulses. They may not work for all types of arrhythmias and can have negative effects.

5.Wearable and mobile health technology: Arrhythmias can be identified via wearable technology, such as smartwatches and mobile health apps, which can monitor heart rate and rhythm. These tools can support real-time cardiac rhythm monitoring for patients and healthcare professionals[23].

PROBLEMS IN EXISTING SYSTEM

Although the outcomes of the predictions are encouraging, these conventional methods are still far from being highly precise and effective.

Although the current procedures are straightforward and efficient, they are incredibly susceptible to change. Furthermore, modern techniques only use one algorithm, which leads to erroneous findings.

ISSUES IN EXISTING METHODOLOGY

Although the outcomes of the predictions are encouraging, these conventional methods are still far from being highly precise and effective. Although the current procedures are straightforward and efficient, they are incredibly susceptible to change. Furthermore, modern techniques only use one algorithm, which leads to erroneous findings. This can cause practitioners to make erroneous assumptions, which would affect the patient's diagnosis and care**.**

III.PROPOSED SYSTEM

Here are some general concepts that might be taken into account when creating a proposed system:

1. Creation of a machine learning-based algorithm capable of accurately identifying arrhythmias from ECG data[24]. To enhance the algorithm's performance, a sizable dataset of ECGs can be used to train it.
2. Development of a decision-support system to help clinicians select the ideal course of action for every patient based on their unique medical history, genetics, and other considerations[25]. Data from numerous sources, including ECGs, genetic tests, and patient electronic health records, can be integrated by this system.
3. The creation of a wearable or implantable gadget that can constantly track heartbeat and spot arrhythmias in real-time. If necessary, this device can alert medical professionals or emergency personnel.
4. Development of a virtual reality-based training course for medical professionals to enhance their expertise in identifying and treating arrhythmias. With the help of this programme, clinical decision-making can be improved by simulating various scenarios and receiving feedback.
5. Creating a mobile health app with capabilities for self-monitoring and self-management of the patient's condition and information regarding arrhythmias. Patients can keep track of their symptoms, prescriptions, and appointments using this app, which can interact with wearable technology.

ADVANTAGES OF THE PROPOSED SYSTEM

The following are the clear benefits that are shown when specific variances are taken into account in the current system: The performance classification of arrhythmia-based diseases is now even better: thanks to a deeper understanding of the various diseases in the field of medicine, it is now much simpler to distinguish between the various types of liver diseases and their occurrence. The process of gathering data has become simpler because to developments in data mining paradigms and software architectures like Hive and R. Depending on the needs of the user, multiple machine learning models can assess time complexity and accuracy, allowing us to measure different parameters: Every prediction system is built around the types of parameters that it is intended to accept, contrast, and then reach a judgement about. As a result, several algorithms are utilised to model the predictive system according to the context. The various machine learning algorithms evaluate the testing parameters and the disease kind.

MACHINE LEARNING

One of the most rapidly developing technologies nowadays is machine learning. The study of methods and models in statistics that are simple enough for machines to grasp to carry out and accomplish particular tasks is known as machine learning. This method has evolved and is now expected in the majority of fields.

OUR MOTIVE

Our goal in this project is to forecast an arrhythmia for a patient as accurately as possible. This dataset was gathered from the Kaggle database, and we used it in our three modules to forecast the arrhythmia disease using different machine learning methods.

PATIENT SAMPLES

The percentage of persons with arrhythmia disease has been rising, and the methods for predicting these disorders are becoming less and less effective. The techniques employed in this project would make the prediction techniques more effective. To achieve this, it is necessary to gather and compile the samples of various patients into a dataset. Then, to increase the accuracy percentage, the samples of the patients we have collected are refined.

INCREASED ACCURACY

The algorithms utilised in our project will boost the accuracy in forecasting the arrhythmia disease despite the fact that there have been many algorithms developed over the years for disease detection.

DATA MINING

The best technique to discover patterns in large informant indices, such as measurements, database frames, and AI crossing point schemes, is through data extraction. Information mining is an interdisciplinary subject of software engineering and measurement that seeks to disentangle data from information collecting (with astute tactics) and transform data into a comprehensible structure for further usage.

DATA PRE-PROCESSING

Preparing the data is a crucial stage in the solution of every machine learning problem. Most of the data sets utilised with machine teaching problems need to be processed in order for a machine learning algorithm to be taught. There are just a few methods that are frequently used for pre-processing, including scaling, categorical coding, and imputation of absence of value. These methods are simple to comprehend. Each dataset is different and presents its own special problems.

DATA COLLECTION

Data collection is essential for these undertakings. We obtained the Arrhythmia dataset from Kaggle, which has 14 distinct attributes.

DATA CLEANING

The word "attributes" refers to a variety of columns. The next step is to remove the null and erratic values from the datasheet before using that dataset for classification. Some columns have null values and some values are erratic.

IV.TRAINING AND VALIDATING

Train set: This dataset is used to train the model and help it discover the traits and patterns that are concealed in the data. The model receives the same training data repeatedly so that it can continue to learn the features of the data. The training package uses a wide range of inputs to train the model in all scenarios and forecast any future appearances of unseen data samples.

Validation set: The performance of our model during training is verified using the validation set, a data package that is independent from the training set. We can modify model setups and hyperparameter settings as a result of the validation process' informational output. It's comparable to a spokesperson informing us whether or not the training is on the proper track. The model is trained using the training package, and it is evaluated using the validation set each time it is run. The major goal of dividing the dataset into a validation set is to avoid overfitting, which occurs when a model becomes exceptionally good at classifying samples from a training set but is unable to generalise and make appropriate classifications on new data.

DEVELOPING TREATMENTS FOR ARRHYTHMIAS

To cure arrhythmias and enhance patient outcomes, create novel medications, gadgets, or other therapies. To do this, it could be necessary to evaluate the safety and efficacy of various treatment alternatives, look for any potential drawbacks or side effects, and perhaps even propose fresh ideas based on cutting-edge technological advancements.

PREDICTING ARRHYTHMIAS

Construct models that can forecast which people, based on their medical histories, genetics, or other variables, are more likely to experience arrhythmias. Large-scale patient data gathering and analysis, testing the precision of various predictive models, and maybe incorporating machine learning or artificial intelligence approaches could all be part of this.

IMPROVING PATIENT OUTCOMES

Build interventions that can enhance quality of life, decrease the frequency of complications, or enhance patient outcomes. This can entail creating patient education materials, putting new clinical protocols or recommendations into practise, or spotting prospective areas where the quality of current care delivery systems could be improved.

In general, the precise objectives and resources at hand will determine the project's overall scope. It might incorporate a mix of clinical trials, data analysis, and laboratory research.

SAMPLE TRAIN SET:

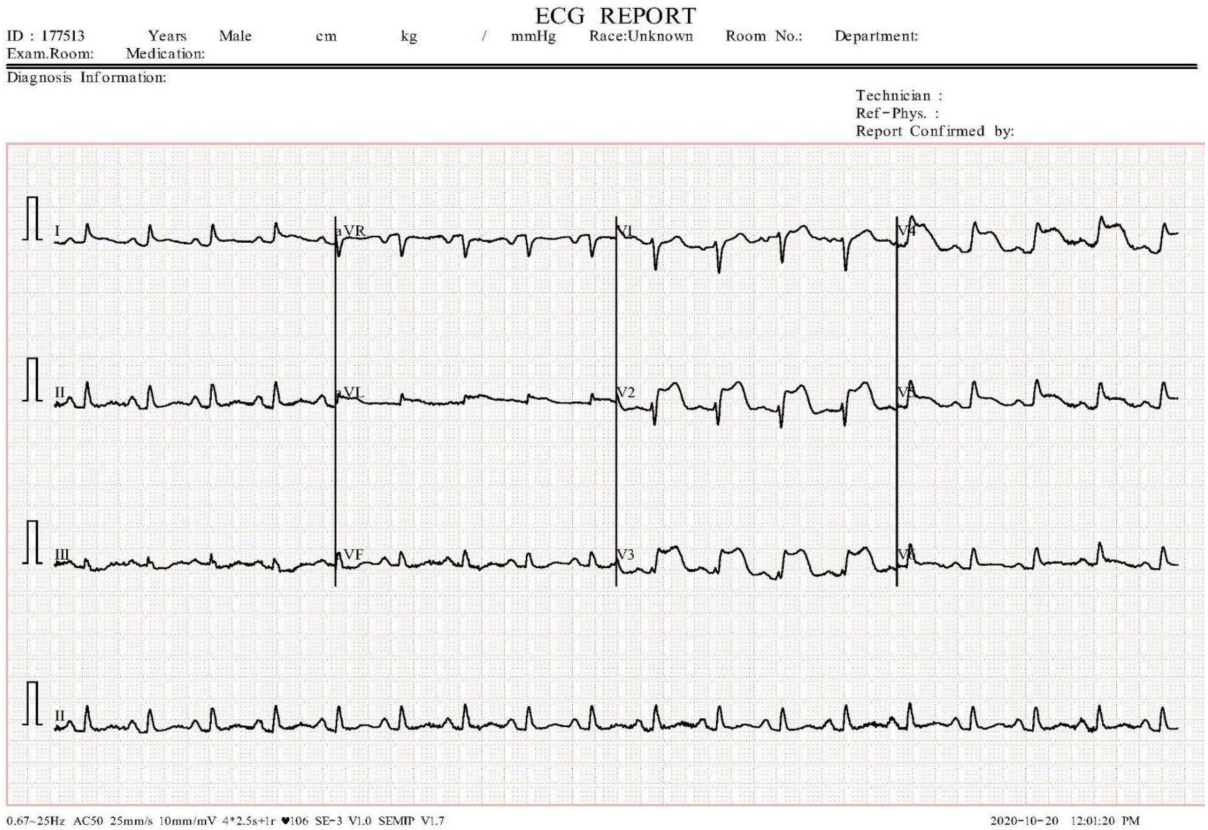

FIGURE 1: ECG of patient with Abnormal heartbeat

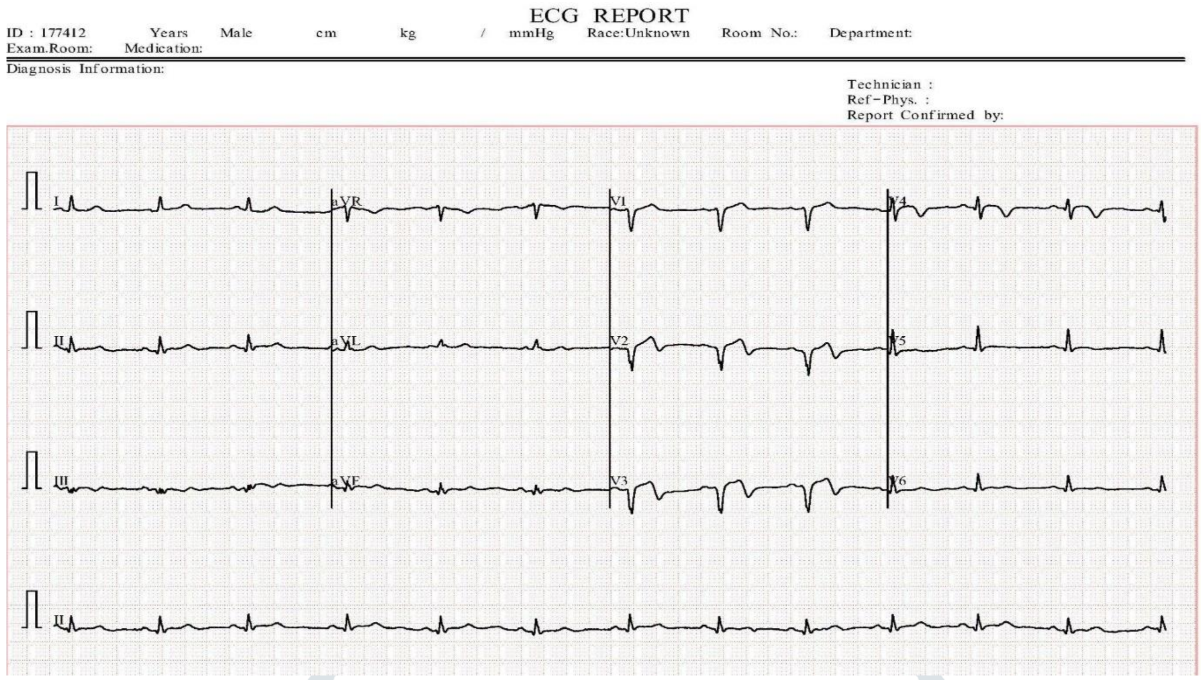

FIGURE 2 : ECG of Myocardial Infarction victims

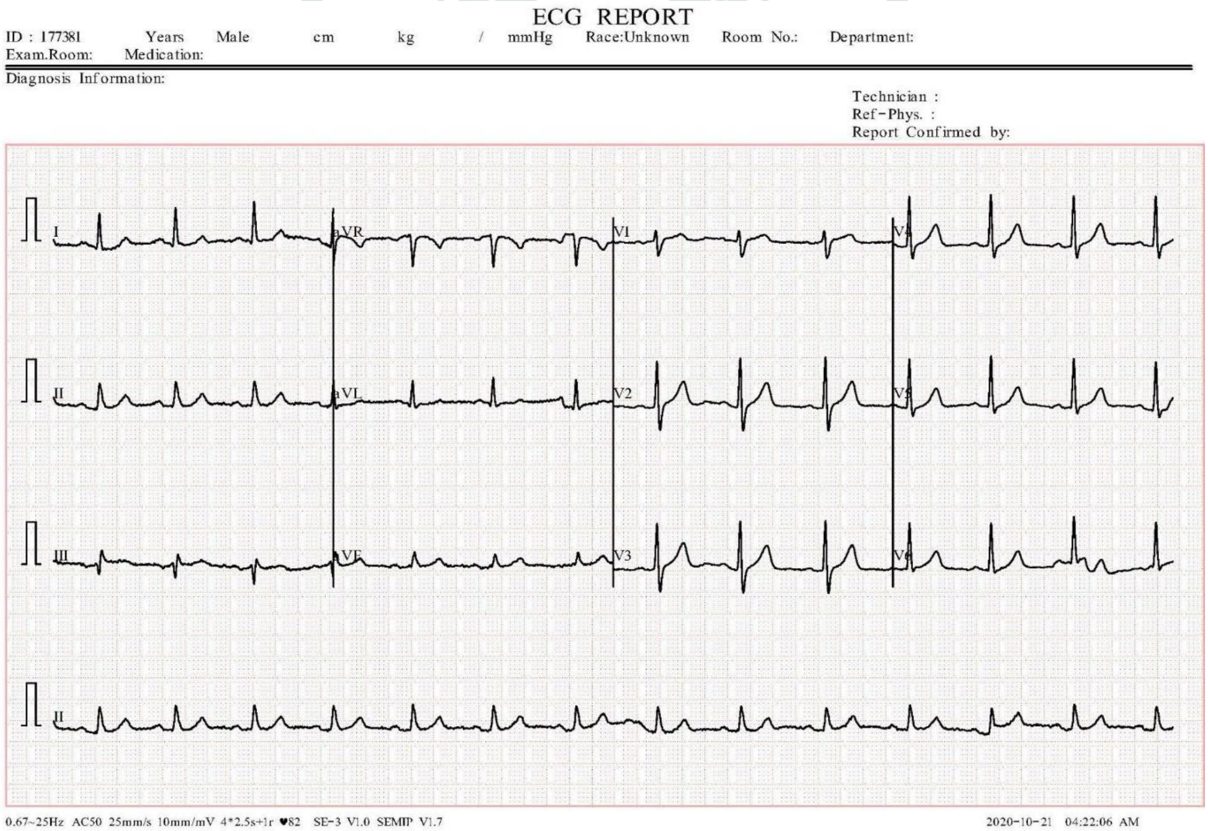

FIGURE 3: ECG of Normal adults

K- NEAREST NEIGHBORS (KNN)

The K-nearest Neighbouring (KNN) algorithm is another name for this approach. The algorithm for lazy learning is another name for it. Its goal is to predict the order of another example point by using a database where information focuses are divided into a few classifications. A KNN sorts a sample to the class that K nearby neighbours agree on the most. The restriction K affects how the categorization algorithms can be changed.

SUPPORT VECTOR MACHINE (SVM)

The SVM algorithm makes an effort to distribute hyper planes and categorise the data. SVM is implemented using the Python scikit-learn library. The pre-processed data is divided into check data and coaching set, which make up separately 25 and 75 percent of the total dataset. In an area with extremely high dimensions, an SVM approach creates hyper planes. In general, the higher the margin, the smaller the generalisation error of the classifier, so a good separation is obtained by the hyper plane that has the most significant distance to the nearest coaching information of any category (so-called intentional margin).

RANDOM FOREST

A supervised machine learning model called random forest is frequently employed in classification and regression issues. On various samples, it constructs decision trees and uses their average for classification and majority vote for regression.

The Random Forest Algorithm's ability to handle data sets with continuous variables, as in regression, and categorical variables, as in classification, is one of its most crucial qualities. For classification and regression tasks, it performs better. In this lesson, we will examine the random forest's operation and apply it to a classification job.

ARCHITECTURE / OVERALL DESIGN OF PROPOSED SYSTEM

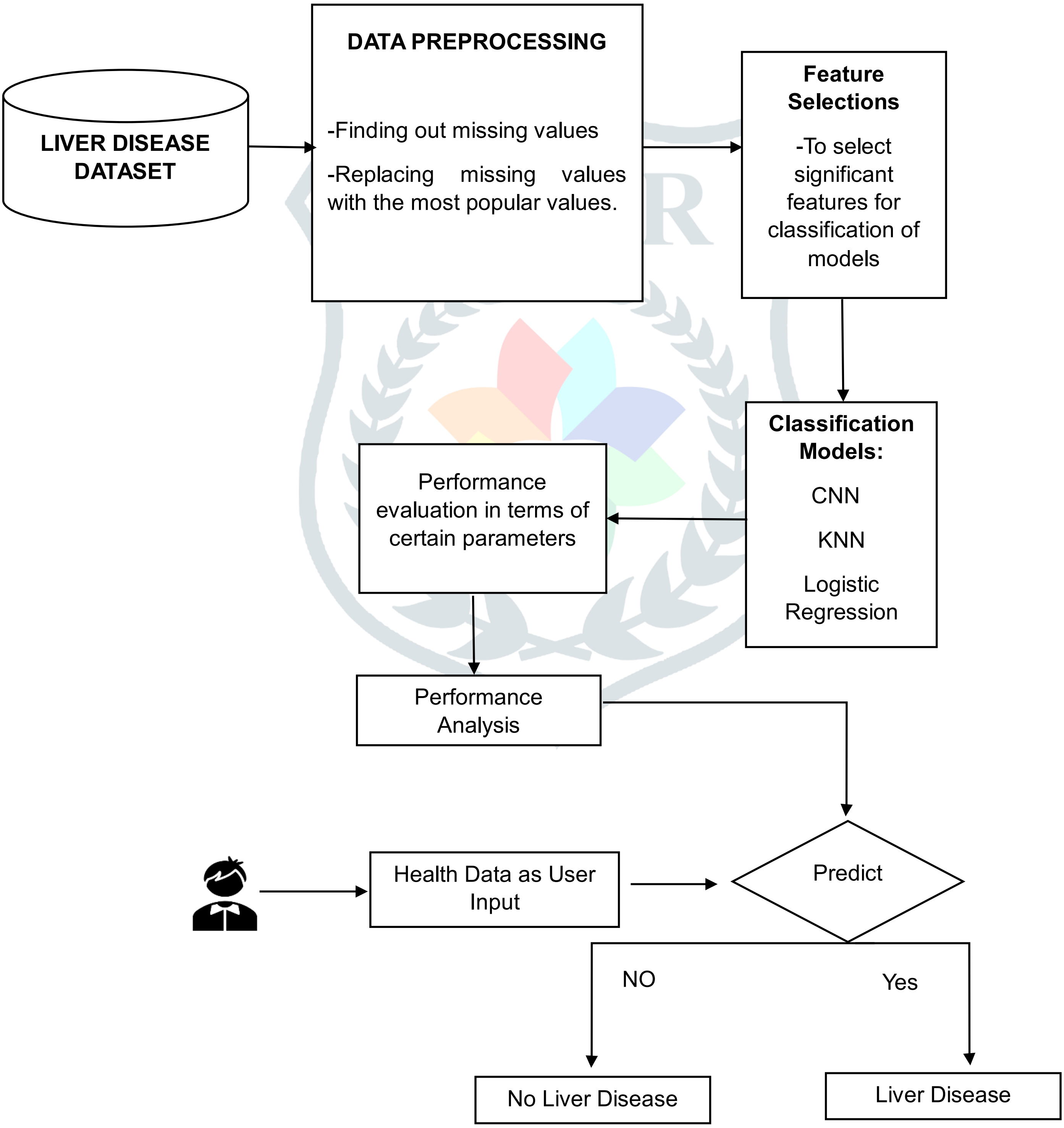

FIGURE 4 : System Architecture

ADABOOST

AdaBoost, also known as Adaptive Boosting, is a machine learning method that is used as an ensemble. Decision trees with one level, or Decision trees with only one split, are the most popular estimator used with AdaBoost. Another name for these trees is Decision Stumps. The gradient boosting algorithm is a different ensemble learning method. Instead than focusing on weights, as in AdaBoost, we strive to reduce error in this technique. However, we will just be concentrating on AdaBoost's mathematical intuition in this essay.

V.RESULTS AND DISCUSSIONS

In this study, an ML model was created to determine whether a user is likely to have arrhythmia disease. 16 sets of human test results make up the input data for this machine learning model. It was confirmed that this dataset was processed and cleaned without any missing or incorrect data values. The data set was divided into x and y, two distinct halves. The output, denoted by the dependent variable y, is a function of the x parameter, which includes the 16 features that correspond to the various test results. The x-axis value was altered using Standard Scaler. For the x and y variables, additional categories called y-test, x-test, x-train, and y-train were made.

The supervised learning techniques K-NN, LR, SVM, and Random Forest were used to learn these y-train and x-train. The accuracy% was calculated using the four approaches using a range of n values. When n = 1 and SVM are used, K Neighbours and SVM both get the highest accuracy. We decided to utilise SVM because it is regarded as one of the most trustworthy forecasting models after doing rigorous experiments.

VI.CONCLUSION

Using ECG data values to categorise arrhythmias is a significant and promising area of research, to sum up. Automated ECG-based arrhythmia classification systems have the potential to increase diagnostic accuracy and efficacy due to the rising occurrence of arrhythmias and the limitations of present diagnostic techniques.

Recent research has demonstrated that machine learning algorithms can be taught on ECG data to correctly classify a variety of arrhythmias, including brady, ventricular, and atrial fibrillation. These devices can analyse ECG signals in real-time, enabling quick diagnosis and timely intervention when necessary.

However, there are still a few issues that must be resolved before ECG-based arrhythmia classification systems can be developed and put into use. These include ensuring the algorithms' precision and dependability, dealing with variations in the quality of the ECG signal, and implementing the systems in clinical settings.

The development of more complex machine learning algorithms that can better distinguish between various arrhythmia types and pinpoint particular characteristics of ECG signals that are indicative of each kind should be the main goal of future research in this field. Research should also look into the possibility of combining ECG data with other types of data to increase accuracy, such as clinical and demographic data.

The creation of algorithms that can forecast the likelihood of developing arrhythmias based on ECG data and other clinical parameters is another crucial field of research. Such algorithms could be used to identify those who are at a high risk of developing arrhythmias and to act sooner to stop or lessen their consequences.

Overall, ECG-based arrhythmia classification systems are a fascinating area of research because of the potential benefits they may have, and future development and improvement of these systems will be crucial for enhancing the diagnosis and management of arrhythmias.

Finally, careful consideration of factors like cost-effectiveness, feasibility, and clinicians' and patients' acceptance will be necessary when integrating ECG-based arrhythmia classification systems into clinical practise. Therefore, future studies should concentrate on addressing these real-world issues and creating plans to make it easier to adopt and use these systems in clinical settings.

To sum up, ECG-based arrhythmia classification algorithms have the potential to increase the precision and effectiveness of diagnosing and treating arrhythmias. Even though there are still issues to be resolved, ongoing research in this field shows significant potential for enhancing the care of people with arrhythmias.

VII.REFERENCES

[1] A book reference of Arrhythmias come in a variety of forms.
https://link.springer.com/chapter/10.1007/978-981-19-7753-4_75

[2] A website reference of group the arrhythmias.
https://dom-pubs.onlinelibrary.wiley.com/doi/full/10.1111/dom.14854

[3] An article reference of morphological arrhythmia.
https://www.sciencedirect.com/science/article/abs/pii/S174680942200876X

[4] A book reference of series of erroneous heartbeats.
https://www.mdpi.com/1424-8220/23/1/260

[5] A website reference of classification of normal heartbeats.
https://journals.sagepub.com/doi/full/10.1177/20552076221150741

[6] An article reference of diagnosing and categorising arrhythmias.
https://iopscience.iop.org/article/10.1088/1742-6596/2449/1/012033/meta

[7] A book reference of ECG records is a potential.
https://www.sciencedirect.com/science/article/pii/S2352906723000088
[8] A website reference of classifying arrhythmias from signals.
https://www.sciencedirect.com/science/article/pii/S0010482523000483
[9] An article reference of determination of the type of heartbeat.
https://www.mdpi.com/2077-0383/12/2/588
[10] A book reference of selecting the preprocessing approaches.
https://pubs.acs.org/doi/full/10.1021/acsomega.2c06610
[11] A website reference of heartbeat segmentation stage.
https://www.sciencedirect.com/science/article/abs/pii/S1746809422005602
[12] An article reference of the issue of ECG demarcation.
https://onlinelibrary.wiley.com/doi/full/10.1002/vms3.985
[13] A book reference of research on ECG-based arrhythmia.
https://www.sciencedirect.com/science/article/abs/pii/S0197457223000411
[14] A website reference of intra-patient and inter-patient.
https://link.springer.com/article/10.1007/s10994-015-5519-7
[15] An article reference of knowledge-based ECG interpretation techniques.
https://www.sciencedirect.com/science/article/abs/pii/S0950705113000609
[16] A book reference of the techniques for ECG signal analysis.
https://link.springer.com/chapter/10.1007/978-981-16-4222-7_42
[17] A website reference of a common diagnostic procedure for arrhythmia,
https://ieeexplore.ieee.org/abstract/document/5446331
[18] An article reference of A tiny device known as an ICD.
https://www.jacc.org/doi/abs/10.1016/j.jacc.2011.06.007
[19] A book reference of treat ventricular arrhythmias.
https://academic.oup.com/eurheartj/article/27/17/2099/2887329
[20] A website reference of cardiac ablation.
https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-8167.2010.01876.x
[21] An article reference of Atrial fibrillation.
https://www.bmj.com/content/311/7016/1361.short
[22] A book reference of Anti-arrhythmic drugs are used to treat.
https://www.sciencedirect.com/science/article/pii/0735109792900036
[23] A website reference of real-time cardiac rhythm monitoring.
https://aapm.onlinelibrary.wiley.com/doi/full/10.1002/acm2.13814
[24] An article reference of arrhythmias from ECG data.
https://www.sciencedirect.com/science/article/pii/S0010482523000483
[25] A book reference of decision-support system.
https://www.ceeol.com/search/article-detail?id=1038880

BIBILOGRAPHY

Nanduri Naveen is studying his 2nd year, Master of Computer Applications in Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, accredited by NAAC. With his interest in python, machine learning and as a part of academic project, he chooses Arrythmia classification using ecg data values . As a result of our analysis, we discovered interesting statistics that can help law enforcement agencies keep society safe. A completely developed project along with code has been submitted for Andhra University as an Academic Project. In completion of his MCA.

Potnuri Gayatri: assistant professor, she received her M Tech in Computer Science & engineering from JNTU Kakinada in January 2015. She received her B Tech Degree from VITAM College of Engineering, JNTU Hyderabad in 2004. she currently working as Assistant Professor, CSE Dept in SVPEC, Andhra Pradesh, India. Her Research Interests include Sensor Networks