

A Survey on various Machine Learning Approaches for ECG Analysis

C. K. Roopa
JSS Research Foundation
Mysuru, Karnataka, India

B. S. Harish
Department of Information
Science & Engineering
Sri Jayachamarajendra
College of Engineering,
Mysuru, Karnataka, India

ABSTRACT

Electrocardiogram (ECG) is a P, QRS and T wave demonstrating the electrical activity of the heart. Feature extraction and segmentation in ECG plays a significant role in diagnosing most of the cardiac disease. The main objective of this paper is to review the various machine learning approaches for diagnosing Myocardial Infarction (heart attack), differentiate Arrhythmias (heart beat variation), Hypertrophy (increase thickness of the heart muscle) and Enlargement of Heart. Further, we also present various machine learning approaches and compare different methods and results used to analyze the ECG. The existing methods are compared and contrasted based on qualitative and quantitative parameters viz., purpose of the work, algorithms adopted and results obtained.

General Terms

Machine Learning, Neural Networks, Fuzzy Systems, Rough Set Theory.

Keywords

Heart Disease, ECG Analysis, Feature Extraction, Segmentation.

1. INTRODUCTION

Diseases affecting heart have become common day by day. Diseases affecting heart is increasing because of present day unhealthy life style increase in medical conditions like diabetes, hypertension and tobacco smoking. Heart can be affected due to multiple conditions.

Most important is “Ischemic heart disease”. Ischemic heart disease occurs due to blockade in coronary arteries (Arteries supplying blood to heart), resulting in lack of nutrition to myocardial (heart muscle) cells. There are three major arteries supplying blood to heart. Blockade in one or more of these blood vessels, can cause Myocardial Infarction (Heart Attack). Consequence of this medical condition can be as grave as death. Diagnosing the above condition at the earliest and treating them immediately gives better scope for survival and quality of life. Next group of disease affecting heart is “Arrhythmia”. Arrhythmia is a group of disorder affecting “Conduction system” (heart beat variation). It can be in the form of slowing or fastening the heart beat. If not identified and treated at the earliest, both are dangerous and can lead to death. Other conditions which can affect the heart are Congenital Heart Disease (Disease by birth), Valvular Heart Disease (diseases affecting heart valves), Infections and Inflammation. Investigation which helps to diagnose or suspect these above conditions are through Electrocardiogram (ECG).

ECG is one of the simple, easily available, less expensive, easily doable, non invasive investigations available at all place including rural areas with minimal infrastructure. Doctors at the remote places with basic medical knowledge may not be well trained to interpret an ECG. Properly diagnosing cardiac (heart related) condition at the earliest is of utmost importance. During these medical conditions, time matters the most. Every minute delay in treating them, trail to more damage in heart muscle resulting in adverse outcome. Hence accurate identification and early diagnosing these heart diseases is very important. Prognosticating patients into low and high risk is very important. Low risk patients can be treated locally at the same hospital with minimal infrastructure itself. High risk patients require early referral to cardiac referral hospitals. ECG confirms or suspects the diagnosis of myocardial infarction, arrhythmias and other conditions. Once confirmed, they are treated with medicines or procedures (surgeries) depending on the patient subtype and the facilities available.

The ECG has been extensively used for diagnosing and prognosticating various cardiac conditions. Present day computerized automated report can diagnose these cardiac conditions and give an automated report on its own with an acceptable accuracy rate. It can diagnose Myocardial Infarction (heart attack), diagnose and differentiate arrhythmias (heart beat variation), diagnose hypertrophym z(increased thickness of the heart muscle), enlargement of heart, using the available algorithm. However the accuracy varies with algorithm used.

The figure1 below shows the basic representation of ECG signal. The Electrocardiogram is a graphical representation of heart muscle Electrical Activation (Depolarization) and Relaxation (Repolarisation). It is represented by various waves, segment and intervals. ‘P’ wave depicts about atrial depolarization (upper chamber of the heart). ‘PR’ interval depicts conduction of electrical impulse from upper chamber to lower chamber of the heart. ‘QRS’ wave depicts about depolarization of Ventricle (lower chamber of the heart). ‘ST’ segment and ‘T’ wave depicts about repolarisation of the Ventricle (lower chamber of the heart).

Predominant part of the ECG which talks about blockage is ‘ST segment’. Most commonly ECG is taken in 12 lead formats. There are many more leads, taken in specific circumstances. Different leads detect cardiac electrical activity from various directions in both horizontal and frontal axis.

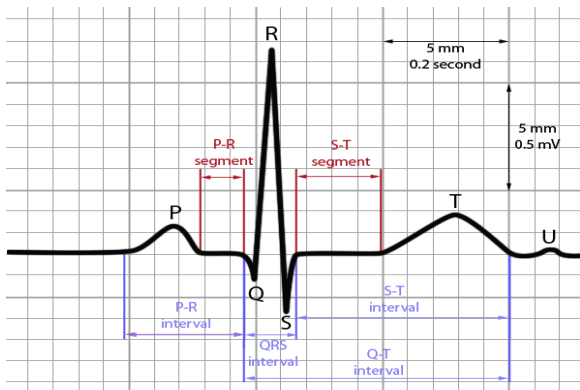


Figure 1: Schematic representation of normal ECG

ECG is essentially a time series signal that reflects the electrical activity of the heart. The signal consists of a series of repetitive and stereotyped complex waveforms with an obvious frequency of approximately 1 Hz. The heartbeat can vary across individuals and within individuals depending on a variety of conditions. The first stage in developing an automated machine learning techniques in ECG classification is to extract characteristic features from the waveform. Some features are considered first order, these features can be derived directly from the data such as RR time (time between the largest peaks that occurs in each heartbeat). Other features are derived from the base signal. These are extracted using signal processing techniques such as Fourier Transforms (FTs) and Wavelets. In a typical supervised classification scheme, the features are labeled with the decision outcome. The block diagram in figure 2 shows the different approaches used in automated machine learning techniques for investigating accurate results from ECG.

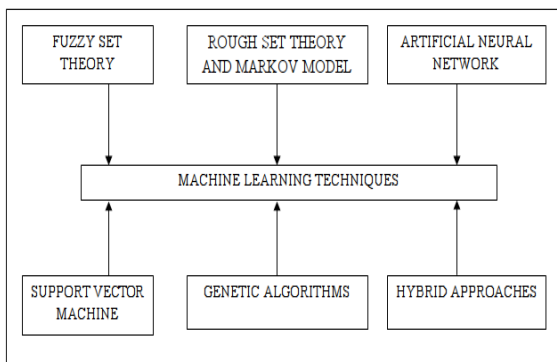


Figure2: Various Machine Learning Approaches used

Present day ECG machines use the above mentioned algorithms and approaches to diagnose various cardiac conditions. Future computerized ECG machines generate reports on its own with reasonably good accuracy.

2. LITERATURE SURVEY

The existing studies and various researchers have investigated the domain of Automated Computerized ECG for effective diagnosis of heart diseases. This review, therefore, is an attempt to critically explore the literature in the area of different algorithms and machine learning techniques in ECG analysis.

2.1 Fuzzy Based Techniques

Numerous methods have been developed to enforce computerized ECG analysis and diagnosis. Fuzzy sets or

fuzzy based machine learning is an approach that is commonly deployed for effective ECG analysis. These techniques utilize smooth variables with membership functions in diagnosing diseases using ECG analysis (Lei et al., 2007). Moreover, in modern medicines, large amount of data is processed pertaining to diseases such as coronary artery diseases, for which, different classifiers are used to analyze the ECG (Pal et al., 2012). For this purposes, Behadada & Chikh, (2013) conducted a study on assessing the cardiac abnormalities by optimizing the classification of cardiac arrhythmias and detection of abnormalities. The researchers used fuzzy decision tree as a methodology to diagnose the cardiac abnormalities. They used the data pertaining to heart beat from MIT-BIH database. The proposed fuzzy tree is able to detect the anomalies of premature ventricular contraction, where the false alarm rate is reduced to 19.76% and the classification rate is 71% (Behadada & Chikh, 2013). However when compared to neural networks and other techniques, the result is unsatisfactory. In alignment to this study, Lei et al., (2007) developed an adaptive fuzzy ECG classifier for enhancing the performance of conventional classifiers. The classifier used MIT-BIH arrhythmia database for evaluation. The developed classifier is effective in terms of self-adaptation as per the inputs of ECG signal. Average correct rate is found to be 88.2%. However, both these approaches lack in terms of comprehending the different types of heart beats. Also, the computational efficiency of the proposed approach (Lei et al., 2007) is significantly high.

ECG is widely used in the diagnosis of heart diseases, where fuzzy inference systems can be used for effective system analysis (Goel et al., 2016). The research on fuzzy based techniques have further developed through the incorporation of neural networks, such as fuzzy inference networks (Lei et al., 2007), that integrates the benefits of neural networks for better learning and fuzzy sets for better human understanding (Guler & Ubeyli, 2004). In a similar manner, Ceylan et al. (2009) used Fuzzy Clustering Neural Network Algorithm for classification in ECG analysis. The researchers aimed at increasing the accuracy rate for ECG beats, for which MIT-BIH databases of 92 patient records are extracted. Also, type 2 fuzzy mean clustering algorithms and back propagation 402 learning is used. The proposed system resulted an accuracy of 99% with acquisition of high clustering performance. It is inferred from the review, that the training time was large in this case, which limited the benefits. Also, for future research, a larger sample can be taken. Therefore, it can be inferred that the fuzzy sets are best used for demonstrating the data under uncertainties, for which the fuzzy logic is efficient in providing an inference mechanism. The output generated from fuzzy logic can be applied to knowledge based systems in ECG analysis (Jeyalakshmi & Robin, 2016).

The fuzzy decision tree provides a certain inferences and knowledge acquisition means to deal with multi-varied data, with complex or missing values (Chen et al., 2009). For the above purposes, it is widely used as interference and expert systems to provide clarity. In medical science stream, Pal et al. (2012) developed a fuzzy expert system for accurate detection of coronary artery diseases and removal from uncertainties emerging from the medical data. The researchers emphasized on the rule organization in tree representation for obtaining efficient search from the database pertaining to Coronary Artery Diseases (CAD). Therefore, it is inferred that fuzzy based techniques are effective in computerized ECG analysis but requires further investigation for attaining higher level of accuracy, especially pertaining to the localization

of culprit coronary artery in ST Elevation Myocardial Infarction (STEMI).

2.2 Rough Set Theory and Hidden Markov Models

Rough Set Theory (RST) came into existence in the early 1980s and was evolved as a strategy to be used for reducing uncertainties and ambiguity from data (Senthilkumaran & Rajesh, 2009). The problems resolved by RST are associated with the reduction of redundant data, detection of data dependencies, and classification of data and assessment of data patterns (Deja & Paszek 2003). Moreover, the advantage of using RST in ECG analysis is that it assists in formulating rules, which are easily understood and leads to accurate extraction of valuable information from complex databases (Tripathy et al., 2011). Rough set theory is used to generate a robust classification model (Barman et al., 2016). Several investigators in the existing literature have utilized rough set theory in computerized ECG analysis in hopes of attaining maximum benefits. One such research conducted by Mitra et al., (2006) demonstrated an optimized approach of rule based rough set for effective diagnosis. A TIFF formatted image database was utilized in the study. The researchers have used differentiation techniques and syntactic approaches, where an automated ECG feature extraction system is developed, noises are removed and knowledge base regarding ECG interpretation is developed. A rule based rough set decision system is developed for the detection of diseases. 100% accuracy is obtained for Myocardial Infarction (MI). The results are appropriate and efficient for trained set but lacks in accuracy pertaining to untrained data set. Further, the sample size taken for the study was just 27 patient records of MI.

Barman et al., (2016) conducted a study to acquire accurate and correct information from MIT-BIT arrhythmia database by using rough set theory and different classification schemes. The data is first pre-processed using MATLAB tool and then RST is applied for adequate classification. The researchers have also used filtering methods for pre-processing. Another work reported by, Setiawan et al., (2009) developed a rule selection method for diagnosing Coronary Artery Disease (CAD). Total samples of 920 patients were collected from different hospitals, which have CAD. Rough Set Theory (RST) is applied by using hybrid approach of support filtering and RST rule importance selection. The proposed technique is efficient at selecting and extracting rules from large number of rules, without reducing the quality of classification. Fuzzy reasoning can be incorporated in the proposed mechanism for future work, so as to increase the accuracy.

A study by Huang & Zhang (2003) revealed that the recognition rules for determining the required data in ECG are condensed with the application of RST. On the same line, Tripathy et al., (2011) conducted an investigation by aiming at identifying characteristics such as ECG, Blood Pressure and Maximum Heart Rate (MHR) that affect heart diseases. Testing data of 250 patients was taken, which was divided into 5 types according to the different heart diseases. Rough sets, decision rules and Formal Concept Analysis (FCA) are used in the study. High accuracy is obtained for determining the characteristic of heart diseases.

The ECG signal is assessed for knowledge acquisition by evaluating the ECG waveform, which can be done with the help of Hidden Markov models (HMMs). An HMM is widely utilized in the application of segmenting the ECG signal into waveforms (Frénay et al., 2009). A study by Andreao et al., (2006) investigated the utilization of HMM in classifying

ECG while performing the functionality of heart beat detection and segmentation. The researchers have addressed the several aspects such as waveform modeling, classification, unsupervised adaptation and segmentation. The study was concluded with effective and accurate detection of heart beats with 99.79% sensitivity.

2.3 Approaches based on Neural Networks

In recent times, several methods have been utilized for effective classification of arrhythmia such as Genetic algorithms, Fuzzy Logic, Self-Organizing Map, Bayesian, Hidden Markov Models and SVMs (Gupta & Chatur, 2012). Considering the study of ECG signal patterns in diagnosis, it is very extensive and time-consuming, thus requiring the computerized means for effective diagnosis of cardiac diseases (Sao et al., 2015). Neural networks are used for classification after pre-processing, detection and feature extraction from the ECG signal. Gao et al., (2004) developed a diagnostic system by using an artificial neural network classifier for cardiac arrhythmias. The classifier is based on a Bayesian framework, which is formulated with the application of back propagation algorithm. The proposed system offers a predication accuracy of 90%, which has further scope for improvement. Also, the system is not well applicable in the practical scenario and requires amendments. In a similar manner, Vishwa et al., (2011) developed an automated Artificial Neural Network (ANN) based classification system for cardiac arrhythmia using multi-channel ECG recordings. An MIT-BIH arrhythmia database consisting of 48 recordings from 1975 and 1979 is used. Also, a Normal Sinus Rhythm (NSR) database is used for the validation of the method. The estimated feed forward ANN and back-propagation learning algorithms are used. The approach shows the accuracy of 96.77% on MIT-BIH database and 96.21% on NSR database. However, the research is unable to classify all 16 arrhythmia classes. The study was concluded with the future scope of pre-processing of ECG signal by developing Multilayer Perceptron (MLP) model. Sao et al., (2015) also used ANN in ECG for diagnosing heart diseases in the most efficient manner. The researchers developed an ANN classifier for the same, wherein the parameters such as Spectral Entropy, Largest Lyapunov exponent and Poincare plot geometry are assessed along with the utilization of back propagation algorithm.

According to Anuradha and Reddy (2008), neural networks can be widely used for optimizing the classification of ECG signals. The authors developed an ANN-based classifier for cardiac arrhythmia. MIT-BIH dataset has been used to retrieve ECG data. Artificial neural network (ANN) is used for the classification of cardiac arrhythmias from ECG signals. The classifier uses a combination of wavelets Back Propagation Algorithms (BPA). An accuracy of 90.56% is achieved and the classifier efficiently detects the abnormal activities. Though the classifier is easy to implement, it lacks in terms of proper validation against different dataset. Another research by Jadhav et al., (2012) aimed at developing an ANN model for diagnosing heart diseases. For this purpose, the researchers used 12 lead ECG signal recordings data from UCI machine-learning repository datasets. Static back propagation algorithm and momentum learning rule are used for classification. Mean Squared Error (MSE), Receiver Operating Characteristics (ROC) and Area Under Curve (AUC) are the measures used to evaluate the performance of the proposed model. The classification accuracy is obtained to be 86.67% and the sensitivity is 93.75%. The developed model is well validated by using different performance

measures, but can be further improved for practical applications.

According to Gupta and Chatur (2012), ECG analysis is highly beneficial for diagnosing the heart diseases. The researchers have applied supervised artificial neural networks and data mining techniques for ECG classification. The application of both the techniques is then compared, where the use of ANN is rendered as more effective. Another research by De Gaetano et al., (2009) developed an algorithm based on supervised neural network to differentiate between the normal beats in ECG signals and ischemic beats. The researchers have used radial basis functions and R-wave segments (RRR interval) for identifying the two types of beats. Later the system is validated on European ST-T database. The proposed classifier is evaluated as reliable when identifying the ischemic beats from ECG data.

2.4 Approaches using SVM

Support vector classifiers, based on statistical learning theory, are widely used in several applications such as image segmentation, ECG signal classification and diagnosis, brain computer interfacing etc (Guler & Ubeyli, 2007). In a study by Kampouraki et al., (2009) the time series of heart beat are classified with the implementation of Support Vector Machines (SVM). The researchers have revealed that the SVM classifier offers a better functionality than those based on neural networks as it classifies the ECG signals with minimal signal to noise ratio. In a similar manner, Polat et al., (2008) developed a technique for classifying arrhythmia from ECG dataset. ECG dataset of healthy persons and diseased persons (arrhythmia) is used in the study. Least square support vector machine (LSSVM) is deployed by the authors as a robust technique for accurate classification. The technique is validated through classification accuracy, sensitivity and ROC curves. On evaluation, the proposed technique achieved 100% accuracy. However, the accuracy obtained is only applicable to atherosclerosis disease and for others, the accuracy was very less. Batra and Jawa (2016) aimed at optimizing the accuracy of detecting Arrhythmia by using ECG. The dataset (452 records) was retrieved from UCI Machine Learning Repository. Information pertaining to Arrhythmia was extracted. The researchers used machine learning algorithms and ECG diagnostic criteria [13]. The machine learning algorithms like neural networks, decision trees, random forest, gradient boosting and support vector machines were used for experiments. However more emphasize was laid upon SVM. With the combination of gradient boost and SVM, an accuracy of 83.04% was obtained. After applying ECG criteria, the accuracy was increased to 84.82%. Though, several techniques have been determined in this research, the combination of only gradient boosting and SVM has been shown. More combinations can also be thought in future.

SVM has been largely used with the integration of other methods to acquire efficiency in ECG classification. For example, Wang and Chiang (2007) used a new approach based on SVM and particle swarm optimization for effective classification and diagnosis of cardiac arrhythmias. On other hand, Nasiri et al., (2009) proposed a new approach for feature extraction and ECG classification for cardiac arrhythmia disease. The SVM classifier is used to optimize the approach by detecting the best subset of feature, which provides the optimal classification of the signal. The researchers have utilized SVM with Genetic Algorithm (GA), where GA is deployed to enhance the generalization

performance of the SVM classifier. The technique proposed by these researchers achieved an accuracy of 93%.

2.5 Approaches using Genetic Algorithms

ECG analysis has been widely deployed by using Genetic Algorithms (GA) to achieve higher computational capabilities with in less time-consumption manner (Omran et al., 2009). However this approach is integrated with others to exploit maximum benefits. Several researchers have used GA with ANN, SVM, and other techniques. On the other hand, the techniques pertaining to the use of GA alone in ECG analysis is scarce. Priyadharshini and Kumar (2015) conducted a research to enhance the classification of Arrhythmia from ECG signals. For this purpose, MIT-BIH Arrhythmia Database is utilized. An Improved Genetic Algorithm (GA) is developed and C4.5 and Naïve Bayes classification algorithm is used in the research. The research revealed that the GA has the most stable accuracy as compared to C4.5 and Naïve Bayes. The scope of the research can be extended to other machine learning algorithm such as SVM, to present the most efficient method.

Omran et al., (2009) used GA with ANN for ECG analysis and diagnosis of heart diseases. The ANN is used to detect the connecting points in ECG signal and GA is used to choose the best values to those points. The genetic algorithms are best utilized for minimizing the error values and improving the connecting points. In this study, after applying GA in the ANN model of detection and classification, the error is reduced to 2.3116e-29 from 5.9487e-28 in 100 generations. El-Dahshan (2010) also utilized genetic algorithm with another approach for gaining higher efficiency in the ECG analysis. The researcher incorporated GA with wavelet transform (WT) for denoising ECG signals, wherein GA is used to determine the best parameters for adequate wavelet denoising. Percentage root mean square and SNR is used to comprehend the performance of the approach.

When considering the use of GA alone, Silveira et al., (2007) developed a compression method using genetic algorithm for minimizing distortion of ECG signal in its analysis. The classical wavelet Db3 is used to compare the performance of the proposed technique. MIT-BIH compression database is used in the experiment. The compression with GA ensures lower values than the classical wavelet Db3 method.

2.6 Hybrid Algorithms

Hybrid techniques for analysis of ECG signals have been extensively investigated in search of improved performance. Dalal and Birok (2016) developed a hybrid classifier by using the Principal Component Analysis (PCA) and Neuro-Fuzzy classifier. Different heart activities are detected with the use of hybrid approach, wherein MIT-BIH database is used. The PCA approach is used to retrieve the required information from the database and then ANN and Fuzzy logic based classifier is deployed to analyze and process the extracted information. The accuracy of the hybrid approach is estimated to be 96%. Another work by, Bensujin et al., (2014) examined and determined the ST Segment Elevation Myocardial Infarction (STEMI) in the ECG signal of a person. 480 patients' records were extracted from MIT BIH database, which are segregated in two sets. Chinese Heart Disease database is used for validation of the method. Hybrid detection methods were used for detection of heart rate, ischemic and extraction of STEMI. K-means clustering algorithm is used for feature extraction. Bacterial Foraging Optimization Algorithm (BFOA) is used for identifying STEMI from ECG signals. The analysis revealed that the

STEMI on the clustered data is identified by 89% and the convergence rate is 91%. The efficiency of the proposed methodology is high.

For gaining higher accuracy, Finlay et al., (2015) determined the impact of electrode placement on the accuracy of identifying coronary artery occlusion. 45 patients were taken, who were undergoing Percutaneous Transluminal Coronary Angioplasty (PTCA) and had 120 lead Body Surface Potential Maps (BSPMs) recording. BSPMs were used to examine the impact of electrode placement on the patients. A computer algorithm is developed for the detection of STEMI and non-STEMI. The researchers established that the lead electrode placement has significant impact on detecting STEMI and 12-lead ECG. However, it is observed that the practical application of the proposed methodology is poor. Also, the number of patients was less due to which the results cannot be relied upon.

Ghosh et al., (2013) examined the accuracy of ECG localization of culprit vessel occlusion in acute STEMI. 21 patients with acute myocardial infarction were selected in the study. The localization of culprit vessel occlusion site in Left Anterior Descending (LAD), Right Coronary Artery (RCA)

and left Circumflex (LCx) coronary arteries were performed by dividing the patients into 3 groups. The results revealed that for LAD, the accuracy of ECG localization was 90.91%; for RCA, it was 100%; and for LCx coronary arteries it was 90%. However, the sample size taken in the study was small. Also, the angiography was not performed immediately, which might have resulted in inconclusive results.

3. CONCLUSION

The present review has addressed different machine learning techniques that are deployed for ECG analysis in classification of heart diseases. The techniques and algorithms that were extensively used are artificial neural networks, support vector machines, fuzzy logic, decision trees, rough set theory, genetic and hybrid algorithms. Though the existing studies are vast, this spectrum of research requires more rigorous investigation in terms of practical application of the classification techniques. Further, the literature review reveals that there are few studies that have addressed the culprit coronary artery in STEMI and its localization. Hence there is a need to address this gap in literature and provide a valuable contribution to this field.

Author (year)	Dataset	Purpose	Methods	Result	Remarks
Fuzzy Based Machine Learning Techniques					
Behadada et al., (2013)	MIT-BIH	To optimize the classification of cardiac arrhythmias and detection of abnormalities.	Fuzzy Decision Tree.	The False Alarm Rate is 19.76%. Classification Rate is 71%.	Feature selection is not accurate.
Lei et al., (2007)	MIT-BIH arrhythmia.	To develop an adaptive fuzzy ECG classifier for enhancing the performance of conventional classifiers.	Fuzzy Sets	Average correct rate = 88.2%.	Computational efficiency is poor
Ceylan et al., (2009)	92 patient records from MIT-BIH database	To increase the accuracy rate for ECG beats.	Fuzzy Clustering Neural Network Algorithm, Type 2 fuzzy means clustering algorithm and back propagation 402 learning	Accuracy = 99%.	Less samples considered
Pal et al., (2012)	Coronary Artery Diseases (CAD) patient database	To develop screening expert system for detecting CAD	Fuzzy Tree and Rule Organization	Sensitivity = 95.85% Specificity = 83.33%	Methodology is not clear.
Rough Set Theory and Hidden Markov Model Based Machine Learning Techniques					
Barman et al., (2016)	MIT-BIT arrhythmia database	To increase accuracy in ECG analysis.	Rough Set Theory and different classification and Filtering Schemes	Proposed mechanism is effective when compared to fuzzy classifiers.	Results are not well presented in terms of accuracy evaluation.
Mitra et al., (2006)	TIFF formatted image database	To develop an optimized approach of rule based rough set for effective diagnosis.	Differentiation techniques and syntactic approaches.	100% Accuracy for Myocardial Infarction (MI).	Less accuracy for untrained sample.
Setiawan et al.,	920 patients' data	To develop a rule	Rough Set Theory	Proposed technique	Accuracy of

(2009)	from different hospitals, having Coronary Artery Disease	selection method for diagnosing CAD.	(RST) with support filtering.	is efficient at selecting and extracting rules (27) from large number of rules (3881).	classification is increased.
Tripathy et al., (2011)	Testing data of 250 heart patients	To identify characteristic such as ECG, Blood Pressure and MHR that affect heart diseases.	Rough Sets, Decision Rules and Formal Concept Analysis (FCA).	High accuracy is obtained for determining the characteristic of heart diseases.	Validation can be further improved.
Andreao et al., (2006)	Two-channel QT database	To increase heart beat detection and segmentation in ECG analysis	Hidden Markov Models	Sensitivity = 99.79%	High Accuracy
Neural Networks Based Machine Learning Techniques					
Anuradha et al., (2008)	MIT-BIH dataset	To develop an ANN-based classifier for cardiac arrhythmia.	Artificial Neural Network (ANN), Wavelets Back Propagation Algorithm	Accuracy =90.56%	Poor validation of methodology
Jadhav et al., (2012)	UCI machine-learning repository datasets	To develop an ANN model for diagnosing heart diseases.	Static Back Propagation algorithm, Momentum learning Rule	Accuracy =86.67% Sensitivity =93.75%.	Only binary classification
Vishwa et al., (2011)	MIT-BIH arrhythmia database consisting of 48 recordings from 1975 and 1979	To develop an automated artificial neural network (ANN) based classification system for cardiac arrhythmia using multi-channel ECG recordings.	Estimated Feed Forward ANN and Back-Propagation learning algorithms	Accuracy =96.77% on MIT-BIH database Accuracy =96.21% on NSR database.	Unable to classify all 16 arrhythmia classes.
De Gaetano et al., (2009)	European ST-T database	To differentiate between the normal beats in ECG signals and ischemic beats	Radial basis functions and R-wave segments (RRR interval)	Provides high reliability.	Lack in Accuracy
Gupta et al., (2012)	Raw ECG signal	To improve ECG signal analysis	Artificial Neural Networks and Data Mining techniques	ANN approach is more productive than data mining.	Classification of only 4 types of heart beats
Support Vector Based Machine Learning Techniques					
Polat et al., (2008)	ECG Arrhythmia dataset	To develop a technique for classifying arrhythmia from ECG dataset.	Least Square Support Vector Machine (LSSVM)	Accuracy =100%	Accuracy obtained is only for atherosclerosis disease.
Batra et al., (2016)	452 records from UCI Machine Learning Repository.	To optimize the accuracy of detecting Arrhythmia by using ECG.	SVM, Neural Networks, Decision Trees, Random Forest, Gradient Boosting. SVM was majorly emphasized.	With SVM, accuracy = 83.04% After applying ECG criteria, the accuracy = 84.82%.	Combined only gradient boosting with SVM.
Nasiri et al., (2009)	MIT-BIH arrhythmia database of 48 records	To improve feature extraction and ECG classification for cardiac arrhythmia disease.	SVM and Genetic Algorithm.	Accuracy = 93%	Accuracy can be further improved.

Genetic Algorithm Based Machine Learning Techniques					
Omran et al., (2009)	ECG Arrhythmia Database	To improve ECG classification and diagnosis	Genetic Algorithm with ANN	Error rate is reduced to 2.3116e-29 from 5.9487e-28 in 100 generations	Higher utilization is yet to be improved.
Priyadarshni et al., (2015)	MIT-BIH Arrhythmia Database	To enhance the classification of Arrhythmia from ECG signals.	Improved Genetic algorithm, C4.5 and Naïve Bayes classifier	Accuracy using GA= 90%, C4.5 = 72% and Naïve Bayes =50%	Genetic Algorithm have the most stable accuracy rate than Naïve bayes and C4.5 algorithm
Dahshan et al., (2010)	MIT-BIH database	Denosing ECG signals effectively	GA with Wavelet Transform (WT)	The approach is successful in eliminating noise from ECG signal	Results are well demonstrated
Hybrid Algorithm Based Machine Learning Techniques					
Finlay et al., (2015)	45 patients undergoing angioplasty and had 120 lead BSPMs recording.	To determine the impact of electrode placement on the accuracy of identifying coronary artery occlusion.	Body Surface Potential Maps (BSPMs) and computerized Diagnostic Algorithm.	Sensitivity = 51.1% Specificity = 91.1%.	Poor practical applications.
Ghosh et al., (2013)	21 patients with acute myocardial infarction	To examine the accuracy of ECG localization of culprit vessel occlusion in acute STEMI	Localization of culprit vessel occlusion site in LAD, RCA and LCx coronary arteries were performed by dividing the patients into 3 groups.	Accuracy for LAD =90.91%, RCA =100%; and for LCx coronary arteries = 90%.	Sample size was small and Coronary Angiography was not done immediately
Dalal et al., (2016)	MIT-BIH database	To improve diagnosis via ECG analysis	PCA and Neuro-fuzzy classifier	Accuracy is 96%	The sample size is small.
Bensujin et al., (2014)	480 patients' records from MIT BIH database	To examine and determine the STEMI in the ECG signal of a person	K-means Clustering Algorithm, Bacterial Foraging Optimization Algorithm (BFOA)	Identification rate = 89% and the convergence rate = 91%	Pre-processing of data is not efficient

4. REFERENCES

- [1] Andreao, R., Dorizzi, B., and Boudy, J. (2006). ECG signal analysis through hidden Markov models. *IEEE Transactions on Biomedical Engineering*, vol.53, no.8, pp.1541-1549.
- [2] Anuradha, B. and Reddy, V. (2008). ANN classification of cardiac arrhythmias, *ARPN Journal of Engineering and Applied Sciences*, vol.3, no.3, pp.1-6.
- [3] Barman, T., Ghongade., and Ratnaparkhi, A. (2016). Rough set based segmentation and classification model for ECG. *IEEE Conference on Advances in Signal Processing (CASP)*, pp.18-23.
- [4] Batra, A. and Jawa, V. (2016). Classification of Arrhythmia using Conjunction of Machine Learning Algorithms and ECG Diagnostic Criteria, *International Journal of Biology and Biomedicine*, vol.1.1, pp.1-7.
- [5] Behadada, O. and Chikh, M.A. (2013). An interpretable classifier for detection of cardiac arrhythmias by using the fuzzy decision tree, *Artificial Intelligence Research*, vol.2, no.3, pp.45-58.
- [6] Bensujin., Vijila, C.K. and Hubert, C. (2014). Detection of ST Segment Elevation Myocardial Infarction (STEMI) Using Bacterial Foraging Optimization Technique, *International Journal of Engineering and Technology (IJET)*, vol.6, no.2, pp.1212-1224.
- [7] Ceylan, R., Ozbay, Y. and Karlik, B. (2009). A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network, *Expert Systems with Applications*, vol.36, no.3, pp.6721-6726.
- [8] Chen, Y., Wang, T., Wang, B., and Li, Z. (2009). A survey of fuzzy decision tree classifier. *Fuzzy Information and Engineering*, vol.1, no.2, pp.149-159.
- [9] Dalal, S., and Birok, R. (2016). Analysis of ECG Signals using Hybrid Classifier. *International Advanced Research Journal in Science, Engineering and Technology*, vol.3 no.7, pp. 89-95.

- [10] De Gaetano, A., Panunzi, S., Rinaldi, F., Risi, A., and Sciandrone, M. (2009). A patient adaptable ECG beat classifier based on neural networks. *Applied Mathematics and Computation*, vol.213, no.1, pp.243-249.
- [11] Deja, A. W., and Paszek, P. (2003). Applying rough set theory to Multi stage medical diagnosing. *Fundamental Informaticae*, vol.54, no.4, pp.387-408.
- [12] El-Dahshan, E. (2010). Genetic algorithm and wavelet hybrid scheme for ECG signal denoising. *Telecommunication Systems*, vol.46, no.3, pp.209-215.
- [13] D. Da Costa, W. J. Brady, and J. Edhouse. (2002) "Abc of clinical electrocardiography: Bradycardias and atrioventricular conduction block," *BMJ: British Medical Journal*, vol. 324, no.7336, pp.535.
- [14] Finlay, D., Bond, R., Kennedy, A., Guldenring, D., Moran, K. and McLaughlin, J. (2015). The effects of electrode placement on an automated algorithm for detecting ST segment changes on the 12-lead ECG, *Computing in Cardiology*, vol. 42, pp.1161-1164.
- [15] Fréney, B., Lannoy, G., and Verleysen, M. (2009). Improving the transition modelling in hidden Markov models for ECG segmentation. *ESANN proceedings, European Symposium on Artificial Neural Networks - Advances in Computational Intelligence and Learning, Bruges (Belgium)*, pp.22-24.
- [16] Gao, D., Madden, M., Schukat, M., Chambers, D., and Lyons, G. (2004). Arrhythmia Identification from ECG Signals with a Neural Network Classifier Based on a Bayesian Framework. *Twenty-fourth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, vol.3, no.3, pp.390-409.
- [17] Ghosh, B., Indurkar, M., and Jain, M.K. (2013). ECG: A Simple Noninvasive Tool to Localize Culprit Vessel Occlusion Site in Acute STEMI, *Indian Journal of Clinical Practice*, vol.23, no.10, pp.590-595.
- [18] Goel, S., Tomar, P., and Kaur, G. (2016). A Fuzzy Based Approach for Denoising of ECG Signal using Wavelet Transform. *International Journal of Bio-Science and Bio-Technology*, vol.8, no.2, pp. 143-156.
- [19] Guler, I., and Ubeyli, E. D. (2007). Multiclass support vector machines for EEG signals classification. *IEEE Transactions on Information Technology in Biomedicine.*, vol.11, no.2, pp. 117–126.
- [20] Guler, J., Ubeyli, E.D. (2004). Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction. *Expert Systems with Applications*, vol.37, no. 27, pp. 323-330.
- [21] Gupta, K. O., and Chatur, P. N. (2012). ECG Signal Analysis and Classification using Data Mining and Artificial Neural Networks. *International Journal of Emerging Technology and Advanced Engineering*, vol.2, no.1, pp. 56-60.
- [22] Huang, X. M., and Zhang, Y. H. (2003). A new application of rough set to ECG recognition. In *Int. Conference on Machine Learning and Cybernetics*, no.3, pp.1729-1734.
- [23] Jadhav, S., Nalbalwar, S.L. and Ghatol, A. (2012). Artificial Neural Network Models based Cardiac Arrhythmia Disease Diagnosis from ECG Signal Data, *International Journal of Computer Applications*, vol.44, no.15, pp. 8-13.
- [24] Jeyalakshmi, M.S, and Robin, C. (2016). Fuzzy based Expert system for sleep apnea diagnosis. *International Journal of Engineering Trends and Technology*, vol.35, no.12, pp. 555-558.
- [25] Kampouraki, A., Manis, G., and Nikou, C. (2009). Heartbeat Time Series Classification with Support Vector Machines. *IEEE Transactions on Information Technology in Biomedicine*, vol.1, no.4, pp. 512-518.
- [26] Lei, W. K., Li, B. N., Dong, M. C. and Vai, M. I. (2007). AFC-ECG: An Intelligent Fuzzy ECG Classifier. In book: *Soft Computing in Industrial Applications*, vol.39, pp.189-199
- [27] Mitra, S., Mitra, M. and Chaudhuri, B.B. (2006). An Approach to a Rough Set Based Disease Inference Engine for ECG Classification, *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 2198-2206.
- [28] Nasiri, J. A., Naghibzadeh, M., Yazdi, H. S., and Naghibzadeh, B. (2009). ECG Arrhythmia Classification with Support Vector Machines and Genetic Algorithm. *Third UKSim European Symposium on Computer Modeling and Simulation*.
- [29] Omran, S. S., Taha, S. M. R., and Awadh, N. A. (2009). ECG Rhythm Analysis by Using Neuro-Genetic Algorithms. *Journal of Basic and Applied Sciences*, vol.1, no.3, pp.522-530.
- [30] Pal, D., Mandana, K., Pal, S., Sarkar, D., and Chakraborty, C. (2012). Fuzzy expert system approach for coronary artery disease screening using clinical parameters. *Knowledge-Based Systems*, vol.36, pp.162-174.
- [31] Polat, K., Akdemir, B., and Gune, S. (2008). Computer aided diagnosis of ECG data on the least square support vector machine, *Digital Signal Processing*, vol.18, pp.25–32.
- [32] Priyadharshini, V. and Kumar, S.S. (2015). An Enhanced Approach on ECG Data Analysis using Improved Genetic Algorithm, *International Research Journal of Engineering and Technology (IRJET)*, vol.2, no.5, pp.1248-1256.
- [33] Sao, P., Hegadi, R., and Karmakar, S. (2015). ECG Signal Analysis Using Artificial Neural Network. *International Journal of Science and Research*, pp.82-86.
- [34] Senthilkumaran, N., and Rajesh, R. (2009). A Study on Rough Set Theory for Medical Image Segmentation. *International Journal of Recent Trends in Engineering*, vol.2, no.2, pp. 236-238.
- [35] Setiawan, N.A., Venkatachalam, P.A. and Fadzil, A. (2009). Rule Selection for Coronary Artery Disease Diagnosis Based on Rough Set, *International Journal of Recent Trends in Engineering*, vol.2, no.5, pp. 198-202.
- [36] Silveira, R. M., Agulhari, C. M., Bonatti, I. S., and Peres, P. D. L. (2007). A genetic algorithm to compress the electrocardiograms using parametrized wavelets. *IEEE*

International Symposium on Signal Processing and Information Technology.

- [37] Tripathy, B.K., Acharjya, D.P. and Cynthya, V. (2011). A Framework for Intelligent Medical Diagnosis Using Rough Set with Formal Concept Analysis, *International Journal of Artificial Intelligence & Applications (IJAA)*, vol.2, no. 2, pp.45-66.
- [38] Vishwa, A., Lal, M., Dixit, S., and Vardwaj, P. (2011). Classification of Arrhythmic ECG Data Using Machine Learning Techniques. *International Journal of Interactive Multimedia and Artificial Intelligence*, vol.1, no.4, pp 67-70.
- [39] Wang, T.Y., Chiang, H.M. (2007). Fuzzy support vector machine for multiclass categorization. *Information Process Management*, vol. 43, no.4, pp.914–929.