IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 08, Issue 4 (April. 2018), ||VI|| PP 48-53

A Study on Various Machine Learning Techniques For ECG Signal Analysis

Sunita Kumari[1], Dr Lini Methew[2]
[1]*M.E. Student, Department of Electrical Engineering, NITTTR Chandigarh (India)*
[2]*Associate Professor Head, Electrical Engineering, NITTTR Chandigarh (India)*

Abstract: ECG signal analysis is very essential for the diagnosis of most of the cardiac diseases. ECG is a recording method of electrical impulses which are generated in the heart. The useful information about the functionality of the human heart is provided by the ECG interpretation. While the ECG is a nonlinear or non-stationary signal, the slight changes in its amplitude and duration are not well explained in time and frequency domains. The intervals and amplitudes of the ECG waves describe the different features required for the ECG signal analysis such as statistical feature, morphological feature, and temporal features etc. The P-QRS-T waves in ECG signal represent one cardiac cycle and the normal heartbeat ranges between 60 to100 beats per minute. The signal processing techniques are an obvious choice to extract the valuable information by using ECG signal for real-time analysis. Whereas, traditional techniques for signal processing are unable to deal with the non-stationary nature of the bio-signals. Further, these extracted features are applied to the classifiers for classification in different categories of cardiac disease. In this proposed paper different techniques are discussed which are proposed earlier for extracting useful features for the analysis of an arrhythmia and interpretation of ECG signals over classical processing technique with different classifiers. This paper also provides a comparison of various methods proposed earlier for classification and feature extraction.

Index Terms: ECG, feature extraction, FFT, STFFT, DWT, PCA, LDA, classification, SVM, ANN etc.

Date of Submission: 22-03-2018 Date of acceptance: 07-04-2018

I. INTRODUCTION

Cardiovascular disease (CVD) is one of the most general causes of death across the world. Approximately 40% of deaths in worldwide and 50% deaths in the high-income class of developed countries are due to cardiovascular disease (CVD) [1][2]. Now the CVD rates are decreases in developed countries, whereas the rates are increasing in every other part of the world, as its mentions in the survey of WHO. [1]

- three-fourth of the world's deaths from CVDs occur in low and middle-income countries or in developing countries.
- As compared to high-income people of developed countries, the low and middle income people of the developing countries do not have the advantage of primary health programs for early detection and cure from the risk of cardiovascular disease (CVD).
- Many people suffer from non-communicable and cardiovascular disease (CVDs); have less right to use the effective and reasonable health care services which respond to their needs. Results of it many people die younger because of late detection of cardiovascular and non-communicable disease.
- In the developing countries, the poorest of the people are most affected by this cardiovascular and non-transitive diseases, there is enough evidence to prove that CVD and other non-transitive diseases contribute to poverty due to fatal health costs and high out-pocket expenses in health care equipments.
- At the macro level, CVD puts heavy burden on economies of developing countries.

Just like a brain, the heart of the human is one of the essential organs of the body. So, the structure and functions of it have become a great source for the research. Even though a lot of research has been done by many researchers in this field, it is still a challenging area. Moreover, ECG is an important guide for the investigation of the physiological situation of the subject.

Like any bio-signals, ECG signals exhibit non-stationary, noise susceptibility, and individual variability characteristics. The special and undesirable aspect of ECG is that they are always affected by noise e.g. low-frequency baseline wander, muscle interferences, power-line electromagnetic interferences, and electromagnetic interferences of impulse character. Still, precise and complete ECG feature extraction is challenging in the field of medicine.

The physiological and pathological condition of any subject is easily found by decoding or extracting the information or required features from ECG signals. Therefore the ECG signal analysis has been found very effective for identify the disease. The feature extraction process can be completed by analyzing the ECG signal visually on paper or screen of the monitor. Moreover, feature extracted by manually are always prone to error. That's why ECG signal processing has become a advanced and successful tool for extracting clinically significant information from the ECG signals. The dependibility on manually collected data of ECG signal is reduced and this will help the physician in making well-founded decisions. ECG signal analysis systems are usually considered to process ECG signals which are measured under particular conditions, like resting ECG interpretation, stress test analysis, ambulatory ECG monitoring and intensive care monitoring of ICU.

The electrical activity of heart in ECG signal is represented by electrical waves followed by peaks and valleys, as shown in Fig. 1.

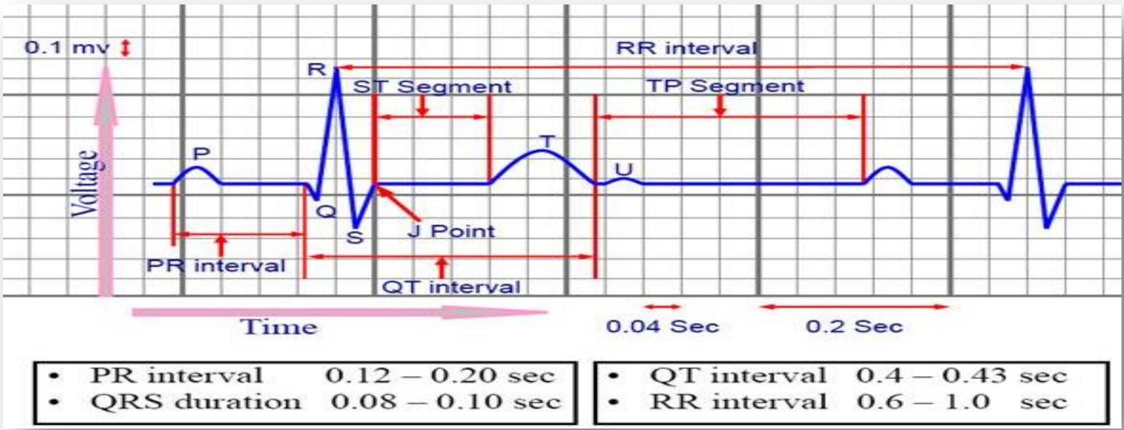

Fig. 1 ECG curve

ECG provides two types of information. One is the time duration of electrical impulse passing through the heart which will decide the regular or irregularity of the electrical activity in the heart . Second is the working of heart muscles which are effected by electrical activity passes through it, this electrical activity helps in finding whether that part of heart muscles activity is too large or it overworked? The heart is divided into four chambers in which the Depolarization and repolarization occurs i.e. the depolarization and repolarization in right and left atria first and then in right and left ventricles. The electrical impulse generated by sinoatrial node (SA) also known as natural pacemaker of the heart is then travelled through the atria's to ventricle causes contraction. The P-wave shown in the Fig. 1 represented the right and left atria depolarization. Similarly, the QRS complex gives the result of the depolarization of the ventricles and repolarization of the atria. The repolarization and depolarization both occur almost at the same time. The ventricular repolarization is represented by the T-wave, whereas the U-wave if it present is normally believed to be the result of afterpotentials in the ventricular muscle. After the U-wave, the heart muscles rest for a mili-seconds. The precious information on the present state of the heart is provided by the shape, duration and time of occurrence. For that reason morphological and temporal feature of ECG is extracted using different techniques.

NEED OF BIO-SIGNAL ANALYSIS

Bio signals means a collective electrical signal acquired from any organ that represents a physical variable of interest. This signals are normally a fuction of time and can be represented in term of its amplitudes and frequency.

In the past fifteen years, A large part of the research has been paying attention on the processing of bio signals. In daily clinical practice, huge amount of bio-signals are generated for the monitoring of patients and for diagnostic purposes. Hence for the clinical data analysis such type of automated signal processing systems are commonly used.

Large amount of data processing can be speed up and simplified by implementing such new automated methods. This will help the physician to frequently decide better diagnosis for the patient's on the basis of its numerical values which is calculated during examination. Examine this volume of data is not always easy and clearly identifiable. Therefore there exist consultation systems that help in reducing human errors [3].

AUTOMATIC DETECTION AND ANALYSIS OF ECG SIGNAL USING DIFFERENT SIGNAL PROCESSING AND FEATURE EXTRACTION TECHNIQUES- AN OVERVIEW

Signal processing mainly used for the analysis, synthesis, and adaptation of any types of signals, which are broadly defined as functions, which conveying, "information about the behavior or attributes of some phenomenon". Signal processing techniques are used to improve the signal strength for transmission fidelity, storage efficiency, and subjective quality, and to highlight or detect components of interest in a measured signal [4], [5].

Feature extraction is done on the original set of data to get the feature which is to be informative and non-redundant, facilitating the succeeding learning and generalization steps, and in many cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. Therefore there are various techniques available for feature extraction [4], [5].

Now a day, different signal processing techniques have developed as an active area of research for automatic detection and analysis of bio-signals, even though their detection and classification remains challenging due to the considerable variation in the morphological and temporal characteristics of ECG signal between different subjects. Among the several techniques investigated in the literature, the some techniques included are time domain analysis [11]-[14], statistical approach [15]-[17], hybrid feature [18],[19], frequency-based analysis [20], and time-frequency analysis [21]-[23] for feature extraction of ECG signals. In the different transform approach, the Fourier transform (FT) (i.e., spectral analysis) is unable to analyze the non-stationary ECG signal, as it fails to provide any information about the occurrence of the frequency component in time. Whereas Short Time Fourier Transform (STFT) overcomes the opacity of Fourier Transform by providing a constant window [28], but the choice of the size of the window remains a challenge. The wavelet transforms family [29] addresses this drawback of fixed window length in STFT for all frequencies. The Wavelet Transform uses longer windows at low frequencies and shorter windows at high frequencies for the analysis of bio-signals. On the other hand, the proper selection of mother wavelet (db4) and sampling frequency for extracting the frequency components is a major challenge, if it fails, then it may result in misleading information. In fact, stock well transforms(S-transform) [28], [29] overcomes the weakness of WT by uniquely combining a frequency-dependent resolution of time-frequency space and absolutely referenced phase information of the input signal. The S-transform localizes the real and imaginary spectra to approximate the local amplitude and phase spectrum [28], [29]. The multi-resolution decomposition of the stock well transform is helpful but corresponds to the redundant representation of the time-frequency plane and that's why is computationally expensive. Whereas, a discrete and non-redundant version of the S-transform, i.e., discrete orthogonal stock well transform (DOST) [28], [29] has made the use of the S-transform more feasible and reliable. It can be represented as an orthogonal set of basic functions [28], [29] that localize the spectrum and maintain the advantageous properties of ST. A new approach involves the application of the DCT in DOST to study the time-frequency distribution of the ECG signal to overcome the drawback of the widely used above mentioned classical signal processing techniques such as FT, STFT and WT(family)[30]. The benefits of using cosines are the lack of discontinuities. In the case of DOST, the input signal is represented periodically in DFT, the signal tends to lose its form during truncation of coefficients. Although, in DCT, the signal can withstand during more truncation of coefficients, i.e., it keeps the desired shape of the input signal. Besides, the DCT is real-valued (i.e., it computes real coefficients) and does not have negative frequencies. Therefore reducing the overall complexity. The application of the DCT in DOST gathers the advantages together and represents the most significant coefficients in the lower frequency components. The coefficients are distributed according to the frequency bandwidths in the time-frequency space and not mirrored from either side (i.e., asymmetric property), which is not the case of DOST. The features extracted using DCT-based DOST represents the time-frequency characteristics of the ECG signal and are asymmetrical in nature [28]-[30].

NEED OF CLASSIFICATION

The classification plays a vital role in biomedical experiments. In clinics, the aim of classification is to differentiate pathology from normal. For example, By monitoring physiological recordings, the physician judge if patients suffer from illness [6]; By watching cardiac MRI scans or ECG, the cardiologists identify which region the myocardium experiences failure [7]; By analyzing gene sequences of a family, the geneticists conclude the likelihood that the children inherit diseases from their parents [8].These above examples illustrate that automatic decision making can be a main step in medical practice. The incorrect disease identification will not only waste diagnosis resources but also cause delay in treatment or some cases cause loss of patients' lives.

Classification faces several difficulties in biomedical: (i) Experts need to take a long time to gather enough knowledge to reliably distinguish between different related cases, such as identifying whether it is normal or abnormal; (ii) Manual classification among these cases is laborious and time consuming; (iii) The most difficult challenge occurs when the signal characteristics are not prominent and hence not easily noticeable by experts. Automated methods for classification in signal processing assure to overcome some of these difficulties and to help out to make best decision in the field of biomedical. An automatic classifier can learn from the experiences i.e. from a database the categorical information, replaces the human operators, and classify imperceptible features without bias [3].

II. CLASSIFICATION

The function of a classifier is to automatically partition a given set of biomedical signals into several subsets. Generally Classifier falls into two categories: Supervised and Unsupervised learning [9], [10]. In Supervised learning, the experts are to be requested to label a small portion of the data. The classifier then propagates this previous knowledge to the remaining unlabeled data. The classes or labels provided by the experts are a good source of information for the classifier to learn, how to make decision? Sometimes, the experts lack confidence in labeling, which gives rise to uncertainty in the classification results. The worst case arises, when the experts mislabel the data and leading the classifier to produce incorrect results. Unlike supervised classification, which is highly sensitive to the past labels, unsupervised classifiers need no previous class labels. In unsupervised learning, the classifier learns by itself the optimal decision rule from the available data and automatically partitions the data set into different classes. The system then matches the two subsets to normal and abnormal states. One of the main disadvantages of unsupervised learning is that the spectral classes do not always correspond to informational classes. The user also has to spend time in interpreting and label the classes, which following the classification. Spectral properties of classes can also change over time, so user can't always use same class information when moving from one image to another in image processing. The all investigated literature uses the supervised learning for the classification instead of unsupervised learning technique [9].

DIFFERENT APPROACHES FOR USING CLASSIFICATION TECHNIQUES:

Feature extracted tools are incorporated with classification algorithms such as Linear Descriminant [11], [12], [24], Neural Networks [17], [21], [23], Neuro-Fuzzy approach [25], and Support Vector Machine (SVM) [15], [16], [18], [26], [27] to provide resourceful detection and analysis of cardiac abnormalities. In the hybrid system, hybrid feature using morphological and dynamic feature containing RR-interval are concatenated to constitute the final feature set, which are utilized to predict their classes using SVM [30]. Further to obtain the optimal parameters and to improve the prediction performance the particle swarm optimization (PSO) technique is used. Each classification techniques have advantages and disadvantages. ANN based classification system produces high confident arrhythmia classification result as compared to other classifier for ECG signal analysis [29]. Based on overall performance for the detection of arrhythmia SVM classifier has best result but it has same limitation. SVM classifier, classified output only into two categories because of that it cannot be used in real time application whereas overall performance of ANN is the best in term of accuracy, sensitivity and predictivity [30]. Also computational speed time is almost negligible compared to other classifiers. It can be the best method for arrhythmia detection in real time classification system due to short processing time and relatively high accuracy. Different neural network architectures required different learning algorithms. ANN trained with the back-propagation algorithm gives the accuracy up to 96.77% [31]. Other methods of ECG beat recognition are neural network classification method [33]-[40], multi-layer perceptron (MLP), they also known as "conventional backpropagation neural networks (BPNN)", they are able to recognize and classify ECG signals more accurately. Some researchers have used neural networks for the ECG beat recognition [33–37]. Nevertheless conventional BPNN suffers from slow convergence to local and global minima and from random settings of initial values of weights, which may show the neural networks have very poor mappings from input to output. Because of that reason, researchers have started to use hybrid structure or system and it has proved that hybrid system gives better and faster result then conventional BPNN. The hybrid system such as fuzzy hybrid neural network, combines their advantages and it is said to be more capable of recognizing other biological signals than conventional BPNN [36], [37],[41].

III. CONCLUSION

The information of heart and cardiovascular functionality is detected by the ECG. To enhance the patient living quality, an appropriate treatment is essential and it can be only achieved by extracting this information by ECG signal analysis. Therefore the feature extraction for ECG signal analysis can be done in time domain as well as in frequency domain. Different feature extraction techniques for ECG signals have been proposed earlier in the literature. As it plays a vital role in cardiac disease diagnosis, the development of accurate and quick method for automatic ECG feature extraction is of major importance. This paper reviewed various methods involved in ECG signal analysis. Also, this paper shows a comparative analysis of the performance of feature extraction and classification techniques.

IV. FUTURE IMPROVEMENT

The previous research implemented some of the feature extraction methods as explained earlier in this paper. Every technique has its own advantages and disadvantages. The future work mainly focus on feature extraction from an ECG signal using more statistical data that are capable of effectively distinguish various diseases properly. It also required relatively short processing time with high accuracy in real-time system. Design a micro controller based electronic hardware that displays the result on the screen after simulation.

REFERENCES

[1]. Fauci AS, Braunwald E, Kesper DL, Hauser SL, Longo DL, Jamesonn JL, Loscalzo J (2008) Harrison's principles of internal medicine, 17th edn. Mc-Graw Hill, New York.

[2]. Park K (2005) Park's textbook of preventive and social medicine, 18th edn. Banarsidas Bhanot publishers, India.

[3]. Hsun-Hsien Chang and Jos´e M. F. Moura "Biomedical Signal Processing" ed. Myer Kutz, in biomedical engineering and design handbook, 2nd edition, Vol. 1, Mc-Graw Hill. 2010, chapter 22, pp. 559-579.

[4]. https://en.wikipedia.org/wiki/Signal_processing.

[5]. https://en.wikipedia.org/wiki/Feature_extraction.

[6]. N. Hazarika, A. C. Tsoi, and A. A. Sergejew, "Nonlinear considerations in EEG signal classification," IEEE Transactions on Signal Processing, vol. 45, pp. 829–836, April 1997.

[7]. H.-H. Chang, J. M. F. Moura, Y. L. Wu, and C. Ho, "Automatic detection of regional heart rejection in USPIO-enhanced MRI," to appear IEEE Transactions on Medical Imaging.

[8]. B. S. Carter, T. H. Beaty, G. D. Steinberg, B. Childs, and P. C. Walsh, "Mendelian inheritance of familial prostate cancer,"Proceedings of the National Academy of Sciences of the United States of America, vol. 89, pp. 3367–3371, April 1992.

[9]. R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification. New York, NY: John Wiley & Sons, second ed., 2001.

[10]. T. Mitchell, Machine Learning. New York, NY: McGraw Hill, 1997.

[11]. P. de Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features", IEEE Trans. Biomed. Eng., vol. 51, no. 7, pp. 1196–1206, Jul. 2004.

[12]. P. de Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," IEEE Trans. Biomed. Eng., vol. 53, no. 12, pp. 2535–2543, Dec. 2006.

[13]. S. Mitra, M. Mitra, and B. B. Chaudhuri, "A rough-set-based inference engine for ECG classification," IEEE Trans. Instrum. Meas., vol. 55, no. 6, pp. 2198–2206, Dec. 2006.

[14]. S. Raj, K. Maurya, and K. C. Ray, "A knowledge-based real time embedded platform for arrhythmia beat classification," Biomed. Eng.Lett., vol. 5, no. 4, pp. 271–280, Dec. 2015.

[15]. F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," IEEE Trans. Inf. Technol. Biomed., vol. 12, no. 5, pp. 667–677, Sep. 2008.

[16]. S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machinebased expert system for reliable heartbeat recognition," IEEE Trans. Biomed. Eng., vol. 51, no. 4, pp. 582–589, Apr. 2004.

[17]. S. Raj, S. Luthra, and K. C. Ray, "Development of handheld cardiac event monitoring system," IFAC-PapersOnLine, vol. 48, no. 4, pp. 71–76, 2015.

[18]. C. Ye, B. V. K. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," IEEE Trans. Biomed. Eng., vol. 59, no. 10, pp. 2930–2941, Oct. 2012.

[19]. S. Raj, K. C. Ray, and O. Shankar, "Cardiac arrhythmia beat classification using DOST and PSO tuned SVM," Comput. Methods Programs Biomed., vol. 136, pp. 163–177, Nov. 2016.

[20]. K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network,"IEEE Trans. Biomed. Eng., vol. 46, no. 2, pp. 179–185, Feb. 1999.

[21]. T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," IEEE Trans. Biomed. Eng., vol. 56, no. 5, pp. 1415–1426, May 2009.

[22]. S. Banerjee and M. Mitra, "Application of cross wavelet transform for ECG pattern analysis and classification," IEEE Trans. Instrum. Meas., vol. 63, no. 2, pp. 326–333, Feb. 2014.

[23]. S. Raj, G. S. S. P. Chand, and K. C. Ray, "Arm-based arrhythmia beat monitoring system," Microprocess. Microsyst., vol. 39, no. 7, pp. 504–511, 2015.

[24]. M. Llamedo and J. P. Martinez, "Heartbeat classification using feature selection driven by database generalization criteria," IEEE Trans. Biomed. Eng., vol. 58, no. 3, pp. 616–625, Mar. 2011.

[25]. T. H. Linh, S. Osowski, and M. Stodolski, "On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network," IEEE Trans. Instrum. Meas., vol. 52, no. 4, pp. 1224–1231, Aug. 2003.

[26]. G. D. Fraser, A. D. C. Chan, J. R. Green, and D. T. Macisaac, "Automated biosignal quality analysis for electromyography using a one class support vector machine," IEEE Trans. Instrum. Meas., vol. 63, no. 12, pp. 2919–2930, Dec. 2014.
[27]. S. Raj and K. C. Ray, "A comparative study of multivariate approach with neural networks and support vector machines for arrhythmia classification," in Proc. Int. Conf. Energy, Power Environ., Towards Sustain. Growth (ICEPE), Jun. 2015, pp. 1–6.
[28]. R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: The S transform," IEEE Trans. Signal Process.,vol. 44, no. 4, pp. 998–1001, Apr. 1996.
[29]. R. G. Stockwell, "A basis for efficient representation of the S-transform," Digit. Signal Process., vol. 17, no. 1, pp. 371–393, May 2006.
[30]. S. M. Jadhav "Artificial Neural Network Based Cardiac Arrhythmia Classification Using ECG Signal Data", 2010 International Conference on Electronics and Information Engineering (ICEIE 2010), vol. 1, pp. 228-231.
[31]. M. D. Ingole "Electrocardiogram (ECG) Signals Feature Extraction and Classification Using Various
[32]. Signal Analysis Techniques", International Journal of Engineering Sciences & Research Technology (IJESRT), vol.3, no. 1,pp. 39-44, January 2014
[33]. G.Krishna Prasad "Classification of ECG Arrhythmias using Multi-Resolution Analysis and Neural Networks".
[34]. Y. Ozbay, B. Karlik, A recognition of ECG arrhythmias using artificial neural network, Proceedings of the 23rd Annual Conference, IEEE/EMBS, Istanbul, Turkey, 2001.
[35]. Y. Ozbay, Fast recognition of ECG arrhythmias, Ph.D. Thesis, Institute of Natural and Applied Science, Selcuk University,1999.
[36]. S.Y. Foo, G. Stuart, B. Harvey, A. Meyer-Baese, Neural network-based ECG pattern recognition, Eng. Appl. Artif. Intell.15 (2002) 253–260.
[37]. V. Pilla, H.S. Lopes, Evolutionary training of a neuro-fuzzy network for detection of P wave of the ECG, Proceedings of the Third International Conference on Computational Intelligence and Multimedia Applications, New Delhi, India, 1999,pp. 102–106.
[38]. M. Engin, S. Demira˘g, Fuzzy-hybrid neural network based ECG beat recognition using three different types of feature set, Cardiovasc. Eng. Int. J. 3 (2) (2003) 71–80.
[39]. S. Haykin, Neural Networks: A Comprehensive Foundation, Macmillan, NewYork, 1994.
[40]. B. Karlık, M.O. Tokhi, M. Alcı, A fuzzy clustering neural network architecture for multifunction upper-limb prosthesis, IEEE Trans. Biomed. Eng. 50 (11) (2003) 1255–1261.
[41]. J.S.R. Jang, C.T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing, Prentice-Hall, Englewood Cliffs, NJ, USA, 1997.
[42]. S. Osowski, T.H. Linh, ECG beat recognition using fuzzy hybrid neural network, IEEE Trans. Biomed. Eng. 48 (11) (2001) 1265–1271.

Sunita Kumari "A Study on Various Machine Learning Techniques For ECG Signal Analysis" .IOSR Journal of Engineering (IOSRJEN), vol. 8, no.4, 2018, pp. 48-53.