Genetic Algorithm based Feature Extraction for ECG Signal Classification using Neural Network

[1]R. Sathya, [2]K. Akilandeswari
[1,2]Research Scholar
[1]Department of Computer Science
[1]Govt. Arts College, Salem-636007, India

***Abstract* - Cardiac Arrhythmia is a key problem faced by many people regardless of age and gender. P wave, QRS complex and T wave forms a complete cardiac cycle. Absence or abnormal appearance of any waves lead to cardiac arrhythmia. If these abnormalities are diagnosed at the earliest stage, appropriate treatment can be provided to the patients. In our research work, classification technique in data mining is used for classifying normal and abnormal patients. Pan Tompkin algorithm is used for de-noising of Electrocardiogram (ECG) signals and to obtain QRS on filtered signal. Genetic algorithm and Neural network classifier are used to achieve high accuracy in classification of signals.**

***Index Terms* – ECG Signal, Preprocessing, Fitness evaluation, Roulette Wheel selection.**

I. INTRODUCTION

ECG is a tool which is used to interpret the electrical activity of human heart. ECG is a form of multichannel signal, because signals are interpreted by placing electrodes on anatomical positions on body surface. Electrical impulses are generated during depolarization (contraction) and re-polarization (relaxation) of cardiac muscles. ECG signals are non-stationary signals which vary from person to person. Some of the ECG signals may contain noisy data which must be pre-processed to perform classification and to improve accuracy of the classifier. There are many techniques for de-noising ECG signals. The most commonly used de-noising technique is classical Pan Tompkin algorithm [1]. Using Pan Tompkin algorithm, it is also possible to extract QRS on filtered ECG signal.

II. LITERATURE SURVEY

ECG signal analysis for detection of Heart Rate and Ischemic Episodes was proposed by sahoo et al [2] where they calculated Heart Rate (HR) and ischemic episodes which follows mainly these stages: pre-processing, beat classification and ischemic episode recognition. HR was calculated from feature extraction of ECG signal and also used for detecting cardiac arrhythmia. The performance of ischemic episode detection resulted in 88.08% sensitivity and 92.42 % positive prediction accuracy.

Efficient Filtering Techniques of ECG Signal Using Fir Low Pass Filter with Various Window Techniques was proposed by Patial et al [3] where they presented the study of FIR filter using window techniques for ECG signal processing. Power spectral density, average power and signal to noise ratio were calculated from ECG signal and compared the performance of different window method used for FIR filter.

A survey of Noise removal Techniques for ECG Signals was proposed by Chandrakar et al [4] where they presented Finite Impulse Response (FIR) filter based on various windows and Infinite Impulse Response (IIR) filters for noise removal of ECG signal and from the results of papers it was seen that kaiser window based FIR filter was better to remove artifacts from ECG signals.

Classification of Arrhythmic ECG Data Using Machine Learning Techniques was proposed by Vishwa et al [5] where they presented an automated ANN based classification system for cardiac arrhythmia using multi-channel ECG recordings. Their experimental results gave 96.77% accuracy on MIT-BIH database and 96.21% on database prepared by including NSR database. Denoising of ECG signal using Undecimated Wavelet Transform was proposed by Baliram et al [6] where they presented elimination of noises occurred in ECG signal using Undecimated Wavelet transform. ECG denoising was carried out using the wavelet coiflet and daubechies filter.

Genetic algorithm for classification and feature extraction was proposed by Pei et al [7] where they presented classification and feature extraction for high dimensionality pattern using genetic algorithm. They combined the effectiveness KNN and Genetic algorithm in order to obtain effective result.

Learning concept classification rule using genetic algorithm was proposed by Jong et al [8] where they explored the use of adaptive search techniques to construct a system GABEL which continuously learns and refines concept classification rules from its interaction with the environment.

III. ECG SIGNAL

Normal cardiac cycle completely differs from abnormal cardiac cycle. Normal cardiac cycle looks like Fig 1.,

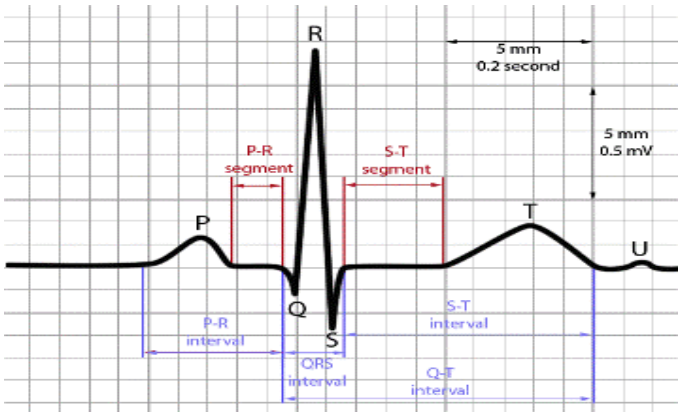

Fig 1: Normal Cardiac Cycle. [GOOGLE SOURCE]

Normal ECG signal consists of P wave, QRS complex, T wave, U wave, PR segment, ST segment, PR interval and QT interval. U wave usually not present in most of the ECG signals.

a. ECG Waves

P wave: The first deflection of the heartbeat is a small upward wave called the P wave which indicates atrial de-polarization (contraction). It activates left and right artria simultaneously. The initial portion of the P wave represents right atrial depolarization and the terminal portion represents left atrial de-polarization. The P waves must not be larger than 0.3mV (3 mm)

Q wave: Downward deflection after the P wave. The normal Q wave represents septal de-polarization. The dead muscle never conducts and never produces current, so the ECG picks up current flowing away from this muscle, producing a strong negative deflection.

R wave: First upward deflection after the P wave, even though Q waves are absent. It represents early ventricular de-polarization.
S wave: First negative deflection after the R wave. It represents the late ventricular de-polarization.

T wave: Last wave in ECG signal which represents re-polarization (relaxation) of the ventricles. T wave is simultaneously followed by U wave which is not seen in many cases.

b. ECG Segments

PR Segment: represents the isoelectric line which occurs immediately after off set of P Wave and terminates at the onset of QRS complex.
ST Segment: Bridge between QRS complex and T wave. It starts at the offset of S wave and ends at the onset of T wave.

c. ECG Intervals

PR interval: Starts at the onset of the P wave and terminates at the onset of the QRS complex. **RR interval**: Represents time measurement between two simultaneous R waves.

IV. METHODS AND MATERIALS

Our research work is fully implemented in MATLAB which stands for matrix laboratory excellent research tool for tackling any kind of complexity. Following Fig 2, shows the systematic flow of neural network classification.

A. Dataset

MIT BIH arrhythmia dataset is used for extracting the features of ECG signal. It contains 4000 long-term Holter recordings that were obtained by the Beth Israel Hospital Arrhythmia Laboratory. It contains 48 inpatient records [9].

B. Genetic Algorithm

Genetic algorithm, one of the Evolutionary algorithms was developed based on the natural inheritance behavior. It is mainly used in solving optimization problem. Discontinuous, stochastic or highly non linear problems can be easily solved using genetic algorithm. In our proposed work two main operations are performed namely ranking and selection. Individuals are taken one at a time and fitness of sample points for individual is calculated via Ranking. Individuals sample points are selected based on the evaluated fitness value. Feature extraction is done after selection process by applying mean and maximum of selected value.

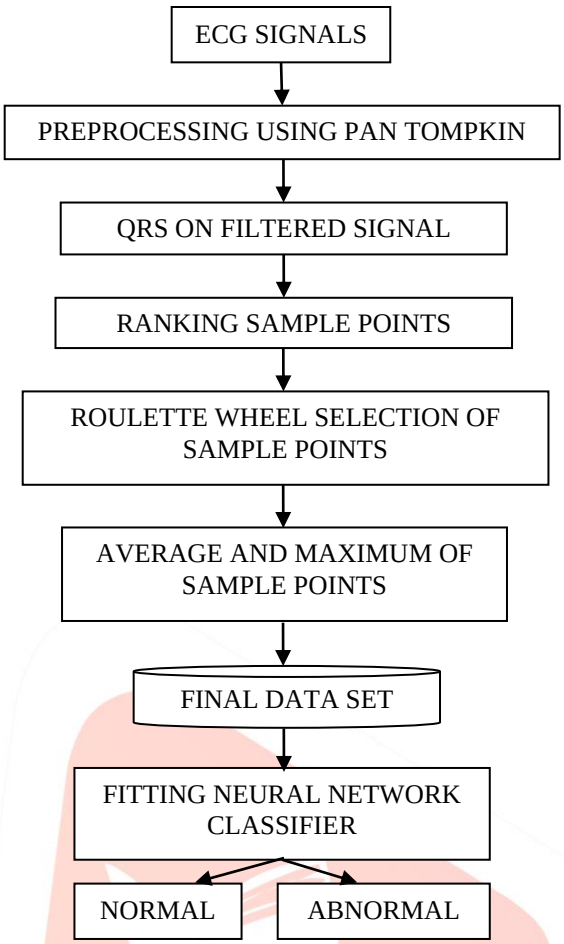

Fig 2: Systematic flow of Classification of Heart abnormalities

C. Neural Network Classification

Classification is a data mining technique used to associate input patterns with the corresponding output pattern. In our research work classification is done using neural network. One of the mathematical models is neural network which trains another model based on user presented input and output data. Neural network has set of layers, were first layer is input layer and last layer is output layer. It may or may not contain hidden layers. Weight adjustments are performed based on the input by the hidden layer neurons. In our proposed system we used fitting neural network for classification. Features extracted using genetic algorithm is passed to fitting neural network to predict normal and abnormal patients. Fitting neural network is a feed forward neural network.

D. Evaluation metrics

Evaluation metrics are used to evaluate accuracy of classifier. Precision, Recall, F-Measure are used to evaluate accuracy of classifier. Various classifier evaluation metrics are listed below [10]:

Precision denotes the probability of selected items to be relevant. It is calculated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall denotes the probability of relevant items to be selected. It is calculated as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

F-Measure denotes the harmonic mean of precision and recall. It is calculated as:

$$\text{F} - \text{Measure} = \frac{2.\,\text{Precision}\,.\,\text{Recall}}{\text{Precision} + \text{Recall}}$$

Accuracy is used for testing performance of classifier and it is calculated as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

V. RESULTS AND DISCUSSION

This research work is implemented using Pan Tompkin algorithm to get QRS on filtered signal, ranking and selection function from genetic algorithm for feature extraction and Fitting Neural Network technique for classification. Initially raw ECG signal is

implemented with Pan Tompkin algorithm to de-noise signal and to obtain QRS. Fitness evaluation using ranking method is performed in genetic algorithm and selecting the sample points using roulette wheel selection. Mean and maximum of sample points are extracted as features. Extracted features are then trained with Fitting neural network and following results are obtained.

Correctly Classified Instances : 47.000000%
Incorrectly Classified Instances : 1.000000%

Table 1: Experimental Report

	Normal	Abnormal
Precision	0.9759	1.0000
Recall	1.0000	0,8889
F-Measure	0.9873	0.9412

Accuracy 0.979167
Sensitivity 1.000000
Specificity 0.888889

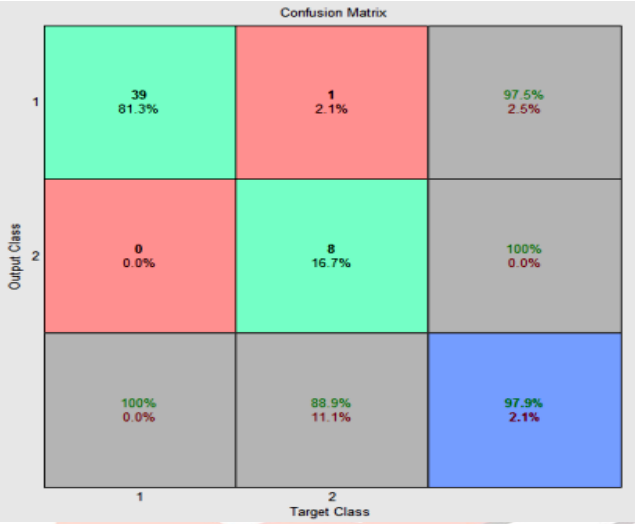

Fig 3: Confusion Matrix

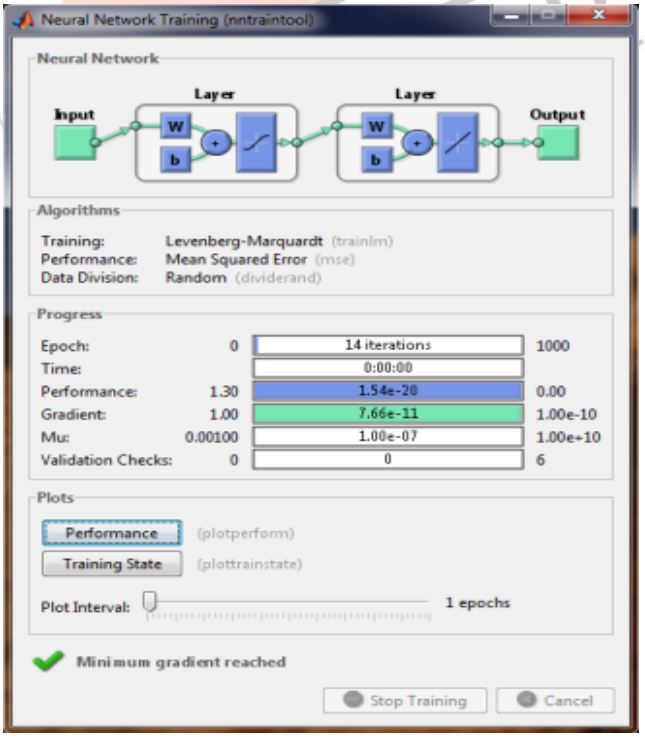

Fig 4: Neural Network Training

VI. REFERENCES

[1] PAN.J, TOMPKINS. W.J,"A Real-Time QRS Detection Algorithm" IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. BME-32, NO. 3, MARCH 1985.

[2] Goutam Kumar Sahoo, Samit Ari and Sarat Kumar Patra, ECG signal analysis for detection of Heart Rate and Ischemic Episodes, International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-3 Number-1 Issue-8 March-2013.

[3] Preeti Patial and Er. Kuldeep singh, Efficient Filtering Techniques of ECG Signal Using Fir Low Pass Filter with Various Window Techniques, international journal of advanced research in computer science and software engineering, volume 4, issue 7, july 2014.

[4] Bhumika Chandrakar, O.P.Yadav, V.K.Chandra, A survey of Noise removal Techniques for ECG Signals, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 3, March 2013.

[5] Classification of Arrhythmic ECG Data Using Machine Learning Techniques, Abhinav Vishwa,Mohit K.Lal, Sharad Dixit and Dr.Pritish Vardwaj, International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 1, Nº 4.

[6] Baliram S. Gayal and Shaikh, Denoising of ECG signal using Undecimated Wavelet Transform, IJAREEIE, Vol. 3, Issue 1, January 2014.

[7] Min Pei, Erik Goodman, William F. Punch and Ying Ding, Genetic algorithm for classification and feature extraction, ww.citeseerx.ist.psu.edu.

[8] Kenneth A De Jong and William M. Spears, Learning concept classification rule using genetic algorithm, www.ijcai.org.

[9] http://www.physionet.org/physiobank/database/mitdb/.

[10] http://en.wikipedia.org/wiki.