

Classification of Heart conditions by Statistical Characterization of ECG Signal

Asma Haque and Md. Abdur Rahman

Abstract— Electrocardiogram (ECG) signal exhibits important distinctive feature for different cardiac issues. Automatic classification of electrocardiogram (ECG) signal can be used for primary detection of various heart conditions. Information about heart and ischemic changes of heart may be obtained from cleaned ECG signals. ECG signal has an important role in monitoring and diacritic of the heart patients. An accurate ECG classification is challenging problem. The accuracy often depends on proper selection of observing parameters as well as detection algorithms. Heart disorder means abnormal rhythm or abnormalities present in the heart. In this research work, we have developed a decision tree based algorithm to classify heart problems by utilizing the statistical signal characteristic (SSC) of an ECG signal. The proposed model has been tested with real ECG signal to successfully (60-98%) detect normal, apnea and ventricular tachyarrhythmia condition.

Keywords— ECG Signal; SSC Parameter; heart disease; classification; decision tree;

I. INTRODUCTION

Electrocardiogram (ECG) is one of the important signal and diagnostic tools to observe heart conditions. It's a technique to measure and records electrical activity of heart [1]. As ECG is a biomedical signal, it tends to be non-stationary. Heart rate frequency is one of the monumental facts of our health and it can be detected from ECG signal. The analysis and classification of the ECG can be utilized in identifying the different types of heart abnormalities. ECG signal consist of P-QRS-T waves. If any deviation or disorder is found in these parameters, it indicates abnormalities present in heart and then it is called abnormal ecg signal [1]. Frequency of the ECG signal is in the range from .5 Hz to 100 Hz. The amplitude and duration of this wave (PQRST) contains useful information about the nature of disease. In ECG signal, there are two phases present, one is repolarization and other is depolarization [2]. Time interval between onset of atrial depolarization and onset of ventricular depolarization is called PR interval. QRS complex is depolarization of ventricle [2]. QT interval is the interval between onset of ventricular depolarization and ending point of ventricular repolarization [2].

Time interval between two QRS complexes is called R-R interval. T wave is called as the ventricular repolarization [2]. Fig 1 shows ECG signal in one cycle with different segment and intervals. For a regular ECG signal this pattern is repeated for the life-time of human.

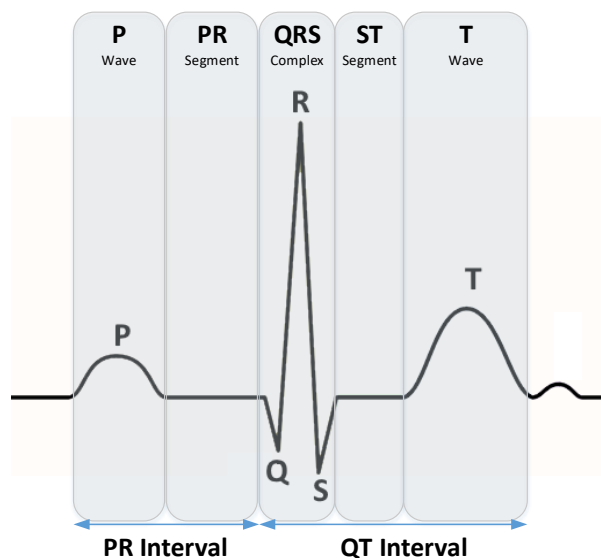


Fig. 1. ECG signal in one cycle

II. LITARATURE REVIEW

A number of methods have been implemented to classify the ECG signal. Artificial immune recognition system, neural network, fuzzy neural network, data mining, wavelet transform and many other techniques are developed for automatic detection and classification of heart disorder [3]. Wavelet transforms approach for feature extraction is a well-known work for ECG signal classification. It extracts the feature from the ECG signal and recognizes the abnormalities of heart. Artificial immune recognition system is another technique that utilizes the well-known neural network (NN) architecture. Multi-layered perceptron (MLP) with back propagation training algorithm and a new fuzzy clustering NN architecture (FCNN) system was used for diagnosing the heart diseases [3]. Another technique was implemented which is known as fuzzy neural network. In this technique, multithreading approach was used for finding out of atrial fibrillation, arrhythmia and normal sinus rhythm [3]. In

Asma Haque

Graduate student, AIUB
Email: haqueasma15@yahoo.com

Md. Abdur Rahman

Professor and Director, FE, AIUB
Email: arahman@aiub.edu

summary, ECG classification systems comprises three sequential steps; Pre-processing of ECG signal, feature extraction and Classification. Accuracy of the classification often depends on the number and quality of the extracted distinctive features. Although few literatures achieved good results [4], the classification accuracy may be further improved by using statistically significant feature. In this paper, we developed a decision tree based algorithm by using Sstatistical Signal characterization (SSC) parameters.

III. SSC PARAMETERS

The ECG wave form is simply repetitive signal of same pattern. For the detection of heart activity, information is mainly extracted from the amplitude and time intervals of the signal. Hence, simple statistical parameters based on time and amplitude should be enough to distinguish the differences. Thus, statistical signal characterization (SSC) parameters [4] are very suitable for this purpose. Extraction of SSC parameters are relatively easier and requires less number of computations compared to other techniques [4]. Four SSC parameters are extracted from the ECG signal. By using SSC technique, a waveform is characterized not only as a function component of amplitude but also as a function of their relative time. But, compared to spectrogram, wavelet transform or Fourier transforms, it requires much less computation with less complexity [4].The SSC parameters are mean amplitude (M_a), amplitude deviation (D_a), mean time (M_t) and time deviation (D_t).At first any given signal is divided into segments based on the extrema. The separation between a maxima and minima of the amplitude is regarded as segment as shown in figure 2. This waveform is mixing of several ranges of frequency components and exhibits a series of extrema over the interval. In this figure segment of the wave form is indicated by the figure's boxed areas being bounded by consecutive extrema [4]. When the waveform crosses a zero-slope condition, we can declare it a segment boundary condition. The waveform figure 2 shows 10 segments.

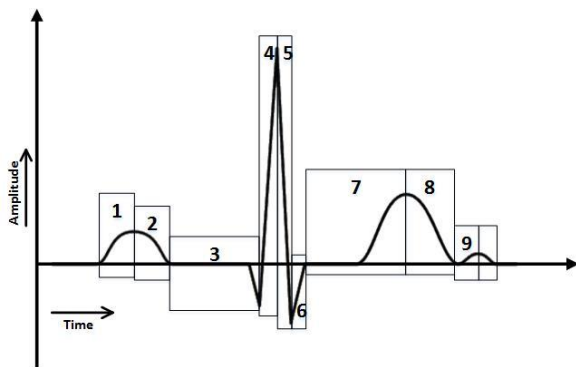


Fig. 2. ECG waveform Segmentation

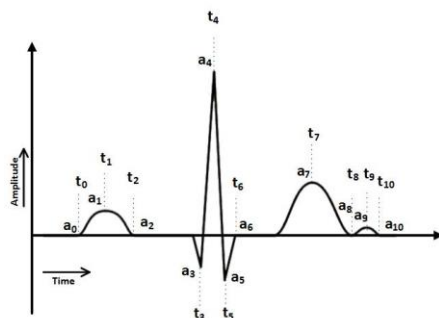


Fig. 3. Segment-boundary of amplitudes and times

As shown in the figure 3, every segment is bounded by minima and maxima. The starting and ending amplitude and time of each segment is termed as a_0 - a_n and t_0 - t_n . Then the amplitude (A_n) and time period (T_n) of n-th segment is calculated as per the equations (1) and (2).

$$A_n = a_n - a_{n-1} \quad \dots(1)$$

$$T_n = t_n - t_{n-1} \quad \dots(2)$$

Where,

A_n =Amplitude of n-th segment,

T_n = period of n-th segment,

a_n = amplitude of concluding extremum,

t_n = Time at the concluding extremum,

a_{n-1} = Amplitude of beginning extremum,

t_{n-1} = Time at the beginning extremum.

Then the four SSC parameters may be calculated by the following equations (3) to (6).

$$M_a = \frac{1}{N_s} \left[\sum_{i=1}^{N_s} (A_i) \right] \quad \dots(3)$$

$$M_t = \frac{1}{N_s} \left[\sum_{i=1}^{N_s} (T_i) \right] \quad \dots(4)$$

$$D_a = \frac{1}{N_s} \left[\sum_{i=1}^{N_s} (A_i - M_a) \right] \quad \dots(5)$$

$$D_t = \frac{1}{N_s} \left[\sum_{i=1}^{N_s} (T_i - M_t) \right] \quad \dots(6)$$

Where,

M_a = amplitude mean;

M_t = period mean;

D_a = amplitude mean deviation, or simply amplitude Deviation

D_t = period mean deviation, or simply period deviation

N_s = number of samples

IV. PROPOSED SYSTEM

ECG signal has an important role in the clinical evaluation of cardiac treatment. It records different potentials of the heart [5]. Analysis and classification system of ECG signal provides us overall idea about cardiac disorder. In this work, ECG Signal has been classified through parameter extraction. Although ECG signal is supposed to be repeated after certain interval, in case of heart disorder the amplitude and time of different segment varies. Hence, it will be beneficial to divide the whole signal into few samples. From experiment, it was observed that, for acceptable results each segment should have at least 10 complete ECG cycles.

In the proposed method, we take same length of segments from raw ECG signal for processing. At first the signal is pre-processed for removing of artifact and noises because ECG signal always contains various artifacts like power line interferences, contact noises, motion artifact, base line drift etc [6]. After pre-processing, actual method of feature extraction and classification has been accomplished. These extracted parameters are used as input to the already trained supervised network for classification.

The system may classify different types of heart conditions present in the human heart. In this work different types of heart beats like Normal ECG signal, Ventricular Tachyarrhythmia, Apnea signal have been analyzed. These heart beats show different types of SSC characteristics and nonlinear behavior. The SSC parameters are very susceptible to noise hence a filter is used to denoise the signal. Butterworth filters provided relatively clean signals. Butterworth filters are having a property of maximally flat response and no ripples in the pass band comparing with chebyshev type I/type II filter or an elliptic filter.

A Butterworth low pass filter is designed using

$$[B, A] = \text{butter}(n, Wn)$$

Where,

n = order of the filter

Wn = cut-off frequency (must be between 0 and 1)

B (numerator) and A (denominator) are the filter coefficients of length $n+1$.

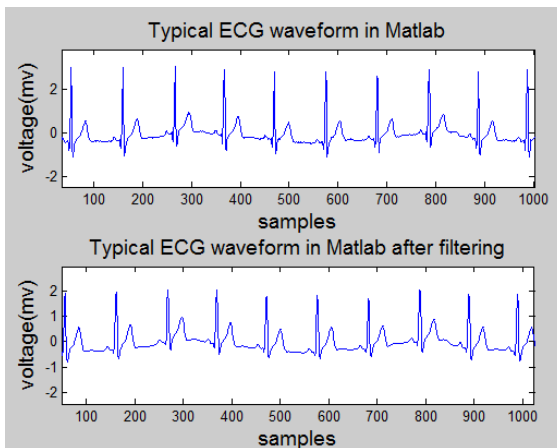


Fig. 4. ECG signal filtering using Butterworth filter

The system is already trained with known signals for all the heart conditions. Hence the extracted SSC parameters may be used classify the signal. This classification falls under the supervised learning algorithm and a decision tree is most suitable for the classification. Weka tool, j48 algorithm is used for data mining and classifying the signal.

The flowchart of the proposed method is presented in figure 5. Here two process of training is classification is clearly explained. During the training the SSC parameters are extracted from the known ECG signals for different heart conditions. These SSC parameters are quite unique for different heart conditions. As a result, the j48 algorithm may construct a decision tree for the classification process.

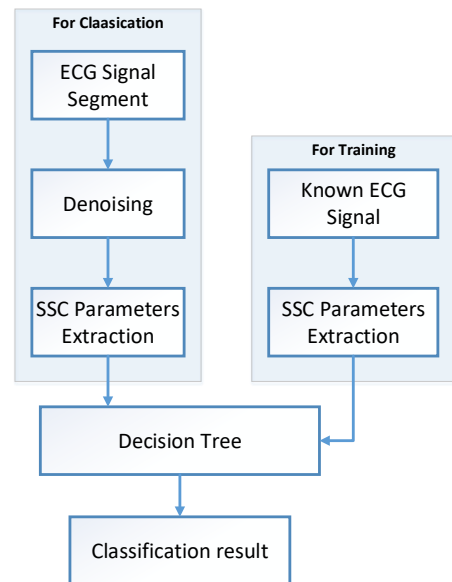


Fig. 5. Typical Block diagram of complete process

In case of classification, the ECG signal is first divided into few segments. These segments are then applied to the system for filtering and parameter extraction. Eventually the parameters are matched to the existing decision tree to find the appropriate heart conditions. As an ECG signal is divided into few segments and each segment is classified independently, finally the classification decision is made majority rule. The process of segmentation and final decisions are presented in figure 6.

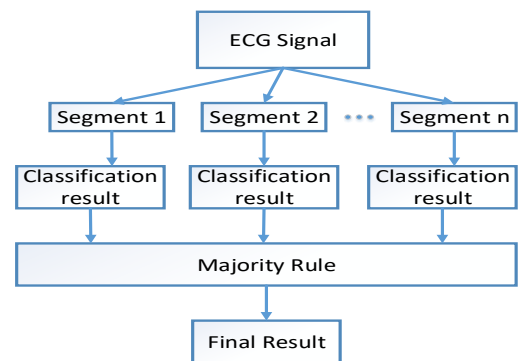


Fig. 6. Classification results by segmentation

V. RESULTS AND DISCUSSIONS

For experiment we used the real-life data from BIO?? From this data we extracted four statistics signal characterization parameters.

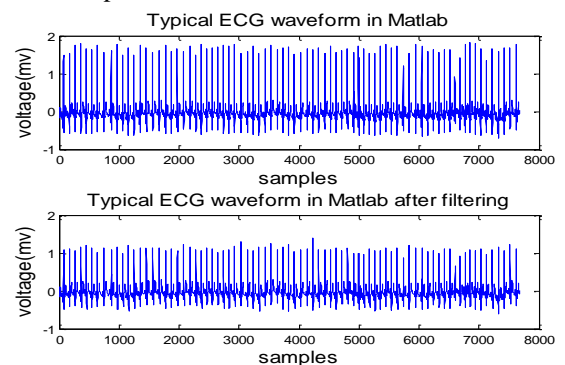


Fig. 7. Typical ECG waveform

At first this coding is used for finding out the normal value of ECG signal. Seven thousand ECG cycles are taken as typical ecg signal. Fig 7 represents a snapshot of normal ecg signal and Fig 8 represents ECG waveform considered as one segment.

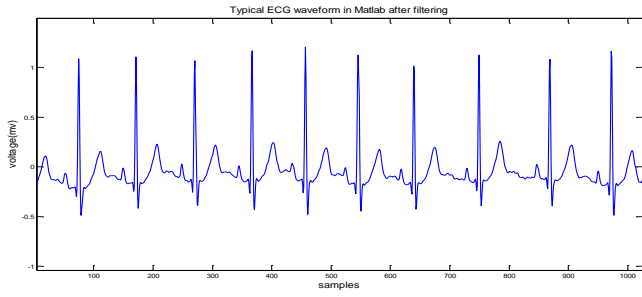


Fig. 8. ECG waveform of one segment

Figure 10 is a snapshot of the ECG signal with ventricular tachyarrhythmia. If we take one thousand samples, then we get Fig11 and it is clearly different than the normal ECG signal. Figure 12 represents the statistics value of ventricular tachycardia. The figure clearly shows remarkable difference in the M_t and D_t .

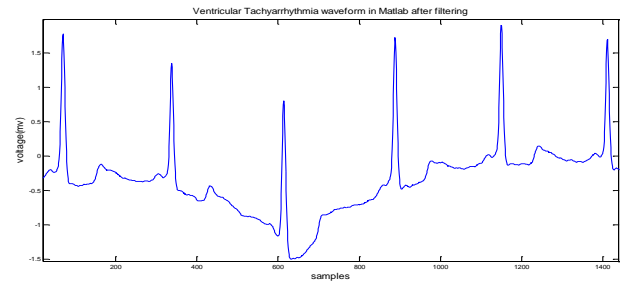


Fig. 11. ECG signal of Ventricular Tachyarrhythmia of one segment

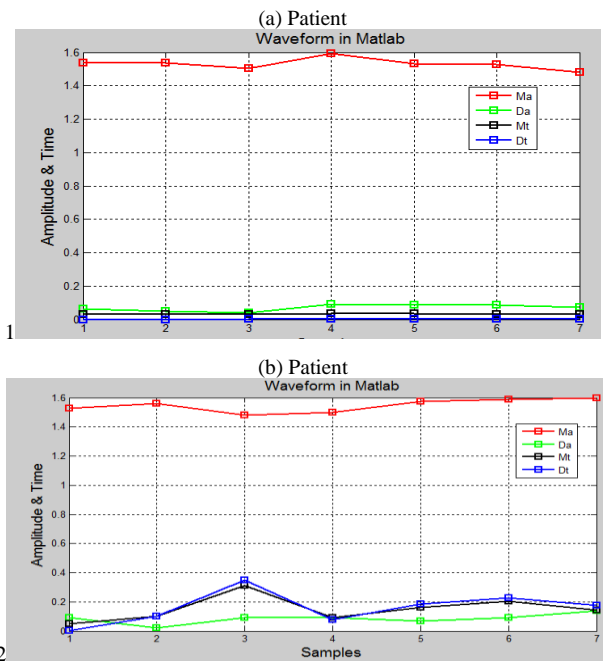


Fig. 9. Statistical parameters of 2 patients

Fig 9 represents the value of mean amplitude, amplitude deviation, mean time and time deviation of normal ECG signals of 2 different persons. As the goal is to analyze the different types of heart beat to find out the abnormalities, the value of ventricular tachyarrhythmia are taken to analyze the data. Fifteen thousand samples are taken to analyze the data.

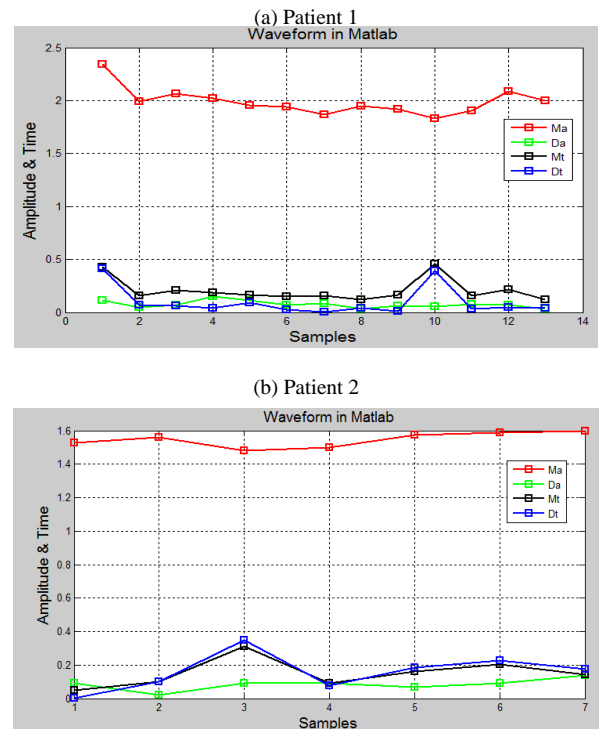


Fig. 12. SSC Parameters of Tachyarrhythmia of 2 patients

Fig 13 and Fig 14 data and segment snapshot for sleeping apnea patient. In this case the variation is not visible in naked eyes. The signal looks similar as the normal ECG signal.

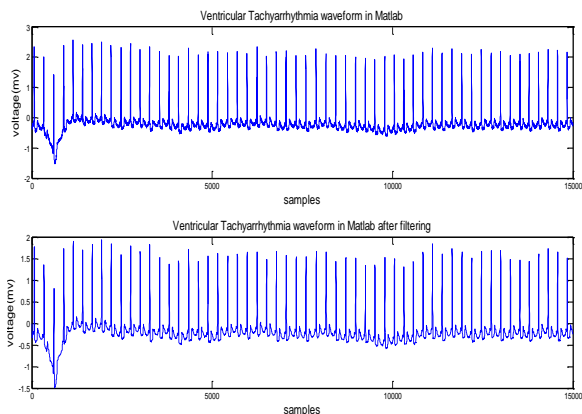


Fig. 10. ECG signal of Ventricular Tachyarrhythmia

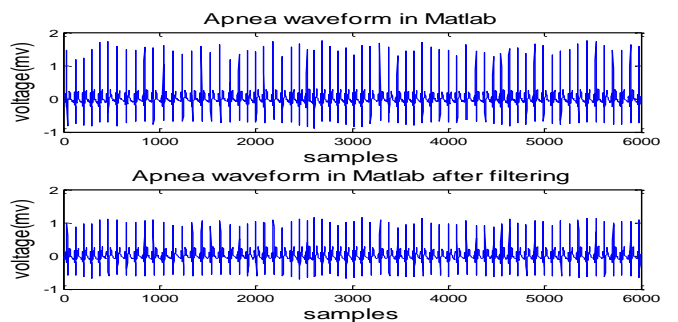


Fig. 13. ECG signal of Apnea

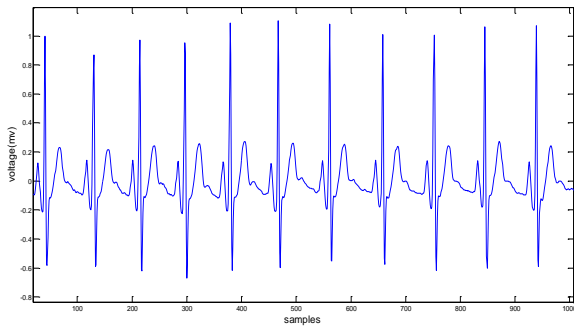


Fig. 14. ECG signal of Apnea in one sample

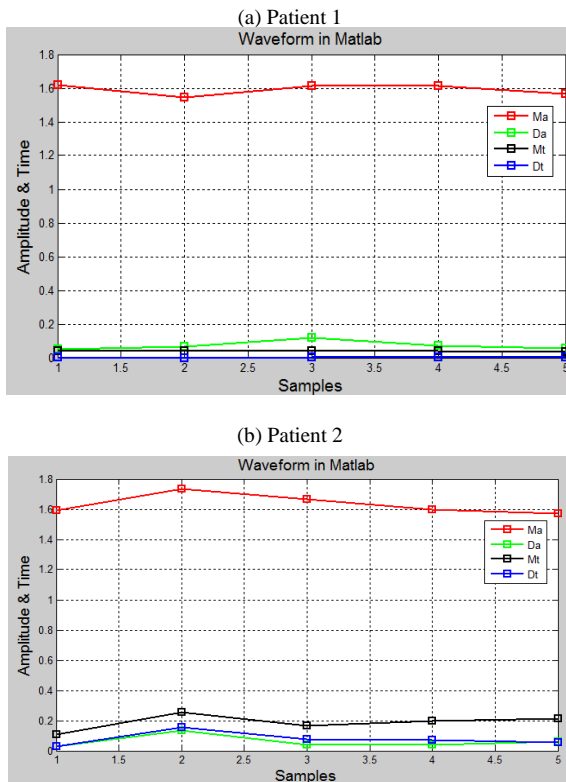


Fig. 15. SSC Parameters of Apnea for 2 patients

Figure 15, represents the SSC parameters of two patients with sleeping apnea. Here, we can see the distinctive pattern in the D_a , M_t and D_t . Hence, we may conclude that the SSC parameters are enough to distinguish among these 3 heart conditions. This technique is repeated to multiple patients for classification results. Finally the classification may be performed by using the j48 decision tree. As the parameters are quite unique for different heart conditions, any decision tree should provide similar performance.

A decision tree is a prediction tool and a data mining technique. Data mining process is used to extract information from a data set and convert it into an understandable model. All the branches between the internal nodes provide us the information about the possible values. Predicted attribute is called the dependent variable. Also there is independent variable in dataset. This independent variable helps in predicting the value of the dependent variable [7].

J48 decision tree classifier is the WEKA development based on simple algorithm of c4.5 developed by Ross Quinlan. At first we need to create a decision tree based which are

available in training data set in order to classify a new item by using j48 algorithm [7]. In this technique a tree is constructed to model for the classification system [8]. J48 algorithm system gives higher classification accuracy.

By checking the entire respective attribute data set and their values with those present in the decision tree model, we can predict the target value of these new instances [8-9].

By using weka tool and j48 algorithm with SSC parameters we get decision tree shown in fig 16. From this visualization tree we can predict the heart condition automatically. When mean amplitude is greater than 1.7929 it detects as a tachyarrhythmia condition. When mean amplitude is less than or equal to 1.7929, then it creates another two branches of time deviation. If the value of D_t is greater than 0.124 then it detects the normal ecg signal and if less than or equal to 0.124 it again create two branches of time deviation. If the value of D_t is greater than 0.0025 then it is apnea signal otherwise it is normal ECG signal.

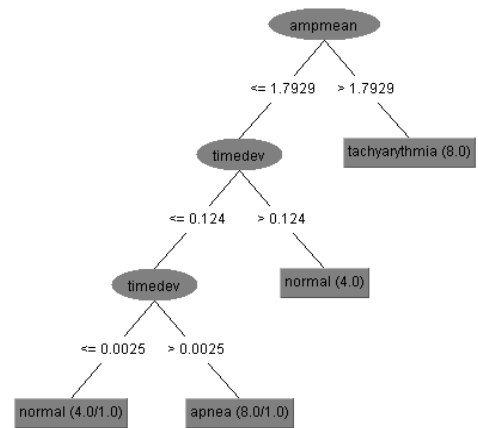


Fig. 16. Generated decision tree

Figure 17 shows the classification results achieved for three heart conditions. We used the proposed method for 100 samples and observed the results shown. The system can classify the normal ECG and ventricular tachyarrhythmia with high accuracy. In case of sleep apnea the parameter values are quite similar with the normal ecg signal. Hence 40% of the samples are classified as normal.

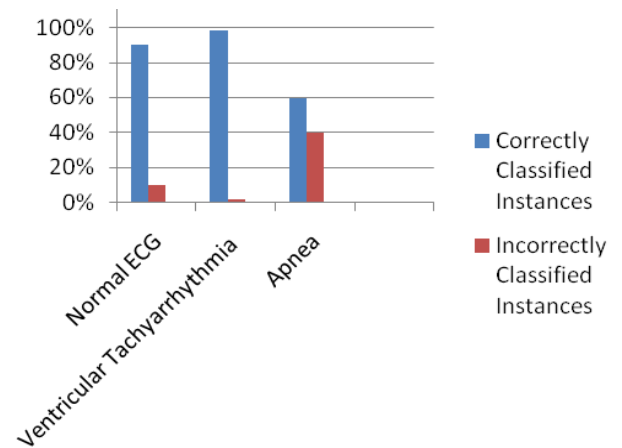


Fig. 17. Comparisons of ECG signal with its classification instances

VI. CONCLUSIONS

Electrocardiogram (ECG) analysis is one of the most common clinical cardiac examinations and it is a useful detection tool for several cardiac abnormalities. In previous so many researcher works on statistical signal characterization to analyses ECG signal. But in this thesis statistical signal is used to develop a decision tree based algorithm which is really a new technique of automatic classification of ECG signal. In this work, the proposed method aims to analyze and classify the ECG signal and detect the heart conditions. We found normal ecg signals are 90% correctly classified and for ventricular tachyarrhythmia the rate is almost 98%. Apnea signals are almost correctly classified 60%. As the parameters value of Apnea and normal ECG are quite close 40% of Apnea signals are classified as normal. This work may provide very good initial results for the SSC parameters that will help increasing the accuracy of classification algorithms. In future ECG signal classification accuracy can be increased by incorporating additional statistical parameters.

ACKNOWLEDGEMENT

We would also like to extend our heartfelt gratitude to our external supervisor **Zeenat Afroze**, Assistant Professor, Faculty of Engineering, American International University - Bangladesh, for her support and enthusiasm. We express thanks and acknowledge the support given by department of Electrical and Electronics Engineering American International university- Bangladesh.

REFERENCES

- [1] Ojha, Durgesh Kumar, and Monica Subashini. "Analysis of electrocardiograph (ecg) signal for the detection of abnormalities using matlab." World Academy of Science, Engineering and Technology, International Journal of Medical, Health, Pharmaceutical and Biomedical Engineering 8.2 (2014): 114-117.
- [2] K .O. Gupta and Dr. P. N. Chatur, *ECG Signal Analysis and Classification using Data Mining and Artificial Neural Networks*, International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 1, January 2012
- [3] Saxena, S. C., A. Sharma, and S. C. Chaudhary. "Data compression and feature extraction of ECG signals." International Journal of Systems Science 28.5 (1997): 483-498.
- [4] Hirsch, Herbert L. "Statistical signal characterization-new help for real-time processing." Aerospace and Electronics Conference, 1992. NAECON 1992., Proceedings of the IEEE 1992 National. IEEE, 1992.
- [5] Islam, M. K., et al. "Study and analysis of ecg signal using matlab&labview as effective tools." International journal of Computer and Electrical engineering 4.3 (2012): 404.

- [6] Carr, Joseph J., and John Michael Brown. Introduction to biomedical equipment technology. Prentice Hall, 1993.
- [7] "Classification Methods", [online] (Accessed on October 4, 2016) Available:http://www.d.umn.edu/~padhy005/
- [8] Chapter5.html. Last Access: December 12, 2016.
- [9] Minas, Graca, Reinoud F. Wolffenbuttel, and Jose H. Correia. "MCM-based microlaboratory for simultaneous measurement of several biochemical parameters by spectrophotometry." Biomedical microdevices 12.4 (2010): 727-736.
- [10] Babikier, MakkiAkasha, et al. "Classification of cardiac arrhythmias using machine learning techniques based on ECG signal matching." (2011).
- [11] ECG features of hyperkalaemia LITFL ECG Library[online] (Accessed on Jul 4, 2016) Available:http://lifeinthefastlane.com/ecglibrary/basics/hyperkalaemia/
- [12] Carr, Joseph J., and John Michael Brown. Introduction to biomedical equipment technology. Prentice Hall, 1993.
- [13] Rao, P. Tirumala, et al. "Distinguishing normal and abnormal ECG signal." Indian Journal of Science and Technology 9.10 (2016).
- [14] Kasar, Smita, Abhilasha Mishra, and Madhuri Joshi. "Performance of Digital filters for noise removal from ECG signals in Time domain." International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering 2.4 (2014).



Asma Haque is a graduate student of Masters in Electrical and Electronic Engineering (MEEE), Dept. of EEE. She received her B.Sc. degree in Electrical and Electronic Engineering from American International University-Bangladesh. Her research interests include signal processing and communication.



Md. Abdur Rahman received his PhD in wireless communications from Tokyo Institute of Technology, Japan in 2013. He received BSc Engg. and M.E. degrees from American International University Bangladesh (AIUB) and Asian Institute of Technology (AIT), Thailand, respectively. He joined AIUB as a lecturer in 2002, and currently employed as Professor and Director of the Faculty of Engineering in the same university.

His recent research interests include cognitive radio, bioinformatics, wireless systems, ICT etc. At present, he is conducting postdoctoral research in the School of Engg. and IT at Federation University Australia. He is a member of IEB and IEEE. During his doctoral studies he lead a team of students formed by both undergraduate and postgraduate students. The team took part in four international research competitions and won four prizes.