

DATA STREAM MINING FOR HEALTH CARE APPLICATION

Satish Thombre*

Snehlata Dongre*

Abstract:

ECG is an electric signal which is generated from human heart. It is used for investigate some of abnormal heart function. Data stream mining plays a key role to analyze the continuous data. The effective and efficient analysis of this data in different forms becomes a challenging task. Developments in sensors device, miniaturization of low power micro- electronics device, and wireless networks devices are becoming a significant opportunity for good quality of health care services. Signals like ECG, EEG, and BP etc. can be monitor through wireless sensor networks and analyzed with the help of data mining techniques. These real-time signals are continuous in nature and abruptly changing hence there is a need to apply an efficient data stream mining techniques for taking intelligent health care decisions. The high speed and large amount of data set in data stream, the traditional classifier and classification technologies are no more time applicable. The important criteria are to solve the 'concept drift' in data streams mining. Data Stream Mining (DSM) process performs data analysis and may uncover important data patterns, knowledge bases data, and scientific and medical research data. To extract the shape of ECG, the discrete wavelet transform is used after filter was applied to remove noise from ECG signal. After that detect the PQRST wave and used the ensemble classification technique to classify the ECG signal.

Keywords: *concept drifts, Data Stream Mining.*

*Department of Computer Science and Engineering, G.H.Raisoni College of Engineering, Nagpur, Maharashtra, India

1. Introduction

The ECG is nothing but the recording of the heart's electrical activity. The deviations in the normal electrical patterns indicate various cardiac disorders. Cardiac cells, in the normal state are electrically polarized. Their inner sides are negatively charged relative to their outer sides. These cardiac cells can lose their normal negativity in a process called depolarization, which is the fundamental electrical activity of the heart [1]. This depolarization is propagated from cell to cell, producing a wave of depolarization that can be transmitted across the entire heart. This wave of depolarization produces a flow of electric current and it can be detected by keeping the electrodes on the surface of the body. Once the depolarization is complete, the cardiac cells are able to restore their normal polarity by a process called re-polarization. This is also sensed by the electrodes [2]. Figure 1 describes human ECG signal for one cardiac cycle.

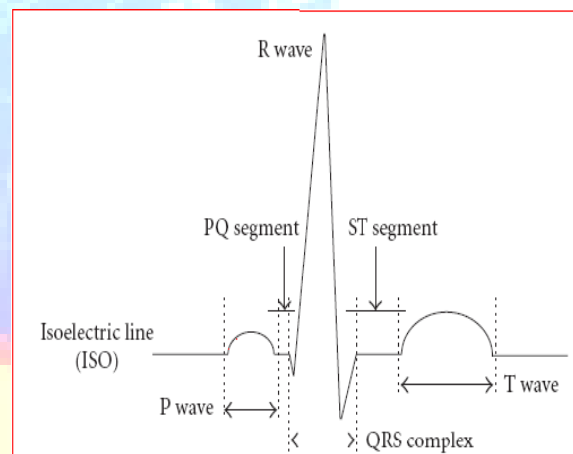


Figure 1 Human ECG signal over one cardiac cycle

Atrial activity

ECG signal is shaped by P wave, QRS complex, and T wave. In the normal ECG beat, the main parameters, including shape, duration, R-R interval, and the relationship between P wave, QRS complex, and T wave components are inspected. Traditionally, the ECG cycle is labeled using the letters P, QRS complex, and T for the individual peaks of the whole cycle's waveform. The QRS complex is useful in diagnosis of cardiac arrhythmias [3], conduction abnormalities, ventricular hypertrophy, myocardial infection, and other disease states [4]. The

electrocardiogram is a graphic recording or display of the time variant voltages produced by the myocardium during the cardiac cycle. The P-, QRS- and T-waves reflect the rhythmic electrical depolarization and repolarization of the myocardium associated with the contractions of the atria and ventricles. This ECG is used clinically in diagnosing various abnormalities and conditions associated with the heart.

Amplitude P-wave — 0.25 mV

R-wave — 1.60 mV

Q-wave — 25% R wave

T-wave — 0.1 to 0.5 mV

Duration P-R interval: 0.12 to 0.20s

Q-T interval: 0.35 to 0.44 s

S-T interval: 0.05 to 0.15 s

P-wave interval: 0.11 s

QRS interval: 0.09 s

The QRS complex represents ventricular depolarization. The QRS rate is classified as:

- 1 Normal – 60–100/minute
- 2 Bradycardic (slow) – < 60/minute
- 3 Severely bradycardic – < 40/minute
- 4 Tachycardic (fast) – between 150 and 200/minute

If the cycles are not evenly spaced, an arrhythmia may be indicated. If the P-R interval is greater than 0.2 seconds, it may suggest blockage of the AV node.

The 'P' wave represents normal atrial activity.

- 1 Are the P waves absent?
- 2 Are they all the same shape and morphology?
- 3 Is the pattern irregular?
- 4 Is there any atrial fibrillation?
- 5 Are there any atrial flutter waves (saw toothed in shape and at a rate of 300 per minute)?

Atrio-ventricular relationship

- 1 Is there 1:1 conduction (i.e. one P wave to one QRS complex), or is it 2:1 or 3:1?

2 Is the P-R interval normal, i.e. between 0.12 and 0.2 seconds (three-to five small squares) and a constant length throughout the rhythm?

2. Wavelet Transform

The wavelet transform is a convolution of the wavelet function $\psi(t)$ with the signal $x(t)$. Orthonormal dyadic discrete wavelets are associated with scaling functions $\varphi(t)$. The scaling function can be convolved with the signal to produce approximation coefficients S . The discrete wavelet transform (DWT) can

be written as

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\psi_{m,n}(t)dt. \quad (1)$$

By choosing an orthonormal wavelet basis $\psi_{m,n}(t)$ We can reconstruct the original [5]. The approximation coefficient of the signal at the scale m and location n can be written as

$$S_{m,n} = \int_{-\infty}^{\infty} x(t)\varphi_{m,n}(t)dt. \quad (2)$$

But the discrete input signal is of finite length N . So the range of scales that can be investigated is $0 < m < M$. Hence a discrete approximation of the signal can be written as

$$x_0(t) = x_M(t) + \sum_{m=1}^M d_m(t), \quad (3)$$

Where the mean signal approximation at scale M is $x_M(t) = S_{M,n}\varphi_{M,n}(t)$ and detail signal approximation corresponding to scale m , for finite length signal is given by

$$d_m(t) = \sum_{n=0}^{M-m} T_{m,n} \psi_{m,n}(t). \quad (4)$$

The signal approximation at a specific scale is a combination of the approximation and detail at the next lower scale.

$$x_{m-1}(t) = x_m(t) - d_m(t) \quad (5)$$

In the present work Daubechies [6] wavelet is chosen although the Daubechies algorithm is conceptually more complex and has a slightly complicated computation, yet this algorithm picks up minute detail that is missed by other wavelet algorithms, like Haar wavelet algorithm.

3. Wavelet Analysis

The wavelet analysis of ECG signal is performed using MATLAB software. MATLAB is a high performance interactive system which allows solving many technical computing problems. The MATLAB software package is provided with wavelet tool box. It is a collection of functions built on the MATLAB technical computing environment. It provides tools for the analysis and synthesis of signals and images using wavelets and wavelet packets within the MATLAB domain.

The normal ECG wave form and the waveforms with Wavelet Transform are shown in Figures 2, 3, 4 and 5.

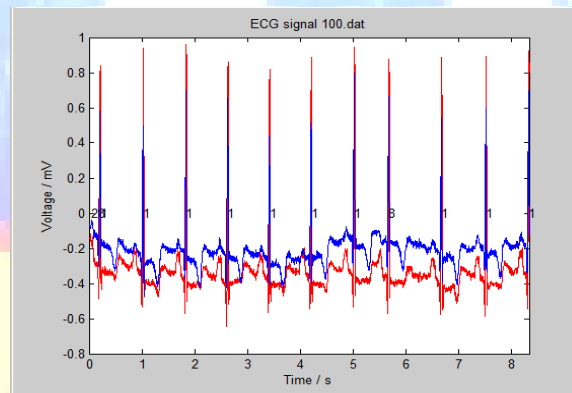


Figure 2 Normal ECG Signal

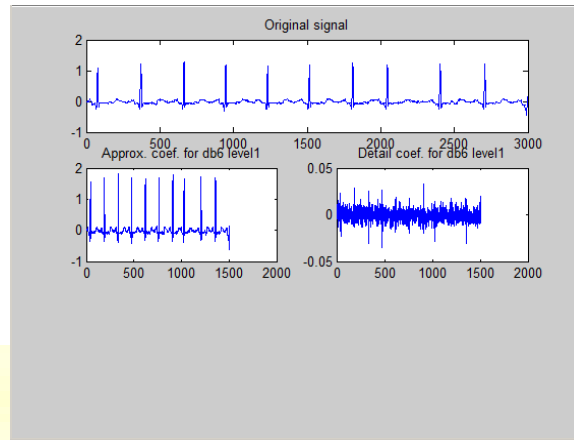


Figure 3 ECG Signal after Wavelet Transform at level 1

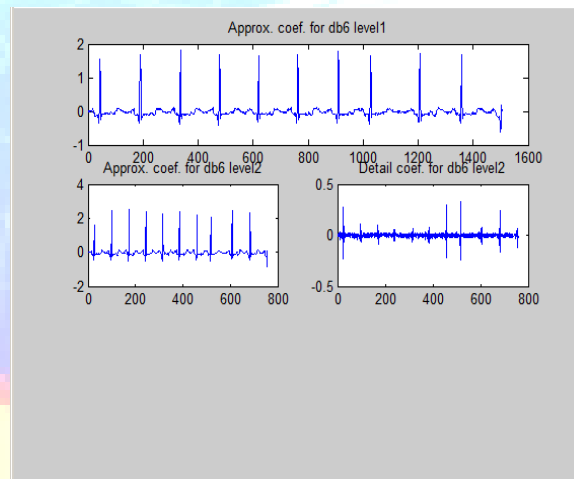


Figure 4 ECG Signal after Wavelet Transform at level 2

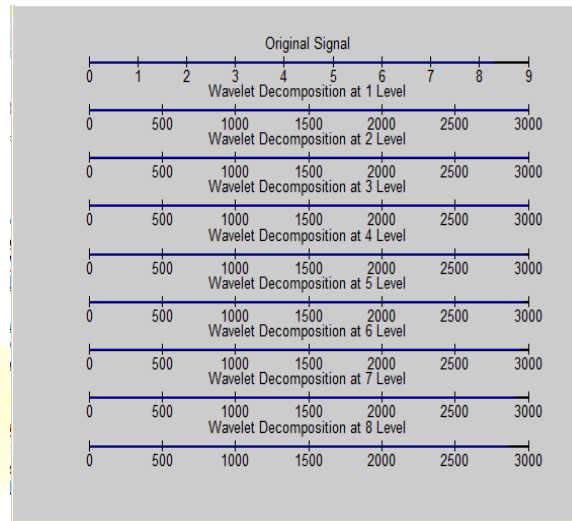


Figure 5 ECG Signal after Wavelet Transform at level 1 to 8

4. QRS Detection

Peak analysis in an ECG (Electro-cardiogram) signal. ECG is a measure of electrical activity of the heart over time. The signal is measured by electrodes attached to the skin and is sensitive to disturbances such as power source interference and noises due to movement artifacts. To find the QRS-complex which is the most prominent repeating peak in the ECG signal? The QRS-complex corresponds to the depolarization of the right and left ventricles of the human heart. It can be used to determine a patient's cardiac rate or predict abnormalities in heart function. The following figure 6 shows the Peak R value in an ECG signal. The R-waves can be detected by thresholding peaks above 0.5mV. Then next detection of the QS-waves, for S-wave detection find the local minima in the signal and apply thresholds appropriately. The local minima can be detected by finding peaks on an inverted version of the original signal. Next, we try and determine the locations of the Q-waves. Thresholding the peaks to locate the Q-waves results in detection of unwanted peaks as the Q-waves are buried in noise. Then filter the signal first and then find the peaks. The following plot shows the Q-waves and S-waves detected in the signal.

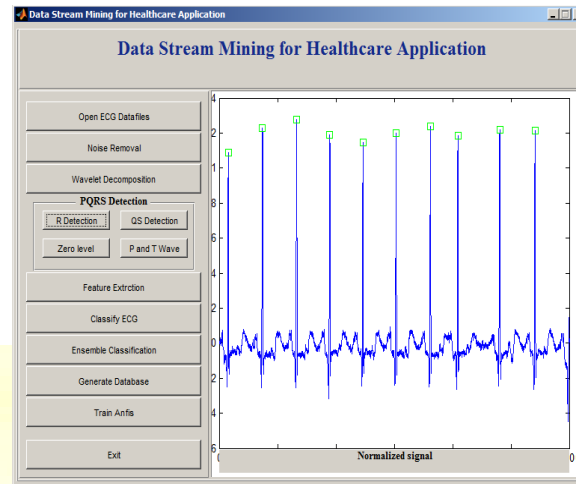


Figure 6 shows R value Detection

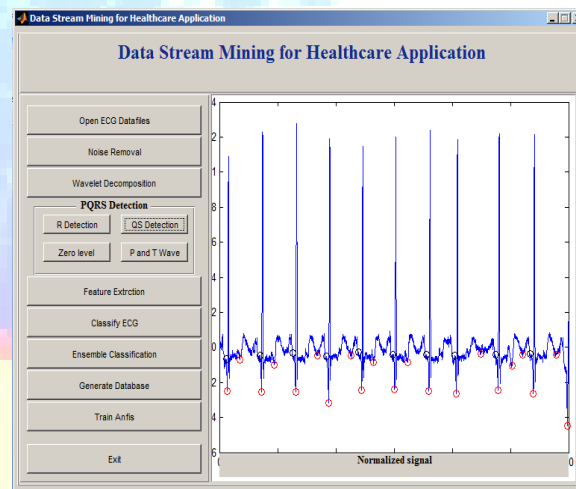


Figure 7 shows QS wave Detection

We perform peak detection on the smooth signal and use logical indexing to find the locations of the Q-waves

5. Fuzzy- Base Classifiers

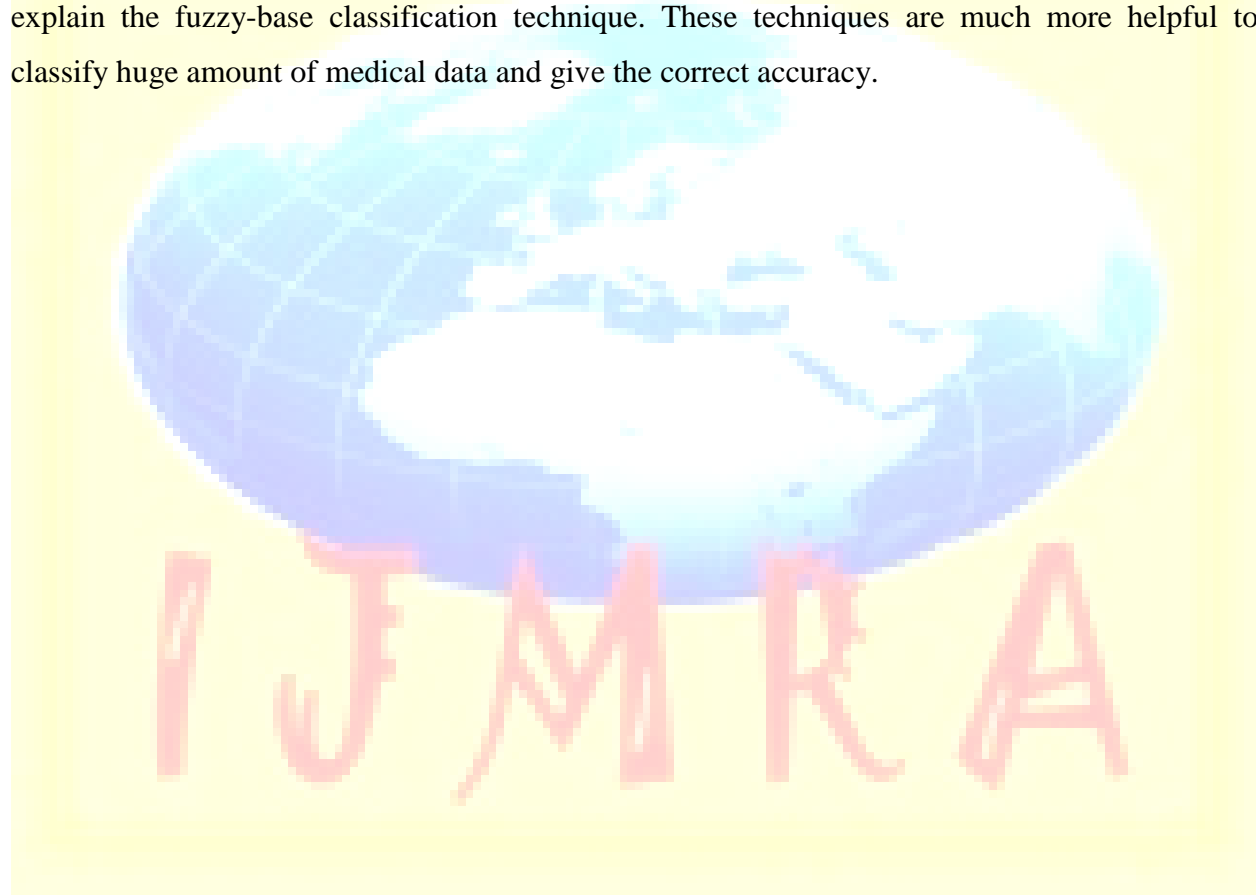
Fuzzy-based classifiers classify medical data by using a collection of “if . . . then . . .” rules. The rule antecedent or condition is an expression made of attribute conjunctions.

The rule consequent is a positive or negative classification. In order to build a rule-based classifier for medical data then follow a direct method to extract rules directly from data. The

advantages of rule-based classifiers are that they are extremely expressive since they are symbolic. But it is difficult for experts to transfer their knowledge into distinct rules, and it needs many rules to make system effectively

6. CONCLUSION

The ECG signal classification technique for data stream mining based on data mining applications. This paper introduced the wavelet transform and wavelet analysis for ECG signal, it also explain the detection of peak value. Here explain the detection of QRS complex. And also explain the fuzzy-base classification technique. These techniques are much more helpful to classify huge amount of medical data and give the correct accuracy.



Reference

- [1] Dipti D. Patil, Shamla Mantri, Himangi Pande, V.M.Wadhai Feature Extraction Techniques For Mining ECG Signals In WBAN For Healthcare Applications International Journal of Advances in Computing and Information Researches ISSN:2277-4068, Volume 1– No.1, January 201
- [2] L. Cromwell, F.J. Weibell, E.A. Pfeiffer (2005) Biomedical Instrumentation and Measurements, Prentice Hall of India, New Delhi.
- [3] Berdakh A, Seo HD. A new QRS detection method using wavelet and artificial neural networks. J Med Syst. DOI 10.1007/s10916-009-9405-3
- [4] Li CC, Zheng CC, Tai. Detection of ECG characteristic points using wavelet transforms. IEEE T Bio-Med Eng. 1995; 42: 21-8.
- [5] J.S. Sahambi, S.N. Tandon and R.K.P. Bhatt (1997) *IEEE Eng. Med. Biol. Mag.* 16 77-83.
- [6] I. Daubechies (1992) *Ten Lectures on Wavelets*, CBMS-NSF Lecture Notes nr.61,SIAM,Philadelphia.