Prediction of Cardiac Arrhythmia using Artificial Neural Network

J. P. Kelwade
B. D. College of Engineering,
RTM University, Nagpur, India

S. S. Salankar
G. H. Raisoni College of Engineering,
RTM University, Nagpur, India

ABSTRACT

Heart Rate Variability (HRV) is a non-invasive way of analysis for estimating the autonomic nervous system and based on measurement of the variability of heart rate signals. HRV signal is mostly noted for cardiac arrhythmia detection and classification. In this paper, artificial neural network (ANN) is used as a classifier to predict cardiac arrhythmias into five classes. The HRV data, RR interval time series is obtained using the Electrocardiogram (ECG) data from the MIT-BIH Arrhythmia Database. The linear and nonlinear parameters are extracted from RR intervals are used to train Multi-Layer Perceptron (MLP) neural network. In our study of neural network for time series i.e. RR interval time series, the prediction of Normal Sinus Rhythm (NSR), Premature Ventricular Contraction (PVC), Atrial fibrillation (AFIB), Left bundle branch block (LBBB) and Second degree block (BII) can be done using proposed algorithm. The 70% of the datasets are used to train MLP neural network. The proposed algorithm is predicted with next 30% of the datasets and satisfactory results obtained with prediction overall accuracy of 97%.

General Terms

HRV, ECG, Prediction

Keywords

Heart Rate Variability, Electrocardiogram, artificial neural network, arrhythmia, prediction

1. INTRODUCTION

Physiological time series are challenging. It requires methods for signal artifacts and providing temporal dynamics (non-stationarity) and nonlinearity. Examples of various physiological time series include a heart rate time series, blood pressure, oxygen consumption, electromyograph (EMG), electroencephalograph (EEG), ECG, etc

Heart diseases are the important health problem and main cause of the death of the patient. However, early diagnosis and medical treatment of heart diseases can prevent sudden death of the patient. One of the tools to diagnose heart diseases is to use ECG signals. ECG is a graphical representation of the electrical potential generated by the heart and has been an invaluable diagnostic tool for the clinician [1].

ECG signals are consist of P wave, QRS complex, and T wave. The regular heartbeat and ECG signal as illustrated in fig.1 [2]. In the normal beat of a heart, the main parameters include the shape, the duration of P wave, QRS complex, T wave and R-R interval. The changes in these parameters indicate an illness of the heart that may occur by any reason. All of the irregular beat phases are generally called arrhythmia and some arrhythmias are very dangerous for patient.
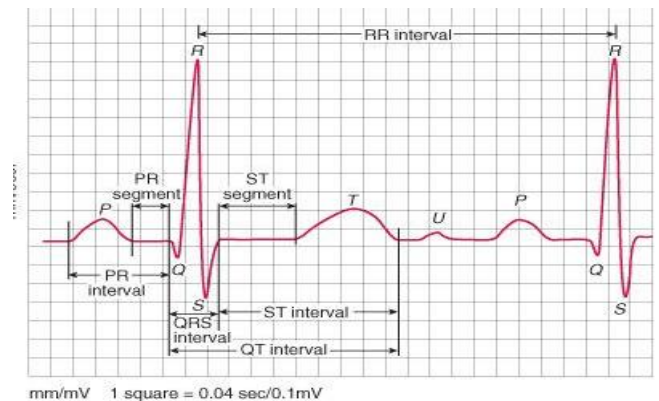

Figure 1.ECG Beats and intervals [2]

Some automatic ECG interpreting systems is available. Moreover, the computer-based interpreter systems are currently being developed to diagnose arrhythmia in time, and various methods are applied to these systems with one of them being Artificial Neural Networks (ANN).

HRV analysis is based on measuring the variability of heart rate signals i.e. variations per unit of time of the number of heartbeats It is also referred to as the RR interval since it is the time interval between consecutive R peaks of the QRS complex of the ECG [1].

2. LITERATURE REVIEW

Deshmukh Rohan et al. [2] proposed the methodology using Empirical Mode Decomposition (EMD) and neural network toolbox from MATLAB environment to classify different types of arrhythmia (Abnormal beats). Maedeh Kiani Sarkaleh et al. [3] proposed an expert system for Electrocardiogram (ECG) arrhythmia classification using discrete wavelet transform and the Multi-Layer Perceptron (MLP) neural network to classify normal and two arrhythmias.

G.K. Jaiswal and Ranbir Paul [4] explain the overall methodology for detection and classification of arrhythmias using ANN by extracting the various features and duration from ECG waveform. ECG signal and heart rate are the parameter used for detection of heart diseases.

V. Khanna et al. [5] suggests the detection and classification techniques for arrhythmic heartbeats namely Second Degree block, Premature Ventricular Contraction (PVC), Ventricular Fibrillation (VFIB), in addition to the Normal Beat using Filter Bank and Fuzzy Classifier.

Mohd Khalid Awang et al. [6] assess the application of artificial neural network in predicting the presence of heart disease mainly the angina in patients. The prediction and detection application involves Heart Disease Management Information System (HDMIS), for data collection and patient management, and Neural Network Simulator (NNS) with the

binary sigmoid activation function using back propagation neural network for training and testing.

Ali Sadr et al. [7] compared the performance of multilayer perceptron network (MLPN) and radial basis function network (RBFN) in terms of ECG signal prediction.

Rahat Abbas et al. [8] presented the Comparison of forecasting of ECG signal using Iterated method and direct method which are based on Backpropagation neural network (BPNN) model.

Sean Franklin et al. [9] described the use of artificial neural network to classify different heartbeat waveforms into 1 of 6 different categories.

X. Tang et al. [10] used rough sets (RS) and quantum neural network (QNN) to recognize electrocardiogram (ECG) signals. Using wavelet transform (WT) features are extracted and results compared with the BP and RBF.

Mehmet Engin [11] studied the application of fuzzy-hybrid neural network for electrocardiogram (ECG) beat classification by using autoregressive model coefficients; higher-order cumulant and wavelet transform variances as features. The algorithm consists of fuzzy c-means classifier and MLP neural network.

Leif Sornmo et al. [12] discusses the design of filters for the purpose of removing baseline wander and power line interference for ECG analysis.

Zarita Zainuddin et al. [13] used the radial basis function network and the wavelet neural network for estimating periodic, exponential and piecewise continuous functions.

Aaron Lewicke et al. [14] demonstrated Learning vector quantization (LVQ) neural network, multilayer perceptron (MLP) neural network, and support vector machines (SVMs) approach to reliably determine sleep and wake using only the electrocardiogram (ECG) of infants.

Jun Yao et al. [15] introduced a pruning algorithm into wavelet neural networks (WNN) for combating the problem of the gradient-descent algorithm for ECG signal classification .A. K. Pandey et al. [16] compared neural network method with various fuzzy time series methods for forecasting wheat production by recording metrological parameters.

Iffat A. Gheyas et al. [17] compared GRNN ensemble with existing algorithms (ARIMA & GARCH, MLP, GRNN with a single predictor and GRNN with multiple predictors) for forecasting univariate time series. Harleen Kaur et al. [18] presented a case study of application of data mining and analysis to data of children with Diabetes mellitus and Diabetes insipidus. Yuehui et al. [19] proposed new time-series forecasting model based on flexible neural tree (FNT). M. R. Shafiee-chafi et al. [20] presented a novel technique based on fuzzy method for chaotic nonlinear time series and validated the proposed method using Mackey-Glass, Lorenz and biological Heart Rate Variability (HRV) time series. Argyro Kampouraki et al. [21] used support vector machines (SVMs) for classification of heartbeat time series. Yüksel Ozbay et al. [22] studied Artificial Neural Networks (ANN) to classify the ECG arrhythmias.

3. METHODOLOGY
Our analysis is carried out in six stages. First ECG data acquisition step involves collecting ECG records. Second stage consists of preprocessing procedure to filter noise and interferences from the ECGs. The third stage is detection of R peaks from QRS complexes. After detection of R peaks, RR intervals time series data sets are obtained in fourth stage by segmenting into small datasets. Each dataset contains 16 RR-intervals. In the fifth stage, frequency domain and nonlinear methods are applied to extract corresponding features. In the final stage the extracted features are used to train a neural network classifier. The block diagram in figure 2 depicts the overall methodology for the arrhythmia prediction used in this study.
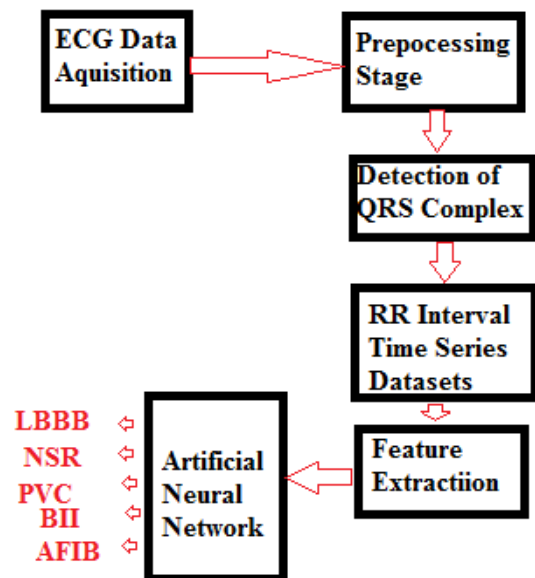

Figure 2.Overall methodology for arrhythmia prediction

3.1   ECG Data Acquisitions
MIT–BIH arrhythmia database [23] was used in this work for training and testing of designed classifier model. This is a standard database containing 48 half-hour excerpts of two channel ambulatory ECG recording, obtained from 47 subjects from 1975 and 1979 containing 360 samples per second per channel with 11 bit resolution over a 10mV range.

3.2   Data Pre-processing
The raw ECG signal contains many types of high frequency artifacts like interspersions and muscle noise and also some of un-wanted low frequency signal those arising from the skin surface and base line wander the power line interference. Figure 3 shows the raw ECG signal from MIT-BIH database. Hence they require pre-processing before actual feature extraction process. For this a band pass filter (Figure 4) with cut-off frequency from 5 to 15Hz is used.
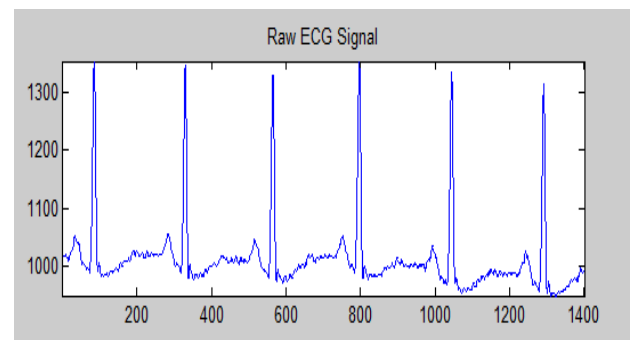

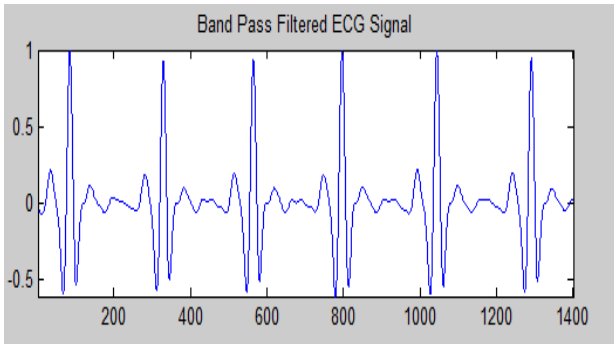**Figure 3.The raw ECG signal**

Figure 4.Band pass filtered signal

3.3 Detection of QRS Complex

In this stage of proposed algorithm, Pan and Thompkin's algorithm is applied to process ECG signal for detection of deflection points i.e. P, QRS and T waves. After filtering, the signal is differentiated to provide the QRS slope information Figure 5 depicts the differentiated signal.

After differentiating, signal is squared for making all data points positive and emphasizing the higher frequencies. The squared signal is shown in Figure 6.

After squaring, the algorithm performs sliding window integration in order to obtain the waveform feature. Signal is averaged with a moving window to get rid of noise (0.150 seconds length).

A temporal location of the QRS is marked from the rising edge of the integrated waveform. In the last step, two thresholds are adjusted. The higher of the two thresholds identifies peaks of the signal. The lower threshold is used when no peak has been detected by the higher threshold in a certain time interval. In this case, the algorithm has to search back in time for a lost peak.
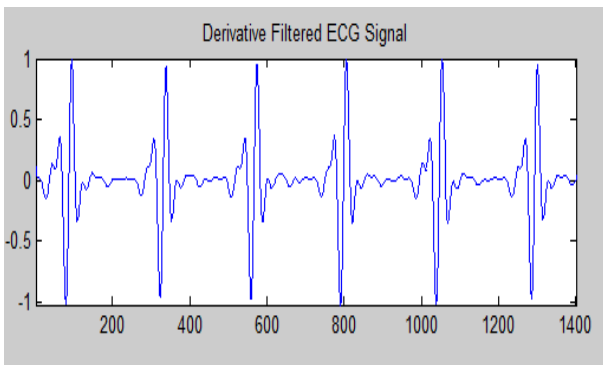

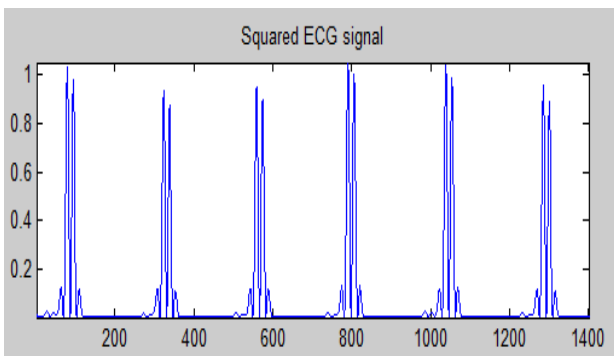**Figure 5.Differentiated signal**

Figure 6.Signal after squaring

In order to detect a QRS complex, the integration waveform and the filtered signals are investigated and different values for the aforesaid thresholds are used [25].
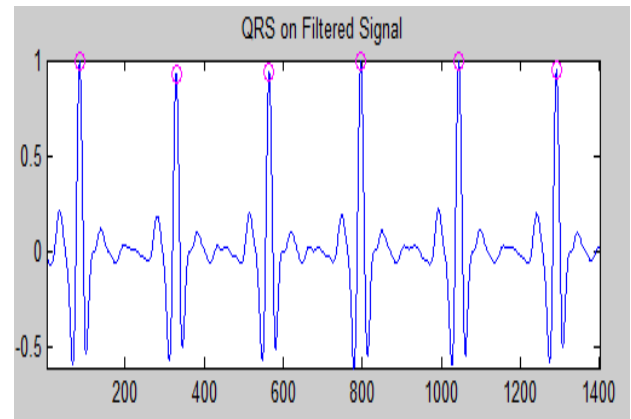

Figure 7.QRS peaks (Red circles) detection

3.4 RR Interval Time Series Datasets

After detection of R peaks, RR intervals time series is obtained. Datasets are obtained in fourth stage by segmenting sequences of RR Interval time series into small datasets. Each dataset contains 16 RR-intervals.

For this algorithm a total of 554 datasets, which are obtained from the MIT-BIH arrhythmia database, are used and it consist of 117 NSR datasets, 51 PVC datasets, 146 AFIB datasets, 154 LBBB datasets and 26 BII datasets.

3.5 Feature Extraction

Various required features can be extracted for estimation of HRV using the methods can be divided into linear (time and frequency domain) and nonlinear methods.

3.5.1 Linear Methods

3.5.1.1 Time domain analysis

The time domain methods are the simplest to perform and various parameters can be extracted from RR interval time series data sets.

3.5.1.2 Frequency domain analysis

Time domain methods are unable to discriminate between sympathetic and parasympathetic contributions of HRV. Hence time domain parameters are not used in this algorithm.

Spectral analysis is the most popular linear technique used for the estimation of HRV signals. Spectral power in the high frequency (HF) (0.15-0.4 Hz) band indicates cardiac vagal activity. Low frequency (LF) (0.04-0.15 Hz) power is related to both vagal and sympathetic systems [31]. Measurement of VLF, LF and HF power components is usually made in absolute values of power (ms2), but LF and HF may also be measured in normalized units (n. u.) which represent the relative value of each power component in proportion to the total power minus the VLF component. The representation of LF and HF in n.u. emphasizes the controlled and balanced behaviour of the two branches of the autonomic nervous system. Moreover, normalization tends to minimize the effect on the values of LF and HF components of the changes in total power [1].

In this algorithm, some of the frequency domain parameters LF/HF Ratio, auto-Regressive (AR) based Normalized units of LF power and Normalized units of HF power are used.

3.5.2 Nonlinear analysis

A cardiovascular system is complex and nonlinear in nature and by considering it as a nonlinear system can lead to better understanding of the system dynamics [32].

Four nonlinear parameters are utilized in this study.

3.5.2.1 Poincare plot geometry

Poincare plot geometry, a technique portrays nature of RR intervals fluctuation. The Poincare plot is one of tool for HRV analysis, which does not require the HRV signal to be stationary. Poincaré plot is a graphical representation of the correlation between successive RR intervals. It is a plot of R*R* (*n* +1) Vs of *RR* (*n*). The Poincaré plot may be analyzed quantitatively by calculating the standard deviations SD_1 and SD_2. SD_1 related to the fast beat-to-beat variability in the data, and SD_2 described the longer-term variability of *RR* interval time series. The ratio SD_1/SD_2 may also be computed to describe the relation between these components [32].

3.5.2.2 Spectral entropy

Spectral entropy quantifies the spectral complexity of the time series. Application of Shannon's channel entropy gives an estimate of the spectral entropy of the process, where entropy is given by

$$H = \sum_f Pf \; \log(\frac{1}{Pf})$$

Where, p_f is the probability density function (PDF) value at frequency *f*. The entropy can be used as a measure of system complexity. The spectral entropy *H* describes the complexity of the HRV [32].

3.5.2.3 Largest Lyapunov Exponent (LLE)

Detecting the presence of chaos in a dynamical system is an important problem that is solved by measuring the largest Lyapunov exponent. Lyapunov exponent λ is simply a measure of how fast two initially nearby points on a trajectory will diverge from each other as the system evolves, thus giving information about the system's dependence on initial conditions [34]. A positive Lyapunov exponent is a strong indicator of chaos. Even though an *m* dimensional system has *m* Lyapunov exponents, in most applications it is sufficient to compose only LLE. The average largest Lyapunov exponent is calculated using Wolf et al. [33] algorithm.

3.5.2.4 Detrended fluctuation analysis (DFA)

In recent years, the DFA method is a technique to determine the fractal scaling properties and the detection of correlations in noisy and non-stationary of short interval RR interval signals. DFA method is very efficient to investigate power-law of time series. This technique is a modification of the root-mean-square analysis of random walks applied to non-stationary signals. The biological signals are, in general fractal type signals. The time series are scale invariant when $x[c.n] = cH.x[n]$. The Hurst exponent (H) is an estimator of self-similarity of time series. The presence of scaling exponent can point to an inner fractal structure of time series. The constant c is a scaling coefficient (c>1, contraction and c<1, dilation). The power law exponent H, is the Hurst exponent and represents a particular kind of the scaling invariant structure. The Hurst exponent is calculated using detrended fluctuation analysis [36].

The linear (frequency domain) and non-linear parameters considered for cardiac arrhythmia prediction using ANN are.

1. Spectral entropy.

2. Poincaré plot geometry.

3. Largest Lyapunov exponent.

4. Detrended fluctuation analysis.

5. LF/HF Ratio

6. Normalized units of LF power

7. Normalized units of HF power

3.6 Artificial Neural Network Classifier

The Artificial Neural Networks (ANN) is the mathematical modeling tool for human cognition or neural biology. ANN is characterized by pattern of connections between the neurons i.e. architecture and determining the weights on the connections i.e. training, or learning algorithm. The Multi-Layer Perceptron (MLP) neural network is the most useful supervised neural network because it requires a desired output in order to learn. In our study, the MLP with two hidden layers are used. The first hidden layer with 35 nodes and second hidden layer with 25 nodes are used for correct prediction of cardiac arrhythmia. A graphical representation of an MLP is shown in Figure 8.
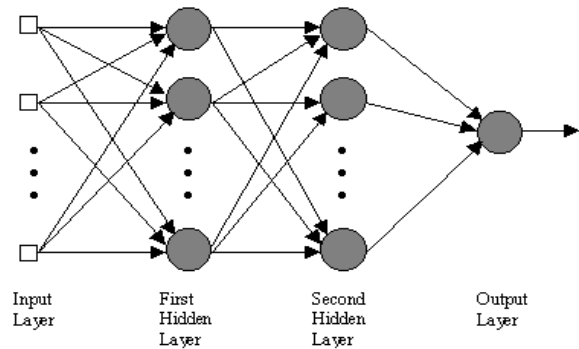

Figure 8.MLP architecture with two hidden layers [3]

This step is called the training phase. In the second step, the trained MLP is used to test using unknown input data. The back-propagation algorithm compares the result that is obtained in this step with the result that was expected. The MLP computes the error signal using the obtained output and desired output. The computed signal error is then fed back to the neural network and used to adjust the weights such that with each iteration the error decreases and the neural model gets closer and closer to produce the desired output. Figure 6 shows this process [3].
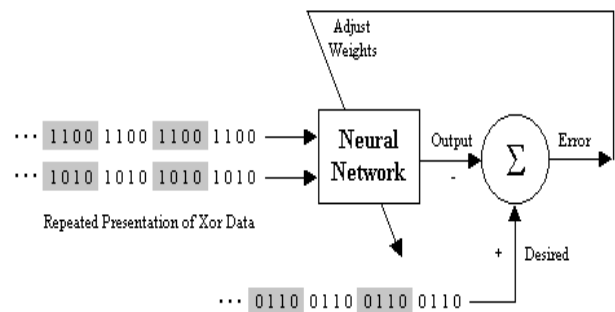

Figure 9.Neural network learning algorithm [3]

There are different training algorithms, while it is very difficult to know which training algorithm is the fastest for a given problem. In this algorithm, MLP neural network is trained with Lavenberg-Marquardt (LM)algorithm, The best

convergent rate can be obtained with the use of LM algorithm achieved the best overall detection accuracy [4].

4. RESULTS AND CONCLUSIONS

In this study of neural network for time series analysis, the heart rate time series i.e. RR interval time series the artificial neural network is used as a predictor to detect the cardiac arrhythmia into five heart diseases. It can be used as diagnostic tool to aid the physician in the analysis of health of the heart. The combination of linear parameters in frequency domains and non-linear parameters derived from the HRV signal are used as input features for the neural network. The proposed algorithm classifier showed satisfactory performances in discriminating five types of arrhythmia. The accuracy of discrimination of NSR, PVC, AFIB, LBBB and BII are 100%, 80%, 100%, 96.70%, and 100%, respectively.

Figure 10.Result of prediction of five cardiac arrhythmias

Table 1.Datasets for training and testing with accuracy

Arrhythmia classes	Training datasets	Testing datasets	Total datasets	Prediction Accuracy %
NSR	124	53	177	100%
BII	18	8	26	100%
PVC	36	15	51	80%
LBBB	108	46	154	96.7%
AFIB	102	44	146	100%
Total	388	166	554	Overall 97%

Table 2.Result of Tested Datasets for arrhythmia

Arrhythmia classes	NSR	BII	PVC	LBBB	AFIB
NSR	53				
BII		8			
PVC	2		12		1
LBBB		2		44	
AFIB					44

The algorithm is implemented in MATLAB Version R2014a. The result of prediction of arrhythmia is shown in figure 10 in confusion matrix. The results show that the proposed algorithm is effective for prediction of cardiac arrhythmias, with an accuracy of 97%. It is shown that the combined set of linear (frequency domain method) and nonlinear features provides the very good prediction. It is concluded that the HRV signal can be used as a reliable indicator of states of heart health. It can further be expanded for prediction of more number of dangerous diseases by acquiring ECG database of dangerous diseases.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Mr. Deshmukh Rohan, Dr. A. J. Patil, Layered Approach for ECG beat classification utilizing Neural Network, International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 6, November-December 2012, pp.1495-1500 1495

[2] Maedeh Kiani Sarkalehm, Asadollah Shahbahrami, Classification of ECG Arrhythmias using Discrete Wavelet Transform and Neural Networks, International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.2, No.1, February 2012

[3] Gaurav Kumar Jaiswal and Ranbir Paul. Artificial neural network for ECG classification, Recent Research in Science and Technology, Vol. 6(1), pp. 36-38, 2014

[4] V. Khanaa, K.P. Thooyamani and R.Udayakumar. Categorization and Forecast of Critical Cardiac Arrhythmias using Filter Bank and Fuzzy Classification Approach, Middle-East Journal of Scientific Research Vol. 18 (12), pp. 1798-1802, 2013

[5] Mohd Khalid Awang and Fadzilah Siraj. Utilization of an Artificial Neural Network in the Prediction of Heart Disease, International Journal of Bio-Science and Bio-Technology Vol. 5, No. 4, August, 2013

[6] Ali Sadr, Najmeh Mohsenifar, Raziyeh Sadat Okhovat. Comparison of MLP and RBF neural networks for Prediction of ECG Signals, International Journal of Computer Science and Network Security (IJCSNS), VOL.11 No.11, pp, 124-128, November 2011

[7] Rahat Abbas, Wajid aziz, Muhhamad Arif. Electrocardiogram signal forecasting using iterated and direct methods based on artificial neural network, Journal of App. Em. Sc., Vol.1. no.1, pp. 72-78, June 2004

[8] Sean Franklin, Joseph Wallcave. Cardiac Condition Detection using Artificial Neural Networks, Senior Project Electrical Engineering Department, California Polytechnic State University, San Luis Obispo, June 2013.

[9] X. Tang, L. Shu. Classification of Electrocardiogram Signals with RS and Quantum Neural Networks, International Journal of Multimedia and Ubiquitous Engineering Vol.9, No.2, pp.363-372, 2014.

[10] Mehmet Engin. ECG beat classification using neuro-fuzzy network, Elsevier, Pattern Recognition Letters 25, 1715–1722, 2004.

[11] Leif SoRnmo, Pablo Laguna. Electrocardiogram (ECG) Signal Processing, Wiley Encyclopedia of Biomedical Engineering, John Wiley & Sons, Inc. 2006

[12] Zarita Zainuddin, Ong Pauline. Function Approximation Using Artificial Neural Networks, International Journal Of Systems Applications, Engineering & Development Issue 4, Volume 1, 2007

[13] Aaron Lewicke, Edward Sazonov, Michael J. Corwin, Michael Neuman. Sleep Versus Wake Classification From Heart Rate Variability Using Computational Intelligence: Consideration of Rejection in Classification Models, IEEE Transactions On Biomedical Engineering, Vol. 55, No. 1, Pp. 108-118, January 2008

[14] Jun Yao, Qiang Gan, Xue-dong Zhang, Jin Li. Pruning Algorithm In Wavelet Neural Network For Ecg Signal Classification, Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 20, No 3, pp. 1482-1485, 1998

[15] A. K. Pandey, A. K. Sinha, and V.K. Srivastava. A comparative study of neural network and fuzzy time series forecasting Techniques –Case study: Wheat Production forecasting. International Journal of computer science and network security (IJCSNS) Vol.8 No.9, pp 382-387, September 2008

[16] Iffat A. Gheyas, Leslie S. Smith. A neural network Approach to Time Series Forecasting, Proc. of the world congress on engineering 2009, London, U. K. Vol. II July 1-3, 2009

[17] Harleen Kaur, Siri Krishan Wasan. Empirical Study on Applications of Data Mining Techniques in Healthcare, Journal of Computer science , pp 194-200, 2006

[18] Yuehui Chen, Bo Yang, Jiwen Dong, Ajith Abraham. Time series forecasting using flexible neural tree model, Elsevier, International Journal of Information Sciences, pp. 1-17, 2004

[19] M. R. Shafiee-chafi, et al. A novel fuzzy based method for heart rate variability prediction, International Journal of Engineering, Vol. 27 no. 7, pp. 1041-1050, 2014

[20] Argyro Kampouraki, George Manis, and Christophoros Nikou, *Member, IEEE,* Heartbeat Time Series Classification With Support Vector Machines, IEEE Transactions On Information Technology In Biomedicine, Vol. 13, No. 4, July 2009

[21] Yüksel Özbay and Bekir Karlik. A Recognition of ECG Arrhythmias Using Artificial Neural Networks, Proceedings – 23rd Annual Conference – IEEE/EMBS Oct.25-28, 2001, Istanbul, TURKEY

[22] MIT-BIH Arrhythmia database: http://www.physionet.org

[23] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals. org/cgi/content/full/101/23/e215]; 2000 (June 13).

[24] Pan. J, Tompkins. W.J, A Real-Time QRS Detection Algorithm, IEEE Transactions on Biomedical Engineering, Vol. Bme-32, No. 3, March 1985

[25] G. E. P. Box, G.M. Jenkins, and G. C. Reinsel. Time series analysis: Forecasting and control. Prentice-Hall Inc., USA, pp 21, 1994

[26] Haykin, S. Neural Networks - a Comprehensive Foundation, Prentice-Hall 1998.

[27] Neural Network Toolbox User's Manual, the Math Works Inc. 1992.

[28] C. K. Peng et al. (1994) Mosaic organization of DNA nucleotides, Physical Review E 49(2), 1685-1689.

[29] R. Weron (2002) Estimating long range dependence: finite sample properties and confidence intervals, Physica1 A312, 285-299.

[30] Laguna P, Jane R, Caminal P. Automatic Detection of Wave Boundaries in Multi-lead ECG Signals: Validation with the CSE Database. Computers and Biomedical Research 27(1):45-60, 1994.

[31] M. Brennan, M. Palaniswami, P. Kamen, Do Existing Measures of Poincaré Plot Geometry Reflect Nonlinear Features of Heart Rate Variability? IEEE Trans Biomedical Eng. 48(11),2001

[32] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, Choo Min Lim, Jasjit S. Suri, Heart rate variability: a review, Med Bio Eng Comput (2006) 44:1031–1051.

[33] Wolf, Jack B. Swift, Harry L. Swinney, and John A. Vastano, Determining Lyapunov exponents from a Time Series, Physica 16D (1985) 285-317

[34] B. Anuradha, V. C. Veera Reddy, ANN for Classification of Cardiac Arrhythmias, ARPN Journal of Engineering and Applied Sciences, VOL. 3, NO. 3, JUNE 2008

[35] Morton L.F. et al., Detrended fluctuation analysis of EEG signals, The 7TH International conference of interdisciplinary in Engineering, Procedia Technology 12 (2014) 125 – 132