PERFORMANCE ANALYSIS OF SOFT COMPUTING TECHNIQUES FOR CLASSIFYING CARDIAC ARRHYTHMIA

R GANESH KUMAR

Research Scholar, Department of CSE, Sathyabama University,
Chennai, Tamil Nadu 600119, INDIA
mailgk74@gmail.com

Dr. Y S KUMARASWAMY

Sr. Professor & HOD, Department of MCA, Dayananda Sagar College of Engineering,
Bangalore 560078, INDIA
yskldswamy2@yahoo.com

Abstract

Cardiovascular diseases kill more people than other diseases. Arrhythmia is a common term used for cardiac rhythm deviating from normal sinus rhythm. Many heart diseases are detected through electrocardiograms (ECG) analysis. Manual analysis of ECG is time consuming and error prone. Thus, an automated system for detecting arrhythmia in ECG signals gains importance. Features are extracted from time series ECG data with Discrete Cosine Transform (DCT) computing the distance between RR waves. The feature is the beat's extracted RR interval. Frequency domain extracted features are classified using Classification and Regression Tree (CART), Radial Basis Function (RBF), Support Vector Machine (SVM) and Multilayer Perceptron Neural Network (MLP-NN). Experiments were conducted on the MIT-BIH arrhythmia database.

Keywords: Cardiac Arrhythmia, Discrete Cosine Transform (DCT), Electrocardiograms (ECG), RR Waves, Soft Computing.

1. Introduction

Cardiovascular diseases kill more people than other diseases. Arrhythmia is a collection term for any cardiac rhythm, which deviates from the normal sinus rhythm. Arrhythmia may be due to a disturbance in impulse formation or conduction, or both, but it is not always an irregular heart activity [1]. Extracting such signal features helps to explain and identify different cardiac arrhythmias. The electrocardiogram (ECG) is an important tool to assess a patient's cardiac status, as this signal represented the difference between two points on body surface, versus time [2]. This becomes difficult as ECG signal size is large and the presence of many noise types contaminates ECG signals [3]. Also manual analysis is time consuming and error prone; thus, extraction of automatic features in ECG signals gains importance.

Many tools / algorithms were proposed for feature extraction from ECG signals like total least squares based. Most techniques involve high consumption and processing time for features extraction / classification. Another disadvantage is limited to arrhythmias which are classified. Hence, a new technique is required to classify many arrhythmias. Also, the suggested technique should be amenable for real time implementation to ensure its use in intensive care units / ECG signal collected. Feature selection is the selection of a features subset occurring in a training set and used as features in classification. Feature selection serves two things; it ensures efficiency in training and applying a classifier through decrease of effective feature size. And secondly, feature selection improves classification accuracy through noisy feature elimination. A noisy feature being one that when represented increases classification error in new data.

The position and magnitudes of the peaks in ECG such as PR interval, ST interval, QRS interval and QT interval are commonly used to classify arrhythmia [4, 5]. The features are either selected from the time domain or frequency domain [6, 7]. Several methods were suggested for cardiac arrhythmias automatic detection/classification in the literature, including: artificial immune recognition system with fuzzy weighted [8], neural networks [9], fuzzy neural networks [10], wavelet transforms [11, 12], combined wavelet transformation [13], Bayesian classifiers [14] support vector machines [15], and Markov models [16].

This paper investigates performance analysis of soft computing techniques for classifying cardiac arrhythmia using Radial Basis Function (RBF), Support Vector Machine (SVM) and Multilayer Perceptron Neural Network (MLP-NN). Experiments were conducted on the MIT-BIH arrhythmia database. Features are extracted from time series ECG data with Discrete Cosine Transform (DCT) computing the distance between RR waves. The feature is the beat's extracted RR interval. The rest of the paper is organized as follows: Section 2 reviews related works available in the literature. Section 3 details the methods and materials used. Section 4 reports the results and Section 5 concludes the paper.

2. Related Works

Karaolis, et al., [17] investigated data-mining systems to assess heart related risk factors and to CHD events. A total of 528 cases from Paphos district in Cyprus was collected for examination, a majority of them being affected with an event. Data-mining analysis was done through the use of c4.5 decision tree algorithm, for the three events applying five different splitting criteria mentioned above. Risk factors considered for classification rules analysis include, 1) MI, age, history of hypertension and smoking; 2) PCI family, hypertension and diabetes histories; 3) for CABG, age, smoking and hypertension history. These risk factors were also extracted by other examiners. For the MI, PCI, and CABG models, classification efficiency was 66%, 75%, and 75% respectively. Hence, data-mining system recognizes subjects high and low risk subgroups, and also decides on therapy selection; (i.e.,) whether medical or surgical.

Kampouraki, et al., [18] examined support vector machines (SVMs) to benefit heartbeat time series (i.e., heart rate signals) classification. Statistical methods and signal analysis techniques extracted signal features. By comparing other neural network-based classification methods, SVM classifier performed eave-one-out type cross validation. Regarding other state-of-the-art classifiers, SVM classifies heart rate signal with reduced low signal-to-noise ratio. Examination was undertaken for two data sets with many features. The first dataset included long-term ECG recordings of juvenile and matured healthy subjects while the second set included long-term ECG recordings of normal subjects classification when compared to the common Heart Rate Variability (HRV) analysis which was unable to classify similar signs correctly. The issue of feature selection is considered in feature space by the classifier.

Pasolli, et al., [19] suggested 3 learning strategies to classify ECG Signals by first selecting from a small, suboptimal training with learning strategies then selecting beat samples additionally from unlabeled large set data. Samples are labeled manually and added to the needed training set. The entire process is iterated till the final training set is constructed. Based on support vector machine classification, the proposed strategies are developed on margin sampling, posterior probability and query by committee principles. All proposed strategies show high performance regarding stability and accuracy in relation to a completely random selection strategy (R).

Faezipour, et al., [20] presented a repetition-detection concept in a patient-adaptive cardiac profiling system, where a wavelet-based beat-detection mechanism first extracts fiducial ECG points following which a new local ECG beat classifier profiles patients normal cardiac behavior. Beat detection is a hybrid of Pan and Tompkins algorithm and wavelet analysis approach. Experiments were undertaken on a MIT-BIH arrhythmia database proving that this system was able to detect beats with 99.5% accuracy and identify abnormalities with 97.42% accuracy.

Kim, et al., [21] proposed an ECG holter system, a signal processing technique comprising of a three-step compression and classification flows with Quad Level Vector (QLV). QLV improved performance with low-computation in ECG processing for both compression and classification flows. Classification was applied to both heartbeat segmentation and R-peak detection. The algorithm's reliability is upped through a noise robust test. Average compression ratio was 16.9:1 with 0.641% percentage root mean square difference value and a 6.4 kbps encoding rate. Without noise, R-peak detection accuracy performance is 100% and with it, -10dB SNR noise 95.63%. Overall processing cost reduction was by 45.3% through this method.

Sufi, et al., [22] presented a new procedure executing real-time CVD (cardiovascular disease) classification. This real-time CVD classification automatically informs emergency personnel/hospital through SMS/MMS/e-mail, when a CVD patient has a life-threatening cardiac abnormality. The method uses data mining techniques, like attribute selection (only selects features from compressed ECG) and expectation maximization (EM)-based clustering. Data mining techniques in a hospital server bear constraints in representing each abnormality. A patient's mobile gets these constraints and a rule-based system that recognizes abnormal beats in real time is used. Experiments undertaken on 50 MIT-BIH ECG entries for cardiac abnormality detection (ventricular flutter/fibrillation, premature ventricular contraction, atrial fibrillation etc.) successfully ended up with 97% average accuracy.

3. Materials and Methods

3.1. *MIT-BIH Arrhythmia Database*

For training and testing of ECG datasets signals for evaluating the performance of the classifiers, Massachusetts Institute of Technology/Beth Israel Hospital (MIT-BIH) arrhythmia database is used [23]. Cardiac Arrhythmias for classification include LBBB (Left Bundle Branch Block), RBBB (Right Bundle Branch Block) and normal beats. Each ECG beat used is a matrix (275x1) for one ECG lead with every ECG lead with every ECG signal having five distinct points (P, Q, R, S and T) used for ECG interpretation.

3.2. *R-R Interval*

It is known that the morphology of the QRS complex along with the instantaneous RR interval (interval between two successive R peaks) play an important role in the heart diseases diagnosis [24]. For the classification of an ECG beat the ratio of the RR interval before it to the one after it is a useful feature. The peaks are detected as follows:

- The moving average of the signal is calculated using a number of records.
- New signal is derived by subtracting the moving average from the original signal.
- Peak of the signal 'R' is found.
- Peaks of P, Q, R, S and T are found by the relative position.

The peak amplitude is measured from k line. The k line is given by:

$$k = max(\theta_i, i = 1, 2, \ldots, 11) + c \tag{1}$$

Where θ is the greatest amplitude, type of heartbeat and c is a constant. Figure 1 shows a typical ECG with various peaks and the RR interval.

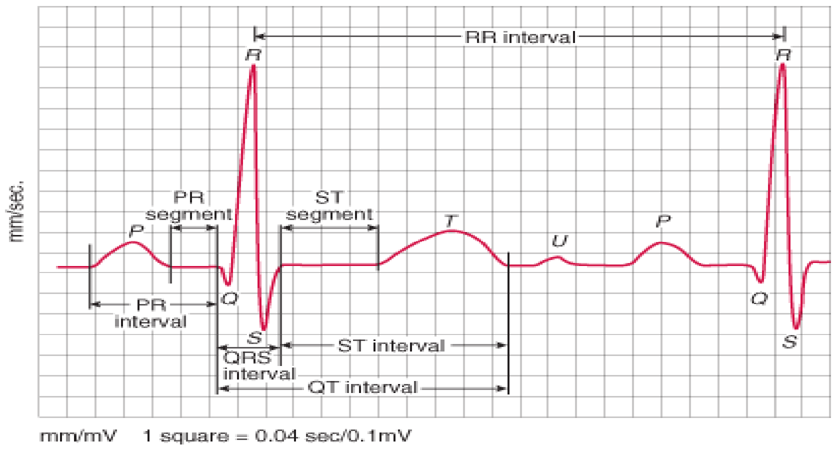

Fig. 1. ECG with R-R interval pattern

3.3. *Discrete Cosine Transform (DCT)*

Discrete Cosine Transform (DCT) translates time series signal into basic frequency components. The data is pre-processed to find RR intervals using Fast DCT. The FDCT [25], of a lsit of *n* real number, *s(x), x = 0,…., n-1*, is the list of length *n* given by the following equation:

$$S(u) = \sqrt{2/n}\, C(u) \sum_{x=0}^{n-1} s(x) \cos \frac{(2x+1)u\pi}{2n} \tag{2}$$

where $C(u) = 2^{-1/2}$ for u=0 or otherwise $C(u) = 1$

The constant factors are chosen so that the basis vectors are orthogonal and normalized. The inverse cosine transform (IDCT) is given as:

$$S(x) = \sqrt{2/n}\sum_{x=0}^{n-1} C(u)s(u)\cos\frac{(2x+1)u\pi}{2n} \tag{3}$$

3.4. *Classifiers*

3.4.1. *Classification and Regression Trees (CART)*

Classification and Regression Tree (CART) based on class label produce output that is either classification or regression. If class label is categorical, CART provides a classification tree and regression tree, if it is numeric. A decision tree is created by rules formed on given data's attribute values. Each rule is based on dependent attribute values. This rule is applied to all nodes till the last node is reached, which is the class node [26].

A Classification/regression tree is a binary tree, that given an input X produces output Ÿ approximating random variable of interest Y, stochastically related to X. this deterministic mapping is as follows. Associated with the tree's every internal node is a binary function of input X and linked to each external node is specific output label Y. Starting at root node, binary function tests input X. When the result is "0", left branch is followed; if it results in "1", then right branch is followed. This is repeated till an external node/leaf is reached when associated label Y is the output. The tree minimizes (approximately) anticipated loss between Y and Ÿ.

3.4.2. *Multi-Layer Perceptron Neural Networks (MLP-NN)*

Neural networks are analytical techniques modelled analogous to learning process in the cognitive system and the brain's neurological functions. They can predict new observations from earlier ones after executing the learning process by using past data. Neural networks or widely used Multilayer Perceptron Neural Networks (MLPNNs) are computational systems comprising of highly interconnected processing elements sets, called neurons, that process information in response to external stimuli which in turn are transmitted from one processing element to another through synapses/interconnection, which are either excitatory or inhibitory. MLPNNs are used in applications like pattern recognition, and classification [27].

3.4.3. *Support Vector Machines (SVM)*

Support Vector Machine (SVM) maps input vectors to higher dimensional spaces where maximal separating hyper planes are constructed. Two parallel hyper planes are constructed on both sides of a data searching hyper plane. The separating hyper plane is that which maximizes distance between two parallel hyper planes as shown in Figure 2. It is assumed that larger the margin between parallel hyper planes, better a classifier's impact. Though SVMs were originally designed for binary classification, it can also be used in multi-class classification issues [28].

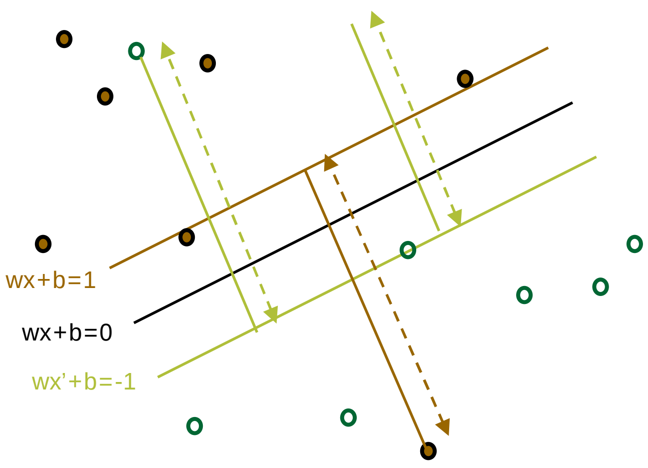

Fig. 2. Representation of Hyper planes

For a training set consisting of instances (x_i; y_i), i = 1, …., i, where $x_i \in R^n$ and $y \in \{1,-1\}^l$, the support vector machines (SVM) has to solve the following optimization problem [19]:

$$\min_{w,b,\xi} \quad \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$

Subject to $\quad y_i \left(w^T \phi(x_i) + b \right) \geq 1 - \xi_i,$

$\xi_i \geq 0$

Polynomial and Radial Basis Function (RBF) is the different kernel functions used in this study.

1] Polynomial:

$$k\,(x, x') = (x, x')^d \tag{4}$$

$$k\,(x, x') = ((x, x') + 1)^d \tag{5}$$

2] Gaussian Radial Basis Function: A RBF kernel models non-linear relation of class labels and features Radial Basis functions most commonly with a Gaussian form

$$k\left(x, x'\right) = exp\left(-\frac{\|x - x'\|^2}{2\sigma^2} \right) \tag{6}$$

4. Results and Discussion

A total of 153 images are used for evaluating the performance of classifiers. The dataset consists of 68 instances of left bunch bundle block, 30 instances of right bunch bundle block and 56 normal instances. In this paper, a method of energy extraction using Discrete Cosine Transform was used and RR interval was extracted and used as feature. Using 10 fold cross validation Classification and Regression tree (CART), Radial Basis Function (RBF), MLP and SVM was tested. Table 1 tabulates the summary of results of classification experiments conducted and the classification accuracy is given in Table 2.

Table 1.Summary of Results for various Classifiers

Properties	**CART**	**SVM RBF kernel**	**SVM Polynomial Kernel**	**MLP NN**
Correctly Classified Instances	135	141	95	135
Incorrectly Classified Instances	18	12	58	18
Kappa statistic	0.814	0.876	0.3688	0.8141
Mean absolute error	0.1079	0.0523	0.2527	0.0883
Root mean squared error	0.2688	0.2287	0.5027	0.2679
Relative absolute error	25.54 %	12.38 %	59.84 %	20.91 %
Root relative squared error	58.52 %	49.78 %	109.44 %	58.33 %
Total Number of Instances	153	153	153	153

Table 2. Classification Accuracy

Classification Method	**Classification Accuracy in %**
CART	88.2353
SVM – RBF kernel	92.16
SVM - Polynomial kernel	62.09
MLP	88.2353

The Figure 3 shows the root mean squared error (RMSE) obtained by various classifiers and the Figure 4 charts the classification accuracy achieved.

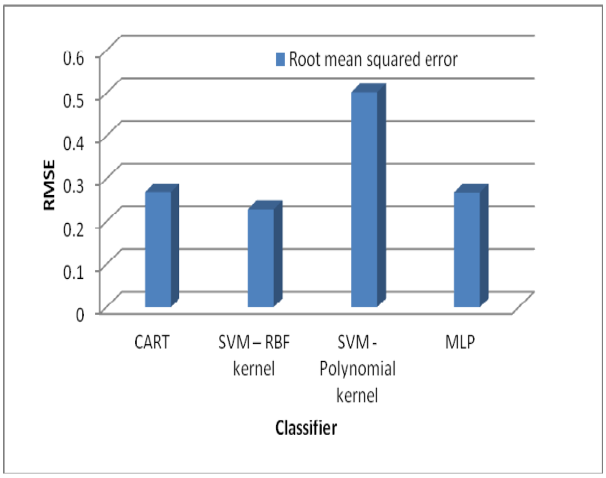

Fig. 3. Root mean squared error (RMSE) achieved

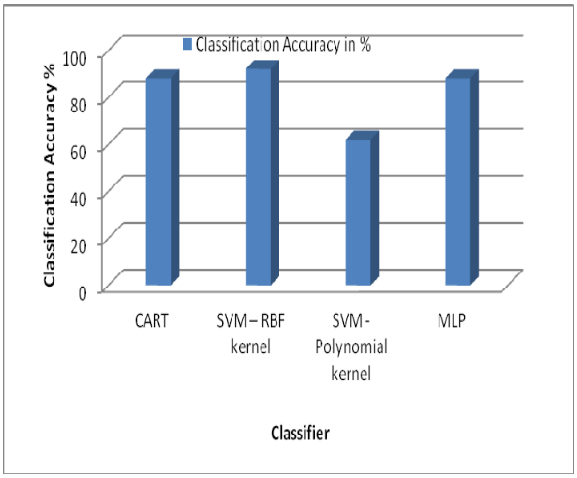

Fig. 4. Classification Accuracy for the classifiers used

It is seen from Figure 4, that the Gaussian RBF kernel performs better than the Polynomial kernel. The classification accuracy obtained by RBF kernel is 92.16%. Table 3 tabulates the precision, recall and F measure achieved by the classifiers. The precision, recall and F measure for the techniques used is shown in Figure 5 and 6 respectively. It is evident that the precision and recall of RBF kernel is much higher than the other classifiers.

Table 3. Precision, Recall and F-Measure

Classification Method	Precision	Recall	F measure
CART	0.884	0.882	0.883
SVM – RBF kernel	0.922	0.922	0.922
SVM - Polynomial kernel	0.429	621	0.494
MLP	0.882	0.882	0.882

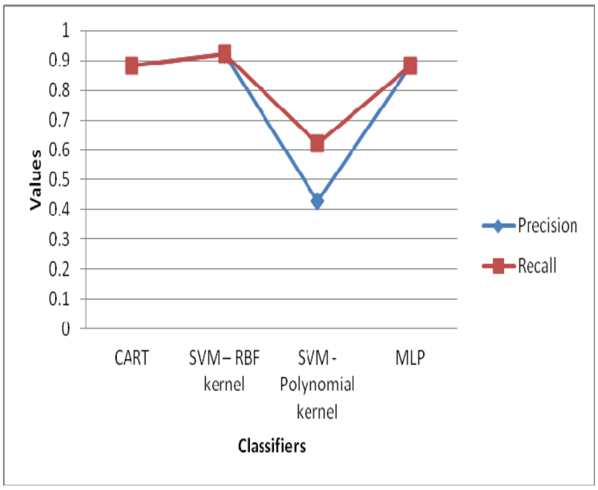

Fig. 5. Precision and Recall

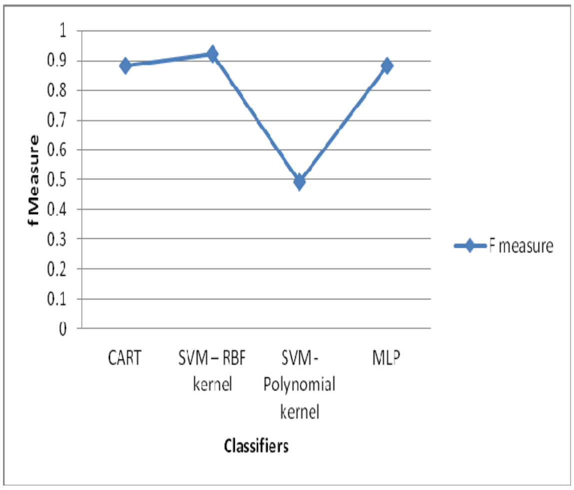

Fig. 6. F Measure

5. Conclusion

Cardiac Arrhythmia is an irregular heartbeat due to irregular rhythm caused by the heartbeat slowing down or increasing. ECG assesses heart arrhythmia. Different arrhythmia types relevant to heart disease are accurately diagnosed through ECG. ECG's automatic arrhythmia assessment is widely researched. Features like RR interval extracted from ECG are used for arrhythmia classification. For feature extraction, MIT-BIH time series data is converted to frequency domain using Discrete Cosine Transform (DCT). Two arrhythmia types, LBB and RBB were classified with normal beats. Experiments demonstrate Support Vector Machine (SVM) with Gaussian Radial Basis Function (RBF) kernel achieving classification accuracy of 92.16%.

References

[1] Sandoe, B. Sigurd (1991): *Arrhythmia – A Guide to Clinical Electrocardiology*, Bingen Publishing Partners Verlags GmBH.

[2] McSharry, P. E., *et al.* (2003): A dynamical model for generating synthetic electrocardiogram signals. *IEEE Trans. Biomed. Eng*., 50(3), pp. 289–294.

[3] Singh, B. N.; Tiwari, A. K. (2006): Optimal selection of wavelet basis function applied to ECG signal denoising. *Dig. Signal Process*., 16(3), pp. 275–287.

[4] Christov, G., *et al.* (2006): Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. *Med. Eng. Phys.*, 28, pp. 876–887.

[5] Chazal, P.; O'Dwyer, M.; Reilly, R. B. (2004): Automatic classification of heart beats using ECG morphology and heartbeat interval features. *IEEE Trans. Biomed. Eng*., 51, pp. 1196–1206.

[6] Martinez, J. P.; Olmos, S.; Laguna, P. (2000): Evaluation of a wavelet-based ECG waveform detector on the QT database. *Comput. Cardiol*., 27, pp. 81–84.

[7] Agante, P. M,; Marques de Sa, J. P. (1999): ECG noise filtering using wavelets with soft-thresholding methods. *Comput. Cardiol*., 26, pp. 535–538.

[8] Polat, K.; Şahan, S.; Güneş, S. (2006): A new method to medical diagnosis: Artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia. *Expert Systems with Applications*, 31(2), pp. 264-269.

[9] Wang, J. S., *et al.* (2012): ECG Arrhythmia Classification Using a Probabilistic Neural Network with a Feature Reduction Method. *Neurocomputing.*

[10] Ceylan, R.; Özbay, Y.; Karlik, B. (2009): A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. *Expert Systems with Applications*, 36(3) pp. 6721- 6726.

[11] Özbay, Y.; Ceylan, R.; Karlik, B. (2011): Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier. *Expert Systems with Applications*, 38(1), pp. 1004-1010.

[12] Khorrami, H.; Moavenian, M. (2010) A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification. *Expert systems with Applications*, 37(8), pp. 5751-5757.

[13] Kim, M. S., *et al.* (2011): A new method of ECG feature detection based on combined wavelet transform for u-health service. *Biomedical Engineering Letters*, 1(2), pp. 108-115.

[14] Jadhav, S. M.; Nalbalwar, S. L.; Ghatol. (2010): ECG arrhythmia classification using modular neural network model. *IEEE EMBS Conference on Biomedical Engineering and Sciences*, pp. 62-66.

[15] Homaeinezhad, M. R.; Ghaffari, A.; Rahmani, R. (2012): Multi-lead Discrete Wavelet-based ECG Arrhythmia Recognition via Sequential Particle Support Vector Machine Classifiers. *Journal of Medical and Biological Engineering*, 32(6), pp. 381-396.

[16] Chang, P.C., *et al.* (2012): Myocardial infarction classification with multi-lead ECG using hidden Markov models and Gaussian mixture models. *J. Applied Soft Computing, Journal*, 12(10), pp. 3165-3175.

[17] Karaolis, M. A.; Joseph A. M; Hadjipanayi, D.; Constantinos, S. P. (2010): Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining With Decision Trees. *IEEE Transactions On Information Technology In Biomedicine*, 14(3), pp. 559-566.

[18] Kampouraki, A.; George Manis; Nikou, C. (2009): Heartbeat Time Series Classification With Support Vector Machines. *IEEE Transactions On Information Technology In Biomedicine*, 13(4), pp. 512-518.

[19] Edoardo Pasolli; Farid Melgani. (2010): Active Learning Methods for Electrocardiographic Signal Classification. *IEEE Transactions On Information Technology In Biomedicine*, 14(6), pp. 1405-1416.

[20] Faezipour, M., *et al.* (2010): A Patient-Adaptive Profiling Scheme for ECG Beat Classification. *IEEE Transactions On Information Technology In Biomedicine*, 14(5), pp. 1153-1165.

[21] Hyejung Kim, *et al.* (2010): ECG Signal Compression and Classification Algorithm With Quad Level Vector for ECG Holter System. *IEEE Transactions On Information Technology In Biomedicine*, 14(1), 93-100.

[22] Sufi, F.; Mahmood, A.; Khalil, I. (2011): Diagnosis of Cardiovascular abnormalities from Compressed ECG: A Data mining based approach. *IEEE Trans. Inf. Technol. Biomed*., 15(1), pp. 33–39.

[23] MIT-BIH Arrhythmia Database. (2010). [Online]. Available: www.physionet.org/physiobank/database/mitdb

[24] Khoureich, *et al.* (2011): ECG beats classification using waveform similarity and RR interval. *J. of Medical and Bio. Engineering,* 32(6), pp. 417-422.

[25] Ahmed, N.; Natarajan, T. (1983): *Discrete-Time Signals and Systems*, Reston Publishing Company.

[26] Berk, R. A. (2008): Classification and regression trees (CART). *Statistical Learning from a Regression Perspective*, New York: Springer, pp. 1- 65.

[27] Gardner, M. W.; Dorling, S. R. (1998): Artificial neural networks (the multilayer perceptron)-A review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14 – 15), pp. 2627- 2636.

[28] Gunn, S. R. (1998): Support Vector Machines for Classification and Regression. *ISIS Technical Report*, 14.