Journal of Soft Computing and Applications 2013 (2013) 1-9

Available online at www.ispacs.com/jsca
Volume 2013, Year 2013 Article ID jsca-00020, 9 Pages
doi:10.5899/2013/jsca-00020
Research Article

Journal of Soft Computing and Applications

A model presented for classification ECG signals base on Case-Based Reasoning

Elaheh Sayari[1*], Mahdi Yaghoobi[1]

(1) Computer engineering department, Mashhad branch, Islamic Azad University, Mashhad, Iran

Abstract

Early detection of heart diseases/abnormalities can prolong life and enhance the quality of living through appropriate treatment; thus classifying cardiac signals will be helped to immediate diagnosing of heart beat type in cardiac patients. The present paper utilizes the case base reasoning (CBR) for classification of ECG signals. Four types of ECG beats (normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat and atrial fibrillation beat) obtained from the PhysioBank database was classified by the proposed CBR model. The main purpose of this article is classifying heart signals and diagnosing the type of heart beat in cardiac patients that in proposed CBR (Case Base Reasoning) system, Training and testing data for diagnosing and classifying types of heart beat have been used. The evaluation results from the model are shown that the proposed model has high accuracy in classifying heart signals and helps to clinical decisions for diagnosing the type of heart beat in cardiac patients which indeed has high impact on diagnosing the type of heart beat aided computer.

Keywords: Heart, Classifying, Case Base Reasoning, Beat, Computer.

1 Introduction

The electrocardiogram (ECG) signal is the recording of the bioelectrical activities of the cardiac system. It provides valuable information about the functional aspects of the heart and cardiovascular system. Early detection of heart diseases/abnormalities can prolong life and enhance the quality of living through appropriate treatment [1]. For effective diagnostics, the study of ECG pattern and heart rate variability signal may have to be carried out over several hours. Thus Conventional methods of monitoring and diagnosing electrocardiographic changes rely on detecting the presence of particular signal features by a human observer. Due to large number of patients in intensive care units and the need for continuous observation of such conditions, several techniques for automated electrocardiographic changes detection have been developed in the past 10 years to attempt to solve this problem [2].

* Corresponding Author. Email address: Elahe.sayari.66@gmail.com, Tel: +989178849311

To classify heart signals, several methods have been reported such as Neural Networks [11], Learning Vector Quantization and Probabilistic Neural Networks [6], Toolbox ANFIS [7], Neuro-Fuzzy Network [10], Fuzzy Inference Systems [9], K- Nearest Neighbor (KNN) [8], etc.
Thus, in this article, in order to diagnose and classify types of heart beat and to help to cardiac patients for early diagnosis of cardiac diseases and the type of their heart beat in emergency cases, sample-based learning system has been used; hence in lack of enough time and for helping cardiac patients, diagnosing the type of their hear beat would be done based-computer-aided as soon as possible.

2 Data description

PhysioBank database [3] is a large and growing archive of well-characterized digital recordings of physiologic signals and related data for use by the biomedical research community. PhysioBank currently includes databases of multi-parameter cardiopulmonary, neural, and other biomedical signals from healthy subjects and patients with a variety of conditions with major public health implications, including sudden cardiac death, congestive heart failure, epilepsy, gait disorders, sleep apnea, and aging. The databases of normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat and atrial fibrillation beat were studied in this work. The waveforms of four different ECG beats (normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat, and atrial fibrillation beat) considered in this study are shown in Figure 1(a)–(d).

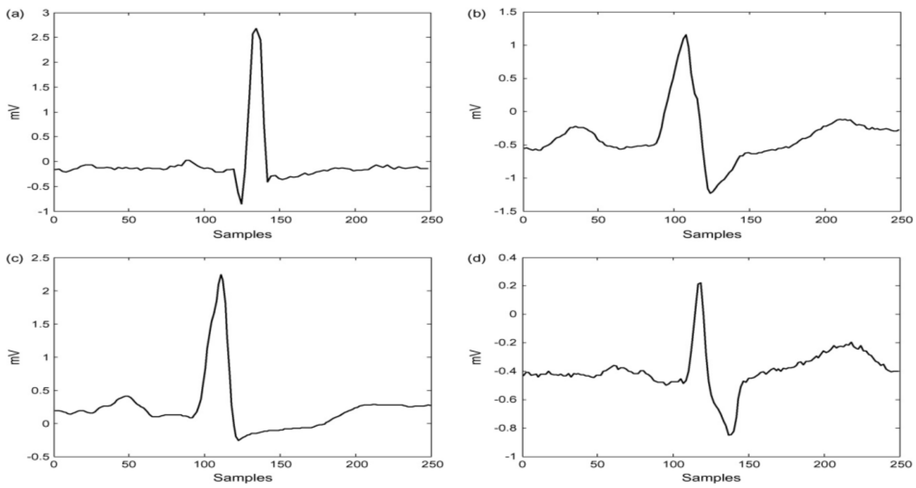


Figure 1: Waveforms of the ECG beats (a) normal beat, (b) congestive heart failure beat, (c) ventricular tachyarrhythmia beat, and (d) atrial fibrillation beat.

The PhysioBank database contains 20 records. The subjects were 25 men aged 32–89 years, and 25 women aged 23–89 years. Each of the 20 records is slightly over 30 min long. In most records, the upper signal is a modified limb lead II, obtained by placing the electrodes on the chest. The ECG signals were divided into two separate data sets- the training data set and the testing data set. In this study, the 100 vectors were used for training and the one vector was used for testing. The highest accuracy was obtained by dividing the data into two equal parts for training and testing. The training data set was used to train the CBR model, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained CBR model for classification of the four classes of ECG signals.

3 CBR system

CBR is one of lazy (learning) methods; i.e. the method which retarded classifying procedure until the new sample entered. We don't carry out any action on training data (collection of issues and remedies) until the new method is entered; after the entrance of new sample, by classifying new sample, we make our new sample class. CBR method depends on previous memory and experiences of studied problems [4].

3.1. Definitions of CBR:

- Case-based reasoning is reasoning by remembering.
- A case-based reasoner solves new problems by adapting solutions that were used to solve old problems.

3.2. CBR assumption(s):

- The main assumption is that:
  - Similar problems have similar solutions.
- Two other assumptions:
  - The world is a regular place: what holds true today will probably hold true tomorrow.
  - Situations repeat: if they do not, there is no point in remembering them.

4 The overall process of CBR system

At first, the new sample is entered to the system and the same problems of new sample recovered from database which is the collection of our samples and in adaptation step, the recovered remedies for the same problem is adapted with a remedy which has been diagnosed for new problem; consequently, the new remedy is suggested and in study phase, if the suggested remedy has had any problem, it is resolved till the new problem and the remedy stored in samples' collection which will be useful for the next required applications.

4.1. CBR system components (R^4 cycle):

- **Retrieve:** the cases from the case-base whose problem is most similar to the new problem.
- **Reuse:** the solutions from the retrieved cases to create a proposed solution for the new problem.
- **Revise:** the proposed solution to take account of the problem differences between the new problem and the problems in the retrieved cases.
- **Retain:** the new problem and its revised solution as a new case for the case-base if appropriate.

The R^4 cycle steps have been shown in Figure 2.

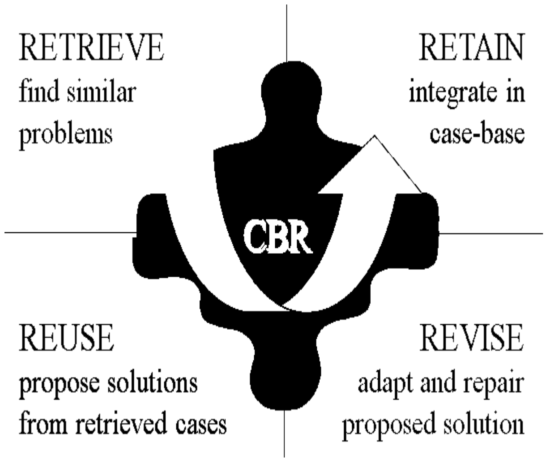


Figure 2: R^4 cycle steps

5 Proposed CBR system for ECG signals

According to the given description in previous part, types of diagnosed heart beat resulted from examining different cardiac patient with their remedies stored in CBR system (case-base) that by referring the cardiac patient who has the same heart beat with the previous referred ones, type of patient's heart beat will be diagnosed immediately and treated on time.

Based on the type of heart beat of a patient who referred to clinics and health care centers, the same heart beats are recovered from the collection of samples (including types of heart beats of cardiac patients who already referred and their remedies) (retrieve step) and a remedy is suggested due to recovered beats; it means that the remedy is chosen which has the most similarity based on allocated numbers between characteristics of recovered and new beats (reuse step). After adaptation stage, new suggested remedy is restudied and its errors identified (revise step); finally new heart beat and its related remedy are stored and kept in collection of samples in the relevant clinic (retain step). Range of numbers which is dedicated to different heart beats to find similarities is between 0-1. The proposed system steps have been shown in figure 3.

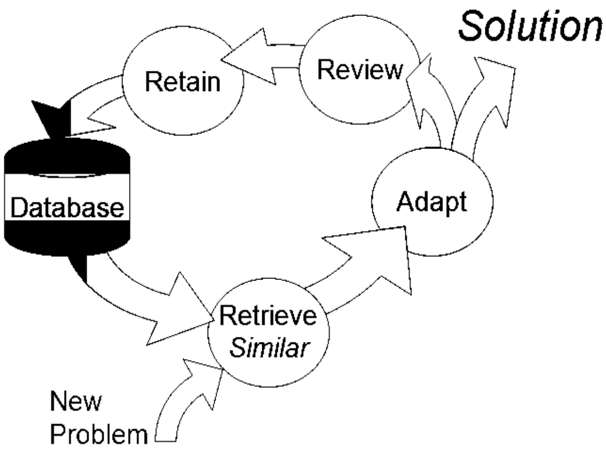


Figure 3: CBR system steps

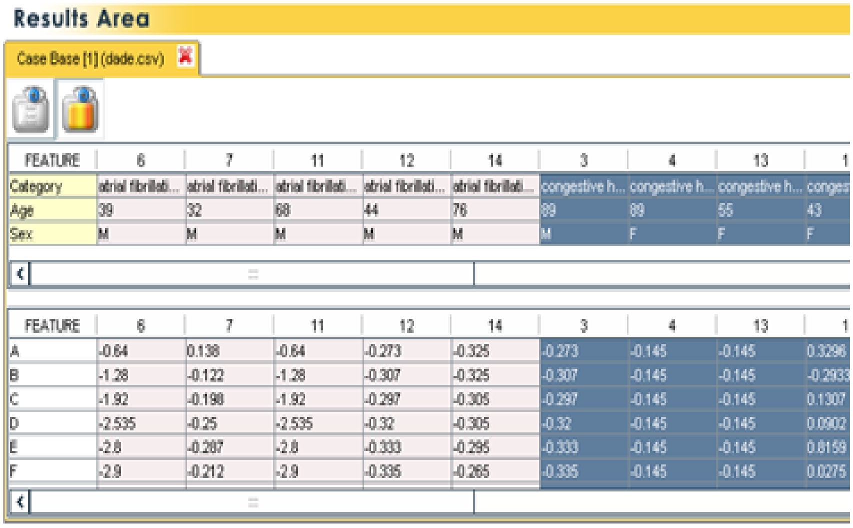
6 Discussion and results

In this article, the simulation is made for training data and test with GeneCBR software that simulation results have been separately shown in following parts.

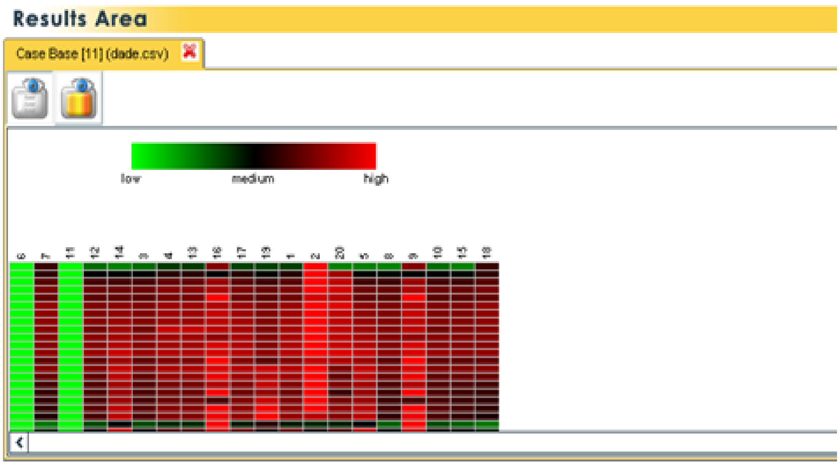
6.1. The simulation Results for training data:

The ECG signals were divided into two separate data sets- the training data set and the testing data set that the 100 vectors (25 vectors from each class) were used for training and the one vector (25 vectors from each class) were used for testing.
In GeneCBR software two working environments exist:
1. Programming environment and special environment for working,
2. Diagnosis environment.

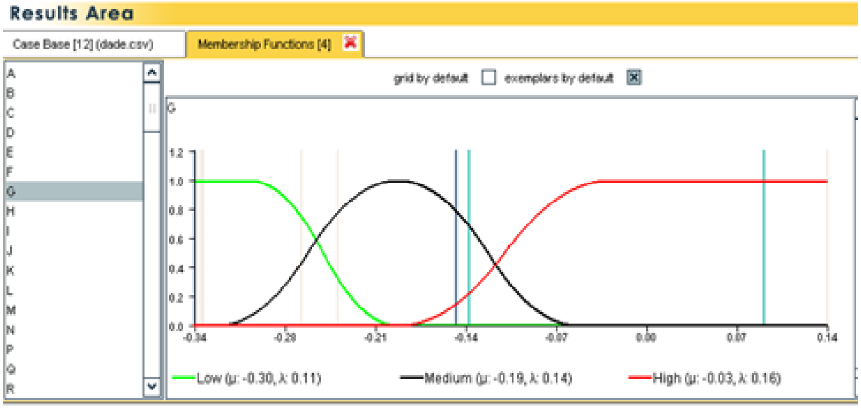
In programming environment we can observe the produced collection of samples in the result normally, in color and membership functions. To produce membership function for characteristics, three-variable linguistic software (low, medium, high) is explained and we can also change these three variables. The results of membership function for one characteristic and collection of produced sample with 100 types of heart beat of different cardiac patients has been shown in figure 4 in different environments of simulation.



(a)



(b)

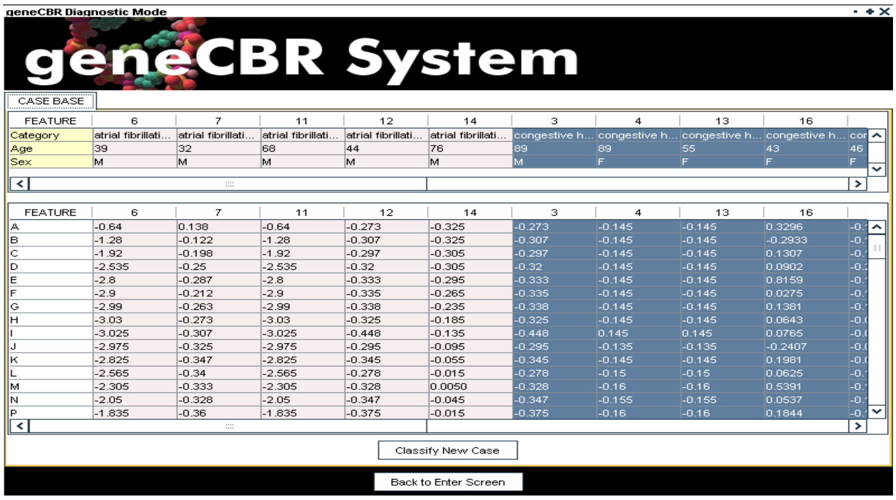


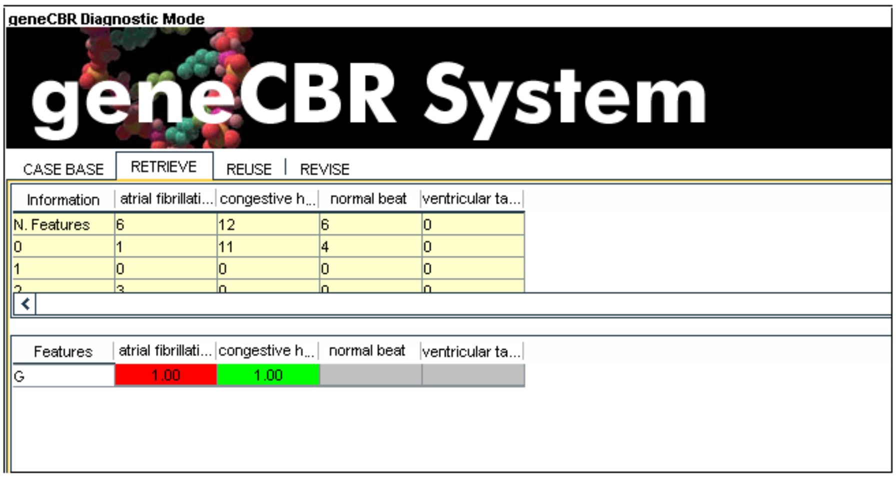
(c)

Figure 4: The obtained results of different environments of simulation with software (a) produced collection of samples in the result normally, (b) produced collection of samples in color, (c) membership functions.

6.2. The simulation results for testing data:

Outcomes resulted from classifying one sample in diagnosis environment has been shown in figure 5 with 4-cycle steps. In diagnosis environment an intelligent system (produced in programming environment) has been used which can be classified patients according to the type of their heart beat to their relevant group.





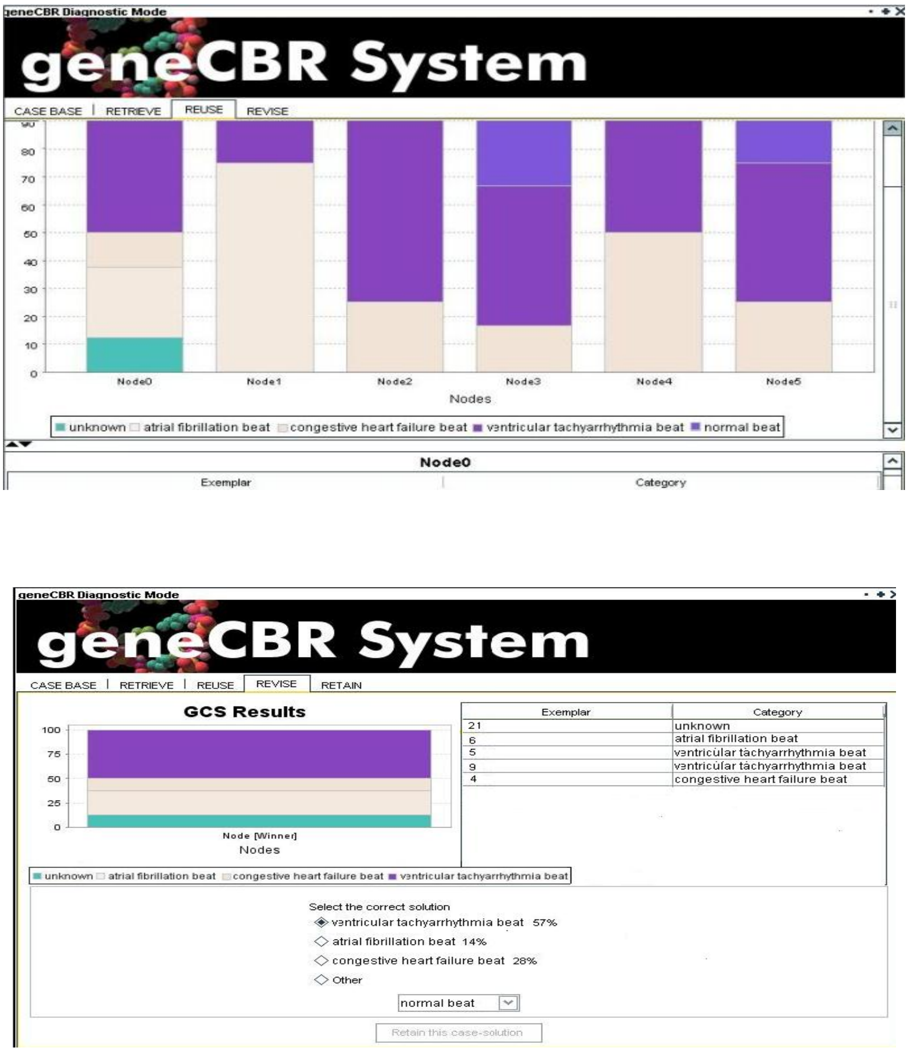


Figure 5: Classifying collection of sample in diagnosis environment with 4-cycle steps.

The confusion matrix showing the classification results of the CBR used for classification of the ECG signals is given in Table 1.

Table 1: Confusion matrix

Output result	Desired result			
	Normal beat	Congestive heart failure beat	Ventricular tachyarrhythmia beat	Atrial fibrillation beat
Normal beat	**98**	**0**	**0**	**1**
Congestive heart failure beat	**2**	**96**	**3**	**1**
Ventricular tachyarrhythmia beat	**0**	**3**	**95**	**2**
Atrial fibrillation beat	**0**	**1**	**2**	**96**

The test performance of the classifiers can be determined by the computation of specificity, sensitivity and total classification accuracy. The specificity, sensitivity and total classification accuracy are defined as:

- **Specificity:** number of true negative decisions/ number of actually negative cases.
- **Sensitivity:** number of true positive decisions/ number of actually positive cases.
- **Total classification accuracy:** number of correct decisions/total number of cases.

The values of the statistical parameters (sensitivity, specificity and total classification accuracy) are given in Table 2. The total classification accuracy of the CBR model was 98.78%.

Table 2: The Values of statistical parameters

ECG beats	Statistical parameters (%)		
	Sensitivity	Specificity	Total classification accuracy
Normal beat	96.64	99.23	98.78
Congestive heart failure beat	93.39	98.64	
Ventricular tachyarrhythmia beat	95.43	98	
Atrial fibrillation beat	97.58	98.99	

In table 3, the comparison between different methods of classifying ECG signals has been shown.

Table 3: The comparison between different methods of classifying ECG signals

Method	Total accuracy (%)	Time (seconds)
Case Base Reasoning (CBR)	98.78	0.32
Pruned fuzzy K-nearest neighbor (PFKNN) [5]	97.63	6.91
Learning Vector Quantization and Probabilistic Neural Networks [6]	97.5	3.45
Toolbox ANFIS [7]	96.39	2.5
K- Nearest Neighbor (KNN) [8]	96.33	1.98
Fuzzy Inference Systems(FIS)[9]	95	10
Neuro-Fuzzy Network [10]	93.5	4.56
Artificial Neural Networks (ANN) [11]	90.56	6.3

7 Conclusion

Classifying heart signals and early diagnosis of heart beat type in cardiac patients will be allowed to diagnose cardiac diseases immediately. Therefore, in this article, to classify types of heart beats CBR system (a lazy method) has been used to store heart beats and their medical remedies and suggested system as a previous experience collected from different physicians will be applied in different situations and times of patient’s refer to clinics and health care centers which indeed has high impact on diagnosing the type of heart beat aided computer. In contrast with other methods of classifying ECG signals mentioned in this article, this method has high accuracy equal to 98.78% and also is remarkably economical in time and the cost of early diagnosis of heart beat type in cardiac patients.

References

[1] S. C. Saxena, V. Kumar and S. T. Hamde, Feature extraction from ECG signals using wavelet transforms for disease diagnostics, International Journal of Systems Science, 33 (13) (2002) 1073-1085.
http://dx.doi.org/10.1080/00207720210167159

[2] Y. C. Yeh, W. J. Wang and C. W. Chiou, Feature selection algorithm for ECG signals using Range-Overlaps Method, Expert Systems with Applications, 37 (4) (2010) 3499-3512.
http://dx.doi.org/10.1016/j.eswa.2009.10.037

[3] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals Circulation, 23 (101) (2000) e215-e220.
http://dx.doi.org/10.1161/01.CIR.101.23.e215

[4] V. Raman, P. Sumari, J. R. Lekha and E. George Dharma Prakash raj, Performance based CBR mass detection in mammograms, IEEE, 10 (23) (2010) 565-568.
http://dx.doi.org/10.1109/ICCCCT.2010.5670776

[5] M. Arif, M. Usman Akramnand, F. AfsarAmirMinhas, Pruned fuzzy K-nearest neighbor classifier for beat Classification, J. Biomedical Science and Engineering, 3 (10) (2010) 380-389.
http://dx.doi.org/10.4236/jbise.2010.34053

[6] A. Fahoum, A. S. I. Howitt, Combined wavelet transform and radial basis neural networks for the classifying life threatening cardiac arrhythmias, Medical & Biological Engineering & Computing, 5 (37) (1999) 566-573.
http://dx.doi.org/10.1007/BF02513350

[7] T. M. Nazmy, H. EL. Messiry, B. AL-Bokhity, Adaptive neuro-fuzzy inference system for classification of ECG signals, Journal of Theoretical and Applied Information Technology, (2010).

[8] I. Christov, I. Jekova and G. Bortolan, Premature, ventricular contraction classification by the K ith nearest neighbors, rule, Physiological Measurement, 1 (123) (2005).

[9] T. P. Exarchos, M. G. Tsipouras Exarchos, C. P. Papaloukas, C. Fotiadis and L. K. Michalis, A methodology for the automated creation of fuzzy expert systems for ischemic and arrhythmic beat classification based on a set of rules obtained by a decision tree, Artificial Intelligence in Medicine, 40 (3) (2007) 187-200.
http://dx.doi.org/10.1016/j.artmed.2007.04.001

[10] M. Engin, ECG beat classification using neuro-fuzzy network, Pattern Recognition Letters, 25 (6) (2004) 1715-1722.
http://dx.doi.org/10.1016/j.patrec.2004.06.014

[11] B. Anuradha and V. C. Veera Reddy, ANN for classification of cardiac arrhythmias, ARPN Journal of Engineering and Applied Sciences, 3 (3) (2008).