



Contents list available at JMCS

Journal of Mathematics and Computer Science

Journal Homepage: www.tjmcs.com



Identification and Classification of Coronary Artery Disease Patients using Neuro-Fuzzy Inference Systems

Saeed Ayat¹, Asieh Khosravanian²

¹ Associate Professor, Department of Computer Engineering and Information Technology, Payame Noor University, IRAN.

² M.Sc, Department of Computer Engineering and Information Technology, Payame Noor University, IRAN.

¹Dr.ayat@pnu.ac.ir, ²khosravanian.a@gmail.com

Article history:

Received July 2014

Accepted September 2014

Available online September 2014

Abstract

In this research patients with coronary artery disease were identified and classified through the neuro-fuzzy network with the capacity of automatically extracting fuzzy rules. Fuzzy expert system was implemented using facilities and functions of MATLAB software (7.12.0 version). Network parameters, introductory and lower parameters, were trained by back-propagation error (gradient descent) method. The proposed method was evaluated through data collected from medical files of 152 patients with coronary angiography in Kowsar Hospital, Shiraz, Iran during September, 2013. The performance indicators of this system were specificity and sensitivity. The indicators, as extracted from testing results, were found to be 0.88 and 1, respectively.

Keywords: Coronary Artery Disease, Neuro-fuzzy Network, Fuzzy Expert System.

1. Introduction

Recently, expert systems have drawn the attention of many researchers for application diagnosis of diseases. These systems, using data mining techniques, can discover patterns of medical data, improve the decision-making process, affect the costs, and increase the quality of health care [1, 2]. Coronary artery disease (CAD) is one of the most prevalent diseases of human populations that, in addition to taking thousands of lives, imposes huge costs on societies [3, 4, 5]. Nowadays, the most important cause of human mortality is the cardiovascular disease [6]. As reported by American Heart Association, more than 12 million Americans suffer from this disease [7]. One of the ways of diagnosing this disease is the coronary angiography procedure, which is an expensive and invasive procedure and might be accompanied by risks such as death and stroke [8].

Different studies show that different methods have been used in disease diagnosis including statistical methods, neural networks, decision trees, genetic algorithms, and fuzzy classifiers. In the fuzzy classifier method with ellipsoidal regions, by changing slopes of membership functions, the necessary fuzzy rules for classification of samples have been obtained [9].

More specifically, many studies on heart coronary have used neural networks, some of which are mentioned in this section. In a study on the application of artificial neural networks (ANN) in determining predictors of inpatient mortality following open-heart surgery and its comparison with logistic regression model [10], an ANN was used with 18 neurons in the input layer, 4 neurons in the hidden layer, and 2 neurons in the output layer as well as back-propagation algorithm and sigmoid transfer function. The objective of present study was to compare the capabilities of two models of ANN and logistic regression so as to predict the risk of in-hospital death following open heart surgery. Statistical population of this study consisted of patients who had open-heart surgery in Shariati Hospital in Tehran, Iran. To show the accuracy of results of this study, the two statistical measures sensitivity and specificity were used. Sensitivity refers to the proportion of individuals diagnosed as patient to total number of patients. Specificity refers to the proportion of healthy individuals diagnosed by the neural system to total healthy individuals.

In the published results of this paper, sensitivity and specificity in training phase were 100%, while they were respectively 99.33% and 100% in the testing phase, but in the logistic regression model, sensitivity and specificity were respectively 99% and 90%, therefore, the results of this study show that neural network is more efficient than logistic regression model in problem-solving.

In another study, comparison of ANN to other statistical methods such as logistic regression for prediction of CAD was done, in which ROC (Receiver Operating Characteristic) curve was used. In ROC curve, the association of sensitivity with specificity was established and the results showed that more area under ROC curve in ANN compared to other statistical methods showed the advantage of ANN in solving such problems [11].

In [12] Mehrabi et al., ANN was used to identify the difference between heart disease and obstructive pulmonary which have similar symptoms among 266 patients in one of Tehran hospitals. MLP (Multi-layer Perceptron) ANNs and RBF neural networks (Radial Basis Networks) were used for diagnosis. In the MLP network, sensitivity and specificity were 83.9% and 86%, respectively, whereas the two variables were 88.4% and 81.8% for the RBF network, respectively.

In this study, an integration of the two fuzzy and neural methods, which use the abilities of learning and parallel processing of neural networks, contributed to designing a fuzzy expert system. In this expert system, the neural network is based on the fuzzy system.

The risk of factors in pre-angiography tests in the population under study included age, creatinine, cholesterol, triglycerides, ejection fraction, gender, smoking, hypertension, diabetes, family history of heart disease, myocardial infarction, exercise test, and echo stress test results, which were all considered influential factors of coronary artery stenosis after consultation with heart specialists. In this system, in addition to fuzzy inference, simplicity and interpretability were added to features of the neural network that makes it possible for, the expert to introduce changes in the system structure and parameters.

2. Neuro-fuzzy network

Neural implementation of a fuzzy system as defined in [13] is described in this section. First, a general neuron is defined. The input and output relationship of this neuron is provided in the following:

$$O_i^k = a(f(u_1^k, u_2^k, \dots, u_n^k, w_1^k, w_2^k, \dots, w_n^k)) \quad (1)$$

where f is central function of neuron, a is neuron activation function, u_i^k is i -th neuron of k -th layer, w_i^k is weight i -th of k -th layer and o_i^k is the output of i -th neuron in k -th layer. This network has 5 layers. The first layer is input. The weights of neurons in this layers is constant and equal to the unity. So, in this layer we have:

$$f = u_i^1, \quad a = kf \quad (2)$$

In the above equation, k is input scaling factor.

The second layer is fuzzification layer in which Gaussian differentiable membership function is used as follows:

$$f = -\frac{(u_i^2 - m_{ij})^2}{\sigma_{ij}} \quad , \quad a = e^{f^2} \quad (3)$$

m_{ij} and σ_{ij} are respectively mean and standard deviation of Gaussian function for the i -th fuzzy set, the adjustment of which was done in the stage of learning membership function.

The third layer is fuzzy inference layer. In this layer, there is a multiplicative neuron for each rule. In this layer, fuzzy rules are defined.

$$f = u_1^3 \times u_2^3 \times \dots \times u_m^3 \quad , \quad a = f \quad (4)$$

In the fourth layer, fuzzy rules defined in previous layer are normalized.

$$f_i = u_i^4 \quad , \quad y_i = a_i = \frac{f_i}{\sum_{i=1}^m u_i^4} \quad (5)$$

The fifth layer is a non-fuzzy layer in which. If mean of centers is considered for defuzzification, the functions of this layer are defined via:

$$f = \sum_{i=1}^m w_i^5 \cdot u_i^5 \quad , \quad y = a = f \quad (6)$$

Sugeno-type fuzzy rules have a simpler structure and easier inference compared with other neuro-fuzzy networks the output of which is defined as follows:

$$f = \sum_{i=1}^m u_i^4 (a_i x_1 + b_i x_2 + c_i) \quad , \quad y = a = \frac{f}{\sum_{i=1}^m u_i^4} \quad (7)$$

In this equation, x_1 and x_2 are inputs of neuro-fuzzy network which are non-fuzzy parameters. a_i , b_i and c_i are consequent parameters of rules which are set up by learning. In this study, parameters of both introductory and consequent sections are taught by gradient descent method and synchronization is performed for each training course.

3. Materials and Methods

The present study is a diagnostic one. The statistical population included individuals who had angiography in Kowsar Hospital, Shiraz, Iran, during September, 2013. 152 individuals were randomly selected and a researcher-made questionnaire was used including demographic information, diseases, diagnostic tests, and the results of angiography. The fuzzy expert system designed based on input variables in this study predict the condition of coronary vessels of the heart. In order to implement this network, MATLAB software (7.12.0 version) was used. 85% of the data were used for the stage of network training, whereas the remaining 15% was used in the stage of network testing. The, 13 clinical variables were chosen as network input and were divided into two categories of continuous and discrete. Continuous variables include age, creatinine, total cholesterol, triglycerides, and ejection fraction, while discrete variables consist of gender, smoking, history of hypertension, history of diabetes, familial history of heart disease, myocardial infarction, abnormal exercise test results, and echo results.

4. The Proposed method

Primarily, based on previous studies and expert opinion of cardiovascular specialists, variables under study were determined, and then a sample of 152 cases of these variables were collected in Kowsar Hospital, Shiraz. The data collected showed that totally 49 individuals had normal angiographic result and 103 individuals had abnormal results (i.e. 40 individuals had stenosis of an artery, 22 individuals had stenosis in two arteries, and 41 individuals had stenosis of 3 arteries). The suggested method involves 2 stages according to [13]. The first stage is data preprocessing which includes normalizing input data so that data is put within a range between 0 and 1. The second stage is processing.

The data had two classes, namely *Normal* and *Abnormal*. So, one neuro-fuzzy network was designed for each class in the processing stage. The desired output for each one of these networks occurs when the input sample belongs to the intended class, as result of which the desired outcome is 1; otherwise, it is 0. The network was trained by 150 epoch of training data. Then, by thresholding the output of each neuro-fuzzy network by genetic algorithm, the network output was discretized so that each sample could be attributed to a class (Table 1).

Table 1. How to assign the output of the two classes

Classification	Output FNN1	Output FNN2
Normal	1	0
Abnormal	0	1

Neuro-fuzzy network interface was used to design a Sugeno-type fuzzy expert system in MATLAB. First, training data of 132 samples was loaded. Then, primary structure of the fuzzy system was created by selection of grid Partition method. In the designed network, to train membership function parameters, back-propagation optimization method was used. Finally, model validation was conducted by using unused data in the training stage. A brief report of the simulation results for statistical population is shown in Table 2.

Table 2. The results of neuro-fuzzy systems

Accuracy in the testing phase	Accuracy in the training phase	The number of membership functions	The structure of the fuzzy system	Number of samples
80	94.6	3	Sugeno	152

One of the reasons for the high rate of classification accuracy is the normalization of input data and using the Sugeno structure. Because this structure uses the input once more in the result section and has more training parameters, it provides a higher classification accuracy. The result of predicting designed fuzzy expert system is shown in Table 3.

Table 3. Prediction Fuzzy Expert System

	Angiographic result	Prediction Fuzzy Expert System
Normal	17	15
Coronary artery disease	6	8
Total	23	23

As shown in Table 3, this system mistakenly diagnosed two healthy individuals as patients and all patients are validly diagnosed. Indicators of this system were sensitivity and specificity. Sensitivity refers to the proportion of diagnosed-as-patient individuals to total number of patients. Specificity, too, refers to the proportion of diagnosed-as-healthy individuals to total number of healthy individuals as shown by fuzzy expert system. In the present study, the above parameters were 1 and 0.88, respectively.

5. Conclusion

In this study, a neuro-fuzzy network was proposed to identify and classify CAD patients by designing a Sugeno-based fuzzy expert system. First, the network was trained by training samples and then it was evaluated by testing samples. In the present study, sensitivity and specificity parameters of this system were 1 and 0.88, respectively, and this fact shows that this network relied on an acceptable degree of accuracy in the classification.

The obtained results through Sugeno structure showed that in regard to sensitivity, the expert system implemented yielded a better accuracy compared to similar cases, while specificity showed less accuracy compared to the findings of some other studies.

The advantages of this system compared to systems that just use ANN are fuzzy inference, simplicity and interpretability, which makes it possible for experts to apply necessary changes to system parameters and structure. In addition, training process and adjusting this system is faster than the neural network. The proposed fuzzy expert system also has the possibility to be used for training medical students.

References

- [1] H.C. Koh, G. Tan, Data mining applications in healthcare, *J Healthcare Info Manag.* 2005 spring; 19(2): 64-72.
- [2] Y.M. Chae, H.S. Kim, K.C. Tark, H.J. Park, S.H. Ho, Analysis of healthcare quality indicator using data mining and decision support system, *Exp Sys Applic.* 2003; 24(2): 167-72.
- [3] G.F. Filetcher, K.R. Oken, R.E. Safford, *Comprehensive Rehabilitation of patients with coronary artery disease*, Braunwald E, Zips Dp, Libby P. *Heart Disease, A text book of cardiovascular medicine* 2001; 6(2): 1406-17.
- [4] U. Fuster, R.W. Alexander, R. O'rouke, *Hurst's The Heart*, Wenger NK. *Rehabilitation of the Patient With coronary heart disease*, 10th edition, Mcgrow-Hill professional publish 2000; 2: 1537-46.
- [5] M.C. Mancini, E.M. Cush, K. Sweatman, J. Dansby, *Coronary artery bypass surgery: are outcomes influenced by demographics or ability to pay?*, *Ann Surg* 2001; 233(5): 617-22.
- [6] S. Mendis, P. Puska, B. Norrving, *Global atlas on cardiovascular disease prevention and control*, Geneva: World Health Organization; 2011.
- [7] American Heart Association. Available at: <http://www.heart.org/HEARTORG>.
- [8] I. Mahmoudi, R. Askari moghadam, M. Moazzam, S. Sadeghian, Prediction model for coronary artery disease using neural networks and feature selection based on classification and regression tree, *J Shahrekord Univ Med Sci.* 2013; 15(5):47-56.
- [9] S. Abe, R. Thawonmas, *A Fuzzy Classifier with Ellipsoidal Regions*, *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 3, Aug. 1997.
- [10] I. Turkoglu, A. Arslan, E. Ilkay, An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks, *Comput Biol Med* 2003; 33(4): 319-331.
- [11] L. Ohno-Machado, M.A. Musen, Sequential versus standard neural networks for pattern recognition: An example using the domain of coronary heart disease, *Comput Biol Med* 1997; 27(4): 267-281.
- [12] S. Mehrabi, M. Maghsoudloo, H. Arabalibeik, Application of multilayer perceptron and radial basis function neural networks in differentiating between chronic obstructive pulmonary and congestive heart failure diseases, *Expert Systems with Applications* 2009; 36: 6956–6959.
- [13] N. Forghany, M. Teshnelab, H. Abrishami Moghadam , M. Nakhjivani, Diagnosis of thyroid diseases using a hybrid Fuzzy Neural Network based on Linear Discriminate Analysis, 14th Iranian Conference on Biomedical Engineering, Tehran, Iran 2008.