

Thyroid and Breast Cancer Disease Diagnosis using Fuzzy-Neural Networks

Canan SENOL¹, Tülay YILDIRIM²

¹ Department of Electronic Eng., Kadir Has University, Cibali, 34083, Istanbul, Turkey
canan@khas.edu.tr

² Department of Electronics and Communication Eng., Yildiz Technical University, Besiktas, 34349, Istanbul, Turkey
tulay@yildiz.edu.tr

Abstract

In this paper a new hybrid structure in which Neural Network and Fuzzy Logic are combined is proposed and its algorithm is developed. Fuzzy-CSFNN, Fuzzy-MLP and Fuzzy-RBF structures are constituted, and their performances are compared. Conic Section Function Neural Network (CSFNN) unifies the propagation rules of the Multilayer Perceptron (MLP) and the Radial Basis Function (RBF) networks at a unique network by its distinctive propagation rules. That means CSFNNs accommodate MLPs and RBFs in its own self-network structure. The proposed approach is implemented in a well-known benchmark medical problem with real clinical data for thyroid and breast cancer disease diagnosis. Simulation results show that proposed hybrid structures outperform both MATLAB-ANFIS and non-hybrid structures.

1. Introduction

Fuzzy Logic and Neural Networks are complementary technologies in the design of intelligent systems. Fuzzy-Neural Networks (Fuzzy-NN) are based on exploiting the learning and decision making capabilities of the Artificial Neural Networks (ANN) and the Fuzzy Inference Systems, respectively.

Fuzzy Logic was introduced by Lotfi A. Zadeh in 1965. A classical set is a set with a crisp boundary. Fuzzy sets eliminate the sharp boundaries that divide members from nonmembers in a group whereas, in the classical set theory, an element strictly either belongs to a given set or does not. Therefore, in the Fuzzy Set Theory, the transition between full-membership and non-membership is gradual and an element can belong to a set partially. The degree of membership is defined by a so-called membership function (MF) $\mu_A(u): U \rightarrow [0,1]$, where U and A are the universe and fuzzy subset of U , respectively.

The basic structure of fuzzy inference systems (FIS) consists of three conceptual components: a rule base which contains a selection of fuzzy rules, a database which defines the membership functions used in the fuzzy rules, and a reasoning mechanism which performs the inference procedure upon the rules and given facts to derive a reasonable output or conclusion. Two types of fuzzy inference systems that have been widely employed in various applications. These are Mamdani and Sugeno type of FIS. Adaptive Neuro FIS (ANFIS) proposed by Jang and included in the Matlab, employs the Sugeno type of FIS. The Sugeno fuzzy model was proposed by Takagi, Sugeno, and Kang in an effort to develop a systematic

approach to generating fuzzy rules from a given input-output data set. A typical fuzzy rule in a Sugeno fuzzy model has the form

if x is A and y is B then $z = f(x, y)$

where A and B are fuzzy sets in the antecedent, while $z = f(x, y)$ is a crisp function in the consequent. In the Sugeno fuzzy model; each rule has a crisp output, the overall output is obtained via weighted average. In practice, the weighted average operator is the weighted sum operator to reduce computation further, especially in the training of a fuzzy inference system [1].

CSFNN was introduced by G. Dorffner in 1994. CSFNN has been used in various applications so far [2, 3, 4].

During the development of the update rule of the hybrid Fuzzy-NN parameters, in addition to MATLAB-ANFIS, some other hybrid algorithms are developed [5, 6]. Rough-ANFIS (RANFIS), backpropagation-genetic and K-SVD hybrid algorithms [7, 8, 9] are only a few types of these hybrid algorithms. As an interesting application, in [9], K-SVD hybrid algorithm derives rules by way of k-means clustering and chooses most dominant rules by Singular Value Decomposition (SVD) method.

In this work Fuzzy-CSFNN, Fuzzy-MLP and Fuzzy-RBF structures are constituted, and their performances are compared. Simulation results show that our proposed hybrid structures outperform both MATLAB-ANFIS and non-hybrid structures.

In section 2, proposed Fuzzy-NN structures is detailed. Simulation results are given in section 3. Section 4 included our comments on the results.

2. Proposed Fuzzy-NN Structures

Block diagram of proposed Fuzzy-NN hybrid scheme is given in Figure 1. Fuzzy part of the scheme must be set up by choosing system parameters such as type and the number of membership functions, defuzzification operators, etc. In Figure-1, $\underline{\alpha}$ weights denote the qualifications of the rules and are obtained by performing fuzzy implication operation for each given input data. Input data \underline{x} together with $\underline{\alpha}$ weights constitute the input data of the NN part. After training, the layer weights \underline{c} of the NN part are obtained.

In the fuzzy part, some parameters of the membership functions are determined according to the input data features. Rule qualification weights α_k 's are found for each of the input data. Here, k represents the rule number index. \underline{x} together with $\underline{\alpha}$ are applied to MLP as input data. RN denotes number of rules, MFN denotes number of membership functions and IN

denotes number of inputs; the number of rules can be calculated by the following equation,
 $RN = (MFN)^{IN}$ (1)

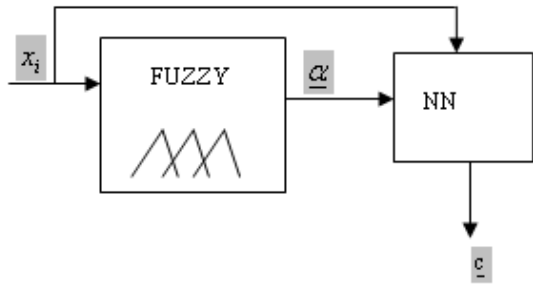


Figure 1. Block Diagram of Fuzzy-NN Hybrid scheme.

First, we consider Fuzzy-MLP hybrid scheme. NN structure was selected as MLP and the Fuzzy-MLP hybrid scheme was constituted as shown in figure 1.

Next, we consider Fuzzy-RBF for performance comparison of the proposed structure. Unlike Fuzzy-MLP, NN part of Fuzzy-RBF was constituted by using an RBFNN. Fuzzy-RBF hybrid structure was trained and tested.

Lastly, Fuzzy-CSFNN was considered as hybrid structure. NN part of Fuzzy-CSFNN was constituted by using a CSFNN.

The idea of the CSFNN is to provide unification between RBF and MLP networks. The new propagation rule (which will consist of RBF and MLP propagation rules) can be derived using analytical equations for a cone.

Let x be any point on the surface of the right circular cone. ω can be any value in the range $[-\pi/2, \pi/2]$, v vertex of the cone and a the unity vector defining the axis of the cone. Thus the

equation of the circular cone is

$$(\bar{x} - \bar{v})\bar{a} = \cos w \|\bar{x} - \bar{v}\| \quad (2)$$

If the coordinates of the points and vectors are defined by $x=(x_1, x_2)$, $v=(v_1, v_2)$ and $a=(a_1, a_2)$ for two dimensional space, equation (2) can be written as below

$$(x_1 - v_1)a_1 + (x_2 - v_2)a_2 = \cos w \sqrt{(x_1 - v_1)^2 + (x_2 - v_2)^2} \quad (3)$$

The propagation rule of conic section function network is described using equation (3). First of all the following form is obtained for n-dimensional input space.

$$\sum_{i=1}^{n+1} (x_i - v_i)a_i = \cos w \sqrt{\sum_{i=1}^{n+1} (x_i - v_i)^2} \quad (4)$$

The center coordinate of the circle, c , can be used instead of the coordinate of vertex v since the distance between the x point and the vertex v equals to the radius of the circle when the opening angle, 2ω , is 90 degrees. Subtracting the right hand side from the left hand side, the propagation rule of the CSFNN is obtained as

$$y_j = \sum_{i=1}^{n+1} (x_i - c_{ij})a_{ij} - \cos w_j \sqrt{\sum_{i=1}^{n+1} (x_i - c_{ij})^2} \quad (5)$$

where a_{ij} refers to the weights for each connection between the input and hidden layer units in an MLP network, and c_{ij} refers to the center coordinates in an RBF network, i and j are the indices referring to the units in the input and hidden layer, respectively, and y_j are the activation values of the CSFNN neurons.

As can be seen easily, this equation consists of two major parts analogous to the MLP and the RBF. The equation simply turns into the propagation rule of an MLP network, which is the dot product when the ω is $\pi/2$. Second part of the equation gives the Euclidean distance between the inputs and the centers for an RBF network. Figure 2 illustrates the structure of a Conic Section Function Neural Network. [12]

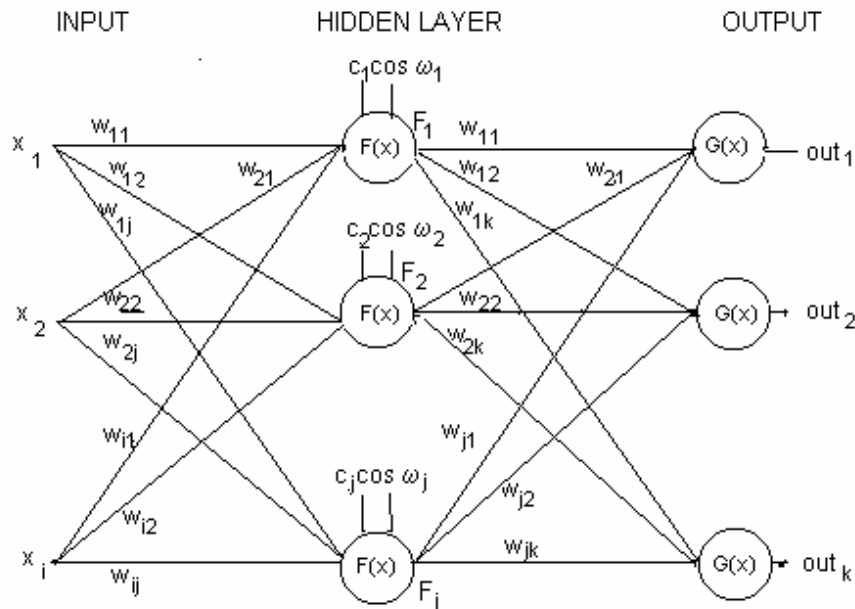


Figure 2. Conic section function neural network structure

3. Simulation Results

The proposed model was trained and tested on two different data sets. These are thyroid and breast cancer data sets. Datasets taken from the UCI machine learning respiratory was used as one of the benchmark datasets for testing classifiers [13,14].

In thyroid data-set, 215 instances have been used for this work. Each instance has five attributes plus the class attribute. All samples have five features. These are:

- T3-resin uptake test. (A percentage)
- Total Serum thyroxin as measured by the isotopic displacement method.
- Total serum triiodothyronine as measured by radioimmuno assay.
- Basal thyroid-stimulating hormone (TSH) as measured by radioimmuno assay.
- Maximal absolute difference of TSH value after injection of 200 micro grams of thyrotropin-releasing hormone as compared to the basal value.

All attributes are continuous. Each of the instances has to be categorized into one of the three classes: Class 1: normal, Class 2: hyper, Class 3: hypo functioning. Out of 215 instances, 47 have been used for training and 215 have been used for testing purposes.

In Fuzzy-MLP hybrid structure, we consider an MLP with an input layer (including 37 neurons), a hidden layer (including 10 neurons) and an output layer. Hybrid scheme was trained by Levenberg-Marquardt back propagation algorithm. We choice the number of MFs as 2 for each input data vector, and we have chosen bell-shaped type MFs. Membership function of output was selected as linear and transfer function was selected as pure linear. Our number of rule was found as 32 from equation 1. The most appropriate learning rate was found as 0,05.

In Fuzzy-RBF structure, regarding same FIS setup parameters for a fair comparison we have chosen the number of MFs as 2 for each input data vector, and we have selected bell-shaped type MFs. Membership function of output was chosen as linear. RBF structure was trained by Orthogonal Least Squares algorithm. We see that the goal as 0, suitable spread value was found as 9.

And lastly Fuzzy-CSFNN was applicated for tyhroid data set. In our proposed Fuzzy-CSFNN structure, the NN part was constituted by CSFNN. Again, we have chosen the number of MFs as 2 for each input data vector, and we preferred bell-shaped type MFs. Membership function of output was selected as linear and transfer function was selected as pure linear. We see that the most appropriate spread value as 4, number of centres as 16, sum square error as 0.001, learning rate as 0.03 [15].

Breast cancer database was attained from the University of Wisconsin Hospitals. The data consist of 683 records taken from patients' breasts. Each record in the database has 9 attributes from a normal state of 1 to 9 (most abnormal state). There are two class variables of breast cancer: malignant (cancerous) and benign (non-cancerous), which is represented numerically by 1 and 2 respectively. There are 239 malignant cases and 444 benign cases. The objective is to classify between malignant and benign cases.

In Fuzzy-MLP hybrid structure, we consider an MLP with an input layer (including 16 neurons) and an output layer. Hybrid scheme was trained by Levenberg-Marquardt back propagation algorithm. We choice the number of MFs as 2 for each input data vector, and we have chosen bell-shaped type MFs.

Membership function of output was selected as linear and transfer function was selected as pure linear. Our number of rule was found as 16 from equation 1. The most appropriate learning rate was found as 0,9.

In Fuzzy-RBF structure, regarding same FIS setup parameters for a fair comparison we have chosen the number of MFs as 2 for each input data vector, and we have selected bell-shaped type MFs. Membership function of output was chosen as linear. RBF structure was trained by Orthogonal Least Squares algorithm. We see that the goal as 0, suitable spread value was found as 1,5.

Finally, Fuzzy-CSFNN was applicated for breast cancer data set. In our proposed Fuzzy-CSFNN structure, the NN part was constituted by CSFNN. Again, we have chosen the number of MFs as 2 for each input data vector, and we preferred bell-shaped type MFs. Membership function of output was selected as linear and transfer function was selected as pure linear. We see that the most appropriate spread value as 0.9, number of centres as 3, sum square error as 0.001, learning rate as 0.005. Table 2 shows the performance comparisons of proposed hybrid cschemes for breast cancer database.

Simulation results of Fuzzy-MLP, Fuzzy-RBF and Fuzzy-CSFNN were compared with ANFIS structure in Matlab. In addition, standard MLP, RBF and CSFNN were used and their results were compared to hybrid structures. In MLP structure, training process was repeated 10 times since it gives different results depending on random initialization of weights in the algorithm. Then, the average of the results was taken. Table 1 and 2 shows our results of proposed hybrid schemes.

Table 1. Results for thyroid database

	Normal	Hyper	Hypo	Average
ANFIS	82	18	17	%71.4
Fuzzy-MLP	108	20	22	%88.53
MLP	110	24	19	%90.09
Fuzzy-RBF	88	20	21	%81.54
RBF	103	21	6	%65.32
Fuzzy-CSFNN	115	24	20	%92.93
CSFNN	103	23	17	%83.91

Table 2. Results for breast cancer database

	Benign	Malignant	Average
ANFIS	41	44	%96.59
Fuzzy-MLP	41	44	%96.59
MLP	42.5	44	%98.3
Fuzzy-RBF	42	44	%97.73
RBF	42	44	%97.73
Fuzzy-CSFNN	43	44	%98.87
CSFNN	43	44	%98.87

Receiver Operating Characteristic (ROC) analysis is commonly used in medicine and healthcare to quantify

the accuracy of diagnostic test. The basic idea of diagnostic test interpretation is to calculate the probability a patient has a disease under consideration given a certain result. Without ROC analysis, it is difficult to summarize the performance of a test with a manageable number of statistics and to compare the performance of different tests.

The diagnostic performance is usually evaluated in terms of sensitivity and specificity. Sensitivity is the proportion of patients with disease whose tests are positive. Specificity is the proportion of patients without disease whose tests are negative. The measures are defined as:

$$\text{sensitivity } y = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}} \quad (6)$$

$$\text{specificity } y = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}} \quad (7)$$

where #true positives and #false negatives are the number of breast cancer correctly classified and incorrectly classified as normal case, respectively. Similarly, #true negatives and #false positives are the number of normal case correctly classified and incorrectly classified as breast cancer case.

ROC Analysis was applied the breast cancer data set and results were presented in Table 3 for Fuzzy-NN hybrid schemes.

Table 3. ROC Analysis results for breast cancer database

	Sensitivity	Specificity
ANFIS	0,9318	1
Fuzzy-MLP	0,9318	1
Fuzzy-RBF	0,9545	1
Fuzzy-CSFNN	0,9772	1

4. Conclusions

Table 1-3 shows the performance comparisons of proposed Fuzzy-MLP, Fuzzy-RBF and Fuzzy-CSFNN hybrid schemes with ANFIS, MLP, RBF and CSFNN schemes. ANFIS uses hybrid learning rule, which combines the gradient method and the least squares estimate to identify parameters. That means, ANFIS contains adaptive algorithms and therefore it is not a fuzzy-neural hybrid schemes exactly. Another important point for CSFNN and Fuzzy-CSFNN schemes, elements of the output vector are considered as the numbers that are very close to 1. Training is faster than the other scheme's. Fuzzy-MLP, Fuzzy-RBF and Fuzzy-CSFNN structures are constituted, and their performances are investigated. Proposed hybrid schemes have better performances than ANFIS structure and non-hybrid schemes.

5. References

- [1] J.S.R. Jang, C.T. Sun, E. Mizutani, "Neuro-Fuzzy and Soft Computing", Prentice Hall, 1997.
- [2] R. Lee, J. Liu, "A Weather Forecasting System Using Intelligent Multiagent Based Fuzzy Neuro Network", *IEEE Transactions on Man&Cybernetics*, Vol.34, No.3, August 2004.
- [3] L. Özyılmaz, T. Yıldırım, "ROC analysis for fetal hypoxia problem by artificial neural networks", *Lecture Notes in Artificial Intelligence, special issue for Artificial Intelligence and Soft Computing*, LNAI Vol. 3070, 2004, pp. 1026-1030.
- [4] L. Özyılmaz, T. Yıldırım, H. eker, "EMG signal classification using conic section function neural networks", *Proceedings of IJCNN'99 International Joint Conf. on Neural Networks*, IEEE publication, USA, vol.5, pp. 3601-3603, 1999.
- [5] K.C. Kwak, D.H. Kim, "Adaptive Neuro-Fuzzy Networks with the Aid of Fuzzy Granulation", *IEICE Trans. INF. & SYST.* Vol.E88-D, No.9, September 2005.
- [6] C.S. Ouyang, S.J. Lee, "A Hybrid Learning Algorithm for Fuzzy Neural Networks", *International Conference on Neural Information Processing*, November 14-18, 2001.
- [7] S. Chandana, R.V. Mayorga, "RANFIS: Rough Adaptive Neuro-Fuzzy Inference System", *International Journal of Computational Inteligence*, Vol.3, No.4, 2006.
- [8] C.H. Lee, Y.C. Lin, "Hybrid Learning Algorithm for Fuzzy Neuro Systems", *IEEE International Conference on Fuzzy Systems*, Budapest, Hungary, 25-29 July 2004.
- [9] H. Seker, M.O. Odetayo, D. Petrovic, R.N.G. Naguib, F.C. Hamdy, "An Intelligent Hybrid Neuro-Fuzzy Rule-Based System For Prognostic Decision Making in Prostate Cancer Patients", *The 4th Annual IEEE Engineering in Medicine and Biology Society Special Topic Conference on Information Technology Applications in Biomedicine*, 24-26 April 2003, Birmingham, UK.
- [10] J.S.R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System", *IEEE Transactions on Systems, Man and Cybernetics*, Vol.23, No.3, Page.665-685, 1993.
- [11] G. Dorffner, "A Unified Framework for MLPs and RBFs: Introducing Conic Section Function Networks", *Cybernetics and Systems*, 25(4), 1994.
- [12] T. Yıldırım, : Development of Conic Section Function Neural Networks in Software and Analogue Hardware, Ph.D. Thesis, Liverpool University, UK, May (1997).
- [13] www.ics.uci.edu/pub/ml-repos/machine-learning-databases/thyroid-disease/
- [14] <http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29>
- [15] C. Senol, T. Yildirim, "A New Hybrid Approach for Fuzzy-Neural Networks", *International Symposium on Innovations in Intelligent Systems and Applications*, June 29-July 1, 2009, Karadeniz Technical University, Trabzon, Turkey.