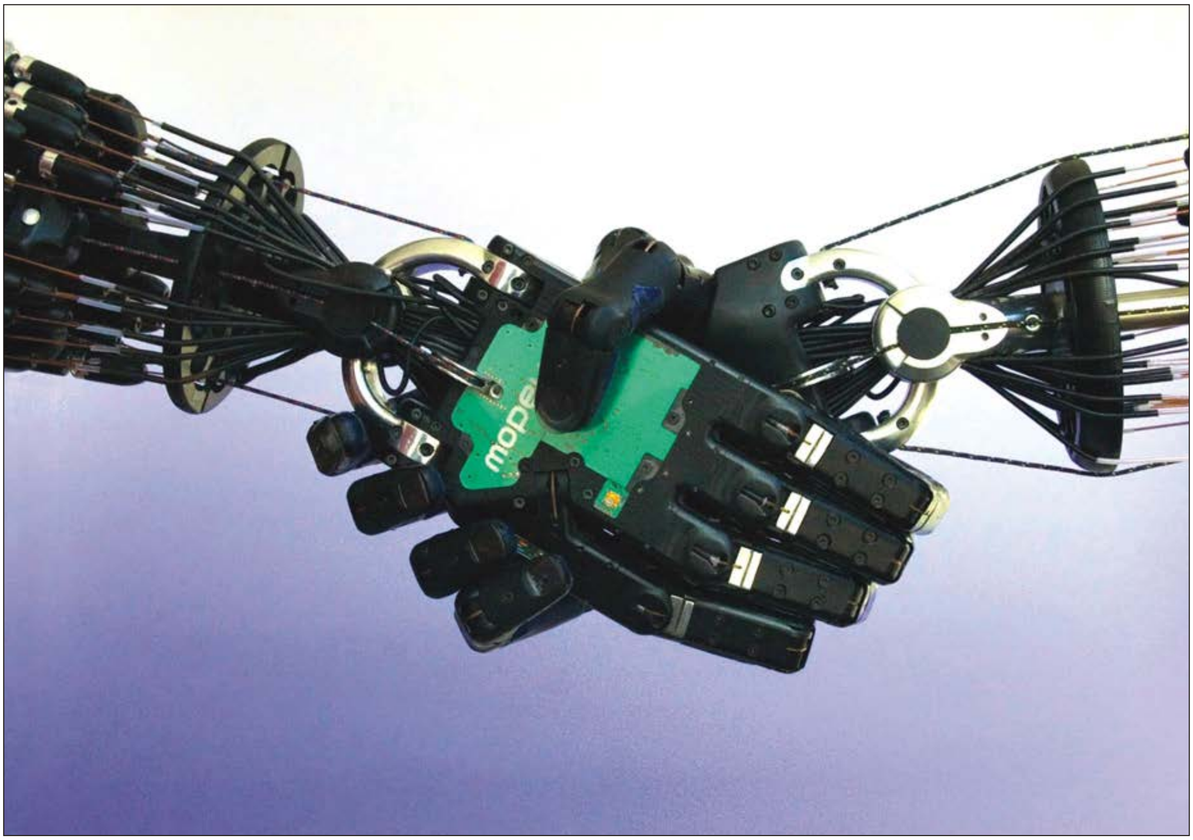

Journal of Artificial Intelligence

ISSN 1994-5450

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Artificial Intelligence 5 (1): 47-55, 2012
ISSN 1994-5450 / DOI: 10.3923/jai.2012.47.55

Diagnosis of Heart Disease using Fuzzy Resolution Mechanism

A.V. Senthil Kumar
Department of MCA, Hindusthan College of Arts and Science, Behind Nava, Coimbatore-641028, Tamil Nadu, India

ABSTRACT

The aim of this study is to combine the neural networks (ANNs) and Fuzzy Logic (FL) to make a powerful tool to diagnosis heart disease. By combining the Fuzzy inference system and neural network, the input values are passed through the input layer (by input membership function) and the output could be seen in output layer (by output membership functions). Training involves iterative adjustment of parameters of the adaptive neuro-fuzzy inference system using a hybrid learning procedure to diagnosis the heart disease. This mechanism presents five layer, each layer has its own nodes. Layer 1 had the input variables with membership function. T-norm operator that perform the AND operator can be used in layer 2. The sum of all rules firing strengths are assigned in layer 3. The nodes in layer 4 are adaptive and perform the consequent of the rules. Single node computes the overall output in layer 5. The proposed method is tested with Cleveland heart disease dataset. The ANFIS approach is implemented using MATLAB. The proposed mechanism can work more effectively for diagnosis of heart disease and also improves the accuracy. The result of the proposed methods is compared with earlier method using accuracy as metrics.

Key words: ANFIS, hybrid neural network, fuzzy resolution mechanism, heart disease

INTRODUCTION

Adaptive neuro fuzzy inference similar to that of a neural network which maps inputs through input membership functions and associated parameters and then through output membership functions and associated parameters to outputs, can be used to interpret the input/output result for diagnosis of heart disease. The learning process can be refined with the parameters associated with the membership functions.

Roan *et al.* (1993) used conventional crisp sets, the concepts of fuzzy sets provides more robust representations of the model of real-world objects. Jang (1993) employed a technique called adaptive neuro-fuzzy inference system (ANFIS). It employs a NN approach to the design of a fuzzy inference system. Kosko (1994) used Learning and adaptation of the Neural Network makes this fuzzy system more systematic and less reliant on knowledge of experts. It has been shown that under proper conditions, ANFIS can be used as a universal approximator. Serpen *et al.*(1997) developed probabilistic potential function neural network algorithm. Haykin (1999), the application of artificial intelligence approaches such as Neural Network (NN) and Fuzzy Logic (FL) do not require an explicit mathematical model and are suitable for nonlinear physiological systems. Moreover, they offer several advantages such as nonlinear input-output mapping, adaptivity and fault tolerance. One of the most commonly used classifier techniques is artificial neural networks. The reason for being commonly used is to present some properties such as learning from examples and exhibiting some capability for generalization beyond the training data (Mukhopadhyay *et al.*, 2002). A Lot of research has also been done about Cleveland heart disease database. A new learning model called granular support vector machines for data classification problems. The accuracy rates were 83.04 and 84.04% for SVM and GSVM, respectively (Tang *et al.*, 2004). Kahramanli and Allahverdi (2008) developed a hybrid system for diabetes and heart diseases using artificial neural network and fuzzy neural network. Liu *et al.* (2010) derived new ANFIS for parameter prediction with numeric and categorical inputs. Forouzanfar *et al.* (2010) oscillometric estimated blood pressure using adaptive neuro fuzzy inference system.

Heart disease is a large health problem. It is the leading cause of death for both men and women. Each year more than 500,000 people die of heart attacks caused by chronic heart failure. The increasing number of death due to heart disease worldwide has drawn the attention to adopt a technology such as Adaptive Neuro-Fuzzy Inference System for research. Experimental results indicate that the proposed fuzzy expert system can work more effectively than other methods can (Tang *et al.*, 2004; Kahramanli and Allahverdi, 2008; Senthil Kumar, 2011).

DIAGNOSIS TOOL FOR HEART DISEASE

Cleveland heart disease dataset: This database has one of the highest known heart disease data. The experimental Cleveland dataset is retrieved from the Internet (http://archive.ics.uci.edu/ml/) and it contains the collected personal data (Fig. 1). Table 1 lists the attributes of Cleveland dataset.

Fuzzification: The conversion from crisp to fuzzy input is known as fuzzification (Fasanghari and Montazer, 2010). If the form of uncertainty happens to arise because of imprecision, ambiguity or vagueness, the variable is probably fuzzy and can be represented by a membership function.

Architecture of the fuzzy resolution mechanism for heart disease: The fuzzy resolution mechanism can take fuzzy inputs but the output produced are always a fuzzy sets. With the crisp inputs and outputs, fuzzy resolution mechanism implements mapping from its input variable to output variable through a number of fuzzy if then rules. Cleveland dataset is taken with fourteen attributes, i.e., age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal are

Table 1: Attributes of cleveland dataset

Abbreviation	Full name
Age	Age in years
Sex	Sex (1 = male; 0 = female)
cp	Chest pain type
Trestbps	Resting blood pressure (mm Hg)
chol	Serum cholesterol (mg dL^{-1})
fbs	Fasting blood sugar >120 mg dL^{-1}
restecg	Resting electrocardiographic results
thalach	Maximum heart rate achieved
exang	Exercise induced angina
oldpeak	ST depression induced by exercise relative to rest
slope	The slope of the peak exercise ST segment
ca	Number of major vessels (0-3) colored by flourosopy
thal	3 = normal; 6 = fixed defect; 7 = revers able defect
Num the predicted attribute	Diagnosis of heart disease (angiographic disease status)

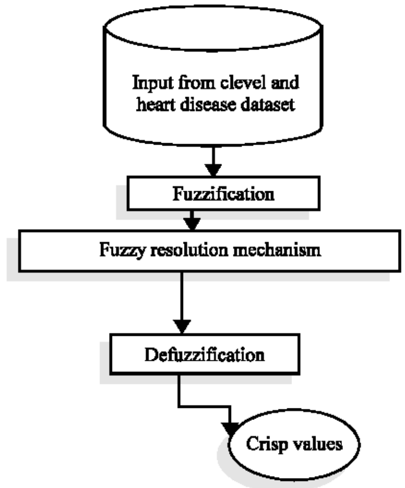

Fig. 1: Architecture of the diagnosis of heart disease

selected as the input fuzzy variables and num as output fuzzy variable are adopted for fuzzy resolution mechanism. For fuzzy resolution mechanism the ANFIS approach learns the rules and membership functions from Cleveland dataset.

The neuro-adaptive learning techniques provide a method for the fuzzy modeling procedure to learn about heart disease from cleveland data set, in order to compute the membership function parameters. An adaptive network is network of nodes and directional links. Associated with the network is a learning rule-hybrid method. Networks are learning a relationship between inputs and outputs. The architecture of the fuzzy resolution mechanism using ANFIS is shown in Fig. 2. The circular nodes represent nodes that are fixed whereas, the square nodes are nodes that have parameters to be learnt. During training, all of the training dataset would be present to network and it tries by learning.

Layer 1: The node function in Layer 1 is the membership of the fuzzy set associated with the corresponding input. The first order Sugeno fuzzy model provides the following rule based structure by Sadighi and Kim (2011).

If Age is young and sex is female and cp is low and trestbps is medium and chol is high and fbs is low and restecg is high and thalach is low and exang is high and oldpeak is medium and slope is high and ca is high and thal is normal then f1 = age.x+sex.y+cp.z+trestbps.a+chol.b+fbs.c+ restecg.d+thalach.e+exang.f+oldpeak.g+slope.h+ca.i+thal.j+t.

where, high, medium, low, young, normal are fuzzy sets for the input, {Age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal} is the consequent parameter set and f1 is the output. The membership functions where parameterized using the generalized Gaussian function (Alves *et al*., 2011):

$$O_i^{1,3} = \mu_{cp}(x) = e^{-\frac{1}{2}\left(\frac{x-c_i^1}{\sigma_i^1}\right)^2} \quad (1)$$

where, c and σ represent the membership function center and width respectively in order to determine coordinates of Gaussian membership function.

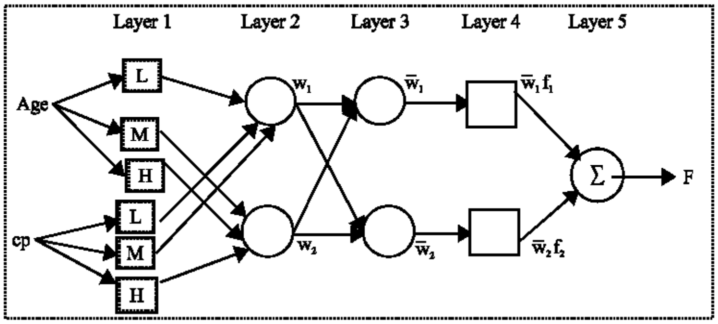

Fig. 2: Architecture of the fuzzy resolution mechanism using ANFIS

Layer 2: T-norm operator that perform the AND operator can be used (Liu *et al.*, 2010). Every node in this layer is a fixed node labeled Prod. The output is the product of all the incoming values. In layer 2, multiplies the inputs from the nodes in layer 1 and generates the firing strength of the rules. The output of this layer is given by:

$$w_i = \mu_{Age}(x)\ \mu_{sex}(y)\ \mu_{trestbps}(\alpha)\ \mu_{chol}(b)\ \mu_{fbs}(c)\ \mu_{restecg}(d)\ \mu_{thalach}(e)\ \mu_{exang}(f)\ \mu_{oldpeah}(g)\ \mu_{slope}(h)\ \mu_{thal}(j)\ i = 1, 2$$

where, w_i is the firing strength of rule i.

Layer 3: Regularize all the rules firing strengths and node function in Layer 3. The ith node calculates the ration of the ith rules firing strength to the sum of all rules firing strengths:

$$= \overline{w}_i = \frac{w_i}{\sum_{i=1}^{m} w_i}$$

where, the output are called normalized firing strengths is of this layer.

Layer 4: The nodes in this layer are adaptive and perform the consequent of the rules:

$$\overline{w}_i f_i = \overline{w}_i (age_i.x+sex_i.y+cp_i.z+trestbps_i.a+chol_i.b+fbs_i.c+restecg_i.d+thalach_i.e+exang_i.f+oldpeak_i.g+slope_i.h+ca_i.i+thal_i.j+t_i)$$

where, $\overline{w^{i}}_{i}$ is a normalized firing strength from layer 3 and {Age_i, sex_i, cp_i, $trestbps_i$, $chol_i$, fbs_i, $restecg_i$, $thalach_i$, $exang_i$, $oldpeak_i$, $slope_i$, ca_i, $thal_i$, t_i} are the parameter set of this node. Parameters in this layer are referred to as consequent parameters.

Layer 5: There is a single node that computes the overall output:

$$\sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

The input vector is fed through the network layer by layer. We now consider how the ANFIS learns the premise and consequent parameters for the membership functions and the rules.

ALGORITHM FOR FUZZY RESOLUTION MECHANISM

Input: Input the fuzzy set for Age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca and thal.

Output: Output the fuzzy set for num(angiographic disease status).

Method: Begin:

- **Step 1:** Input the crisp values for Age, sex.cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca and thal
- **Step 2:** Set first order sugeno fuzzy model, common rule set with fuzzy if-then rules

Input the rule as {Rule 1,2…..k}

- **Step 3:** ANFIS is executed by Sugeno method
- **Step 4:** Layer 1-every node is an adaptive node with node function

$$O_i^{1,1} = \mu_{age}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,2} = \mu_{sex}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,3} = \mu_{cp}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,4} = \mu_{trestbps}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,5} = \mu_{chol}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,6} = \mu_{fbs}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,7} = \mu_{restecg}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,8} = \mu_{thalach}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,9} = \mu_{exang}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,10} = \mu_{oldpeak}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,11} = \mu_{slope}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,12} = \mu_{ca}(x), \text{ fori} - 1,2,3$$
$$O_i^{1,13} = \mu_{thal}(x), \text{ fori} - 1,2,3$$

where, x is input to node and Age, sex.cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca and thal is a linguistic label associated with this node.

- **Step 4.1:** Set the Gaussian function membership function for the fuzzy number with Eq. 1
- **Step 5:** In layer 2, multiplies the inputs from the nodes in layer 1 and generates the firing strength of the rules. T-norm operator that perform the AND operator is used
- **Step 6:** Layer 3 contains fixed nodes. The i^{th} node calculates the ration of the i^{th} rules firing strength to the sum of all rules firing strengths

$$= \overline{w_i} = \frac{w_i}{\sum_{i=1}^{m} w_i}$$

- **Step 7:** In Layer 4, the nodes in this layer are adaptive and perform the consequent of the rules
- **Step 8:** There is a single node here that computes the overall output

$$\sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

- **Step 9:** Present the knowledge in the form of human natural language

End.

Defuzzification: Defuzzification process is conducted to convert aggregation result into crisp value for num output. In this process the single number represent the outcome of the fuzzy set evaluation. The final combined fuzzy conclusion is converted into a crisp value by using the weighted average method (Alves *et al*., 2011).

EXPERIMENTAL RESULTS

The proposed Fuzzy Resolution Mechanism for heart disease was implemented with the MATLAB. The experimental environment was constructed to evaluate the performance of the proposed approach with Cleveland data set. The first experiment shows sets of results in Fig. 3 and 4, indicating that the proposed approach automatically supports the analysis of the data.

The Decision can be taken from the about the status of angiographic disease. Neural network was developed with the training data, with rule, parameter and membership function.

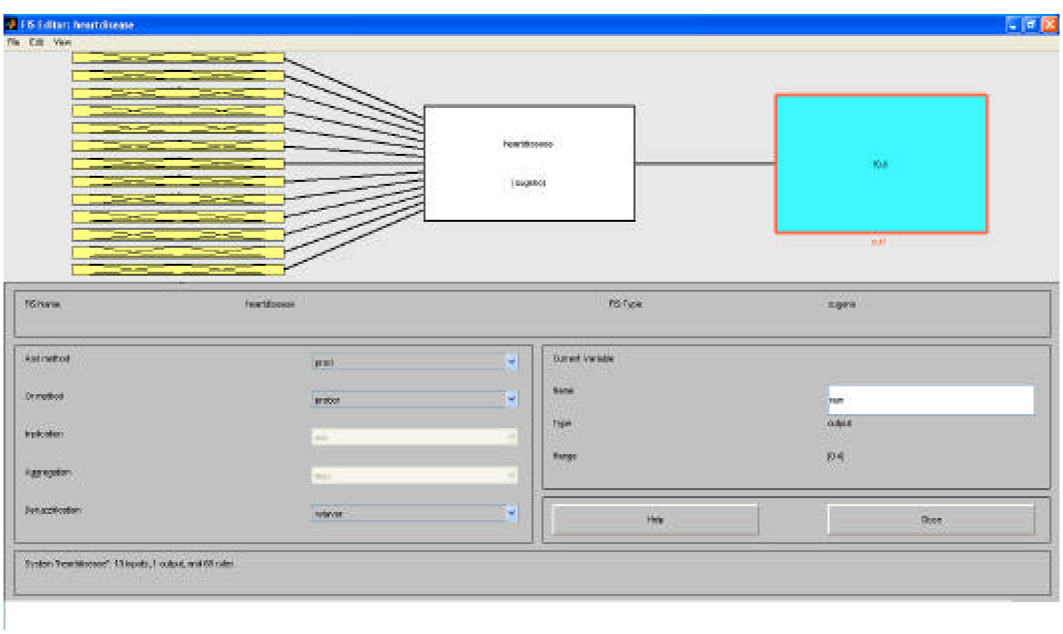

Fig. 3: ANFIS modeling framework

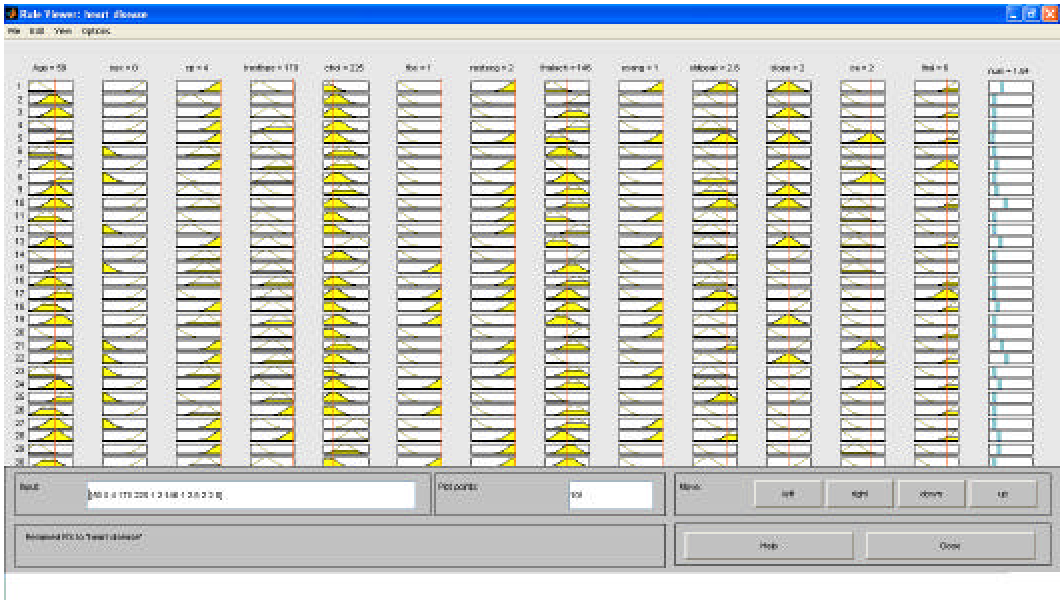

Fig. 4: Result obtained from MATLAB

Table 2: Comparison of proposed method accuracy with earlier methods

Method	Accuracy (%)	Author
Current study	91.83	Senthil Kumar (2011)
Adaptive Neuro-Fuzzy Inference System for Heart Disease diagnosis	91.18	Senthil Kumar (2011)
IncNet	90	Jankowski and Kadirkamanathan (1997)
Hybrid system	86.8	Kahramanli and Allahverdi (2008)
28-NN, stand, Euclid, 7 features	85.1±0.5	WD/KG (Dich *et al.*, 1998)
LDA	84.5	Ster and Dobnikar (1996)
Fisher discriminant analysis	84.2	Ster and Dobnika (1996)
k = 7, Euclid, std	84.2±6.6	WD, GhostMiner (Duch and Gridzinski, 1998)
16-NN, stand, Euclid	84±0.6	WD/KG (Dich *et al.*, 1998)
FSM, 82.4-84% on test only	84.0	Adamczak *et al.* (1998)
k = 1:10, Manhattan, std	83.8±5.3	WD, GhostMiner (Duch *et al.*, 1998)
Naive Bayes	82.5-83.4	Ster and Dobnikar (1996)

EVALUATION OF SYSTEM PERFORMANCE

The second experiment evaluates the performance of the system. Accuracy is the common performance metrics used in medical diagnosis tasks. The measure of the ability of the classifier to produce accurate diagnosis is determined by accuracy. So that accuracy Loo and Rao (2005) is given by Eq. 2:

$$\text{Accuracy} = \frac{\text{Total number of correctly diagnosed cases}}{\text{Total number of cases}} \tag{2}$$

The final experiment compares the accuracy of the proposed method with results of studies involving the Cleveland heart disease dataset (Tang *et al.*, 2004; Kahramanli and Allahverdi, 2008;

Senthil Kumar, 2011). Comparing these methods, as listed in Table 2, reveals that the proposed method achieves the first highest accuracy values based on the proposed Fuzzy Resolution Mechanism.

CONCLUSIONS AND FUTURE RESEARCH

Fuzzy Resolution Mechanism is used to diagnosis the heart disease. The experimental, Cleveland heart disease dataset, is initially processed and the crisp values are converted into fuzzy values in the stage of fuzzification. The Fuzzy Resolution Mechanism undergoes five layers to execute rules, to make a decision on the possibility of individuals suffering from heart disease. Defuzzification process is conducted to convert the result into crisp value for angiographic disease status. Experimental results indicate that the proposed method can analyze data more efficiently than other methods. Future works should test the Fuzzy Resolution Mechanism approach used herein for other similar tasks or other related data sets to evaluate its capability to produce a similar accuracy.

REFERENCES

Alves, M.C., E.A. Pozza, J.C.B. Costa, L.G. de Carvalho and L.S. Alves, 2011. Adaptive neuro-fuzzy inference systems for epidemiological analysis of soybean rust. Environ. Modell. Software, 26: 1089-1096.

Dich, W., K. Grudzinski and G.H.F. Diercksen, 1998. Minimal distance neural methods. Proceedings of the World Congress of Computational Intelligence, May 4-9, 1998, Anchorage, Alaska, USA., pp: 1299-1304.

Duch, W. and K. Grudzinski, 1998. A framework for similarity-based methods. Proceedings of the 2nd Polish Conference on Theory and Applications of Artificial Intelligence, September 28-30, 1998, Lodz, Poland, pp: 33-60.

Duch, W., R. Adamczak, K. Gr¹bczewski and G. Zal, 1998. A hybrid method for extraction of logical rules from data. Proceedings of the 2nd Polish Conference on Theory and Applications of Artificial Intelligence, September 28-30, 1998, Lodz, Poland, pp: 61-82.

Fasanghari, M. and G.A. Montazer, 2010. Design and implementation of fuzzy expert system for Tehran stock exchange portfolio recommendation. Expert Syst. Appl., 37: 6138-6147.

Forouzanfar, M., H.R. Dajani, V.Z. Groza, M. Bolic and S. Rajan, 2010. Adaptive neuro-fuzzy inference system for oscillometric blood pressure estimation. Proceedings of the IEEE International Workshop on Medical Measurements and Applications, April 30-May 1, 2010, Ottawa, Canada, pp: 125-129.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. 2nd Edn., Prentice-Hall International Inc., New Jersey, USA., ISBN-13: 9780139083853, Pages: 842.

Jang, J.S.R., 1993. ANFIS: Adaptive network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern., 23: 665-685.

Jankowski, N. and V. Kadirkamanathan, 1997. Statistical control of RBF-like networks for classification. Proceedings of the 7th International Conference on Artificial Neural Networks, October, 1997, Lausanne, Switzerland, pp: 385-390.

Kahramanli, H. and N. Allahverdi, 2008. Design of a hybrid system for the Diabetes and heart diseases. Expert Syst. Appl., 35: 82-89.

Kosko, B., 1994. Fuzzy systems as universal approximators. IEEE Trans. Comput., 43: 1329-1333.

Liu, M., M. Dong and C. Wu, 2010. A new ANFIS for parameter prediction with numeric and categorical inputs. IEEE Trans. Autom. Sci. Eng., 7: 645-653.

Loo, C.K. and M.V.C. Rao, 2005. Accurate and reliable diagnosis and classification using probabilistic ensemble simplified fuzzy ARTMAP. IEEE Trans. Knowledge Data Eng., 17: 1589-1593.

Mukhopadhyay, S., C. Tang, J. Huang, M. Yu and M. Palakal, 2002. A comparative study of genetic sequence classification algorithms, neural networks for signal processing. Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing, September 4-6, 2002, Martigny, Switzerland, pp: 57-66.

Roan, S.M., C.C. Chiang and H.C. Fu, 1993. Fuzzy RCE neural network, fuzzy systems. Proc. IEEE Int. Conf. Fuzzy Syst., 1: 629-634.

Sadighi, A. and W.J. Kim, 2011. Adaptive-neuro-fuzzy-based sensorless control of a smart-material actuator. IEEE/ASME Trans. Mechatron., 16: 371-379.

Senthil Kumar, A.V., 2011. Adaptive neuro-fuzzy inference system for heart disease diagnosis. Proceedings of the International Conference on Information System, Computer Engineering and Application, November 28-29, 2011, Singapore.

Serpen, G., H. Jiang and L.G. Allred, 1997. Performance analysis of probabilistic potential function neural network classifier. Proc. Artif. Neural Networks Eng. Conf., 7: 471-476.

Ster, B. and A. Dobnikar, 1996. Neural networks in medical diagnosis: Comparison with other methods. Proc. Int. Conf. EANN, 1: 427-430.

Tang, Y., B. Jin, Y. Sun and Y.Q. Zhang, 2004. Granular support vector machines for medical binary classification problems. Proceedings of the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, October, 7-8, 2004, USA., pp: 73-78.