Work-In-Progress: An intelligent diagnosis influenza system based on adaptive neuro-fuzzy inference system

Sheng-Ta Hsieh[1] and Chun-Ling Lin[2,*]

[1]Department of Communication Engineering, Oriental Institute of Technology, New Taipei City, Taiwan, R.O.C.
[2]Department of Electrical Engineering, Ming Chi University of Technology, New Taipei City, Taiwan, R.O.C.
ginnylin@mail.mcut.edu.tw

Abstract—**This study combines adaptive neuro-fuzzy inference system (ANFIS) with greedy forward feature selection to develop the intelligent diagnosis system. Two different membership functions (MFs), Trapezoidal and Gaussian, are adopted during the training process of ANFIS in order to compare the diagnosis accuracy of Trapezoidal MF with one of Gaussian MF. The comparison of ANFIS values with simulated data indices that adoption of both Trapezoidal and Gaussian MF in proposed system achieve satisfactory accuracy (>96%). Furthermore, the accuracy of ANFIS with Gaussian MF is above 98%. Hence, the intelligent diagnosis system can provide a preliminary result to physicians so that the doctor could quickly and accurately decide whether patient have cold or influenza.**

Keywords-component; Adaptive neuro-fuzzy inference system (ANFIS), Membership function (MF), Greedy forward feature selection, Intelligent diagnosis system

I. INTRODUCTION

Influenza is an acute respiratory disease caused by "flu virus". Early constitutional symptoms of influenza are very severe, including fever, headache, muscle pain, fatigue, rhinorrhea, sore throat and cough, which affect the work performance. In the high-risk populations of patients, it may result in serious health problems such as pneumonia and hospitalizations and even cause deaths. According to Centers for Disease Control and Prevention (CDC) reports [1], in approximately 5,000-10,000 patients, one person could die from the influenza.

When you sneeze and cough a lot, and furthermore have fever and muscle pain, how do you know whether you have cold or influenza? Although both of cold and influenza are viral infection, cold and influenza have different pathogeny and the degree of symptoms. The impact of influenza has been not only economic but also serious health problems such as pneumonia and hospitalizations. Therefore, it is certainly important to distinguish between cold and influenza. Since patients' descriptions of their own symptoms should in uncertainty and vagueness, vague terms of symptoms are unavoidable. Hence, fuzzy set is an efficient tool because it is able to work with problems where no problem definitions are possible. The imprecise descriptions of symptoms may be dealt with using fuzzy rules in a fuzzy inference system, where the conditional part and/or the conclusions contain linguistic variables. By this approach, the expert and /or physicians does not need to know the precise symptoms.

ANFIS is a class of adaptive networks that is functionally equivalent to fuzzy inference systems (FIS) [2]. ANFIS rely on combining neural networks with fuzzy logic to create powerful expert decision systems. ANFIS can construct an input-output mapping based on both human knowledge (in the form of fuzzy if-then rules) and stipulated input–output data pairs. ANFIS offers a powerful framework for tackling practical classification problems and is able to create class boundaries that reduce its misclassification rates. In addition, ANFIS should be able to learn in a short time, soft and hard decisions. Thus, many previous studies adopted the ANFIS to develop an intelligent diagnosis system for different diseases [3-6].

The aim of this study is to create a reliable and effective methodological tool, ANFIS, for diagnosing normal, cold and influenza. The output of ANFIS can provide a preliminary result to physicians so that the doctor could quickly and accurately decide whether patient have cold or influenza.

II. MATERIAL

This study creates 60 simulated people including 22 patients with cold, 25 patients with influenza, and 13 normal person. In the simulated data, there are eight types of symptoms (i.e., period of onset, temperature, cough, headache, muscle pain, fatigue, nasal congestion and sore throat) and then those symptoms have the corresponding disease (normal, cold and influenza). The experienced doctor (Division of Family Medicine, Far Eastern Memorial Hospital) decides whether simulated people have cold, influenza or normal according to the degree of eight types of symptoms. The definitions of degree of eight type symptoms are shown in Table I.

TABLE I. THE DEFINITION OF DEGREE OF EIGHT TYPE SYMPTOM

Symptoms	Degree
Period of onset	progress step by step or suddenly (1, 2)
Temperature	$<38^o$ or $\geqq 38^o$ (1, 2)
Cough	Ten degree of symptom (0-9)
Headache	Ten degree of symptom (0-9)
Muscle pain	Ten degree of symptom (0-9)
Fatigue	Ten degree of symptom (0-9)
Nasal congestion	Ten degree of symptom (0-9)
Sore throat	Ten degree of symptom (0-9)

III. AN INTELLIGENT DIAGNOSIS INFLUENZA SYSTEM

A. Feature selection

In this study, eight types of symptoms could be as the input data in ANFIS. Although ANFIS can create powerful expert decision systems, the size of the input-output data set is very crucial. The computation of ANFIS become expensively and complexly when the dimension of input data grows up. Furthermore, there are a very large number of features and only few of them are relevant for distinguishing between cold and influenza. Feature selection is the task of selecting a small and adaptive set out of a given set of features that capture the relevant properties of the data. In this study, greedy forward feature selection [7] are adopted to avoid the drawbacks of ANFIS. The greedy forward feature selection repeatedly iterates the feature set and updates the set of chosen features. After applying the greedy forward feature selection, appropriate features are presented and fed into the ANFIS.

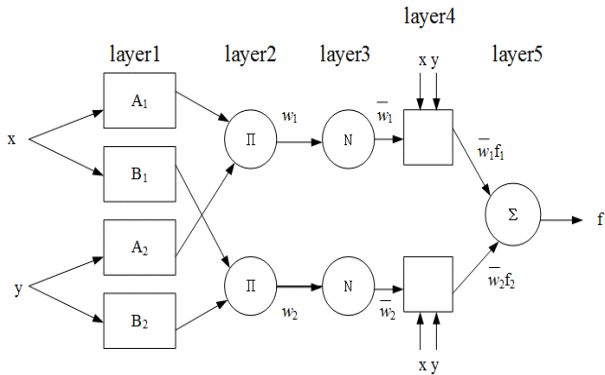### B. Adaptive neural fuzzy inference system (ANFIS)

Figure 1. ANFIS architecture

The ANFIS is a class of adaptive multi-layer feed-forward neural networks that are functionally equivalent to fuzzy inference systems [2]. The ANFIS is typically used to formalize a systematic approach to generating fuzzy rules from an input-output dataset. The parameters for training the fuzzy inference system (FIS) membership function are determined by the ANFIS and using the Matlab function "anfis" (Matlab fuzzy logic toolbox). Given an input/output dataset, the "anfis" function constructs an FIS whose membership function parameters are adjusted using the back propagation algorithm with a least squares method [2, 8]. This allows the FIS to learn from input data. The procedure is briefly described as follows. First, the Matlab function "genfis1" (Matlab fuzzy logic toolbox) generated a Sugeno-type FIS structure [9] as the initial conditions (initialization of membership function parameters) for the function "anfis". For simplicity, we assume a network with two inputs, x and y, and one output f. Eq. (1) shows a common Sugeno-type FIS structure and Fig. 1 delineates the regular framework of ANFIS

$$\text{Rule1: if } x_1 \text{ is } A_1, \text{and } x_2 \text{ is } B_1, \text{then } f_1 = \alpha_1 x_1 + \beta_1 x_2 + \gamma_1$$
$$\text{Rule2: if } x_1 \text{ is } A_2, \text{and } x_2 \text{ is } B_2, \text{then } f_1 = \alpha_2 x_1 + \beta_2 x_2 + \gamma_2$$

$$(1)$$

To generate the Sugeno fuzzy inference system (FIS), two MFs are used: (1) Trapezoidal (function "trapmf", Matlab fuzzy logic toolbox) and (2) Gaussian (function "gaussmf", Matlab fuzzy logic toolbox) membership function. For each type of membership function, three input of MFs are used for each variable category. The input matrix of training data included input data (selected feature set from the degree of eight types of symptoms) and output data (normal, cold or influenza). When the functions "genfis1" are defined, the "anfis" function started training. The parameters associated with MFs are updated during the learning processes. The convergence of these parameters (or adjustments) is facilitated by a gradient vector v_{ij}. This gradient vector assessed how well the FIS modeled data for a given set of parameters. When the gradient vector is acquired, optimization routines in Eq. (2) – (4) are applied to adjust parameters to reduce error (E). Error (E) is defined as the sum of squared differences between actual output f and desired output y_d.

$$\text{E} = (y_d - f)^2 \qquad (2)$$

$$\Delta v_{ij} = -\eta \frac{\partial E}{\partial v_{ij}} \qquad (3)$$

$$v_{ij}(h+1) = v_{ij} + \Delta v_{ij}(h) \qquad (4)$$

where v_{ij} is used to adjust parameters, with a step of Δv_{ij}.

The "anfis" function either used back propagation or back propagation in combination with least squares estimation to estimate the parameters of membership function. The parameters for the FIS are set according to a minimum training-error criterion. Output variables for the FIS are estimated by fuzzy inference calculation function "evalfis" (Matlab fuzzy logic toolbox), which has two arguments, including parameter "fis" and a training dataset.

C. The structure of intelligent diagnosis influenza system

In order to develop the intelligent diagnosis system with adaptive feature set, this study applies both greedy forward feature selection and ANFIS. The leave-one-subject-out

(LOSO) cross-validation [10] is used to train and test the intelligent diagnosis influenza system. This study adopts the accuracy (ACC) and sensitivity to assess the performance of intelligent diagnosis influenza system. ACC indicates proximity of measurement results to the true value, precision to the repeatability or reproducibility of the measurement. Sensitivity is a statistical measure of the performance of a binary classification test. Sensitivity can reflect the ability to identify positive result.

IV.   RESULTS AND DISCUSSION

After applying both greedy forward feature selection and ANFIS, the adaptive feature set is period of onset and muscle pain because the performance combining ANFIS with those two feature as input data can highest than other feature sets regardless of type of membership functions. That is, the intelligent diagnosis influenza system can be structure by ANFIS with two input data (period of onset and muscle pain).
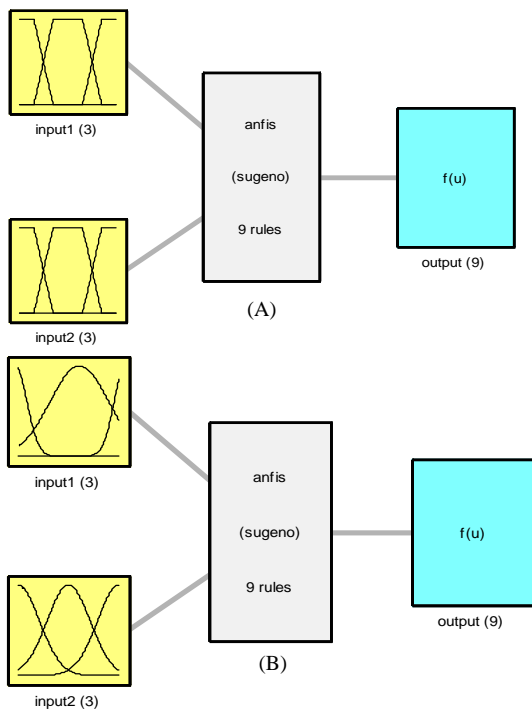

Figure 2 (A) The fuzzy inference system using Trapezoidal MF. (B) The fuzzy inference system using Gaussian MF

In this study, two MFs are used: (1) the Trapezoidal MF and (2) Gaussian MF. In this study, ANFIS is trained and tested using LOSO cross-validation. That is, a total of 59 sets of data are selected from total of 60 sets obtain in simulated people for the purpose of training in ANFIS. In addition, a remnant simulated people is used as testing data to verify the accuracy of the ANFIS. This validation method repeats 60 times until each simulated people is used as the testing data. This training adjusts the parameters of two MFs. Fig. 2 shows the final fuzzy inference system of two MFs after training. It seems that Gaussian MF have significant modifications through the learning process. In order to confirm the modifications of Gaussian MF, Fig. 3 shows the initial and final Gaussian MF before and after training.
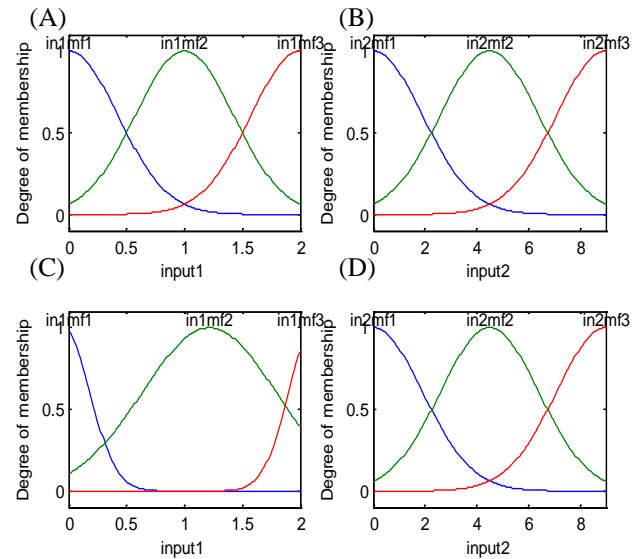

Figure 3 (A) and (B) are the Gaussian MF before training. (C) and (D) are the Gaussian MF after training

During training in ANFIS, ANFIS learning numbers with two MFs are as follow: (1) Trapezoidal MF=nodes, 35; linear parameters, 27; nonlinear parameters, 24; and fuzzy rules, 9; (2) Gaussian MF=nodes, 35; linear parameters, 27; nonlinear parameters, 12; and fuzzy rules, 9. Table II shows accuracy and sensitivity of two MFs for the testing data. When Trapezoidal MF is used in ANFIS, the accuracy is 96.67%. When the Gaussian MF is adopted, the accuracy is 98.33%. Furthermore, the sensitivity of cold in Gaussian MF is more than in Trapezoidal MF. These results indicated that the training of ANFIS with Gaussian MF obtained a higher diagnosis accuracy than Trapezoidal MF.

TABLE II.   THE ACCURACY AND SENSITIVITY OF TWO MFs FOR THE TESTING DATA.

Membership functions	Accuracy (%)	Sensitivity (%)		
		Normal	Cold	Influenza
Trapezoidal	96.67	100	95.45	96
Gaussian	98.33	100	100	96

This study adopted ANFIS to develop an intelligent diagnosis influenza system. However, there is still plenty of room for improvement of this intelligent diagnosis influenza system. Support vector machine (SVM) is considered as a novel machine learning method based on statistical learning theory [11]. It can provide a unique solution with a strongly regularized characteristic which is quite appropriate for ill-posed classification problems. In future work, we will combine ANFIS with SVM to develop our diagnosis influenza system.

V.   CONCLUSION

This study combines ANFIS with greedy forward feature selection to develop the intelligent diagnosis system. The 60 simulated people with different symptoms are used to train and test the intelligent diagnosis system. Two different MFs, Trapezoidal and Gaussian, are adopted during the training

process of ANFIS. The comparison indicates that the Gaussian MF in ANFIS achieve slightly higher diagnosis accuracy than Trapezoidal MF. The accuracy of ANFIS with Gaussian MF is above 98%. Thus, the intelligent diagnosis system can provide a preliminary result to physicians so that the doctor could quickly and accurately decide whether patient have cold or influenza.

REFERENCES

[1] S. A. Harper, K. Fukuda, T. M. Uyeki, N. J. Cox, and C. B. Bridges, "Prevention and control of influenza," *Mmwr,* vol. 53, pp. 1-40, 2004.

[2] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *Systems, Man and Cybernetics, IEEE Transactions on,* vol. 23, pp. 665-685, 1993.

[3] M. A. Abushariah, A. A. Alqudah, O. Y. Adwan, and R. M. Yousef, "Automatic Heart Disease Diagnosis System Based on Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) Approaches," *Journal of Software Engineering and Applications,* vol. 7, p. 1055, 2014.

[4] J. Nourozi, M. M. Mazdeh, and S. A. Mirbagheri, "The effects of the underlying disease and serum albumin on GFR prediction using the Adaptive Neuro Fuzzy Inference System (ANFIS)," *Journal of Health Management and Informatics,* vol. 1, pp. 46-50, 2014.

[5] J. Zhang, W. Ma, and L. Ma, "A fault diagnosis method based on ANFIS and bearing fault diagnosis," in *Information Science, Electronics and Electrical Engineering (ISEEE), 2014 International Conference on*, 2014, pp. 1274-1278.

[6] K. Polat and S. Güneş, "An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease," *Digital Signal Processing,* vol. 17, pp. 702-710, 2007.

[7] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research,* vol. 3, pp. 1157-1182, 2003.

[8] J.-S. R. Jang, "Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm," in *AAAI*, 1991, pp. 762-767.

[9] J. Jang, "Fuzzy inference systems," *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence,* pp. 73-91, 1997.

[10] J. A. Rice and B. W. Silverman, "Estimating the mean and covariance structure nonparametrically when the data are curves," *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 233-243, 1991.

[11] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 42, pp. 1778-1790, 2004.