IEEE Int. Workshop VLSI Design & Video Tech.
Suzhou, China, May 28~30, 2005

CREDIT RISK EVALUATION WITH FUZZY NEURAL NETWORKS ON LISTED CORPORATIONS OF CHINA

Xiong Zhi-Bin and Li Rong-Jun
College of Business Administration, South China University of Technology
Guangzhou city, Guangdong Province, 510640, P.R.China

ABSTRACT*

Neural networks (NNs) have been widely used to evaluate credit risk because of their excellent performances of treating non-linear data with learning capability. However, the shortcoming of neural networks is also significant due to a "black box" syndrome and the difficulty in dealing with qualitative information, which limited its applications in practice. To overcome these drawbacks of NNs, in this study we suggested an adaptive network-based fuzzy inference system (ANFIS), a kind of fuzzy neural network models, to evaluate credit risk on the Chinese listed corporations. The results of this study indicate that the predictive accuracies of ANFIS model is much better than NNs model. An illustrative example is given for demonstration.

Keywords: Neural network; Fuzzy system; Credit risk.

1. INTRODUCTION

Credit risk is one of the three major risks that are usually faced by banks and financial institutions. The traditional methods used to evaluate credit risk, such as discriminant analysis, *KMV* model, Credit-Risk+, Credit-metrics, Credit -Portfolio-View and *VaR* etc, are all based on probability theory and statistical analysis, in which some kinds of distributions are often assumed previously. As a matter of fact, these assumptions are not reasonable and non-realistic. Besides, the statistical models are more or less lack of discrimination in some cases due to the non-structured characteristics of credit risks. Therefore, neural network models are developed and applied quickly to credit risk evaluation in recent years because of their excellent performances of treating non-linear data with learning capability (Weat Robert Craig, 1985, Tam and Kiang 1992, Coats and Fant 1993, Tan 1996, Patricia et al, 2000 and Caideron et al, 2002). However, the shortcoming of neural networks is also significant due to a "black box" syndrome and the difficulty in dealing with qualitative information, which limited its applications in practice (Arnold F. Shapiro, 2002, Selwyn Piramuthu, 1999). Besides, a relatively slow convergence speed is also the disadvantage of NNs model. On the other hand, fuzzy logic (FL) as a rule-based development in artificial intelligence can not only tolerate imprecise information, but also make a framework of approximate reasoning. The disadvantage of fuzzy logic is lack of effective learning capability. Thus, a number of fuzzy neural networks have been developed to overcome the drawbacks mentioned above (Selwyn Piramuthu, 1999, Rashmi Malhotra, 2002).

In this study an adaptive network-based fuzzy inference system (ANFIS), a kind of fuzzy neural network model, is suggested to evaluate credit risk on the Chinese listed corporations. The results of the comparative research conducted in this paper indicate that the predictive accuracies of ANFIS model are much better than the predictive accuracies of NNs model in credit risk evaluation.

The rest of this paper is organized as follows: Section 2 describes the data source and methodology; Section 3 explains the structures of NNs model and FNN model; An empirical analysis is shown in section 4 and finally some concluding remarks are drawn from section 5.

2. DATA SOURCE AND CLASSICATION

The data of samples used in this study are selected from the listed corporations of China. Those corporations can be

* The research is supported by Chinese NSF(no.70371029).

divided into two categories: one category consists of a set of ST (special treated) corporations and the other one consists of non-ST corporations. Suppose that ST category is considered as default (insolvency) corporations and non-ST as non-default (solvency) corporations.

According to the proportion of ST corporations in China stock market, a set of data with 225 observations is selected from 1998 to 2001, including 45 ST and 180 non-ST corporations. The detail of the number of selected corporations is shown in Table 1.

Table 1 Number of selected corporations

Periods	1998	1999	2000	2001	Sum
ST-corp.	7	17	11	10	45
Non-ST	28	68	44	40	180
Total					225

The data set is divided into two subsets: one is a training sample set with 150 corporations including 30 ST corporations and 120 non-ST corporations, used to design the NN model and the FNN model; another is a test sample set with 75 corporations including 15 ST corporations and 60 non-ST corporations to test the performance of models.

Altman financial ratios have been used widely to assess corporate financial distress, which include:

(1) Working capital / total assets ($X1$),

(2) Retained earnings / total assets ($X2$),

(3) Earnings before interest and taxes / total assets ($X3$),

(4) Sales / total assets ($X4$),

(5) Market value of equity / book value of total debts($X5$)

Since the stock market of China is not fully effective, in which there exists serious equity segmentation, and the prices of shares do not reflect the actual situations of the listed corporations sometimes, $X5$ is then dropped out and only $X1$, $X2$, $X3$ and $X4$ are used as input variables.

3. MODEL STRUCTURE

3.1 Neural networks model

A feed-forward NNs model consists of layers of processing units, where the input signals are fed forward from the input layer through the networks to the processing units in output layer. The back propagation (BP) algorithm has been one of the most popular learning algorithms.

In this study, a NNs model with one hidden layer is designed and shown in Fig.1. The number of processing units can be selected by trial and error or by heuristics.

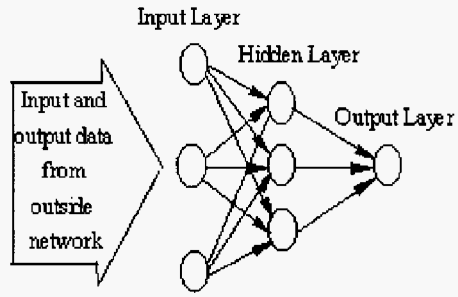

Fig.1. Structure of NNs model

3.2 Fuzzy Neural Networks Model

In this study, fuzzy neural networks model is a class of adaptive networks that are functionally equivalent to fuzzy inference system, i.e. adaptive network-based fuzzy inference system (ANFIS) (Jan et al, 1997). ANFIS is a network-type structure that maps inputs through input membership functions and associated parameters, and then through output membership functions and associated parameters to outputs, to interpret the input/output map.

3.2.1 Model architecture

ANFIS is based on Sugeno fuzzy model (Sugeno and Kang, 1988), To illustrate the architecture of ANFIS (Jan et al, 1997), it is assumed that the fuzzy inference system under consideration has two inputs, denoted by x and y, and one output, denoted by z. The architecture presented here can be extended to ANFIS model with four inputs and one output used in this study. For a first order Sugeno fuzzy model, a set of rules is given as follows:

Rule 1: if x is A_1 and y is B_1, then $f_1 = p_1 x + q_1 y + r_1$

Rule 2: if x is A_2 and y is B_2, then $f_2 = p_2 x + q_2 y + r_2$

Fig.2 illustrates the reasoning mechanism for Sugeno fuzzy model and Fig.3 shows the corresponding equivalent ANFIS architecture. As shown in Fig.2, ANFIS network has five layers. The output of the k-th node in layer k is denoted by $o_{k,i}$. Each layer processes information as follows:

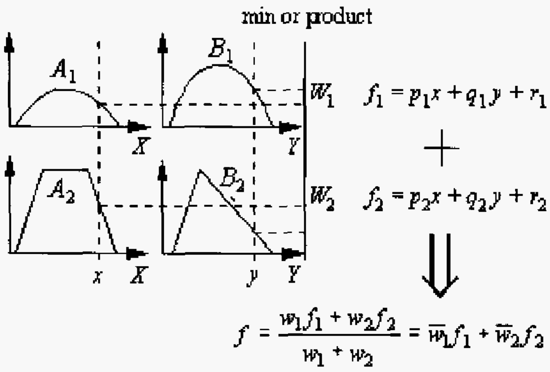

Fig. 2. A Sugeno fuzzy model with two rules

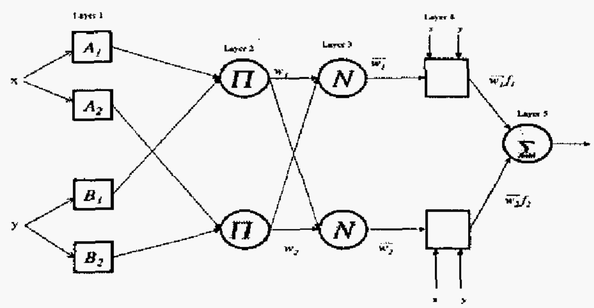

Fig. 3. ANFIS Architecture

Layer 1: Every node in this layer is an adaptive node with node function:

$$O_{1,i} = \mu_{A_i}(x), i = 1,2 \quad \text{or} \quad O_{1,i} = \mu_{B_{i-2}}(y), i = 3,4$$

where x (or y) is the input of node i, A_i (or $B_{i\text{-}2}$) is a linguistic label (such as "small" or "big") associated with node i, $O_{1,i}$ is the membership grade of fuzzy set A_i or $B_{i\text{-}2}$ specifying the degree to which the given input x (or y) satisfies the quantifier A. The membership function for A can be any appropriate parameterized membership function such as the generalized bell function used in this study:

$$\mu_A(\mathrm{x}) = \frac{1}{1 + \left|x - c_i/a_i\right|^{2b}}$$

where $\{a_i, b_i, c_i\}$ is the parameter set. As the values of these parameters change, the bell-shaped function varies accordingly, thus exhibiting various forms of membership functions for fuzzy set. Parameters in this layer are referred to as premise parameters.

Layer 2: Every node in this layer is a fixed node labeled Π, whose output is the product of all the incoming signals:

$$O_{2,i} = \mu_{A_i}(\mathrm{x})\ \mu_{B_i}(\mathrm{y}),\ i = 1, 2.$$

The output of each node represents the active strength of a rule.

Layer 3: Every node in this layer is a fixed node. The i-th node calculates the ratio of the i-th rules' active strength: $O_{3,i} = \overline{w} = w_i/(w_1 + w_2)$, $i = 1, 2$. The outputs of this layer are called normalized active strengths.

Layer 4: Every node in this layer is an adaptive node with node function:

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i)$$

where $\overline{w}_i$ is a normalized active strength from layer 3, and $\{p_1, q_1, r_1\}$ is the parameter set of this node. Parameters in this layer are referred to as consequent parameters.

Layer 5: The single node in this layer is a fixed node labeled Σ, which computes the overall output as the sum of all incoming signals:

$$\text{Overall output} = O_{5,i} = \sum w_i f_i = \frac{\sum_i \overline{w}_i f_i}{\sum_i w_i};$$

3.2.2 Model learning algorithm

The ANFIS model, applied in this study, uses the hybrid learning algorithm (Jang et al, 1997), a combination of least square estimation and back-propagation (the Gradient-Descent Model), for membership function parameter estimation. For the ANFIS architecture in Fig. 3, when the values of premise parameters are fixed, the overall output can be expressed as a linear combination of the consequent parameters. In symbols the output f can be written as:

$$\begin{aligned} f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \\ &= \overline{w}_1 (p_1 x + q_1 y + r_1) + \overline{w}_2 (p_2 x + q_2 y + r_2) \qquad (1) \\ &= (\overline{w}_1 x) p_1 + (\overline{w}_1 y) q_1 + \overline{w}_1 r_1 \\ &+ (\overline{w}_2 x) p_2 + (\overline{w}_2 y) q_2 + \overline{w}_2 r_2 \end{aligned}$$

which is a linear function of parameters p_1, q_1, r_1, p_2, q_2, and r_2.

The Hybrid Learning Algorithm is shown as follows. Assuming that the adaptive network under consideration has only one output represented by

$$O = f(i,S), \tag{2}$$

where i is the vector of input variables, S is the set of parameters, and f is the overall function implemented by the adaptive network. If there exists a function H such that the composite function $H \bullet f$ is linear for some elements of S, then these elements can be identified by the least-squares method. As illustrated before, if the parameter set S can be divided into two sets, that is

$$S = S_1 \oplus S_2 \tag{3}$$

where $\oplus$ represents the direct sum such that $H \bullet f$ is linear for the elements of S_2, then upon applying H to equation. (2), we have

$$H \bullet O = H \bullet f(i,s) \tag{4}$$

which is linear in the elements of S_2. Assuming H is identity, then equation (1) and equation (4) are equivalent. Given values of elements of S_1, we can put P training data in equation (4) and obtain a matrix equation:

$$A\theta = y \tag{5}$$

where θ is an unknown vector, whose elements come from parameters in S_2. This is a standard linear least-squares problem, and the best solution for θ is the least-squares estimator (LSE) θ^*:

$$\theta^* = (A^T A)^{-1} A^T y \tag{6}$$

where A^T is the transpose of A and $(A^T A)^{-1} A^T$ is the pseudo-inverse of A if $A^T A$ is non-singular. Furthermore, we can use the recursive LSE formula to calculate θ^*. If the i-th row vector of matrix A is a_i^T and the i-th element of y is y_i^T, then θ can be calculated iteratively as follows:

$$\begin{aligned} \theta_{i+1} &= \theta_i + p_{i+1} a_{i+1} (y_{i+1}^T - a_{i+1}^T \theta_i) \\ P_{i+1} &= P_i - \frac{P_i a_{i+1} a_{i+1}^T P_i}{1 + a_{i+1}^T P_i a_{i+1}}, i = 0,1,\ldots, P-1, \end{aligned} \tag{7}$$

where LSE θ^* is equal to θ_P. The initial conditions needed to bootstrap equation (7) are $\theta_P = 0$ and $P_0 = \gamma I$, where γ is positive large number and I is the identity matrix of dimension $M \times M$.

The hybrid-learning algorithm combines Gradient-Descent Model and LSE to update the parameters in an adaptive network. For hybrid learning to be applied in a batch mode, each epoch is composed of a forward pass and a backward pass. In the forward pass, after an input vector is presented, we calculated the node outputs in the network layer by layer until a corresponding row in the matrices A and Y in Eq. (5) is obtained. This process is repeated for all the training data pairs to form the complete A and Y. The parameters in S_2 are identified by using Eq. (6) or Eq. (7).after the parameters in S_2 are identified, we compute the error measure for each training data pair. In the backward pass, the error signals propagate from the output end towards the input end. The error gradient is accumulated for each training data entry, and at the end of the backward pass for all training data, the parameters in S_1 are updated by the error gradient. For the given fixed values of the parameters in S_1, the parameters in S_2 thus found are guaranteed to be the global optimum point in the S_2 parameter space because of the squared error measure. Moreover, in the forward pass, node outputs go forward until layer 4 and the consequent parameters are identified by the least-squares method. In the backward pass, the error signals propagate backward and the premise parameters are updated by gradient descent. Table 2 summarizes the actions in each pass.

Table 2 Hybrid learning procedures of ANFIS

	Forward pass	Backward pass
Premise parameters	Fixed	Gradient descent
Consequent parameters	Least-squares estimator	Fixed
Signals	Node outputs	Error signals

4. EXPERIMENT RESULTS

4.1 Actual models for credit evaluation

Two actual models are constructed in this study based on the training samples. One is a 3-layer feed - forward neural network with a hidden layer. As shown in Fig. 4, the structure of NNs is a 4-4-1 network model. Another is a fuzzy neural network with four-input and one output. The structure of the FNN model is shown in Fig. 5. Bell membership functions are pictured in Fig. 6 – Fig. 9.

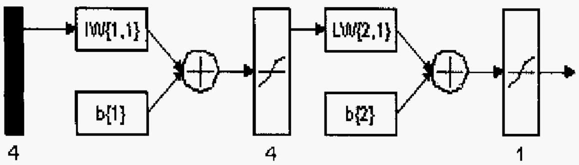

Fig. 4. Actual structure of NNs

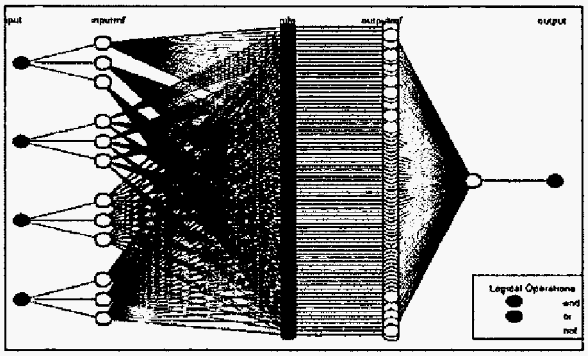

Fig. 5. Actual ANFIS network structure

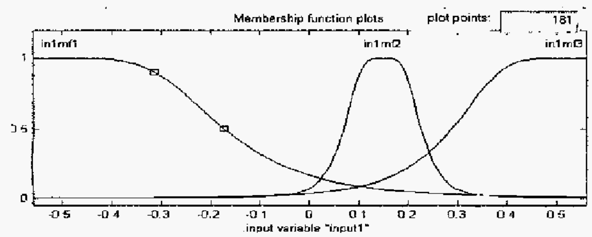

Fig.6. Final membership function for *X*1

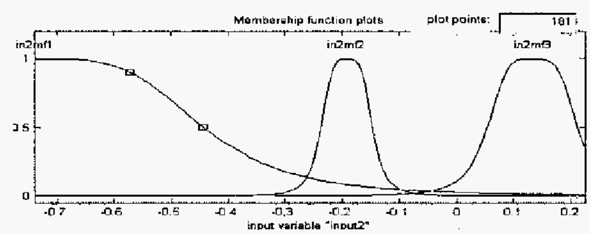

Fig.7. Final membership function for *X*2

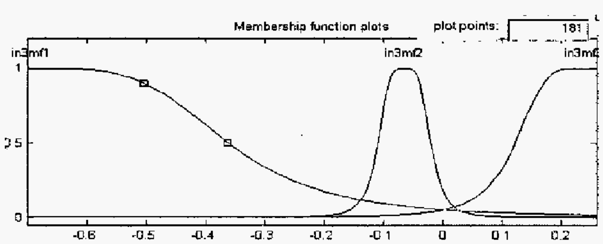

Fig.8. Final membership function for *X*3

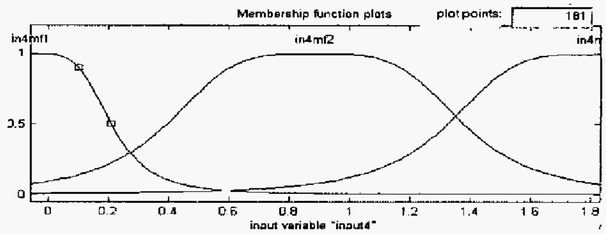

Fig.9. Final membership function for *X*4

4.2. Empirical results

Table 3 summarizes the empirical results of the NNs model and ANFIS model on credit risk evaluation. It is noticed that the performances of both models are fully satisfied with no misclassification for the training data. However, The difference of results given by the test data is quite significant. The total misclassification rate of NNs model is 5.33%, while the total misclassification rate of ANFIS model is 4.00%. Specially, the type 1 error rate of the NNs model is 20%, while the type 1 error rate of ANFIS is only 13.33%. It is clearly shown that the accuracy of credit risk evaluation of ANFIS model is much better than the NNs model. More importantly, the performance of ANFIS model is superior to the NNs model in minimizing type 1 errors and shows considerable promise to screen the corporation's credit risk.

Table 3 Misclassification rate of listed corporations

Samples/Errors		NNs	ANFIS
Training samples (30ST 120non-ST)	Type 1 error	0(0.00%)	0(0.00%)
	Type 2 error	0(0.00%)	0(0.00%)
	Overall	0(0.00%)	0(0.00%)
Test samples (15ST 60non-ST)	Type 1 error	3(20.00%)	2(13.33%)
	Type 2 error	1(1.67%)	1(1.67%)
	Overall	4(5.33%)	3(4.00%)

5. CONCLUDING REMARKS

This study investigated the modeling and performances of neural network and fuzzy neural network on a problem of credit risk evaluation, based on a set of financial data selected from China listed corporations. To compare the two networks, suppose that ST corporations denote the default (insolvency) corporations and non-ST corporations denote non-default (solvency) corporations. The proportion of ST corporations in total sample is designed as 20% based on the actual ratio. The empirical results indicate that the performance of fuzzy neural network is much better than neural network in credit risk evaluation. In particular, fuzzy neural network models involve lower errors in type 1 and have an in-built reasoning mechanism in the form of IF-THEN rules, which makes users understand the results given by the fuzzy neural network model more easily.

REFERENCES

[1] Altman, Edward I., Giancarlo, Marco, Varetto, Franco, Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks, Journal of Banking and Finance 18 (1994), 502-529.

[2] Arnold F. Shapiro, The merging of neural networks, fuzzy logic, and genetic algorithms, Insurance: Mathematics and Economics 31(2002), 115-131.

[3] Eisenbeis R.A., Pitfalls in the application of discriminant analysis in business, finance and economics, Journal of Finance 32(1977), 875-900.

[4] Hornik,K., Approximation capabilities of multilayer feedforward networks, Neural Networks 4(1991), 251-257.

[5] Hornik,K., Stinchcombe, M., White, H., Multilayer feedforward networks are universal approximators, Neural Networks 2(1989), 359-366.

[6] Jain B.A., Nag B.N., Performance evaluation of neural network decision models, Journal of Manage Information Systems 14 (1997), 201-216.

[7] Jang, J-S.R., Sun, C.T., Mizutani, E., Neuro-Fuzzy and Soft Computing, Matlab Curriculum Series, Prentice-Hall, Englewood Cliffs, N.J., 1997.

[8] M.B.Gordy, A comparative anatomy of credit risk models, Journal of Banking and Finance 24(2000), 119-149.

[9] Michel Crouhy, Dan Galai, Robert Mark, A comparative analysis of current credit risk models, Journal of Banking and Finance 24(2000), 59-117.

[10] Patricia Jackson, Wikkiam Perraudin, Regulatory implications of credit risk modeling, Journal of Banking and Finance 24(2000), 1-14.

[11] Rashmi Malhotra, D.K.Ma;hotra, Differentiating between good credits and bad credits using neuro-fuzzysystems, European Journal of Operational Research 136(2002), 190 – 211.

[12] Selwyn Piramuthn, Financial credit=risk evaluation with neural and neurofuzzy systems, European Journal of Operational Research 112(1999), 310 – 321.

[13] Skapura, D., Building Neural Networks, Addison Wesley, New York (1), 1996

[14] Sugeno,M., Kang G.T., Structure identification of fuzzy model, Fuzzy Sets and Systems 28(1988), 329-349.

[15] Tam K.Y., Kiang M, Managerial applications of neural networks: the case of bank failure predictions, Management Sciences 38(1992), 926-947.

[16] Thomas G.Calderon, John J.Cheh, A roadmap for future neural networks research in auditing and risk assessment, International Journal of Accounting Information Systems 3(2002), 203-235.