Can Banks Learn to Be Rational?

Haider A. Khan
University of Denver
Denver
Co. 80208 USA
Tel. 303-871-4461/2324
Fax 303-871-2456
e-mail:hkhan@du.edu
Revised,March 2002
and
Visiting Professor,
CIRJE,
Graduate School of Economics
University of Tokyo,
Tokyo, Japan
Tel.03-5841-5642
Fax 03-5841-8294
Revised,March 2002
April, 2002

I would like to thank Prof. Hori, Prof. Kamoike, Prof.Akita , Prof. Ozaki and the other participants of the Graduate School of Economics seminar at Tohoku University on April 13, 2002. All remaining errors are mine.

Abstract

Can banks learn to be rational in their lending activities? The answer depends on the institutionally bounded constraints to learning. From an evolutionary perspective the functionality (for survival) of “learning to be rational” creates strong incentives for such learning without, however, guaranteeing that each member of the particular economic species actually achieves increased fitness. I investigate this issue for a particular economic species, namely, commrercial banks.

The purpose of this paper is to illustrate the key issues related to learning in an economic model by proposing a new screening model for bank commercial loans that uses the neuro fuzzy technique. The technical modeling aspect is integrally connected in a rigorous way to the key conceptual and theoretical aspects of the capabilities for learning to be rational in a broad but precise sense. This paper also compares the relative predictability of loan default among three methods of prediction---discriminant analysis, logit type regression, and neuro fuzzy--- based on the real data obtained from one of the banks in Taiwan.The neuro fuzzy model, in contrast with the other two, incorporates recursive learning in a real world, imprecise linguistic environment. The empirical results show that in addition to its better screening ability, the neuro fuzzy model is superior in explaining the relationship among the variables as well. With further modifications,this model could be used by bank regulatory agencies for loan examination and by bank loan officers for loan review. The main theoretical conclusion to draw from this demonstration is that non-linear learning in a vague semantic world is both possible and useful. Therefore the search for alternatives to the full neoclassical rationality and its equivalent under uncertainty---rational expectations--- is a plausible and desirable search, especially when the probability for convergence to a rational expectations equilibrium is low.

Keywords:rationality, bounded rationality, recursive learning,screening model, discriminant analysis, logistic regression, neuro fuzzy

1. Introduction

Can a bank learn to be more rational than it actually is, at any particular time in its historical trajectory? In this paper, the question is answered affirmatively in the concrete context of banks managing their default risks.[1] The problem in this specific context is how to effectively manage the default risk of commercial loans. This problem has been one of the main questions posed in the accounting and finance literature for years. The question is significant for both academic and practical purposes. Looking at practical aspects, the evaluation process for commercial loan can be divided roughly into two stages. The first is the screening part before the loan is approved. This can be called 'the credit scoring model'. Each applicant is assigned a credit score after being evaluated according to some prespecified criteria. Whether a case is accepted or not is based on the score received by applying such criteria.. The second stage is the continuous monitoring after the loan has been approved. This can be called 'the bankruptcy prediction problem'. After the commercial loan has been approved, one of the most important questions for the bank to answer is whether the debtor company will go bankrupt or not. Therefore, a warning system to predict the chances of bankruptcy is needed. These two successive stages raise prediction questions that can be viewed as problems of dichotomous choices: either accepting or rejecting a loan application;and if accepted then further continuing or stopping the commercial loan. This paper focuses on the screening of the new commercial loan applicants as a specific context for learning to be (more) rational. This naturally leads to a focus on the problem of the predictability of the default loan before the loan is approved.

In section 2, I review the past literature related to the credit scoring model in the context of recursive learning in an imprecise environment. Section 3 describes the existing screening process in Taiwan's commercial banks, and shows how a neuro fuzzy model can be constructed for enhancing learning to be (relatively more) rational. In terms of the major theoretical question,(namely, can banks learn to be rational?), the answer lies precisely in the ability to construct an appropriate knowledge base, and learning in a recursive way to predict the default loans better than other competing models of predicting default loans . The relative success in constructing

---

[1] The empirical illustration presented later in the text draws heavily upon my joint work-in-progress with C.-S. Lin without implicating him in any of the substantive, epistemological or ontological interpretations that I have presented in this paper.

and using the knowledge base in an imprecise, fuzzy world shows that our economic rationality is both bounded and capable of expansion. This is consistent with a realist epistemology, and lends support for economic learning under a realist ontology[2] of firms and their environments The empirical results are shown in section 4. Finally, the paper ends with a summary and the implications of the research efforts in section 5.

2.Some Received Views: a critical review

The past literature related to commercial loans can be classified according to the tools of analysis that are used. Some researchers use discriminant functions and logit type regression to construct the predictive model by using the financial ratios (Zavgren 1985; Blum 1974; Collins and Green 1982; Dietrich and Kaplan 1982; West 1985; Srinivasan and Kim 1987). The problem is that the relationship among the variables can be more complicated than just the postulated linear relationship. Some researchers have proposed the expert system in order to construct the predictive model (Chorafas1987; Duchessi, Shawky, and Seagle 1988; Romaniukk and Hall 1992; Yang et al. 2001). However, the knowledge base of the expert system is hard to derive. Some researchers have used neural networks to model the bankruptcy prediction problem(Quinlan 1993; Altman, Marco, and Varetto 1994; Boritz and Kennedy 1995; Boritz, Kennedy, and Albuquerque 1995; Atiya 2001; Coats and Fant 1993; Lenard, Alam, and Madey 1995; Lacher et al, 1995; Sharda and Wilson 1996; Tam and Kiang 1992; Wilson and Sharda 1994; Yang 1999). While empirical studies show that neural networks produce better results for many classification or prediction problems, they are not always uniformly superior (Quinlan 1993; Altman, Marco, and Varetto 1994; Boritz and Kennedy 1995; Boritz, Kennedy, and Albuquerque 1995). Besides, the mapping process is too complicated to explain the relationships among the variables. It could only be seen as a black box. The learning processes need to be specified better in line with the recent advances in cognitive psychology, artificial intelligence and related field.[3]

---

[2] This is indeed crucial, and sets the present work apart from both the ‘as if…’ variety without ontological commitments and the ‘satisficing’ ‘school which is essentially empiricist in a phenomenological way. Usually, the theorists in both schools have not confronted the difficult philosophical issues directly.

[3] This is in fact the ultimate aim of the project. In this paper the “black box” problem is not solved except to show that neural networks with hidden layers can learn satisfactorily under some circumstances. The moving euilibria may not necessarily

Furthermore, a sharp classification or unnatural approximating values during the evaluating process may result in unreasonable or incorrect outcomes. In other words, replacing an inherently fuzzy classification and measurement system with a non-fuzzy one actually leads to a loss of predictive precision. Given the inherent fuzziness in credit rating for commercial loans, the fuzzy approach is, therefore, a more reliable technique than the apparently more precise non-fuzzy ones. This line of thinking has led some researchers to develop a more reasonable credit-rating procedure by using fuzzy techniques. Zimmermann and Zysno (1983) used fuzzy operators to aggregate evaluation results from a four-level hierarchy of criteria. Levy et al. (1991) developed a computer based system to evaluate a company's financial position based on fuzzy logic in determining whether to grant or deny the loan application. Besides, several studies employ the fuzzy integral (Sugeno, Nishiwaki, Kawai, and Harima 1986; Tahani and Keller 1990; De Neyer, Gorez, and Barreto 1993; Leszczynski, Penczek, and Grochulskki 1985; Chen, and Chiou, 1999) as a tool of information fusion to aggregate the credit information of loan applicants.

In addition to the literature related to the different techniques used, Edmister (1988) argued that numerical financial ratios and human credit analysis could be combined to produce more accurate evaluation results. Neglecting the information provided by these qualitative factors may result in undesirable consequences. Marais, Patell, and Wolfson (1984) suggest that market information such as commercial paper ratings or stock price variability can be an effective substitute for extensive financial statement analysis.

In sum, it is possible to draw the following critical conclusions based on the above literature review.

1. The discriminant function and the logistic regression mainly deal with the linear relationships among the independent and dependent variables. If the true relationships among the variables are nonlinear, then these two methods are not appropriate.
2. The discriminant function and the logit type regression ignore the interaction between the variables in general. Therefore, a more detailed modeling of the relationship---particularly, the reasoning relationships--- among the variables cannot be obtained through these two traditional statistical tools.
3. The expert system is a good approach to construct a warning system. However, it is really hard to get the correct "knowledge base", and decide the

---

converge to the rational expectations equilibrium in the "stationarystate", however.

relative importance of each rule. Even the expert cannot tell the relative importance of each rule.

4. Neural network is a good tool to get the mapping function between the independent and dependent variables. However, it cannot explain the causal relationship among the variables by itself. Further causal specification and testing are necessary conditions for a deeper non-Humean causal analysis.
5. In addition to the financial ratios, some qualitative variables such as the general management and some other perspectives and characteristics of the company are also helpful in evaluating the case. A tool capable of dealing with the qualitative variables and their interrelations is needed.

In order to solve the above problems, I propose the use of neuro fuzzy technique combined with a fuzzy set theoretic approach. At the same time, I offer a somewhat novel interpretation of learning(in neuro fuzzy setting) to be (more) rational in a world where rationality is bounded, but can also be improved through learning. The qualitative variables in the real world can be dealt with through the membership function of the fuzzy logic. The functionality of fuzzy logic can be used to describe the vague[4] ordinary language definitions and relationships among the variables. The learning ability of neural network can be used to adjust the relative importance of each decision rule. Finally, the knowledge base obtained from this technique can be used as a diagnostic system to see the heuristic reasoning process behind the screening result. This advances the project of understanding how banks can be conceptualized as members of an economic species in a competitive market setting with capacity to learn--- but not all banks learn at the same time or at an equal rate.

The concrete purpose of this paper is to propose a screening model for commercial loans as an illustrative example of the more ‘general learning to be rational’ class of models and theories. The model presented here can not only predict the default loan successfully, but it can also explain at least partially how the decision is made. The model can be used by bank regulatory agencies for loan examination and by bank loan officers for loan review after some practical modifications.

[4] The classic discussion of vagueness beyond Wittgenstein’s idea of “ family resemblances” is Max Black(1937). See also, Birkhoff and von Neumann(1936). For more recent discussions, see Rescher(1969) and Kosko(1992).

3. Methodology

3.1. The credit-rating system.

Bank loan officers need to evaluate the loan risk before approving a particular loan. Presumably, this evaluation should be based on some relevant criteria. In Taiwan, the evaluation process is based on the " Credit-Rating Table for Commercial Loan" made by Taiwan Bank in 1987. The evaluation items consist of three main categories: financial conditions, general management, and characters and perspective. Basically, the criteria listed under "financial conditions" can be measured quantitatively. However, the other two categories are evaluated according to the loan officers' subjective judgements.

Each category has several indicators. Each indicator contains some evaluation criteria with various points based on the different satisfaction levels of these criteria. The basic structure of the evaluation variables is listed in table 1. The scores of quick ratio (FC11) and current ratio (FC12) add up to liquidity ratios (FC1). Similarly debt ratio (FC21) and long-term asset efficiency ratio (FC22) add up to financial structure ratios (FC2). Finally, liquidity ratios (FC1), financial structure ratios (FC2), profitability ratios (FC3), and efficiency ratios (FC4) add up to financial conditions (FC).[5] For each company, the total score for financial condition (FC), general management (GM), and characters and perspectives (CP) can be obtained by simply adding up all the scores.

Loan officers perform a credit-rating process via quantitative methods to examine a company's financial position based on the previously determined evaluation criteria in assessing a company's credit level on the whole. The problem is how to make the decision based on these three scores more scientific and effective. A neuro fuzzy(NF) approach is proposed in order to model the decision process and the NF approach is compared with discriminant analysis and logit type regression. In the next section, a fuzzy logic system will be introduced followed by the neuro fuzzy model in the section after that.

3.2 Fuzzy Logic System

Fuzzy logic mainly deals with the extent to which an object belongs to a (fuzzy) set. Usually the functional $\mu_A(x)$ is used to denote the extent to which object x belongs

---

[5] For the detailed formulas for each item please refer to Chen and Chiou (1999).

to fuzzy set A. A fuzzy logic system is constructed syntactically by introducing the logical relation of implication or the ”IF-THEN” rules to describe the relationship among independent and dependent variables.There are mainly two families of logical inferences covered under the names *modus ponens* and *modus tollens* in classical logic. In terms of modern logic the two forms can be described as follows:

If p and q are two *well-formed formulas(wff)* connected by the logical connective ‘if…then’, as ‘if p then q’, then *modus ponens* is simply the form of argument: p, therefore q. *Modus tollens* is: not q, therefore not p. In fuzzy logic, p and q can refer to ‘vague’ linguistic terms in a precise, *possibilistic* way.[6] The only difference between fuzzy logic and traditional expert system is that the variables used in fuzzy logic are linguistic terms rather than numeric values as in the traditional expert system. Let FC, GM, CP and SCORE denote the financial conditions, general management,

Table 1 Explanatory Variables

Financial conditions (FC):
- Liquidity ratios (FC1):
  - quick ratio (FC11)
  - current ratio (FC12)
- Financial structure ratios (FC2)
  - Debt ratio (FC21)
  - long-term asset efficiency ratio (FC22)
- Profitability ratios (FC3)
  - net sales ratio (FC31)
  - profit margein before tax (FC32)
  - return on net worth before tax (FC33)
- Efficiency ratios (FC4)
  - inventory turnover (FC41)
  - receivables turnover (FC42)
  - total assets turnover (FC43)

General management (GM)

[6] However, there are still some unsolved logical problems---the most important being the Duhem-Quine problem of joint hypotheses. Roughly, neither inference form can work because both p and q(but particularly p) are very rarely, if ever, single hypotheses.

Administrator’s personal credit (GM1)
Administrator’s management experiences (GM2)
Stockholders’ structure type (GM3)
Average sale growth rate during the last three years (GM4)
Conditions of capital increment during the last three years (GM5)
Outstanding check records in banks (GM6)

Characters and Perspectives (CP)
Equipment and technologies (CP1)
Product marketability (CP2)
Collateral (CP3)
Economic conditions of the industry in the next year (CP4)

---

character and perspectives, and the credit score of an applicant. A typical rule in a traditional expert system, for example, is stated as follows:

If FC > 30, GM >20, and CP >16, then SCORE is 10. (1)

A fuzzy logic rule is stated instead as follows:

If FC is **high**, GM is **low**, and CP is **high**, then SCORE is **medium.** (2)

where FC, GM, CP, and SCORE are called *linguistic variables* and high, medium, and low are called *linguistic terms.* Basically there are three main steps in building a fuzzy logic system: fuzzification, construction of knowledge base, and defuzzification.

3.2.1 Fuzzification

Fuzzy logic uses linguistic terms to describe the characteristics of an object. For example, we use low, medium, and high to describe the extent of financial condition (FC), general management (GM), and character and perspective (CP) of an applicant. Each linguistic term is defined by a membership function. Figures 1a, 1b, 1c, and 1d are the membership functions for FC, GM, CP, and SCORE respectively. If the measurements of an applicant are {FC, GM, CP}={30, 20, 16}, for example, then the corresponding values of each term can be seen from figure 1a, 1b, and 1c as follows.

Figure 1 a. Membership function for linguistic variable FC

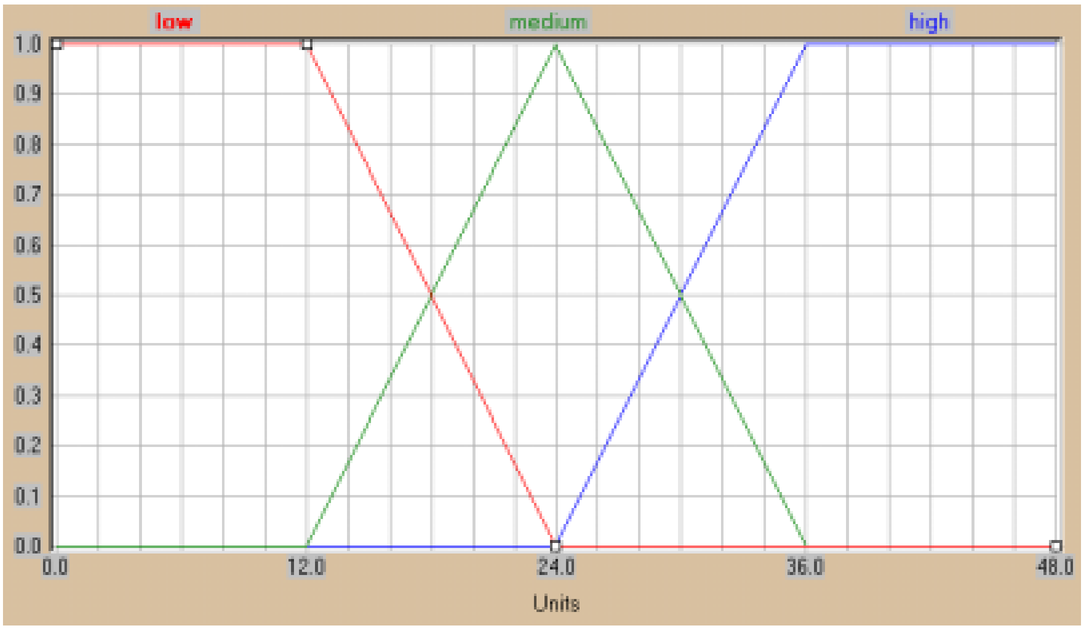


Figure 1 b. Membership function for linguistic variable GM

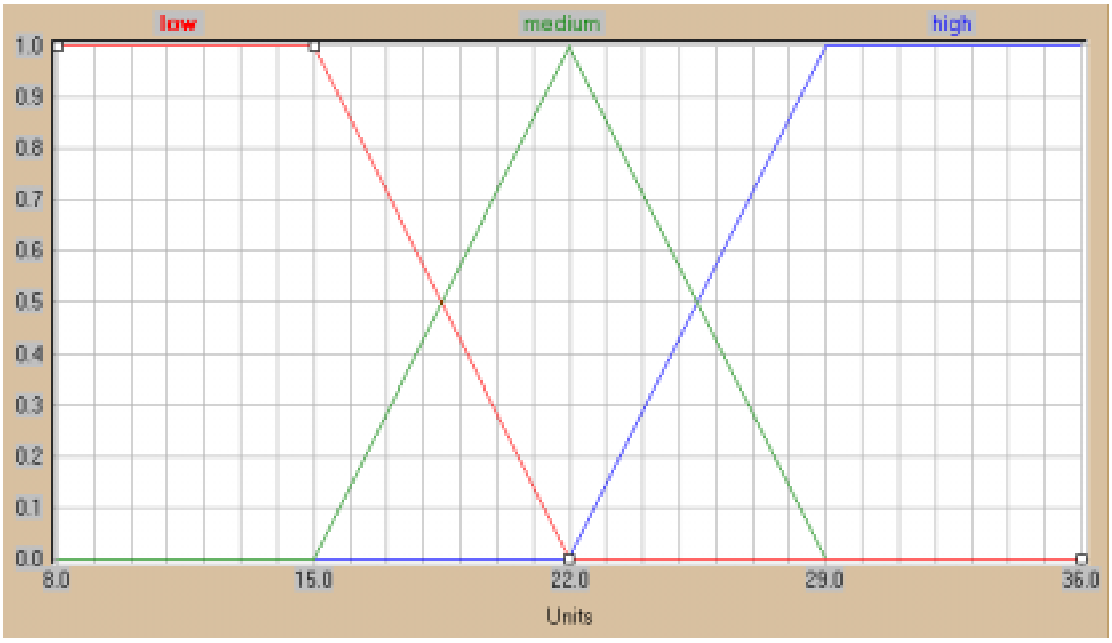


Figure 1 c. Membership function for linguistic variable CP

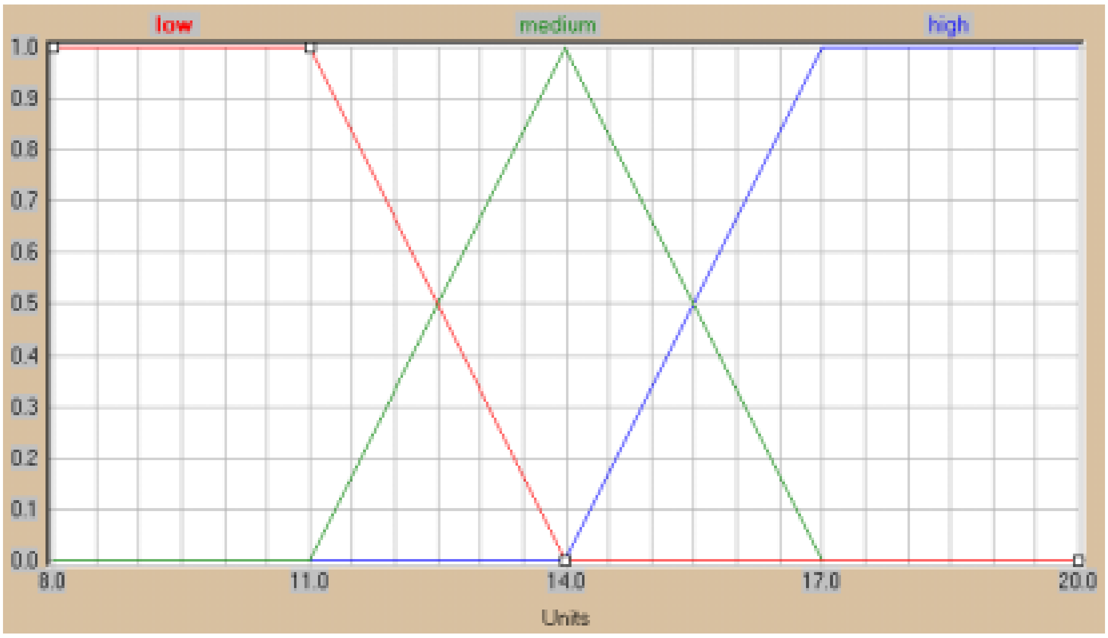
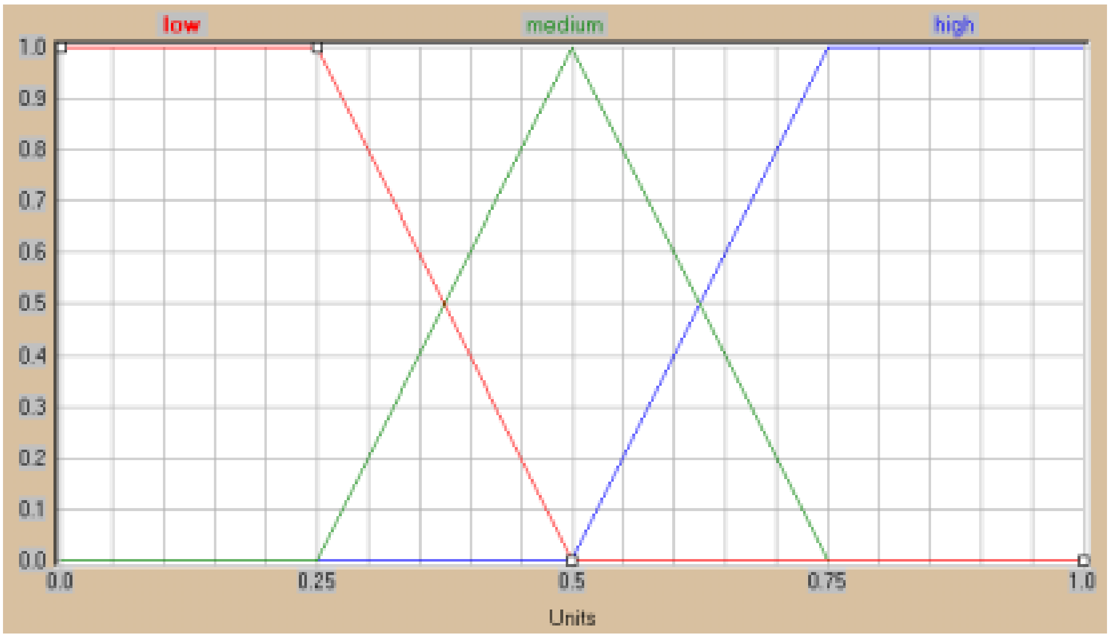


Figure 1 d. Membership function for linguistic variable credit score



FC：μ_{low} （30） = 0, μ_{medium} （30） =0.5, μ_{high} （30） = 0.5

GM: μ_{low} （20） = 0.28, μ_{medium} （20） =0.72, μ_{high} （20） = 0

CP：μ_{low} （16） = 0, μ_{medium} （16） = 0.32, μ_{high} （16） = 0.68

In other words, the corresponding values can be written as follows.
FC： { low, medium, high} = {0.00, 0.50, 0.50}.
GM： { low, medium, high} = { 0.28, 0.72, 0.00}.
CP： { low, medium, high} = {0.00, 0.32, 0.68}.

An applicant with FC equal to 30 has membership function values for low, medium, and high equal to 0.00, 0.50, and 0.50 respectively. Since each linguistic variable after mapping can have different membership function values for different linguistic term, it breaks the traditional binary logic that a case can only belong to or not belong to a category. This process is what we call *fuzzification*. The most commonly used membership functions are linear and spline functions[7], Table 2 lists all the linguistic variables, their linguistic terms, and variable types used in this paper.

Table 2. Linguistic variables, linguistic terms, and variable types

Linguistic variable	Linguistic terms	Type
FC	Low、medium、high	Input
GM	Low、medium、high	Input
CP	Low、medium、high	Input
SCORE	Low、medium、high	output

3.2.2 Towards a model of bounded, but expanding rationality: the construction of knowledge base:

Knowledge base is constructed by the “IF-THEN” rules. Each rule has two parts, “IF” and “THEN” parts. “IF” part measures the extent to which the object satisfies the logical antecedent condition, “THEN” part is the response of the system. Of course, in the implication relation of any symbolic or mathematical logical system, say that of a first order predicate calculus, the then part is the “consequent”. It is supposed to logically follow from the sufficient logical conditions subsumed under the antecedent. Empirically, and in fuzzy logic the validity degree of the response depends on the satisfaction extent of the “IF” part.Thus in terms of symbolic logic, under fuzzification the sharp or strict sufficiency of the antecednt in a well formed formula(wff) is lost, but partial consequences are still deducible through the validity

[7]please refer to Zimmermann(1991) for details.

degree relation. Take equation (2) for example,

IF FC is high, GM is low, and CP is high, then SCORE is medium, (2)

the validity degree of the then part depends on the *minimum* extent of each linguistic term in the if part according to the definition of Zimmerman and Thole (1978). In other words, $\mu_{A\cap B} = \min\{\mu_A;\mu_B\}$. The satisfaction extent of the "IF" part of the above rule is the minimum of the validity of "FC is high", "GM is low", and "CP is high". That is the validity extent of "IF" part is equal to min{0.50, 0.28, 0.68}=0.28, which is the validity extent of the response. In other words, the response of this system is "the SCORE is medium" with validity extent equal to 0.28.。

3.2.3 Defuzzification

After fuzzification and inference procedure, each applicant will have a corresponding value for each linguistic term of the output variable. For example the corresponding value of the linguistic term "the SCORE is medium" is 0.28 for equation (2) for the above example. Assume that the corresponding values for the other linguistic terms are "SCORE is low" is 0.1, and "SCORE is high" is 0.2. The procedure to transform these linguistic values into the numeric output is called defuzzification. Basically it consists of two main steps. The first step is to find out the representative value for each linguistic term. Usually it is the value with membership function equal to 1. The second step is to summarize these linguistic outputs. For the second step, we do the weighted sum of the representative values and its corresponding extent values. For example, assume the representative values for each linguistic term of the output variable is {0.25,0.5,0.75} as depicted in Figure 1d, then with the corresponding validity extent of each linguistic term {0.10,0.28,0.20}, the final output value is equal to 0.10*(0.25) + 0.28*(0.5) + 0.20*(0.75) = 0.315. In other words, the final credit score for this case is 0.315.

This defuzzification method is called the method of gravity, which is one of the most commonly used defuzzification methods.[8] .

Although fuzzy logic has been applied to many fields successfully, there still exists two main shortcomings associated with this method. The first is how to decide the membership function for each linguistic term. The second is how to decide the relative importance of each rule. Therefore, some effective approach is needed to improve this method. One of the possible ways is to use the learning ability of neural network to do the modification of the membership function and the

---

[8] Please refer to Tong and Bonissone (1984) and Zimmermann (1987) for the other defuzzification methods. And please refer to Klir and Yuan (1995) for the detailed discussion of fuzzy logic

relative importance of each rule. The technique to combine the knowledge base of fuzzy logic and the learning ability of neural network is called the neuro fuzzy technique.

3.3 Neuro Fuzzy Technique

Basically neuro fuzzy technique(from here on called simply neuro fuzzy) is a fuzzy logic system with the learning ability of neural network to modify its parameters, including the parameters of the membership function and the relative importance of each rule. There are different ways to combine these two techniques (Buckley and Hayashi, 1994；Nauck and Kruse, 1997；Lin and Lee, 1996). These methods turn out to be not so different from one another in practice. This paper adopts the FAM (fuzzy associative memory ; FAM) proposed by Kosko (1992). Each rule is viewed as a neuron, the weight of each rule is represented as the weight of each edge in the neural network. For each data point there is a predicted value generated by the system associated with a realized value. The training process will stop until the error between the predicted value and realized value is less than a certain threshold value. The general neural network model is:

The general neural network model used is:

$$\text{Output}\, Trend_{t+1} = F_2(w_2 F_1(w_1 x))$$

where F_1 and F_2 are the transfer functions for hidden node and output node, respectively. The most popular choice for F_1 and F_2 are the sigmoid function, $F(x) = \frac{1}{1+e^{-\alpha x}}$, representing the activation function adopted in the calculation process, w_1 and w_2 are the matrices of linking weights from input to hidden layer and from hidden to output layer, respectively, x is the vector of input variables.

3.4 Research Model

Based on the scores derived from the table made by Taiwan Bank, the research model of neuro fuzzy is depicted as Figure 2. In addition to the FC, the qualitative variables, GM and CP, are included in the evaluation process. We use the fuzzy logic to construct the knowledge base to describe the relationship among the independent and dependent variables. The crucial step here is to fine tune the knowledge base through the learning ability of neural network based on the training data set. Finally we use the testing data set to validate the obtained model.

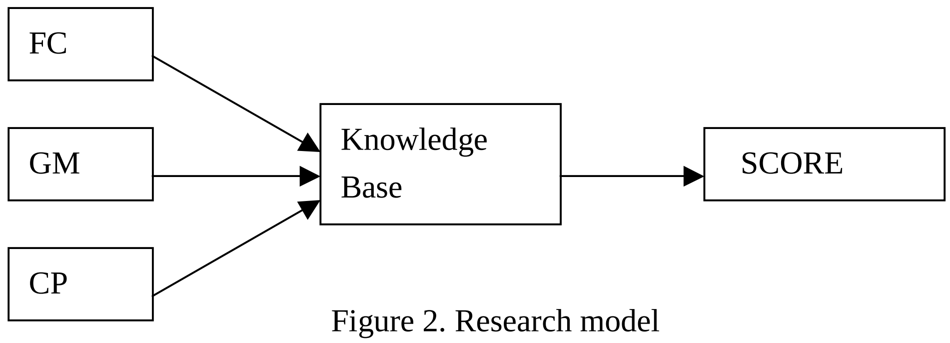


Figure 2. Research model

Since discriminant function and logistic analysis are now well documented, for example, see Pindyck and Rubinfeld (1998) and Sharma (1996), a detailed description will not be given here.

4. Empirical Results

4.1 Data Sample

This data set comes from one of the commercial banks in Taiwan. Among all the borrowers who still were in contract with this bank on September 30, 2000, we select out all the borrowers who have incurred default before. “Default” is defined as those borrowers who failed to pay the interest within one week after the due date. There are a total number of 76 cases. In addition, we randomly choose other 195 normal counterparts.

The data set is divided into two parts, training data set for model construction and testing data set for model validity testing. Three different combinations of normal cases and the default cases are formed for the training data set. Table 3 lists the composition of each sample for training data set and testing data set. For sample 1, the proportion of normal cases to default cases is 1:1. There are 50 normal cases and default cases for training data set, and 145 normal cases and 26 default cases for testing data set for sample 1. Sample 3 has 150 normal cases and 50 default cases for training data set and 45 normal cases and 26 default cases for testing data set.

Table 3. The composition of different testing samples

	Training data set		Testing data set		Total
Sample	No. of normal cases	No. of default cases	No. of normal cases	No. of default cases	No. of cases
Sample 1 (1:1)	50	50	145	26	271
Sample 2 (2:1)	100	50	95	26	271
Sample 3 (3:1)	150	50	45	26	271

4.2 Empirical Results

Instead of setting the credit score as the dependent variable, we set the probability of default as the dependent variable. The value of the dependent variable is set to 1 for default cases and 0 for normal cases. The predicted value represents the probability that the case will be a default case. The higher the value predicted, the more probable it is that the event loan default will happen. In this paper, three different criteria are tested to decide the threshold values to determine whether a case is normal or not. These are: maximizing the total classification accuracy, maximizing the testing power, and minimizing the misclassification cost. In the following, MDA, Logit, and NF are used to denote the multivariate discriminant analysis, logit type regression, and neuro fuzzy respectively.

4.2.1 Maximizing total classification accuracy

The model maximizing the total classification accuracy for training data set is used to test the testing data set. The simulation results for training data set and testing data set are shown in table 4 and 5 respectively. The values in table 4 and 5 represent the total classification accuracy of each method for different samples.

Table 4. Total classification rate for training set

	Sample 1	Sample 2	Sample 3
MDA	64%	62%	61%
Logit	63%	67%	75%
NF	99%	72%	78%

Table 5. Total classification rate for for testing set

	Sample 1	Sample 2	Sample 3
MDA	64%	60%	62%
Logit	64%	78%	61%
NF	70%	79%	65%

It can be seen that NF can reach 99% precision for the training data set for sample 1. Although NF can reach only 70% for sample 1 of testing data, it still is the best one among these three methods.

4.2.2 Maximizing the testing power

Maximizing the testing power means to detect the default cases as best as possible. Usually it goes with increasing the probability of misclassifying the normal cases into the default cases while decreasing the threshold value to increase the testing power. Table 6 and 7 shows the simulation results based on the criterion to maximize the testing power. The values in table 6 and 7 represent the testing power of each method for different samples. It can be seen that NF can always detect the default cases for training data set as table 6 depicted. For testing data set, NF can also detect all the default cases for sample 1 and sample 2 as shown in table 7. Still NF is the best one among these three methods in terms of the testing power.

Table 6. Testing power for training set

	Sample 1	Sample 2	Sample 3
MDA	56%	58%	62%
Logit	98%	78%	28%
NF	100%	100%	100%

Table 7. Testing power for testing set

	Sample 1	Sample 2	Sample 3
MDA	50%	58%	58%
Logit	100%	73%	19%
NF	100%	100%	85%

4.2.3 Minimizing the misclassification cost

Since the misclassification cost is different for classifying a normal case into a default case and classifying a default case into a normal case, it is not appropriate to view them as equal. The criteria to maximize the total classification accuracy or maximizing the testing power is not quite appropriate from the practical point of view. What the loan manager cares about is to find out the decision with the least cost. Assume that the misclassification cost for classifying the normal case into default one be A, and the misclassification cost for classifying the default case into a normal one be 30A. The simulation results are listed in table 8 and 9 respectively for training data set and testing data set. The values represent the cost associated with each method for different samples. It can be seen that NF can find the decision with least cost among these three methods either for training data set or for testing data set.

Table 8. Total costs for training set

Ratio	Sample 1	Sample 2	Sample 3
MDA	740A	400A	686A
Logit	82A	414A	1217A
NF	38A	69A	112A

Table 9. Total costs for testing set

	Sample 1	Sample 2	Sample 3
MDA	477A	400A	379A
Logit	141A	278A	702A
NF	113A	198A	103A

4.3 Sensitivity Analysis

Since the ratio of the misclassification cost is difficult to determine from the practical point of view, 8 different ratios are tested. Let the misclassification cost for classifying the default case into a normal one be 5A, 10 A, 15A, 20A, 30A, 40A, and 50A. The simulation results are listed in Table 11, 12, and 13 for training data set. Table 14, 15, and 16 list the simulation results for testing data set.

Table 11 The sensitivity analysis results for training data set with sample 1

	1：1	1：5	1：10	1：15	1：20	1：30	1：40	1：50
MDA	36A	124A	234A	344A	454A	674A	894A	1114A
Logit	37A	54A	59A	64A	69A	79A	89A	99A
NF	28A	38A	38A	38A	38A	38A	38A	38A

Table 12 The sensitivity analysis results for training data set with sample 2

	1：1	1：5	1：10	1：15	1：20	1：30	1：40	1：50
MDA	57A	141A	246A	351A	456A	666A	876A	1086A
Logit	49A	106A	161A	216A	271A	381A	491A	601A
NF	42A	69A	69A	69A	69A	69A	69A	69A

Table 13 The sensitivity analysis results for training data set with sample 3

	1：1	1：5	1：10	1：15	1：20	1：30	1：40	1：50
MDA	78A	154A	249A	344A	439A	629A	819A	1009A
Logit	50A	209A	389A	569A	749A	1109A	1469A	1829A
NF	45A	112A	112A	112A	112A	112A	112A	112A

Table 14 The sensitivity analysis results for testing data set with sample 1

	1：1	1：5	1：10	1：15	1：20	1：30	1：40	1：50
MDA	61A	113A	178A	243A	308A	438A	568A	698A
Logit	63A	141A	141A	141A	141A	141A	141A	141A
NF	51A	113A	113A	113A	113A	113A	113A	113A

Table 15 The sensitivity analysis results for testing data set with sample 2

	1：1	1：5	1：10	1：15	1：20	1：30	1：40	1：50
MDA	48A	92A	147A	202A	257A	367A	477A	587A
Logit	27A	82A	117A	152A	187A	257A	327A	397A
NF	25A	86A	106A	126A	146A	186A	226A	266A

Table 16 The sensitivity analysis results for testing data set with sample 3

	1：1	1：5	1：10	1：15	1：20	1：30	1：40	1：50
MDA	48A	71A	126A	181A	236A	346A	456A	566A
Logit	28A	114A	219A	324A	429A	639A	849A	1059A
NF	25A	47A	57A	67A	77A	97A	117A	137A

It can be seen that the difference among these three methods can become more obvious as the cost proportion increases. NF is the best one among these three methods except for the testing data with cost proportion 1:5. On the other hand, the performance of MDA and Logit is inconclusive.

4.4 Discussion

The Logit functions obtained from three different training data set are as follows.

Sample 1:

Ln (p/1-p) = 3.0793 + 0.0071 FC – 0.0539 GM – 0.1497 CP (3)

s	1.6689	0.0229	0.0426	0.0920
p-value	0.0650	0.1036	0.2061	0.7565

Sample 2

Ln (p/1-p) = 2.1980 + 0.0116 FC – 0.0436 GM – 0.1599 CP (4)

s	1.3733	0.0205	0.0378	0.0725
p-value	0.1095	0.5699	0.2483	0.0275

Sample 3:

Ln (p/1-p) = 1.3861 + 0.0141 FC – 0.0336 GM – 0.1506 CP (5)

s	1.2321	0.0190	0.0346	0.0672
p-value	0.2606	0.4582	0.3316	0.0251

An applicant is classified as being in default if the probability is greater than the threshold value. Equation 3, 4, and 5 indicate that the higher the value of GM and CP, the less the probability for the applicant being in default. It is the in the same direction as what we expected. However, the FC factor has sign that is different from what we expected. The reason can be the window dressing of the financial ratios, which makes this ratio no more effective in predicting the default case.

Fisher's linear classification functions obtained from these three different training data sets are obtained as follows.

Sample 1:

SCORE = -32.650 + 0.217*FC+0.847*GM+2.817*CP (for normal case) (6)

SCORE =-29.646+0.224*FC+0.794*GM+2.671 *CP (for default case) (7)

Sample 2:

SCORE = -32.623+0.228*FC+0.964*GM+2.604*CP (for normal case) (8)

SCORE =-29.687+0.239*FC+0.921*GM+2.443*CP (for default case) (9)

Sample 3:

SCORE = -30.469 +0.212*FC + 0.853*GM +2.521*CP (for normal case) (10)

SCORE = -27.923+0.225*FC+0.819*GM+2.368*CP (for default case) (11)

A case is classified as default if the value calculated from the default equation is greater than the value calculated from the normal equation. It can be seen from these equations that the effect of variable FC is the least among these three variables in determining the credit score. And CP has the most influence. In other words, FC is the least significant among these three variables in predicting default.

Finally, we can take a look at the knowledge base obtained from the neuro fuzzy technique. Table 10 lists the rules with relative importance (Dos) greater than 0.9.

Table 10. Knowledge base with relative importance greater than 0.9

	IF				THEN
No.	CP	FC	GM	DoS	SCORE
1	low	low	medium	1	High
2	medium	medium	medium	1	Medium
3	medium	high	high	1	Medium
4	high	high	high	0.91	Low

It can be seen that these rules do make sense. For rule 1, IF CP is low, FC is low, and GM is medium, THEN SCORE is high, the reasoning process follows what we expect. Similarly, for rule 4, IF CP is high, FC is high, and GM is high, THEN SCORE is low, it is in accordance with our intuition. As for the other situations, the rules can show the interactions among the variables.

Basically, MDA and Logit should perform equally well if the independent variables have multi-normal distribution/density functions and the relationships among the variables are linear. The difference between the Logit and NF implies the existence of the nonlinear relationship among the variables. The empirical results show that the NF can really provide the loan manager with the “true” explanations behind the screening result in addition to the good prediction result in screening the applicants. Three different criteria to decide the threshold values show the robustness of the proposed NF model. Although the different compositions of the training data set will lead to different performance, however, NF still performs the best for different data samples.

5. Conclusions

This paper has attempted to advance towards a new conception of rationality by departing from the assumption of neoclassical maximizing behavior under the usual preference relations. Concretely, I have proposed a screening model for the commercial loan credit rating based on the variables derived from the credit-rating table by using the neuro fuzzy technique, which combines the functionality of the fuzzy logic and the learning ability of neural network. The empirical results show that in addition to the better prediction results than can be obtained from discriminant analysis and logit type regressions, the proposed model can also show the complex relationship among the variables through the knowledge base. This model can also be applied to the problem of bankruptcy prediction after some modifications. Ultimately, the hope is to create a more general model of learning embedded in the neuro fuzzy structure that will contain the present model as a special case. Although that goal is still elusive, enough progress has been made in the present context to demonstrate that it is rational to pursue the task of building a general 'learning to be rational model' in a world where rationality is bounded but amenable to gradual improvement under favorable circumstances.

References

1. Altman, E., Marco, G., and Varetto, F., "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks," J. Banking and Finance, 18, 505-529, 1994.
2. Atiya, A. F., "Bankruptcy Prediction for Credit Risk using neural networks: A survey and new results", IEEE Transactions on neural networks, 12,4, 929-935,2001.
3. Birkhoff, G. and von Neumann, J., "The logic of quantum mechanics", Annals of Mathematics, vol. 37, no.4, 823-843,1936.
4. Black, M., "Vagueness: an exercise in logical analysis", Philosophy of Science, Vol. 4, 427-455, 1937.
5. Blum M. "Failing company discriminate analysis", Journal of Accounting Research 1974, 12 , p. 1-25.
6. Boritz, J., and Kennedy, D., "Effectiveness of neural network types for prediction

of business failure," Expert Syst. Appl., 9, 504-512, 1995.

7. Boritz, J., Kennedy, D., and Albuquerque, A.,"Predicting corporate failure using a neural network approach," Intelligent Syst. In Accounting, Finance, and Manag., 4, 95-111, 1995.
8. Buckley, J. J. and Hayashi, Y. (1994) "Fuzzy Neural Networks: A Survey", Fuzzy Sets and Systems, 66, 1-13.
9. Chen, L.H. and Chiou, T.W., "A fuzzy credit-rating approach for commercial loans : a Taiwan case", OMEGA, International Journal of Management Science, 1999, 27, p. 407-419.
10. Chorafas, D. N. "Expert system at the banker's reach", International Journal of Bank Marketing, 5, 72-81,1987.
11. Coats, P., and Fant, L.,"Recognizing Financial Distress Patterns Using a Neural Network tool", Financial Management, 22, 142-155, 1993.
12. Collins R. A. and Green R. D. "Statistical method for bankruptcy forecasting", Journal of Economics and Business, 1982, 34, p. 349-354.
13. De Neyer, M., Gorez, R., Barreto, J. "Fuzzy integral action in model based control systems", Second IEEE International Conference on Fuzzy Systems, 1, 172-177, 1993.
14. Dietrich J. R. and Kaplan R. S. "Empirical analysis of the commercial loan classification decision", The Accounting Review, 1982, 57(1),p. 18-38.
15. Duchessi, P., Shawky, H, and Seagle, J. P., "A knowledge-engineered system for commercial loan decision", Financial Management, 17, 57-65, 1988.
16. Edmister R. O. "Combining human credit analysis and numerical credit scoring for business failure prediction", Akron Business Economic Review, 1988, 19(3), p. 6-14.
17. Khan, Haider A.(1993), "Econometrics", in Outhwaite and T. Bottomore eds. *Encyclopaedia of Twentieth Century Social Thought*, Oxford, Basil Blackwell.

------------------(1997), "Does Japanese Bilateral Aid Work? Foreign Aid and Fiscal Behavior in a Bounded Rationality Model," <u>Regional Development Studies</u>, Vol. 3 (Winter, 1996/97), pp. 283-97.

18. Klir, G. J. and Yuan, B. (1995) Fuzzy Sets and Fuzzy Logic, Prentice-Hall.
19. Kosko, B., Neural Networks and Fuzzy Systems. A Dynamical Systems Approach to Machine Intelligence, 1992, Prentice Hall, Englewood Cliffs, NJ.
20. Lacher, R., Coats, P., Sharma, S., Fant, L., "A neural network for classifying the financial health of a firm", European Journal of Operational Research 85 (1), 53-65, 1995.
21. Lenard, M. J., Alam, P., and Madey, G. R.,"The Application of Neural Networks and a Qualitative Response Model to the Auditor's Going Concern Uncertainty

Decision", Decision Science, 26, 209-227, 1995.

22. Levy J., Mallach, E., and Duchessi, P. "a fuzzy logic evaluation system for commercial loan analysis", OMEGA Interanl Journal of Management Science, 1991, 19(6), p. 651-669.
23. Leszczynski, K., Penczek, P., and Grochulskki, W. "Sugeno's fuzzy measure and fuzzy clustering", Fuzzy Sets and Systems, 15, 147-158, 1985.
24. Lin, C.-T and Lee, C. G. (1996) Neural Fuzzy Systems. A Neuro-Fuzzy Synergism to Intelligent Systems. Prentice Hall, New York.
25. Marais, M. L., Patell, J. M., and Wolfson, M. A., "The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications", Journal of accounting Research, 22, 87-114, 1984.
26. Nauck, D. and Kruse, R. (1996) "Neuro-Fuzzy Systems Research and Applications outside of Japan", Fuzzy-Neural Networks, Soft Computing series, pp. 108-134. Asakura Publ., Tokyo.
27. Pindyck, R. S. and Rubinfeld, D. L., Econometric models and economic forecasts, McGraw-Hill, Fourth edition, 1998.
28. Quinlan, J. R., Comparing connectionist and symbolic learning methods, in : G. Hanson, G. Drastal, R. Rivest (Eds.), Computational Learning Theory and Natural Learning Systems: Constraints and Prospects, MIT Press, Cambridge, MA, 1993.
29. Romaniuk, D. G., Hall, L. O. "Decision making on creditworthiness, using a fuzzy connectionish model", Fuzzy Sets and Systems, 1992, 48, 9. 15-22.
30. Sharda, R., and Wilson, R. L., "Neural network experiments in business-failure forecasting: Predictive performance measurement issues", International Journal of Computational Intelligence and Organizations, 1,2, 107-117, 1996.
31. Sharma, S., Applied multivariate techniques, John Wiley & Sons, Inc., 1996.
32. Srinivasan, V., Kim, Y. H. "Credit granting: a comparative analysis of classification procedures", Journal of Finance, 1987, 42(3), p. 665-683.
33. Sugeno, M, Nishiwaki, Y., Kawai, H., and Harima, Y. "Fuzzy measure analysis of public attitude towards the use of nuclear energy". Fuzzy Sets and Systems, 20, 259-289, 1986.
34. Tahani, H., and Keller, J. M. "Information fusion in computer vision using fuzzy integral", IEEE Transaction of System, Man and Cybernetics, 20, 3, 733-741, 1990.
35. Tam, K. Y., and Kiang, M. Y., "Managerial applications of neural networks: The case of bank failure predictions", Management Science, 38, 7, 926-947, 1992.
36. Tong, R. M. and Bonissone, P.P."Linguistic Solutions to Fuzzy Decision problems," in Zimmermann, H.-J., Zadeh, L. A., and Gaines, B. R. (Eds.), Fuzzy

Sets and Decision Analysis, Amsterdam (1984), p. 323-334.
37. Rescher, N. Many-Valued Logic, McGraw-Hill, New York, 1969.
38. West, R.C. “A factor-analysis approach to bank condition”, Journal of Banking and Finance, 1985, 9, p. 253-266.
39. Wilson, R. L., and Sharda, R., “Bankruptcy prediction using neural networks”, Decision Support Systems, 11, 545-557, 1994.
40. Yang, B., Li, L. X, Ji, H., Xu, J.,”An early warning system for loan risk assessment using artificial neural networks”, Knowlede-Based Systems, 14, 303-306, 2001.
41. Yang, Z. R., “Probabilistic Neural Networks in Bankruptcy Prediction”, Journal of Business Research, 44, 67-74, 1999.
42. Zavgren, C. V. “Assessing the vulnerability to failure of American industrial firms: a logistic analysis”, Journal of Business Finance and Accounting, 12, 19-45, 1985.
43. Zimmermann, H. J. and Thole, U. (1978) “On the Suitability of Minimum and Product Operators for the Intersection of Fuzzy Sets,” Fuzzy Sets and Systems, Vol. 2, pp. 173-186.
44. Zimmermann, H. J. and Zysno, P. “Decision and evaluation by hierarchical aggregation of information”, Fuzzy Sets and Systems, 10, 243-260,1983.
45. Zimmermann, H.-J., Fuzzy Sets, Decision Making, and Expert Systems, Boston: Kluver Academic Publisher (1987). ISBN 089838-149-5.
46. Zimmermann, H. J. Fuzzy set theory and its applications. Dordrecht: Kluwer, 1991.