

NEURO-FUZZY MODELING IN BANKRUPTCY PREDICTION*

D. VLACHOS

*Department of Mechanical Engineering
Aristotle University of Thessaloniki, Thessaloniki, Greece*

Y. A. TOLIAS

*Telecommunications Laboratory
Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki, Greece*

Communicated by Byron Papathanassiou

Abstract: For the past 30 years the problem of bankruptcy prediction had been thoroughly studied. From the paper of Altman in 1968 to the recent papers in the '90s, the progress of prediction accuracy was not satisfactory. This paper investigates an alternative modeling of the system (firm), combining neural networks and fuzzy controllers, i.e. using neuro-fuzzy models. Classical modeling is based on mathematical models that describe the behavior of the firm under consideration. The main idea of fuzzy control, on the other hand, is to build a model of a human control expert who is capable of controlling the process without thinking in a mathematical model. This control expert specifies his control action in the form of linguistic rules. These control rules are translated into the framework of fuzzy set theory providing a calculus, which can stimulate the behavior of the control expert and enhance its performance. The accuracy of the model is studied using datasets from previous research papers.

Keywords: Neuro-fuzzy, bankruptcy.

1. INTRODUCTION

The ability to predict firm bankruptcies has been extensively studied in the accounting literature. Creditors, auditors, stockholders and senior managers all have a vested interest in utilizing and developing a methodology or a model that will allow them to monitor the financial performance of a firm. There exist extensive studies in this area using statistical approaches and Artificial Intelligence, most of which use

* Presented at 6th Balkan Conference on Operational Research

financial ratios as inputs in a forecasting model. Discriminant Analysis, proposed by Altman [1] has been used most frequently among statistical approaches in bankruptcy prediction. Previous empirical results show that neural network models provide higher predictive accuracy than statistical methods. [2] [4] [8] [11]

Expert systems (ES) represent a field of study within the Artificial Intelligence, which has earned the attention and commitment from business and information systems managers, as well as knowledge engineers. They have become critical components in many products and services, as well as in many decision-making processes. There are different types of ES depending on the mathematical tool used to induce knowledge from the available data. One of the major categories of ES is rule-based ES where knowledge is expressed through "if ... then ..." rules.

The first attempt of using expert systems for bankruptcy prediction was the one of Messier and Hansen [5]. The objective of the proposed 'data-driven' method was to take firms of known classes (bankrupt/non-bankrupt) described by a fixed set of attributes (financial ratios), and then to generate a production system using attributes which correctly classify all the firms of the sample. The rules at each stage (i.e. the variable and the cut-off score) were defined by using measures of entropy and selecting the minimum entropy rule. A decision tree was derived from the production system rules. Messier and Hansen's study was based on a sample of 23 firms (8 bankrupt and 15 non-bankrupt). The classification accuracy of the production system was encouraging in this small case.

Some other ES implications are presented in the review paper of Dimitras et al. [3]. The same paper summarizes most of the bankruptcy prediction methods.

In this paper, we present a rule-based expert system developed to predict the probability of business failure, where the rules are induced through an approach, which combines neural networks and fuzzy logic, usually referred in literature as neuro-fuzzy approach. Specifically, we use the Adaptive Network-based Fuzzy Inference System (ANFIS), which is a technique proposed by Roger Jang [7]. Section 2 briefly reviews a generalized model for fuzzy inference systems. Section 3 describes ANFIS, while section 4 presents the results and the performance evaluation of its implementation in bankruptcy prediction. Finally, section 5 advances some conclusions and recommendations for ES developers and managers to proceed.

2. A GENERALIZED MODEL FOR FUZZY INFERENCE SYSTEMS

Fuzzy if-then rules or *fuzzy conditional statements* are expressions of the form IF **A** THEN **B**, where **A** and **B** are *linguistic values of fuzzy linguistic variables*. Each linguistic variable is characterized by a *universe of discourse* and the *membership functions* of its values, defined on the same universe of discourse. The linguistic values are fuzzy sets. An example that clarifies the aforementioned definitions is the following:

IF **HEIGHT IS TALL**, THEN **WEIGHT IS HEAVY**.

Here **HEIGHT** and **WEIGHT** are the linguistic variables while **TALL** and **HEAVY** are the linguistic values of the linguistic variables, being characterized by appropriately defined membership functions. By using fuzzy if-then rules we can capture and encode expert knowledge in the form of a fuzzy inference system. For more details on the theory of fuzzy sets and the definition and use of linguistic variables the interested reader is referred to Zadeh's original work [12], [13].

A *fuzzy inference system* (also known as *fuzzy rule-based system*) is composed of five functional blocks, a *rule-base* that contains a number of fuzzy if-then rules, a *database* that defines the membership functions of the fuzzy sets used by the fuzzy rules, a *decision-making subsystem* that performs the inference operations on the rules, a *fuzzification interface* that transforms crisp measurement to degrees of membership to different fuzzy sets and finally, a *defuzzification interface* that transforms the fuzzy results into a crisp output (e.g. a control signal, a predicted value, etc). The block diagram of a fuzzy inference system is shown in Figure 2.1.

A fuzzy inference system performs the following processing steps on the given inputs:

1. *Fuzzification*: Compare the input variables with the membership functions that constitute the database on the premise part to obtain the membership values of each linguistic label.
2. *Determination of the firing strength of each rule* by combination of the membership values on the premise part.
3. *Generation of the consequent of each rule* depending of the firing strength.
4. *Defuzzification*: Aggregation of the consequents to produce a crisp output.

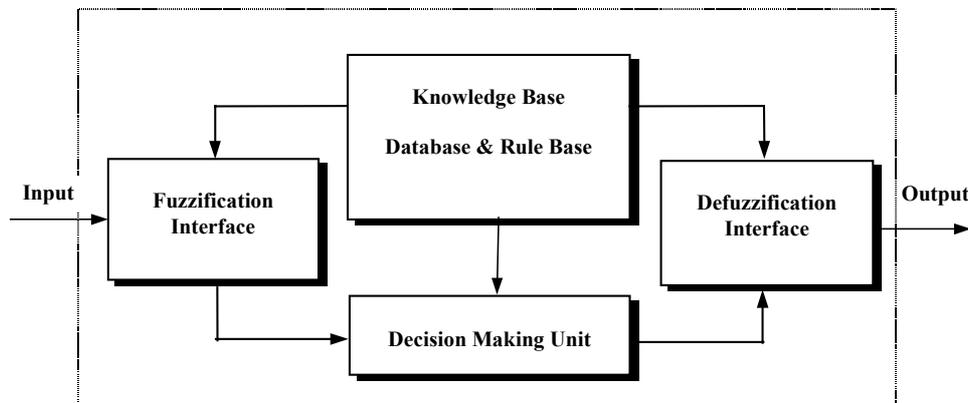


Figure 2.1: The block diagram of a fuzzy inference system

Depending on the type of fuzzy reasoning and the fuzzy if-then rules that are used, there are three types of inference systems. The overall output in the first type is the weighted average of each rule's crisp output induced by the rules firing strength and output membership functions. The output membership functions must be non-decreasing functions in the universe of discourse. In the second type, the overall fuzzy output is derived by applying the maximum operator to the fuzzy outputs and the final crisp output is calculated using an appropriate defuzzification method (area, bisector of

area, center of mass, and others). Finally, the third type of fuzzy inference systems uses the Sugeno type approach [9] and the fuzzy output for each rule is a linear combination of input variables with an additional constant term. The final output is the weighted average of each rule's output.

3. THE ADAPTIVE NETWORK-BASED FUZZY INFERENCE SYSTEM (ANFIS)

The Adaptive Network-based Fuzzy Inference System (ANFIS) that was proposed by Roger Jang [7] is one of the most commonly used fuzzy inference systems. In this section we highlight ANFIS's architecture. For additional details the reader is referred to [7].

ANFIS is a 5-layer feed-forward network in which each node performs a particular function in incoming signals as well as a set of parameters pertaining to the node. Let us suppose that the fuzzy inference system under consideration has two inputs x and y and one output z . Suppose that the rule base contains the following two Sugeno-type fuzzy if-then rules:

$$\text{R1: IF } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ THEN } f_1 = p_1x + q_1y + r_1,$$

$$\text{R2: IF } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ THEN } f_2 = p_2x + q_2y + r_2.$$

These rules correspond to the third category of fuzzy inference systems mentioned in [7]. The architecture of the equivalent ANFIS system is shown in Figure 3.1.

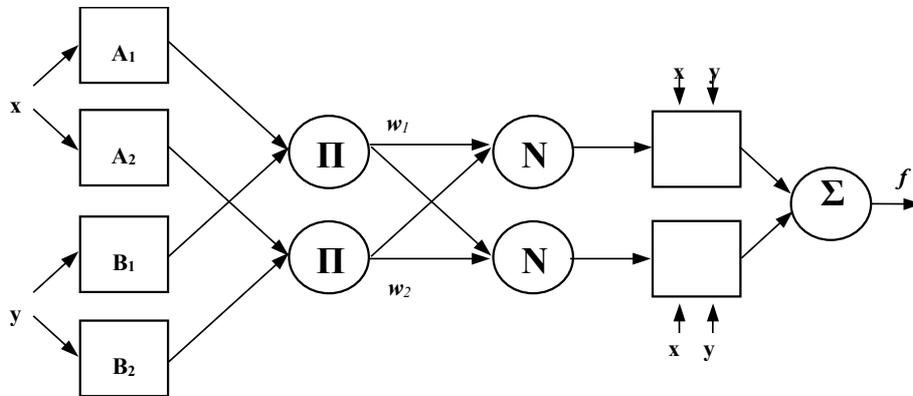


Figure 3.1: The architecture of ANFIS

Every node i that belongs to the first layer O^1 is characterized by the membership function $\mu_{A_i}(x)$, where A_i is the linguistic value of the input variable x . Usually, the membership functions are generalized bell functions, of the form:

$$O_i^1 = \mu_{A_i}(x, \alpha_i, \beta_i, \gamma_i) = \frac{1}{1 + \left[\left(\frac{x - \gamma_i}{\alpha_i} \right)^2 \right]^{\beta_i}} \quad (1)$$

The values $\{\alpha_i, \beta_i, \gamma_i\}$ constitute the parameter set that is adjusted using the learning algorithm.

The nodes O^2 belonging to layer 2 calculate the firing strength of a rule by finding the product of the membership values of the nodes of layer 1 that belong to a rule:

$$O_i^2 = w_i = \mu_{A_i}(x) \cdot \mu_{B_{i2}}(x), \quad i = 1, 2. \quad (2)$$

This operation is essentially the application of the product T -norm on the premise memberships.

The nodes O^3 belonging to layer 3 normalize the rules' firing strengths, according to the equation:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (3)$$

The nodes O^4 in layer 4 multiply the rules' firing strengths with the consequent parameters $\{p_i, q_i, r_i\}$,

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i). \quad (4)$$

Finally, the single node that constitutes layer 5 calculates the overall output, a crisp value, which is defined as follows:

$$O_1^5 = \frac{\sum_i w_i f_i}{\sum_i w_i}. \quad (5)$$

Learning in ANFIS is implemented by employing a hybrid learning rule that combines the gradient-descent method and the least squares estimate to identify the sets of parameters $\{\alpha_i, \beta_i, \gamma_i\}$ and $\{p_i, q_i, r_i\}$. The back propagation learning rule is applied to tune the parameters in the hidden layers and the parameters in the output layer are identified by the least squares method. Each epoch of this hybrid procedure is composed of a forward and a backward pass. In the forward pass the premise parameters are kept fixed and the consequent parameters are estimated by the least squares method. In the backward pass, the consequent parameters are kept fixed and the premise parameters are calculated by gradient descent. Again, for more details the reader is referred to [9].

4. NUMERICAL INVESTIGATION

The first methodological step is the determination of input variables (ratios) for the model. Since the purpose of this study was to examine the value of applying

expert systems in bankruptcy prediction and not to evaluate the appropriate input variables, we used the simple financial ratios employed by Altman [1]. These ratios were:

- Working Capital/Total Assets (WC/TA)
- Retained Earnings/Total Assets (RE/TA)
- Earning before Interest and Taxes /Total Assets (EBIT/TA)
- Market Value of Equity/Total Debt (MVE/TD)
- Sales /Total Assets (S/TA)

The sample under study consists of firms that either were in operation or went bankrupt between 1975 and 1982. It was obtained from Moody's Industrial Manuals, and included a total of 129 firms, out of which 65 went bankrupt during the period and 64 non bankrupt firms matched on industry and year. Data used for the bankrupt firms were taken from the last financial statement issued before the firms declared bankruptcy. The sample was chosen the same used by Wilson and Sharda [10] and Rahimian et al. [6] to make possible direct comparison. The numerical values of the entire data set are shown in Figure 4.1. The horizontal axis in the figure represents the firm index. The first 65 index numbers refer to bankrupt firms.

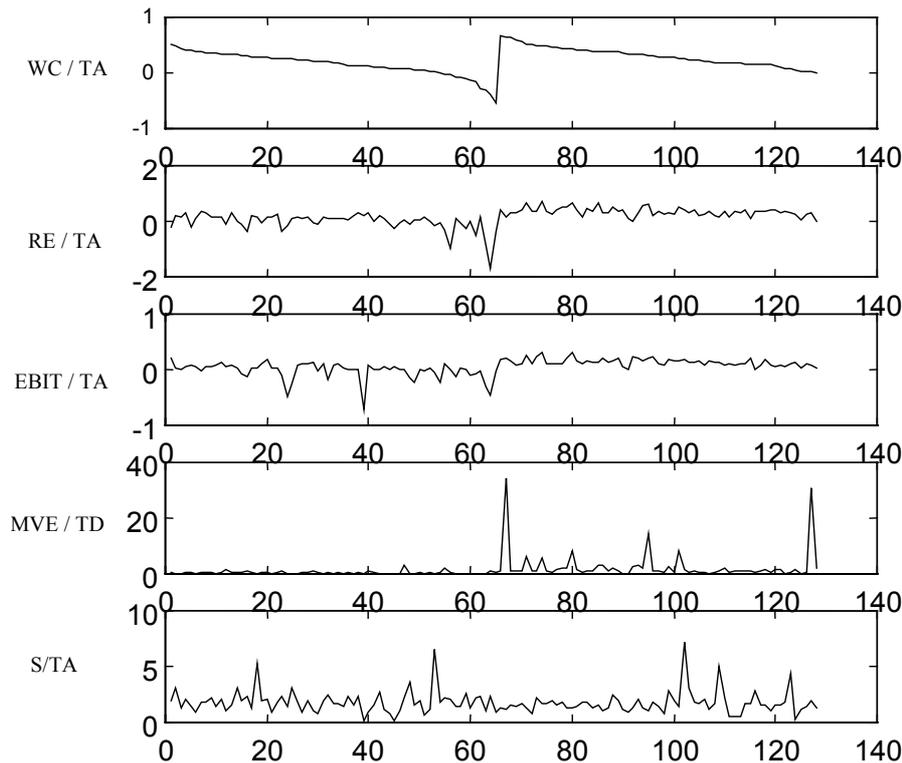


Figure 4.1: The numerical values of the entire dataset used in this paper

This method was implemented in a Pentium II PC using Matlab (The Mathworks, Natick, MA, USA).

First, two data sets were generated; a training data set consisting of 74 vectors that were picked in random and a testing data set consisting of the remaining 55 data vectors. For the training data set, the covariance matrices and the mean value vectors were calculated for the bankrupt and the non-bankrupt samples, \mathbf{C}_b , \mathbf{C}_{nb} , \mathbf{m}_b and \mathbf{m}_{nb} , respectively.

For every bankrupt/non bankrupt company vector, a covariance transform was applied in order to provide uncorrelated, zero mean training samples:

$$\mathbf{y}_{b/nb} = (\mathbf{x}_{b/nb} - \mathbf{m}_{b/nb})\mathbf{C}_{b/nb}^{-1} \quad (6)$$

The dataset \mathbf{y} that results from this covariance transform is shown in Figure 4.2.

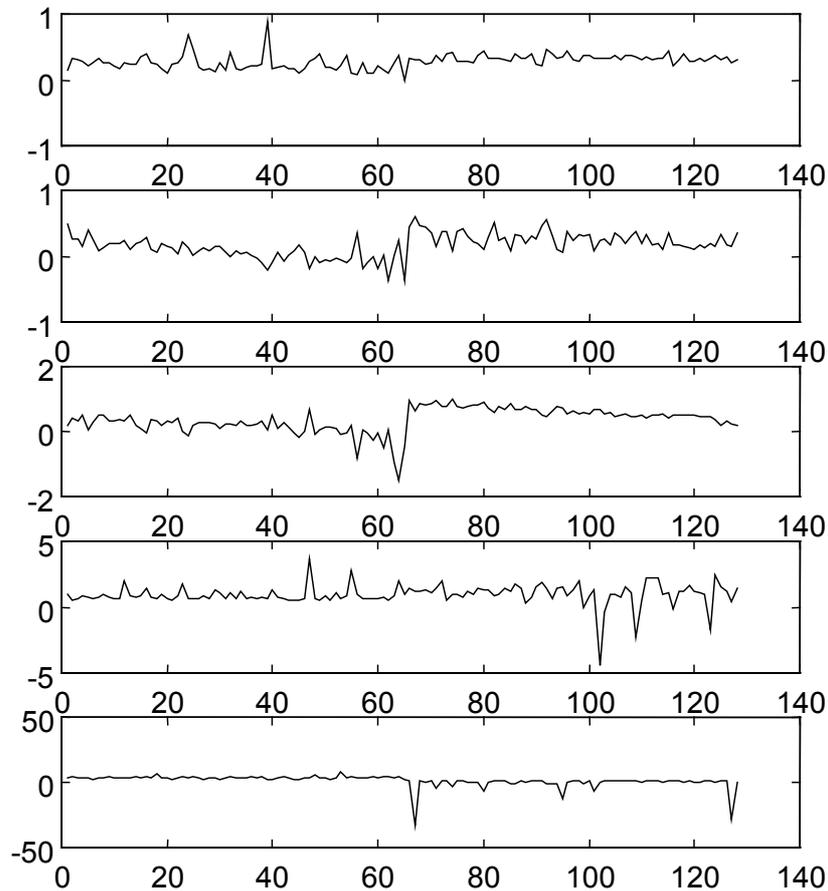


Figure 4.2: The dataset after the covariance transformation

We used the transformed training data set to train ANFIS and obtain a fuzzy inference system having 5 input variables (the transformed ratios). Its output shows the possibility of not going bankrupt, i.e. the **Non Bankrupt Index (NBI)**. ANFIS required 100 training epochs to reach a Root-Mean-Squared-Error (RMSE) of 0.034. The plot of the training error per epoch is shown in Figure 4.3. ANFIS produced 32 fuzzy rules of the form mentioned in Sec. 3.

The system was evaluated using the testing data set, by utilizing the following approach:

Since we are not aware whether a testing vector belongs to bankrupt or non-bankrupt company, we need to generate two hypotheses for every testing vector.

That is,

H0: The vector corresponds to a company that is **not** going to get bankrupt, and

H1: The vector corresponds to a company that is going to get bankrupt.

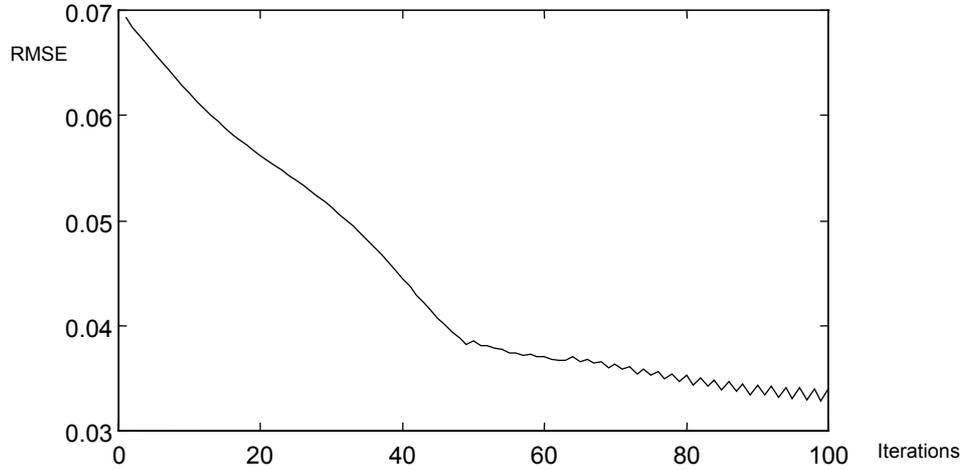


Figure 4.3: The plot of the RMSE per training epoch

For each hypothesis, a covariance transformation is applied using the corresponding vectors and matrices:

$$\mathbf{y}_{test/H0} = (\mathbf{x}_{test} - \mathbf{m}_{nb})\mathbf{C}_{nb} \quad (7)$$

$$\mathbf{y}_{test/H1} = (\mathbf{x}_{test} - \mathbf{m}_b)\mathbf{C}_b \quad (8)$$

A different value of the Non-Bankrupt Index is calculated by ANFIS for each vector in (7,8), i.e., $NBI_{test/H0}$ and $NBI_{test/H1}$. From these values we select the ones having the lowest fuzziness, according to the rule:

$$NBI = \begin{cases} NBI_{test/H0}, & \text{if } |NBI_{test/H0} - 0.5| \geq |NBI_{test/H1} - 0.5| \\ NBI_{test/H1}, & \text{if } |NBI_{test/H1} - 0.5| > |NBI_{test/H0} - 0.5| \end{cases} \quad (9)$$

The results of the application of the proposed scheme to the entire dataset are shown in Figure 4.4.

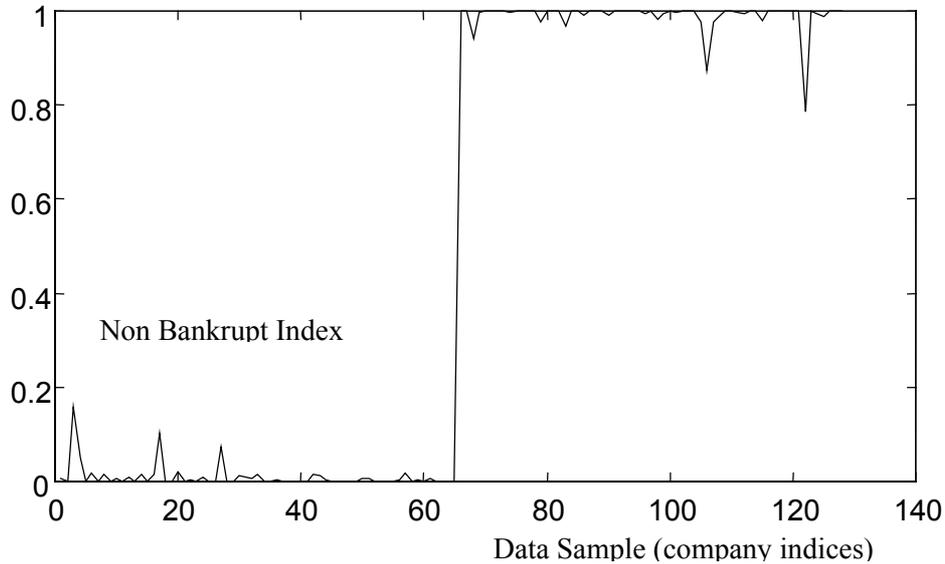


Figure 4.4: The Non-Bankrupt Index for the entire data set

From Figure 4.4 we conclude that the entire dataset (training and test data) are correctly classified by the Non-Bankrupt Index. So, ANFIS achieved 100% correct classifications, while the results of previous research were worse. Specifically, discriminant analysis gives 85% correct classifications and the best neural network approach of Wilson and Sharda [10] gives 93%. The paper of Rahimian et al. [6] provides analytic results of applying different methods in the same dataset.

5. CONCLUSIONS

This paper has compared the predictive capability of neuro-fuzzy techniques with that of previous research paper within the context of forecasting firm bankruptcies on the basis of a small number of financial ratios. In this study the proposed system ANFIS outperformed both classical methods and modern approaches. Although the method is tested with a small sample of firms, the 100% correct classification of bankrupt and non-bankrupt firms shows that this is a robust and promising approach in the prediction of firm stability. Thus, the research on bankruptcy prediction with neuro-fuzzy methods must continue.

REFERENCES

- [1] Altman, E.I., "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *Journal of Finance*, (1968) 589-609.
- [2] Coats, P.K., and Fant, F., "Recognizing financial distress patterns using a neural network approach", *Financial Management*, 22(3) (1992) 142
- [3] Dimitras, A.I., Zanakis, S.H., and Zopounidis, C., "A survey of business failures with an emphasis on prediction methods and industrial applications", *European Journal of Operational Research*, 90 (1996) 487-513.
- [4] Flecher, D., and Goss, E., "Forecasting with neural networks-an application using bankruptcy data", *Information and Management*, 24 (1993) 159-167.
- [5] Messier, W., and Hansen, J., "Inducing rules for expert system development: an example using default and bankruptcy data", *Management Science*, 34 (1988) 12.
- [6] Rahimian, E., Singh, S., Thammachote, T., and Virmani, R., "Bankruptcy prediction by neural network", in: R., Trippi, and E. Turban (eds.), *Neural Networks in Finance and Investing*, Irwin, 1996.
- [7] Roger Jang, J.S., ANFIS: "Adaptive network-based fuzzy inference system", *IEEE Transactions on Systems, Man and Cybernetics*, 23 (3) (1993) 665-685.
- [8] Salchenberger, L., Mine Cinar, E., and Lash, N., "Neural networks: a new tool for predicting thrift failures", *Decision Sciences*, 23 (1992) 899-916.
- [9] Takagi T., and Sugeno, M., "Fuzzy identification of systems and its application to modeling and control", *IEEE Transactions on Systems, Man and Cybernetics*, 15 (1985) 116-132.
- [10] Wilson, R., and Sharda, R., "Bankruptcy prediction using neural networks", *Decision Support Systems*, 11 (1994) 545-557.
- [11] Yoon, Y., Guimaraes, T., and Swales, G., "Integrating neural networks with rule-based expert systems", *Decision Support Systems*, 11 (1994) 497-507.
- [12] Zadeh, L.A., "Fuzzy sets", *Information Control*, 8 (1965) 338-353.
- [13] Zadeh, L.A., "Outline of a new approach to the analysis of complex systems and decision processes", *IEEE Transactions on Systems, Man and Cybernetics*, 3 (1) (1973) 28-44.