Applying Internal Analysis Data and Non-Linear Genetic Algorithm in Developing a Predicting Pattern of Financial Distress

Zahra Poorzamani
Department of Accounting, Assistant Professor, Central Tehran Banch, Islamic Azad University, Tehran, Iran,
E- mail: zahra.poorzamani@yahoo.com
Mostafa Nooreddin
Master of Accounting, Central Tehran Branch, Islamic Azad University, Tehran, Iran
E- mail: Mostafa.noreddin@yahoo.com

Abstract: Bankruptcy is an event with strong impacts on management, shareholders, employees, creditors, customers and other stakeholders, so as bankruptcy challenges the country both socially and economically. The aim of this study is to make a financial distress predicting model for listed companies' in Tehran stock exchange using financial proportions and artificial intelligent techniques. So financial information relevant to time period 1992 to 2011 is compiled and expected financial proportions' are extracted and neural network patterns (ANN), principal component analysis combination, and Non-Linear Genetic Algorithm (PCA +NON-LIN) have been compiled to predict the financial distress. Then according to obtained results, these patterns have been compared and the best pattern has been chosen. In accordance with the results, It is distinguished that the neural net work using the information One year before financial distress occurring has more efficiency in predicting the financial distress of the companies rather than the other technique in this research.

[Zahra Poorzamani. **Applying Internal Analysis Data and Non-Linear Genetic Algorithm in Developing a Predicting Pattern of Financial Distress.** *Life Sci J* 2013;10(7s):58-63] (ISSN:1097-8135). http://www.lifesciencesite.com. **10**

Key words: Financial Distress, Financial Variables, Non-Linear Genetic Algorithm, Neural Network

1. Introduction

Competition intensification at industry level has led to bankruptcy of many firms and their removal from the competition field. This has given rise to some concerns among shareholders, managers, creditors and in general the whole society. Investors by estimation of financial crises and bankruptcy of firms try to prevent loss of their capital. If management of business unit is timely informed of bankruptcy risk, it can take preventive actions. Creditors are very sensitive about loss of their principal amount and interest in the granted loans and credits to potential and current customers, and since it imposes heavy economic and social bankruptcy costs on the society, it is also interesting from macroeconomic point of view, because the lost resources in the distressed economic unit could have been used for other profitable opportunities. Given the importance of bankruptcy, all people and stakeholders are interested in bankruptcy prediction before its actual occurrence.

Financial crisis or bankruptcy prediction using historical financial data is well known. Although the first effort in this relation dates back to 1930, but from 1966 and following the research carried out by Beaver on this topic it took a more serious form (Dimitras et al, 1996). Beaver (1966) is one of the first researchers who investigated prediction of financial crisis or bankruptcy and is regarded as one of the leading academic researchers in this field. After him, Altman (1968) using advanced statistical techniques succeeded in achieving significant results. Today, numerous prediction models are introduced by researchers. These models, according to the model construction method, number of model's variables, definition of bankrupt firms and other firms, are classified into various groups, such as traditional model versus Artificial Intelligence models, or univariate models versus multivariate models.

This research, in addition to setting the prior research in financial crisis prediction as the departure point which is based on data of one or several years, intends to construct and design a financial crisis prediction model based on financial variables during the understudy years using the Artificial Intelligence and Genetic Algorithm corresponding to Iranian economic condition.

2. Research literature review

Bankruptcy prediction techniques according to their nature are classified into three groups of (classical) statistical approaches, Artificial Intelligence techniques, and Theoretical models. Data mining models and Artificial Intelligence Techniques performs tasks similar to human's knowledge, intelligence and logic. In fact, the AI is a system which learns and improves performance of its problem solving given the past experiences. AI application in finance and particularly in bankruptcy prediction does not have a long record, yet due to its high efficiency and being free from the existing restrictive assumptions in statistical methods, it has been widely accepted by the researchers. These models are mainly focused on signals of commercial failure, are generally multivariate and the used variables in them are derived from the information available in the firm's accounts. Intelligent techniques are composed of neural networks, genetic algorithms, hard sets, Support Vector Machine, reasoning based on Fuzzy and logic and issues. Many studies have been carried out on application of these techniques for prediction of businesses failure among which it can be referred to Etemadi, Rostami and Farajzadeh Dehkordi (2009), Huang, Tsai, Yen and Cheng (2008), Hung and Chen (2009), Lin et al (2009), Min an Jeong (2009), Min and Lee (2008), Ravi and Pramodh (2008), Sun and Li (2008), and Wu (2010).

For structuralization of Computer systems neural networks, human learning process and inference pattern are followed. Architecture of neural networks in general is consisted of three input layers including input information, throughput (hidden) layer, and output layer. Identification of the best architecture for problem solving is a complex and difficult task, and the best architecture is obtained by trial and error.

Neural Networks (NN) which are also known as Artificial Neural Networks (ANN) are a data mining technique used to solve many problems. Among advantages of Neural Networks relative to other methods, according to the rules observed in data mining and artificial intelligence systems, it can be referred to the following ones:

Since Neural Networks do not need a knowledge base for being structuralized, their use in problems about which there is little knowledge is useful.

Processing in NN can be performed at high speed and accuracy relative to traditional methods, because these networks simultaneously examine all the information existing in one problem and the processing units or the neurons function along each other (Back et al, 1996).

In these networks, the type of data distribution or communication structure of the existing variables does not require to be considered as the basic assumption (Wu et al, 2006). If the input data are incomplete and disturbed or have a high correlation with each other, or have not been already observed, traditional systems are hardly able to extract rules and patterns, but in the same conditions, neural networks provide reasonable answer, and after learning and adaptation, they will be able to generalize the results to similar instances.

However, NNs have some disadvantages as well. Their main limitation is lack of a definitive method for specification of optimal architecture. To design an NN model in solving a problem of classification type, there is no systematic principle and method, as a result, the best network topology is specified by trial and error. Many factors such as hidden layers, number of neurons in hidden layers, data normalization and learning algorithm can affect the network's performance. In defining network's architecture, one has to take this fact into account that a greater number of layers leads to more complexity of the network and to a problem called 'over-fitting' and non-usability of new data. In general, with increase in number of the middle neurons, the network's power in identification of existing complexities increases, but this may reduce the network's generalizability. In other words, if number of neurons in the middle layer is too large, the network memorizes in place of learning. Another shortcoming of these networks concerns their performance as a black box. Comprehension and verification of the mode these networks classify data and quality of relationships in layers' structure is not possible for the user. NN method does not specify significance of each variable in the final classification, and the assigned weights in this regard are not interpretable (Fallahpour, 1383). From among the researchers who employed NNs for prediction of bankruptcy and financial crisis in different countries it can be referred to Coates and Fant (1992), Serrano and Cinca (1996), and Shah and Murtaza (2000).

3. Theoretical Framework and Research Hypotheses

By contemplation on bankruptcy prediction models, it can be found that all of them are somehow the heritage of statistical techniques. For instance, AI models generally use both unvaried and multivariate techniques so as they can be considered as the children of mechanized statistical techniques. Similarly, theoretical models often are derived from application of an appropriate statistical technique and are not directly derived from theoretical principles.

Artificial Neural Networks (ANN) is flexible and non-parametric modeling tools. They are able to execute every complex function with a satisfactory accuracy. The first effort for use of ANNs for bankruptcy prediction was made by Adam and Sharda (1999).

Franco Varreto (1998) used Genetic Algorithm for bankruptcy prediction. His sample included 500 companies, comprising 236 bankrupt firms and 264 non-bankrupt firms. Results of this research indicated a prediction accuracy of 93% one year ahead of the bankruptcy event and a prediction accuracy of 91.6% three years prior to bankruptcy. In addition, in this research, by comparing Genetic Algorithm with traditional prediction models, it was voted for superiority of genetic process, because these models in addition to being free from the restrictive assumptions, relative to traditional methods have a higher accuracy. In traditional models, with increase of time horizon from bankruptcy event, the model's accuracy significantly decreases, while this accuracy reduction is far less with GA models. Among other studies on this subject, it can be referred to Shin and Lee (2002) and McKee and Lensberg (2002).

Despite the large number of researches on the issue of bankruptcy prediction, few desirable results have been found (Plat & Plat, 1990). The created models did not come successfully out of the Test of Robustness. This may be due to several factors. One reason can be that the statistical models tend to use Matched Pairs of bankrupt and non-bankrupt firms. Next, data of the variables specifying cutoff points (thresholds) have been used for distinction of bankrupt from non-bankrupt firms. The used data for extraction of thresholds were year specific, and were tried for correct identification of firms in the hold-out sample (similar time periods). These cutoffs have less efficiency when used in different time periods (including their use in next studies).

Farajzadeh Dehkordi (2005), in his master thesis, investigated bankruptcy prediction modeling of the listed companies on the stock exchange using two models of Multiple Discriminant Analysis (MDA) and Genetic Planning. To construct the above models, first, he prepared a full list of financial ratios (93 ratios) and after study of the ratios, eventually, he extracted 42 financial ratios for construction of the models, and using Independent Samples Test (IST) of the two societies he constructed the two intended models. In fine, the Genetic Planning and MDA models succeeded in correct classification of the firms present in the training set with an accuracy of 94% and 77%, and firms present in the hold-out set with an accuracy of 90% and 73%, respectively.

Saee (2008) investigated application of Support Vector Machine in prediction of firms' financial insolvency using financial ratios. His research was mainly focused on the use of Support Vector Machine in prediction of firms' insolvency. The results obtained from this model were compared with those obtained from the Logistic Regression model and Support Vector Machine was found superior to Logistic Regression model.

Therefore in the present research, based on literature review and given the purposes, the main hypothesis is as follows:

Prediction power of Neural Networks (NN) models is greater than Non-linear Genetic Algorithm (PCA + NON-LIN) which are based on internal analysis

4. Research methodology

This study is of applied type which employs a survey-exploratory methodology based on correlation. The information regarding the under study financial ratios was extracted from the TSE software 'RAHAVARD'.

Statistical population in this research includes the listed companies on the TSE that from 1996 through to 2008 have reported their financial ratios to the stock exchange. The statistical sample of the under study firms are divided into two major groups:

First group: includes 73 financially healthy firms or firms without financial crisis. The main criterion in selection of this group of firms was non-subjection of these companies to article 141 of the Commercial Law during the under study period.

Second group: includes 69 firms in financial crisis, and the main criterion in their selection was subjection of this group to article 141 during the period in question.

Data analysis method: first using RAHAVARD software, financial statements of all listed firms on the stock exchange during 1992 to 2011 were extracted. Next, using Excel software, the intended financial ratios were calculated and given article 141 of the Commercial Law, the firms were divided into two groups of financially healthy and distressed companies.

In the group of financially healthy firms, two random selection phases were used; first, from the firms' total population, the sample firms were selected, next, the intended fiscal year was randomly selected from the period 1992 to 2011.

In the group of financially distressed firms, given limited number of the firms, random selection was not possible, so all the firms that for several successive years or only for one year have been subjected to article 141 of the Commercial Law all financial information of whom was fully available were placed in the sample.

After selection of the sample firms, each group of companies were once again randomly divided in two training and hold-out groups, as presented in table 1.

Table 1: Sample division

Firm	Samples	Training samples	Hold-out samples
Financially healthy	73	58	15
Financially distressed	69	54	15
Total number	142	112	30

It should be noted that in all models, the above sample structure has been used.

In this research, dependent variables are both financially healthy and distressed companies. Independent variables in this research include financial ratios as presented in table 2.

Table 2: Used Financial Ratios (independent variables)

Variable	Financial ratio	Variable	Financial ratio
X1	Working capital to equity	X13	Total debt to accumulated profit or loss
X2	Working capital to sales	X14	Total debt to total asset
X3	Working capital to total debt	X15	Accumulated profit or loss to total asset
X4	Working capital to total asset	X16	Operational margin to sale
X5	EBIT to equity	X17	Financial cost to gross profit
X6	EBIT to sales	X18	Current asset to total asset
X7	EBIT to total debt	X19	Current asset to current debt
X8	EBIT to total asset	X20	Net profit to sale
X9	Equity to total debt	X21	Net profit to total asset
X10	Equity to total asset	X22	Current debt to total asset
X11	Sales to total debt	X23	Current debt to equity
X12	Sales to total asset		

5. Research Findings

5-1. Results obtained from Neural Networks (NN)

Table 3 represents the obtained results from the empirical test of ANN model for prediction of financial crisis. This model succeeded in correct classification of the firms present in the training, hold-out, and total sample into financially healthy and distressed firms with a general accuracy of 100%, 95.83% and 99.19%, respectively, so as 100, 24 and 124 firms present in the training, hold-out and total sample, 100, 23, and 123 firms have been correctly classified.

Study of the results obtained from this model in the training data indicates that ANN model had an accuracy of 100% in correct classification of financially distressed firms in this set (i.e. from 50 financially distressed firms present in this set, 50 firms have been correctly classified). In addition, this model had an accuracy of 100% in correct classification of financially healthy firms in this set (i.e. from among 50 financially healthy firms in this set, 50 firms have been correctly classified).

Study of the results obtained from this model in the hold-out data indicates that ANN model had an accuracy of 100% in correct classification of financially distressed firms in this set (from among 12 financially distressed firms in this set, 12 firms have been correctly classified). In addition, this model has an accuracy of 91.67% in correct classification of financially healthy firms in this set (from 12 financially healthy firms in this set, 11 firms have been correctly classified).

Study of total results from this model (in training and hold-out sets) indicates that ANN model has an accuracy of 100% in correct classification of financially distressed firms in this set (from among 62 financially distressed firms in this set, 62 firms have been correctly classified). In addition, this model has an accuracy of 9839% in correct classification of financially healthy firms in this set (from 62 financially healthy firms in this set, 61 firms have been correctly classified).

Table 3: The ANN experimental results

result	Training samples			Hold-out samples			Total samples		
	0	1	Total	0	1	Total	0	1	Total
Number 1	0	50	50	0	12	12	0	62	62
Number 0	50	0	50	11	1	12	61	1	62
Percentage 1	%0	%0	%100	%0	%100	%100	%0	%100	%100
Percentage 0	%100	%100	%100	%91.67	%8.33	%100	%98.39	%1.61	%100
General percentage	%100	%100	%100	%91.67	%100	%95.83	%98.39	%100	%99.19

5-2. Results obtained from combined model of Principal Constituents Analysis and Non-Linear Genetic Algorithm (PCA.NON-LIN)

Table 4 presents the results obtained from PCA + NON-LIN experimental model for prediction of financial crisis. This model succeeded in correct classification of firms present in training, hold-out and total samples into two groups of financially distressed and healthy firms with a general accuracy of 89%, 79.17%, and 87.10%, respectively, so as from 100, 24 and 124 firms present in the training, hold-out and total sample, 89, 19, and 108 firms have been correctly classified.

Study of the results obtained from this model in the training data indicates that PCA + NON-LIN model had an accuracy of 94% in correct classification of financially distressed firms in this set (i.e. from 50 financially distressed firms present in this set, 47 firms have been correctly classified). In addition, this model had an accuracy of 84% in correct classification of financially healthy firms in this set (i.e. from among 50 financially healthy firms in this set, 42 firms have been correctly classified).

Study of the results obtained from this model in the hold-out data indicates that PCA + NON-LIN model had an accuracy of 83.33% in correct classification of financially distressed firms in this set (from among 12 financially distressed firms in this set, 10 firms have been correctly classified). In addition, this model has an accuracy of 75% in correct classification of financially healthy firms in this set (from 12 financially healthy firms in this set, 9 firms have been correctly classified).

Study of total results from this model (in training and hold-out sets) indicates that PCA + NON-LIN model has an accuracy of 91.94% in correct classification of financially distressed firms in this set (from among 62 financially distressed firms in this set, 57firms have been correctly classified). In addition, this model has an accuracy of 82.26% in correct classification of financially healthy firms in this set (from 62 financially healthy firms in this set, 51 firms have been correctly classified).

Table 4: Results of PCA + NON-LIN experimental test

result	**Training samples**			**Hold-out samples**			**Total samples**		
	0	1	Total	0	1	Total	0	1	Total
Number 1	3	47	50	2	10	12	5	57	62
Number 0	42	8	50	9	3	12	51	11	62
Percentage 1	6%	94%	%100	16.67%	83.33%	%100	8.06%	91.94%	%100
Percentage 0	84%	16%	%100	75%	25%	%100	82.26%	17.74%	%100
General percentage	84%	94%	89%	75%	83%	79.17%	82.26%	91.94%	87.10%

6. Conclusion

In the last several decades, extensive researches on bankruptcy especially bankruptcy prediction have been carried out. In this regard, several bankruptcy prediction models have been introduced. Generally, reapplication of these models to the set of data different from the initial data (different data in terms of time or the data regarding different economic systems) has not been able to repeat the earlier successes. Therefore, correct prediction of bankruptcy is of high importance in the financial world. However, in the course of time, conditions change, as a result, the used variables in the models lose their efficiency (Haber, 2006). In addition, the economic systems based on which these models are designed vary per sector or country (Grice and Dugan, 2001). Therefore, design of a model corresponding to economic condition of each particular country and use of variables fitting its economic and financial system are necessary.

This research intends to investigate financial crisis prediction power using models based on Neural Networks and to compare it with Non-Linear Genetic Algorithm. The obtained results according to tables 5 regarding prediction power indicate superiority of ANN to the other model as well as extraction of identical results from all the other models. Therefore the main hypothesis confirmed.

Table 5: Models accuracy test results

Data/ models		Total samples	ANN	PCA + NON-LIN
Training samples	Number	100	100	89
	Percentage		100%	89%
Hold-out samples	Number	24	23	19
	Percentage		95.83%	79.17%
Total samples	Number	124	123	108
	Percentage		99.19%	87.10%

Acknowledgement

This article is based on the research plan, "Compiling the prediction patterns for financial distress by artificial intelligent techniques (two viewpoints of comparative and interned analysis)", that profited from advice, encouragement and support of Central Tehran Branch, Islamic Azad University, Tehran, Iran.

References

1- Altman, E.I. (1968), “Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy”, Journal of Finance, 23, pp. 589–609.
2- Beaver, W.H. (1966), "Financial Ratios as Predictors of Failure", Journal of Accounting Research 4, Empirical Research in Accounting: Selected studies, pp. 71-111.
3- Etemadi, H., Rostamy, A. & Dehkordi, H. (2009), "A Genetic Programming Model for Bankruptcy Prediction: Empirical Evidence from Iran", Expert Systems with Applications 36(2), pp. 3199–3207.
4- Farajzadeh Dehkordi, H, Etemadi, H, & Rostami, A (2005), “Application of Genetic Algorithm in Bankruptcy Prediction Modeling”, Master Thesis in Accounting, Faculty of Human Sciences, Teachers Training University, p.80
5- Haber, J. (2006), “Theoretical Development of Bankruptcy Prediction Variables”, the Journal of Theoretical Accounting Research, 2, pp. 82-101
6- Huang, S., Tsai, C. F., Yen, D. & Cheng, Y. (2008), "A Hybrid Financial Analysis Model for Business Failure Prediction", Expert Systems with Applications 35(3), pp. 1034–1040.
7- Hung, C., & Chen, J. (2009), "A Selective Ensemble based on Expected Probabilities for Bankruptcy Prediction", Expert Systems with Applications 36(3), pp. 5297–5303.
8- McKee, T.E., Lensberg, T. (2002), “Genetic Programming and Rough Sets: a Hybrid Approach to Bankruptcy Classification”, European Journal of Operational Research, 138, pp. 436-51.
9- Min, J., & Jeong, C. (2009), "A Binary Classification Method for Bankruptcy Prediction", Expert Systems with Applications 36(3), pp. 5256–5263.
10- Min, J.H., & Lee, Y.C. (2008), "A Practical Approach to Credit Scoring", Expert Systems with Applications 35(4), pp. 1762–1770.
11- Ravi, V., & Pramodh, C. (2008), "Threshold Accepting Trained Principal Component Neural network and Feature Subset Selection: Application to Bankruptcy Prediction in Banks", Applied Soft Computing 8(4), pp. 1539–1548.
12- Saee, R & Fallahpour, S, (2008), “Application of Support Vector Machine in Prediction of Firms’ Financial Insolvency using Financial Ratios”, Quarterly Journal of Accounting and Audit Studies, 53: 17-34
13- Serrano, Cinca, C. (1996-97), "Feed-Forward Neural Networks in Classification of Financial Information", the European Journal of Finance 3, pp. 183-202.
14- Shah, J.R., Murtaza, M.B. (2000), "A Neural Network Based Clustering Procedure for Bankruptcy Prediction", American Business Review 18(2), pp. 80-86.
15- Shin, K., Lee, Y. (2002), “A Genetic Algorithm Application in Bankruptcy Prediction Modeling”, Expert Systems with Applications, 23 (3), pp. 321-8.
16- Tam, K.Y., & Kiang, M.Y. (1992), "Managerial Applications of Neural Network: The case of Bank Failure Predictions", Management Science 38(7), pp. 926-947.
17- Wu, W.W. (2010), "Beyond Business Failure Prediction", Expert Systems with Applications 37(3), p. 2371–2376

1/8/2013