INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES

HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY
EXPLORING INTELLECTUAL CAPITAL

www.hrmars.com
ISSN: 2225-8329

The Bankruptcy Prediction by Neural Networks and Logistic Regression

Ahmad Ahmadpour Kasgari, Seyyed Hasan Salehnezhad, Fatemeh Ebadi

To Link this Article: http://dx.doi.org/10.6007/IJARAFMS/v3-i4/386 DOI:10.6007/IJARAFMS /v3-i4/386

Received: *10* October *2013,* ***Revised:*** *11* November 2013*,* ***Accepted:*** *29* November 2013

Published Online: 16 December 2013

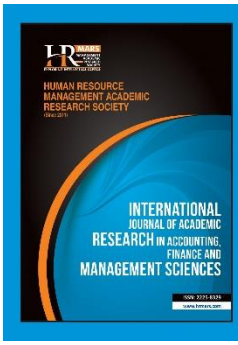
In-Text Citation: (Kasgari et al., 2013)

To Cite this Article: Kasgari, A. A., Salehnezhad, S. H., & Ebadi, F. (2013). The Bankruptcy Prediction by Neural Networks and Logistic Regression. *International Journal of Academic Research in Accounting Finance and Management Sciences. 3(4),* 191 – 199.

Vol. 3, No. 4, 2013, Pg. 191 - 199

http://hrmars.com/index.php/pages/detail/IJARAFMS | JOURNAL HOMEPAGE

Full Terms & Conditions of access and use can be found at
http://hrmars.com/index.php/pages/detail/publication-ethics


INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES
HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY
EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2225-8329

The Bankruptcy Prediction by Neural Networks and Logistic Regression

Ahmad Ahmadpour Kasgari[1], Seyyed Hasan Salehnezhad[2], Fatemeh Ebadi[3]

[1]Department of Accounting, Faculty of Economics and Administrative Science, Mazandaran University, Iran, [2]Department of Management, Economics and Accounting, Payame Noor University, I.R. of Iran, [3]Department of Accounting, Science and Research Branch, Islamic Azad university, Mazandaran, Iran

Email: ahmadpour@umz.ac.ir, shs_489@yahoo.com, f.ebadi66@yahoo.com

Abstract

Today, the intensity of industry competition has led many companies going bankrupt and pulling out of race. The early warning against the possibility of bankruptcy enables the managers and investors to take pre-emptive actions when it is necessary. The bankruptcy prediction models reveal the latent problems in financial structures like a warning bell and provide timely feedback to managers and investors as well as other people who benefit from this. The bankruptcy of manufacturing companies in Tehran Stock Exchange Market has been predicted in this study using artificial neural network in this respect. It has been also used the logistic regression to do compare with neural network as well. All information which has been used here is related to time periods from 2001 to 2011 and the bankrupt groups have been selected on the basis of Article 141 of the Commercial Code of Iran. In the years before bankruptcy, the financial management has the chance to predict the probability of bankruptcy by using this model and take necessary actions in this regard since the results derived from the neural network predictions are very consistent with reality. Moreover, this model is more accurate than that of logistic regression in prediction process.

Keywords: Bankruptcy Prediction, Artificial Neural Networks, Logistic Regression, Back Propagation Learning Algorithm (BP)

Introduction

Financial decision is the most important part of financial management. The purpose of financial and accounting information preparation is to provide a basis for economic and financial decision as well. Naturally, the more the information about the issue, the more the accurate results will derive from. It has also to be taken in to account that the bankruptcy is considered as an important concept in financial management area. Predicting that the activity of economic units has a continuation in the future periods is what makes it one of the important elements in investment decision making. Investors and creditors are interested in predicting the bankruptcy of companies because of the possibility of imposing a lot of costs in a case of bankruptcy. Each prediction model has its own strengths and weaknesses (Adnan

& Humayon, 2006). In last 40 years, the bankruptcy prediction of companies has changed into a principal research issue in financial literature. Furthermore developed countries, also in developing countries a lot of research have done for improve the models of bankruptcy prediction (Rahimpoor et al., 2012).

The prediction models serve as functions to predict the continuation or cessation of business units' activities using the financial ratios in this regard. Among the different methods used to predict the bankruptcy, the ratio analysis is also considered as one of them. The probability of bankruptcy, in this method, is estimated by a group of financial ratios. Hence, the bankruptcy issue is discussed in this research using two bankruptcy prediction models. The bankruptcy prediction model by providing necessary warnings can make the Companies aware of the occurrence of bankruptcy and help investors in identifying investment opportunities. The artificial neural network is a popular method in bankruptcy prediction study which uses the benefits of technology and needs no special requirements for predictor variables. The present study is about using artificial neural network as one of the bankruptcy prediction models and its purpose is to compare this model to that of logistic regression statistical modeling as well.

The first research on bankruptcy prediction was begun in 1900 by Thomas Woodlock. He performed a classic analysis in the railroad industry and presented the results of his research in an article titled"The anatomy of a railroad report and ton-mile cost". Two types of analyses i.e. univariate and multivariate analyses are used to study better the bankruptcy prediction models. Three stages in the development of financial distress measures exist: univariate analysis, multivariate analysis, and logit analysis. Univariate analysis assumes "that a single variable can be used for predictive purposes" (Cook & Nelson, 1998). Beaver (1966) presented empirical evidence that certain financial ratios, most notably cash flow/total debt, gave statistically significant signals well before actual business failure. However, this approach had limitations in practice encouraging later researchers to overcome the facing problems.

The best-known, and most-widely used, multiple discriminant analysis method, is the one proposed by Edward Altman, Professor of Finance at the Stern School of Business, New York University. Altman's z-score, or zeta model, combined various measures of profitability or risk. The resulting model was one that demonstrated a company's risk of bankruptcy relative to a standard. Altman's initial study proved his model to be very accurate (Altman 1968).

Since the multivariate discriminant analysis approach has certain limitations, Ohlson (1980) applied an alternative statistical method i.e. logit analysis (generalized linear model with a logit link function) in predicting corporate failure. He considered the discontinuity point equal to 5.0 in order to distinguish the bankrupt companies from non-bankrupt ones. He could prove the accuracy of logit classification model by relying on his observation from 105 bankrupt firms and 2058 non-bankrupt firms which have been studied using this point of discontinuity and three different periods (one, two and three years before bankruptcy) during 1970 to 1976.

Odom and Sharda (1990) were the first researchers to use neural networks for bankruptcy classification in non-experimental studies and found that neural networks were at least as accurate as discriminant analysis. They state that the neural network approach has significant advantages over other prediction methods and is able to analyze the complicated plans much more efficiently than statistical-based ones. It has also no need to restrictive statistical assumptions. This allows the neural network model to provide a higher level of accuracy.

The predictive ability of neural network models and multiple discriminant analysis were compared by Charalambous et al (2000) in an experimental study. They used a secondary experimental research including 139 matched-pairs of bankrupt and non-bankrupt firms over the period 1983-1994. In study conducted by these researchers it seemed that there was a gap in comparing the neural network methods and those of traditional. They used different algorithms for training neural networks. Based on results derived from research, they arrived at the conclusion that the new model of neural networks provides better results in prediction model compared to back propagation learning model.

To predict the bankruptcy of manufacturing companies accepted in Tehran Stock Exchange Market, the following assumptions can be made using artificial neural network in this regard:

- The prediction-based neural network models increase the ability of financial management to deal with failures.

The artificial neural network models have greater accuracy than the logistic regression statistical ones in predicting the bankruptcy.

The Research Methods

The present study is a quasi-experimental research based on post-event approach using the historical data to derive better results from. The population studied in this research has consisted of manufacturing firms listed in Tehran Stock Exchange during the period from 2002 to 2011 (the nine-year period has been considered as best one because of data limitations, time interval between previous executive researches in Iran and the mean change of financial ratios over the very long-term period). Accordingly, a list of companies has been prepared to study including 54 bankrupt ones which were subject to Article 141 of the commercial code. It also should be noted that using the information related to one and two years before bankruptcy for each company required them to be studied during the periods from 2000 to 2011. After selecting 54 bankrupt companies, it needed to consider a number of non-bankrupt ones as a second group. To do that they were selected using elimination method (screening) and following criteria:

1. In order to homogenize the statistical sample during the research period, They should have been accepted in Tehran stock exchange before 2000 and not have been resigned from stock exchange by the end of 2011.
2. The financial year should be ended in March to increase the comparability of them.
3. Their financial year must remain unchanged during the research period.
4. Not being related to financial or investment companies (like banks and insurance institutions).

As per this regard, 120 companies listed on the Stock Exchange of Tehran were selected to study, according to the above criteria.

The variables used in this study are as follows:

- **Dependent Variable**: the probability of bankruptcy of a firm;
- **Independent Variables**: the financial ratios, like:

1. Current ratio, shows the proportion of current assets of a business in relation to its current liabilities.

2. Quick ratio, is the ratio of the sum of cash and cash equivalents, marketable securities and accounts receivable to the current liabilities of a business. It measures the ability of a company to use its near cash or quick assets to extinguish or retire its current liabilities immediately.

3. The return-on-asset ratio (ROA), which is the ratio of net income to total assets, measures a company's effectiveness in deploying its assets to generate profits.

4. The debt ratio, which is defined as the ratio of total debt to total assets.

5. The ratio of working capital to total assets, net working capital is calculated as current assets minus current liabilities.

The SPSS and Matlab soft wares have also been used to analyze data.

Neural Network and Logistic Regression

Neural networks are the nonlinear computational algorithms designed for numerical data processing. They are computational models inspired by animal central nervous systems (in particular the brain) that are capable of machine learning and pattern recognition (Rezaiedolatabadi *et al.*, 2013). A neural network is a collection of neurons that are connected to each other. It consists of a set of highly interconnected entities, called nodes or units. Each collection of these neurons is called a layer. Several characteristics of neural network like the internal dynamics of neural networks in prediction process, data error changes and the advantage of not needing to additional information for input data are what make it attractive to use in many areas. The multilayer Perceptron (MLP) is one of neural network structures. The basic concept of perceptron was introduced by (McCulloch and Pitts, 1943). As mentioned in the research history McCulloch and Pitts (1943) produced the first neural network, which was based on their artificial neuron. A bias input for perceptron has been shown in figure 1:

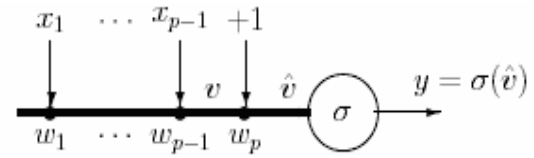


Figure 1. A bias input for perceptron

A multi-layer perceptron shows a nonlinear relationship between input and output vectors. In fact, it is a nonlinear function that maps vector-valued input by means of numerous hidden layers to vector-valued output. This is done by connecting each neuron node i.e. each node in a layer is connected to all the nodes from the previous and next layers. Output neuron is multiplied by the weighting coefficients and given to non-linear activation function as an input. The perceptron is trained (i.e. the weights and threshold values are calculated) based on an iterative training phase involving training data. Training data are composed of a list of input values and their associated desired output values. In the training phase, the inputs and related outputs of the training data are repeatedly submitted to the perceptron. Normally, when a neural network is trained, the network weights are adjusted to minimize the error between the predicted output and the target output vectors or that the number of training repetition reaches a predetermined maximum value. Then, a set of non-experimental inputs is applied to network in order to assess the training accuracy. These entries must be different from those used for training network. Neural network training is a very time consuming and complicated task having an optimization problem with a large number of variables and uncertain parameters (Kaviani et al., 2007).

A variety of learning algorithms have been used for neural network training in which back-propagation algorithm is one of the most important ones. During learning, for each input, the desired target output is computed with the actual output of the network and the disparity between the two is computed. In this algorithm the computed output value is

compared to actual one and the weights and biases of the network are modified based on errors derived from, so that at the end of each iteration, the size of the resulting error is less than the amount obtained in the previous iteration (Rumelhart, Hinton and Williams, 1986).

In order to produce the output (y) in a mathematical model of artificial neurons, the calculations are to be performed based on the following relationship:

$$y_j = \sum_{i=1}^{n} f\left(w_{ij} x_i\right) + b_j \quad (1)$$

In the above equation, W_{ij} is the weight coefficient of neuron *i* which is connected to neuron *j* (similar to synaptic weights in normal neurons). Where *n* is a number of inputs for each neuron of the network and b_j is a bias vector for neuron *j*. The bias vector accelerates the learning process by adding a fixed amount to the sum of the product of the weights of input vectors, where *f* expresses the activation function in this equation. The neural network used in this study is a multilayer perceptron which has been trained as an error back-propagation algorithm including a three-layer feedforward neural network with a neuron arrangement of (5:15:2).

Logistic regression (LR) is a widely used multivariable method for modeling dichotomous outcomes. It is a multiple discriminant analysis which is capable of addressing all of the predictive factors simultaneously. Another concept which is used in logistic regression model is odds ratio. The term "odds" is defined differently according to the situation under discussion, but it is the ratio of the probability of occurrence of an event (P_i) to that of non-occurrence ($1\text{-}P_i$). The logarithm of this ratio is calculated by the following expression. This model is also known as the Logit model (Mo'meni and Fa'al Qayyomi, 2007):

$$Ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_i x_i \quad (2)$$

The LR model has been designed for prediction purpose using logistic regression, SPSS software and annual ratios information of t, t-1 and t-2 related to 120 companies in this regard. The extraction function is given in the next section.

Results and Discussion

The neural network training process, as described before, is implemented in MATLAB environment. The company's bankruptcy and health status, in the other words, is predicted based on three years of financial ratios. It is worthy to mention that, 84 companies among 120 companies have been selected in this study to be trained as neural network.

Table 1
Prediction results by ANN for year t

t	Bankrupt companies	Non-bankrupt companies	
Bankruptcy prediction	51	4	92.7%
Healthy prediction	3	62	95.7%
Accuracy	93.9%	94.4%	94.2%

Table 2
Prediction results by ANN for year t-1

t-1	Bankrupt companies	Non-bankrupt companies	
Bankruptcy prediction	49	6	89.1%
Healthy prediction	5	60	92.3%
Accuracy	90.7%	90.9%	90.8%

Table 3. Prediction results by ANN for year t-2

t-2	Bankrupt companies	Non-bankrupt companies	
Bankruptcy prediction	45	9	83.3%
Healthy prediction	9	57	86.4%
Accuracy	83.3%	86.4%	85%

As can be seen in tables 1 to 3, the neural network model is able to predict the company's bankruptcy or health status much more accurately than the other methods so that the prediction process can be done in base year with the accuracy equal to 94.2 percent. This shows that the neural network model has placed 51 companies in correct class from the number of 54 bankrupt ones i.e. only three of them have been improperly placed. From 66 non-bankrupt companies, 62 of them have placed in correct class as well. Similarly, there were 90.8 and 85 percent accuracy of prediction according to data derived from one and two years ago, respectively.

The fit logistic regression model has been used in this section to predict the failure of firms in each year of the study. Consequently, the results of model fitting can indicate the proper selection of these models during the years of study. The estimation results of coefficients, estimation error and the statistical significance of indicators in the model have been presented in table 4:

Table 4
The estimated coefficients of the regression models

time	Variable	Coefficient	Estimated Error	Wald	Sig.
Year t-1	V_1	-4.198	1.709	6.037	.014
	V_2	-11.176	4.254	6.901	.009
	V_3	-9.249	2.049	20.379	.000
	V_4	16.353	4.652	12.358	.000
	V_5	13.088	3.194	16.792	.000
Year t-2	V_1	-4.809	1.935	6.177	.013
	V_2	-15.218	6.005	6.422	.011
	V_3	-7.820	2.126	13.530	.000
	V_4	15.967	4.951	10.402	.001
	V_5	12.874	3.127	16.952	.000
Year t	V_1	-4.081	1.302	6.125	.021
	V_2	-13.116	5.953	6.376	.008
	V_3	-8.321	3.257	11.956	.001
	V_4	14.873	3.267	11.324	.011
	V_5	12.021	3.211	15.934	.000

As it can be seen from the regression results, there is a significant relationship between all variables and bankruptcy probability so that all variables show a significant impact on this level of error. The regression models can be symbolized, during these two years, as follows:

The model of regression for year t:

$$\log\left(\frac{p}{1-p}\right) = -4.081V_1 - 13.116V_2 - 8.321V_3 + 14.873V_4 + 12.021V_5 + \varepsilon \quad (3)$$

The model of regression for year t-1:

$$\log\left(\frac{p}{1-p}\right) = -4.198V_1 - 11.176V_2 - 9.249V_3 + 16.353V_4 + 13.088V_5 + \varepsilon \quad (4)$$

The model of regression for year t-2:

$$\log\left(\frac{p}{1-p}\right) = -4.809V_1 - 15.218V_2 - 7.820V_3 + 15.967V_4 + 12.874V_5 + \varepsilon \quad (5)$$

Where p is equal to the probability of enterprises' bankruptcy and ε is stochastic or random disturbance, according to the above models.

Table 5 shows the ability of the regression models to explain and correctly predict the bankruptcy of companies.

Table 5
Prediction results by LR

year	Observations	Prediction		Accuracy
		Non-bankrupt	Bankrupt	
t-1	Non-bankrupt	58	8	87.9%
	Bankrupt	5	49	90.7%
	The percentage of all correct prediction			89.2%
t-2	Non-bankrupt	53	13	80.3%
	Bankrupt	8	46	85.2%
	The percentage of all correct prediction			82.5%
t	Non-bankrupt	61	5	92.4%
	Bankrupt	6	48	88.9%
	The percentage of all correct prediction			90.8%

Conclusion

Financial managers need different efficient models to predict, during decision making, the condition especially when it seems critical. One of these models is Bankruptcy prediction model which is one of the most important issues in financial decision-making. This model can be used to increase the financial management capacity for dealing with the bankruptcy problems.

It can be said from the results of tables that the neural network method has higher accuracy than that of logistic regression. For example, based on logistic regression model, the probability values of bankruptcy prediction are 87.9 and 90.7 percent respectively for non-bankrupt and bankrupt companies in one year before bankruptcy (and totally 89.2 percent for both of them) while these values are more accurate in neural network model. In this model

compared to that of regression, the bankrupt and non-bankrupt companies have been predicted 90.9 and 90.7 percent respectively which shows that each company has placed in correct class with high accuracy.

It is noteworthy that, according to the results predicted by both methods, especially neural network, Type I Error is less than Type II Error, in other words, the designed model has distinguish the bankrupt companies from those of non-bankrupt with high accuracy in this respect. The necessity to reduce the type I error is more than type II error depending on the type of study and the need to recognize the bankrupt companies before the bankruptcy began. Note that a type I error occurs when the model places a company in a bankrupt group when it is really non-bankrupt one and a type II error occurs when the model places a company in a non-bankrupt group when it is really bankrupt one.

References

Altman, E. (1968). Financial ratios, Discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance. 4 (September), pp. 54-89.

Aziz, M., Adnan, and Humayon, A. D. (2006). Predicting corporate bankruptcy: where we stand?, Corporate governance, *6*(1), pp. 18-33.

Beaver, W. H. (1966). Financial ratios as predictors of bankruptcy. Journal of Accounting Research 4, pp. 71-111.

Charalambous, C., Charitou, A., & Kaourou, F. (2000). Comparative analysis of artificial neural network models: Application in bankruptcy prediction. Annals of operations research, 99(1-4), pp. 403-425.

Cook, R. A., & Nelson, J. L. (1998) A Conspectus of Business Failure Forecasting. Small Business Institute Director's Association. Available at http://sbaer.uca.edu/research/sbida/1988

Kaviani, K. A., Kani, P. M. S. A., Ardakani, J. A. (2007). The Multilayer Neural Network Training by PSO Algorithm. First Joint Congress on Fuzzy and Intelligent Systems Ferdowsi University of Mashhad, Iran, 29-31 Aug.

Mo'meni, M., and Ghayyomi, F. M. (2007). Statistical analysis using SPSS. Ketab e No Press. Iran, Tehran.

Odom, M., and Sharda, R. (1990). A neural network model for bankruptcy prediction, Proceedings of the IEEE international conference on neural networks, pp. 163–168.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 18(1).

Rahimpoor, A., Ansari, A., and Sarokolaie, A. M. (2012). A Comparative Study of Bankruptcy Prediction Models of Fulmer and Springate in Accepted Companies in Tehran Stock Exchange. Caspian Journal of Applied Sciences Research, 1(13), pp. 76-81.

Rumelhart, D. E., Hinton, E., Williams, J. (1986). Learning internal representation by error Propagation. Parallel Distributed Processing 1, pp. 318-362.

Woodlock, T. F. (1900). The anatomy of a railroad report and ton-mile cost (Vol. 2). SA Nelson.