http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (23): 7684-7690, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.7684.7690

Enterprise Financial Distress Prediction Based on BPNN: A Case Study of Chinese Listed Companies

[1]Ying Zhang, [2]Chong Wu and [2]Xin-ying Zhang
[1]Department of Admission and Employment,Harbin University of Science and Technology, Harbin, 150000, China
[2]School of Economics and Management, Harbin Institute of Technology, Harbin, 150000, China

Abstract: Enterprise financial distress prediction has been the attention focus in the theory study and the business community. To build a scientific, fast and effective model for financial crisis prediction of the Chinese listed companies, 11 key financial indicators are chosen for the financial distress prediction model. The factor analysis is used to extract five common factors and therefore, to get the comprehensive score of each sample. The traditional ST and non-ST classification criteria are abandoned; the score intervals of the enterprise financial status are divided in a novel way-health, concern and distress. Finally, the five common factors are trained and tested as the input and the financial status as the output with the prediction model based on the backpropagation neural network. The result shows that the proposed model is accurate and can provide a great assistance for enterprises, investors and decision-makers.

Key words: Financial distress prediction, classification method, BPNN, Chinese listed companies

INTRODUCTION

As a key to effectively prevent the enterprise failure, enterprise financial distress prediction has been the focus of the attention in the theory study and the business community. The enterprise financial distress refers to the comprehensive precarious or emergent financial state caused by severe external setbacks or internal control failures of financial activities. The financial distress prediction model is designed to provide the warning of the financial states or the enterprise poor management strategies. Nowadays, enterprise failure identification and early warnings of impending financial crises have become increasingly significant and necessary for analysts, practitioners as well as regulators who are interested in China's capital market because of the following factors.

First of all, the progress toward the bankruptcy process for the listed enterprises has aroused great attention among Chinese market participants since the enforcement of the delisting system in China's stock market started in 2001; the concerns with enterprise failures and their impacts to the society are undoubtedly sincere. However, due to the fact that some Chinese listed companies are state-owned, their stock price can't reflect the actual financial distress. Moreover, Chinese government encourages banks to develop their own distress prediction models. Secondly, as a useful tool to support enterprise internal control, a good predication model enables management to identify potential financial problems early enough to prevent critical situations. Thirdly, creditors, fund managers and stockholders can use the prediction model to screen out undesirable investments, or reduce losses by withdrawing investment or collect receivables from those unhealthy firms. Fourthly, the government and the market authorities can use the predication model as a guideline to increase the transparency of regulatory objectives. Also it can be used as a barometer to measure the economic condition of firms in different sectors.

Consequently, a sound and reliable model to predict the financial distress of the Chinese listed companies is extremely important for the Chinese market, the enterprises, investors and researchers as well.

LITERATURE REVIEW

Literature on enterprise financial crisis prediction: The literature on enterprise financial crisis prediction is very rich. Experts and scholars have done a large number of quantitative researches of enterprise financial crisis since the 1930s (FitzPatrick, 1932; Gerke *et al*., 2012). The study and the investigation of the financial crisis prediction methods develop from the univariate analysis to the multivariate prediction, from the traditional statistical

Corresponding Author: Ying Zhang, Department of Admission and Employment, Harbin University of Science and Technology, Harbin, 150000, China

methods to the statistical analysis based on the artificial intelligence (Zhang *et al.,* 2010). Especially from the 1980s, along with the rapid development of the computer technology, the advantages of artificial intelligence technology has become prominent and the artificial intelligence methods based on data mining have been successfully applied to the financial crisis prediction which greatly broadens the general theory of financial crisis prediction(Thorne and Porter, 2012).

Literature on back propagation neural network (BPNN): Artificial Neural Networks (ANNs) are a brain-like intelligent information processing system intending to mimic the human brain structure and function (Huang, 2012). Artificial intelligence prediction models are powerful tools for nonlinearities predictions. These mathematical models comprise individual processing units called neurons that resemble neural activities. Each processing unit sums weighted inputs and then applies a linear or nonlinear function to determine the output. The neurons are arranged in layers and are combined through excessive connectivity.

BPNN is one of the most effectively applied artificial neural network models. Its architecture is a Multilayer Perception (MLP) and it uses a learning algorithm known as Error Back Propagation (EBP) (Pant and Wadhwani, 2011). The MLP has a hidden layer between the input and output layers; the hidden layer can be one or more levels. BPNN has the incomparable superiorities as follows:

- **Nonlinear programming capacity:** Depending on the nature of the application and the strength of the internal data patterns, BPNN is trained quite well. This applies to problems where the relationships may be quite dynamic or non-linear. BPNN provides an analytical alternative to conventional techniques which are often limited by strict assumptions of normality, linearity, variable independence etc. Because a BPNN can capture many kinds of relationships, it allows the researchers to quickly and relatively easily model phenomena which otherwise may have been very difficult or impossible to explain otherwise
- **Generalization ability:** Once a BPNN is trained to a satisfactory level, it may be used as an analytical tool on other data and no special training is needed again. New inputs are filtered into the input pattern and are processed by the middle layers as though training is taking place. However, at this point, the output is retained. The output is the predicted model for the data which can then be used for further analysis and interpretation
- **Error tolerance:** BPNN is universal approximation and it works best if the system has a high tolerance to errors.The adjustment process of the weight matrix is to extract the statistical properties form a large number of samples. And the knowledge that reflects the laws comes from all the samples. Therefore, the error cannot affect the adjustment of the weight matrix. So BPNN allows the sample with large errors and even individual errors (Wu *et al*., 2011)

With high-speed computing power, fast recall speed, high learning accuracy and error tolerance, BPNN works well for capturing associations or discovering regularities within a set of patterns; where the volume, number of variables or diversity of the data is very great; the relationships between variables are vaguely understood; or, the relationships are difficult to describe adequately with conventional approaches. Accordingly, BPNN fits for all the requirements of the issue of financial distress prediction of the Chinese listed companies.

Literature on criterion of financial status: In 1998, China's Securities Regulatory Commission (CSRC) began to differentiate those firms with financial difficulties as Special Treatment (ST) firms. These “ST” firms include: (1) Companies that have negative cumulative earnings over two consecutive years or its Net Asset Value (NAV) per share is below its par value (book value); (2) Companies that have a negative earning for one year, but in the current year, the shareholders’ equities are below its registered capital; and (3) Companies that receive the auditors’ “going concern opinions”.

Several researchers studied the corporate financial distress classification problem. Sun and Li (2008) predicted the Chinese listed companies’ financial distress and assorted the tested companies as normal and distressed ones (Sun and Li, 2008). Zhang *et al.* (2009) constructed a forewarning model of the listed companies against financial crisis based on Bayesian discrimination (Zhang *et al*., 2009). Kumar and Anand (2013) assessed the financial health of Indian firms using Altman’s original and revised z-score models (Kumar and Anand, 2013). Moradi *et al.* (2013) predicted the financial distress of Iranian companies (Moradi *et al*., 2013). Many researchers applied various methods to predict the financial distress in different countries. But their criterion to describe the enterprise financial status is the same, i.e. all the enterprises fall into either distressed group or healthy (normal) group.

Summary of the literature review: The work above forms a base to support further studies in such field in China.

However, it can be found through the literature study and the comprehensive review of the financial distress prediction methods that the previous studies are based on ST and non-ST classification. According to this "one or the other" category, the predicted ST enterprises with a foregone financial distress don't have enough adjustment time to avoid a crisis situation which reduces the practical significance of the model.

Additionally, there is a lack of a solid theoretical basis or a deep understanding of what factors really affect the enterprise finance. Most researchers select variables based on the prior studies that conducted in the western countries. However, the situations of Chinese listed companies are quite different, especially when it comes to the accounting rules, the quality of data, due diligence and the equity structure. So their models might not truly reflect the actual situation of the Chinese companies.

Based on the analysis and summaries of home and abroad theoretical and empirical studies of enterprise financial distress prediction, the authors construct the models based on BPNN and conduct the empirical research on the prediction model. In this paper, a more realistic and feasible classification method is proposed to classify the financial statuses of enterprises into three categories: Health, concern and distress. Therefore, the predicted financial concerned enterprises can pay attention to internal improvements in advance to conduct accurate and reasonable adjustments and avoid the potential financial predicament.

METHODOLOGIES

Data collection: As listed companies are the interest focus of the society, the financial data of the listed companies are comparable, open and normal. Therefore, it's feasible to select the listed companies as a research object. In this study, the listed companies of A Share in Shanghai and Shenzhen Stock Exchange are selected as the main sources of data. Based on the industry representation and the asset size, the authors select 100 valid samples in the latest annual report of 2009 and 2010 from the authoritative security websites like Sohu Security (Available at http://stock.sohu.com/gegufengyun/).

Based on the definition of financial distress and the previous researches, this study establishes an objective and rational index system for the financial distress prediction model. In order to simplify the network input and reduce information redundancy, the authors process the indicators using the Principal Component Factor Analysis (PCA) to reduce the BPNN input dimensions. At the same time, the authors derive the composite score of the enterprises' finance status using Factor Analysis (FA) and check the distribution of ST enterprises to determine their initial financial status. Then the authors analyze the listed company's highly timely data in 2009 and 2010. The authors take the common factors of each enterprise gotten from FA as the input and the financial status of each sample as the output. The authors train the samples according to the neural network training algorithm using Matlab to determine the network parameters and construct the model and the finally test the model.

According to the China's disclosure system of listed companies' annual reports, the deadline for the listed companies to publish their annual report is on April 30 of the next year. Therefore, a listed company's annual report of the year (t-1) and the ST decision in the year t happen almost simultaneously. That means a company's annual report information of the year (t-1) can decide whether the company will be in the ST due to "the financial anomalies". Therefore, it's not practical to predict whether the company will be in the ST in the year t by using the information of year (t-1). In order to avoid this problem, this study uses a listed company's annual financial report of year (t-2) to its financial status in year t. Thus, the authors can predict the general financial status of the next year at the end of the previous year which will help the enterprise to improve the business-related measures to avoid the financial problems the next year. This method plays a vital role in the early prevention and can help managers and supervisory staff with the scientific and effective decision-making.

Construction of prediction index system: The construction of this index system is based on the principles of systematical design, scientific feasibility, expansibility, independence and objectivity. Meanwhile, learning the financial crisis evaluation system from domestic and foreign scholars and taking into account the data availability and the special financial situations of Chinese enterprises, the authors select the following 11 indicators as the index system of financial distress prediction model for a listed company, namely, Current ratio, Quick ratio, Asset-liability ratio, Profit rate to net worth(average weighed), Net profit margin, Earning per share, Return on total assets ratio, Inventory turnover, Total asset turnover, Fixed asset turnover and Accounts receivable turnover.

EMPIRICAL STUDY

Data process and analysis: Based on the industry representation and the asset size, the authors select 100 valid samples in the data of 2009 and 2010 and use SPSS18.0 to standardize the raw data. As the correlation of the indicators between the variables increases the probability of the information overlap, there is unnecessary trouble in the statistical analysis. FA is very effective to solve this problem (Lu *et al*., 2011). Under the premise of the least lost information, FA minimizes the original data base to several types of composite variables. These variables in this basic structure are named common factors. In order to get the common factors among the indicators, the authors process the samples using PCA for the factor analysis.

The authors process the samples in KMO and Bartlett's Test using SPSS18.0. KMO value which is 0.571, greater than 0.5, is suitable for PCA; the accompanied probability of the Bartlett's Test which is 0.000, less than the significant level of 0.05, rejects the null hypothesis of the Bartlett's Test. That means the sample data are suitable for PCA. Table 1 shows the KMO and Bartlett's Test.

Selection and analysis of common factors: In the total variance table of PCA (Table 2), there are five factors whose eigenvalues are greater than 1, namely, the common factors. The cumulative variance contribution rate of their eigenvalues reaches 79.121%. By observing the gravel figure (Fig. 1) for the scree test, the two lines converge at the fifth factor and the scree line has a sudden rise at the fifth factor. Therefore, it is reasonable to reserve these five factors.

The following conclusions are obtained from the rotated component matrix in SPSS (Table 3). Different indicators are chosen to explain the five different factors.

Based on the coefficient matrix of the component score out of obtained from SPSS (Table 4), the scores of the five common factors are calculated as follows:

$$F_1 = 0.346\times X_1 + 0.276\times X_2 - 0.069\times X_3 - 0.079\times X_4 - 0.043\times X_6 \\ +0.278\times X_7 + 0.353\times X_8 - 0.054\times X_9 - 0.043\times X_{10} + 0.013\times X_{11} \quad (1)$$

$$F_2 = -0.057\times X_1 - 0.033\times X_2 + 0.388\times X_3 + 0.399\times X_4 - 0.005\times X_5 \\ -0.305\times X_6 - 0.046\times X_7 + 0.005\times X_8 + 0.058\times X_9 + 0.013\times X_{10} - 0.037\times X_{11} \quad (2)$$

$$F_3 = -0.025\times X_1 + 0.147\times X_2 + 0.016\times X_3 + 0.070\times X_4 - 0.051\times X_5 \\ +0.092\times X_6 - 0.215\times X_7 + 0.035\times X_8 + 0.519\times X_9 - 0.045\times X_{10} + 0.603\times X_{11} \quad (3)$$

$$F_4 = 0.014\times X_1 - 0.182\times X_2 - 0.030\times X_3 + 0.840\times X_5 - 0.041\times X_6 \\ +0.113\times X_7 + 0.024\times X_8 + 0.338\times X_9 + 0.092\times X_{10} - 0.209\times X_{11} \quad (4)$$

$$F_5 = -0.066\times X_1 + 0.125\times X_2 + 0.130\times X_3 + 0.026\times X_4 + 0.104\times X_5 \\ +0.158\times X_6 - 0.138\times X_7 + 0.004\times X_8 - 0.276\times X_9 + 0.864\times X_{10} + 0.113\times X_{11} \quad (5)$$

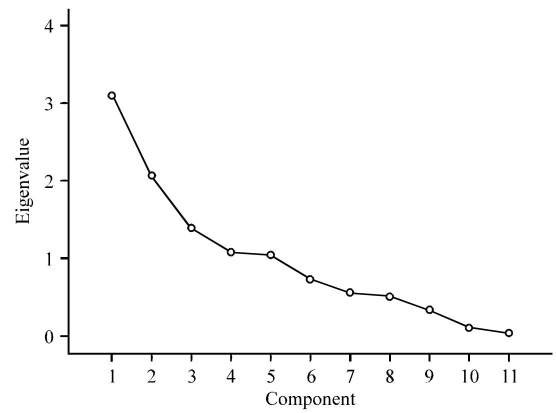

Fig. 1: Gravel figure of the principle component analysis

Table 1: KMO and Bartlett's Test

Kaiser-Meyer-Olkin measure of sampling adequacy 0.571		
Bartlett's test of sphericity approx	Chi-Square	1139.152
	df	55
	Sig.	0.000

Table 2: Total variance of PCA

	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
Component	Total	Variance (%)	Cumulative variance (%)	Total	Variance (%)	Cumulative variance (%)	Total	Variance (%)	Cumulative variance (%)
1	3.105	28.224	28.224	3.105	28.224	28.224	2.570	23.368	23.368
2	2.073	18.846	47.069	2.073	18.846	47.069	2.532	23.023	46.391
3	1.396	12.689	59.758	1.396	12.689	59.758	1.411	12.825	59.216
4	1.083	9.844	69.602	1.083	9.844	69.602	1.095	9.954	69.169
5	1.047	9.519	79.121	1.047	9.519	79.121	1.095	9.952	79.121
6	0.736	6.695	85.816						
7	0.566	5.145	90.961						
8	0.515	4.686	95.648						
9	0.337	3.062	98.710						
10	0.105	.953	99.664						
11	0.037	.336	100.000						

Table 3: Rotated component matrix

	Indicator	Z-score					
Factor	Name	Code	Component 1	Component 2	Component 3	Component 4	Component 5
F1	Return on total assets ratio	X9	0.919	0.201	0.108	0.053	0.071
	Profit rate to net worth (average weighed)	X1	0.844	0.041	0.018	0.04	-0.012
	Earning per share	X2	0.725	0.071	0.248	-0.165	0.198
	Net profit margin	X6	0.638	0.033	-0.259	0.126	-0.119
F2	Quick ratio	X4	0.027	0.966	0.072	-0.022	0.025
	Current ratio	X3	0.055	0.948	0.005	-0.057	0.136
	Asset-liability ratio	X5	-0.235	-0.796	0.146	-0.022	0.17
F3	Total asset turnover	X11	0.117	-0.099	0.847	-0.173	0.178
	Fixed asset turnover	X8	-0.053	0.074	0.725	0.405	-0.263
F4	Accounts receivable turnover	X7	0.065	-0.057	0.01	0.916	0.117
F5	Inventory turnover	X10	0.051	0.015	0.013	0.101	0.935

Table 4: Component score coefficient matrix

Z-score component	1	2	3	4	5
Profit rate to net worth (average weighed)	0.346	-0.057	-0.025	0.014	-0.066
Earning per share (Yuan)	0.276	-0.033	0.147	-0.182	0.125
Current ratio (Time)	-0.069	0.388	0.016	-0.030	0.130
Quick ratio (Time)	-0.079	0.399	0.070	0.000	0.026
Accounts receivable turnover (Times)	0.000	-0.005	-0.051	0.840	0.104
Asset-liability ratio (%)	-0.043	-0.305	0.092	-0.041	0.158
Net profit margin (%)	0.278	-0.046	-0.215	0.113	-0.138
Return on total assets ratio (%)	0.353	0.005	0.035	0.024	0.004
Inventory turnover	-0.054	0.058	0.519	0.338	-0.276
Fixed asset turnover	-0.043	0.013	-0.045	0.092	0.864
Total asset turnover	0.013	-0.037	0.603	-0.209	0.113

Table 5: Financial condition classification of enterprises

Intervals of F value	Financial condition	No. of enterprises	ST enterprises	Proportion of ST enterprises (%)
(-∞,-0.2)	Distress	19	16	84.2
(-0.2,0.13)	Concern	54	16	29.6
(0.13,+∞)	Health	27	1	3.7

The composite score F, that is, the sum of each factor and its contribution that was got out of total variance are:

$$F = F_1 \times 0.28224 + F_2 \times 0.18846 + F_3 \times 0.12689 + F_4 \times 0.09844 + F_5 \times 0.09519 \tag{6}$$

Specific classification of financial status: By analyzing the F value of the enterprises and taking the ST or non-ST status as the reference, the authors divide the intervals of F value and determine the final classification as shown in Table 5.

As can be seen in Table 5, when F is greater than 0.13, the very small proportion of ST enterprises can be ignored. Therefore, it can be regarded as a healthy status interval; when F is between -0.2 and 0.13, the proportion of ST enterprises is much smaller than the proportion of distressed enterprises but cannot be ignored. Therefore, it can be regarded as a concerned status interval; when F is smaller than -0.2, ST enterprises account for a large part. Therefore, it can be regarded as a distressed status interval. In summary, based on the analysis above, the financial status of the enterprises can be divided into three categories, i.e., distress, concern and health. Through the above analysis and processing, the financial status of the sample enterprises can be seen in Appendix A where the input is the common factor of each enterprise obtained through FA.

Empirical analysis: Since five common factors represent the 11 original data indicators, the number of the input nodes of BPNN is five. In this study, the enterprise financial status can be divided into three situations: health, concern and distress which can be represented by the available state values (1,0,0), (0,1,0) and (0,0,1). So the number of the output nodes of this network is 3. Referring to Formula (7) (Qi and Kang, 1998), the number of the hidden nodes is 7. In training, the learning rate is 0.05, the required training accuracy is 0.0001 and the training times are 4000.

$$n = n_i + 0.618 \times (n_i - n_o) \tag{7}$$

where, n_i is the number of the input nodes, n_o the number of the output nodes, n the number of the hidden nodes.
85 enterprises are selected as the training samples, 15 as the testing samples. Matlab7.11 is used and the programs

Table 6: Prediction analysis

Financial condition	No. of enterprises	Correctly predicted enterprises	Accuracy
A	3	3	1
B	10	9	0.9
C	2	2	1

Table 7: Prediction results of the enterprise financial distress based on SVM

Financial condition	No. of enterprises	Correctly predicted enterprises	Accuracy
A	3	3	1
B	10	8	0.8
C	2	2	1
Total	15	13	0.87

are in Appendix B. The prediction results of the enterprise financial distress based on BPNN are shown in Table 6.

In order to better test the accuracy of the BPNN model, the same samples are used in the well-trained Support Vector Machine (SVM) model and the results are shown in Table 7.

The results of BPNN exhibit a high accuracy rate at 93%. The training results show that the error ratio of BPNN is relatively smaller than that of SVM and the results are more satisfactory which means that BPNN is very appropriate to predict the financial distresses. Enterprises may easily identify their financial status by using this model.

CONCLUSIONS

In this study, a financial distress prediction model of the Chinese listed companies is built based on BPNN and an empirical study is conducted. The main features of this study are as follows:

- Innovation in the classification of enterprise financial status, that is, to divide the financial status into three categories: health, concern and distress. This can help the enterprises to take the preventive measures as early as possible, to make more scientific and rational decision-making and to avoid the financial crisis
- Simplification of the index information. To reduce the overlapped and redundant data, PCA is used to identify the common factors that can maximally reflect the information of the index system. The selected five common factors are the irrelevant linear combinations of the original 11 indicators
- Practicality of the training model. As the indicator variables chosen for this study can be obtained from public audited annual report of the listed companies, the method is quite feasible. Meanwhile, as the annual report and ST decisions of the enterprises are released almost at the same time, this research uses a listed company's annual financial report of year (t-2) to predict its financial status which can increase the scientific nature and the practicality of the training model

Future research may focus on the application of the proposed model in a specific industry. Moreover, in order to improve the prediction accuracy, a random selection method can be used to determine the training samples and the test samples.

ACKNOWLEDGMENTS

This study is funded by the National Natural Science Foundations of China (71271070 and 91024028).

REFERENCES

FitzPatrick, P.J., 1932. A comparison of the ratios of successful industrial enterprises with those of failed companies. Certified Public. Account., 10: 598-605.

Gerke, R., F. Hammermann and V. Lewis, 2012. Robust monetary policy in a model with financial distress. J. Macroecon., 34: 318-325.

Huang, H.C., 2012. Artificial intelligenc technology in travel agency operating revenue forecasts. J. Theor. Applied Inform. Technol., 4: 270-280.

Kumar, M. and M. Anand, 2013. Assessing financial health of a firm using altman's original and revised Z-score models, A Case of Kingfisher Airlines Ltd (India). J. Applied Manage. Investments, 2: 36-48.

Lu, W.Z., H.D. He and L.Y. Dong, 2011. Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis. Build. Environ., 46: 577-583.

Moradi, M., M. Salehi, M.E. Ghorgani and H.S. Yazdi, 2013. Financial distress prediction of Iranian companies using data mining techniques. Organizacija, 46: 20-27.

Pant, S. and A.K. Wadhwani, 2011. A real time GRS detection system used error back propagation neural network. Int. J. Elec. Elctron. Comput. Eng., 1: 48-54.

Qi, D.H. and J.C. Kang, 1998. Architecture of BPNN. Comput. Eng. Des., 19: 48-54.

Sun, J. and H. Li, 2008. Listed companies' financial distress prediction based on weighted majority voting combination of multiple classifiers. Expert Syst. Appl., 35: 818-827.

Thorne, D. and K. Porter, 2012. Debtors' assessments of bankruptcy financial education. Consu. Knowl. Financ. Decisions Int. Ser. Consu. Sci., 2: 197-214.

Wu, S.X., D.L. Luo, Z.W. Zhou, J.H. Cai and Y.X. Shi, 2011. A kind of BP neural network algorithm based on grey interval. Int. J. Syst. Sci., 42: 389-396.

Zhang, L., E.I. Altman and J. Yen, 2010. Corporate financial distress diagnosis model and application in credit rating for listing firms in China. Front. Comput. Sci., 40: 220-236.

Zhang, L.J., Y. Wang and J.H. Liu, 2009. Forewarning model of listed companies against financial crisis based on Bayesian discrimination. Commer. Res., 04: 112-114.