

# Choosing Credit Risk Index for Listed Companies Based on Grey Clustering Method

Lei Hua<sup>1</sup>, Zhigang Song<sup>2</sup>, and Ying Li<sup>3</sup>

<sup>1</sup> Wuhan University of Technology/School of Economics, Wuhan, P.R.China

Email: istphualei@126.com

<sup>2</sup> Henan University of Finance and Economics/Chenggong College, Gongyi, P.R.China

Email: conferencepaper@126.com

<sup>3</sup> Zhongnan University of Economics and Law/Wuhan College, Wuhan, P.R.China

Email: istplying@126.com

**Abstract**—The credit standing of listed companies is a key index that investors usually place high importance on, and thus many methods are developed for analyzing it. However, no matter what method is used, we have to face the problem of choosing proper indexes before analyzing and forecasting. This article adopts grey clustering method for index selection, then uses neural network for forecasting, and finally it compares the forecasting effects, which show that the forecasting ability of neural network after using grey clustering improves noticeably.

**Index Terms**—credit, grey clustering, factor analysis, neural network

## I. INTRODUCTION

Credit risk refers to a loss risk caused by nonperformance of contract, namely, a debtor fails to pay the debt according to schedule and thus this default brings about business risk for the corresponding creditor. Credit risk for listed companies is about those behaviors like default and illegal raising money when listed companies transact with other organizations and individuals and when they borrow money or issue bond for financing from others in order to maintain or expand business<sup>[1]</sup>. Credit risk has become much more important due to scientific progress and volatile economic environment abroad in the past 20 years. In addition, artificial factors are playing more and more key significant role in credit risk. In USA, since the Enron accounting scandal in 2001, the accumulating company scandals have been disintegrating investor confidence in economy. The index of each stock exchange is crumbling which gives rise to credit crisis of investors. Some people even think this is an event bigger than the September 11<sup>th</sup> event. In China, Yinchuan Guangxia event, Lantian share event, Zhengzhou Baiwen event and so on all make people feel surprised about accounting fraud behavior.

At present, the credit analysis and evaluation technology in China are still traditional ratio analysis, banks mainly make credit risk evaluation by calculating loan risk degree. But in the abroad, Altman came up with a multi-variable model named Z-score model based on accounting data and market value in 1968. After that,

logistic model, clustering model, artificial neural network model and many other methodologies were developed. In recent years, some scholars explored the credit standing of domestic enterprises by using the above models, and many of their researches are limited to discussions about model applicability or a little bit improvement<sup>[2][3]</sup>. But few of them focus on how to choose variables and the problems are: (a) without considering industrial differences when choosing variables for different industries; (b) without considering historical differences when choosing variables for a certain industry. Some scholars only choose available variables in view of data availability and some scholars only conduct factor analysis for having the common factors.

This article will study 26 listed companies in IT industry and 40 available variables of each company by applying grey clustering method to them so as to screen variables, and then make forecasting by constructing neural network. Finally, it will compare the forecasting effects.

## II. GREY CLUSTERING METHODOLOGY

Since Professor Deng Julong published his article “The Control Problem of Grey System” that represents the most classical works in the field of Grey System Theory in 1982, this theory has developed widely and deeply, and has been applied to many professional fields successfully. Grey clustering is a method that can divide some observable variables or objects into several definable types according to grey correlation matrix or grey whitening weight function. Each clustering can be recognized as a set of the same observable objects. The nature of this methodology is to substitute known information for unknown and uncertain information completely and reasonably, and to recognize and classify the attribute of grey system, and finally to show an objective and reliable result in a quantitative way. Grey clustering includes grey correlation clustering and grey whitening weight function<sup>[4]</sup>. Grey correlation clustering is mainly used for classifying the same factors in order to simplify the complicated system. This article will use this method for research.



Investment Income/Total Profit, Net Profit/Total Profit, Net Profit after Adjusting the Extraordinary Items, Main Business Income and Net Profit. Profitability has Gross Profit Margin, Profit Rate of Main Business, Net Profit Margin on Sales, Return on Total Assets and Return on Equity. Operating Capacity includes Receivable Turnover Ratio, Inventory Turnover Ratio, Fixed Assets Turnover Ratio, Equity Turnover Ratio and Total Assets Turnover Ratio. Credit Capacity is composed of Current Ratio, Quick Ratio, Interest Cover Ratio and Equity/Fixed Assets. Capital Structure consists of Assets Liability Ratio, Shareholder's Equity, Equity Ratio and Fixed Assets Ratio. Development Capacity includes Main Business Income Growth Ratio, Operating Profit Growth Ratio, Net Profit Growth Ratio, Total Profit Growth Ratio, Net Assets Growth Ratio as well as Total Assets Growth Ratio. And finally Cash Flow includes Net Cash Flow/Net Profit and Cash Received from Sales of Goods/Main Business Income.

### B. Sample Unitization

The aim of unitization is to change sample value of each variable into non-dimensionalized data. It is necessary to determine the variable polarity before unitization. Usually, there are three types of variable polarity, including: benefit-type index which means the bigger the better; cost-type index which means the smaller the better; target-type index which means the closer to the ideal value the better. Their formula for each type is as follows:

#### 1) Benefit-type Index

$$x_{ij}^* = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \quad (1)$$

#### 2) Cost-type Index

$$x_{ij}^* = \frac{\max(x_{ij}) - x_{ij}}{\max(x_{ij}) - \min(x_{ij})} \quad (2)$$

#### 3) Target-type Index

$$x_{ij}^* = \frac{\left| x_{ij} - \frac{1}{n} \sum_{j=1}^n x_{ij} \right|}{\max \left\{ \left( \max_{j=1,2,\dots,n} x_{ij} - \frac{1}{n} \sum_{j=1}^n x_{ij} \right), \left( \frac{1}{n} \sum_{j=1}^n x_{ij} - \min_{j=1,2,\dots,n} x_{ij} \right) \right\}} \quad (3)$$

### C. Grey Correlation Matrix

By calculating the absolute grey correlation degree  $\varepsilon_{ij}$  between  $X_i$  and  $X_j$ , we can have the upper triangular matrix. For instance, this matrix in 2003 is (we hereby omitted the matrixes in 2004 and 2005):

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$\dots$	$x_{37}$	$x_{38}$	$x_{39}$	$x_{40}$
1	0.67	0.99	0.58	0.54	$\dots$	0.95	0.59	0.64	0.60
	1	0.68	0.70	0.60	$\dots$	0.66	0.72	0.87	0.75
		1	0.58	0.54	$\dots$	0.93	0.59	0.64	0.60
			1	0.76	$\dots$	0.58	0.96	0.77	0.90
				1	$\dots$	0.54	0.73	0.64	0.70
					$\ddots$				
						1	0.58	0.63	0.60
							1	0.79	0.94
								1	0.84
									1

### D. Variable Classification

Suppose the critical value  $r = 0.8$ , if  $\varepsilon_{ij} \geq r (i \neq j)$ , and then we take  $X_j$  and  $X_i$  as the same class and adjust variables in a proper way. Hence, we can get the following classification (see Table II):

TABLE II. THE RESULTS OF GREY CLUSTERING

The Three Years before Crisis	The Two Years before Crisis	The Year before Crisis
$X_1/X_3/X_{11}/X_{12}/X_{14}$ $X_{17}/X_{27}/X_{37}$	$X_1/X_2/X_3/X_{18}/X_{28}$ $X_{30}/X_{36}/X_{39}$	$X_1/X_2/X_3/X_{12}/X_{14}/X_{25}$ $X_{34}/X_{35}/X_{36}$
$X_2/X_8/X_{15}/X_{16}/X_{18}$ $X_{19}/X_{21}/X_{22}/X_{23}/X_{39}$	$X_4/X_5/X_7/X_9/X_{13}$ $X_{17}/X_{20}/X_{21}/X_{22}/X_{23}$ $X_{24}/X_{25}/X_{26}/X_{29}/X_{31}$ $X_{33}/X_{37}/X_{40}$	$X_4/X_5/X_6/X_7/X_{13}/X_{20}$ $X_{21}/X_{22}/X_{24}/X_{33}$ $X_{38}/X_{40}$
$X_4/X_6/X_{20}/X_{29}/X_{30}$ $X_{31}/X_{38}/X_{40}$	$X_6/X_{12}/X_{14}/X_{19}/X_{27}$ $X_{34}/X_{35}/X_{36}$	$X_8/X_{10}/X_{11}/X_{15}/X_{16}/X_{17}$ $X_{29}/X_{31}$
$X_5/X_7/X_9/X_{13}/X_{24}/X_{25}$ $X_{26}/X_{34}$	$X_8/X_{10}/X_{11}/X_{15}/X_{16}$ $X_{32}/X_{38}$	$X_9/X_{14}/X_{18}/X_{19}/X_{26}$ $X_{32}$
$X_{10}/X_{32}/X_{35}/X_{36}$		$X_{23}/X_{27}/X_{28}/X_{30}/X_{37}/X_{39}$
$X_{28}/X_{33}$		

In Table II, we choose variables that have the smallest subscript in each class as their typical variables.

### IV. CONCLUSION

After creating a simple three-layer back-propagation neural network, we can use it to study and forecast. The number of neuron at input-layer is equal to the number N of input variables, the number of neuron at hidden-layer is 2N, and the number of neuron at output-layer is 1. We choose some financial data from 18 listed companies at random and use them to train neural network. We use log-sigmoid transfer function at input-layer, use hyperbolic logarithm transfer function at hidden-layer, and use linear transfer function at output-layer. Meanwhile, we use conjugate gradient back-propagation training function based on Fletcher-Reeves algorithm. For the left 8 listed companies, we use their data to forecast and test the forecasting ability of neural network. The results are shown in Table III:

TABLE III. COMPARISON OF FORECASTING RESULTS

	Without Grey Clustering	With Grey Clustering
The Three Years before Crisis	62.5%	75%
The Two Years before Crisis	37.5%	62.5%
The Year before Crisis	75%	75%

The clustering results show that the forecasting ability of neural network after using grey clustering improves noticeably.

In the meantime, when using grey clustering analysis method, some factors should be given more consideration. For instance, first, to choose reasonable and proper indexes; second, to determine the grey whitening function correctly; third, to determine reasonable clustering weight and clustering coefficient; fourth, to set up a correct model so as to ensure the reliability and accuracy of assessment. For the above goals, on one hand, it is necessary to know all related major influencing factors comprehensively and systematically by investigating a large number of process examples, and to form clustering indexes through scientific analysis and summary; on the other hand, it is necessary to set up proper grey whitening function by using correct mathematical tools when processing data, so that the clustering result would be in accordance with reality. This article only consider some of those major factors, if we can taking account of other factors more comprehensively and systematically, then we can have

results that are close to real situations. Besides, grey clustering belongs to multi-factor assessment, it is the result of whitening, but is different from grey type, it is convenient for both scientific analysis and management planning. It is an effective way for assessing reliability. It can avoid calculating comprehensive indexes and can determine the reliability degree directly according to standard. Also, this method is easy to master and clustering coefficients have a strong discreteness. The result has a high-resolution ratio and evaluation method can be realized with the help of computer.

#### REFERENCES

- [1] Qu Yanmei, "Financial credit evaluation index system of public companies," *Commercial Research*, vol. 2, pp. 104-106, 2006.
- [2] Zhang Huijun, Su Ning, Chen Lirong, "Research on the financial early-warning system from the choice of financial indexes," *East China Economic Management*, vol. 18(2), pp. 171-173, 2004.
- [3] Zhao Xiaoju, Liu Yongming, *Credit Management of Financial Institutions*, Beijing: China Fangzheng Press, 2004.
- [4] Liu Sifeng, Dang Yaoguo, Fang Zhigeng, *Theory and Application of Grey Systems*, Beijing: Science Press, 2004.
- [5] Xu Lingfeng, *A Ressearch on the Credit Risk Measurement of China's Listed Companies*, Hangzhou: Zhejiang University, 2003.