Research of Financial Early-warning for Listed Companies Based on SVM

Hailun Huang[1], Wuxue Jiang[2], Shi Wang[2]

[1] School of Business, Wuhan Vocational College of Software and Engineering, Wuhan, 430205,China

[2] Department of Computer Engineering, Dongguan Polytechnic, Dongguan, 523808, China

Keywords: Financial early-warning, Support vector machine (SVM), Model.

Abstract. In this paper, the support vector machine (SVM) is applied for the early warning of financial crisis of listed companies, financial early-warning indexes and quantitative methods are analyzed, the early-warning ability of SVM model is verified by combining with the financial data of listed companies, and the real evidence has proved the feasibility of SVM model used for financial early-warning, finally, the possibility of using multi-mode classification to identify the alert degree of financial early-warning.

Introduction

Economic cycle theory is the precondition of the problems of financial early-warning, and the changes in the economic cycle can produce periodic financial risks to enterprises. As an enterprise, the accumulated risks in the long-term operating process can continue to conduct and influence the enterprise's financial information and finally form a financial crisis. Through collecting this kind of financial information and using certain methods for classification, identification and evaluation, the enterprise's financial risks could be pre-warned, which can assist the enterprise to avoid and scatter financial risks effectively [1]. This paper uses the binary discrimination model of support vector machine (SVM) and selects appropriate early-warning indexes to obtain the known early-warning information of the alert situation company to identify the new early-warning samples of the unknown alert situation. The research shows that on the accuracy of early-warning and the forward early-warning ability, SVM model is better than traditional early-warning methods [2].

SVM model

The model of a nonlinear support vector machine can be established through the following quadratic programming [3].

$$\text{Maximize } L(a)=\sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} a_i a_j y_i y_j (x_i^T x_j) \tag{1}$$

Subject to $0 \le a_i \le C$, $\sum_{j=1}^{l} y_i a_i = 0$, i,j=1,2,…,l

The corresponding classification function is:

$$f(x)=\text{sgn}\left\{\sum_{i=1}^{l} a^*_i y_i K(x_i, x) + b^*\right\} \tag{2}$$

In the above function, xi is the support vector, xj is the unknown vector, a is the optimal solution, and b is the classification threshold. Select Gaussion RBF kernel function as the inner product function of modeling, and the form of Gaussion RBF kernel function is:

$$K(x_i, x_j)=\exp\left\{-\frac{\|x-x_i\|^2}{2\sigma^2}\right\} \tag{3}$$

Sample collection and index selection

Sample collection

In order to improve the accuracy and representation of the model, the samples in this paper are limited to the manufacturing industry, and total 84 listed companies of Shanghai and Shenzhen under special treatment in 2002 and 2003 as well as the companies with normal finance at the same time. The sample data is in table 1.

Table 1: Sample statistics

	ST	Non ST	Summation
2002	28	28	56
2003	14	14	28
Summation	42	42	84

Index selection

This paper chooses 4 categories of 17 financial indexes including profitability, debt paying ability, asset management ability and ascending ability as alternative variables, and the specific indicator is as table 2.

Table 2: Alternative variables of financial indexes

Attribute of index	Name of index	Attribute of index	Name of index
Profitability	Earnings per share （X_1） Return on equity (X_2) Net profit rate of assets（X_3） Net profit margin on sales (X_4) Profit margin of main business (X_5) Operational index of net income ($X_{6)}$	Debt paying ability	Current ratio （X_7） Asset-liability ratio (X_8) Equity ratio (X_9) Number of times interest earned (X_{10})
Asset management ability	Rate of stock turnover (X_{11}) Turnover of account receivable (X_{12}) Turnover of total assets (X_{13})	Ascending ability	Year-on-year growth rate of earnings per share（X_{14}） Year-on-year growth of net income（X_{15}） Year-on-year growth rate of total assets (X_{16}) Year-on-year growth rate of net assets (X_{17})

The result of Wilcoxon rank test shows that the sample data approximately meets the assumption of normal distribution and T assumption checking can be used for screening financial indexes [4]. Through the analysis of the T test results (table 3) of the difference among the financial ratio indexes of ST and non-ST companies, this paper has eliminated six indexes including number of times interest earned, Rate of stock turnover, turnover of account receivable, year-on-year growth rate of earnings per share, year-on-year growth of net income, and year-on-year growth rate of net assets, and the rest 11 indexes are remained as the modeling variables.

Table 3: Sample descriptive statistics and T test results

Variable	Means		T test		Wilcoxon rank test	
	Non-ST	ST	t-value	p-value	z-value	p-value
（X1）	0.1981	-0.5936	-6.167	0.000	-6.358	0.000
(X2)	2.7228	-174.5741	-2.490	0.015	-5.976	0.000
（X3）	2.3361	-18.6441	-5.371	0.000	-5.768	0.000
(X4)	-5.1916	-538.8405	-2.782	0.007	-4.637	0.000
(X5)	23.2118	2.1002	-2.912	0.005	-3.520	0.000

(X6)	0.6848	-0.7486	-3.191	0.002	-4.020	0.000
（X7）	1.5012	0.8254	-5.712	0.000	-5.371	0.000
(X8)	43.9399	114.6624	2.462	0.016	-6.261	0.000
(X9)	134.4116	377.2814	2.125	0.037	-2.688	0.007
(X10)	12.8110	-1.7612	-0.666	0.507	-5.346	0.000
(X11)	3.8333	5.6958	1.454	0.150	-0.633	0.527
(X12)	41.8394	21.0958	-0.716	0.476	-3.473	0.01
(X13)	0.6376	0.4177	-4.155	0.000	-4.666	0.000
（X14）	-146.7901	21.2766	1.285	0.203	-2.086	0.037
（X15）	-145.7357	42.0404	1.587	0.116	-2.383	0.017
(X16)	16.1405	-16.2020	-5.315	0.000	-4.890	0.000
(X17)	37.9262	-101.9581	-1.674	0.102	-3.907	0.000

Note: T test value is the pair test value of heteroscedasticity: the standard value is p-value that is 0.05 or less.

The null hypothesis is: there's no difference between the financial indexes of ST companies and the financial indexes of non-ST companies in the significance level of 0.05.

Training evaluation of the early-warning model

In the phase of training evaluation of the early-warning model, it is mainly to train the acquired sample data and determine the model parameter C and σ [5]. In this paper, libsvm software tool has been used to automatically realize the whole process.

Determination of model parameters

In order to test the classification of the model itself, the training and the test samples in this paper use the same data. 84 companies' financial data samples have been taken as the training and evaluation samples, which include isometric ST samples and non-ST samples, and LIBSVM tool is used to automatically conduct the training evaluation of the samples. The training phase includes that: train the input training samples to get the initial value of the model; and then use the above algorithm to extract the effective and relevant data for key training to get the final model; the task of the input phase of training parameters is mainly to determine the parameter c and σ of SVM model. After the test of parameter training, C=arclg(C)=8192.0 , $\sigma = \sqrt{arc\ \lg(gamma)}$ =0.125， and CV rate =92.5%. Input the known and classified training samples into the model for inspection as the test samples, in the 84 samples, 83 of them have been classified correctly, and the accuracy has reached, from which we can see that this model has a good classification ability in financial early warning.

Analysis of the model of early warning

After the parameters of the early-warning model have been determined, 32 sample companies' financial data has been chosen as the test samples in this paper, and the companies' financial data in the year of T, T+1, T+2 and T+3 have been respectively input into the model test. And the results are in table 4.

Table 4: Classification analysis of the test samples

Year (s)	Group	Test sample	
		Actual number	Classification number
T	ST company	18	18
	Non-ST company	14	14
	Classification accuracy	100%	
T+1	ST company	18	21
	Non-ST company	14	11
	Classification accuracy	90.625%	
T+2	ST company	18	13
	Non-ST company	14	19
	Classification accuracy	84.375%	
T+3	ST company	18	10
	Non-ST company	14	22
	Classification accuracy	75%	

We can see from table 4 that the classification accuracy in the year of T, namely the year when the financial crisis happens, is 100%, and the reason is that the data in the year of T comes from one part of the determined samples of model parameters. The earlier the year is, the gradually lower the classification accuracy is, which is because the data of the model comes from the year of T, and the earlier the year is, the difference between the financial crisis information included in the test of sample data and the sample information data of the model training will be widened, eventually, the classification accuracy will reduce. In order to improve the early-warning classification accuracy of medium and long-term financial crisis, the training sample appropriately including the data of the previous years of the crisis can obviously increase the medium and long-term classification accuracy.

Conclusion

SVM has good fitting precision and generalization ability, and it is widely used in the aspect of pattern classification and regression analysis. In this paper, SVM model's two types of classification of financial early warning are discussed and SVM supports multi-class classification, which have provided technical support to us for the establishment of recognition of alert degree. The problems of how to select financial indexes and how to identify training samples when setting up recognition of alert degree deserve further researches.

Acknowledgements

This work was financially supported by the fund project of Dongguan Polytechnic (No.2014a06, No.2014a07), and by by the project of Guangdong higher vocational education research association(No.GDGZ14Y07), and by the College students off-campus practice teaching base construction projects of Guangdong Province in 2014, and by the project of China higher vocational education research association(No. GZYLX1213279), and by the Key Teaching Reform Project of Dongguan Polytechnic (JGXM2014020).

References

[1] Jae Joon Ahn, Kyong Joo Oh, Tae Yoon Kim, Dong Ha Kim. Usefulness of support vector machine to develop an early warning system for financial crisis. *Expert Systems with Applications,* 38(4), pp. 2966-2973,2011.

[2] Tsui-Chih Wu, Ming-Fu Hsu, Credit risk assessment and decision making by a fusion approach. *Knowledge-Based Systems*, 35, pp. 102-110, 2012.

[3] Boshoff, W. H.. The transmission of foreign financial crises to South Africa: A firm-level study. *Studies in Economic and Econometrics,* 30(2), pp. 61-85, 2006.

[4] Huang Li, Jia Shuyuan, Jia Jihan and Chen Hong, 2014, Innovative Applications of IE Technological in the Service Sector: the Case Study- A Case of Panzhihua University Canteen Service System, Journal of applied science and engineering innovation, Vol.1 No.5, pp.338-344.

[5] Tennant, D., & Kirton, C.. Assessing the impact of financial instability: The Jamaican case study. *Iberoamericana Nordic Journal of Latin American and Caribbean Studies*, 36(1), pp. 9–36, 2006.

[6] L.C. Jen, C.H. Chou, G. Allenby. A Bayesian approach to modeling purchase frequency. *Marketing Letters,* 14 (1), pp. 5–20, 2003.

[7] Ping Hu, Zujun Ma, 2014, Travel Efficiency in Urban Traffic Networks Based on Routing Strategies ,Journal of applied science and engineering innovation, Vol.1, No.5, pp,289-296.