INTERNATIONAL JOURNAL ON MANAGEMENT AND EDUCATION HUMAN DEVELOPMENT

J. Management & Education Human Development

ISSN: ISSN: 2775 - 7765
web link: http://www.ijmehd.com

Research on Financial Risk of Chinese Manufacturing Listed Companies Based on SVM

Yuan Yan[1], Xiahui Che[2]

[1]Graduate School, Jose Rizal University, Manila 0900, Philippines
[2]School of Management Science and Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China

Received: 20/08/2021 Accepted: 26/12/2021 Published: 15/03/2022

Representative e-Mail: chexiahui@yeah.net

ABSTRACT

Financial crisis early warning is of great significance for the company's management, investors and government regulators to make correct crisis identification. Up to now, many scholars at home and abroad have used neural networks, support vector machines, and logistic regression to conduct financial crisis early warning research on listed companies, and have obtained many valuable research results. However, for the financial crisis early warning of listed companies in China's manufacturing industry, there is no financial crisis early warning model that considers non-financial indicators such as the nature of equity and the nature of EVA. This article takes the manufacturing companies listed on the Shanghai and Shenzhen A-share main boards as a sample, uses principal component analysis to reduce dimensionality, and select 7 indicators from 30 financial and non-financial indicators as typical indicators of financial crisis early warning, the establishment of SVM, Logistic and KNN three model algorithms for empirical research, the research results found: return on assets, total asset net profit rate, quick ratio, fixed asset turnover rate, total asset turnover rate, cash re The investment ratio and the total asset EVA ratio accounted for 81.427% of the predicted financial crisis. The first three years of these indicators data can be used to judge whether the company will have a financial crisis. At the same time, it also proves that the support vector machine has a better financial early warning ability.

Keywords: *Support Vector Machine, Logistic Regression, Financial Crisis Warning, Principal Component Analysis*

I. INTRODUCTION

With China's reform and opening up and its accession to the World Trade Organization, China's economic strength has increased day by day, its gross national product has continued to increase, and its gross industrial output has also been among the highest in the world. However, with the promotion of globalization and technological advancement, market competition has become increasingly fierce, coupled with some natural hazards brought about by global warming, companies are facing more and more uncertain factors, leading to increasing business and financial risks of companies. Therefore, it is becoming more and more important for companies to effectively predict financial crises. Research on financial crisis early warning has been relatively in-depth in foreign countries, and the models and methods are relatively mature. However, my country's financial early warning research started relatively late and there is no complete scientific standard system. The research content has developed from a single univariate analysis to multiple variable analysis, early warning indicators are gradually added from pure financial indicators to non-financial indicators. Taking into account the particularity of China's market economy, completely copying foreign financial early-warning systems may not be suitable for my country's actual conditions. Therefore, it is urgent to explore a set of reasonable financial crisis early-warning modeling methods to provide stakeholders with decision-making suggestions and techniques support.

II. RESEARCH METHOD

This article is a quantitative analysis article. Through a sample with binary classification, statistical models and artificial intelligence models are used for empirical analysis, and relevant graphics and tables are obtained. First of all, this paper makes a systematic analysis of financial crisis early warning at home and abroad，which is roughly divided into three aspects: selecting sample objects, selecting early warning indicators, and adopting early warning models. The more representative one is Fitzpatrick (1932)(Fitzpatrick, 1932) used univariate analysis; Beaver (Beaver, 1966) used statistical methods to study the problem of corporate financial failure; American scholar Altman (Altman, 1968)used the multi-discriminant analysis (MDA) for the first time for early warning of corporate financial crisis, and established a multi-discriminant Z-score model, which was widely accepted. However, the data analyzed by MDA must be normally distributed, and Martin (Martin, 1977)used the Logistic model to predict bank bankruptcy, and achieved good results, gradually replacing the mainstream position of MDA in financial crisis warning. There is the support vector machine theory first proposed by Vapnik (Vapnik, 1999)in the artificial intelligence model, its application is the most popular; Jeffrey Hinton and his student Ruslan Salahdinov(Hinton, Osindero, & Teh, 2006) formally proposed in 2006 the concept of deep learning has been the most widely used until now. Secondly, determine the research object and indicator system of this article, select the listed Chinese manufacturing companies that have been specially dealt with for the first time in 2019-2020, and the matching companies are the normal manufacturing listed companies with similar industries and total equity during this period; the indicator system is selected 23 financial indicators and 7 non-financial indicators, using principal component analysis method to reduce the dimensionality of these indicators. Finally, the model that this article focuses on is the support vector machine model, which was first applied by Fan and Palaniswami (Fan & Palaniswami, 2000) to the support vector machine theory in Australian companies; Min and Lee (Min & Lee, 2005)used it in South Korean companies, the effect is better than neural network, logistic and other models. In this paper, the model is used in special companies and normal companies in the manufacturing industry, and the data is fully utilized by the five-fold crossover. The grid search method is used to obtain the specific parameter indicators of various models, and the kernel function is used as a non-linear classifier; At the same time, in order to compare the application effects of support vector machines, logistic regression and KNN in the statistical analysis model were also selected as conpare models to compare which model is more suitable for predicting the financial crisis of the manufacturing industry.

III. DISCUSSION

3.1 The definition of financial crisis

In China, due to our special national conditions, there are few companies that go bankrupt. Once a listed company has a financial crisis, companies will resort to seeking political help or inviting a third party to participate in capital injection, or let other companies buy out the backdoor listing, so The definition of financial crisis in this article refers to the China Securities Regulatory Commission s document No. 6 "Notice on the Special Handling of Stocks During the Period of Abnormal Status of Listed Companies", The Shanghai Stock Exchange and Shenzhen Stock Exchange are required to treat listed companies with abnormal financial or other conditions within the Exchange for special treatment of their stocks, and the listed companies subject to special treatment should add the word "ST" before the company name, the date of stock quotation the price limit is 5%.

3.2 Samples

The samples in this article are from manufacturing companies listed on the main board of China’s A-share Shanghai and Shenzhen Stock Exchanges. The selected crisis sample objects are listed companies that have been specially dealt with by the China Stock Exchange for the first time in 2019-2020. The selection is based on 1) Recent Two consecutive years of losses or the current year’s net profit disclosed in the audited annual report is negative; 2) The audit result of the most recent fiscal year shows that its net assets are negative; 3) The most recent fiscal year’s financial accounting report is issued and cannot express an opinion or deny it opinions of the audit report. A total of 78 sample companies that were ST for the first time in two years were selected, and then 78 normal manufacturing listed companies were selected for matching in following the principle of the same industry and similar total equity.

3.3 Indicators

Indicator in the previous studies of financial crisis early warning by scholars have shown that they have an important impact on the accuracy and results of early warning, and the indicator system generally selected by scholars is mainly based on financial indicators. Ohlson (Ohlson, 1980) selects six financial indicators including the ratio of working capital to total liabilities, the ratio of total liabilities to total assets, the ratio of current liabilities to current assets, the ratio of capital reserves to total liabilities, the ratio of net income to total assets, and the growth rate of net income. And the size of the enterprise is used as control variables, and the asset-liability ratio and net income of the past two years are selected as dummy variables. Almamy ,Aston, & Ngwa (Almamy, Aston, & Ngwa, 2016)added a new financial indicator variable cash flow, which will be substituted into the original Z-score model of Altman (1968) to predict the financial crisis of financial crisis companies and healthy companies. The research results show that the predictive ability of the new model has been improved to 82.9%. Sun Xiao-lin(Sun, 2013) built a dynamic early warning system for financial crises based on the state-space model and selected 30 financial indicators for time series analysis and testing. The results show that the state space model can effectively describe the cumulative variation of the company's financial status over time and realize the dynamic early warning system recursive update and real-time prediction. Wang Xiaoyan & Jiahan (Xiaoyan & Jiahan, 2020) and others added 51 financial indicators for early warning of the financial crisis of listed companies, and adopted the GB method to screen 14 indicators into the model. When performing classified forecasting, the clustering GB model achieved good forecasting results. This article adds some new non-financial variables on the basis of previous studies. The indicator system of the article including solvency indicator, operating ability indicator, profitability indicator, cash ability indicator and development ability indicators belong to the financial indicator system; the equity structure and the EVA indicator structure belong to the non-financial indicator system. The preliminary determinations are 23 financial indicators and 7 non-financial indicators, the specific indicators are shown in Table 1 and Table 2.

Table 1 Statistics of financial indicators

Indicator name	Indicator code	Variable name	Indicator name	Indicator code	Variable name
Solvency index	X1	Current ratio	Cash capacity indicator	X15	Net cash content of net profit
	X2	Quick ratio		X16	Cash content of operating income
	X3	Interest coverage ratio		X17	Net cash content of operating profit
	X4	Assets and liabilities		X18	Operating Index
	X5	Equity ratio		X19	Cash reinvestment ratio
Operational Capability Index	X6	Accounts Receivable Turnover Rate	Development Ability Index	X20	The growth rate of total assets
	X7	Inventory turnover		X21	Net profit growth rate
	X8	Liquid assets turnover rate		X22	Operating profit growth rate
	X9	The turnover rate of fixed assets		X23	Operating income growth rate
	X10	The turnover rate of total assets			
Profitability indicator	X11	Return on assets			
	X12	The net profit margin of total assets（ROA）			
	X13	Roe			
	X14	Operating profit margin			

Table 2 Statistics of non-financial indicators

Indicator name	Indicator code	Variable name	Indicator name	Indicator code	Variable name
Nature of equity	G1	The largest shareholder's shareholding ratio	EVA properties	G4	EVA per share
	G2	The shareholding ratio of the top ten shareholders		G5	EVA rate
	G3	The separation rate of two rights		G6	EVA rate of total assets
				G7	EVA rate of net assets

3.4 Data

Listed companies will publicize their financial statements every quarter, but only the fourth-quarter reports, namely the annual report and the second-quarter report, that is, the interim report will be reviewed by the accounting firm. This indicates that the data of these two quarters is relatively authoritative, while the interim report only has incomplete data for half a year. Therefore, this article uses annual report data for financial early warning analysis. The year in which the listed company is specially treated is assumed to be year T, and the special treatment period and the financial report (T-1) of the previous year are in the same period, which means that it is meaningless to select the data of the previous year to predict the financial crisis of the company; generally, under the circumstances, researchers will not choose the data of the T-2 period to predict the financial status of T year. Because according to the listed company guidelines that A shares are specially processed, listed companies that have lost money for two consecutive years will be ST, for example, a company is profitable in year T-1, so no matter whether it is profitable or loss-making in year T-2, it will not be ST in year T. Therefore, use the data of year T-3 to predict the financial status of the listed company in year T, the theoretical prediction effect is the best, and the above-mentioned problems will not exist. Therefore, this paper selects T-3 data from the specially processed company, that is, the data from 2016 to 2017 for early warning research. The data comes from the

Wind database and the CSMAR database.

3.5 Statistical Analysis

There are many types of financial crisis indicator systems, and the situation of each indicator is different, and the correlation between the indicators is also uncertain. If all indicators are incorporated into the model, it will inevitably lead to a large number of calculations in the model. The probability of errors increases. T

o improves the representativeness and significance of the indicators; it is necessary to perform statistical analysis on the above 30 indicators.

3.5.1Normal distribution test.

The normal distribution test is a test to determine whether there is a significant difference between the background population represented by the same and the theoretical normal distribution. It has the most important significance and is the most widely used test method. It is the premise of parametric statistical analysis. The KS test is commonly used normal test method. The formula of K-S test is:

$$K=Max\left[F_n(x)-F_0(x)\right] \quad (1)$$

Among them, $F_0(x)$ represents the cumulative probability of the normal distribution, $F_n(x)$ represents the cumulative probability of the variable, set the significance to P, when P>0.05, then $F_0(x)$ obeys the normal distribution, when P<0.05, then $F_0(x)$ does not obey the normal distribution .The following uses SPSS24.0 software to perform K-S test on the sample data of listed manufacturing companies, and the normal distribution results obtained are shown in Table 3.

Table 3 Normality test

	Shapiro-wilk				Shapiro-wilk		
	statistical	df	significance		statistical	df	significance
X1	0.842	78	0.000	X16	0.948	78	0.003
X2	0.837	78	0.000	X17	0.242	78	0.000
X3	0.459	78	0.000	X18	0.751	78	0.000
X4	0.971	78	**0.074**	X19	0.321	78	0.000
X5	0.503	78	0.000	X20	0.307	78	0.000
X6	0.650	78	0.000	X21	0.348	78	0.000
X7	0.844	78	0.000	X22	0.440	78	0.000
X8	0.848	78	0.000	X23	0.279	78	0.000
X9	0.211	78	0.000	G1	0.969	78	**0.053**
X10	0.712	78	0.000	G2	0.988	78	**0.660**
X11	0.824	78	0.000	G3	0.620	78	0.000
X12	0.863	78	0.000	G4	0.963	78	0.023
X13	0.839	78	0.000	G5	0.736	78	0.000
X14	0.655	78	0.000	G6	0.863	78	0.000
X15	0.591	78	0.000	G7	0.886	78	0.000

It can be seen from the results in Table 3 that three of all 30 indicators obey a normal distribution, namely the asset-liability ratio (X4), the shareholding ratio of the largest shareholder (G1), and the shareholding ratio of the top ten shareholders (G2), the rest of the indicators do not obey the normal distribution.

3.5.2 Significance test

The following uses the T difference test method to test the significance of the indicators that obey the normal distribution, and the Mann-Whitney U nonparametric test for the variables that do not obey the normal distribution. The SPSS24.0 software is used for statistics, and the T test statistical results are shown in Table 4:

Table 4 Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
X4	Equal variances assumed	2.179	0.142	3.778	154.000	0.000	0.122	0.032	0.058	0.186
	Equal variances not assumed			3.778	150.035	0.000	0.122	0.032	0.058	0.186
G1	Equal variances assumed	1.736	0.190	(3.397)	154.000	0.001	(0.079)	0.023	(0.125)	(0.033)
	Equal variances not assumed			(3.397)	152.161	0.001	(0.079)	0.023	(0.125)	(0.033)
G2	Equal variances assumed	0.520	0.472	(1.761)	154.000	0.080	(0.044)	0.025	(0.094)	0.005
	Equal variances not assumed			(1.761)	153.980	0.080	(0.044)	0.025	(0.094)	0.005

According to Table 4, the bilateral significance of the asset-liability ratio (X4) and the largest shareholder's shareholding ratio (G1) is less than 0.05, indicating that these two indicators can effectively distinguish special treatment companies from normal companies, but the top ten the significance of shareholder shareholding ratio (G2) is 0.08, which is greater than 0.05, so this indicator cannot effectively distinguish between special handling companies and normal companies, and this indicator should be deleted.

For the remaining 27 indicators that do not obey the normal distribution, the Mann-Whitney U non-parametric test is used for testing. The results are shown in Table 5:

Table 5 Mann-Whitney U non-parametric test results

Indicator code	Sig. (2-tailed)	Indicator code	Sig. (2-tailed)
X1	0	X16	**0.224**
X2	0	X17	0.021
X3	0.009	X18	**0.131**
X5	0	X19	0
X6	0.004	X20	**0.662**
X7	**0.28**	X21	0.01
X8	0	X22	0.002
X9	0.02	X23	**0.288**
X10	0	G3	**0.788**
X11	0	G4	0
X12	0	G5	0
X13	0	G6	0
X14	0.004	G7	0
X15	0.04		

It can be seen from Table 5 that the 27-index data include inventory turnover rate (X7), operating income cash content (X16), operating index (X18), total asset growth rate (X20), operating income growth rate (X23) And the two-weight separation rate (G3), the significance test value of a total of six indicators is greater than 0.05, indicating that the indicators cannot clearly distinguish between special treatment companies and normal companies, and these indicators should be deleted. The remaining variables that can effectively distinguish the two types of companies are X1, X2, X3, X5, X6, X8, X9, X10, X11, X12, X13, X14, X15, X17, X19, X21, X22, G4, G5, G6 , G7, a total of 21 indicators can be used to build a financial crisis early warning model.

3.6 Factor analysis

A total of 23 indicator variables have been screened out in this paper. If so many variables are brought into the model, a large amount of calculation will be generated, and the correlation between indicators is not conducive to the test of the model. Therefore, 23 indicators need to be tested. Factor analysis to filter out variables with high information content and low relevance. Using SPSS24.0 software to perform factor analysis on 23 variables, the KMO value obtained was 0.654, which was greater than 0.5, and the sample data passed the KMO test. The Bartlett sphericity test statistic is 2278.40, and the significance is 0, which is less than 0.05. The result shows that the sample data is suitable for principal component analysis.

For the selected 23 variables, using principal component analysis, extract the common factors, select the principal components whose feature value is greater than 1 and the cumulative variance contribution rate reaches about 80%; then calculate the score of each principal component. In this paper, there are 7 indicators that meet the requirements of eigenvalues, which contain 81.427% of the information and integrate most of the information of the original indicators. Therefore, the samples can be explained better.

Table 6 Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.259	35.910	35.910	8.259	35.910	35.910
2	3.392	14.747	50.657	3.392	14.747	50.657
3	2.006	8.722	59.379	2.006	8.722	59.379
4	1.518	6.599	65.978	1.518	6.599	65.978
5	1.358	5.906	71.884	1.358	5.906	71.884
6	1.147	4.986	76.870	1.147	4.986	76.870
7	1.048	4.556	**81.427**	1.048	4.556	81.427

To explain the extracted 7 principal components, it is necessary to obtain the factor loadings of 23 original variables to 7 common factors. This paper uses the maximum variance method to perform orthogonal rotation to obtain the factor loading matrix. The results are shown in Table 7:

Table 7 Component matrix a after rotation

	Component						
	1	2	3	4	5	6	7
X1	0.032	**0.944**	0.032	-0.003	-0.069	-0.051	-0.018
X2	0.053	**0.945**	-0.045	0.054	-0.05	-0.016	0.017
X3	0.221	0.548	0.15	-0.206	0.026	0.023	0.014
X4	-0.231	-0.758	0.072	0.058	0.413	-0.16	0.24
X5	-0.137	-0.374	-0.071	-0.035	**0.821**	-0.16	0.15
X6	0.289	0.18	0.554	0.29	-0.005	0.022	0.036
X8	0.321	-0.064	**0.892**	0.031	-0.116	0.023	-0.034
X9	-0.009	-0.132	-0.006	-0.065	-0.084	-0.032	**0.899**
X10	0.139	0.022	**0.903**	0.098	-0.019	-0.124	0.054
X11	**0.936**	0.174	0.127	0.155	-0.047	-0.048	-0.018
X12	**0.922**	0.293	0.1	0.112	-0.086	-0.017	-0.011
X13	**0.907**	-0.048	0.121	0.198	0.146	-0.072	0.143
X14	0.635	0.055	-0.412	-0.199	-0.212	-0.082	-0.08

X15	-0.019	0.215	-0.119	0.16	0.039	**0.797**	0.079
X17	-0.047	-0.164	0.042	-0.142	0.008	**0.8**	-0.097
X19	0.051	0.017	-0.048	0.023	**0.867**	0.179	-0.191
X21	0.486	-0.157	0.113	**0.76**	-0.047	0.009	-0.062
X22	0.355	-0.168	0.174	**0.825**	-0.008	-0.028	-0.081
G1	0.126	0.157	0.188	0.441	0.095	0.038	0.39
G4	0.843	0.111	0.283	0.096	0.015	0.019	0.001
G5	**0.93**	0.039	0.168	0.261	0.007	0.004	-0.005
G6	**0.953**	0.113	0.159	0.145	0.004	-0.011	-0.002
G7	0.903	0.052	0.214	0.204	-0.121	0.06	-0.018

From the factor rotating load matrix, we can get: The variance cumulative load ratio of the principal component 1 reached 35.910%, accounting for more than one-third of the total ratio. Its dominant variables include the return on assets, the net profit rate of total assets, and the return on net assets , EVA rate, total asset EVA rate, these indicators reflect the profitability and EVA ability of the company, so F1 can be interpreted as profitability and the nature of EVA; the current ratio and quick ratio in factor F2 have a larger weight, so they mainly reflect the solvency of the company; the factors F3 and F7 together reflect the company's operating capacity, the representative indicators are total asset turnover and fixed asset turnover respectively. The higher the asset turnover, the better the efficiency of asset utilization and the stronger the company's operating capacity; The net profit growth rate and operating profit growth rate in the F4 factor account for the largest weight, which can reflect the company's development capabilities; factor F5 and factor F6 both reflect the company's cash capacity, and the representative indicators are the cash reinvestment ratio and operating Net cash content of profits. Therefore, through the above principal component factor analysis, this paper selects 7 new variables after eliminating the collinearity between the indicators to replace the original 23 indicator variables, and will use them as input variables to establish a model and predict.

F1= 0.081 * X1 + 0.085 * X2 + 0.101 * X3 - 0.124 * X4 - 0.102 * X5 + 0.184 * X6 + 0.189 * X8 - 0.018 * X9 + 0.143 * X10 + 0.332 * X11 + 0.330 * X12 + 0.305 * X13 + 0.149 * X14 + 0.008 * X15 - 0.039 * X17 - 0.011 * X19 + 0.225 * X21 + 0.195 * X22 + 0.107 * G1 + 0.306 * G4 + 0.333 * G5 + 0.334 * G6 + 0.328 * G7

The information of the 7 main factor calculation formulas thus obtained is as listed in the above formula. The other F2-F7 factor equations are similar in format content, but there is a difference in the number ratio, which is omitted here and in the actual software calculations and results there are detailed presentations.

3.7 Evaluation method of model results

After the model is established, the model needs to be evaluated. Different types of models have different evaluation standards. For example, regression evaluation indicators include relative and absolute errors, average absolute errors, relative square root errors, etc., and classification evaluation indicators include recognition accuracy, precision, and recall, F1 value, AUC value, etc. This article is a classification model. The model evaluation methods used include accuracy, recall, measuring the effectiveness of the binary model (AUC), and precision. We classify the financial crisis samples as the (ST) class as a positive class, and the normal company samples as the (non-ST class) as a negative class. The four cases in which the classifier predicts correctness on the test set are respectively denoted as: TP—The positive class is predicted as the number of positive classes, FN—predicts the positive class as the number of negative classes, FP—predicts the negative class as the number of positive classes, and TN—predicts the negative class as the number of negative classes.

Accuracy. The accuracy rate refers to the ratio of the correct sample to the total sample in the classified sample. The calculation formula is:

$$\frac{TP+FN}{TP+TN+FP+FN} \times 100\% \qquad (2)$$

Recall. The recall rate is for our original sample. It indicates how many positive examples in the sample are predicted correctly. There are also two possibilities, one is to predict the original positive class as a positive class (TP), and the other is to predict the original positive class as a negative class (FN):

$$Re\,c\,all = \frac{TP}{TP+FN} \times 100\% \qquad (3)$$

Precision. The accuracy rate is based on our prediction results, and it indicates how many of the samples whose predictions are positive are truly positive samples. Then there are two possibilities for the prediction to be positive, one is to predict the positive class as a positive class (TP), and the other is to predict the negative class as a positive class (FP), which is:

$$Precision = \frac{TP}{TP+FP} \times 100\% \qquad (4)$$

3.8 Model construction

In the current domestic literature on financial crisis early warning, many scholars are good at using logistic regression in econometrics and support vector machines in artificial intelligence system methods, and some scholars widely use BP neural networks and Z-score models, so this article ssing the same data sample set, substituting the data set into the support vector machine, KNN and Logistic models, compare which of the three models is more suitable for the financial crisis prediction of the manufacturing sample. This article uses the five-fold cross-validation method, and selects the average value as the result, and mainly uses Python 3.7 for programming analysis. The experimental results are shown in Table 8, Figure 1, Figure 2 and Figure 3.

Table 8 Results of the three models

	Logistic	KNN	SVM
Average ACC	0.616	0.654	0.686
Average AUC	0.702	0.687	0.739
Average Precision	0.614	0.665	0.699
Average Recall	0.682	0.614	0.649
Average AUPR	0.737	0.675	0.726

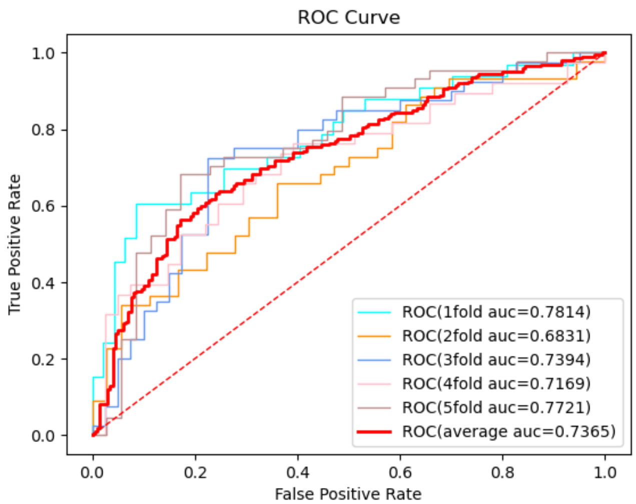

Figure 1 SVM model results

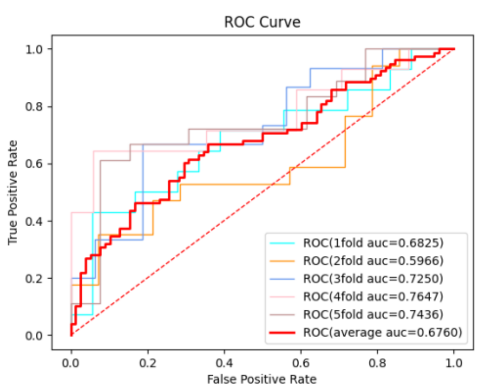

Figure 2 Logistic regression results

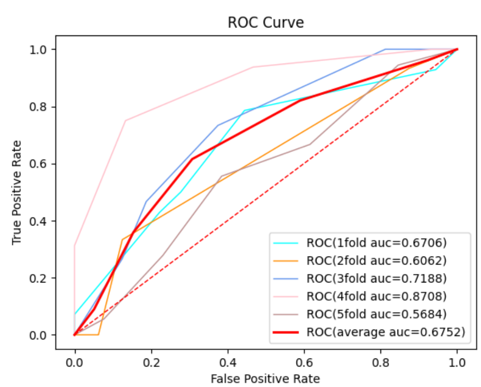

Figure 3 KNN model results

It can be seen from the results in Table 8 that the overall index performance of SVM is better than logistic regression and KNN among the average values obtained by using five-fold cross-validation. The theoretical and technical characteristics of the above model can also support the results of this experiment. SVM is based on the structure-based model requires a relatively small number of samples. It is good at coping with the linear inseparability of sample data. It is mainly achieved through kernel functions and slack variables. For classification samples, the classifier is only determined by the support vector, which can be effective Avoid over-fitting; Logistic model is more sensitive to multicollinearity data, usually the accuracy rate is not very high, because the form is very simple (very similar to linear model), it is difficult to fit the true distribution of the data; KNN needs to calculate each the distance between the test point and the training set. When the training set is large, the amount of calculation is quite large and the time complexity is high, especially when the number of features is relatively large. Therefore, it can be seen that for the manufacturing samples in this article, SVM is still the most suitable, the accuracy of the obtained model is high, and the obtained model has a strong generalization ability.

From the above three graphs, it can be found that the average AUC of each model method is above 0.6, and the distance is higher than the 0.5 dividing line. According to the AUC standard for judging the quality of the classifier (predictive model), the scope of this article is 0.5 <AUC<1 means that the classifier has predictive value and is better than random guessing. The red line represents the average value of the five-fold cross-validation. The best effect is SVM with a value of 0.7365. From this, it can be concluded that the discriminative ability of the support vector machine is stronger than that of the other two models.

IV. CONCLUSIONS

This article uses 156 companies in the Shanghai and Shenzhen A-share main board manufacturing industry as a sample to conduct an empirical study. The samples are matched at a ratio of 1:1, and a five-fold cross-validation is used to make full use of the sample data. By selecting seven aspects of corporate solvency, operating capability, profitability, cash capability, development capability, equity nature and EVA nature, 30 indicators were initially selected, and principal component analysis was used to reduce dimensionality, and 7 indicators were obtained as a typical indicator, finally, three algorithms with artificial intelligence algorithms and traditional statistical analysis algorithms as comparative models are constructed for empirical research. The research results obtained are as follows:

1. From the point of view of indicator variables, the indicator variables that affect the company's financial crisis mainly include return on assets, total asset net profit rate, quick ratio, fixed asset turnover rate, total asset turnover rate, and total asset EVA rate. Their distribution represents the company's profitability, solvency, operating capacity and EVA nature, and changes in these indicators can continue to predict the company's financial status, allowing the company's management and the board of directors to do risk prevention work three years in advance to avoid the company's continued financial crisis or even bankruptcy.
2. When predicting the financial crisis, considering that the financial indicators are all made by the company's management, to avoid the concealment or modification of financial indicators, this article adds the non-financial indicators EVA rate and equity structure. The results show that the EVA rate has an impact on the financial crisis early warning has a good predictive effect.
3. Comparing the SVM model, the Logistic model and the KNN model, it is found that the predictive classification ability of the SVM model is stronger than the other two types of models，the overall generalization ability and effect of the model are the best, and it can better show that the artificial intelligence algorithm is in the manufacturing industry the financial crisis early warning ability is stronger than traditional statistical analysis methods.

REFERENCES

Almamy, J., Aston, J., & Ngwa, L. N. (2016). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance, 36*, 278-285.

Altman, E. I. (1968). FINANCIAL RATIOS, DISCRIMINANT ANALYSIS AND THE PREDICTION OF CORPORATE BANKRUPTCY. *Journal of Finance, 23*(4), 589-609.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, 71-111.

Fan, A., & Palaniswami, M. (2000). *Selecting bankruptcy predictors using a support vector machine approach.* Paper presented at the Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium.

Fitzpatrick, P. J. (1932). A Comparison of the Ratios of Successful Industrial Enterprises with Those of Failed Companies.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation, 18*(7), 1527-1554.

Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of banking & finance, 1*(3), 249-276.

Min, J. H., & Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert systems with applications, 28*(4), 603-614.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109-131.

Sun, x. (2013). Dynamic Early Warning Model Based on a Financial Distress State Space:An Empirical Study in China. *Chinese soft science*(4), 8.

Vapnik, V. (1999). *The nature of statistical learning theory*: Springer science & business media.

Xiaoyan, W., & Jiahan, Y. (2020). Financial risk early warning model with penalty constraint based on cluster analysis. *Statistics and Decision making*(2), 4.