Research on Financial Early Warning Analysis Based on Big Data Technology

Xiangxia Shao(✉)

Xingzhi College Zhejiang Normal University, Business College, Jinhua, China
sxx@zjnu.cn

Abstract. Financial early warning analysis is of great significance to the survival and development of enterprises, and the application of big data technology helps to collect and process non-financial data in the financial early warning analysis model. This article first introduces the characteristics of big data and big data mining technology, and then analyzes the characteristics of financial early warning analysis in the era of big data. Finally, the analytic hierarchy process is used to construct a financial early warning analysis model and select listed companies for empirical analysis. The financial early warning analysis model introduces big data indicators as non-financial indicators, which can make a more accurate judgment on the company's financial status.

Keywords: Big data mining · Financial Early Warning Model · Analytic Hierarchy Process

1 Introduction

Big data refers to the term developed under the birth of information technologies such as the Internet of Things, mobile Internet, and cloud computing. With the rapid development of the Internet in the modern society of the 21st century, McKinsey & Co. Proposed the concept of big data, and society has entered the era of big data. The salient features of big data are the huge amount of information and various types. In this context, the financial analysis work of enterprises has been expanded and extended. As a link in financial analysis that must be highly valued, financial early warning analysis aims to discover the possible and existing financial risks of an enterprise, and issue warnings promptly so that operators can take corresponding countermeasures. Therefore, the establishment of a financial analysis model is very important for the early warning analysis of enterprises. In addition to financial data indicators, there are also non-financial data indicators that affect corporate financial risks. Big data technology provides new channels for the acquisition of non-financial data, and financial early warning analysis also has more diversified technical support due to the Internet. At the same time, the quantification of non-financial indicators provides new opportunities for the research of big data indicators.

G. Guan et al. (Eds.): ICBDSS 2022, AHCS 8, pp. 949–957, 2023.
https://doi.org/10.2991/978-94-6463-064-0_97

2 Big Data and Big Data Mining Technology

2.1 Characteristics of Big Data

The concept of big data was first proposed by McKinsey in the United States. McKinsey pointed out that "big data refers to data clusters whose size exceeds the capture, storage, management and analysis capabilities of traditional database software tools." Big data has five major characteristics as follows.

One is Volume, that is, the volume of data is huge and cannot be managed with existing technology, and it has increased geometrically.

The second is Variety, that is, the data types are diverse. The data types include structured data in the form of tables, as well as unstructured data such as logs, pictures, audio, and video.

The third is Value. Mass data contains a great value, but the value density is low.

The fourth is Veracity, which is authenticity. The authenticity of big data refers to the accuracy and reliability of the data and represents the quality of the data.

The fifth is Velocity, that is, the processing speed is fast. When an enterprise or system faces a rapidly increasing amount of data, it must process it at a high speed and respond quickly.

The massive characteristics of big data lead to the need for intelligent processing of big data, and the hidden characteristic information can be mined from the massive data, to provide financial personnel with decision-making, analysis, and identification capabilities through big data.

2.2 Big Data Mining Technology

Big data mining technology uses machine learning and data mining algorithms to find out the internal relationships and association rules hidden in massive data and mines valuable information. The core of data mining technology is data mining algorithms. Data mining algorithms perform data mining by establishing models Analysis. These models are based on statistical theory. Data analysis and mining can be better completed in the face of massive data. Data mining algorithms mainly include association analysis, clustering, classification, and prediction [1].

Association analysis is to extract meaningful association rules from data through mathematical models and make decisions through association rules. In association rule mining, the most commonly used is the Apriori algorithm. Through the clustering method, some unknown data can be classified and the corresponding results can be obtained. The clustering algorithm classifies the data through a certain distance calculation. Euclidean distance is often used when calculating distance. The distance calculation formula is as follows.

$$dist = \sqrt{(x_2 - x_1)^2 (y_2 - y_1)^2} \tag{1}$$

Classification and prediction: Through training labeled data, a model for classification and prediction is obtained, and some unknown labeled data are classified and predicted through the model. Common classification and prediction models include perceptrons, neural networks, logistic regression, and support vector machines.

3 The Characteristics of Financial Analysis in the Era of Big Data

3.1 Data Sources are Wider

The traditional financial early warning analysis mainly refers to the analysis of financial statements. The basis of the analysis is the statement data such as the balance sheet, the income statement, and the cash flow statement. The report data comes from the company's internal, measured in currency, and reflects the company's financial status and operating conditions. There are a lot of data outside the company that will affect the company's production and operation decisions. These data can also be used as a data source for financial early warning analysis. According to statistics from the China Mobile Internet database, as of December 2020, there are 1.158 billion mobile Internet users in China, and the average person spends about 6.4 h on the mobile Internet every day, for a total of 7.4 billion hours per day.

Various behaviors on the mobile Internet are recorded, and data generation becomes very easy. Everyone generates massive amounts of data every day, such as video data, e-commerce data, and social data. The US Internet Data Center has pointed out that the data on the Internet will increase by 50% every year. At this rate, the data doubles every two years. The center predicts that by 2025, the total installed volume of global Internet connection equipment is expected to reach 75.44 billion, and the scale of the global data circle will reach 175ZB. These data, especially the Internet public opinion data obtained from social networks, provide a broad data source for financial early warning analysis and then provide support for decision-making.

3.2 Can Be Analyzed in Rea- Time

In the era of big data, data collection and analysis technologies have developed rapidly. Companies can use big data technology to quickly and efficiently collect and perform real-time financial risk early warning analysis. Real-time early warning analysis enables companies to pay close attention to trends in financial risks and helps companies make decisions to deal with financial risks.

3.3 Improve the Effectiveness of Financial Early Warning Analysis

It is precise because of the use of big data technology that financial analysts can obtain various non-financial data in the course of business operations. Comprehensive consideration of financial data and non-financial data, seeking the linkage relationship between the two, constructing a model is conducive to analysis and evaluation of financial risks and then formulate corporate financial risk response strategies. This kind of comprehensive analysis is more effective than simply analyzing financial data. For example, statistical analysis of online customer reviews, forwarding and evaluation, and product search data can realize the understanding of market feedback, and then analyze the risks of the enterprise.

4 Financial Early Warning Analysis Model Based on Big Data

4.1 Selection of Early Warning Indicators

The financial early warning analysis indicator system is composed of a series of related indicators that can reflect all aspects of the business status of the enterprise. There is no mature indicator standard yet [2]. This article selects the construction of financial indicators and non-financial indicators. Financial indicators include indicators that reflect corporate profitability, operating capacity, debt solvency, and development capabilities. Non-financial indicators choose sentiment indicators calculated based on online reviews. Among them, the source of financial indicator data comes from the financial statements of the enterprise, while the source of non-financial indicator data is obtained from network evaluation by big data mining technology.

4.2 Analytic Hierarchy Process to Determine the Weight

There are many ways to determine the weight, and the analytic hierarchy process is one of them. The Analytic Hierarchy Process, abbreviated as AHP, is a decision-making method proposed by the American operations researcher T.L.Satty in the 1970s. It uses qualitative and quantitative methods to express and process human subjective judgments in quantitative forms. The core idea of the Analytic Hierarchy Process is to decompose the decision-making problem into different levels in the order of general goals, sub-goals, evaluation standards, and specific measures, and use the eigenvectors of the judgment matrix to solve the problem and use the weighting method to find the relative importance of lowest level and the highest level [3].

1) **Build a hierarchical model**

The financial early warning analysis index system is composed of the general target level A, the sub-target level B, and the specific index level C. As shown in Table 1.

2) **Construct a judgment matrix**

A scaling method from 1 to 9 and its reciprocal is used to indicate its importance [4]. As shown in Table 2.

According to Table 2, experts are asked to compare and assign values for every pair of indicators to construct a judgment matrix. Among them, $a_{ij} = 1 / a_{ij}$, $a_{ij} > 0$, as shown in Table 3.

Calculate the geometric mean of each row element of the judgment matrix and perform normalization processing to obtain the weight of each indicator. As shown in Table 3.

Table 1. Financial Early Warning Analysis System

Overall goal A	Sub-Goal B	Specific Index C
Financial Early Warning Analysis System A	Profitability B1	Operating net profit margin C1
		Return on assets C2
	Operational ability B2	Inventory turnover rate C3
		Accounts receivable turnover rate C4
	Solvency B3	Asset-liability ratio C5
		Current ratio C6
	Development Ability B4	Net profit growth rate C7
		Total assets growth rate C8
	Sentiment indicator B5	Positive sentiment index C9

Table 2. Quantification standard table

Scale a_{ij}	Definition	Description
1	Equally important	I factor and j factor is equally important
3	Slightly important	I factor is slightly more important than the j factor
5	Relatively important	I factor is relatively more important than factor j
7	Very important	I factor is more important than the j factor
9	Absolutely important	I factor is absolutely more important than factor j
2,4,6,8	Between adjacent levels	The intermediate state of the above two judgments

3) **Perform a consistency check**

Find the maximum characteristic root λ max of the judgment matrix, and calculate the consistency evaluation index CI.

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1} \tag{2}$$

Calculate the agreement ratio CR.

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{3}$$

RI is the degree of freedom index. RI is a fixed value, which can be found from the average random consistency test standard value table. As shown in Table 4.

If CR < 0.1, it is considered that the consistency of the judgment matrix can meet the requirements, and the resulting ranking vector is valid. If CR > 0.1, it is considered that the consistency of the judgment matrix cannot meet the requirements, and the

Table 3. Judgment matrix, weights, and test indicators

index	**matrix**					**Weight**	
A	B1	B2	B3	B4	B5		
B1	1	5	3	3	5	0.469	
B2	1/5	1	1/2	2	3	0.1433	$\lambda max = 5.2715$
B3	1/3	2	1	1	4	0.1932	$CI = 0.0679$
B4	1/3	1/2	1	1	3	0.1382	$CR = 0.0606$
B5	1/5	1/3	1/4	1/3	1	0.057	
B1	C1	C2					
C1	1	3				0.75	
C2	1/3	1				0.25	
B2	C3	C4					
C3	1	1/3				0.25	
C4	3	1				0.75	
B3	C5	C6					
C5	1	1/7				0.125	
C6	7	1				0.875	
B4	C7	C8					
C7	1	3				0.75	
C8	1/3	1				0.25	
B5	C9						
C9	1					1	

Table 4. Degree of Freedom Index

n	1	2	3	4	5	6
RI	0	0	0.58	0.90	1.12	1.24

judgment matrix must be rebuilt. For the first-order matrix and the second-order matrix, the corresponding RI value is 0, and the vector values can meet the requirements, and no consistency check is required. The calculation results are shown in Table 3.

4) **Build a hierarchical model**

According to the above calculation, the weight of the enterprise financial early warning index is obtained, as shown in Table 5.

Table 5. Financial warning indicator weight

Overall goal A	Sub-Goal B		Specific Index C	
A	B1	0.4549	C1	0.3412
			C2	0.1137
	B2	0.152	C3	0.0556
			C4	0.0964
	B3	0.1959	C5	0.0245
			C6	0.1714
	B4	0.1402	C7	0.1052
			C8	0.035
	B5	0.057	C9	0.057

Through the above analysis, the following financial early warning model relations can be obtained.

$$Z = 0.3412X_1 + 0.1137X_2 + 0.0556X_3 + 0.0964X_4 + 0.0245X_5 + 0.1714X_6 + 0.1052X_7 + 0.035X_8 + 0.057X_9 \quad (4)$$

Xi is the index value of each index. The larger the Z, the smaller the financial risk, and vice versa, the larger the financial risk.

4.3 Empirical Analysis

1) **Financial indicator**

This article selects the listed pharmaceutical manufacturing company "ST Guanfu" as the research object. The financial index data comes from the CSMAR economic and financial database. This article selects its 2016–2020 period-end financial data to calculate the company's financial indicators. As shown in Table 6.

2) **Non-financial indicators**

For non-financial indicators, sentiment indicators use web crawler technology to obtain online comment data from various financial and stock forums [5]. Multi-angle and multi-dimensional information related to enterprise development obtained through online public comments. This information is obtained by obtaining the evaluations of many Internet users on the company. The evaluation information of many netizens is characterized by real-time nature, multiple sources of information, and complex structure. They can find out the status of business activities from the network evaluation of people who are concerned about the development of the enterprise.

Table 6. 2016–2020 financial indicators and non-financial indicators of "ST Guan Fu"

	2016	2017	2018	2019	2020
Operating net profit margin C1	0.2790	0.0286	−0.1935	0.0420	0.0081
Return on assets C2	0.0528	0.0470	−0.3584	0.0966	0.0097
Inventory turnover rate C3	2.0103	12.8107	10.2566	10.7088	9.8863
Accounts receivable turnover rate C4	3.3719	18.3281	40.5093	38.0966	30.6836
Asset-liability ratio C5	0.3558	0.3495	0.6814	0.5955	0.6260
Current ratio C6	1.4255	1.3316	0.6902	0.7479	0.5642
Net profit growth rate C7	2.1346	1.1522	−50.5121	−0.9285	−2.3447
Total assets growth rate C8	0.4615	0.1532	−0.0660	0.0694	0.1185
Positive sentiment index C9	0.2620	0.2010	0.1332	0.2050	0.1973

Table 7. Z value of "ST Guan Fu"

	2016	2017	2018	2019	2020
Z	1.0467	2.8690	−0.8050	4.3525	3.3922

Through the analysis and processing of the crawled text information, the positive emotions in the network emotions are recognized. Then calculate the positive sentiment index. The positive sentiment index is equal to the number of positive sentiment online reviews divided by the total number of online reviews. The calculation results are shown in Table 6.

Substituting the company's 2016–2020 data into the above model, the Z indicator can be calculated for each year. The calculation results are shown in Table 7.

From the calculation results, it can be seen that the company's Z indicator declined in 2018, indicating that its financial risk is the greatest. It was also in 2018 that the company was ST. Since 2019, as the company's performance has increased and its operating capabilities have increased, its Z indicator has increased significantly and financial risks have decreased.

5 Conclusion

The rapid development of big data technology has provided new ideas and new technologies for research in many fields. Therefore, big data technology can also be used as an auxiliary forecast tool in financial early warning to improve forecast accuracy. Big data technology is mainly used to collect and process rich network data information, and to quantify non-financial indicators in early warning models. Considering both financial indicators and non-financial indicators in the financial early warning analysis model helps to make a more comprehensive and objective evaluation. According to the analytic hierarchy process, the weight of each indicator is determined and then the financial early

warning analysis model is constructed. According to the financial early warning analysis model, enterprises can pay attention to the situation of financial risks in real-time, which is conducive to the timely response of enterprises.

Acknowledgment. First of all, I want to thank my family for their understanding and support. In the process of writing the thesis, my family contributed a lot to me so that I had time to write the thesis. Secondly, I would like to thank my colleagues and friends for giving me a lot of help and constructive suggestions.

References

1. Jieying Lin, "Research on Accounting and Financial Analysis under the Background of Big Data Era," J. Journal of Jinan Vocational College, pp. 97–110, December 2015.
2. Ling Zheng, Jie Song, "Weight design of financial early warning indicators based on analytic hierarchy process (AHP)", J. China Market, pp. 111–113, October 2014.
3. Jiahui Yang, Jian He, "Research on Financial Risk Early Warning of Big Data Enterprises Based on AHP Model", J. Sci-Tech & Development of Enterprise, pp. 206–207+210, March 2021.
4. Wenxue Xu, Xuqiang Yang, "Application of AHP in Weight Design of Financial Early Warning Indicators", J. Friends of accounting, pp. 20–21, January 2010
5. Chunyao Luo, " Research on enterprise financial early warning based on big data—Take ABC Company as an example", D. Nanjing University of Science and Technology,2020.