Hindawi

Research Article

Analysis of Financial Risk Early Warning Systems of High-Tech Enterprises under Big Data Framework

Maotao Lai [ID]

Finance Department, Minxi Vocational & Technical College, Longyan 364000, China

Correspondence should be addressed to Maotao Lai; 82003001@lyun.edu.cn

With the further development of China's market economy, the competition faced by companies in the market has become more intense, and many companies have difficulty facing pressure and risks. Among the many types of enterprises, high-tech enterprises are the riskiest. The emergence of big data technologies and concepts in recent years has provided new opportunities for financial crisis early warning. Through in-depth study of the theoretical feasibility and practical value of big data indicators, the use of big data indicators to develop an early warning system for financial crises has important theoretical value for breaking through the stagnant predicament of financial crisis early warning. As a result of the preceding context, this research focuses on the influence of big data on the financial crisis early warning model, selects and quantifies the big data indicators and financial indicators, designs the financial crisis early warning model, and verifies its accuracy. The specific research design ideas include the following: (1) We make preliminary preparations for model construction. Preliminary determination and screening of training samples and early warning indicators are carried out, the samples needed to build the model and the early warning indicator system are determined, and the principles of the model methods used are briefly described. First, we perform a significant analysis of financial indicators and screen out early warning indicators that can clearly distinguish between financial crisis companies and nonfinancial crisis companies. (2) We analyze the sentiment tendency of the stock bar comment data to obtain big data indicators. Then, we establish a logistic model based on pure financial indicators and a logistic model that introduces big data indicators. Finally, the two models are tested and compared, the changes in the model's early warning effect before and after the introduction of big data indicators are analyzed, and the optimization effect of big data indicators on financial crisis early warning is tested.

1. Introduction

With the further development of China's market economy, the competition faced by companies in the market has become more intense, and many companies have difficulty facing pressure and risks. If a corporate crisis fails to be effectively controlled and adjusted, it will not only affect a single company, but also affect other players in the market, such as creditors, investors, securities, and the banking system. Therefore, the financial crisis of enterprises is a problem that affects the stability of the capital market [1, 2].

Among the many types of enterprises, high-tech enterprises [3, 4] are the riskiest. One of the characteristics of this type of enterprise is that high risks bring high returns. In international competition, China's high-tech enterprises have to face the impact of foreign enterprises' new technologies and strong capital, and the instability of enterprises is greater than that of ordinary traditional enterprises. However, high-tech enterprises are an industry that any country must vigorously promote and support. This industry represents more advanced technology and development trends and is more in line with the themes and trends of the times, while the high pollution and high energy consumption of traditional enterprises will gradually take part in the competition [5, 6].

Based on the review and summary of the theoretical methods of financial crisis early warning systems both domestically and internationally [7–9], this paper puts forward the reasons for the financial crisis of high-tech enterprises in the Chinese market and then establishes the

early warning of this type of enterprise based on the logistic regression method in the financial early warning model. The meaning of this article's topic selection is as follows:

(1) Through modeling and analysis, we put forward the research ideas for establishing China's high-tech enterprise model. Foreign research on financial risk early warning has been considerable and mature, but it is difficult to directly apply foreign early warning mechanisms to the Chinese market. After all, the economic environment and the companies thmeselves are different. There is also domestic research in this field, but this article carries out research and analysis of high-tech enterprises in order to obtain a more targeted financial early warning system and operating mechanism.

(2) We establish a targeted early warning model to improve the level of high-tech enterprises in dealing with financial crises. Under the market economy [10], business operations are full of risks. While enjoying the high returns brought by risks, companies are also suffering from the same crises. How to make enterprises have a place in the market competition without going into financial distress requires a set of practical crisis early warning systems to assist enterprises. The financial crisis early warning system not only give signals to companies before the crisis, but also helps companies strengthen their own management and operation level, actively discover the abnormal financial phenomena of the company in daily operations, and work hard to propose solutions to deal with risks or crises.

(3) It is conducive to the supervision and management of enterprises by relevant state departments and organizations. First of all, a large number of high-tech enterprises are state-owned enterprises. These enterprises use state-owned capital, so the state as the owner has the right to supervise and manage the operation status and operating results of the enterprise. The financial early warning system is an effective tool which can encourage enterprises to have the pressure and sense of mission to work hard to maintain and increase the value of state-owned capital. Secondly, the early warning system is also a new way of supervision for the China Securities Regulatory Commission. In the current securities market, one of the evaluation and management methods for listed companies [11, 12] is to carry out special treatment (such as ST); then, the CSRC can also try to use the financial early warning model as a company evaluation standard to measure the crises and risks threatening a company.

Key contributions of the proposed study include the following:

(i) An accurate early financial risk warning system is proposed for the high-tech enterprises

(ii) The key influential big data indicators are identified for accurate modeling

(iii) Regression modeling of the key indicators was used

The rest of the paper is structured as follows: Section 2 describes the related work in the subject area. Section 3 describes the proposed methodology adopted for achieving the objectives. Section 4 validates the proposed methodology using experimentation. Section 5 summarizes the work done.

2. Related Work

The research of big data in the field of financial crisis warning originated from the research of big data in the financial market. Enormous research conducted in this area by predecessors focused on the impact of big data on the financial market. The big data generated by network information can be quantified to accurately predict financial indicators. A large number of studies on big data in the financial field have aroused some scholars' interest. Scholars around the world have done less research on big data. There have been studies on the feasibility of introducing big data into the field of financial crisis early warning. From the perspective of the application of big data in the financial market, existing research generally collects information on network platforms and discovers the emotional tendencies and development trends of netizens from the information. Ranco et al.'s research shows that many seemingly irregular phenomena are displayed on the Internet platform, and through collection and analysis, it is found that a large number of Internet users' interactions are closely related to changes in the financial market. The research verifies that sentiment analysis of Twitter platform information can accurately determine the accumulation [16]. Based on this, Zheludev et al. also analyzed the Twitter platform information statistics and reached new conclusions. The indicators derived from the sentiment judgment of netizens' comments have improved the prediction accuracy of the Standard and Poor's 500 [17]. Zhang and others (from MIT) also selected the most active Twitter platform as the indicator collection source. After judging the sentiment index of netizens, they found that the rise of negative sentiment would lead to the decline of the Dow Jones Index, S&P 500 Index, and Nasdaq Index [18]. Tobias Preis et al. used the search behavior of netizens to predict the number of investors' buying and selling. The method adopted was to count the retrieval volume of corresponding words in a certain period of time, which is also a type of big data indicator application [19]. Domestic scholars also have similar research in this area. Li Jinhai and others collect online comment information, analyze the emotional tendencies of netizens through text mining, and design an early warning system for corporate Internet word-of-mouth to help companies understand the current status and future trends of online public opinion [20, 21]. In summary, researchers attach great importance to the impact of big data in the financial field. Bian Hairong believes that online information big data affects investor sentiment and is reflected in stock transactions because the information contained in online information big data has the characteristics of timeliness,

comprehensiveness, and diversity [22, 23]. Network information big data contains many sources of information, including reports and data disclosed by listed companies themselves; authoritative media reports on listed companies; and opinions, evaluations, and trend forecasts from the media and industry experts. Among the many network platforms, the stock bar platform is the most active and valuable for research. The stock bar platform is an important medium for information transmission between companies, media, institutional experts, and investors. The big data it generates is extremely valuable for studying the changes in the stock price and financial status of companies. While the impact of big data on financial position has been studied, there is a lack of evidence to support it. Whether it can help improve the effect is a very valuable research topic [24]. Therefore, judging from the feasibility, the cause is inseparable from the interaction of corporate stakeholders on various channels or platforms. The traces of people's behavior on the Internet will continue to form big data that will bring endless value to academic research. The financial status of an enterprise can be reflected through network information. The stakeholders of an enterprise express their opinions through the Internet. The emotional tendency formed by a large number of opinions represents the management and financial status of the enterprise. Through related theoretical sociological analysis, it can be considered that network information big data has strong objectivity. The results obtained from the total statistical analysis should be unconscious. Such objective and scientific data can provide help for financial crisis early warning. Not only is big data closely related to the financial status of enterprises, but also the quantitative processing is performed through computer natural language processing technology, and the results are more objective. Therefore, the indicators formed by the quantitative processing of big data can solve the one-sided, subjective, and difficult-to-quantify nonfinancial indicators in the past. The problem is of great significance to the study of financial crisis early warning. In summary, the information in social platforms is closely related to social and economic activities. Big data on network information provides new ideas and methods for exploring the status and trends of social economy. The introduction of big data can improve the selection of nonfinancial indicators. Although relevant research is still in its infancy, big data can make up for the limitations of nonfinancial indicators.

In terms of big data acquisition technology, network information data has various forms and large capacity. In the traditional mode, the search engine can significantly expand the collection's geographic reach, but the relevance and value of the information are low, and filtering information with low value density requires manual supervision and repeated screening, which is extremely inefficient. Scholar DuAn once pointed out that precise positioning of data can help to mine highly relevant and valuable data [25]. Therefore, intelligently locking the subject and collecting it constitute an effective way to obtain big data. Web crawlers can crawl relevant topic information in the network and can filter out irrelevant information according to requirements [26]. Compared with the traditional model, the web crawler covers the complete range of information, locks the target subject, and can collect and organize information while ensuring relevance and value, greatly improving efficiency. In terms of the quantitative analysis technology of big data, existing research mainly quantifies the big data generated by network information into two indicators: information volume and emotional tendency. The quantification of the amount of information is relatively simple, that is, the number of texts of network information is counted. The analysis of emotional tendency is more difficult. The big data information text collected on the Internet can generally be divided into the three forms of vocabulary, sentence, and document from the text content. The sentiment tendency analysis of the text requires different algorithms according to the form of the corpus. Big data of network information has various language styles, grammatical forms, and emotional characteristics. For this kind of corpus, scholar Tang Xiaobo pointed out that the research can be carried out by expanding the emotional dictionary. First, the emotional tendency of the words in the text is judged through the dictionary comparison, and then the emotional tendency of the entire sentence or the entire document is calculated by the algorithm [27]. This paper studies the introduction of big data indicators into the field of financial crisis early warning and collects a large amount of relevant information on the Internet for early warning. Collection efficiency is an important prerequisite. With limited resources, it is necessary to obtain as effectively as possible relevant information on corporate topics to adapt to financial crises. For the indicator requirements of the early warning model, web crawlers are the best choice. In addition, the collected network information includes three fields: basic vocabulary, financial professional vocabulary, and network terms. According to previous research, it is planned to expand on the basis of the basic emotion dictionary.

3. Method

There are many existing research methods: unary and multiple discriminant models, multiple logistic regression models (logistic regression), and artificial neural network models. Through comparison, we can see that although the unary decision model is widely applicable, its accuracy is not high. The independent variables of the multivariate discriminant model are required to be normally distributed, and the variables can be converted into categorical variables for processing through statistics. The accuracy is much higher than that of the unary linear discriminant, but the scope of application is limited. The multiple logistic regression model does not require normality for the distribution of the independent variables and has a wider range of applications. The assumptions and preconditions are also looser than those of the multiple discriminant model, and its accuracy is higher. However, some approximations are needed in the calculation process. The artificial neural network model is more complicated, in which the calculation is difficult and the degree of stability needs to be mentioned. After comparison, this article chooses a more stable and rigorous logistic regression model as the modeling

method. The basic form of the logistic regression model is as follows:

$$\ln\left(\frac{P}{1-P}\right) = a + \sum b_i x_j. \tag{1}$$

In the model, P is the probability of a company having a financial crisis under the variable $x_i(x_1, x_2, x_3 \ldots, x_m)$. The corresponding $1 - P$ is the probability of financial normal. Use $x_i(x_1, x_2, x_3 \ldots, x_m)$ as the i-th factor that affects financial risk prediction, and $a, b(j = 1, 2, 3, \ldots m)$ are estimated parameters. The basic model can be deformed to get the probability of a financial crisis P:

$$P = \frac{e^{a+\sum b_i x_j}}{1 + e^{a+\sum b_i x_j}}. \tag{2}$$

This article selects $P = 0.5$ as the cut of the model (cutoff point). When the calculated model value $P > 0.5$, it means that the possibility of being specially processed is relatively high. The company is considered to be a financial crisis company. The assumptions of the logistic regression model are as follows: (1) the dependent variable y is a dichotomous variable; (2) sample selection is random; (3) y is nonlinear with the independent variable x_i; (4) there is no collinearity between the independent variables.

This article takes high-tech companies as the research object takes the A-share listed companies of this type as a sample, and defines companies that are first in the forecast year by ST or ∗ST as financially distressed companies, while non-ST companies are defined as financially normal companies. This article first constructs an indicator system of financial early warning by selecting financial indicators and then uses statistical methods to build a logistic regression financial risk early warning model. Finally, the accuracy of the model is tested.

This article is a research on the improvement of the financial crisis early warning model with the introduction of big data indicators. It is necessary to reasonably quantify the indicators and analyze the forecast effects of the model after the introduction of new indicators. The specific research design ideas include 6 parts: First, select training samples. Select the training samples for establishing the model from the listed high-tech companies in China's A-share market. In addition, the remaining listed companies in the high-tech industry are used as test samples after the model is established to test the prediction accuracy of the model. Second, select early warning indicators. The early warning indicators studied in this paper include two parts: financial indicators and big data indicators. The selection of financial indicators refers to the existing literature and conducts a preliminary selection from the financial indicator analysis database of listed companies in China, paying attention to the comprehensiveness and systematicness of financial indicators. The selection of big data indicators is based on the above theoretical analysis, collected from a stock discussion community, paying attention to the relevance and availability of big data indicators. Third, select the

model method. The research in this paper introduces an early warning indicator system and uses statistical analysis methods to determine whether a company is a financial crisis company or a normal company. This is a typical binary classification problem. The definition of the critical point of the scope can determine the type of company, which matches the research of this article. Therefore, the logistic model analysis method is selected to establish an early warning model. Fourth is the treatment of financial indicators. SPSS software is used as an analysis tool to detect the significance of the abovementioned primary financial indicators. After SW normal distribution test, T parameter test, or Mann–Whitney U nonparametric test, the main indicators are selected from 66 financial indicators as early warning variables to clearly distinguish between the two sample groups of financial crisis companies and nonfinancial crisis companies. Fifth is the processing of big data indicators. Use web crawler software to build a stock bar review corpus, apply python software to analyze the sentiment orientation of the collected stock bar review information, and extract the sentiment orientation and opinion enthusiasm from millions of comments. Sixth is the construction of the financial crisis early warning model. Construct a logistic regression model with processed financial indicators as variables and record it as the original model. Then, introduce big data indicators and financial indicators together as variables to build a logistic regression model, which is recorded as an improved model. Compare and analyze the two models to verify the optimization effect of the improved model. The overall research and design ideas are also shown in Figure 1.

4. Experiments and Results

4.1. Financial Indicator Processing

4.1.1. Normal Distribution Test. In this paper, SPSS software is used for data processing. The Shapiro–Wilk method (SW test) in the software can make a judgment on the normal distribution. It is suitable for nonlarge sample detection. This article is used to judge the normal distribution of 66 financial indicators. There are some basic principles: first, arrange the index data with normal distribution characteristics in order of category, and calculate the theoretical cumulative probability of each layer distribution; second, arrange the index data to be tested in order of category, and obtain the empirical cumulative probability; third, compare and analyze the two sets of data, calculate the maximum deviation value, and confirm whether it belongs to the normal distribution according to the confidence. Assuming that all indicators obey a normal distribution, the significance standard value is determined to be 0.05, and the SW value and Sig value of all indicators are obtained by running the software. There should be no acceptance of the null hypothesis when the Sig value is less than 0.05, and the indicator should be considered to have failed normal distribution
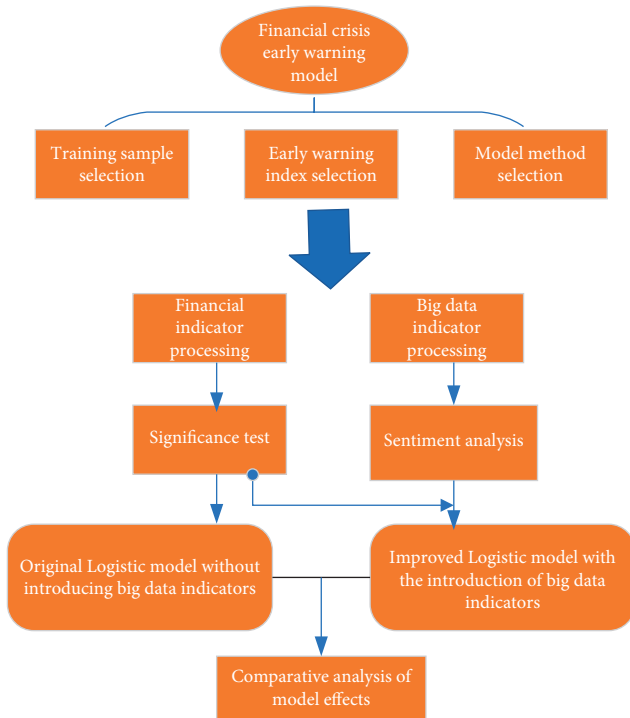
Figure 1: Financial model research and design ideas.

test. On the contrary, if it is greater than 0.05, it is considered to pass the normal distribution test. From the analysis and test results of the 66 financial indicators, we draw the following conclusions: First, using the significance level of 0.05 as the standard, we have a total of 12 financial indicators: X2, X3, X5, X20, X28, X32, X34, X35, X36, X38, X52, X56. If the Sig value is greater than 0.05, the 12 financial indicators generally conform to the normal distribution through the test, and the T parameter test method can be used for the significance test. Second, with the significance level of 0.05 as the standard, the remaining 54 financial indicators such as X1, X4, X6, X7, X8, X9, X10, X11, X12 have Sig values less than 0.05, finally failing to pass the test, which means that they do not obey the normal distribution, and the Maim–Whitney U nonparametric test method should be used for the significance test.

4.1.2. T Parameter Inspection. The indicators that pass the normal distribution test are tested for significance using T parameter analysis. From the results of the previous section, 12 indicators suitable for T-test are obtained. The test of such indicators can obtain indicators that can clearly distinguish between nonfinancial crisis samples and financial crisis samples. To begin, establish a null hypothesis that the indicator mean values of the two samples of financial crisis firms and nonfinancial crisis companies are not significantly different from one another. Make sure that the significance standard is 0.05. After the software is run, the T value and Sig value are sorted. If the value is less than 0.05, the null hypothesis should be rejected, and the indicators between the sample

groups should be clearly different. The null hypothesis should be accepted if the value is higher than or equal to 0.05 and the indicators between the sample groups have no obvious difference. Insignificant indicators are not suitable for judging the financial crisis and should be removed. The test results are shown in Table 1.

The analysis and test results lead to the following conclusions. First, with the significance level of 0.05 as the standard, the test values of X2, X3, X5, X32, X34, X35, a total of 6 indicators, are below the judgment standard value (<0.05), and these 6 indicators pass the test, which shows that the two types of samples can be fully distinguished and can be used in the model. Second, taking the significance level of 0.05 as the standard, the test values of the remaining 6 indicators, X20, X28, X36, X38, X52, X56, are above the judgment standard value (≥0.05). These 6 indicators have not passed the test, and the explanation is not sufficient to distinguish the two types of samples and cannot be used in the model.

4.1.3. Mann–Whitney U Test. For indicators that did not pass the normal distribution test, the Mann–Whitney U nonparametric analysis was used for the significance test. From the above processing results, 57 indicators suitable for nonparametric testing are obtained. Some of these indicators can significantly affect the occurrence of corporate financial crises, while some indicators have less significant impact on the occurrence of financial crises. In order to prevent the impact of insignificant indicators on the early warning effect of the model, it is necessary to conduct a significant analysis of the two samples of financial crisis companies and nonfinancial crisis companies and eliminate insignificant indicators. First, make the null hypothesis that there is no significant difference between the two samples index mean values, and determine the significance standard to be 0.05. After the software runs, the MW value and Sig value are sorted. If the value is less than 0.05, the null hypothesis should be rejected and the indicators between the sample groups should be rejected. There is a significant difference, and if it is greater than or equal to 0.05, the null hypothesis should be accepted and there is no significant difference in the indicators between the sample groups; such indicators should be removed. Analyzing the test results, we draw the following conclusions: First, using the significance level of 0.05 as the standard, X1, X4, X7, X12, X14, X15, X16, X17, X18, X19, X21, X24, X25, X29, X37, X42, X45, X54, X55, X59, X60, X63, and others, a total of 24 indicators of inspection values, are being judged to be below the standard value (<0.05). These 24 indicators pass the test and show that the two types of samples can be fully distinguished and can be used in the model. Second, taking the significance level of 0.05 as the standard, the test values of the remaining 32 indicators such as X5, X6, X8, X9, X10, X11, X13 are above the judgment standard value (≥0.05), and these 32 indicators have not pass the test and show that the two types of samples cannot be fully distinguished and cannot be used in the model.

TABLE 1: *T*-test result.

Index	F	F sig.	T	df	Sig. (double-sided)	Sig.
X2	0.92	0.26	−3.52	32	0.25	Y
X3	0.88	0.87	−2.95	32	0.29	Y
X5	0.65	0.62	−2.56	32	0.08	Y
X20	0.39	0.98	−1.29	32	0.87	N
X28	1.68	1.35	−0.98	32	0.28	N
X32	1.85	0.66	−0.87	32	0.76	Y
X34	1.20	0.82	−0.38	32	0.09	Y
X35	0.52	0.58	−2.67	32	0.12	Y
X36	0.38	0.37	−0.35	32	0.46	N
X38	0.98	0.86	−0.86	32	0.89	N
X52	1.02	0.74	−1.74	32	0.77	N
X56	1.56	0.28	−2.01	32	0.63	N

4.2. Big Data Indicator Processing

4.2.1. Construction of Sentiment Dictionary for Big Data Indicators.
The specific dictionary composition includes the following aspects: first, the basic dictionary is mainly based on HowNet's emotional dictionary; second, the domain dictionary is mainly based on the securities operation vocabulary and the ready-made financial field emotional dictionary; third, the network term dictionary is mainly based on BosonNLP sentiment dictionary and SnowNLP sentiment dictionary, combined with other stocks sentiment dictionaries, and is derived from the annotation analysis of millions of articles in forums, post bars, and online media, as well as online comments. Through the above-mentioned emotional dictionary expansion, an emotional dictionary that meets the quantitative processing of big data indicators is obtained. In addition to the basic negative adverbs and degree adverbs, the emotional vocabulary has reached more than 500,000. Examples of emotional words are shown in Table 2.

4.2.2. Big Data Indicator Processing Results.
Big data indicators come from the stock bar reviews, which include two indicators, sentiment value indicator and information popularity indicator. Through the above analysis, this article considers the big data indicators of T-3 and T-2 at the same time. Therefore, there are 4 subindicators, namely, the sentiment value index of T-3 year, the information popularity index of T-3 year, T-2 year's sentiment value index, and T-3 year's information heat index. The calculation of the enterprise's information popularity index is relatively simple, which is statistics of review information in the year, and the logarithm of the data value is too large. However, the quantitative processing of corporate sentiment value indicators is more complicated, including processes such as text segmentation, word frequency statistics, dictionary comparison, and sentiment judgment. Both indicators reflect the social network's attitude and evaluation of the overall status of the enterprise to varying degrees and reflect the management and financial status of the enterprise from another perspective, which is of great significance for supplementing the lack of financial indicators. After the above analysis and processing, two big data indicators can be obtained as shown in Table 3.

4.3. Effect Test and Comparative Analysis of the Two Models.
The big data indicators obtained through the above quantitative processing can be used to improve the research of financial crisis early warning. According to the analysis of the concept of big data indicators in Section 2, the big data indicators in this article are quantified from the comments of mass stock bars, and the majority of stakeholders of the enterprise behaviors and emotions show a certain value under a large number of repeated interactions, reflecting the financial status and development trend of the enterprise. The core of financial crisis early warning is the early warning function. Whether big data indicators are the cause of the financial crisis is still uncertain, but they cover a wide range of information and have a symptomatic effect on the early warning of corporate financial crisis. The comprehensive use of big data indicators and financial indicators can reflect the state of the enterprise to the greatest extent and can make up for and correct the shortcomings of simply using financial indicators in the early warning of financial crises. In the following, we will use the logistic method to analyze the early warning effect of the model before and after the introduction of big data indicators by testing sample data. First, establish a logistic financial crisis early warning model that uses financial indicators only, and then integrate big data indicators and financial indicators to establish a logistic model that introduces big data indicators. Finally, the early warning effects of the two models are tested and a comparative analysis is made to verify the important role of big data indicators in financial crisis early warning.

(1) The logistic early warning model that does not introduce big data indicators checks the forecasting effect of the test model, making its financial crisis early warning function more realistic. Substituting the relevant data of T-3 years of 36 test samples into the early warning model and taking 0.5 as the split point for judgment, we find that when $p > 0.05$, it is a financial crisis company. When $p < 0.05$, it is a company with normal financial status. The 36 test samples contain 12 ST samples and 24 non-ST samples, and the actual effect of the early warning model without the introduction of big data indicators is obtained, as shown in Figures 2–4.

TABLE 2: The sentiment dictionary of big data indicators.

Vocabulary category	Basic dictionary	Domain dictionary	Network dictionary
Positive emotions	Love, respect, etc.	Rise, daily limit, etc.	Reliable, pretty, etc.
Negative emotions	Regret, mean, etc.	Crash, clearance, etc.	Stupid, copycat, etc.
Neutral emotions	Hunch, weather, etc.	Blue chips, option, etc.	Friend, bamboo, etc.

TABLE 3: Big data index result processing.

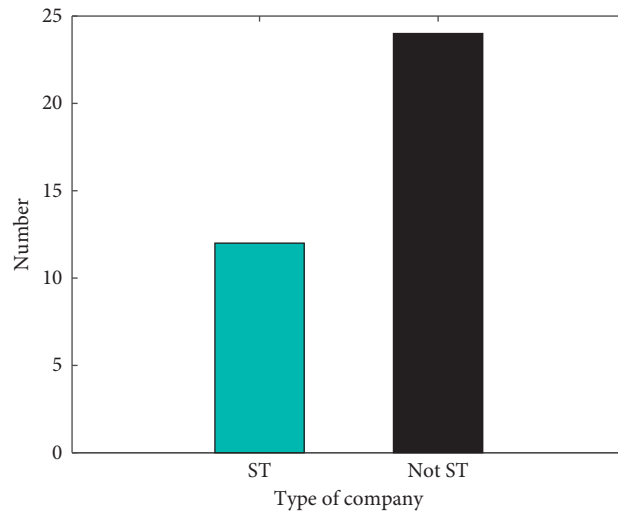| Big data index	Index name	Remark
B1	Emotion value 1	T-3 annual data
C1	Information fever 1	T-3 annual data
B2	Emotion value 2	T-2 annual data
C2	Information fever 2	T-2 annual data

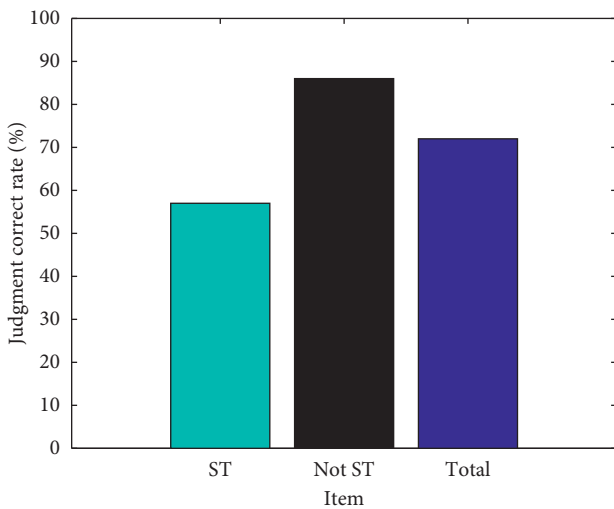FIGURE 2: Comparison of ST and non-ST companies in the total sample.

FIGURE 3: The accuracy of the original model without introducing big data indicators.

FIGURE 4: The misjudgment rate of the original model.

According to the test results, it can be found that the overall prediction accuracy rate is 77.78%. Specifically, the test sample consists of a total of 12 ST 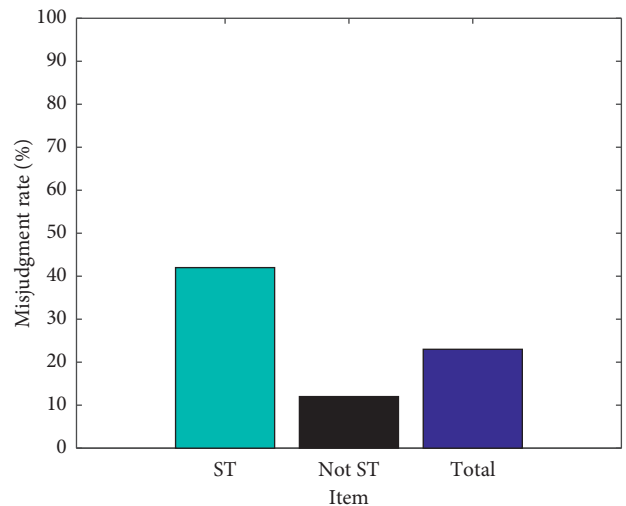samples and 24 non-ST samples. The original model that did not introduce big data indicators judged 7 ST samples and 21 non-ST companies and judged ST companies and non-ST companies The accuracy rates are 58.33% and 87.50%, respectively.
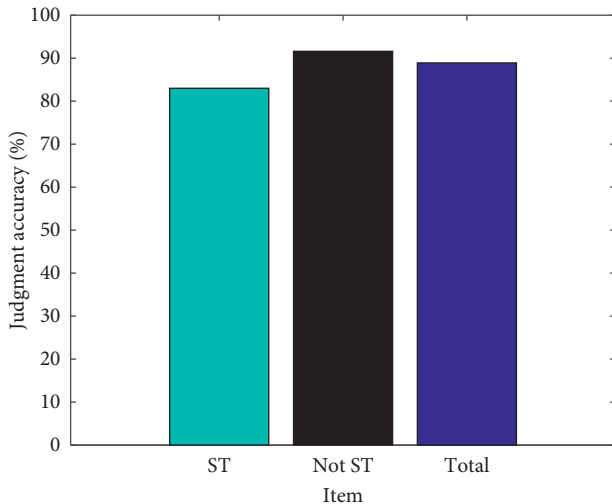
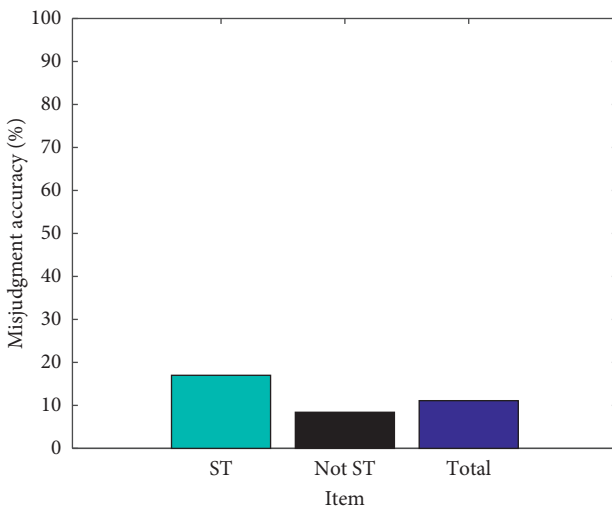FIGURE 5: The accuracy of the improved model with introducing big data indicators.

FIGURE 6: The misjudgment rate of the improved model.

(2) Introduce the (T-2) year big data indicators of 36 test samples into the original model, and build an improved model together with financial indicators. The same judgment standard is used to test the effect, and the actual effect of the early warning model with the introduction of big data indicators is obtained, as shown in Figures 5 and 6.

It can be found that the overall prediction accuracy rate is 88.89%. Specifically, the test sample has 36 samples, 12 ST samples and 24 non-ST samples. The improved model that introduces big data indicators judges 10 ST samples and 22 non-ST samples. The company's judgment accuracy rate was 83.33% and 91.67% respectively, and the overall forecast accuracy rate was 88.89%. In summary, comparing the test results of the two models, we can clearly see that in terms of overall prediction accuracy, the improved model with big data indicators has an accuracy rate of 11.11% higher than that of the original model without big data indicators. In terms of prediction accuracy, after the introduction of big data indicators, the model's accuracy in predicting ST samples increased by 25%. The test results show that the forecast accuracy of the financial crisis early warning model has indeed been improved after the introduction of big data indicators.

5. Conclusion

With the rapid development of big data thinking and technology, big data has penetrated almost all fields, affecting the development of all walks of life, and hence became the focus of attention in academia. This article is based on the industry effect of high-tech enterprises and the enthusiasm of netizens' exchanges, and it can well reflect the situation of enterprises through early warning indicators. Therefore, we determine China's high-tech listed companies as the training samples of the model, select and quantify the big data indicators and financial indicators, design the financial crisis early warning model, and verify its accuracy. The following are the main work and research conclusions of this article. First, the introduction of big data indicators is an effective way to improve the effect of financial crisis early warning. The financial indicator information is too single, and the existing nonfinancial indicators are too subjective, which seriously impacts the effect of early warning. The development of big data has enabled more and more information to be collected and processed. The existing results show that the review information of stakeholders on a company has had an important impact on the value of the company. The company is in an increasingly complex market environment and is involved in social activities. Financial crisis early warning applications should pay attention to the development and changes of network interactive information. It is difficult to obtain the complete and true financial situation of a company only by the information required by the regulatory authorities. The quantified big data indicator information can fully and objectively reflect the status of the company, which is an important way to improve the effect of the financial early warning model. Second, we build a financial crisis early warning model that introduces big data indicators and find that the predictive effect of the financial crisis early warning model has been improved after the introduction of big data indicators. We design enterprise-related topics to collect network information big data, and judge the emotional tendency of the review text through the emotional dictionary to obtain the emotional value, that is, the big data indicator. Combining financial indicators, the logistic model is used to establish the financial crisis early warning model without the introduction of big data indicators. Through the effect test of inspection samples, the results show that the financial early warning model with big data indicators can improve the accuracy of financial crisis early warning.

Data Availability

The datasets used are available from the author on reasonable request.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

[1] W. Bessler and P. Kurmann, "Bank risk factors and changing risk exposures: capital market evidence before and during the financial crisis," *Journal of Financial Stability*, vol. 13, pp. 151–166, 2014.

[2] M. Prochniak and K. Wasiak, "The impact of the financial system on economic growth in the context of the global crisis: empirical evidence for the EU and OECD countries," *Empirica*, vol. 44, no. 2, pp. 295–337, 2017.

[3] L. X. Liang and H. F. Zhang, *An Empirical Study on the R&D Inputs Performance of High-Tech Enterprises*, Journal of Central South University, vol. 2, 2005.

[4] Z. Yuan, H. Yan, and L. V. Sumin, *An Empirical Study of the Impact of Patent on the Future Performance of High-Tech Enterprises*, Science of Science and Management of S & T, 2010.

[5] J. Liu, Y. Wang, H. Zhang, and J. Han, "A review of the eco-industrial park evaluation of high energy consumption and high pollution industry," in *Proceedings of the 2015 International Power, Electronics and Materials Engineering Conference*, Dalian, China, January 2015.

[6] S. M. Tehrani, A. R. Karbassi, J. Ghoddosi, and S. M. Monavari, "Prediction of energy consumption and urban air pollution reduction due to the traditional shopping with e-shopping substitution," *Journal of Food Agriculture and Environment*, vol. 7, no. 3, pp. 898–903, 2009.

[7] Y. Wang, *The Establishment of an Early Warning Model for Financial Crisis*, Journal of Xi'an Petroleum Institute, china, 2003.

[8] W. Peng, W. Zhong, and Y. Wu, "A comparison study and modeling suggestion on financial crisis prediction at home and abroad," in *Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, October 2010.

[9] C. Dai, *Development Status and Problem Analysis of Financial Early Warning System at Home and Abroad*, Science Mosaic, 2007.

[10] R. I. Mckinnon, *The Order of Economic Liberalization: Financial Control in the Transition to a Market Economy*, Johns Hopkins University Press, Charles Village, Baltimore, 1991.

[11] Y. C. Liu, W. N. Hu, and B. W. Sun, "Comprehensive evaluation of listed companies in power industry on the operation performance based on fuzzy borda combination method," in *Proceedings of the International Conference on Management Science & Engineering*, September 2013.

[12] A. K. Sharma and S. Kumar, "Economic value added (EVA)-literature review and relevant issues," *International Journal of Economics and Finance*, vol. 2, no. 2, pp. 200–220, 2010.

[13] Q. Zhang, Q. Zhang, and D. Sornette, "Early warning signals of financial crises with multi-scale quantile regressions of log-periodic power law singularities," *PLoS One*, vol. 11, no. 11, Article ID e0165819, 2016.

[14] K. C. Chiou, M. M. Lo, and G. W. Wu, "The minimizing prediction error on corporate financial distress forecasting model: an application of dynamic distress threshold value," in *Proceedings of the 2017 8th International Conference on Awareness Science and Technology*, November 2017.

[15] Y. Dang, *The Nature of the Firm and Corporate Governance: Dynamic Corporate Governance Oriented towards Innovation [J]*, Research of Institutional Economics, vol. 10, no. 4, pp. 64–91, 2012.

[16] G. Ranco, D. Aleksovski, G. Caldarelli, M. Grčar, and I. Mozetič, "The effects of twitter sentiment on stock price returns," *PLos One*, vol. 10, no. 9, pp. 1–21, 2015.

[17] I. Zheludev, R. Smith, and T. Aste, "When can social media lead financial markets," *Scientific Reports*, no. 4, pp. 74–89, 2014.

[18] X. Zhang, H. Fuehres, and P. Gloor, "Predicting stock market indicators through Twitter "I hope it is not as bad as I fear","" *Procedia - Social and Behavioral Sciences*, vol. 26, pp. 1–8, 2010.

[19] P. Tobias, "Quantifying trading behavior in financial markets using google trends," *Scientific Reports*, vol. 3, p. 1684, 2013.

[20] J. Li, Y. He, Y. Ma et al., *Research of Corporate IWOM Crisis Early Warning Based on Online Reviews Mining in the Era of Big Data*, Journal of Intelligence, 2015.

[21] J. R. Tang, Z. Y. Ding, X. U. Wen-Ting et al., *The Effect of Electronic Word-Of-Mouth on Corporate Performance: A Case of Online Tourism Industry*, Entific Decision Making, vol. 2017, no. 1, 2017.

[22] Z. Min, G. Hui, F. Yan, and D. Nie, "The research on analyzing emotional tendency of sentence based on watershed algorithm," in *Proceedings of the The 3rd International Conference on Information Sciences and Interaction Sciences*, June 2010.

[23] Z. Wei and Z. Chen, *Principal Components Analysis on Financial Crisis in Listed Company of China*, Zhengzhou Conference on Management of Technology, pp. 712–716, 2009.

[24] B. Song, J. M. Zhu, and X. Li, "The research of enterprise financial early warning based on big data," *Journal of Central University of Finance & Economics*, no. 06, pp. 55–64, 2015.

[25] A. n Du, B. X. Fang, and M. Z. Hu, "A decision tree-based web mining system for Chinese pages," *Advances of Search Engine and Web Mining in China [M]*, (in Chinese), p. 03, Higher Education Press, Beijing, Higher Education Press, 2003.

[26] H. Garcia Molina, "Crawling the hidden web," in *Proceedings of the 27th International Conference on Very Large Data Bases*, vol. 09, San Francisco CA USA, September 2001.

[27] X. B. Tang and G. C. Liu, "Research review on fine-grained sentiment analysis," *Library and Information Service*, vol. 61, no. 5, pp. 132–140, 2017.