Research on Financial Risk Early Warning Model Based on Big Data

Fenglian Zhou

Sichuan Vocational College of Cultural Industries, Chengdu, China
741924611@qq.com

Keywords: Big Data, Financial Risk, Early Warning Model, Decision Tree

Abstract: The application of big data (BD) technology can help enterprises improve the processing efficiency of financial data(FD), enhance the persuasion of decision-making information and enhance the use value of FD. It can monitor the financial situation of enterprises in real time, analyze the financial risk(FR) faced by enterprises, and give early warning of potential FR, so that enterprises can formulate strategies to deal with FR in time. This paper mainly studies the FR early warning model based on BD. By analyzing the relevant theories of FR, this paper studies the impact of BD on enterprise FR. A decision tree FR early warning model based on BD is proposed, and the data of listed companies are selected to train the early warning model. From the model training results, we can know that the BD FR early warning model proposed in this paper has good prediction performance and can help enterprises deal with FR timely and effectively.

1. Introduction

In recent years, the word BD more and more by people, with the development of the society and the Internet widely applied, BD arises at the historic moment, because of its diversity, access speed, acquisition and storage characteristics such as low cost, wide sources, can be applied to various fields, the study of large data has been increasingly in-depth. BD makes it possible for enterprises to achieve efficient and high-quality financial management. BD can not only improve the use value of financial data, but also improve the processing efficiency of financial data and provide managers with useful information for decision-making [1-2]. In BD environment, a series of changes have taken place in the financial risk of the enterprise management, information access to become widely make the enterprise can have more reference in decision-making, access to information faster enable enterprises to focus on real-time have an impact on financial risk in all kinds of information, the application of cloud computing, mobile computing technology to improve the efficiency of the financial analysis, at the same time, Applications of BD is of great significance to the enterprise financial risk early warning work, companies use data mining technology, in the vast amounts of data in the selected valuable information to the enterprise, and analyzing these information, predict potential risks according to the results of the analysis, this method is more and more applied to our country enterprise financial management, It helps enterprise managers to more accurately grasp the current situation of enterprises and take timely countermeasures to reduce the possibility of financial risks and minimize losses [3].

Foreign scholars have carried out early research on FR early warning and achieved outstanding results. In terms of dealing with FR, some foreign scholars put forward systematic practical suggestions and sorted them into books for publication. He stressed that in practical management activities, enterprises need to understand risk, manage risk and improve the quality of data information based on different dimensions, while simple data classification methods will not achieve ideal risk management results [4]. Other scholars emphasized the importance of strengthening management based on the problems existing in the enterprise's FR, and required to establish an independent risk management department within the enterprise, establish and improve the supervision and monitoring mechanism, and improve the enterprise's ability to predict and recognize the risk, so as to facilitate the discussion and formulation of solutions in advance [5]. Some scholars in order to solve the problem of low accuracy and poor adaptability of traditional financial analysis methods. A FR assessment model based on BD is proposed. This method uses quantitative analysis method to analyze the explanatory variable model and control variable model of FR assessment. Simulation results show that this method can provide high accuracy of FR [6]. According to the above literature, it can be concluded that the research of foreign scholars on enterprise FR early warning has been very sufficient and the research conclusion is very mature, but the literature of domestic FR early warning is relatively few.

This paper constructs the company's FR early warning model based on BD, in order to realize the early warning analysis of the company's FR status, which will help to provide some methods and ideas for relevant research.

2. Role of BD in FR Early Warning

2.1 FR Related Theory

FR are ubiquitous in the daily operation of enterprises, mainly including unreasonable business approval process, chaotic management structure, improper financial management and inappropriate forms of investment and financing. In terms of capital operation, FR can be divided into five types:

(1) Financing risk: the risk of whether the enterprise can reasonably use funds to create more value to provide the sustainable development of the enterprise, and whether it can repay debts and pay interest on time.

(2) Investment risk: due to uncertain factors in the internal and external environment of enterprise development, the expected return of enterprise investment has uncertain FR.

(3) Liquidity risk: whether the enterprise has sufficient operating cash flow may lead to risk in the development and production of the enterprise.

(4) Income distribution risk: due to the relative contradiction between managers and shareholders on the development of the enterprise, the enterprise managers may make mistakes in the distribution scheme specified for profit distribution and affect the enterprise's decision-making.

(5) Business risk: the uncertainty risk caused by whether the enterprise purchases appropriate raw materials, sells goods and sells on credit. The risk of enterprise operation and management may be in all aspects. Only by fully understanding the possibility of risk in the enterprise operation mechanism and controlling the occurrence of risk at all times can the enterprise develop healthily.

Enterprise financial crisis is a process of gradual accumulation, with uncertainty and predictability. Through prediction, it can give clear crisis warning to enterprise managers, so that enterprises can adjust their business strategies, so as to avoid the occurrence of financial crisis or reduce the intensity of financial crisis. Enterprise financial early warning is to analyze and judge enterprise financial activities, establish statistical early warning model by using FD, find potential risk in enterprise operation, and make managers adjust business strategies [7-8].

Enterprise financial crisis is the result of the comprehensive action of many factors. The external

environment and internal organizational structure of the enterprise management system interact and jointly affect the business process of the enterprise. Enterprise managers can avoid financial crisis only if they understand the integrity of the business system [9-10].

2.2 Impact of BD Environment on FR

(1) Change of evaluation method

In the BD environment, massive data enables enterprises to conduct in-depth analysis of FR and evaluate them from various levels and perspectives based on a more macro, comprehensive and multi-dimensional level. In the BD environment, the evaluation of FR is no longer only through the analysis of various financial indicators, using the traditional variable model to consider and finally draw a conclusion. Instead, various external factors are also included in the risk evaluation system. Instead of relying only on the analysis of structural data, non structural data will be more valuable. In the BD environment, enterprises comprehensively evaluate the main incentives, elements and impacts of FR, draw more accurate conclusions, rank them according to the severity of risk, take different measures, focus on the risk with the greatest negative impact, and implement targeted strategies based on the analysis conclusions.

(2) Recasting of prevention and Control Countermeasures

Accurately predicting risk, changing strategies in time or taking corresponding measures are effective means to prevent risk. Under the BD environment, how to establish an effective FR early warning system is the focus of current prevention and control work. The construction of enterprise risk early warning system should be based on BD technology, and apply the model to various economic activities of enterprises, timely identify potential FR, process and analyze enterprise related data based on data mining, so as to provide valuable reference for decision-making, so as to reasonably avoid risk. Under the background of BD, it is very necessary to study how to improve the accuracy of risk control for the healthy development of enterprises. The proposal of risk prevention and control countermeasures should be based on data mining, data analysis and mathematical statistics. Only in this way can the conclusions be more targeted and accurate. In addition, the judgment of the external environment should also be based on BD technology, based on the evaluation of external data, and timely formulate prevention and control countermeasures for current problems.

(3) Characteristics of financial management

The accuracy of financial information management is high. In the past, due to the insufficient level of financial management of the enterprise, the lack of consolidated data classification standards and accurate analysis results, was unable to provide a basis for decision-making by managers, and the value of the information resources could not be fully utilised. After the preparation of financial statements, the enterprise's FD had no use value. However, in the BD environment, data mining and data processing technologies provided technical support for enterprises to comprehensively process data, After comprehensively analyzing all kinds of information, the accuracy of analysis results can be improved, which greatly improves the accuracy of financial management.

The efficiency of financial management of enterprises is higher. In the environment, businesses have more channels to obtain information and the speed of data update is accelerating. The technology-based data analysis system enables efficient data processing. At the same time, it also reduces the pressure of financial personnel on monitoring and analyzing data, provides more time to think for business managers before making decisions, and significantly improves efficiency.

Data is more accessible and more accessible. In addition, with the development of the financial environment, enterprises can no longer acquire more basic data from the financial environment,

such as large data mining. In addition, with the development of the financial environment, enterprises can no longer collect more basic data from the financial environment, The process of data collection is more efficient, so it is easier for enterprises to obtain the required data.

3. Early Warning Model based on BD

3.1 Construction of Decision Tree Model

Decision tree algorithm is a mainstream supervised machine learning algorithm. It can construct a set of decision variables through learning a group of training samples. It has the advantages of high speed, high precision and easy to understand. Therefore, it has strong applicability in dealing with binary classification problems. The main operation steps are divided into the generation of decision tree and the pruning of decision tree.

In the decision tree algorithm, the minimum number of samples of the number of leaf nodes can significantly affect the complexity of the model. Therefore, it is necessary to find the optimal minimum number of samples first. Because the total number of training samples is 10000, this paper sets the upper limit of the minimum number of samples to 1000, regroups the training sample set, and tests the efficiency of the decision tree through cross validation to obtain the optimal parameter value.

After setting the optimal parameters, in order to improve the prediction efficiency outside the sample and avoid over fitting, it is necessary to prune the optimal decision tree to improve the effectiveness and robustness of the early warning model. The algorithm steps of pruning are as follows:

set up $k=0, T=T_0, \alpha=+\infty$;

Calculate C (Tt), |Tt| for each internal stage T from bottom to top, and:

$$\begin{cases} g(t) = \dfrac{C(t) - C(T_t)}{|T_t| - 1} \\ \alpha = min(\alpha, g(t)) \end{cases} \quad (1)$$

Where Tt represents the subtree with t as the root node, C (Tt) is the prediction error of training data, and |Tt| is the number of leaf nodes of Tt.

Top down access to internal node T, if g (t)= α, Prune and determine the class of leaf node T by majority voting to obtain the number t. set up

$$k = k + 1, \alpha_k = \alpha, T_k = T \quad (2)$$

If T is not a tree separately composed of root nodes, return to step (3);
The cross validation method is used to select the optimal subtree in the subtree sequence.

3.2 Sample Selection and Data Source

Research on the annual financial statements and other sources of CSR of the listed company. "ST" shares refer to the shares in which China Stock Exchange conducts special treatment for financial or other abnormalities of all listed companies. This paper takes whether the listed company is st as the standard of whether the enterprise has FR.

In this paper, the sample of normal listed companies selects the companies that have never had st events in the listed companies this year, excludes the listed companies in the financial industry, and selects 942 companies with a market value of less than 3 billion. The sample of abnormal

companies (i.e. companies with FR) selects companies in recent three years and deletes some samples. Finally, there are 204 remaining companies. The excluded samples include: delete the audit reports with negative opinions and unable to express opinions issued by the audit companies, and the listed companies that fail to disclose the company's financial annual report within the legal period and are subject to st for two consecutive years.

4. Training Results of FR Early Warning Model

Use the initial data set prepared earlier, generate the final sample set through data preprocessing and Feature Engineering, and then divide the test set and training set, as shown in Table 1, which are the test set and training set divided in this paper.

Table 1: Test Set and Training Set

	Number of samples	Abnormal sample	Normal sample
Training set	942	37	905
Test set	204	8	196

Based on the training set and test set divided in Figure 1, a stable decision tree model is finally obtained through continuous iteration and hyperparametric optimization.
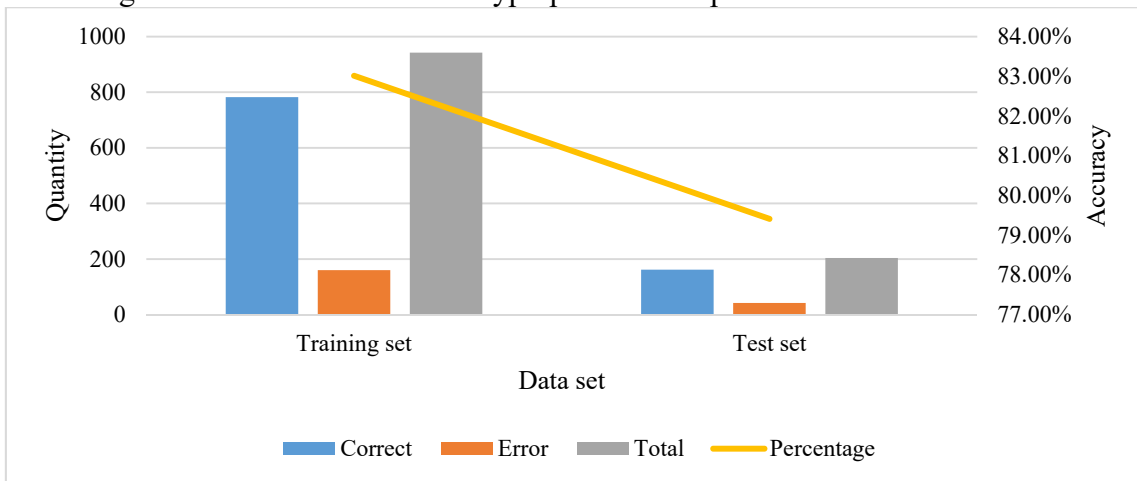
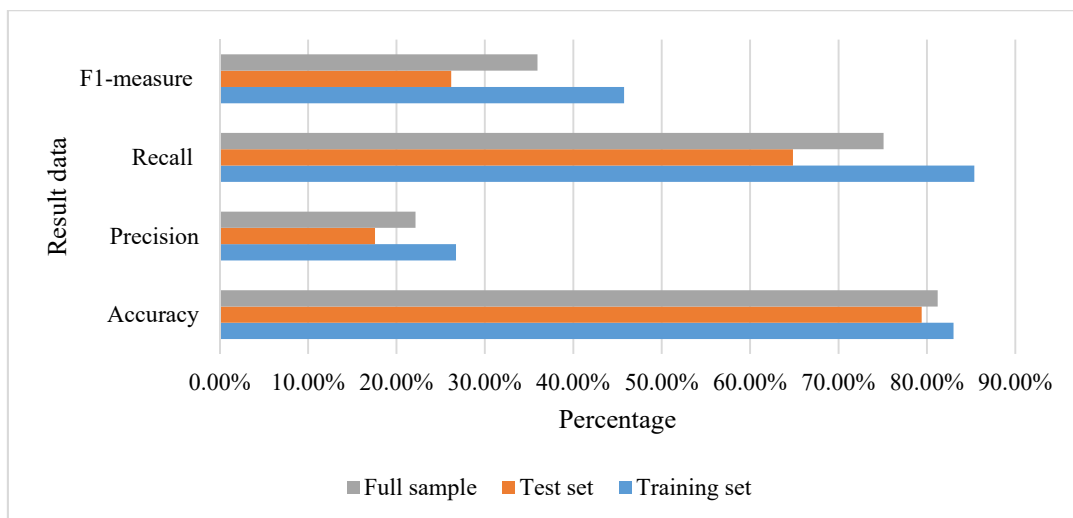

Figure 1: Accuracy of decision tree model

Figure 2: Performance evaluation of decision tree model

As shown in Figure 2, the accuracy of decision tree model in training set and test set is 83.01% and 79.41% respectively.

BD analysis provides a set of scalable and deepening methods for the construction of enterprise risk early warning model, and can reveal the value of data with high quality and efficiency. This paper introduces the BD analysis method into the enterprise FR early warning, constructs the enterprise risk early warning index system and establishes the risk early warning model, so that enterprises can carry out risk early warning more scientifically and effectively. Because the collected samples are not comprehensive and the amount of data is not sufficient, the model needs to be further optimized and verified. At the same time, for each business link and data involved in the enterprise, it is necessary to use the method of BD for risk early warning, and constantly explore and try the application of various BD analysis methods. In order to better achieve the goal of risk early warning, enterprises need to further optimize their governance structure, institutional arrangements, tax planning, corporate culture, personnel control and so on.

5. Conclusions

With the expansion of the scale of e-commerce and express the tells the development of the industry, city logistics distribution business also got great development, urban distribution business needs "loose" and "small", the urban road traffic I complex, how to scientifically plan distribution path has become each big city to the enterprise important means to reduce the cost, improve market competitiveness. In this paper, the method of logistics distribution path is optimized by improving the quantum genetic algorithm, and the physical distribution path planning system is designed. However, due to the limited time and personal ability, some work still needs to be further studied, which can be improved from the following aspects in the future: This paper considers that customer service is not in order, but in real life, important customers need to be served first, which can be used as a research direction.

References

[1] Rodriguez Gonzalez M, Basse T, Kunze F, et al. Early warning indicator systems for real estate investments: Empirical evidence and some thoughts from the perspective of FR management. Zeitschrift für die gesamte Versicherungswissenschaft, 2018, 107(4): 387-403.

[2] Ashraf S, GS Félix E, Serrasqueiro Z. Do traditional financial distress prediction models predict the early warning signs of financial distress? Journal of Risk and Financial Management, 2019, 12(2): 55.

[3] Hidayat T, Masyita D, Nidar S R, et al. Early Warning Early Action for the Banking Solvency Risk in the COVID-19 Pandemic Era: A Case Study of Indonesia. Economies, 2021, 10(1): 6.

[4] Diks C, Hommes C, Wang J. Critical slowing down as an early warning signal for financial crises? Empirical Economics, 2019, 57(4): 1201-1228.

[5] De Nicolò G, Lucchetta M. Forecasting tail risk. Journal of Applied Econometrics, 2017, 32(1): 159-170.

[6] Nurfalah I, Rusydiana A S, Laila N, et al. Early warning to banking crises in the dual financial system in Indonesia: The markov switching approach. Journal of King Abdulaziz University: Islamic Econ., 2018, 31(2): 133-156.

[7] Utami W, Nugroho L, Mappanyuki R, et al. Early warning fraud determinants in banking industries. Asian Economic and Financial Review, 2020, 10(6): 604-627.

[8] Holopainen M, Sarlin P. Toward robust early-warning models: A horse race, ensembles and model uncertainty. Quantitative Finance, 2017, 17(12): 1933-1963.

[9] Ridwan M, Ulupui I G K A, Ahmad G N. The impact of total equity, risk-based capital ratio, adequacy investment ratio, direct premium receivable ratio, liquidity ratio, and loss ratio toward the early warning system of general insurance. The International Journal of Social Sciences World (TIJOSSW), 2022, 4(1): 11-27.

[10] Gupta N, Kumar A. Artificial neural networks for developing early warning system for banking system: Indian context. International Journal of Economics and Business Research, 2022, 23(2): 229-254.