Development and Application of Enterprise Financial Risk Analysis System Based on Data Mining Technology

Ming Liu, Guoqiang Wu*

School of Economy and Management, Huainan Normal University, Huainan City, Anhui Province, 232038, China

32610899@qq.com; *442134376@qq.com

Abstract. With the rapid rise of the digital economy, the overall informatization level of enterprises is getting higher and higher, and massive data information is transformed into productivity, which gives enterprises opportunities for development and brings new challenges to internal financial management. Faced with numerous financial risk problems, traditional risk management tools and technologies have obvious shortcomings in application scope, identification efficiency and control ability. In this regard, this paper puts forward a set of construction scheme of enterprise financial risk analysis system based on data mining technology, aiming at making use of the practical advantages of digital information technologies such as big data and machine learning, and putting forward new solutions for enterprise financial risk management. The system takes Hadoop cluster as the data management and processing server, MapReduce as the data mining engine, and combines Javaweb technology to form a comprehensive application service platform integrating online application, intelligent processing, visual analysis and other functions. Practice has proved that the system constructs the corresponding enterprise financial risk identification and measurement model through data mining algorithms such as Logistic, which meets the enterprise's demand for financial risk management and improves the enterprise's ability to resist risks.

Keywords: data mining; Enterprise financial risk; Machine learning algorithm; Hadoop; Computer software application

1 Introduction

In the era of digital economy, a series of new digital technologies, such as Internet, big data, cloud computing and artificial intelligence, have reshaped the whole process of information collection, transmission, storage, analysis and application, and also transformed massive data information into important production factors to promote economic and social development, and promoted the transformation and upgrading of industrial structures in various industries and fields. [1] At the same time, the changes in the external environment of enterprises and the adjustment of internal control methods can easily expand the financial risks of enterprises, thus affecting the normal

A. Rauf et al. (eds.), *Proceedings of the 3rd International Conference on Management Science and Software Engineering (ICMSSE 2023)*, Atlantis Highlights in Engineering 20,
https://doi.org/10.2991/978-94-6463-262-0_9

operation and operation of enterprises. Faced with this situation, enterprises usually adopt traditional financial risk management methods, and use manual operation to analyze financial statements, financial indicators or Delphi survey to realize financial risk identification, evaluation, control and supervision. However, in the traditional mode, financial risk management focuses on the content of financial statements, and the data and content abundance have obvious limitations, and the work efficiency also lags far behind the actual application requirements. [2] In view of this, this paper believes that enterprises should re-examine the changes and innovations brought about by digital transformation and upgrading, reposition the functions and objectives of financial management, and promote the transformation of enterprise financial management mode to digital and intelligent. [3] For enterprise financial risk, we will give full play to the application advantages of big data technology, data mining technology and computer application technology, and build an enterprise financial risk analysis system with Hadoop cluster as the data analysis and processing server and Javaweb as the interactive application system, and put forward a set of practical and comprehensive solutions from the aspects of data and content sources, risk handling tools and technologies, and overall business process control. It enhances the enterprise's ability to manage financial data and deal with risks, and forms a working mechanism integrating financial risk identification, evaluation and measurement, control and feedback, which makes a useful attempt to improve the effectiveness of enterprise internal management.

2 System construction

Complete the configuration and deployment of the overall development environment according to the actual application requirements of the enterprise financial risk analysis system. The content of system development is divided into two parts. One is to combine Hadoop and Javaweb, two complementary powerful tools, to build the structural framework of the system. The second is the development and deployment of each functional service module, and the logical relationship between them is determined to complete the packaging and release of the whole system.

Firstly, the data management server is built based on Hadoop framework. Hadoop framework is deployed in cluster, involving seven functional nodes, named as Master1, Master2, Slave1, Slave2, Slave3, Slave4 and Slave5. The hardware configuration of each functional node includes two 4-core CPUs with a main frequency of 2.5GHz, 4TB disk drives, 36GB memory and Gigabit Ethernet connection to meet the distributed storage requirements of various types of data. [4] As for the software program, Linux is selected as the bottom operating system of each node, and the version is CentOS 6.8 (x86_64), jdk-1.8.0_201-linux-x64 and Hadoop is selected as v2.6.1, which are installed on each node respectively. Yarn, HDFS, Zookeeper, HBase, Kafka, Flume, Sqoop and other components are also deployed in each node, thus forming the foundation of enterprise financial data information collection, transmission, storage and management.

Secondly, the Web application design of the system will mainly rely on the Javaweb technology system. Under the Javaweb technology system, the display layer is the user interactive interface, which is led by JSP technology and supplemented by HTML, CSS and JavaScript to complete the page development and deployment. The business logic layer aims to complete the definition and declaration of various functions of the system. The data access layer pays more attention to all kinds of data interfaces and method classes, which provides convenience for data calling and processing. [5] Java is selected as the basic development environment of Web application, MyEclipse V 2022 as the integration tool, Tomcat 8.0 as the Web server and MySQL 5.7 as the database server. Complete the configuration of Tomcat in the Preference option under MyEclipse. Then, based on the Spring architecture, the integration and encapsulation of the whole system are completed.

Finally, the development of online application, intelligent processing, visual analysis and other functional applications of the system core will be fully combined with the structural framework of the system. Online application depends on the call and control of business logic layer, intelligent processing needs MapReduce distributed computing framework under the data management server as the engine, and visual analysis will choose D3.js to complete the deployment in the display layer of the application. Among them, many data mining algorithms involved in intelligent processing should be defined and trained under each node in Hadoop cluster in advance, and encapsulated into a Hadoop task class. The system can run the main method of this class without jar and pass the necessary parameters to it, so that users can call MapReduce jobs from Web applications with Java API to complete the analysis and mining operation of accounting data information. [6] Through the introduction of the above key technical theories, the overall environment of system development, the configuration of related software and tools are determined, and the technical feasibility of the overall project of enterprise financial risk analysis system is also clarified.

3 Functional implementation

3.1 Selection of financial data samples

The system can rely on Hadoop data management server to connect business data and financial data of enterprises, and collect all kinds of data including financial system, business system and other external sources, thus providing comprehensive data support for financial risk management and control. [7] Massive data information will be input into HDFS through collection technologies such as Sqoop, Flume and Kafka to complete distributed storage, which greatly improves the data capacity and strengthens the transmission and sharing of data information. Before initiating enterprise financial risk management, users can select samples according to key attributes such as year, department and category.

3.2 Financial risk identification

Under this function module, enterprise risk management will be completed according to the steps of risk indicator selection, risk identification model selection and risk identification implementation. First of all, risk indicators will be selected according to the actual operating conditions of the enterprise and the abundance of sample data. Based on the traditional financial indicators, the system adds cash flow indicators and non-financial information indicators to form a brand-new financial risk identification index system, as shown in Table 1. Among them, profitability, solvency, operational ability, development ability and cash flow ability are financial indicators, while ownership structure, agency status and audit opinions are non-financial indicators.

Table 1. Financial risk identification index system

Profit	Solvency	Operational	Development	Cash flow	Other
Profit margin A_1	Liquidity ratio A_3	Inventory turnover ratio A_7	Continuous growth rate A_9	Cash flow ratio A_{13}	Share ratio of major shareholders A_{16}
Net profit margin A_2	Quick ratio A_4	Fund turnover ratio A_8	Profit growth rate A_{10}	Income cash ratio A_{14}	Management fee rate A_{17}
	Debt ratio A_5		Income growth rate A_{11}	Cash recovery rate A_{15}	Opinion type A_{18}
	Capital loan ratio A_6		Capital accumulation rate A_{12}		

There are many indicators in the financial risk identification index system, so it is necessary to further screen out indicators with strong sensitivity and high correlation to improve the working efficiency of the financial risk identification model of the system. After T-test and nonparametric test, 10 indicators, including A_1, A_2, A_6, A_8, A_9, A_{12}, A_{13}, A_{16}, A_{17} and A_{18}, are finally selected as the input values of the risk identification model.

Secondly, the system will automatically input the selected 10 indicators into the identification model, and select the forward step-by-step method to fit the model according to the maximum accuracy of the band. As shown in Formula 1, it is a Logistic regression model, e is a natural logarithm, which is about 2.7182818, and Z is the sum of the product of the input indicator value and the weight variable. [8] For the output result of Logistic regression analysis model, take 0.5 as the benchmark value, when the output value is greater than 0.5, it can be concluded that this indicator will face financial risks, otherwise, this indicator is currently in a normal state.

$$P = \frac{e^Z}{1+e^Z}, Z = -1.231 + 1.574A_6 - 3.774A_1 + 3.244A_8 - 3.731A_9 \tag{1}$$

Finally, in order to verify the system's identification results of financial risks, a set of comparative experiments will be constructed to calculate the accuracy of Logistic regression analysis model. T represents the year in which financial risks occur, and

the results of judging accuracy are shown in Table 2. The results show that the system functions have obvious advantages in accuracy and can identify the financial risks of enterprises.

Table 2. Accuracy of the Logistic regression analysis model

Year	Accuracy rate	Traditional mode verification accuracy rate
T-3 year	70.03%	66.35%
T-2 year	89.25%	80.42%
T-1 year	94.13%	90.77%
T year	99.33%	92.64%

3.3 Financial risk assessment

After the identification of enterprise financial risk is completed, the enterprise financial risk will be further evaluated as a whole and the size of financial risk in specific links will be determined. [9] On the basis of risk identification, the system uses factor analysis to build a financial risk assessment model and automatically calculates the score coefficient of each factor, as shown in Table 3.

Table 3. Score matrix of each analysis index

	Profit	Debt service	Development	Operating	Cash flow
A_1	-0.031	0.337	0.016	0.026	-0.034
A_2	-0.036	0.345	0.020	0.021	-0.015
A_6	-0.042	0.344	0.015	0.016	0.064
A_8	0.279	-0.008	-0.054	-0.009	0.036
A_9	0.285	-0.048	-0.051	0.026	-0.019
A_{12}	0.274	0.002	-0.052	0.006	0.018
A_{13}	-0.020	0.047	0.016	0.485	0.126

According to the score matrix of analysis indicators, the principal component factors F_1, F_2, F_3, F_4 and F_5 are converted into variance contribution rates to get the corresponding weight values, and then the calculation formula of enterprise financial risk assessment is obtained by weighted calculation. [10] As shown in Formula 2, where F represents the comprehensive score for financial risk assessment. In the system, the value of f is preset in four ranges, which correspond to four criteria for judging financial conditions and financial risks, as shown in Table 4.

$$F = 0.3054F_1 + 0.2523F_2 + 0.1772F_3 + 0.1434F_4 + 0.1211F_5 \tag{2}$$

Table 4. Evaluation criteria for the comprehensive score of financial risk assessment

F value	Financial status	Financial risk
F>0.5	Excellent	Low
0<F<0.5	Good	Lower
-0.5<F<0	Ordinary	Higher
F<-0.5	Poor	High

In order to verify the evaluation results of the system on financial risks, a simulation test is conducted with real corporate finance. The test results are shown in Table 5. The results show that the system can complete the evaluation of enterprise risks, and the ability of enterprise financial risk analysis and processing can be further improved by combining the above financial risk identification.

Table 5. Financial risk test results

	F_1	F_2	F_3	F_4	F_5	F
F value	-0.4311	-0.4188	-0.2501	-1.7311	0.9805	-0.4033
Risk assessment	Higher	Higher	Higher	High	Low	Higher

4 Conclusions

In order to promote the reform of enterprise financial risk management mode, this paper constructs a Web-based enterprise financial risk analysis system. The system completes the sharing and interaction of all kinds of data. With a brand-new data analysis system, it promotes the efficiency of financial risk handling and increases the value embodiment and reasonable application of financial data. In the follow-up research, it will further enrich the system's support for other algorithm models and provide necessary technical support for enterprise financial risk management.

References

1. Wang Yilin. Present Situation and Prospect of Digital Economy Development[J]. Contemporary County Economy .2023.05.
2. Liu Qinjin. Present Situation and Measures of Enterprise Financial Risk Management[J]. Modern Economic Information.2019.09.
3. Yang Yang, Zhou Xinyu. Transformation of Enterprise Financial Management Mode under the Background of Digital Economy[J]. Modern Enterprise.2023.04.
4. Zhou Shaoke, Wang Lei, et al. Fully Distributed Cluster Deployment of Big Data Hadoop Technology[J]. Industrial Control Computer.2021.08.
5. Shi Feng. Development and Application of JavaWeb Based on MVC Pattern[J]. Electronic Technology.2021.05.
6. Jiao Xiangyu, Huang Kanghui, et al. Hadoop+JavaWeb Big Data Analysis Visualization System[J]. Management & Technology of SME.2020.02.
7. Ding Yanan. Early Warning and Prevention of Financial Risks in the Era of Big Data[J]. China Market.2022.12.
8. Wang Wanrong. Financial Risk Early Warning of Listed Companies Based on logistic Model[J]. China Water Transport.2021.01.
9. Liu Mei, Yan Ruixia, et al. Financial Risk Assessment of Enterprises from the Perspective of Big Data[J]. Intelligent Computer and Applications.2020.01.
10. Tian Guiyuan. Research on Financial Risk Assessment and Prevention of Company D[D]. Heilongjiang University.2022.05.