

Research Article

A Method for Financial System Analysis of Listed Companies Based on Random Forest and Time Series

Chi Zhang , Huaigong Zhong, and Aiping Hu

Nanjing Audit University Jinshen College, Nanjing, Jiangsu 210046, China

Correspondence should be addressed to Chi Zhang; zhangchi@naujsc.edu.cn

Received 12 March 2022; Revised 3 April 2022; Accepted 11 May 2022; Published 31 May 2022

Academic Editor: Mian Ahmad Jan

Copyright © 2022 Chi Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The world economy has recently moved in a fresh era, where the financial world is rapidly developing. Various economic crises, such as banking, economic, and currency crises, impose high economic costs, and harm the entire society. This necessitates the creation of an early warning system for financial crisis that can be adaptively analyzed using past information. Early warning systems could prevent the occurrence of business and economic crises by providing a systematic prediction of unfavorable events. Early warning systems are mainly used to detect crises before they do damage and to reduce false alarms of impending crises. Because of the above, this paper studies early warning of the financial crisis of listed companies based on random forest and time series. Besides, it constructs a random forest and Boruta-Random forest (BRF) model with Benford factor to deal with the impact of financial data quality on the financial risk early-warning model. Our model can effectively improve the prediction accuracy of the financial early warning model. The experiments show that, in comparison to RF, BRF can increase the accuracy of financial risk early warning, expand the applicability of RF, as well as provide a fresh perspective for research on listed company financial risk early warning.

1. Introduction

A financial crisis is an economic process in which an enterprise fails to refund progressing liabilities or expenditures. A company's business can be greatly expanded through globalization communication and the world Internet. If a company experiences a financial crisis, it implies that the extent of the crisis caused by the crisis is expanding. As a result, developing a successful financial crisis early warning framework becomes critical. Even more significant, the goal is to be able to bring out an early warning, tracking, and resolution of the enterprise's economic crisis in a quick, precise, and timely manner, which has significant theoretical and practical application value. As a result, an early warning system is an essential component of tragedy or other critical event prevention [1]. They can be implemented in any area where it is beneficial to obtain identifiers of future, generally negative, occurrences. The Financial Early Warning Mechanism (FEWM) is a system that provides both warning and control. Modern information technology serves a

significant role in achieving one of the most important goals in the development of FEWM. It identifies and addresses various problems in the process of enterprise financial management at an early and timely stage. In-depth analysis of various financial data and information collected during the enterprise's operation process enables financial problems that occur during the actual operation process. These problems need to be reported to the enterprise's management promptly by issuing a warning, after which an in-depth analysis of the factors that contribute to financial problems is carried out [2]. This allows for the development of effective solutions, which in turn provides a very reliable basis for the organization to make business decisions. To increase their overall competitiveness, modern firms should develop a comprehensive set of FEMs, identify any existing challenges, and propose solutions [3]. China's economic growth has hit a new plateau in recent years, putting a strain on the market competitiveness. It is imperative for businesses that they monitor their development in real-time, comprehend and grasp their financial data, and enhance the efficiency with

which they conduct their operations if they are to remain competitive and gain a firm foothold in the market's tidal wave of change. In terms of risk management capabilities and levels, an enterprise's financial crisis is usually caused by a breakdown in cash flow, failure to pay liabilities on time, or a negative result of the liquidation of the company's assets [4]. If a company experiences a financial crisis and is unable to resolve the problem on time, it may be delisted. To ensure the smooth operation and development of the firm, it is necessary to build a dynamic FEWM and take appropriate antifinancial-crisis measures before the occurrence of the crisis [5].

Even though the government has consistently increased the oversight of publicly traded corporations in recent years, the phenomena of financial fraud continue to occur regularly. We must investigate a more effective FEWM. Financial fraud refers to the behavior of businesses that intentionally overstate assets or underestimate liabilities by manipulating financial indicator data in a variety of ways to falsely claim profits to improve overall performance and appearance. Financial fraud can easily cause financial data to be distorted, the quality of financial data to be reduced, and the prediction effect of FEWM to be severely harmed, all of which have negative consequences. To increase the accuracy of FEWM prediction, it is vital to fully examine the potential influence of financial data quality issues when generating FEWM [6]. FEWM can be classified into two types: single model and combined model, depending on the approach used to generate the model. Discriminant analysis and Logistic regression are examples of classic statistical methods. Machine learning methods such as neural networks, random forest, and genetic algorithms are examples of methods that can be used to create FEWMs. Combinatorial models are those that are created by combining two or more ways to create FEWMs. Some researchers have combined fuzzy mathematical theory with RF to develop a fuzzy RF model that can be used to assess the problem of early warning of corporate financial risk. Several additional researchers merged the concepts of K-fold cross-validation and RF to develop the K-fold RF algorithm, which was designed to improve the selection procedure for financial FEWM. Alternatively, some researchers have combined Benford's law with the logistic model to improve the accuracy of the FEWM forecast [7–9].

Based on the current state of FEWM in the United States and overseas, it appears that the actual early warning effect of a single model is frequently insufficient. The univariate analysis approach, for example, offers just a limited number of financial indicators, making it difficult to anticipate a significant number of financial hazards using it. The predictive validity of the logistic model is highly dependent on the final set of financial indicator variables that are chosen. Therefore, variable selection is essential before modeling can begin. The models used in FEWM research, such as neural networks and genetic algorithms, are highly sophisticated, have large processing costs, and have limited explanatory power. When using an ensemble approach to integrate several weak classifiers and Bootstrap resampling, which overcomes the problem of overfitting and considerably boosts the accuracy of FEWM [10–12], the classification and

early warning impact of RF is more resilient than that of a single decision tree. When compared to a single model, the combined model takes advantage of the advantages of various approaches to maximize the accuracy of FEWM calculations. Therefore, the RF-based combination model will be the primary research emphasis for the FEWM type. The data quality of financial indicators is a critical factor in determining the effectiveness of the financial risk early warning model (FEWM) in terms of prediction. The fact that there are few publications on the combination model that consider financial data quality, and that there is no study on the RF forest model of FEWM that considers financial data quality, should be emphasized [8, 13].

In the financial data quality inspection field, Benford's law is among the most-often utilized methods of examination. The failure of a financial indicator to comply with Benford's law is interpreted as indicating that the indicator carries a high risk of financial fraud and that the financial data may have quality issues. According to some researchers, the application of Benford's law to the field of auditing has been extended, and they have discovered that when the sample size is large enough, the distribution of real financial data is consistent with Benford's law, whereas it is difficult for data containing financial fraud to comply with this law. The law of numerical manipulation of net profit indicators of Chinese listed firms is summarized by other academics, and the efficiency of Benford's law in identifying financial fraud in Chinese listed companies is tested by other researchers as well. Some researchers have created the FEWM, which greatly enhances the accuracy of the financial risk early warning logistic model by taking into account the quality of financial indicators and combining Benford's law with the logistic model. For the sake of conclusion, Benford's Law can be used to successfully identify the possibility of fraud in financial data [14]. Based on the foregoing research backdrop, this study will introduce Benford's law based on the RF model, create the BRF model, and apply it to FEWM research of Chinese A-share and US stock market listed businesses. The key contributions of this paper are listed under:

- (1) First, this paper tests the data quality of financial indicators using Benford's law. After that, it determines whether there is a statistically significant difference among the actual frequency of the first digit of the indicator and the theoretical frequency of Benford's law.
- (2) It constructs the Benford factor to identify sample points that may be at risk of financial fraud. To develop an RF early warning model, this paper uses the Benford factor as a variable and the enterprise financial risk early warning variable index.
- (3) After that, by using empirical analysis, the effectiveness of the BRF and RF models is compared to one another. The empirical findings indicate that BRF can give information for identifying sample locations that are at risk of financial fraud, and that BRF has greater prediction accuracy than other methods.

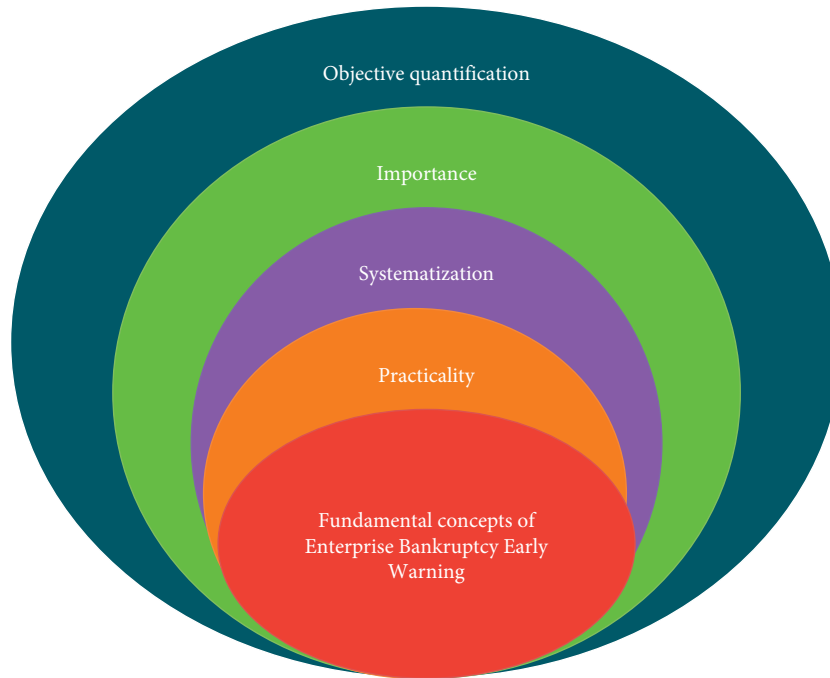


FIGURE 1: Fundamental concepts of enterprise bankruptcy early warning.

- (4) Finally, this work contributes to FEWM research on publicly traded corporations by offering a fresh point of view.

The other sections of this research paper are planned as follows: an overview of enterprise FEWM and its outstanding problems can be explained in Section 2; our planned methodology can be explained in Section 3, while Section 4 consists of our experimental work and simulations and this paper are concluded in the last section such as Section 5.

2. Overview of Enterprise FEWM and Outstanding Problems

2.1. Basic Concept. FEWM, also known as “Enterprise Bankruptcy Early Warning,” is a micro-level extension of financial risk prevention management that is used to detect and avoid business bankruptcy [15]. Accounting and marketing are used by businesses to detect existing or potential financial problems in financial management on time. Advanced enterprise management concepts based on ratio analysis methods and other methods are used to provide corporate management with a premature warning signal system as well as to provide a relevant basis for corporate decision-making oversight system. Through comprehensive analysis of changes in corporate financial-related indicators, forecasts, and reflection of corporate operating conditions, early warning signal systems are issued to corporate management, and the corresponding basis for the corporate decision-making monitoring system is provided. Figure 1 depicts the four fundamental concepts that enterprises should follow throughout the process of developing and implementing a FEWM.

2.1.1. Practicality. One of the most important goals of developing FEWM is to provide an early warning system if problems arise during the organization’s actual operation and management process. Therefore, a high sensitivity and strong practicability should be the fundamental principles of enterprise-wide financial risk management.

2.1.2. Systematization. The current scenario of an enterprise’s development is the essential basis for developing FEWM. Take into account the overall situation and develop a set of systematic FEWM, thereby contributing to a significant improvement in the enterprise’s management level.

2.1.3. Importance. When developing a FEWM, businesses should distinguish between primary and secondary contradictions, understand the major issues, and apply targeted treatment strategies. The primary goal of prevention should be highlighted, and consideration should be given to the cost-effectiveness of interventions.

2.1.4. Objective Quantification. When businesses are constructing FEWM, they must include relevant indicators in their plans. These indicators are also helpful in ensuring the effectiveness of the FEWM. Because of this, it is possible to achieve measurable FEWM indicators, reflect and display forecast data intuitively, and so encourage the better development of organizations more efficiently.

2.2. Outstanding Issues. In the current time, there are still numerous lingering problems in the FEWM mechanism as shown in Figure 2 that must be addressed as soon as possible.

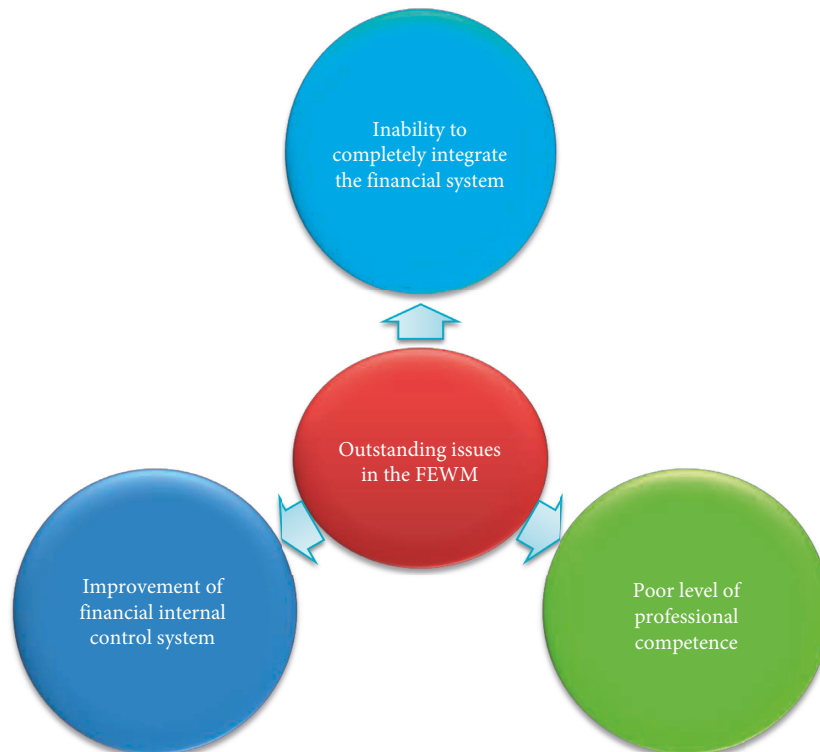


FIGURE 2: Outstanding issues in the FEWM mechanism.

These issues are mostly reflected in the following components of the mechanism.

2.2.1. Inability to Completely Integrate the Financial System with the Early Warning System. The financial information of an organization can be continuously improved by allowing the financial operations of an enterprise based on modern information technology. It promotes the exchange of internal data and information, primarily enterprise financial, manufacturing, and sales data. However, in many firms, FEWM has not been able to provide the aforementioned information more comprehensively, making it impossible for FEWM to fulfill its proper role in these organizations.

2.2.2. There Is a Poor Level of Professional Competence among Financial Management Staff. When it comes to the actual procedure, enterprise fund management is a critical duty. It is necessary due to the enormous number of linkages, significant coordination among financial management personnel, and the need to master a wide variety of data. It shows that efforts should be made to raise the degree of professional excellence among those in charge of financial management. However, a significant issue now facing business finance management staff is that the quality level of many of them is generally low, and they have not undergone systematic and comprehensive professional training.

2.2.3. The Financial Internal Control System Must Be Improved and Perfected Even Further in the Future. As a result

of the widespread adoption of modern technology in enterprise capital management, internal control is now exposed to significant risks. Internal control in modern organizations has gradually evolved from manual control to information control as the financial information system of the company has improved. In the current period, the financial internal control of contemporary organizations has steadily shifted away from the conventional method of checking accounts and toward the current financial management system control, which was previously used. In this regard, the internal control of modern organizations has experienced significant changes not only in terms of structure but also in terms of material and methodology, resulting in a significant increase in financial risks. Modern enterprise financial internal control necessitates the proper integration of the financial system with other departmental management systems in order to facilitate the enterprise's integrated management. Aside from that, the quality of corporate financial management staff is generally considered to be inadequate.

The critical significance that financial risk relationships play in the operation of businesses is not well understood. The person in charge of the enterprise must take a strong leadership role in the procedure of developing a financial early warning mechanism. The enterprise should also take proactive and effective measures to further improve and perfect its financial early warning mechanism while also providing timely feedback to the enterprise's management or decision-making level on any problems that are discovered during its construction. This is necessary in order to ensure that each business activity of the enterprise operates

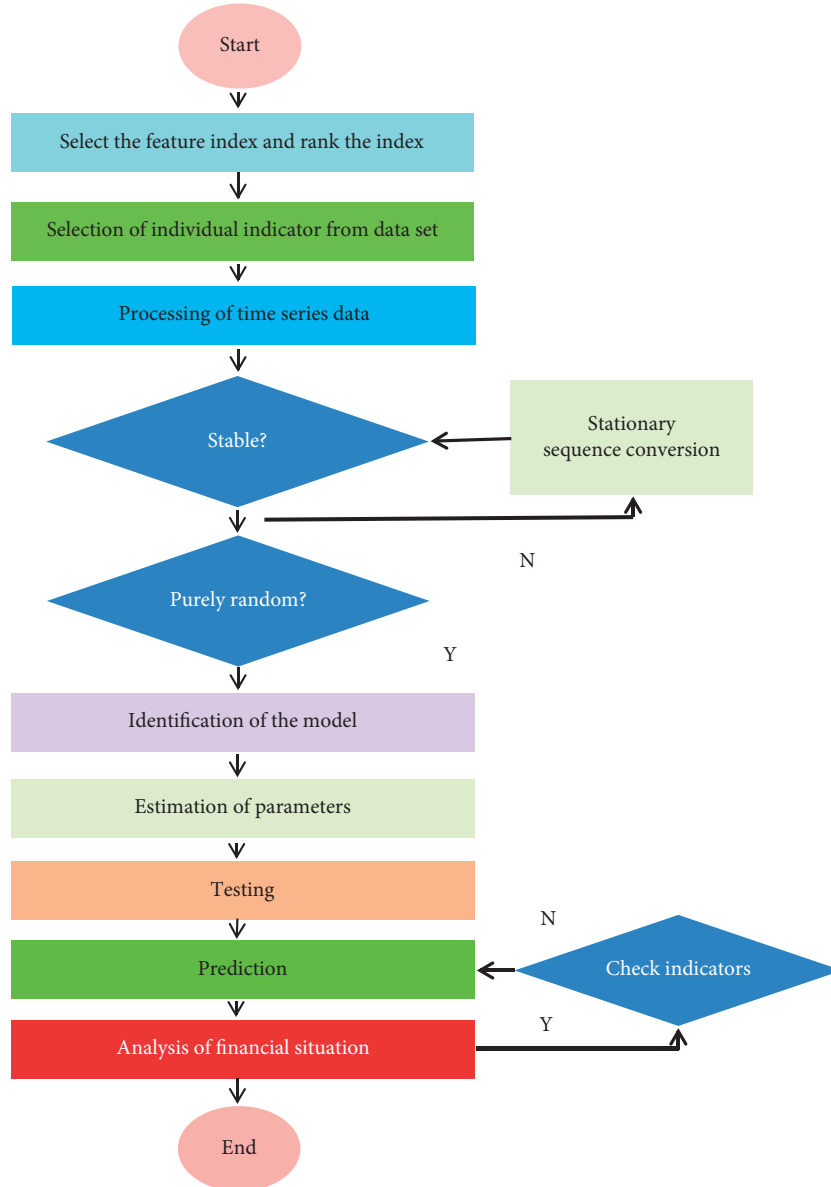


FIGURE 3: Flowchart of our early warning algorithm for financial crises.

efficiently. As a result, efficient actions can be developed in a timely manner to successfully prevent the occurrence of various risks and, eventually, to ensure that the organization continues to operate normally.

3. Boruta-Random Forest (BRF) Method

Figure 3 explains the flowchart of our proposed model. Our model selects the feature index and rank index. After this, it processes time-series data by selecting the individual indicator from our data set and checking whether it is stable or not. If it is not stable then stationary sequence conversion is performed, otherwise, it checks whether it is purely random or not. If it is purely random then the model estimates the parameters by identifying the model and predicting after testing. Finally, it checks analysis of the financial situation; if

it is not then prediction performs again otherwise, analysis is done.

This intrinsic rule of data states that the likelihood that the first digit of all-natural data is one of the numbers from one to nine is stable and that the probability distribution of this probability is on a monotonically decreasing trend. The likelihood that the initial digit D is the letter d is as given in equation.

$$F_{B,d} = \ln\left(1 + \frac{1}{d}\right). \quad (1)$$

It was discovered that the first digit of many financial data sets obeyed Benford’s rule, but that the first digit of financial data sets with high-fraud risk did not comply to Benford’s law very often. If the distribution law of the first digit in financial data differs greatly from Benford’s law, it

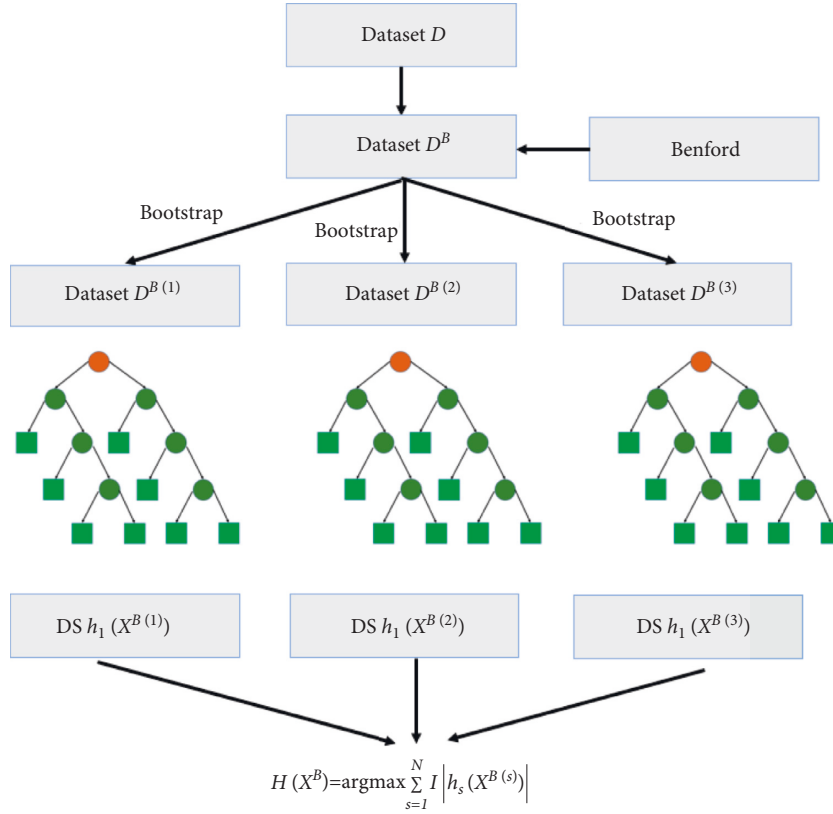


FIGURE 4: Structure of BRF.

TABLE 1: FEWC symbols.

Class	Symbol	Definition
Profitability	X_1	Net profit/(net assets + total assets)
Solvency	X_2	Present assets/current liabilities
Growth ability	X_3	(Total assets at the end of the year - total assets at the beginning of the year) / total assets at the beginning of the year
Operating ability	X_4	Main business income / (fixed assets + total assets)
Cash flow	X_5	Liquidity / current liabilities

indicates that the data has been manipulated or tampered with, that the quality of financial data is poor, and that the financial risk is high as a result. It is the most commonly used method to define if the distribution of the initial digits of a collection of data complies with Benford's law, and it is also the most accurate method. The test statistic is given in equation (2):

$$\chi^2 = N \sum_{d=1}^9 \left[\frac{(F_d - F_{B,d})^2}{F_{B,d}} \right], \quad (2)$$

where N is the overall amount of samples, and F_d denotes the frequency of occurrence of the initial digit d of the information to be checked. If $\chi^2 > \text{Critical}$, the null hypothesis is rejected, and the frequency of the first digit in the data set does not satisfy Benford's law, the data set may have been manipulated

or tampered with. It is impossible to determine whether a given sample point is at risk because this test method can only evaluate the overall quality of the data. Referencing the existing research literature, the Benford factor is developed by Benford's law to determine whether or not there is a data quality problem in a certain sample point, as well as the likelihood of financial fraud occurring in that sample point.

Let $X_j \{j = 1, 2, \dots, k\}$ represent financial indicator variables. Let $r_d^{(j)}$ represent the difference between the observed frequency of the first digit d of the indicator X and the theoretical frequency of Benford's law, then the expression of $r_d^{(j)}$ is given in equation (3):

$$r_d^{(j)} = F_d^{(j)} - F_{B,d}^{(j)}, \quad (3)$$

where $\sum_{d=1}^9 F_d^{(j)} = 1$, $\sum_{d=1}^9 F_{B,d}^{(j)} = 1$, then we have

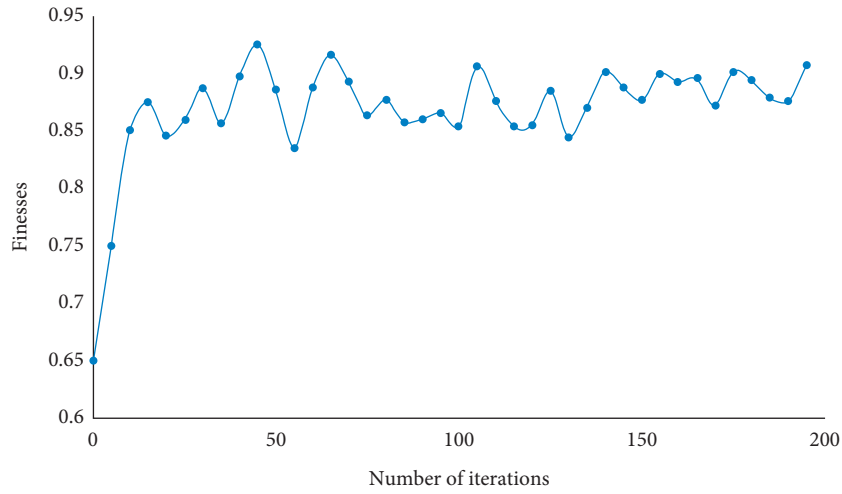


FIGURE 5: Model parameters tune chart.

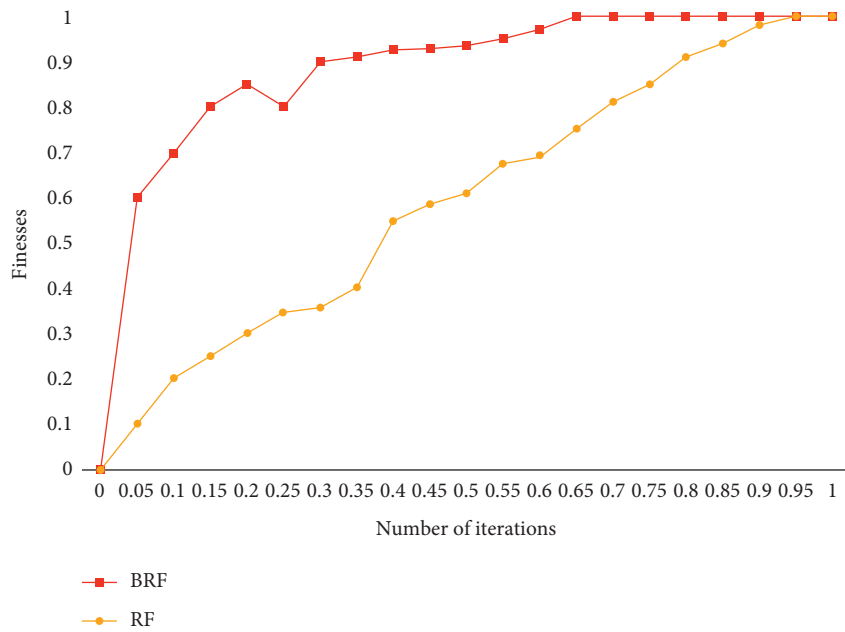


FIGURE 6: AUC results.

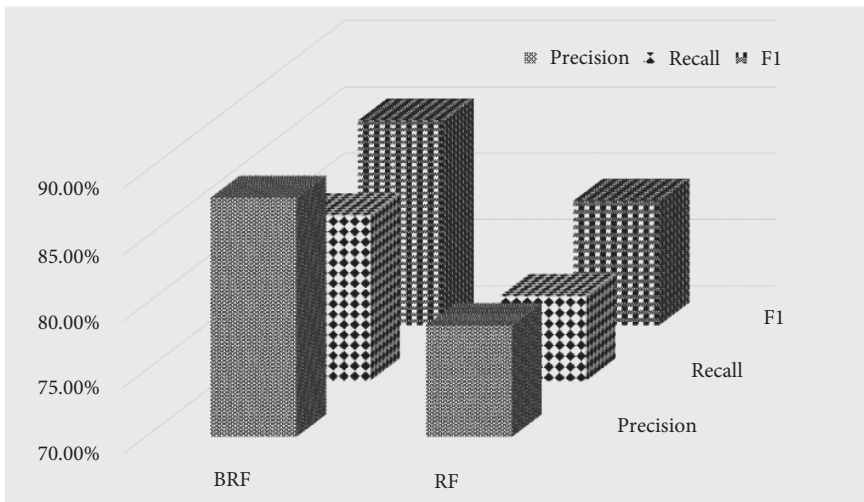


FIGURE 7: Comparison of prediction effects of different models.

$$\sum_{d=1}^9 r_d^{(j)} = 0. \quad (4)$$

When the data is modified or edited, the observed frequency of the first digit of indicator X_j $\{j = 1, 2, 3, \dots, k\}$ will be different from the theoretical frequency. The larger the absolute value of the difference $r_d^{(j)}$, the higher the risk of financial fraud at the sample point. This paper focuses on the first digit with the largest absolute value of the difference between the observed frequency and the theoretical frequency. Let the number with the largest absolute value of the first digit frequency difference be $a^{(j)}$, given in equation (5)

$$a^{(j)} = \operatorname{argmax}_d |r_d^{(j)}|. \quad (5)$$

Then, we have

$$B_{i,j} = \begin{cases} 1, & \text{first digit is } a^{(j)}, \\ 0, & \text{else.} \end{cases} \quad (6)$$

First, construct the Benford factor, and the independent variable X_i is as follows:

$$X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,k}). \quad (7)$$

And the categorical variables is Y_i , the data D is as follows:

$$D = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}. \quad (8)$$

Add the constructed Benford factor to the model as a new independent variable, then we have

$$\begin{aligned} X_i^B &= (X_{i,1}, X_{i,2}, \dots, X_{i,k}, B_{i,1}, B_{i,2}, \dots, B_{i,k}), \\ D^B &= \{(X_1^B, Y_1), (X_2^B, Y_2), \dots, (X_N^B, Y_N)\}. \end{aligned} \quad (9)$$

A key advantage of bootstrap sampling is that it ensures that there is a difference between the extracted samples, which is critical for improving the performance of ensemble learning. The $D^{B(s)}$ values are assigned to the retrieved n sample data sets. Set the initial decision tree to 50, 100, 20, or any other number that corresponds to the sample size before modeling. Learning curves and grid search are used to modify the parameters after they have been determined. With the sluggish running speed of grid search and limited parameter interpretability, this article uses the learning curve for tuning, which is more intuitive and accurate than the traditional tuning method while also having a higher level of operability.

Suppose n sample data sets are trained to obtain n decision trees (DS), and record the DS model sequence as $\{h_1(X_1^B), h_2(X_2^B), \dots, h_N(X_N^B)\}$. The difference in the classification results of the DS can be further improved. The generalization ability of RS.

The final classification result can be expressed as in equation (10):

$$H(X^B) = \operatorname{argmax}_{s=1}^N I|h_s(X^{B(s)})|. \quad (10)$$

Figure 1 depicts the process of building the aforesaid BRF model during its construction. Benford's law and the RF

model are combined in the BRF model, which has the advantages of both. Benford's law is used to assess the data quality of financial indicators, and the data of sample points that may be at risk of fraud are detected. The Benford factor is then designated as a representative variable of data quality and included in the RF forest model. The RF model employs an ensemble-learning algorithm in conjunction with Bootstrap resampling, resulting in a significant improvement in the generalization ability of the final ensemble model as a result of the difference between DS. The features listed above ensure that the BRF model has a practical use. Structure of BRF is shown in Figure 4.

4. Experimental Work and Simulations

Before our experimental work, we first, establish the following FEWM system, which mainly includes five aspects: profitability, solvency, growth ability, operating ability, and cash flow. The specific variables and definitions are shown in Table 1.

Cross-validation is a technique in which 80% of a data set is randomly picked as the training set and the remaining 20% is utilized as the test set, according to the theory behind it. When the sample size of the dataset is taken into consideration, the initial number of decision trees is set to 100. By using the training set, the initial BRF model is constructed, and the prediction accuracy of the test set is used to evaluate the model's advantages and cons, as well as its strengths and weaknesses. When it comes to accuracy, it refers to the likelihood that all classifications are correct. The BRF model, which was developed using financial data from Chinese A-share listed businesses, has an initial prediction accuracy of 92.34% based on the data available so far. To fine-tune model parameters, take advantage of the learning curve.

Figure 5 depicts the outcomes of the learning curve experiment. The abscissa of the learning curve represents the number of DS trees used in the model, while the ordinate represents the accuracy with which the model predicts. The parameter with the highest accuracy in predicting the outcome is chosen as the most optimal parameter. For example, as shown in Figure 2, when the model parameter value is around 45–60, the model's prediction accuracy in the test set is very high. A further comprehensive study of the learning curve reveals that the prediction accuracy of the test set is at its maximum when the model parameter value is 40, with an accuracy of 92.25% when the model parameter value is 40. The accuracy of the prediction is enhanced by 2.49 percentage points once the parameters have been tuned.

The AUC curves for the BRF model and the RF model are depicted in Figure 6 as well. It can be seen that the AUC value of the BRF model is closer to one, indicating that the BRF model is more effective than the conventional model.

Using data from China's A-share listed businesses, Figure 7 illustrates a comparison of the prediction impacts of the BRF model and the RF model developed from scratch. Clearly, the BRF model outperforms the classic RF model across the board.

5. Conclusions

The financial risks that businesses suffer are increasing daily as globalization continues to advance. The objective requirement of market competitiveness and an essential condition for the development and survival of organizations is to be able to predict financial crises of businesses in a timely and efficient manner. Since early warning systems help organizations reduce losses and are critical before an emergency happens, that is extremely important. Based on the integration of the BRF and the RF models, this study presents an early warning system for business financial crises. Furthermore, this research uses financial data from Chinese A-share listed companies to create a BRF model. Furthermore, the FEWM prediction accuracy was found to be significantly influenced by data quality, and the Benford factor can make full use of information accurate information to efficiently identify specific sample points associated with high-financial risks. The BRF model employs the time series analysis model's ability to make short-term projections of historical information and the time series analysis model's capability to estimate recently built financial index data. The accuracy rate of the economic crisis-warning model based on random forest algorithms and time series is 92.25%, which indicates that the model is efficient and workable, according to experimental results.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This paper was supported by General Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province, construction of manufacturing cost management model system based on value chain theory, No. 2020SJA2258.

References

- [1] S. Trzeciak, "Emergency department overcrowding in the United States: an emerging threat to patient safety and public health," *Emergency Medicine Journal*, vol. 20, no. 5, pp. 402–405, 2003.
- [2] M. Jiang and X. Wang, "Research on intelligent prediction method of financial crisis of listed enterprises based on Random Forest algorithm," *Security and Communication Networks*, vol. 2021, pp. 1–7, Article ID 3807480, 2021.
- [3] M. I. Hoffert, K. Caldeira, G. Benford et al., "Advanced technology paths to global climate stability: energy for a greenhouse planet," *Science*, vol. 298, no. 5595, pp. 981–987, 2002.
- [4] E. C. Appel, "Burlesque, tragedy, and a (potentially) "yuuuge" "breaking of a frame": donald trumps rhetoric as "early warning"?" *Communication Quarterly*, vol. 66, no. 2, pp. 157–175, 2018.
- [5] V. Galaz, H. Österblom, Ö. Bodin, and B. Crona, "Global Networks and global change-induced tipping points," *International Environmental Agreements: Politics, Law and Economics*, vol. 16, no. 2, pp. 189–221, 2014.
- [6] C. Zietsma and T. B. Lawrence, "Institutional work in the transformation of an organizational field: the interplay of boundary work and practice work," *Administrative Science Quarterly*, vol. 55, no. 2, pp. 189–221, 2010.
- [7] G. Yao, X. Hu, T. Zhou, and Y. Zhang, "Enterprise credit risk prediction using supply chain information: a decision tree ensemble model based on the differential sampling rate, synthetic minority oversampling technique and adaboost," *Expert Systems*, vol. 24, 2022.
- [8] J. I. Sordo Sierpe, M. Del Rio Merino, A. Pérez Raposo, and V. Vitiello, "Algoritmos de Random Forest como alerta temprana para la predicción de insolvencias en empresas constructoras = Random Forest algorithms as early warning tools for the prediction of insolvencies in construction companies," *Anales de Edificación*, vol. 7, no. 1, p. 9, 2022.
- [9] W. Wei and B. Li, "Analysis and risk assessment of corporate financial leverage using mobile payment in the era of digital technology in a complex environment," *Journal of Mathematics*, vol. 2022, pp. 1–9, Article ID 5228374, 2022.
- [10] J. Luo, "Modeling of data mining technology in financial data recognition mining and forecasting," in *Proceedings of the 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, 20–22 Jan. 2022.
- [11] N. Gupta and A. Kumar, "Artificial Neural Networks for developing early warning system for banking system: Indian context," *International Journal of Economics and Business Research*, vol. 23, no. 2, p. 229, 2022.
- [12] I. Bou-Hamad, A. L. Anouze, and I. H. Osman, "A cognitive analytics management framework to select input and output variables for data envelopment analysis modeling of performance efficiency of banks using random forest and entropy of information," *Annals of Operations Research*, vol. 308, no. 1–2, pp. 63–92, 2021.
- [13] J. Park and M. Shin, "An approach for variable selection and prediction model for estimating the risk-based capital (RBC) based on machine learning algorithms," *Risks*, vol. 10, no. 1, p. 13, 2022.
- [14] N. Arora and P. D. Kaur, "GeoCredit: a novel fog assisted IOT based framework for credit risk assessment with behaviour scoring and geodemographic analysis," *Journal of Ambient Intelligence and Humanized Computing*, vol. 76, 2022.
- [15] H. Etemadi, A. A. Anvary Rostamy, and H. F. Dehkordi, "A genetic programming model for bankruptcy prediction: empirical evidence from Iran," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3199–3207, 2009.