Research Article

Simulation Model of Enterprise Financial Management Activities Based on the Random Forest Algorithm

Fang Zhao

School of Accounting, Henan Logistics Vocational College, Zhengzhou, Henan 450000, China

Correspondence should be addressed to Fang Zhao; 2111602035@e.gzhu.edu.cn

Received 5 July 2022; Revised 26 August 2022; Accepted 8 September 2022; Published 20 September 2022

Academic Editor: Chi Lin

Copyright © 2022 Fang Zhao. 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.

In the era of information and data, big data and cloud computing have been widely used in all walks of life, especially for promoting the transformation of business model and the financial administration model of corporations. Building various management systems based on the financial sharing center can realize centralized management of business finance and improve enterprise management efficiency. Enterprises must establish a set of perfect financial administration systems to prevent and resolve financial crisis, reduce financial losses and ensure the normal operation of corporations. Based on situational learning theory and simulation experiment technology, this paper constructs a comprehensive simulation experiment system of financial administration and focuses on three key technologies: case decomposition and task-driven, business simulation expansion, and scientific evaluation. Based on the random forest algorithm, a preparation setup model of accounts receivable and payable audit is constructed, and its application in case corporations is analyzed, which provides new ideas for the modern audit work mode and creates a dynamic and open business simulation environment for the financial administration experiment. The results show that the classification accuracy of the improved K-fold random forest algorithm is 1.54% higher than that of the traditional random forest algorithm, and the accuracy of the crisis preparation setup model in predicting the actual enterprise financial crisis reaches 90.327%, which is feasible and effective in the preparation setup of enterprise financial administration and provides a new research idea for the construction of the preparation setup model of enterprise financial administration. Financial management is the core part of enterprise management activities, and it is also an important factor affecting the comprehensive competitiveness of corporations. Therefore, on the basis of fully studying the objective regularity of financial activities, improving the financial administration algorithm, applying the random forest algorithm to enterprise management, and building a perfect management system can comprehensively control the financial administration activities of corporations and realize the innovation of financial administration activities from the aspects of data collection, fund management, and internal control, which is conducive to promoting the continuous improvement of the efficiency of enterprise financial administration.

1. Introduction

With the rapid development of the market economy, the market competition environment faced by corporations is becoming increasingly fierce. In order to enhance their comprehensive competitiveness, corporations must continuously improve their internal management models and improve management efficiency. Financial management is the core part of enterprise management activities and also an important factor affecting the comprehensive competitiveness of corporations. In the traditional internal audit, the audit of accounts receivable and accounts payable needs to use the confirmation method most of the time, but due to the extremely low response rate of the confirmation, it is impossible to obtain the necessary external audit evidence. In recent years, through the policy guidance and guidance of government departments such as the Ministry of Finance, the financial shared service model has caused a new upsurge. Financial Shared Service Center is from the perspective of the market, using modern information technology, to provide professional financial services for the internal and external customers of the group enterprise, it is a new type of financial administration model and is the Internet economy and enterprise management sharing concept a new
The random forest algorithm cited in this paper is to randomly construct a number of unrelated decision trees to form a forest. Each decision tree makes a decision on the attribute category of each data put into the random forest and finally predicts the financial risk level of each data according to the decision results of all decision trees. Real-time data collection and timely uploading of output audit warning results in the process of enterprise production, operation and financial administration can be realized by the management through the financial sharing center. The results show that the classification accuracy of the improved K-fold random forest algorithm is 1.54% higher than that of the traditional random forest algorithm, and the prediction accuracy of the crisis preparation setup model to the actual enterprise financial crisis reaches 90.327%, which is feasible and effective in the enterprise financial administration preparation setup and provides a new research idea for the construction of the enterprise financial administration preparation setup model [2].

While fully considering the openness and dynamics of the financial sharing environment, this paper studies and solves the key scientific problem of the audit preparation setup of accounts receivable and payable in the enterprise financial sharing center and constructs the audit preparation setup framework of accounts receivable and payable from the perspective of classification by using the random forest theory and develops the implementation framework in combination with the evidence theory, which verifies the correctness and effectiveness of the evaluation method in the case analysis. The model is verified by using R software, and the conclusion is drawn, which provides an effective method and a reasonable way for the enterprise financial sharing center to study the audit preparation setup of a/P [3]. The innovation lies in the following:

1. The financial preparation setup model of accounts receivable and payable under the financial sharing mode based on the random forest is constructed. On the basis of in-depth analysis, first build a financial preparation setup framework and then analyze the current situation and existing problems of the company’s receivables and payable management, select financial preparation setup indicators and then use the random forest algorithm to select features and then build a random forest model. Finally, the financial warning threshold is obtained. According to the financial preparation setup model established by a random Senli algorithm, the application of the model is explained in combination with the actual situation of the enterprise, and the validity of the model is proved.

2. Taking an enterprise under the financial sharing model as an example, a new idea and scheme of financial preparation setup is introduced, which is conducive to the development of internal finance and provides new ideas for the construction of an intelligent financial system [4]. On the basis of in-depth analysis of the features of the enterprise financial sharing mode, the financial preparation setup model of accounts receivable and payable is analyzed in detail. The data source before the implementation of the on-site financial procedures is collected and cleaned through big data technology, and the big data technology is further used. Analysis of massive financial data, compared with traditional financial sampling, expands financial scope, improves financial efficiency, and provides new ideas for financial preparation setup methods of accounts receivable and payable, so as to realize the random forest method in internal financial research innovation.

This paper is divided into six chapters from the organizational structure.

The first chapter is the introduction. This part analyzes the application status of the business model and financial administration model, summarizes the causes of problems and compares new technologies and algorithms to improve the efficiency of enterprise management. The second chapter mainly summarizes the relevant literature, summarizes its advantages and disadvantages and puts forward the research ideas of this paper. The third chapter introduces in detail the objectives of enterprise financial administration activities and the establishment of the financial administration simulation system. The fourth chapter introduces the enterprise financial administration activity simulation model based on the random forest algorithm and the financial crisis preparation setup based on the improved K-fold random forest algorithm. The fifth chapter describes the data processing and the construction of financial indicators. The sixth chapter is the conclusion, which summarizes the research results of the full text.

2. Related Work

From the perspective of financial practice, Wu Jing compared the financial risk preparation setup and the financial preparation setup, constructed the framework of the financial preparation setup and explained the important points in the construction process and put forward some suggestions. Constructive comments [5]. Wu pointed out that a new type of financial model is online finance, an important part of which is financial preparation setup. According to the features of China’s online finance, a financial preparation setup model is constructed, and the most important part is the collection and detection of databases and incomplete data [6]. Starting from the principles of financial administration, Wu and Chen finds out errors and frauds and makes preparation setups according to financial goals, studied the necessary elements for building a financial preparation setup system, designs a financial preparation setup framework flow standard and constructs a tax preparation setup system framework through five aspects and pointed out the functional role and limitations of this indicator [7]. Forsaith uses the Granger causality test for the financial administration index system to determine the
weight of management indicators and designs a financial preparation setup index. The direct invisible risk has the greatest impact on finance [8]. Anna selects tax risk management from the macro level and the micro level [9].

Since the concept of the random forest algorithm was proposed, scholars have conducted a lot of research on it and put forward many improved algorithms on this basis. Jie summarized the main theoretical decision trees and integrated learning of random forests and then proposed an improved algorithm for the classification of ultra-high dimensional attribute data. The weight of random forest features is determined according to the size of statistics, which can improve the classification accuracy [10]. Tu compared the similarities and differences between the cart algorithm and the C4.5 algorithm and created a new splitting algorithm by combining the two algorithms, focusing on the construction of weighted nodes of binary numbers. He applied this algorithm to the random forest algorithm and proposed the concept of the hybrid random forest algorithm [11]. Scholars have also proposed many improved methods for the classification of random forests. Wang and Liu proposed that after building the decision tree, the decision tree should be sorted according to its classification ability, and the decision tree with strong classification ability should be selected to build a new random forest [12]. Li and Wang pointed out that different weights are allocated according to the classification ability of the decision tree, and then the classification is carried out. In the classification process, the weighted voting method should be used, that is, the weight of the decision tree is given to each ballot [13]. Liu and Song used a vector space model and potential topic representation to construct a decision tree and classified the random forest composed of these two features with experimental data. The final classification result was calculated by the weighted average formula [14].

3. Objectives of Corporate Financial Administration Activities

3.1. Goal Theory. The objectives pursued by enterprise financial administration activities can be roughly divided into the following three aspects:

(1) Maximize profits. It is generally believed that profits represent the wealth created by corporations. The more profits, the more wealth the enterprise has. As an independent entity, corporations must pursue profits driven by the principle of self-interest. Therefore, the greater the profit, the closer to the goal of the enterprise. In recent years' the view of profit maximization has been questioned and denied. The reason is that the profit maximization financial administration goal does not consider the time value and risk of money. The relationship between input and output is not considered.

(2) Development of the view that earnings per share maximization is profit maximization. To a certain extent, it makes up for the lack of profit maximization. It is generally believed that this view not only reflects the pursuit of profits by corporations but also links the profits of corporations with the capital invested by shareholders. However, the risk and time of earnings acquisition are still not considered [15].

(3) Enterprise value maximization measures the overall value of the enterprise. It represents the judgment and evaluation of entrepreneurs made by capital market participants. It is generally believed that this view comprehensively and fully considers the company's current and future possible profitability, time factors, risks, and other factors.

3.2. Financial Management Simulation System. First, collect and preprocess the dynamic financial data of the time series of corporations and then conduct the stationarity test and the pure randomness test. If it fails the test, it will be stabilized by differential operation and transformed into a stationary time series and then select a mature model to complete the time series modeling. The modeling process of time series is shown in Figure 1.

The research object of the financial administration comprehensive simulation system is the enterprise, that is, how to flexibly formulate financial administration strategies in the process of operation and development and use internal and external resources to carry out various financial accounting and decision-making activities, so as to maximize the economic benefits of the enterprise. It integrates financial strategy formulation, accounting operation, financial decision simulation, business operation simulation, and other modules, with huge data stock and flow [16]. Based on the basic idea of comprehensive experiments, a virtual financial business environment is constructed through computer simulation technology in the network environment, and a variety of positions and roles of the financial administration department are set for overall operation or decision-making. An experimental platform system, such as shown in Figure 2.

4. Simulation Model of Enterprise Financial Administration Activities

4.1. Introduction to the Random Forest Algorithm. Random forest is to plant many decision trees that are not related to each other by using random methods. These decision trees form a forest, which is called a random forest. After the random forest is established, assuming that there is a new sample, it is put into the random forest, and then each decision tree in the random forest is allowed to make a decision on the attribute category of the sample to be entered. Each tree has one vote. The minority obeys the majority method. The category with the largest number of votes in the decision tree is the final classification result of the sample [17]. The working principle of random forest algorithm can be divided into the following four steps, as shown in Figure 3.

Step 1: Randomly sample the training set \( P \) times with replacement and form a subset of the training set with the obtained \( P \) samples and use it as a new training set \( P \).
Step 2: Randomly extract \( k \) features of the training set from \( P \) to form a subset and use this subset to perform a decision tree. During the decision-making process, let the decision tree grow sufficiently without pruning it [18].

Step 3: Repeat steps 1 and 2 until \( m \) decision trees are generated to establish a random forest.

Step 4: The test sample to be classified is given to the decision tree in the forest for classification decision, and the classification voting results of each decision tree are counted, and the category with the most votes in the decision tree is used as the final classification result of the test sample.

4.2. Financial Crisis Warning based on the Improved K-Fold Random Forest Algorithm. Use the time series model to make a short-term forecast on the historical financial index data of
the enterprise and then use the K-fold random forest algorithm to make a financial analysis on the forecast data and finally realize the preparation setup of the enterprise’s financial crisis [19]. Test the pure randomness of the time series data to ensure that the time series data have correlation and can be effectively predicted through the historical data. The test original hypothesis function is shown as follows:

\[
\rho_1 = \rho_2 = \cdots = \rho_m = 0, \forall m \geq 1. \tag{1}
\]

In the formula, \( \rho \) represents the autocorrelation coefficient of the sequence samples. For large sample data, the QBP statistic is used to judge the randomness, and the function is expressed as follows:

\[
Q_{BP} = n \sum_{k=1}^{m} \rho_k^2 \sim \chi(m). \tag{2}
\]

In the formula, \( n \) represents the sequence observation period, \( m \) represents the specified delay period, and \( \chi(m) \) represents the quantile. When faced with small sample data, the function of the QLB statistic is expressed as follows:

\[
Q_{LB} = n(n + 2) \sum_{k=1}^{m} \left( \frac{\rho_k^2}{(n-k)} \right) \sim \chi(m). \tag{3}
\]

The time series model selection is carried out according to the features of the autocorrelation diagram and the partial autocorrelation diagram of the time series. When the autocorrelation diagram has tailing phenomenon, and the partial autocorrelation diagram has the tailing phenomenon, the AR autoregression model is used, and the function is expressed as follows:

\[
X_i = \sum_{j=1}^{q} \theta_j \alpha_{i-j} + \alpha_i + c. \tag{5}
\]

In the formula, \( \theta_j \) represents the moving average coefficient, and \( q \) represents the moving average order. When the autocorrelation graph and the partial autocorrelation graph both show tailing phenomenon, the ARMA autoregressive moving model or the ARIMA autoregressive difference moving average model is used, and the ARIMA model is used for the sub-differentiated time series. The ARMA function is expressed as follows:

\[
X_i = \sum_{j=1}^{p} \phi_j \alpha_{i-j} - \sum_{j=1}^{q} \theta_j \alpha_{i-j} + \alpha_i + c. \tag{6}
\]

The improved K-fold random forest algorithm is used to analyze the financial situation of corporations and realize the preparation setup of financial crisis. Random forest algorithm is based on classification regression tree, combined with a bagging algorithm and a random subspace method. It has the features of high accuracy, good scalability and parallelism, low demand for prior knowledge, less vulnerable to the negative impact of interfering data, wide applicability, and simple and clear structure [20]. The random sampling method is used to select \( y \) training sample sets from \( N \) original data samples, and the optimal splitting feature is selected to split the classification regression tree for training.
$Y$ classification regression trees are generated through the learning of the sample sets. Finally, the simple majority voting method is used to obtain the final classification results.

In the process of random forest sample extraction, taking bagging idea as the core, the self-help sampling method with the random selection method of putting back is adopted. After repeated operations for many times, the number of training sample sets is consistent with the number of original sample sets, and the out of pocket data is a part of the sample data that has not been selected. Add noise to a single feature and measure the importance of the feature by reducing the change of prediction accuracy. First, use the corresponding out of pocket data to calculate the out of pocket data error of each classification regression tree, carry out random noise interference on each sample feature $x_i$ in the out of pocket data and calculate the out of pocket data error again. Repeat the operation for many times until the two out of pocket data error calculations of all classification regression trees are completed. Finally, calculate the importance of the feature, and the calculation function is shown as follows:

$$T_{x_i} = \frac{\sum_{j=1}^{Q} OOB2(x_i) - OOB1(x_i)}{Q}. \quad (7)$$

In the formula, $Q$ represents the number of decision trees, $OOB1(x_i)$ represents the out-of-bag data error obtained for the first time, $OOB2(x_i)$ represents the out-of-bag data error after interference, $T(x_i)$ represents the importance of the feature $x_i$, and the value of $T(x_i)$ is positively correlated with the importance, when $T(x_i)<0$, the noise of feature $x_i$ has a negative disturbing effect on the model.

4.3. Data Dimensionality Reduction Processing. The data dimensionality reduction process first compares the algorithms before and after dimensionality reduction and uses decision tree, random forest, and support vector machine algorithms for comparative analysis. The first 770 data are used as training samples, and the last 30 data are used as test samples. First normalize the data, and then perform matrix analysis and calculate the MSE values before and after dimensionality reduction. The calculation formula is as follows:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - y_d)}{n}. \quad (8)$$

The results obtained from the calculation are shown in Table 1.

It can be seen that the error rate has decreased significantly and the accuracy has been improved after dimensionality reduction by factor analysis.

In order to obtain the audit warning threshold suitable for the enterprise, the enterprise should use its own data to verify and optimize the set financial threshold: generally, the enterprise should make a financial evaluation once a month or every quarter, and the financial department should make a comprehensive evaluation at the end of the year. First, the actual financial data are used to determine the actual financial early-warning interval of the enterprise, then the actual early-warning interval is compared with the set threshold interval, the financial differences between them and the reasons for the differences are analyzed, and the early-warning interval of each financial supervisor is adjusted. Finally, the financial audit early-warning threshold of the enterprise is comprehensively verified, adjusted, and optimized in combination with the enterprise’s financial status and the market environment. Finally, we get the most suitable audit warning threshold for ABC corporations.

Let function $f(x) = f(x_1, x_2, x_3, \cdots, x_n)$, where $(x_1, x_2, x_3, \cdots, x_n)$ is the independent variable in random forest. The partial correlation of function $f(x)$ on $x_j$ is

$$f_j(x_j) = E_x[f(x)]. \quad (9)$$

For the $k$ classification problem, let $\rho_k(x)$ be the vote probability of category $k$, then the partial correlation function corresponding to category $k$ is

$$f_k(x) = \log \rho_k(x) - \frac{1}{k} \sum_{j=1}^{k} \log \rho_j(x). \quad (10)$$

In this paper, a similar integration idea is applied to the data level, and the data mining software is $R$ language. $R$ language is not only compatible with various operating systems but also has many mature and stable data mining models, which is conducive to the accuracy of audit warning indicators. The training set used in each round of the fitting process of the model is a new data set obtained by random reorganization on the original data set, and the process of data reorganization is completely randomized and independent of the generation process of the previous training set. The final prediction result of the algorithm in this paper is also estimated using the average value, which is similar to the random forest.

5. Data Processing and Financial Indicator Construction

In order to carry out the enterprise financial crisis preparation setup more comprehensively and accurately, based on the classification features and the feature importance measurement method of the K-fold random forest algorithm, a new financial preparation setup index system is established, and the financial crisis analysis that has a high impact on the enterprise is screened out. Indicators are established to improve the prediction accuracy of crisis preparedness setting. Based on the financial statements of listed manufacturing corporations, starting from the features of the financial situation of the financial crisis of the enterprise, the preliminary screening of financial indicators is carried out from the five aspects of the enterprise’s profitability, solvency, operating ability, development space, and cash flow. It includes 26 financial indicators such as current ratio, asset-liability ratio, and total asset turnover ratio, which are numbered from $A$ to $Z$. Using OOB out-of-bag data to carry out importance measurement analysis, the top six financial indicators in the order of importance are shown in Table 2, and a new financial
In this part, we use the decision tree model and random forest model introduced earlier to analyze and process the financial data, and test the classification effect of each model.

(1) Decision tree classification. First, we set the parameters of the decision tree, including the maximum number of features, the maximum depth, and other attributes. After SMOTE sampling, the optimal hyperparameters are found, and the model is trained and evaluated by 10-fold cross-validation. Then we verify the model on the test set. After modeling, we found that the maximum depth of the decision tree was 8, and the accuracy rate was 87.63%. Too deep a decision tree will cause the data to be overfitted, which will lead to the decline of judgment results. The specific maximum depth traversal results are shown in Figure 4.

(2) Random forest classification. The random forest model also needs to do each of the parameters first. Considering that each decision tree in the random forest requires cross-validation, its fold does not have much impact. Considering the calculation performance, we will randomly select the number of cross-validation folds in the forest as 3. At the same time, we traverse the total number of decision trees and the maximum depth in different situations. The specific values are shown in Table 3.

The results of data analysis showed that the best judgment rate was 800 decision trees at the maximum depth of 5 layers of random forest, and the value was 88.27% (See Figure 5 for details).

According to the above experimental results, we have constructed 14 financial indicators based on financial data, which are related to the five aspects of the enterprise’s solvency, operating ability, profitability, development ability, and risk level. Based on these data, a decision tree model and a random forest model are constructed. From the empirical results, the random forest model has a higher analysis accuracy than the decision tree model. The results based on the model need to be checked manually in the follow-up, and the judgment accuracy rate close to 90% can effectively reduce the time cost of manual checking.

Algorithm Accuracy Analysis. In order to verify the effectiveness and optimization of the improved K-fold random forest algorithm, three data sets in the UCI database were used for simulation experiments, 80% of which were used as training data sets and 20% as test sets. Experiments are carried out in the Anaconda3+ VCode environment, and the classification accuracy of the traditional random forest algorithm and the K-fold random forest algorithm under different numbers of classification and regression trees is compared. The comparison results are shown in Figure 6.

The input data in this model is the financial information of listed corporations. In this chapter, the data processing process and the construction of financial indicators will be introduced. There are many methods to judge the financial irregularities of listed corporations, among which we adopt a more credible method, that is, to judge whether there are any irregularities in the financial reports of the year through the statistical caliber of the problems and results of the handling of irregularities disclosed by CSRC. We downloaded all the processing file announcements of A stock exchange and B stock exchange since 2017, marked them according to the name of the offending enterprise and the offending time.
disclosed therein and completed the processing of the forecast variables, as shown in Figure 7.

Therefore, it is particularly important to establish a number of decision tree individuals with independence and distinctiveness in building a random forest model. If we simply use the original data to judge the data, it is difficult to realize a diversified decision tree model, and it is also difficult to reflect the advantages of integrated learning. Therefore, when building a random forest model, we need to sample the data according to certain rules, and each decision tree uses part of the total sample for training, which can not only reduce the repetition of the training data but also make the database of each decision tree different and increase the difference of individual decision trees.

In the era of big data, the financial sharing center provides great convenience for the data collection of the financial preparation setup model. The corporate account information data can be directly extracted from the business module of the financial sharing center. In addition to the internal financial data of the enterprise, the average data of various indicators in various industries and the related information data of the bank-enterprise interconnection system should also be collected. The data required by the financial preparation setup model must be authentic, accurate, and complete in order to ensure the accuracy of the financial risk preparation setup output by the financial model. However, the information and data initially collected are numerous and complicated and need to be sorted and processed uniformly. The information data collected initially needs to firstly screen out the data required by the financial preparation setup model and digitally process and store it in the database; secondly, use the data mining technology to convert the data in the database and the financial preparation setup indicators correspondingly and eliminate the correlation and noise. Big data is as shown in Figure 8; finally, the transformed and eliminated data is copied to the DBMS management system, and the Hadoop platform is used to normalize the data.

<table>
<thead>
<tr>
<th></th>
<th>Whole school (%)</th>
<th>Whole class (%)</th>
<th>Student A (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math</td>
<td>82.30</td>
<td>79.30</td>
<td>90.60</td>
</tr>
<tr>
<td>Accounting</td>
<td>84.20</td>
<td>91.80</td>
<td>94.20</td>
</tr>
<tr>
<td>Programming</td>
<td>83.60</td>
<td>79.30</td>
<td>84.20</td>
</tr>
<tr>
<td>Management</td>
<td>84.30</td>
<td>81.20</td>
<td>77.20</td>
</tr>
<tr>
<td>Computer application</td>
<td>87.20</td>
<td>86.10</td>
<td>87.30</td>
</tr>
</tbody>
</table>

Figure 5: Accuracy of the random forest model in different situations.

Figure 6: Comparison of the classification accuracy of the two algorithms.

Figure 7: Statistics on violations.
6. Conclusion

The new economic development trend requires companies to monitor their financial status in real time and strengthen their risk management and control capabilities. In order to realize the timely and accurate preparation setup of enterprise financial crisis, an enterprise financial crisis preparation setup model of improved K-fold random forest algorithm was constructed based on time series and real financial data. The results show that, compared with the traditional random forest algorithm, the classification accuracy of the improved K-fold random forest algorithm is improved by 1.54%, and the classification accuracy rate reaches 82.6%. K-fold cross-validation idea is optimized for the random forest algorithm. In the application of enterprise financial crisis preparation setup, the prediction accuracy of the preparation setup model is 90.327%. The positive and negative classes are 91.869% and 88.785%, respectively. The prediction accuracy is high. There are differences in the actual data of financial indicators of corporations in different industries, and it is difficult to use a unified indicator standard for financial crisis preparation setup. The research limits the type of sample data and can further design the financial preparation setup indicator system in different industries to improve the adaptability of the preparation setup model.

Data Availability

The data used to support the findings of this study are available within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References


