Research Article

Financial Account Audit Early Warning Based on Fuzzy Comprehensive Evaluation and Random Forest Model

Yuting Zhao¹ and Yupeng Fang²

¹Department of Accountancy Accounting, School of Shandong Youth University of Political Science, Jinan 250000, China
²General Affairs Office, Shandong Institute of Commerce and Technology, Jinan 250000, China

Correspondence should be addressed to Yuting Zhao; 080266@sdyu.edu.cn

Received 12 January 2022; Revised 8 February 2022; Accepted 10 February 2022; Published 22 March 2022

Academic Editor: Miaochao Chen

Copyright © 2022 Yuting Zhao and Yupeng Fang. 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.

With the continuous and rapid development of China’s economy, the operating environment of listed companies has become more and more complex, and the increasing pressure of international competition among companies has made the issue of financial risks of listed companies more severe. If you do not pay attention to the financial risk status of the enterprise, it will cause the financial risk to accumulate and eventually cause a financial crisis, which will be marked by ST. Therefore, this paper proposes an early warning model of enterprise financial accounts combining fuzzy sets and random forest trees, which specifically includes the following steps. First, the dataset is analyzed, selected and initially constructed by the training prediction sample. It is further explained by the data labels, and is charged whether the label is marked by ST or not. Then, the method of fuzzy mathematics is used to fuzzify the training sample data, and the two-category label is converted into a multiclass label; then, the random forest model is used to train the above-mentioned fuzzified sample data. Obtain the trained random forest model. Finally, input the prediction sample data into the trained random forest model to make decisions on the scene application. At the same time, the invention is applied to the enterprise financial risk early warning, which demonstrates the practicability and effectiveness of the invention Sexuality and scientificity. The significant advantage of the present invention is that the two-class decision making is converted into the multiclass decision making by combining the fuzzy set and the random forest model, which greatly improves the prediction accuracy, efficiency, and data rationality.

1. Introduction

Since 2009, the ChiNext, as a supplement to the main board market, has developed rapidly since its launch in the capital market. It has greatly developed China’s capital market and has also provided more opportunities for some small and medium-sized start-up companies. However, while it brings opportunities, it also brings challenges. In recent years, there have been instances of GEM companies causing financial crises due to poor management. The frequent financial crises of listed companies will not only cause the public to gradually lose confidence in listed companies but also affect the development of China’s capital market [1–9].

Therefore, this article takes listed companies as examples to analyze their liquidity risks and find out the root cause of their financial crisis. We further analyze the loopholes in its liquidity risk management, on this basis, put forward suggestions for improving its liquidity risk management for the case company, and build appropriate liquidity risk assessment and early warning models, so as to strengthen risk awareness and improve risk management for the company [10–15]. Corporate financial crisis is shown in Figure 1.

Corporate financial crisis will not break out suddenly. The occurrence of crises must be a periodical feature. There may be many internal reasons. This shows that the rapid development of computer technology has also brought more solutions to this field. Corporate financial risk early warning experts and scholars in the field are also working hard to find more useful early warning model algorithms that are more suitable for the characteristics of actual corporate financial risks.
composition of internal and external factors makes the financial crisis very complex, and it also determines that the use of a simple linear function for early warning will be very poor. This shows that the relationship between corporate financial crisis and influencing factors is not a linear relationship but a non-linear relationship affected by many aspects. From the beginning, the scholars of the predecessors used a simple single variable to describe the financial risks of the enterprise, and now they use a set of comprehensive index system to use the random forest algorithm to evaluate, which is the continuous innovation of the method and the continuous improvement of the accuracy rate. At present, random forest is a very commonly used and very flexible algorithm. Its advantages allow it to be applied in many fields. It can be used in the marketing field to predict the source of users, and it can also be used in the medical field by extracting diseases. The characteristics and modeling of these data can predict the probability of contracting the disease and classify patients. In recent years, in various data mining competitions at home and abroad, the players who use the random forest algorithm to build data mining models accounted for a large proportion. This shows that the random forest algorithm has a very wide range of actual data mining scenario applications. Random forest is composed of a set of decision trees with the same properties, so it is still subject to ensemble learning in essence. Random forest is a very flexible algorithm. It can process sample data of high-dimensional features without reducing its dimensionality. It can also evaluate the importance of each vector feature in the classification results. It also handles default values. Good results can be obtained. Compared with other classification prediction algorithms used, the random forest algorithm is not only accurate but also suitable for more scenarios. Many experts and researchers apply it in various fields, which has proved that random forest has a very significant effect on the description of multidimensional complex functions. Therefore, this article applies the algorithm of random forest to the early warning of corporate financial risks. There have been a large number of applications in other fields before this. Compared with other algorithms, random forest will have a more prominent performance in the early warning of financial risks [16–22].

Adair and Hutchiso constructed a financial risk evaluation scale for financial risk issues and used actual companies as examples to assess their financial risk levels. Altman used a multivariate discriminant analysis method to analyze case studies of real estate companies that have experienced financial crises from 1969 to 1999. The study found that it is imperfect to use the Z model to identify companies that have financial crises. Therefore, the financial crisis prediction model is further improved. Charitou et al. used listed companies in the UK as a case and selected UK bankrupt companies in the two periods of 1988–1994 and 1995–1997 to conduct a comparative analysis, using neural networks and logical models to find the cash flow’s risk to corporate funds. The importance of duarte triggers is considered as the analysis basis of the ratio analysis method, and the financial ratio indicator system calculated through historical data can only reflect the status of the company’s past transactions or events and cannot make predictions about the future. Based on such a situation, the evaluation system constructed by relying on financial indicators cannot be a basis for the future development of the enterprise. Sally constructed a joint forecasting model by effectively combining four independent financial risk early warning research methods. Empirical research shows that under the same conditions, this model is significantly better than a single method research model [23–27].

Random forest (RF) is a representative algorithm in the field of data mining. It can dig out a lot of information from limited data. The RF algorithm uses the bootstrap re-sampling method to obtain training samples. The basic idea is to construct a multidecision tree model. The advantages of high prediction accuracy, controllable generalization error, fast convergence speed, and few adjustment parameters can effectively avoid the occurrence of overfitting and are especially suitable for high-level data calculations. There have been studies comparing RF algorithm with SVM and ANN algorithms, which proved the superiority of RF algorithm. At present, there are only two types of decision making for various scenarios through the random forest tree algorithm, which is also the shortcoming of the random forest algorithm.

Fuzzy mathematics is a new subject, which has been initially applied to various aspects such as fuzzy control, fuzzy recognition, fuzzy fruit analysis, fuzzy decision making and fuzzy judgment, system theory, information retrieval, and so on. Fuzzy representation of training samples through fuzzy mathematics, conversion of two-category labels into multicategory labels through category membership value features, and multicategory decision making combined with random forest trees greatly improve the accuracy and breadth of decision making and are suitable for decision making in various scenarios. This is also the key point of this article. The application fields of fuzzy mathematics research are also very broad, such as medicine, biology, engineering, artificial intelligence, society, psychology, and other disciplines.

With the rapid development of China’s economy, the capital market system has become more and more perfect. At present, China is in a critical period of economic transformation, and it is necessary to always be aware of financial risks brought by the external environment or within the enterprise. Only by controlling the risks can the enterprise operate better. Improving the early warning of
corporate financial risks is an indispensable control stage for the continuous development of enterprises. Many researchers have carried out long-term research and exploration on how to improve the accuracy of financial risk early warning. At the same time, this article uses financial risk early warning as one of the scenarios to demonstrate the feasibility, effectiveness, and scientificity of this article [28, 29].

### 2. Fuzzy Comprehensive Evaluation and Random Forest

#### 2.1. Fuzzy Comprehensive Evaluation

Fuzzy means that the boundary is unclear, and the boundary cannot be distinguished in the field of theory. In real life, many concepts are vague. For example, there is no clear boundary between tall and short. Some people think that 170–180 cm is a medium height, but some people think that 180 cm is a body height. It can be seen that such a limit is often unclear. The fuzzy comprehensive evaluation method is a mathematical method based on fuzzy mathematics. This kind of evaluation method is to quantitatively deal with qualitative problems, and the process of quantitative processing is based on the theory of membership degree. In this way, a complex thing or system can be decomposed into many factors and levels, and finally a comprehensive overall evaluation can be carried out.

Fuzzy mathematics was first proposed by the American expert Professor Zadeh LA. He published his research results in the form of published papers, and thus fuzzy mathematics was officially born. The title of this well-known academic paper is "Fuzzy Sets," which first proposed the concept of fuzzy and also made a related introduction to quantitative representation. Fuzzy mathematics is a method of studying fuzzy phenomena. It is a relatively new subject of mathematics. It expands the application scope of mathematics from the deterministic field to the fuzzy field, that is, from the precise phenomenon to the fuzzy phenomenon. In various scientific fields, various quantities can be divided into two categories: certainty and uncertainty. Uncertainty can be divided into random and fuzzy. Fuzzy mathematics is a mathematical method to study fuzzy uncertainty. The idea of fuzzy mathematics is to use precise mathematical methods to describe, model, and calculate a large number of fuzzy concepts and fuzzy behaviors in the real world and to deal with them reasonably. In these years, the applications of fuzzy mathematics research are also very wide, such as medicine, biology, engineering, artificial intelligence, society, psychology, and other disciplines.

Fuzzy comprehensive evaluation method is to use some concepts of fuzzy mathematics to provide some evaluation methods to actual comprehensive evaluation problems. Specifically, fuzzy comprehensive evaluation is based on fuzzy mathematics, applying the principle of fuzzy relationship synthesis to quantify some unclear borders and difficult to quantify factors and comprehensively evaluate the status of the subject’s affiliation from multiple factors. It has the characteristics of clear results and strong system. It can solve fuzzy and difficult to quantify problems. It is suitable for solving various non-deterministic problems. The basic principle is as follows: first determine the set of factors (indicators) and evaluation (rank) of the evaluated object; then determine the weights of each factor and their membership degree vectors to obtain the fuzzy evaluation matrix; finally, the fuzzy evaluation matrix and the weight of the factors are determined. The vector is subjected to fuzzy calculation and normalization, and the comprehensive result of fuzzy evaluation is obtained.

#### 2.2. Random Forest Model

The learning method of the decision tree model is based on an inductive learning algorithm, which focuses on how to find the features that form a decision tree from a set of examples in an irregular order. These feature rules are usually used to construct a classification model or prediction model, and then through this model, some unknown data can be predicted or classified. To build a decision tree model, you need to go through the following steps: the first step is to select a suitable training sample dataset and then use these data to build a decision tree model through a certain training algorithm. The process of building a decision tree model can be understood as mining the internal rules of the sample through a certain algorithm. The training is mainly carried out in two stages: the first stage is tree building, and the second stage is pruning. After the decision tree model is trained, proceed to the second step. This process is to use the established decision tree model to perform a judgment analysis on the new sample. Random forest model (RFC) is a strong classification model formed by a group of decision tree models with the same properties through a certain combination strategy. The vectors in the parameter set are independent of each other. Under the input variable $X$, each decision tree in the set has the right to vote on the result based on a certain strategy. The basic idea of the RFC model is as follows. The first step is to use the bootstrap method to form from the sample data used for training. There are two sample sets, and the amount of data inside each sample set is the same. The second step is to train a decision tree separately for each sample set in this sample set. The third step is to make a final vote based on the results of this group of decision tree classification to determine the final classification prediction result.

Decision tree is an algorithm that can classify and predict new data (or test data) by measuring training data (historical data). These data are analyzed to find the characteristics or rules in the data, and these are used as a basis to predict the results of the new data. Simply put, the purpose of establishing a decision tree is to construct a suitable model based on the values of several input variables to predict the values of the target and output variables and present them in a tree structure. The algorithms used in the decision tree mainly include ID3, C4.5, and CART algorithms. The ID3 selection attribute uses the information gain of the subtree, which is the change value of entropy; the C4.5 algorithm uses the information gain rate, while the CART algorithm and the Gini indicator are selected and used in this article.
Random forest is an algorithm that integrates multiple trees through the idea of ensemble learning, as shown in Figure 2. Its basic unit is a decision tree, and its essence belongs to a major branch of machine learning-ensemble learning methods. In machine learning, a random forest is a classifier that contains multiple decision trees, and the output category is determined by the mode of the category output by the individual tree. Use the same training number to build multiple independent classification models and then make the final classification decision based on the principle of minority obeying the majority through voting. For example, if you train 5 trees, 4 of them are True, and 1 of them is False; then, the final result will be True. A standard decision tree will be based on the impact of each feature on the prediction result the degree is sorted to determine the order in which different features are constructed from top to bottom. In this way, all decision trees in the random forest will be affected by this strategy and constructed completely the same, thus losing the diversity. Therefore, during the construction of the random forest classifier, each decision tree will abandon this fixed sorting algorithm and select features randomly.

The basic core idea of RF is to increase the internal differences of the model by constructing some independent and connected decision trees. These internal differences allow the random forest to make a correct judgment on complex data. Obtain the group dataset through the bootstrap sampling method and then train the data in the row group dataset to obtain a set of decision trees. These decision trees constitute a combined classification model system. The final result of the system is determined by voting a minority.

Compared with traditional classification algorithms, random forest has many advantages, such as fewer parameters that need to be adjusted, which means it can efficiently process large sample data and no need to worry about overfitting or strong noise tolerance, which can effectively prevent data sparse problem of decision tree.

3. Enterprise Financial Early Warning Model Based on Fuzzy Comprehensive Evaluation and Random Forest Model

Fuzzy mathematics is a new subject, which has been initially applied to various aspects such as fuzzy control, fuzzy recognition, fuzzy fruit analysis, fuzzy decision making and fuzzy judgment, system theory, information retrieval, and so on. The following is a detailed description of the technical solution for applying this article to financial risk early warning with reference to Figure 3. It should be noted that the description here only takes the financial risk early warning of listed companies as an example. The invention is suitable for decision making in various scenarios. The steps to apply this article to the financial risk warning of listed companies are as follows.

**Step 1.** Analyze and select the corporate quarterly data released by more than 3,000 listed companies in the Shanghai and Shenzhen stock exchanges in the past seven years and initially construct a training and forecast sample set; further, the data include the income statement, cash flow statement, and total assets and liabilities. The mark of the twenty-six-dimensional data shows whether it is marked by ST or not.

Step 2. Use fuzzy mathematics to fuzzify the training sample data, and convert whether the ST two markers are low risk (A file), mild risk (B file), moderate risk (C file), and extremely high risk. Risk (D file), the fuzzy table conversion method further includes the following steps:

1. For a given sample, each sample has dimensional original input, and the input sample data constitute $M$ rows and $N$ arrays. The characteristics of these $N$ dimensions include the profit performance fund flow statement and the financial indicators in the balance sheet.
2. Due to the large difference between financial characteristics, in order to obtain a more accurate training model, the calculation formula of financial ratio is required to standardize the financial data of different numerical types into a unified percentage. The finance used includes, but is not limited to, the rate of return on assets, the rate of return on total assets, and so on. The $M \times X$ dimension matrix is calculated by calculating the $M \times N$ matrix mentioned above. The number of samples represented represents the number of characteristics of the financial ratio calculated above, and the final data characteristic show whether it is marked by ST or not.
3. The total sample includes more than 80 industries, and different industries have different characteristics of financial ratios. This article needs to make corresponding fuzzy representations through fuzzy mathematics for these more than 80 industries and then take the A industry as an example to describe the process of fuzzy representation in detail. Assuming that industry A has K pieces of data, the size of the sample matrix of this industry is $K \times X$, and each dimension feature needs to be fuzzified. Let the sample set be expressed as

$$[S_1, S_2, \ldots, S_k], \quad (1)$$
where $S$ is a sample, and the data dimension characteristics of each sample are expressed as

$$[T_1, T_2, \ldots, T_x],$$

(2)

where $T$ is the characteristic value.

Here you need to perform fuzzy calculations on each feature to get the corresponding fuzzy value. Taking special $T_1$ as an example, select all the $T_1$ feature values in formula (1) and sort it in ascending order to get the feature set:

$$[T_{11}, T_{12}, T_{13}, \ldots, T_{1x}],$$

(3)

Assume that this feature is negatively correlated with financial risk, that is, the larger the value, the lower the degree of financial risk. According to the sorted eigenvalue set, the markers whose zero boundary point $T_{1m}$ is greater than this value are found to be non-ST, so that formula (3) can be divided into sets $T_{11}, T_{12}, T_{13}, \ldots, T_1 (w-1)$, and $T_1$; according to $T_1$, set $[T_{1w}, T_{1x}]$, assuming that the lower and upper limits of the set are $a_1$ and $a_2$, respectively; a semi-trapezoidal distribution is used to determine the degree of membership, and the membership function is as follows. Ascending semi-trapezoidal distribution:

$$\mu(x) = \begin{cases} 
0, & x \leq a_1, \\
\frac{x - a_1}{a_2 - a_1}, & a_1 \leq x < a_2, \\
1, & x > a_2,
\end{cases}$$

(4)

where $a_1$ and $a_2$ are the lower limit and upper limit of the set, respectively. Falling half trapezoidal distribution:

$$\mu(x) = \begin{cases} 
1, & x \leq a_1, \\
\frac{x - a_1}{a_2 - a_1}, & a_1 \leq x < a_2, \\
0, & x > a_2.
\end{cases}$$

(5)

Through this membership function, the membership value of each dimension feature in formula (2) can be calculated. Assuming that sample $A$ obtains the membership value $[A_1, A_2, A_3, \ldots, A_x]$ of all the features of the sample after calculation, add one-dimensional total membership value to the sample dimension, and the expression is

$$A_{x+1} = \sum_{k=1}^{x} A_k.$$ 

(6)

Among them, $A$ is the sample, the total membership value feature is selected, and the upper and lower limits of the feature are set to $b_1$ and $b_2$, respectively, and the ascending trapezoidal distribution is used to determine its final membership value. The membership function is

$$\mu(x) = \begin{cases} 
A, & x \geq b_1, \\
B, & \frac{b_1 - b_2}{2} \leq x < b_1, \\
C, & b_2 \leq x < \frac{b_1 - b_2}{2}, \\
D, & x < b_2.
\end{cases}$$

(7)

Fuzzy representation of training samples through fuzzy mathematics, conversion of two-category labels into multiclass labels through category membership value features, and multiclass decision making combined with random forest trees greatly improve the accuracy and breadth of decision making and are suitable for decision making in various scenarios. This is also the key point of this article. The total membership value feature can be converted into low risk ($A$ file), mild risk ($B$ file), moderate risk ($C$ file), and extremely high risk ($D$ file) by performing the above-mentioned fuzzification processing, and mark the feature as the new data. The convergence is compared in Figure 4.

**Step 3.** Use the random forest model to train the above-mentioned fuzzy processed sample data to obtain the trained random forest model: the construction method of the random forest model is further explained, including the following steps: (a) given the attributes of the training sample $S$, the number $Mw$ is an integer greater than 0 and less than $M$; (b) use the bootstrap method to sample the fuzzified training sample data, and randomly generate $k$ training subsets $[S_i \cdot S_k]$; use the bootstrap method to reduce the dimensionality and to resample the training sample data, which is to further say that the set $S$ contains $n$ different samples, $[X_i, X_m]$ extracts a sample from the set $S$ with replacement and draws $n$ times, and summarizes the extracted
samples to form a new set S'; use each training subset to generate the corresponding decision tree:

\[ \text{cart}_1, \text{cart}_2, \ldots, \text{cart}_n. \]  

Before selecting an attribute on each non-leaf child node, randomly select \( m \) attributes from the \( M \) attributes as the split attribute set of the current node, split the node by the best split method among the attributes, and grow in the entire forest. The value of \( m \) remains unchanged during the process. The evaluated data are shown in Figure 5. (d) Each tree grows completely without reducing branches until the training is completed. (e) For the test sample \( X \), use each decision tree to test to obtain the corresponding decision tree:

\[ \text{cart}_1(x), \text{cart}_2(x), \ldots, \text{cart}_n(x). \]  

\( \text{cart}() \) means a voting strategy on the results of these \( n \) decision trees and obtains the final predicted township result corresponding to the test sample. The final prediction result is further expressed as:

\[ H(x) = \arg \max_i \sum_{j=1}^k I(h_i(x) = Y), \]  

where \( H(x) \) is the predicted output, \( k \) is the number of decision trees \( 1 \) = trees, and \( h_i(x) \) is the model of the \( i \)-th decision tree.

**Step 4.** Input the prediction sample data into the trained random forest model and predict the corresponding test data, the output of which is the financial risk level of the data to be predicted. Analyze and select the data that affect the financial risk of the enterprise, as shown in Table 1.

The original corporate financial data are specifically cash flow statements, balance sheets and income statements published by more than 5,000 listed companies in Shanghai and Shenzhen from 2013 to 2017, with a total of more than 180 monetary funds, accounts receivable, etc., divided by quarter eigenvalues. From some of the financial indicators shown in Table 1, it is not difficult to see that there are great differences between different indicators, and the characteristic dimension is relatively high. It is necessary to use the calculation formula of the financial ratio to standardize the financial data of different numerical types into a unified percentage, which can effectively reduce the dimensionality of the training data, thereby improving the computational efficiency. Figure 6 shows examples of financial ratio indicators.

Analyzing Figure 6, it is not difficult to get the standardized data obtained after calculating the financial ratio. The dimension has been reduced from more than 180 dimensions before processing to 22 dimensions, and the financial ratio indicator can better describe the characteristics of the financial risk, and this step can significantly reduce the size of the model and the running time of the model. Fuzzy mathematics method is introduced to transform the data. The specific process of data fuzzification refers to Step 2 above. Figure 6 shows the results obtained after partial data fuzzification. There have been studies comparing RF algorithm with SVM and ANN algorithms, which proved the superiority of RF algorithm. At present, there are only two types of decision making for various scenarios through the random forest tree algorithm, which is also the shortcoming of the random forest algorithm.

As shown in Figure 7, after the original financial data are expressed in a fuzzy manner, binary data are converted into a multiclass problem, which also reflects a gradual characteristic of financial risk. The random forest model is used to train the above-mentioned fuzzy training samples, and the trained model is used to make predictions. The number of attributes of a given training sample \( S \) is an integer greater than 0. Before the model training, the initial setting \( N_{\text{tree}} = 30 \), which represents the number of decision trees. It can be adjusted according to the prediction error in the future:

\[ M_{\text{try}} = \log_2 (M + 1). \]  

Among them, \( M_{\text{try}} \) represents the number of nodes when the decision tree is split and will not change during the experiment. First, the bootstrap method is used to sample the original financial data. Through this process, \( K \) datasets are randomly generated \( [S_1, S_2] \); use this training dataset to generate the corresponding decision tree \( \text{CART}1, \text{CART}n \), and the decision tree model is defined as \( h(x) \), where \( x \) represents the input vector; before selecting attributes on each non-leaf node, it is randomly selected During \( m \) from attributes \( M \). During the training, each decision tree will grow completely, and no pruning operation will be performed in this process until the training is completed. Compared with the data \( X \) used for testing, the data are input to each decision tree for testing, and the result is obtained. The voting strategy is adopted on these \( k \) results, and the final prediction result is
### Table 1: The financial risk of the enterprise.

| Company | Time       | Funds         | Accounts receivable | ... | Undistributed profit | ST or not |
|---------|------------|---------------|---------------------|-----|----------------------|-----------|
| A       | 2017/12/31 | 169593.87     | 36083.44            | ... | 170469.36            | No        |
| B       | 2016/12/31 | 12193401      | 12802844.3          | ... | 5441107.2            | No        |
| C       | 2017/12/31 | 3524.01       | 38659.75            | ... | −119953.55           | Yes       |
| D       | 2014/6/30  | 9839.99       | 396.03              | ... | −92647.31            | Yes       |
\[ H(x) = \arg \max \sum_{i=1}^{k} I(h_i(x) = Y), \]

where \( t \) is the \( t \)-th decision tree. In order to evaluate the pros and cons of the model, the average relative error is used to analyze the predicted results. The average relative error is

\[ \text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - y_f}{y_f} \right| \times 100\%, \]

where \( y_f \) represents the predicted value, \( t \) represents the actual value, and \( N \) represents the total amount of training data. MAPE is used to evaluate the deviation between the predicted value and the actual value. The smaller the value, the better the prediction effect. In order to analyze the prediction effect of the present invention, the present invention uses the fuzzy mathematics processing sample and the model trained by the original financial index to compare the prediction results. The predicted value is shown in Figure 8.
Analyzing the data in the table, it can be known that the financial risk prediction results of the listed companies using the present invention are basically consistent with the actual results, which also shows the feasibility of the present invention; compare whether the prediction results of fuzzy mathematics and the characteristics of the financial risks in the table are basically similar, and the fuzzy mathematics model can further refine the financial risks to a certain level, which fully reflects the effectiveness of the present invention. In summary, the application of the present invention can better predict financial risks and has high engineering use value. The evaluated data are shown in Figures 9 and 10. The advantages of high prediction accuracy, controllable generalization error, fast convergence speed, and few adjustment parameters can effectively avoid the occurrence of overfitting and are especially suitable for high-level data calculations.

4. Conclusion

With the continuous and rapid development of China’s economy, the operating environment of listed companies has become more and more complex, and the increasing pressure of international competition among companies has made the issue of financial risks of listed companies more severe. If you do not pay attention to the financial risk status of the enterprise, it will cause the financial risk to accumulate and eventually cause a financial crisis, which will be marked by ST. Therefore, this paper proposes an early warning model of enterprise financial accounts combining fuzzy sets and random forest trees, which specifically includes the following steps. First, the dataset is analyzed, selected and initially constructed by the training prediction sample. It is further explained by the data labels, and is charged whether the label is marked by ST or not. Then, the method of fuzzy mathematics is used to fuzzify the training sample data, and the two-category label is converted into a multiclass label; then, the random forest model is used to train the above-mentioned fuzzified sample data. Obtain the trained random forest model. Finally, input the prediction sample data into the trained random forest model to make decisions on the scene application. At the same time, the invention is applied to the enterprise financial risk early warning, which demonstrates the practicability and effectiveness of the invention. The significant advantage of the present invention is that the two-class decision making is converted into the multiclass decision making by combining the fuzzy set and the random forest model, which greatly improves the prediction accuracy, efficiency, and data rationality.

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 that they have no conflicts of interest.

Acknowledgments

This study was supported by the Shandong Youth University of Political Science-Level Project (Research on the Construction and Practice Path of Modern Service Industry Internal Audit Personnel Training System) (no. xxpy-yyyyyb13).

References

[1] Y. Yu, C. Yang, Q. Deng, T. Nyima, S. Liang, and C. Zhou, "Memristive network-based genetic algorithm and its application to image edge detection," Journal of Systems Engineering and Electronics, vol. 32, no. 5, pp. 1-9, 2021.
[2] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," Procedia Computer Science, vol. 60, no. 1, pp. 1472-1481, 2015.
[3] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," IEEE Trans. on CAD of Integrated Circuits and Systems, vol. 37, no. 9, pp. 1774–1787, 2018.
[4] J. Pais, "Random matching in the college admissions problem," Economic Theory, vol. 35, no. 1, pp. 99–116, 2018.
[5] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," Decision Support Systems, vol. 28, no. 4, pp. 291–384, 2020.
[6] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," Journal of Knowledge Management, vol. 21, no. 3, 2017.
[7] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," Tourism Management, vol. 50, no. 10, pp. 31-40, 2015.
[8] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 37, no. 2, pp. 324–336, 2018.
[9] C. Koch and S. P. Penczynski, "The winner’s curse: conditional reasoning and belief formation,” Journal of Economic Theory, vol. 174, pp. 57–102, 2018.
[10] C. K. Karl, "Investigating the winner’s curse based on decision making in an auction environment," Simulation & Gaming, vol. 47, no. 3, pp. 324–345, 2016.
[11] D. Ettinger and F. Michelucci, "Creating a winner’s curse via jump bids," Review of Economic Design, vol. 20, no. 3, pp. 173–186, 2016.
[12] J. A. Brande and E. J. Egan, “The winner’s curse in acquisitions of privately-held firms,” The Quarterly Review of Economics and Finance, vol. 65, pp. 249–262, 2017.
[13] Z. Palmowski, “A note on var for the winner’s curse,” Economics/Ekonomia, vol. 15, no. 3, pp. 124–134, 2017.
[14] B. R. Routledge and S. E. Zin, “Model uncertainty and liquidity,” Review of Economic Dynamics, vol. 12, no. 4, pp. 543–566, 2009.
[15] D. Easley and M. O’Hara, “Ambiguity and nonparticipation: the role of regulation,” The Review of Financial Studies, vol. 22, no. 5, pp. 1817–1843, 2019.
[16] P. Klibano, M. Marinacci, and S. Mukerji, “A smooth model of decision making under ambiguity,” Econometrica, vol. 73, no. 6, pp. 1849–1892, 2005.
[17] Y. Halevy, "Ellsberg revisited: an experimental study," Econometrica, vol. 75, no. 2, pp. 503–536, 2017.
[18] D. Ahn, S. Choi, D. Gale, and S. Kariv, “Estimating ambiguity aversion in a portfolio choice experiment,” Working paper, vol. 5, no. 2, pp. 195–223, 2019.

[19] T. Hayashi and R. Wada, “Choice with imprecise information: an experimental approach,” Theory and Decision, vol. 69, no. 3, pp. 355–373, 2010.

[20] K. Zima, E. Plebankiewicz, and D. Wieczorek, “A swot analysis of the use of bim technology in the polish construction industry,” Buildings, vol. 10, no. 1, 2020.

[21] P. Sun, B. Liu, and T. Sun, “Injury status and strategies of female 7-a-side rugby players in Anhui Province,” Sports Boutique, vol. 38, no. 3, pp. 72–74, 2019.

[22] P. Guild, M. R. Lininger, and M. Warren, “The association between the single leg hop test and lower-extremity injuries in female athletes: a critically appraised topic,” Journal of Sport Rehabilitation, vol. 30, no. 2, pp. 1–7, 2020.

[23] U. G. Inyang, E. E. Akpan, and O. C. Akinyokun, “A hybrid machine learning approach for flood risk assessment and classification,” International Journal of Computational Intelligence and Applications, vol. 19, no. 2, Article ID 2050012, 2020.

[24] Q. Liu, S. Du, B. Wyk, and Y. Sun, “Double-layer-clustering differential evolution multimodal optimization by speciation and self-adaptive strategies,” Information Sciences, vol. 545, no. 1, pp. 465–486, 2021.

[25] H. R. Medeiros, F. D. Oliveira, H. F. Bassani, and A. Araujo, “Dynamic topology and relevance learning SOM-based algorithm for image clustering tasks,” Computer Vision and Image Understanding, vol. 179, pp. 19–30, 2019.

[26] Y. Deng, D. Huang, S. Du, G. Li, and J. Lv, “A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis,” Computers in Industry, vol. 127, Article ID 103399, 2021.

[27] J. J. Chan, K. K. Chen, S. Sarker et al., “Epidemiology of Achilles tendon injuries in collegiate level athletes in the United States,” International Orthopaedics, vol. 44, no. 3, pp. 585–594, 2020.

[28] W. Li, G. G. Wang, and A. H. Gandomi, “A survey of learning-based intelligent optimization algorithms,” Archives of Computational Methods in Engineering, vol. 28, pp. 1–19, 2021.

[29] G. G. Wang, A. H. Gandomi, A. H. Alavi, and D. Gong, “A comprehensive review of krill herd algorithm: variants, hybrids and applications,” Artificial Intelligence Review, vol. 51, no. 1, pp. 119–148, 2019.