Research Article

Research on Audit Data Analysis and Decision Tree Algorithm for Benefit Distribution of Enterprise Financing Alliance

Jie Bai [1] and Tian He[2]

[1]*Shanxi Vocational & Technical College of Finance & Trade, Shanxi 030031, Taiyuan, China*
[2]*Shanxi University of Political Science and Law, Shanxi 030031, Taiyuan, China*

Correspondence should be addressed to Jie Bai; 201705060125@sust.edu.cn

When traditional methods analyze the audit data of enterprise financing alliance, there are some problems, such as long algorithm modeling time and low accuracy of interest distribution algorithm of enterprise financing alliance. Therefore, this paper proposes an analysis method of interest distribution of enterprise audit data financing alliance based on the decision tree algorithm. The audit data collection process of enterprise financing alliance is given, and the continuous attributes of audit data are discretized by the C4.5 algorithm. We perform enterprise financing alliance audit data analysis, remove inconsistencies from audit data through data cleaning, and finally realize enterprise financing alliance audit data analysis based on the improved C4.5 algorithm. The experimental results show that this method can shorten the modeling time and improve the accuracy of interest distribution algorithm of enterprise financing alliance. We achieved an average accuracy of 84.7% with the C4.5 algorithm while 84.35% with NBTree.

1. Introduction

Studies have shown that establishing a financing cooperation alliance between small and medium-sized enterprises (SMEs) is an effective way to solve the financing difficulties of SMEs. However, in practice, with the continuous development of SME financing alliance, the problem of instability of financing alliance is gradually exposed. Among them, the imbalance of benefit distribution of financing alliance has become one of the factors leading to the instability of financing alliance [1]. At this stage, most of the research on the distribution of interest has evolved from the game, for example, using the expected income of the pairwise game under different situations of the three parties to study whether banks, small and medium-sized enterprises and guarantee institutions participate in cooperation under the condition of incomplete information.

With the intensification of economic globalization and technology, fundamental changes have taken place in the survival and development of enterprises in many aspects such as the changing technical level, the rapidly changing market environment, the shortened product life cycle, the more personalized, diversified, and differentiated customer demand, and the higher requirements for product quality and service [2]. Facing the complex and fierce competitive environment, if enterprises want to get a place and improve their competitive advantage, innovation has gradually become the key to enhance their competitiveness. The Fifth Plenary Session of the 18th CPC National Congress proposed that innovation is the first driving force leading development, and innovation driven development is a major strategy for enterprises facing the future [3]. In terms of the current situation, a single enterprise has limited resources. In order to meet the personalized needs of customers, the product life cycle needs to be shortened, and risks need to be reduced. It is difficult for enterprises to fully rely on their own strength to achieve their goals, which is necessary to combine with other enterprises with complementary resources. In order to analyze the audit data of benefit distribution of enterprise financing alliance, this paper introduces the decision tree algorithm. The ID3 algorithm forms a decision tree and uses mutual information to find

the attribute field with maximum amount of information in database. The C4.5 algorithm is used to create a data mining model. This model is used to predict the classification authenticity, and the records are considered suspicious if the predicted results are different from the actual distribution.

The rest of the paper is organized as following. Section 2 discusses the computer-based audit process. In Section 3 and the following sections, we present an overview of the decision tree algorithm and audit data analysis and talk about data acquisition. In Section 4, we talk about audit data analysis based on the C4.5 algorithm. In Section 5, we discuss the experiments and the results, and Section 6 is the conclusion of the proposed work.

2. Computer Audit

Computer audit is also known as computer-aided audit. Since the emergence of the word "computer audit," there has been a vague understanding of the concept of computer audit. With the continuous development of theory and practice, people have a deeper understanding of computer audit. At present, it mainly includes two aspects: first, auditors use audit software to check the accounting system. It is an audit of whether the economic activities and accounting materials of the audited unit are true, correct, legal, and effective, which is basically consistent with the traditional audit content, but the means have changed, which is generally called computer-aided audit. Second, testing and evaluation of the electronic information system itself refers to the collection and evaluation of audit evidence to judge the security of resources and assets related to the information system of the auditee, the integrity of data and system, etc. [4]. Figure 1 describes the main process of using the computer audit technology to audit the electronic data of the audited unit.

Pretrial investigation: by grasping the organizational structure of the units being audited, the general situation of the distribution and use of the information system of the unit is mastered. Determine the key subsystems according to the audit objectives, conduct further in-depth investigation on them, and put forward the audit data requirements.

Audit data collection: according to the data requirements of pretrial investigation, determine the objects and means of data collection, and grasp the overall information of audit objects [5].

Data collation: due to the complexity of the auditee's information system and possible deliberate concealment and creation, the collected data must be cleaned and transformed. If necessary, the data from multiple data sources must be integrated to ensure the integrity and consistency of the data and lay a necessary foundation for the subsequent audit work [6].

Audit analysis: for the integrated, complete, and consistent data stored in the database, establish the corresponding analysis model according to the audit focus and analyze the data from different layers and different angles.

Extension, implementation, and evidence collection: since the analysis process is subjectively processed by people, it is inevitable that there will be analysis errors or even errors. According to the results of audit analysis, collect evidence for
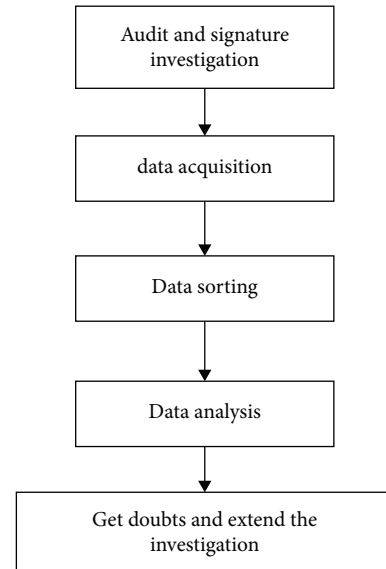

FIGURE 1: Process of computer aided audit.

the problems found and further verify and implement the clues found [7].

From the above process, it can be seen that the sorting and modeling analysis of audit data are two key links closely related to the audit results. Their implementation will directly affect the establishment quality of audit database and further have a significant impact on audit analysis [8].

At present, there is no precedent for authenticity prediction in public rental housing audit. After discussion and analysis, we decided to collect the information of vehicles, industrial and commercial registration, and endowment insurance in addition to the real estate information and provident fund loan information, so as to generate a decision tree and to extract rules to predict the authenticity. In terms of specific operation, the first step is to randomly divide all public rental housing information into two parts. One part is the test data set, and the other part is the training data set. The second step is to use the C4.5 algorithm, which is the training data set to create the data mining model. The third step is to use the created model to "predict" the classification authenticity of the former, that is, the test data set, and obtain records that the prediction results are different from the actual distribution. The fourth step is to find the idea of improving the C4.5 algorithm through the research on the improvement of the decision tree algorithm and finally get the improved result, compare it with the original result, and evaluate its accuracy. Step 5: if there are records that the predicted results are different from the actual distribution, it can be considered as suspicious data, treated as abnormal conditions, and conducted in-depth audit, analysis, and investigation.

3. Audit Data Analysis Based on Decision Tree Algorithm

3.1. ID3 Decision Tree Algorithm. Decision tree is a tree structure similar to binary tree or multi tree, which is used for predictive modeling of discrete and continuous

attributes. The decision tree uses the attribute of the sample as the node and the value of the attribute as the branch, which is a process similar to the flow chart. Each internal node represents the test on an attribute, each branch represents a test output, and each leaf node represents a class or class distribution. It adopts the recursive method of top-down, divide and conquer. If all examples in the training sample set belong to the same class, they will be regarded as a leaf node and identified as this class [9]. If not, first determine a test attribute according to a certain method and divide the sample set into multiple subsets—internal nodes according to different values of the test attribute—so as to ensure that the samples have the same attribute values on the same subset. Then, each subset is processed repeatedly until the classification attributes satisfying the conditions are obtained. The root node is the attribute with the largest amount of information in all samples, the intermediate node is the attribute with the largest amount of information in the sample subset contained in the subtree with this node as the root, and the leaf node of the decision tree is the category value of the sample [10].

ID3 learning algorithm, which forms a decision tree through the selection window, uses the mutual information (information gain) in information theory to find the attribute field with the maximum amount of information in the database. It establishes a node of the decision tree and then establishes the branch of the tree according to the different values of the attribute field. Repeat the process of establishing the lower nodes and branches of the tree in each branch subset [11]. The pseudocode is as follows:

Generate_decision_tree(samples, attribute): a decision tree is generated from the given training data.

> **Input**: training samples, represented by discrete value attribute; Collection of candidate attributes attribute_ list.
>
> **Output**: a decision tree.
>
> **Method**: Generate_decision_tree(samples, attribute_list)

(1) Create node N;

(2) if samples All in the same class C then//The value of class label attribute is C, its candidate attribute values are not considered.

(3) return N is the leaf node, marked with class C;

(4) if attribute_list Empty then.

(5) return N, as a leaf node, is marked as the most common class in samples//The one with the largest number of class label attribute values.

(6) Select attribute_ Best attribute with the highest information gain in list_ attribute; //Find the best partition attribute.

(7) Mark node N as best_ attribute;

(8) for each best_ Unknown value ai in attribute//set samples as best_ Attribute

(9) A condition of best grows from node N_ Attribute = branch of ai;

(10) Let si be the best in samples_ Attribute = set of samples of ai//a partition.

(11) if si Empty then.

(12) Add a leaf and mark it as the most common class in samples//Find out the one with the largest number of class labels from the sample as the tag of this node.

(13) Else plus one by generate_ decision_ tree(si, attribute_list–best_ Attribute) returned node//Call the data subset Si recursively. At this time, the candidate attribute has been deleted_ attribute.

The advantages of ID3 algorithm are as follows:

(1) The basic principle of the algorithm is clear

(2) Fast classification speed

(3) A practical example learning algorithm

The ID3 algorithm also has some disadvantages such as the following:

(1) There is a bias problem, and the number of values of each characteristic attribute affects the amount of information

(2) A little problem with the training data will make the results different and sensitive to noise

(3) The probability of error is proportional to the increase of category

3.2. Application of Decision Tree Algorithm in Audit. According to the work plan deployment of the audit department, auditors will audit the authenticity of a public rental housing distribution. At present, the judgment standard of public rental housing allocation is mainly whether there is a real estate, the size and purpose of the real estate, and the number of family members. The specific situation often needs field visit and investigation. However, although some audit doubts have been found according to this standard, many doubts have been ignored. For example, some people who enjoy affordable housing give their real estate to their relatives or parents by means of ownership transfer in the recent period of audit in order not to be found out in the audit process [12]. In order to improve the efficiency of field visit and investigation and prevent false information reporting (such as real estate information fraud), relevant data of multiple units should be collected, the classification rules of identified doubtful points of affordable housing allocation should be excavated, and the authenticity of normally allocated personnel should be predicted, so as to select some suspicious personnel for extended audit [13].

At present, there is no precedent for authenticity prediction in public rental housing audit. After discussion and analysis, we decided to collect the information of vehicles, industrial and commercial registration, and endowment insurance in addition to the real estate information and provident fund loan information, so as to generate a decision tree and to extract rules to predict the authenticity.

In terms of specific operation, the first step is to randomly divide all public rental housing information into two parts: one part is the test data set and the other part is the training data set.

The second step is to use the C4.5 algorithm, which is the training data set to create data mining model [14]. The third step is to use the created model to "predict" the classification authenticity of the former, that is, the test data set, and obtain records that the prediction results are different from the actual distribution. The fourth step is to find the idea of improving the C4.5 algorithm through the research on the improvement of decision tree algorithm and finally get the improved result, compare it with the original result, and evaluate its accuracy [15]. In Step 5, if there are records that the predicted results are different from the actual distribution, it can be considered as suspicious data This is treated as abnormal conditions, and in-depth audit, analysis, and investigation were conducted.

3.3. Data Acquisition. Audit data collection is the working premise of data mining in audit application and an important step to obtain the electronic data of the audited unit [16]. The audit practice shows that the acquisition can be realized in two steps: the first step is to accurately select the electronic data to be collected. The second step is to effectively collect the target data. The process is shown in Figure 2.

3.4. Data Selection Strategy. In order to efficiently select the data concerned by auditors from various complex data of the auditee, the collection shall follow the following three principles.

(1) The selected data shall meet the requirements of the audit scheme. The audit scheme describes the main work contents and requirements of the audit. Through the analysis of the audit scheme, master the audit contents and requirements and determine the object and scope of data collection according to these contents and requirements, which is the first step of data collection [17].

(2) Data selection shall be based on a full understanding of the auditee's information system and its business process. By analyzing the generation and use relationship between business processes and data, combined with the needs of audit business, reasonably determine the object and scope of collection.

(3) Data selection should not be limited to specific auditees. Using external data for correlation analysis is an effective means. Generally, there are specific relationships between the data of some departments, such as between enterprise reports and the records of tax departments and between enterprise sales data and Golden Tax engineering data. In the audit process, not only can external data be used to verify the data of the auditee but also external data can be used to replace the data of the auditee for audit when necessary [18]. For example, in this audit, in addition to collecting the data of the audited units such as real estate and provident fund loans, in order to conduct effective audit analysis, relevant data of vehicle, industrial and commercial registration, endowment insurance, and other departments were also collected.

4. Audit Data Analysis

4.1. Data Cleaning. Audit data cleaning is an important part of computer data audit. Its main work is to remove attributes and redundant attributes irrelevant to mining topics, smooth noise data, and fill in vacancy values, so as to serve for subsequent audit data analysis. Although most current classification algorithms (such as C4.5 algorithm) integrate the function of processing noise and vacancy values, however, doing this step well before data mining helps to reduce the deviation in the learning process. Data cleaning can be performed before or after data conversion. Audit data cleaning mainly includes confirming the input data, modifying the wrong value, replacing the null value (e.g., "0"), ensuring that the data value falls into the definition domain, eliminating redundant data, and resolving conflicts in data. This will have a great impact on the results of data mining [19]. According to different data conditions, there are mainly three technical methods for processing:

(1) Completing the incomplete data

(a) Ignore tuple method: it can be used when the class label is missing, or when multiple attributes of the tuple are missing values.

(b) Fill in the missing value manually: this method is applicable when there are few missing values. If the workload is too large, this method may not work.

(c) Fill the missing values with a global constant: a constant can be used to replace the missing attribute value. For example, in this affordable housing audit, the number 0 is used to replace the missing data in the number plate type.

(d) Populates missing values with the average value of the attribute.

(e) Fill in the missing values with the most likely values. Use reasoning-based tools to determine, such as regression or decision tree induction.

(2) Processing noise data

(a) Binning: by examining the "neighbors" (i.e., the surrounding values) to smooth the value of the stored data, the box average smoothing or box boundary smoothing is mainly used to divide the stored values into "bucket" or "box."

(b) Clustering: detection of outliers by clustering [20]. The data are clustered unsupervised. The records that cannot be clustered are considered as outliers and can be treated as noise.

(c) Combination of computer and manual inspection: computer detection and manual detection
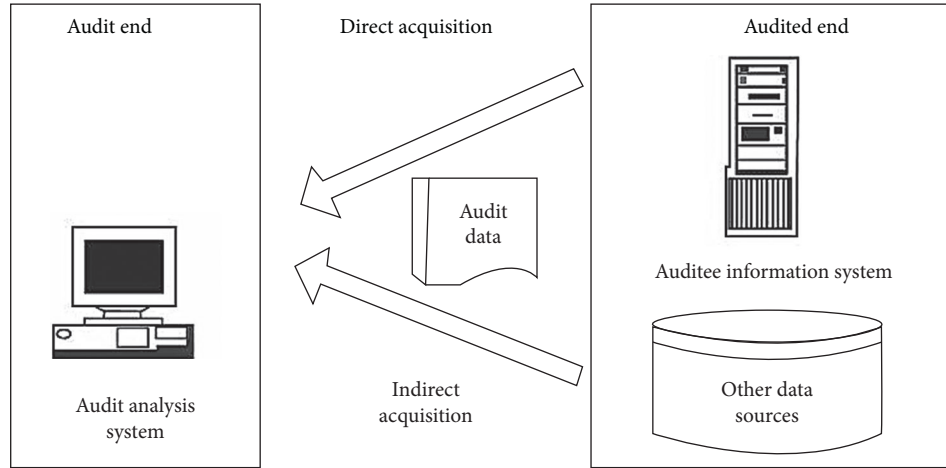
FIGURE 2: Audit data collection process.

methods are combined to find outliers and other data in the data set and clean them appropriately.

(3) Inconsistent data correction

In addition to missing data, there are data inconsistencies in the original data. In fact, by analyzing the correlation between data attributes, integrity constraints can be defined to detect inconsistencies and correct these inconsistent data. For example, by observing and analyzing the intermediate table of public rental housing audit data, it is found that some attributes are vacant and lack attribute values. For example, dkye (loan balance) attribute lacks corresponding data values. At this time, the method of ignoring tuples can be used to delete these records. If the attribute value of dkrq (loan date) is earlier than that of KHRQ (account opening date) in the table, such inconsistencies in the data shall be noted and corrected to clean up the data.

4.2. Data Conversion.

In the original data obtained in this paper because in the process of input, the personal habits of the staff cause the data format of many data to be inconsistent. In order to better carry out data mining, this paper first needs to unify the data format through the method of data conversion.

For example, the formats such as issuing date and house purchase date are uniformly converted to date format.

In this paper, the C4.5 algorithm can be used to discretize continuous attributes. In the existing public rental housing information table, the payment base and construction area are continuous, so it needs to be discretized. Discretization technology can also provide hierarchical or multidimensional division of attribute values, which can be called conceptual stratification. The main methods are: box division, histogram analysis, cluster analysis, entropy-based discretization, and data segmentation through natural division. In this paper, we use the box method to discretize the attribute values. The payment base is the base of endowment insurance. In order to better carry out data mining, the payment base and construction area can be divided into three boxes. The implementation process of discretization will be carried out in Weka.

4.3. Audit Data Analysis Based on the Improved C4.5 Algorithm.

In this paper, the C4.5 algorithm is improved accordingly. The degree of interest is introduced in the weighting process λ. It is called the user's interest in uncertain knowledge and is a fuzzy concept. It usually refers to the prior knowledge about a certain transaction, including domain knowledge and expert suggestions. Its size is determined by the decision-maker according to the prior knowledge or domain knowledge. The improved C4.5 algorithm modifies the original attribute selection criteria. When calculating the information gain of an attribute, the introduction of positive interest will reduce the support of the attribute for decision-making, that is, if the entropy increases, the information gain will be small, and the information gain rate will also be reduced. On the contrary, the introduction of negative interest will increase the support of the attribute, that is, when the entropy decreases, the information gain will increase, and the information gain rate will also increase.

Due to

$$P(V_i) = \frac{|T_i|}{|T|}. \tag{1}$$

Regulations:

$$-|T_i| \leq \lambda \leq |T_i| \, (i = 1, 2, \ldots, n), \tag{2}$$

among λ. The size is determined by $|T_i|$ of the attribute. Introducing interest λ, formula (2) becomes

$$H_\lambda\left(\frac{C}{V}\right) = -\sum_j p(V_j) + \frac{\lambda}{|T|} \sum_i \left(\frac{C_j}{V_i}\right) \log_2 \, p\left(\frac{C_j}{V_i}\right) = \text{info}_{v\lambda}(T), \tag{3}$$

or

$$I_\lambda\left(\frac{C}{V}\right) = H(C) - H_\lambda\left(\frac{C}{V}\right) = \text{Info}(T) - \text{Info}_{v\lambda}(T) = gain_\lambda(v). \tag{4}$$
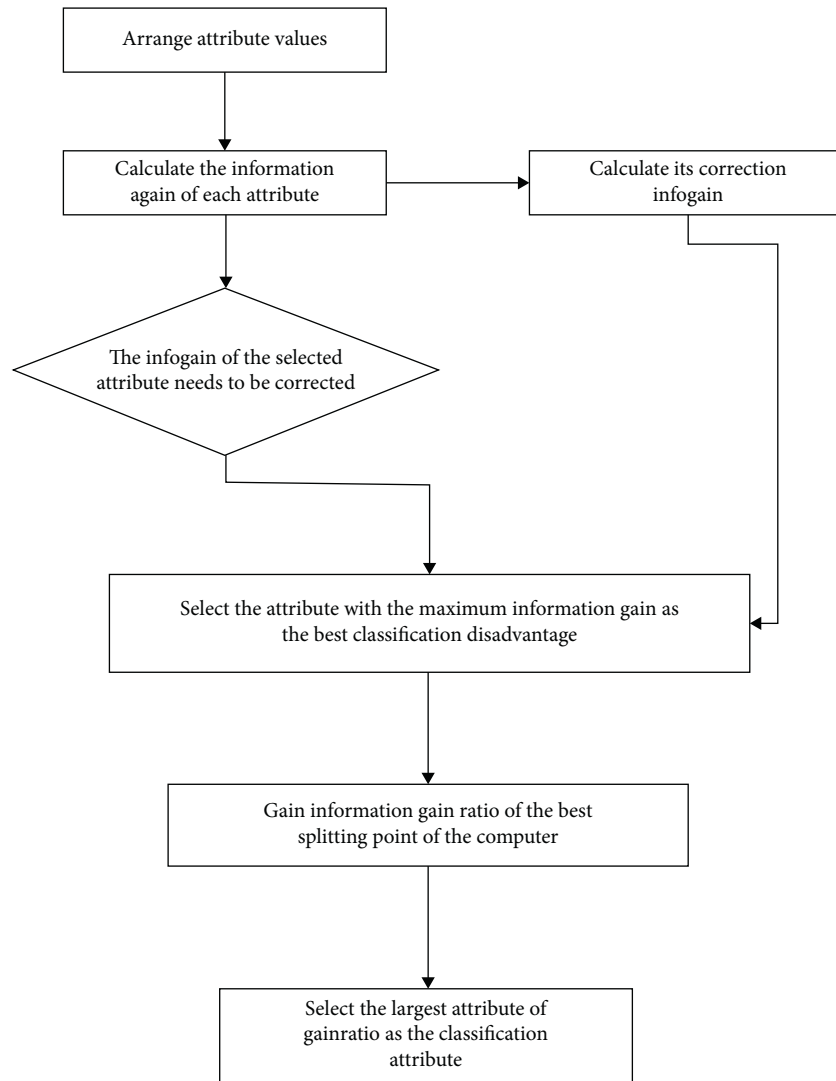
Formula (1) becomes

FIGURE 3: Flowchart of the improved C4.5 algorithm.

$$\text{Gain\_ratio}(v) = I_\lambda \frac{(C, V)}{H(V)} = \frac{\text{gain}_\lambda(V)}{\text{split\_Info}(v)}. \qquad (5)$$

The degree of interest can improve or reduce the support of the selected attributes for decision-making, in which the degree of interest is positive $\lambda > 0$. When using, the information gain of appropriate attributes should be corrected, and the correlation changes dynamically with $|T_i|$ of different attributes and the size of $|T|$ of training samples.

The selection steps of classification attributes by the improved C4.5 algorithm are shown in Figure 3

5. Performance Analysis Experiment of Decision Tree Algorithm

We selected 11 sample data sets from the data set published on Weka website to analyze and compare the algorithm performance of the decision tree algorithm C4.5, cart, and NBTree algorithm in Weka. In the selection of sample data sets, the data size of these 11 sample sets is different, and they all contain both discrete attributes and continuous

attributes, so they can comprehensively test the performance of each algorithm. The experimental sample data is shown in Table 1.

Among them, meta, optdigits and aggregation data sets are large. The above 11 groups of data are classified and processed by the above three algorithms on Weka.

5.1. Evaluation Index of Decision Tree Algorithm. Among many decision tree algorithms, some deal with discrete data, some deal with continuous data, some consider the shortest modeling time, and some consider the best classification effect. Different decision tree algorithms may have many differences in processing data types, selection of modeling mechanism, construction methods of decision tree, expression methods of classification rules, and so on. So, in the specific affordable housing audit process, how to select the appropriate classification method is to improve the audit efficiency? Next, we analyze the performance of decision tree algorithm based on Weka platform.

TABLE 1: Experimental sample data.

Sample set	Number of records	Number of attributes
E. coli	214	10
Fruitfly	286	10
Gascons	24	5
Meta	810	20
Optdigits	1000	21
Pollution	768	9
Servo	351	35
Sonar	150	3
Strike	57	17
Vowel	683	36
Wisconsin	435	17

TABLE 2: Comparison of algorithm accuracy.

Accuracy (unit%)	C4.5	Cart	NBTree
E. coli	66.82	70.56	68.22
Fruitfly	75.52	69.23	70.98
Gascons	83.33	79.17	70.83
Meta	95.73	95.07	93.87
Optdigits	70.5	73.9	75.3
Pollution	73.83	75.13	73.57
Servo	91.45	89.74	88.89
Sonar	96	95.33	94.67
Strike	76.68	78.95	85.96
Vowel	91.51	91.07	91.51
Wisconsin	96.32	95.4	95.63
Average	84.69	84.35	85.68

There are basically two indexes used to compare and evaluate the performance of decision tree algorithm:

(1) *Accuracy*. The accuracy of classification is the primary condition to judge the advantages and disadvantages of the algorithm, and high accuracy can provide people with reliable classification information

(2) *Calculation Speed*. The computational cost of generating and using the model is directly proportional to the efficiency of the algorithm

5.2. Experimental Result

5.2.1. Accuracy. Table 2 presents the results obtained from 11 groups of sample datasets. The average accuracy of the algorithm C4.5 is 84.69%, that of cart is 84.35%, and that for NBTree is 85.68%. Therefore, NBTree and C4.5 algorithms have good classification results for this kind of data.

By analyzing Table 2, we can see that the accuracy of interest distribution algorithm of enterprise financing alliance is different under different methods.

5.2.2. The Modeling Speed of the Algorithm. As shown in Table 3, C4.5 algorithm has the fastest modeling speed, cart algorithm takes the second place, and NBTree is the slowest. When the data set is large, such as meta, optdigits, and aggregation, the time required for cart increases rapidly, while C4.5 algorithm is relatively stable.

TABLE 3: Comparison of algorithm modeling time.

Modeling time (unit: s)	C4.5	Cart	NBTree
E. coli	0.06	0.09	1.68
Fruitfly	0.09	0.16	1.02
Gascons	0.01	0.06	0.01
Meta	0.97	1.15	22.95
Optdigits	0.53	1.14	0.83
Pollution	0.13	0.23	0.26
Servo	0.16	0.32	5.9
Sonar	0.01	0.02	0.17
Strike	0	0.03	0.44
Vowel	1.28	2.1	39.88
Wisconsin	0.14	0.43	1.96

6. Conclusion

This paper presents an audit data analysis method of interest distribution of enterprise financing alliance based on the decision tree algorithm. In this paper, the decision tree model obtained by improving the C4.5 algorithm is more in line with the benefit distribution rules of enterprise financing alliance. Experiments also verify the effectiveness and rationality of the improved algorithm. However, in the experiment, it is found that while the accuracy of the improved model is better, it also increases the complexity of the decision tree.

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 declares no conflicts of interest.

References

[1] S. Junxin, L. Qin, and L. Xinan, "Research on PPP financing benefit coordination mechanism based on evolutionary game," *China collective economy*, vol. 635, no. 15, pp. 106–109, 2020.

[2] B. Nan, "Research on the impact of equity financing demand on dividend distribution of Listed Companies in China," *Finance*, vol. 16, no. 14, pp. 94-95, 2020.

[3] D. U. Jingxin and T. Niancheng, "Research on income distribution of receivables financing based on B2B E-commerce platform," *Logistics Sci-tech*, vol. 042, no. 012, pp. 156–159, 2019.

[4] Z. D. Dan and G. G. Jun, "E-TOPSIS improved Shapley value of receivables financing income distribution," *Manufacturing Automation*, vol. 42, no. 5, pp. 41–48, 2020.

[5] L. Ying, "Research on income distribution of cross-border financial leasing from the perspective of Supply Chain Finance – based on the improved Shapley value analysis method," *Logistics technology*, vol. 42, no. 04, pp. 158–165, 2019.

[6] L. L. hua and D. C. lie, "Multi-layer classification storage simulation of complex network user privacy data," *Computer Simulation*, vol. 37, no. 3, pp. 410–444, 2020.

[7] Q. Wan and W. Li, "Audit data analysis based on improved apriori algorithm," in *Proceedings of the 2020 5th International*

Conference on Control, Robotics and Cybernetics (CRC), Wuhan, China, October 2020.

[8] N. Tojiboyev, D. Appelbaum, A. Kogan, and M. Vasarhelyi, "Basics of SQL for audit data retrieval and analysis," *Journal of Emerging Technologies in Accounting*, vol. 16, no. 23, pp. 67–83, 2021.

[9] X. Cai, Y. Qian, Q. Bai, and W. Liu, "Exploration on the financing risks of enterprise supply chain using Back Propagation neural network," *Journal of Computational and Applied Mathematics*, vol. 367, pp. 412–457, 2019.

[10] S. Wiyono, D. S. Wibowo, and M. F. Hidayatullah, "Comparative study of KNN, SVM and decision tree algorithm for student's performance prediction," *International Journal of Computing Science and Applied Mathematics*, vol. 6, no. 2, pp. 50–67, 2020.

[11] J. Li, J. Cai, F. Khan et al., "A secured framework for sdn-based edge computing in IOT-enabled healthcare system," *IEEE Access*, vol. 8, pp. 135479–135490, 2020.

[12] B. Lalithadevi, "Novel technique for price prediction by using logistic, linear and decision tree algorithm on deep belief network," *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 5, pp. 1751–1761, 2020.

[13] Q. Chen, Y. Zhang, and L. Chen, "A study of internet development and enterprise financing in China," *Networks and Spatial Economics*, vol. 54, no. 34, pp. 66–87, 2021.

[14] M. A. Jan, F. Khan, R. Khan et al., "Lightweight mutual authentication and privacy-preservation scheme for intelligent wearable devices in industrial-CPS," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5829–5839, 2020.

[15] W. Li, Y. Chai, F. Khan et al., "A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system," *Mobile Networks and Applications*, vol. 26, pp. 234–252, 2021.

[16] J. Liu, "Research on the prevention and control of supply chain financial risks from the perspective of enterprise financing," *Financial Marketing*, vol. 6, no. 1, pp. 49–57, 2021.

[17] M. Krutwig and A. Tantau, "A data analysis of a SME energy audit program in Germany," *XIX International Multidisciplinary Scientific GeoConference SGEM*, vol. 19, no. 23, pp. 87–98, 2019.

[18] C. Benítez, C. C. Castro, and D. Domínguez, "Data analysis as an alternative for evaluating the impact for social programs," vol. 65, no. 34, pp. 87–98, 2020.

[19] W. He, "The relationship between enterprise financing and enterprise life cycle," *Advances in Management and Applied Economics*, vol. 10, no. 25, pp. 55–68, 2020.