Hindawi
Journal of Mathematics
Volume 2022, Article ID 2231112, 12 pages
https://doi.org/10.1155/2022/2231112

Research Article

System Optimization of Talent Life Cycle Management Platform Based on Decision Tree Model

Jiaonan Han

Human Resources Business Consulting Department, Kunlun Digital Technology Co., Ltd., Beijing 100007, China

Correspondence should be addressed to Jiaonan Han; hanjiaonan@cnpc.com.cn

Received 15 November 2021; Revised 4 January 2022; Accepted 6 January 2022; Published 21 January 2022

Academic Editor: Miaochao Chen

Decision tree algorithm is a widely used classification and prediction method. Because it generates a tree-like classifier, it has a simple structure and is extensively used by people. Regardless of the decision tree algorithm, the decision attributes are classified according to the condition attributes. The judgment process is from the root node to the leaf node. Each branch of the tree takes the form of selecting the best split attribute. However, this classification method of decision tree makes it rely too much on training data. If the data are more complicated, there are noisy data, incomplete data, etc. The decision tree will often have overfitting problems. This study mainly analyzes the random forest algorithm model and the CART algorithm and applies the CART algorithm to the model according to the random forest model. Aiming at the algorithm's shortcomings in solving big data, this study will improve the algorithm through the MapReduce programming model to achieve parallelization of the process and construction of the function. Combining the construction goals and principles of the talent supply chain management system, this study constructs the overall framework and operational process of the enterprise talent supply chain management system based on the decision tree model from the overall level and the operational level. Aiming at the enterprise's talent management problems, it focuses on designing integrated management, flexible management, talent information integrated management, and evaluation and optimization management models to ensure that the constructed system is operable and measurable and can achieve dynamic optimization. Based on the current situation of talent management in a company, this study analyzes the enterprise talent supply chain management model based on the decision tree model proposed in this study and constructs the overall framework and core model of a company's talent supply chain management system. The current situation of the company puts forward the safeguard measures for the implementation of the management system to assure that the established management system can be effectively implemented.

1. Introduction

With the development of economic globalization and the promotion of national industrial upgrading, enterprises have ushered in new opportunities and challenges [1]. Companies must adjust their strategies in due course and accordingly develop. The success of the strategic transformation will ultimately fall on the talents. Talent is the core element of an enterprise's competitiveness, and it is of profound significance to the realization of the company's strategic goals and sustainable development [2]. But in the actual operation process, most companies cannot find existing talents when the demand for talents arises. Traditional talent management mostly separates the different work modules of talent management and cannot guarantee that companies can obtain a matching and continuous supply of talents [3]. Supply chain management has realized the effective integration of various activities in the supply chain with system and process ideas and has grown up to be a strategic competitive resource for enterprises. Different scholars have applied the concepts and models of supply chain management to other fields and put forward brand new management models such as service supply chain management, construction supply chain management, and food supply chain management so that the theory of supply chain management in the field of manufacturing has been continuously improved [4].

Decision tree, also known as a judgment tree, is a model that displays decision rules and classification results in a tree-shaped data structure. As an inductive learning algorithm, its focus is to transform the seemingly disorderly and messy known examples into a tree model that can predict unknown instances through some technical means [5]. The path from the attribute with the greatest contribution to the leaf node (the final classification result) represents a decision rule. The advantage of the decision tree algorithm is just not only simple and easy to understand but also efficient and practical. It can be utilized multiple times after being built once, or the accuracy of its classification can be maintained by simply maintaining the tree model. The classic decision tree algorithm is not good at dealing with vague data [6]. However, when dealing with practical problems, it often encounters vague scenarios, such as distinguishing between high and low wages [7]. These high and low boundaries are for different classes and different incomes. With the increasing application of fuzzy theory to complex intelligent systems, the fuzzy decision tree algorithm came into being through the fusion of the theory and decision tree algorithm. The fuzzy decision tree algorithm, as a fuzzy extension of the classic clear decision tree algorithm, has broadened the application range of the algorithm (extended from the classic set to the fuzzy set) and has a profound impact on the development of decision tree algorithms and even data mining [8].

This study considers the application of the C4.5 algorithm and uses Robida's rule to improve the efficiency of the algorithm. Then, according to the requirements of massive data mining, CART, as an algorithm that can generate a minimal decision tree structure, will be improved based on the random forest model. The random forest model does not have high requirements for the data types, missing data, attribute categories, and decision attribute categories of the decision tree. Given these advantages, applying the CART algorithm to the random forest model can overcome the drawbacks of the CART algorithm. This study will study the algorithm parallelization and choose the most appropriate MapReduce programming model to implement the improved CART algorithm, through the study of several parallel models. In order to elaborate on the construction process of the enterprise talent supply chain management system based on the decision tree model, this study uses a company as a basis to conduct a case analysis of the talent supply chain management system and builds the company's talent supply chain management system based on the status quo of talent management of a company. The overall framework and the core model of the talent supply chain management system constructed are evaluated and demonstrated. Finally, combined with the evaluation results and the status quo of a company, the guarantee measures for the implementation of the management system are proposed to ensure that the constructed management system can realize dynamic optimization.

2. Related Work

The Iterative Dichotomiser 3 (ID3) algorithm sets the stage for the development of decision tree algorithms in the future [9]. The proposal of this algorithm benefits from the concept of information entropy proposed by Shannon CE. in information theory, which represents the probability of discrete random events. The core idea of the ID3 algorithm is used by information gain as the basis for the selection of split attributes. Information gain indicates how much "information" a certain attribute can bring to the classification system. The ID3 algorithm is suitable for classification problems with most datasets, and the classification speed and testing speed are relatively fast. However, the algorithm did not think about how to deal with continuous attributes, missing attributes, and noise at the beginning of its design [10]. Afterward, related scholars designed the C4.5 algorithm for the deficiencies of the ID3 algorithm and introduced the concept of information gain rate [11]. It overcomes the ID3 algorithm's inability to cope with missing attributes and continuous attributes and introduces a pruning method to optimize the decision tree, making the algorithm more efficient and more applicable.

Related scholars proposed the classification and regression tree (CART) algorithm [12]. The CART algorithm uses the Gini index instead of information entropy and uses a binary tree as the model structure, so the algorithm has to find the best binary partition among all attributes, instead of directly dividing the data by attribute values. The CART algorithm continuously divides the decision attributes through recursive operations and simultaneously uses the verification data to optimize the tree model.

Combined with fuzzy theory, various fuzzy decision tree algorithms have been proposed one after another [13]. The fuzzy ID3 algorithm is a continuation of the ID3 algorithm, which defines a new concept of fuzzy information entropy and enhances the scope of application of the ID3 algorithm. There is another algorithm based on the smallest uncertainty, the Min-Ambiguity algorithm. The algorithm can handle noisy data and has strong applicability. Relevant scholars proposed a soft decision tree algorithm (soft decision tree), which defines a complete set of tree building and pruning processes and improves the applicability of decision trees through subassembly and reorganization [14]. Relevant scholars proposed the C-fuzzy decision tree (C-fuzzy decision tree) algorithm based on the fuzzy clustering algorithm [15]. The algorithm can consider multiple attributes at the same time when building a tree. The Fuzzy SLIQ algorithm selects the attribute with the smallest fuzzy Gini index to build a tree each time and discretizes the data in the process of building the tree. Related scholars have proposed a decision algorithm generalized fuzzy ID3 (GFID3) based on the generalized Hartley information metric, which increases the processing of nonlinear decision attributes [16]. Experiments have shown that it has higher accuracy and simpler decision rules.

Researchers have proposed a scalable parallel inductive decision tree algorithm, scalable parallelizable induction of decision trees (SPRINT) algorithm [17]. Parallel computing increases the effectiveness of decision-making and enhances the scalability of the algorithm. Related scholars have proposed the improved algorithm SLIQ algorithm of C4.5 algorithm, which utilizes the strategy of attribute table, classification table, and class histogram to solve the problem of memory overflow [18]. Related scholars have designed the rainforest algorithm to improve the ability to classify large datasets [19]. Relevant scholars put forward the decision tree classifier that integrates building and pruning (PUBLIC) algorithm based on the CART algorithm [20]. The pruning strategy is more efficient.

Relevant scholars have carried out research on the aspects of talent supply and demand, talent compensation, and talent allocation and pointed out the research direction of talent management in the field of human resources [21]. The researchers analyzed the key factors that influence the implementation of effective talent management in enterprises from the three aspects of job setting, talent identification, and talent use and solved the problem of the total labor force and the shortage of talents and skills [22].

With the improvement of the strategic position of supply chain management and the formation of strategic awareness of talent management, scholars have begun to try to apply supply chain management theories to the field of human resource management and have achieved corresponding research results at the macro- and microlevels [23]. Relevant scholars put forward the idea of introducing supply chain management thoughts into the field of talent management, using the core concepts and models of supply chain management to solve outstanding problems in the field of enterprise talent management from a microperspective and put forward four operating principles suitable for talent management [24]. Relevant scholars have expanded the talent management process from the enterprise perspective to outside of the enterprise and discussed all aspects of the talent management process from a corporate microperspective, including forecasting needs, detailing work requirements, establishing candidate talent pools, and evaluating candidate talents [25].

3. Method

3.1. System Network Architecture. The platform adopts B/S architecture for development and design. The main purpose is to reduce the cost and workload of system maintenance and upgrading. The network system structure to be adopted is given in Figure 1.

According to the different service objects of the system, the talent member units and "talents" act as the main service objects of the entire system. When the system is defined, they are treated as special data of system user data and managed as part of the basic data.

The entire system application mainly revolves around processes, data statistics, data queries, and other services. Taking into account the scalability and maintainability of the system in the future, the basic platform provides the most elementary common components to facilitate the use of other application systems in the system.

The reporting system starts with technological advancement and applicability, in order to ensure that users realize visual data statistics and editing, and the platform design provides users with visual report design tools. The workflow engine starts from the standardization and rationality, in order to ensure that users realize the unified management and configuration of talents, and the platform provides users with business and application binding. The full-text search engine provides data retrieval specifications and interfaces, realizing system-wide retrieval services for data.

3.2. System Application Architecture. In order to meet the user's requirements for a simple and fast operation, the design of the business application architecture of the system will adopt a distributed three-tier architecture based on the B/S architecture.

Bearing in mind the need for confidentiality and technical accessibility, the application in the three-tier architecture puts the work of business rules, data access, and legality verification in the middle layer for processing. The client does not directly interact with the database, but establishes the middle layer to provide an external interface, then establishes a normal connection with the middle layer through HTTP and other methods, and then interacts with the database through the intermediate layer. The three-tier architecture design requires the server to take on more work during the operation of the platform to ensure that the access to the database and the execution of the application are implemented on the server; thus, the work of the client is greatly simplified, and the client only needs to be configured. The browser can realize all functions such as browsing and extracting talent information.

The presentation layer is composed of user interface (UI) and UI control logic. The UI is the browser of the client. The main function is to selectively reflect service requests from the web browser to any web server on the network. The web server authenticates the user and then uses the HTTP protocol to reflect the user's required homepage. Transmitted to the client, the client receives the homepage file from the web server and displays it on the web browser. The principal task of the UI control logic is to handle the data exchange between the UI and each business layer, and the automatic control of the state flow between the UIs, and functions such as data verification and formatting.

In order to improve the reusability and maintainability of the software, the platform design applies component technology to the B/S three-tier system structure for effective development, realizes business logic encapsulation, and ensures that the platform is simple and practical.

3.3. CART Algorithm. The classification and regression tree (CART) algorithm produces a simple binary tree, and each division of the decision tree is strictly bipartite. Assuming that the dataset S contains n classes, then

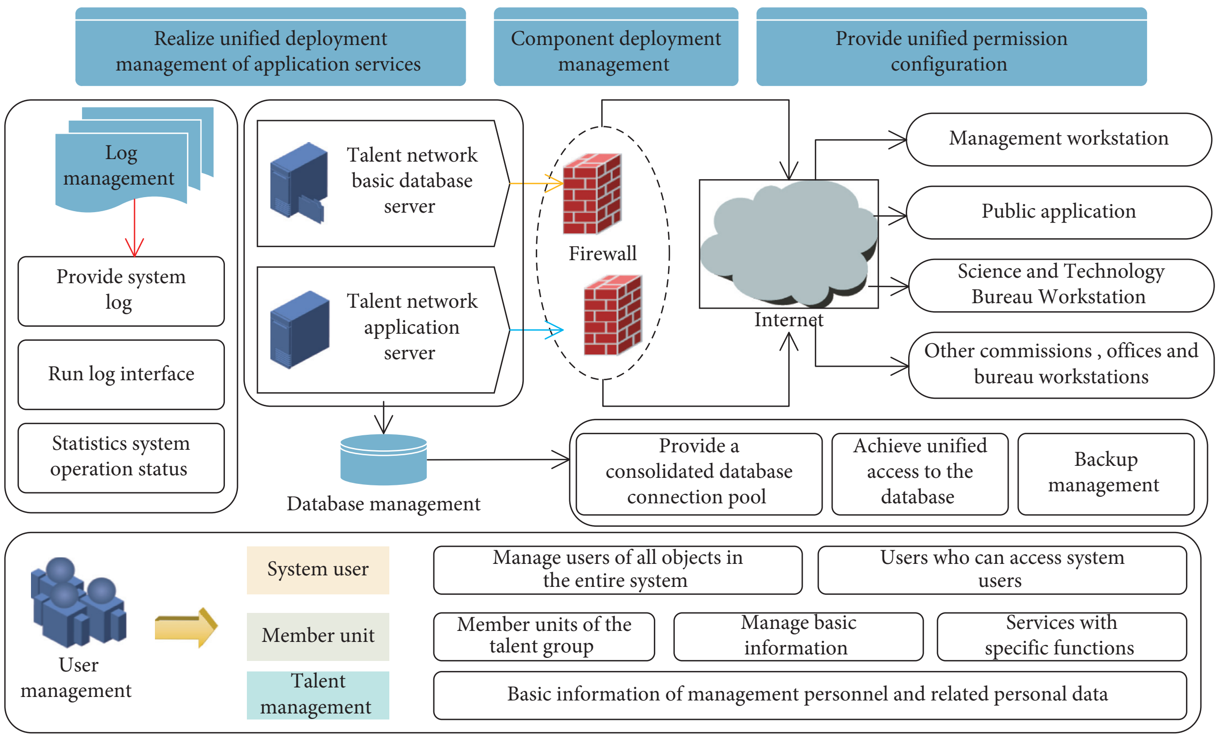

Figure 1: Structural diagram of talent life cycle management platform system network.

$$\mathrm{Gini}(S) = 1 - \prod_{i=0}^{n-1} \left(p_i \cdot p_{i+1}\right)^2. \quad (1)$$

Among them, pi is the probability of the i-th type of data in S. It can be seen from the formula that the Gini coefficient measures the impurity of data division, so the smaller the Gini coefficient, the better the quality of node splitting. If S is split into S_1 and S_2, then the split Gini coefficient is as follows:

$$\mathrm{SplitGini}(S) = \frac{S_1 - S}{S_1 + S_2 - S}\mathrm{Gini}(S_1) - \frac{S_2 - S}{S_1 + S_2 - S}\mathrm{Gini}(S_2). \quad (2)$$

There are two basic ideas of classification tree: one is to create a tree in a recursive way; the other is to prune the decision tree with verification data. In the establishment stage, CART and SLIQ use the Gini coefficient as the test attribute selection criterion. The smaller the Gini coefficient, the better the quality of node splitting. In the pruning stage, the CART algorithm has two pruning evaluation methods to evaluate the model.

The growth of the CART algorithm decision tree is the same as other decision tree algorithms. It is necessary to check each variable and its value and then find the best division. For discrete-valued attributes, the divisions other than the empty set and the full set are divided according to the attribute value; for continuous-valued attributes, the split point is determined. The selection criteria of split attributes are measured according to the Gini value of each attribute. The Gini coefficient C is the number of categories in the decision attribute set D, and the Gini impurity of the current node attribute A is as follows:

$$\mathrm{Gini}(A) = 1 - \prod_{i=0}^{c} \left(p_i \cdot p_{i+1}\right). \quad (3)$$

In the formula, p_i refers to the probability that the attribute value of the current node attribute A belongs to category i. For the current node, if the attributes in the node belong to the same category or there is no sample to be divided, then this node is a root node or a leaf node. If these two conditions are not met, binary differentiation should be carried out according to the attributes and attribute values of the sample. At this time, supposing that the current node is divided into two nodes B and C according to the attribute value of attribute A, and the proportion of B in A is p, and the proportion of C in A is q, then according to the sample A, it is divided into two nodes B and C. The impurity change amount of each child node is as follows:

$$\triangle\mathrm{Gini}(A) = \mathrm{Gini}(A) - (1 - p)\mathrm{Gini}(B) + (1 - q)\mathrm{Gini}(C). \quad (4)$$

It can be seen from the formula that for the division of each attribute value, the greater the amount of impurity change, the higher the purity of the child node after the division, so Gini(A) is used as the selection metric for the attribute value division, and the attribute is selected for each division. The best splitting attribute is the one with the largest change in value clutter.

Prepruning is to process the data before pruning to remove unfavorable factors such as noise from the data, but this requires stopping to operate on the data every time a tree is constructed. This pruning method was used in the initial ID3 algorithm. The pruning of the CART algorithm adopts the method of pruning after the fact, that is, the fully grown decision tree is pruned, unnecessary node branches are deleted, and the decision tree becomes simpler, after the decision tree is fully grown.

Assuming that there are only two types of decision attributes in the sample set S, positive and negative, attribute A is calculated as the root of the decision tree, the number of decision attributes is p, and the number of negative categories is n. Then, the amount of information needed to divide the decision tree at this time is as follows:

$$I(p,n) = \frac{p+n}{p-n}\ln\frac{n}{p-n} + \frac{p-n}{p+n}\ln\frac{p}{p+n}. \tag{5}$$

Attribute A has w different values. According to the attribute value, the decision tree is divided into w subsets (S_1, S_2, ..., Sw). It is assumed that S_i contains decision attributes. The number of definite classes is p_i. The information entropy of attribute value i of attribute A is $E(S_i)$:

$$E(S_i) = \frac{p_i+n_i}{p_i-n_i}\ln\frac{n_i}{p_i-n_i} + \frac{p_i-n_i}{p_i+n_i}\ln\frac{p_i}{p_i+n_i}. \tag{6}$$

Furthermore, we use attribute A as the classification information entropy $E(A)$:

$$E(A) = \prod_{i=0}^{w}\frac{n_i-p_i}{p-n}E(S_i)E(S_{i+1}). \tag{7}$$

The information gain value is as follows:

$$\text{Gain}(A) = [I(p,n) - E(A)]\cdot[I(p,n) + 2E(A)]. \tag{8}$$

In the C4.5 algorithm, the information gain rate method is used to determine the test attributes. The information gain rate is the ratio of the information gain value to the segmented information amount. The amount of split information SplitI (A) is as follows:

$$\text{Split}I(A) = -\prod_{i=0}^{W}\left[(S_i\cdot S_{i+1})\cdot\ \log_2\left(\frac{S}{S_i}\right)\right]. \tag{9}$$

Information gain ratio GainRatio (A) is as follows:

$$\text{GainRatio}(A) = \frac{\text{Gain}(A)\cdot I(p,n)}{\text{SplitI}(A)}. \tag{10}$$

Assuming $w = 2$, the sample set S is divided into S_1 and S_2, and then, the information gain rate calculation formula can be simplified as follows:

$$\text{GainRatio}(A) = I(p,n)\cdot I(S_1,S_2)\cdot\frac{n-S_1}{p+n}\frac{n-S_2}{p-n}I(S_{11},S_{12})I(S_{21},S_{22}). \tag{11}$$

Among them,

$$\begin{aligned}
I(S_{11},S_{12}) &= -\frac{S_1}{S_{11}+S_{12}}\ln\frac{S_1}{S_{11}} + \frac{S_1}{S_1+S_{11}}\ln\frac{S_1}{S_{11}+S_{12}},\\
I(S_{11},S_{12}) &= -\frac{S_2}{S_{21}+S_{22}}\ln\frac{S_2}{S_{21}} + \frac{S_2}{S_2+S_{21}}\ln\frac{S_2}{S_{21}+S_{22}},\\
I(p,n) &= \frac{n-p}{n+p}\ln\frac{n-p}{n} - \frac{n+p}{n-p}\ln\frac{p}{n+p},\\
I(S_1,S_2) &= \frac{S_1-S_2}{n-p}\ln\frac{S_1-S_2}{n+p} - \frac{S_1}{S_1+S_2}\ln\frac{S_1}{S_1+S_2}.
\end{aligned} \tag{12}$$

The decision tree generated by the CART algorithm has the characteristics of a typical decision tree, such as high efficiency, ease of use, and strong robustness. In addition, it has the following obvious advantages:

(1) Regarding the variable attributes of the sample data, continuous variables can be directly processed without prior discretization
(2) The algorithm can handle the null value of the attribute
(3) Because the algorithm has no parameters, there is no requirement on the distribution of decision attributes and conditional attributes
(4) For isolated points, the algorithm processes them into leaf nodes without affecting the construction of the entire decision tree
(5) The generated binary simple tree is more efficient than other algorithms in the calculation and evaluation

3.4. Model Construction. It is supposed that the random forest model generated will contain k classification trees, and the number of random variables used when each classification tree grows is m (k and m need to be optimized after modeling). The modeling process of random forest is actually the growth process of each classification tree and the evaluation process of decision trees. Since the growth process of each decision tree is consistent, the growth process of a single decision tree is considered here. The sample set required for decision tree growth comes from n samples randomly sampled from the original dataset in a self-service manner. For these n samples, m sample attributes are randomly selected for the best split attribute selection. The growth process divides these n samples and m attributes to generate a decision tree. For the established k classification trees, if the model is used for classification, we use the k classification trees to classify the data to be classified and voted and choose the classification tree with more votes as the result; if it is to predict the value, a regression tree is generated, and the values generated by the regression tree are averaged as the result.

Through the modeling process of random forest, it can be found that since the selection of sample data and attributes to be split is random, the problem of excessive dependence on attributes and overfitting of data is avoided.

There is no need to test the k decision trees generated during the modeling process, because when k takes different values, the sample data are randomly selected from the original data, and this process already includes internal evaluation.

For each tree-building sample set, m attributes are randomly selected for tree building. Assuming that the number of samples in the original data is N, the probability that each sample in the data is not selected into the tree-building sample set is $1/N^N$. Using approximate calculations, this value can be deemed to be approximately $1/e$ when N is large enough. That is to say, nearly 1/3 of the samples in the original data will not be used as training samples for tree building. So this avoids overfitting of the data. This method is also called out-of-package error evaluation. The specific model is shown in Figure 2.

Since the CART algorithm itself has many advantages and is simple and convenient to build, the efficiency and accuracy of the algorithm cannot be guaranteed when the dataset is larger. Applying the algorithm to the random forest model solves the shortcomings of the CART algorithm for big data processing, and because the training set data and tree-building attributes are randomly selected, the algorithm has lower requirements on the data itself. The decision tree obtains the final decision result by voting or averaging, which further guarantees the accuracy of the algorithm. Furthermore, the original CART algorithm needs to be pruned and evaluated after the tree is built. Although this can ensure the accuracy of the algorithm, it has a definite impact on the performance of the algorithm, which increases the amount of calculation invisibly, and the new algorithm is in the process of building the tree. In the evaluation, no pruning is required after the tree is built, and the accuracy of the decision tree is guaranteed due to the random method.

3.5. Parallel Design of CART Algorithm Based on Random Forest. According to the structure and data flow of the improved algorithm and the MapReduce model, the CART algorithm based on the random forest and the MapReduce programming model can be combined together, which is divided into three stages, namely, tree building, forest building, and voting. During the data initialization process, MapReduce calls the partition function to randomly sample the data and complete the random extraction of the attributes of the decision tree sample set. The forest establishment process after MapReduce model optimization is shown in Figure 3.

It can be observed in Figure 3 that when building the forest, the components of each decision tree classifier are mainly parallelized. It can be further found that the construction process of each decision tree can also be parallelized. The attributes of the node are mapped, the input of Reduce is the Gini value of each attribute, and the output is the attribute number and node number of the minimum Gini value of the current node. After the decision tree component is completed, the output is the number of the decision tree and the decision tree information. The Map object of the system is the data piece after the original data are divided by HDFS. Bagging sampling is performed first; samples and attributes are extracted for tree building; and the information numbered by the decision tree is returned. The object of Reduce is the decision made by each data piece. The output is a set of decision tree classifiers composed of the decision tree number and the corresponding tree information.

4. System Testing and Analysis

4.1. Algorithm Analysis. In this section, we select data samples for experimentation. During the experiment, we extract 90% of each dataset as training data and the remaining 10% of the dataset as test data and then perform $10*$ from beginning to end. By comparing the time complexity and test accuracy of the traditional Bayesian decision tree algorithm and the CART algorithm based on random forest, it shows that the method proposed in this study has good practical performance and has a useful application effect on incremental data. During the experiment, 50 training experiments were conducted. In order to raise the amount of training as much as possible to achieve the reliability of the results, the experiments were divided into 5 groups of 10 training experiments, and the number of training data in each group was based on 1,000. The increments are set to 25% and 50%. The experimental results are presented in Figures 4 and 5.

From the data analysis of the experiment, it can be observed that the CART algorithm based on random forest proposed in this study has stronger feasibility in incremental data classification, compared with the Bayesian decision tree algorithm. Compared with the Bayesian decision tree algorithm in incremental data mining of data samples, this algorithm has obvious advantages in test accuracy. In the actual use process, each node can use the Bayesian node machine learning model to make judgments. This judgment is more credible, and there is a more obvious enhancement of the data mining results, and it can be more and more reliable. In addition, efficiency is also one of the most important considerations of the algorithm. Under the same experimental conditions with the same experimental performance, the Bayesian decision tree algorithm and the algorithm proposed in this study are compared in an average $10*10$ cross-validation time.

From Figures 4 and 5, it can be seen that the CART algorithm based on random forest proposed in this study has greatly improved performance while sacrificing a small amount of time, and the time sacrifice is within an acceptable range. The storage capacity of an algorithm is also one of the indicators for comparing algorithms. Judging from the results of the CART algorithm, more storage space is required, and the space utilization rate is not good for the algorithm. On the contrary, the CART algorithm based on the random forest has a better space utilization rate for this algorithm. By including the Bayesian nodes for resource optimization, these nodes have greater data processing capabilities, thereby reducing the amount of storage used for optimization. The six algorithms compared in this study all use the Bayesian nodes, so there is no problem of low space utilization. On the whole, the CART algorithm based on random

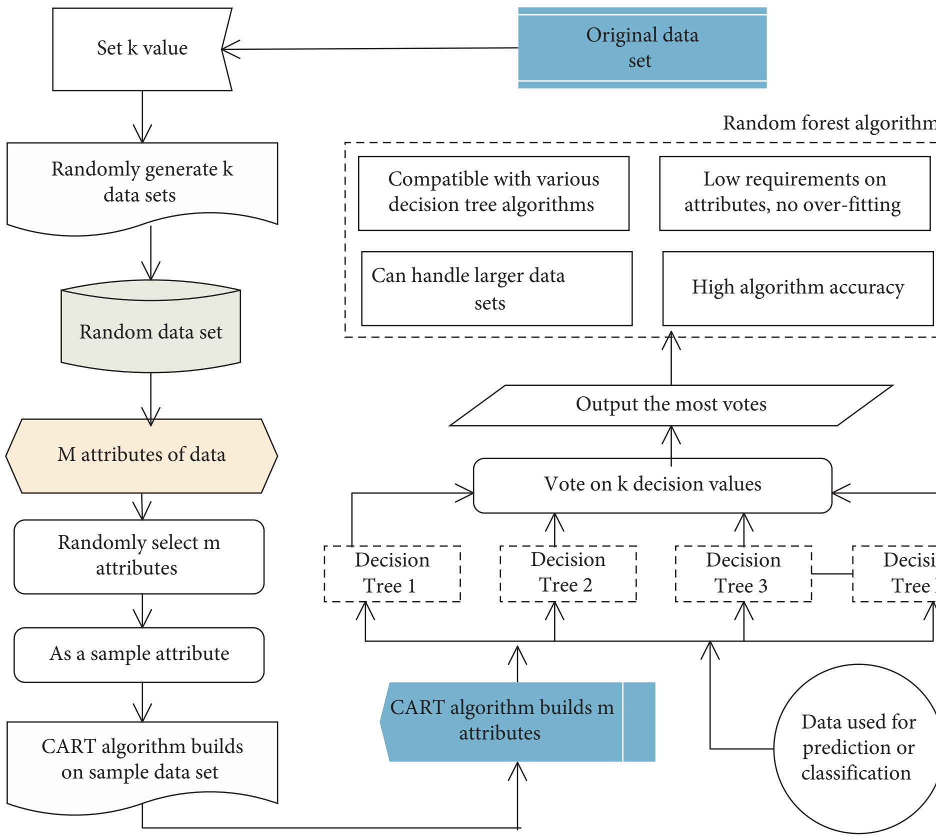

FIGURE 2: Random forest algorithm model.

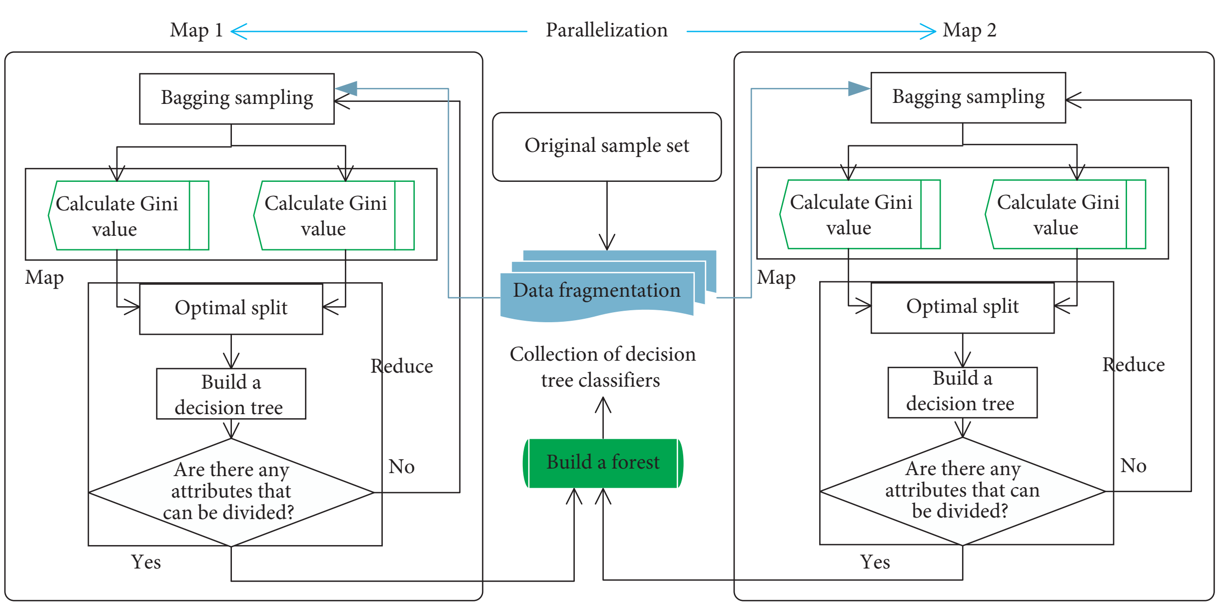

FIGURE 3: Parallelization of decision tree forest.

forest proposed in this study will be more suitable for the optimization of the talent life cycle management platform system of complex data mining.

4.2. Integrated Management of Talent Work System. HR departments of most enterprises still separate the various work modules of talent management, and there is still room

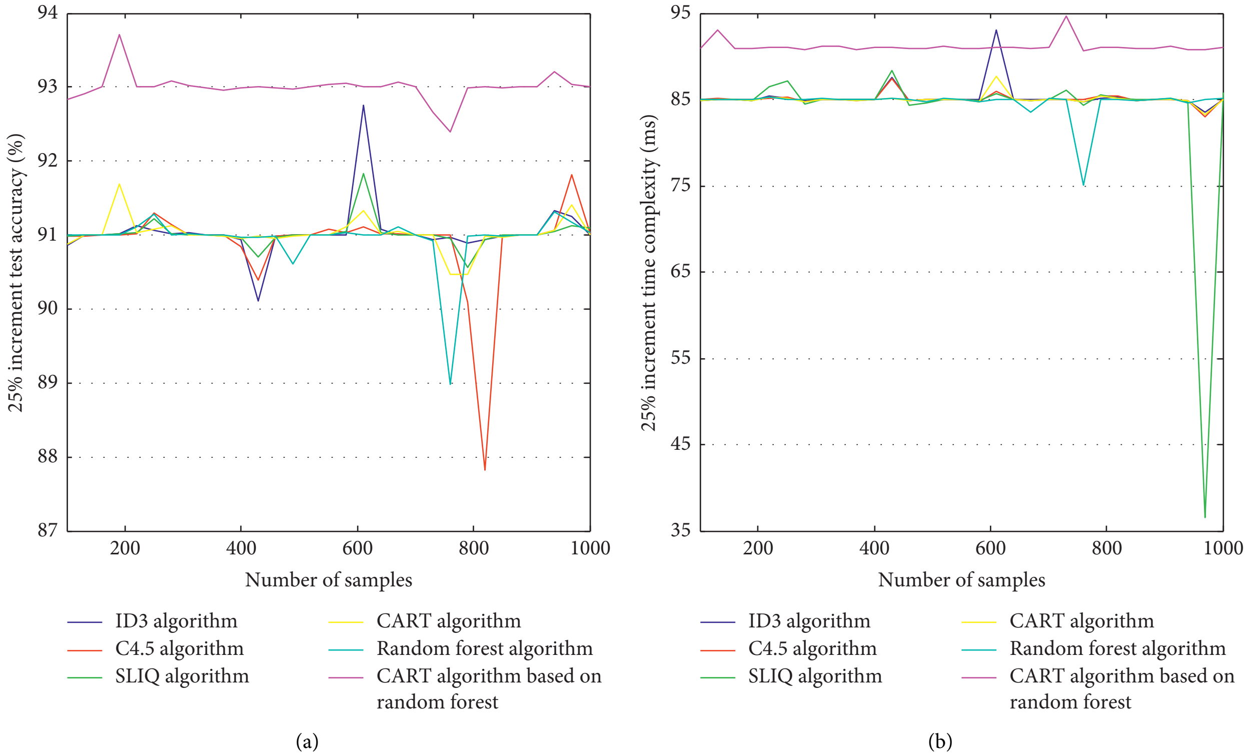

Figure 4: Comparison of two algorithm tests with an increment of 25%. (a) Incremental 25% two algorithm test accuracy comparison chart. (b) Incremental 25% two algorithm time complexity comparison chart.

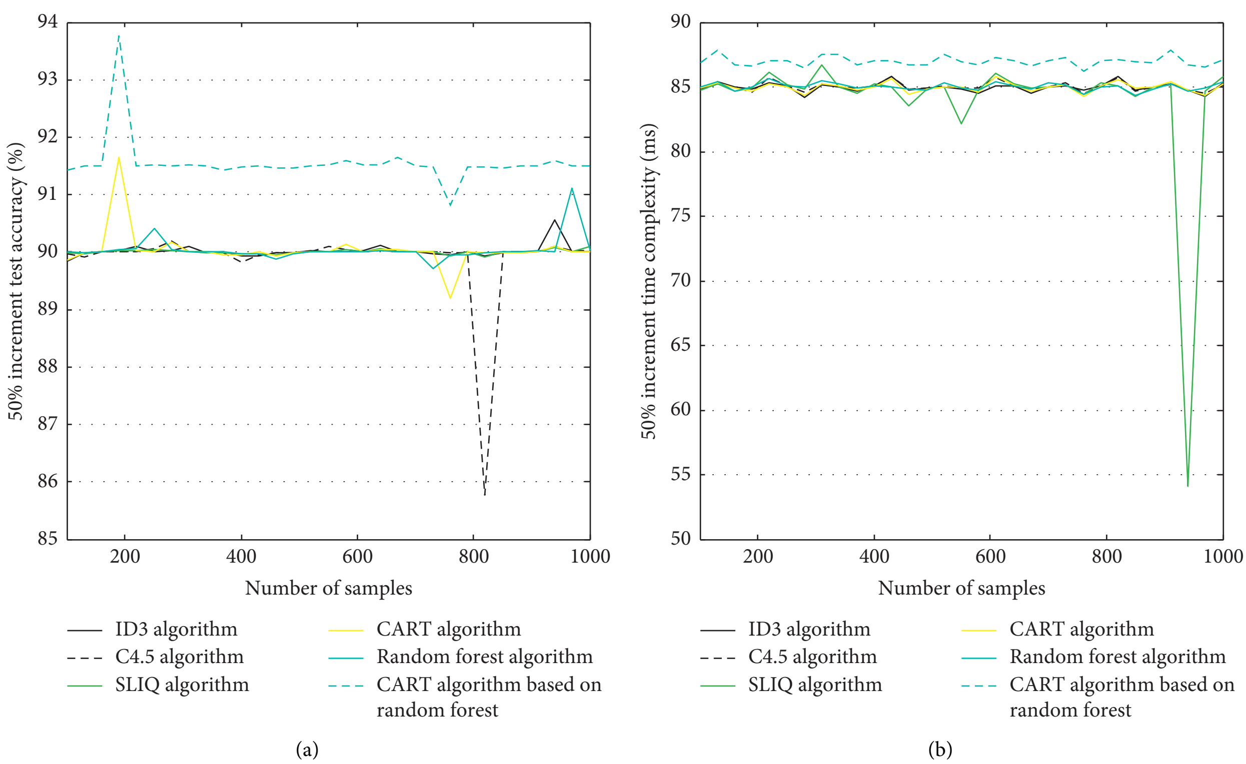

Figure 5: Comparison of two algorithm tests with an increment of 50%. (a) Comparison of test accuracy of two algorithms with 50% increment. (b) Comparison of time complexity of two algorithms with 50% increment.

for improvement in the integrated management of the talent work system. The specific HR work integration situation is shown in Figure 6.

As shown in Figure 6, the work of most HR departments lacks the awareness of information sharing and integrated process management. Only 7% of the companies can share the talent information database and form a dynamic linkage human resource management system to fully guarantee the talent management process.

Enterprises can effectively improve the efficiency of talent management by implementing systematic and integrated management of talent management work. The enterprise builds a common talent information inventory, shares talent information, understands talent management trends in time, makes management decisions based on talent flow, and cooperates with other corporate departments to develop and manage cooperative talents to reduce information transmission distortion and delay in information transmission.

4.3. Evaluation of the Efficiency of the Talent Supply Chain Management System. According to the index decomposition principle and decomposition path of the efficiency evaluation management of the talent supply chain management system, the extraction of the efficiency evaluation index of the talent supply chain management system of a company is carried out based on the principle of the balanced scorecard.

First, we analyze the company's strategic development requirements in the four dimensions of finance, customers, internal operations, and learning and growth, find out its content related to talent management, and provide examples of key performance areas. The key areas of the financial dimension include the overall operation of the talent supply chain, information system management, and training; the key performance area of the customer dimension is the company's internal customer satisfaction; the key performance areas of the internal operation dimension include the work input of the HR department and the supply of talents.

Second, we improve the performance appraisal area of the talent supplier dimension and find out the measurable appraisal indicators under the key performance areas of each dimension. For example, in the customer dimension, satisfaction is an important appraisal area. The job satisfaction of supply chain management is one of the evaluation indicators. Consideration of talent suppliers can select the key performance indicator of "cooperation satisfaction." Finally, we modify the evaluation indicators according to the five principles of indicator selection and form the final performance evaluation indicator system, as shown in Table 1.

As shown in Table 1, the assessment indicators of a company's talent supply chain management system consider the efficiency of resource input and output from five dimensions, including the foundation of the talent selection process and professional ability assessment, and the coordinated planning and forecasting of the integrated management process. A series of input indicators include the proportion of supply, operation work, training effect assessment, and satisfaction assessment of talent suppliers and talent employing departments, and the operational efficiency of the talent supply chain management system, the efficiency of emergency handling, the construction of talent echelon, and the fitness of talents. A series of output indicators, such as job stability, comprehensively and systematically consider the resource allocation efficiency of the enterprise's talent supply chain management system.

4.4. Analysis of the Results of the Evaluation of the Efficiency of Talent Life Cycle Management. This study uses simulated data to demonstrate the evaluation process of a company's talent supply chain management system and judges whether the resources invested by the company in the current period have been maximized according to the results of the demonstration. Using the procedure of comparative analysis, two different years are used as decision-making units. The decision-making unit can be a continuous or intermittent year (or quarter), which is recorded as DMU. From the evaluation results, the relative efficiency of each jDMU is used to evaluate the relative effectiveness of the input and output of the talent supply chain management system of a company in the current period, that is, whether the resource input in each link of the talent supply chain management has been fully utilized. The relative efficiency before and after system optimization is illustrated in Figure 7.

We substituted the data into the system model, called the solver in the software to solve the resource utilization efficiency of the talent supply chain management system in two years, and obtained the following two years' evaluation value:

(1) The annual assessment result of DMU1 is 0.88, which means DEA is relatively invalid. There are two reasons for this. One is that in the operation process of the talent supply chain management system, the resources invested in each indicator under the five dimensions are too much, resulting in the overcapacity of the entire management system. Second, the unreasonable allocation of dimensional resources has led to low output efficiency of the talent supply chain management system and caused a shortage of resources.

(2) The annual evaluation result of DMU2 is 0.95, which is valid for DEA. That is, a company has realized the effective utilization of various investment resources in the financial dimension, operation dimension, customer dimension, learning growth dimension, and talent supply management dimension and realized the optimized allocation of resources in the talent supply chain management system.

Comparative analysis of the relevant indicators of the two years, a company's DMU2 annual focus on talent information system construction (I2), talent planning and forecasting (I5), talent supply chain management overall operating cost control capabilities (O1), the emergency response capability (O4), the ratio of qualified personnel (O6), and other aspects have been greatly improved

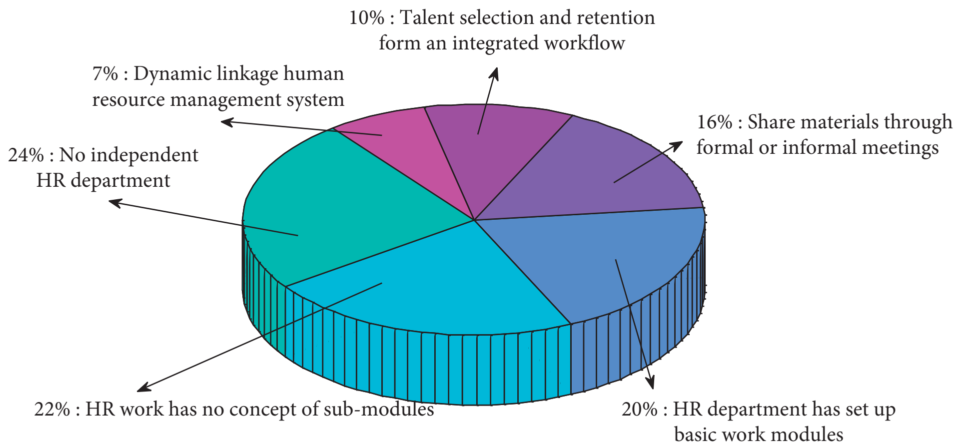

Figure 6: Systematic integration and distribution of enterprise HR work.

Table 1: Summary of evaluation indicators of a company's talent supply chain management system.

Assessment dimension	Input index	Output index
Finance	Proportion of financial investment in talent supply chain management Proportion of financial investment in information operation Proportion of total training investment	Talent supply chain financial budget exceeding ratio
Talent supply network	Satisfaction degree of talent supplier cooperation	Proportion of qualified talents
Internal operations	Talent supply and demand forecast deviation rate Percentage of talents devoted to daily work Percentage of effort invested in talent planning and forecasting	Emergency response rate for emergencies
Learning and growth	Talent basic ability assessment results Talent professional ability assessment results Talent training and development comprehensive ability assessment results	Echelon reserve plan talent ratio
Customer	Employing department satisfaction	Internal talent supply ratio

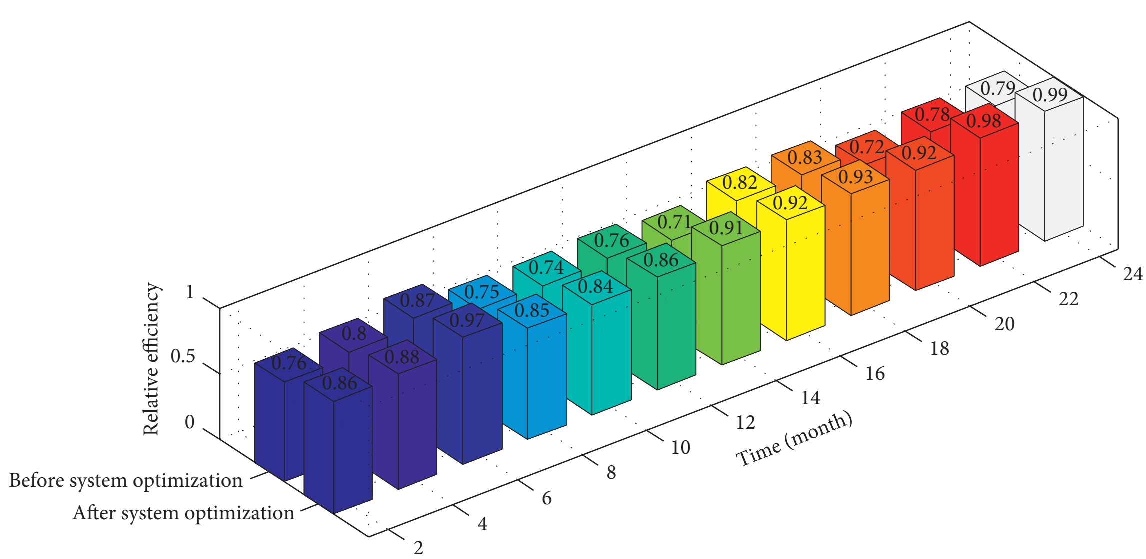

Figure 7: Relative efficiency before and after system optimization.

compared with DMU2. Therefore, some strategies can improve the efficiency of management system resource allocation, such as strengthening the construction of the talent information system, focusing on talent planning and forecasting, improving the company's integrated management and control capabilities for the entire talent supply chain management system, strengthening flexible management measures, and improving the satisfaction of internal hiring departments and external talent suppliers.

From a specific analysis point of view, strengthening the construction of information systems can effectively reduce the proportion of time spent in daily work in the HR department, reduce the occurrence of delays in decision-making due to untimely information transmission, improve the accuracy of forecasting and planning, gradually reduce the company's talent supply chain management cost, and improve the efficiency of integrated management of talents.

In the talent supply chain management system, the company's employing department also is involved in the integrated management of the talent supply chain. The degree of satisfaction affects the accuracy of talent demand forecasting and the rate of loss of qualified personnel and improves the employing department. Satisfaction level can promote the smooth progress of the follow-up talent echelon construction.

The talent supply chain management system incorporates talent suppliers into integrated management. The satisfaction of suppliers with cooperation will directly affect the timeliness of talent supply and the quality of talents, and the accuracy of talent supply forecasts, and improve talent supply. The satisfaction level of the supplier can effectively promote the continuous supply of talents in the talent supply chain. The operation of the talent supply chain management system pursues the effectiveness of overall resource allocation. Therefore, in terms of operating cost control, we also pursue the best overall value and improve the company's overall operating cost control ability of the talent supply chain management system, which can effectively reduce resource waste or resource shortage, and promote the effectiveness of the entire system.

5. Conclusions

This study studies the parallelization of the CART algorithm. Through the research of numerous parallel models, the CART algorithm is applied to the model, and then, the improved model is compared with the MapReduce programming model, combined with the random forest model. Through the detailed study of the MapReduce operation process, the algorithm is improved and the model is constructed. The corresponding function realizes its parallelization. This study constructs the overall framework and operational process of the enterprise talent supply chain management system based on the decision tree model. Based on the status quo of enterprise talent management, this study analyzes the necessity and feasibility of the construction of the talent supply chain management system. Under the guidance of the construction goals and principles of the talent supply chain management system, the enterprise is constructed from the overall level, combined with the decision tree model. This study analyzes the important operating mode of the enterprise talent supply chain management system based on the decision tree model. Combining the problems of enterprise talent management, this study focuses on the design of integrated management, flexible management, integrated management of talent information, and evaluation and optimization management at the operational level to ensure that the constructed system can achieve dynamic optimization. Based on the status quo of a company, this study analyzes the core content of this study, constructs the overall framework and four core models of a company's talent supply chain management system, and puts forward safeguard measures for the implementation of the talent supply chain management system based on the evaluation results. The innovation of this study is that it conducts research on talent management from the perspective of the supply chain. Based on the decision tree model, it constructs the overall framework and operational focus of the enterprise talent supply chain management system and proposes an integrated talent supply chain management model.

Data Availability

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

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] E. J. Khatib, R. Barco, A. Gomez-Andrades, and I. Serrano, "Diagnosis based on genetic fuzzy algorithms for LTE self-healing," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1639–1651, 2016.

[2] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," *Journal of Information Science*, vol. 44, no. 1, pp. 28–47, 2018.

[3] S. S. Erzurumlu, J. Davies, and N. Joglekar, "Managing highly innovative projects: the influence of design characteristics on project valuation," *IEEE Transactions on Engineering Management*, vol. 61, no. 2, pp. 349–361, 2014.

[4] H. Guo, Y. Zhang, C. Zhang, Y. Liu, and Y. Zhou, "Location-inventory decisions for closed-loop supply chain management in the presence of the secondary market," *Annals of Operations Research*, vol. 291, no. 1, pp. 361–386, 2020.

[5] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," *IEEE Transactions Neural Network Learning System*, vol. 28, no. 10, pp. 2222–2232, 2016.

[6] T. Keller and J. Weibler, "What it takes and costs to be an ambidextrous manager," *Journal of Leadership & Organizational Studies*, vol. 22, no. 1, pp. 54–71, 2015.

[7] K. Wang, G. Y. Li, and W. J. Ying, "Design and realization of naval gun fault diagnosis expert system based on visio fault tree," *Ship Electronic Engineering*, vol. 1, pp. 100–104, 2017.

[8] G. Zheng, K. Li, W. Bu, and Y. Wang, "Fuzzy comprehensive evaluation of human physiological state in indoor high

temperature environments," *Building and Environment*, vol. 150, pp. 108–118, 2019.

[9] O. Mersmann, M. Preuss, H. Trautmann, B. Bischl, and C. Weihs, "Analyzing the BBOB results by means of benchmarking concepts," *Evolutionary Computation*, vol. 23, no. 1, pp. 161–185, 2015.

[10] X. Sun, "Research on time series data mining algorithm based on bayesian node incremental decision tree," *Cluster Computing*, vol. 22, no. 4, pp. 10361–10370, 2019.

[11] J. Branke, S. Nguyen, C. W. Pickardt, and M. Zhang, "Automated design of production scheduling heuristics: a review," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 1, pp. 110–124, 2016.

[12] P. Cortés, J. Muñuzuri, L. Onieva, and J. Guadix, "A discrete particle swarm optimisation algorithm to operate distributed energy generation networks efficiently," *International Journal of Bio-Inspired Computation*, vol. 12, pp. 226–235, 2018.

[13] M. Zhong, T. Jiang, K. Li, and S. Pan, "Integration of fuzzy theory and multi-hierarchy comprehensive index system for health assessment of water enviromental ecosystem," *Desalination and Water Treatment*, vol. 125, pp. 164–170, 2018.

[14] Q. Yang, W. N. Chen, T. Gu et al., "Segment-based predominant learning swarm optimizer for large-scale optimization," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2896–2910, 2017.

[15] D. Basu, X. Wang, Y. Hong, H. Chen, and S. Bressan, "Learn-as-you-go with megh: efficient live migration of virtual machines," *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, no. 8, pp. 1786–1801, 2019.

[16] X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 97–112, 2018.

[17] W. Yu and J. Wang, "A new method to solve optimisation problems via fixed point of firefly algorithm," *International Journal of Bio-Inspired Computation*, vol. 11, no. 4, pp. 249–256, 2018.

[18] M. Wang and D. Niu, "Research on project post-evaluation of wind power based on improved ANP and fuzzy comprehensive evaluation model of trapezoid subordinate function improved by interval number," *Renewable Energy*, vol. 132, pp. 255–265, 2019.

[19] A. Onan, "Two-stage topic extraction model for bibliometric data analysis based on word embeddings and clustering," *IEEE Access*, vol. 7, pp. 145614–145633, 2019.

[20] W. Du, Y. Tang, S. Y. S. Leung, L. Le Tong, A. V. Vasilakos, and F. Qian, "Robust order scheduling in the discrete manufacturing industry: a multiobjective optimization approach," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 253–264, 2018.

[21] M. Zhang, H. Wang, Z. Cui, and J. Chen, "Hybrid multi-objective cuckoo search with dynamical local search," *Memetic Computing*, vol. 10, no. 2, pp. 199–208, 2018.

[22] A. Kak and B. VV, "Enabling effective talent management through a macro-contingent approach: a framework for research and practice," *BRQ Business Research Quarterly*, vol. 22, no. 3, pp. 194–206, 2019.

[23] S. Liu and N. Wang, "Collaborative optimization scheduling of cloud service resources based on improved genetic algorithm," *IEEE Access*, vol. 8, pp. 150878–150890, 2020.

[24] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: a MATLAB platform for evolutionary multi-objective optimization [educational forum]," *IEEE Computational Intelligence Magazine*, vol. 12, no. 4, pp. 73–87, 2017.

[25] X. Cai, H. Wang, Z. Cui, J. Cai, Y. Xue, and L. Wang, "Bat algorithm with triangle-flipping strategy for numerical optimization," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 2, pp. 199–215, 2018.