

## Research Article

# The Evaluation of Online Education Course Performance Using Decision Tree Mining Algorithm

Yongxian Yang 

*Department of Education Science, Ningxia Normal University, Guyuan 756000, Ningxia, China*

Correspondence should be addressed to Yongxian Yang; 82001023@nxnu.edu.cn

Received 22 February 2021; Revised 22 March 2021; Accepted 29 March 2021; Published 9 April 2021

Academic Editor: Wei Wang

Copyright © 2021 Yongxian Yang. 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 development of “Internet + Education”, online learning has become a hot topic of concern. Decision tree is an important technique for solving classification problems from a set of random and unordered data sets. Decision tree is not only an effective method to generate classifier from data set, but also an active research field in data mining technology. The decision tree mining algorithm can classify the data, grasp the teaching process of the teacher, and analyze the overall performance of the students, so as to realize the dynamic management of the educational administration and help the educational administration personnel to make the right decision, with more reasonable allocation of resources. This paper evaluates students’ academic performance based on the learning behavior data of online learning, so as to intervene in students’ learning in advance, which is the key problem that needs to be solved at present. Taking students’ learning attitude, completion of homework, and attendance as factors, the paper uses decision tree technology to analyze the factors affecting students’ performance, and evaluates students’ performance. Firstly, this paper collects the high-dimensional behavioral characteristic data of students’ online learning and conducts correlation analysis after preprocessing the behavioral characteristic data. Then, the decision tree C4.5 algorithm is used to construct a performance evaluation model. Students’ performance is evaluated by the model, and the evaluation accuracy is about 88% compared with actual performance. Finally, through the model analysis, it is concluded that the video task point completion is the most influential in students’ achievement, followed by chapter test completion and chapter test average score, and the course interaction amount and homework average score are the least influential in students’ achievement, which has a practical reference value for effectively serving online learning and teachers’ teaching.

## 1. Introduction

By January 2021, the number of online education users in China reached 452 million, accounting for 49.8% of the total number of Internet users, an increase of nearly 200 million compared with January 2020. The 2020 Government Work Report points out that the government will strengthen and improve the construction of the “Internet + Education” model and promote the equitable development and quality improvement of education, which reflects the government’s determination to strongly support online education. Obviously, online education has become an important part of the development of education. With the expansion of the users of online education platform, a large amount of learning behavior data has been generated in the background. How to apply these data to the performance evaluation has become a

hot topic of current research. The key purpose of this study is to adopt reasonable algorithm, establish scientific performance evaluation model, and mine learning behavior data. Based on the analysis results of the model, the key factors affecting students’ online learning are obtained, which provide a theoretical basis for the feedback and guidance of students’ online learning activities, the intervention of teachers, and the implementation of educational evaluation by educational administrators. It is of great significance to guide students to learn online more effectively.

Achievement is an important standard to measure students’ performance and an important basis in the process of teaching quality evaluation. The process of teaching is also a process of data accumulation. By using relevant technologies to analyze and mine the data, we can fully analyze the rules contained in the performance data, carry out

quantitative and qualitative analysis of the results, and accurately analyze various aspects of the results according to different students' learning conditions. The use of data visualization technology can not only show the results intuitively but also more clearly show the influence and relationship between various disciplines. According to the relationship between these factors, the educational administration department of the school can modify the teaching plan pertinently, improve the teaching quality of the school, and ensure the results and effects of teaching.

The biggest advantage of data mining is to be able to make a comprehensive analysis of a very large amount of data and to extract rules that can arouse the interest of relevant personnel and some unknown and potential knowledge that is very beneficial to decision-making. Using the tree structure to express the results of data analysis, users can better understand the relationship between various factors. With the development of visualization technology, artificial intelligence technology, and machine learning technology, many data mining methods have been widely used in real life. Achievement, as important data of educational administration in higher vocational colleges, can reflect a student's learning situation and can also evaluate the teachers' teaching quality. The decision tree method enables students to show the connection between different professional courses and practices through various professional achievements and also helps teachers improve teaching methods. Therefore, in teaching work, more targeted teaching methods can be used to improve the learning efficiency of students and the teaching level of teachers.

Section 2 reviews and discusses the relevant literature in the field. Section 3 summarizes the decision number algorithm and describes the basic form and characteristics of the decision tree mining algorithm. Section 4 constructs the basic model of the decision tree mining algorithm, completes the online performance evaluation based on the decision tree mining algorithm, and carries out the corresponding data analysis. Section 5 summarizes the whole paper.

## 2. Related Works

Decision tree is one of the commonly used algorithms in the field of data mining research. It has the advantages of high precision, high efficiency, fast calculation speed, and ability to process multiple types of data [1–3]. At present, the research on online education using decision tree mining algorithm mainly focuses on the evaluation of learning effect and the research of learning influencing factors. In the study of using decision tree analysis to predict student performance in colleges and universities, Hamoud A et al. used the decision tree C4.5 algorithm to build a classification model for student English graduation test performance evaluation, with an evaluation accuracy rate of 81.62% [4]. Di and Xu evaluated the improved decision tree algorithm and its application, and the highest evaluation accuracy rate was 81% [5]. According to the research results of the above researchers, although the evaluation accuracy rate is about 81%, there is still a gap with the ideal data, indicating that [6]

still has room for improvement. In terms of studying influencing factors, Weinberg and Last adopted the interpretable decision tree induction method under the big data parallel framework and found that the factor that has the greatest impact on student performance is the basic knowledge mastery table [7]. Wu adopts the MOOC learning behavior analysis and teaching intelligent decision support method based on the improved decision tree C4.5 algorithm, uses the decision tree mining algorithm to evaluate student performance, and finds that the factors that affect the total unqualified score are homework scores, discussion, communication, etc. [8]. Yaacob et al. propose the supervised data mining approach for predicting student performance, which concludes complex decision tree classifier in their experiments [9]. Fiarni et al. designed an academic decision support system for choosing information systems submajors programs using decision tree algorithm [10]. The influencing factors of the research can affect the academic performance of students, but some key factors that ignore the current online learning affect learning, such as the completion of the video point to complete the task and the length and quantity of the video to watch the partial test completion. Therefore, it is necessary to further data mining online learning. Behavior, find a more comprehensive achievement factor, the influence of online education, to meet the needs of the current rapid development of online education.

In summary, the algorithms used for performance evaluation have problems such as insufficient use of learning behavior data, low accuracy of evaluation results, and inaccurate mining of influencing factors. Therefore, the main purpose of this article is to use the decision tree mining algorithm to better solve the above problems.

## 3. Overview of Decision Tree Mining Algorithms

*3.1. General Form of Decision Tree Mining Algorithm.* The algorithm of decision tree learning is usually a process of recursively selecting the optimal feature and segmenting the training data according to the feature, so as to have the best classification of each subdata set. This process corresponds to the division of feature space and the construction of decision tree. To begin, build the root node and place all the training data at the root node. Select an optimal feature, and divide the training data set into subsets according to this feature, so that each subset has a best classification under the current conditions. If these subsets can be classified basically correctly, then build leaf nodes and divide these subsets into corresponding leaf nodes. If there are subsets that cannot be classified basically correctly, then select new optimal features for these subsets, continue to segment them, and build corresponding nodes. This process continues until all subsets of training data are classified roughly correctly, or there are no appropriate features. Finally, each subset is assigned to a leaf node, so you have a specific class. This generates a decision tree [11–13].

It can be seen from the above process that the generation of decision tree is a recursive process. In the basic algorithm

of the decision tree, there are three scenarios that result in recursive returns: the samples of the current node all belong to the same category and need not be divided. The current attribute set is empty, or all samples have the same value on all attributes and cannot be partitioned. The sample set contained in the current node is empty and cannot be partitioned.

In the second case, we mark the current node as a leaf node and set its category as the category where the node contains the most samples. In the third case, the current node is also marked as a leaf node, but its category is set to the category whose parent node contains the most samples. In the second case, the posterior distribution of the current node is used, while in the third case, the sample distribution of the parent node is taken as the prior distribution of the current node.

The decision tree generated by the above method may have a good classification ability for training data but may not have a good classification ability for unknown test data; that is, overfitting phenomenon may occur. We need to prune the generated tree from the bottom up to make the tree simpler and thus more generalizable. Specifically, it removes the oversegmented leaves, regresses them back to the parent or even higher nodes, and then changes the parent or higher nodes to new leaves. If the number of features is large, the features can also be selected at the beginning of decision tree learning, leaving only those features with sufficient classification ability for training data.

It can be seen that the decision tree learning algorithm includes feature selection, decision tree generation, and decision tree pruning process. Since the decision tree represents a conditional probability distribution, the different depth of the decision tree corresponds to the probability model of different complexity. The generation of the decision tree corresponds to the local selection of the model, and the pruning of the decision tree corresponds to the global selection of the model. The generation of decision tree considers only local optimum, and the pruning of decision tree considers global optimum [14–16].

Decision tree model has a tree structure. In classification problem, it represents the process of classifying instances based on features. It can be considered as a set of if-then rules, or as a conditional probability distribution defined in feature space and class space. The classification tree has the advantages of good readability and fast classification speed. When training the classification tree, the training data are used to establish the classification tree model according to the principle of minimizing the loss function. When forecasting, the new data are classified by the classification tree model. Decision tree learning usually includes three steps: feature selection, decision tree generation, and decision tree pruning.

A decision tree can be viewed as a collection of if-then rules: a rule is constructed from each path from the root of the decision tree to the leaf, where the characteristics of the internal nodes correspond to the conditions of the rule, and the classes of the leaf nodes correspond to the conclusion of the rule. The path to a decision tree or its

corresponding set of if-then rules has an important property: they are mutually exclusive and complete. That is, each instance is covered by one path or one rule. Here, the so-called coverage refers to the conditions that the features of the instance are consistent with those on the path or that the instance satisfies the rules.

- (1) The content of decision tree mining algorithm: decision tree mining algorithm has a lot of different methods after a long time of development, usually used in ID3, C4.5, CART, etc. This method will form a decision tree after data mining, through the use of the decision tree, to help users to make the right decision. In the process of tracking the generation of the decision tree mining algorithm, it needs to go through the process of decision tree pruning and decision tree generation. The algorithm can process the unknown data and evaluate the future development direction of an item.
- (2) Principle of decision tree mining algorithm: decision tree mining algorithm will use certain rules to analyze and sort out data during data mining, which is widely used in actual life and production. After the data is preprocessed, the induction method can be used to generate the decision tree, the related classification work can be carried out, and then the further data analysis can be conducted. This method can be used in project risk assessment to analyze the feasibility of the project. The algorithm also uses machine learning techniques to analyze the relationships between different values to evaluate and extrapolate the future. A decision tree is composed of nodes and forks. The nodes represent various objects, the forks represent various attributes, and the paths passed by each leaf stage represent the attribute values of objects. Forks are generally expressed in the plural form, allowing for relatively independent data processing. Therefore, decision trees can be used for both data evaluation and data analysis.

Decision tree mining algorithm can conduct in-depth analysis of all the features of the sample, so as to find out the features with decisive significance, then display the analysis results, determine the most significant feature as the root node of the entire decision tree, and then proceed to analyze the significance of other features to build an inverted tree. The decision tree mining algorithm can also be specially designed for nonnumerical data, especially for students' grades. Since there are not only numerical data but also a lot of nonnumerical data, the nonnumerical data can be processed according to the characteristics of students' grades. The processing model of the decision tree mining algorithm is shown in Figure 1.

- (3) Advantages of decision tree mining algorithm: (1) It is conducive to users' understanding of rules. Normally, academic administrators conduct data

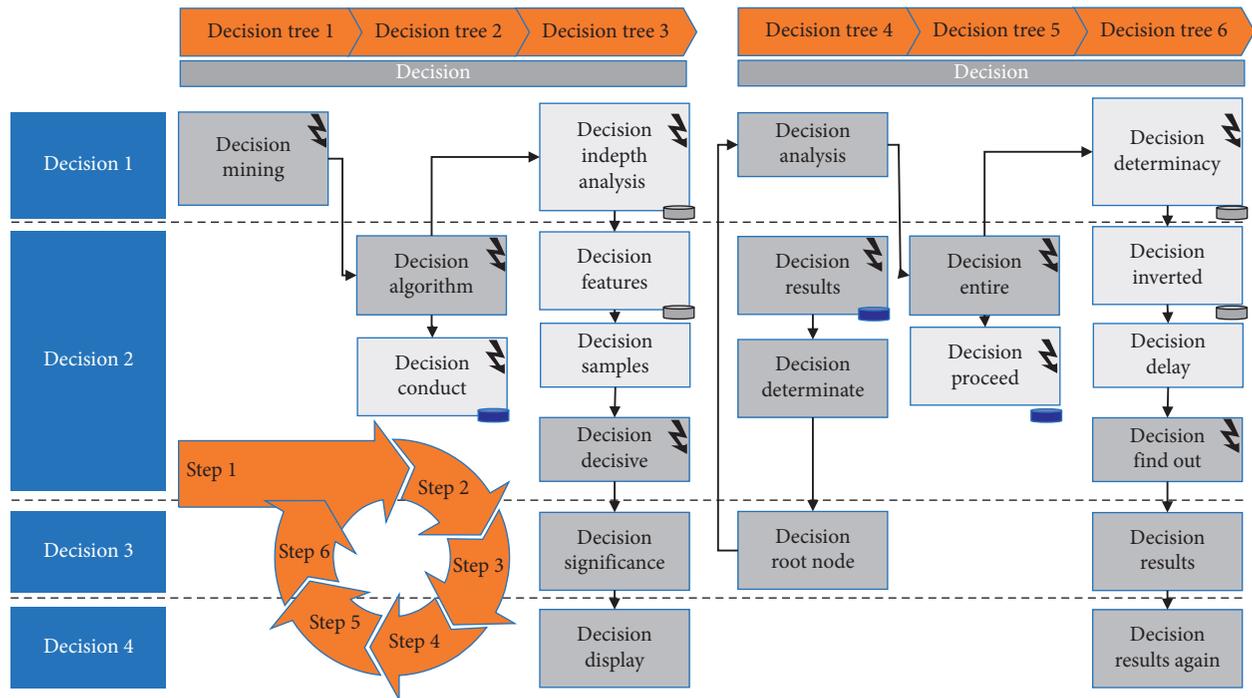


FIGURE 1: The processing model of the decision tree mining algorithm.

analysis on the results of students. As the educational administrators do not have a high level of understanding of data mining technology, they need to mine data to ensure the interpretation results of data mining. The decision tree mining algorithm can classify the tree structure of the decision tree and generate “If form”, which can avoid the difficulty of educators to understand. (2) The amount of calculation is relatively small. The data analysis of the educational administration system must have strong practicability, which must ensure that the data analysis has higher efficiency. In decision tree mining algorithm, compared to the data mining of various algorithms, calculation is relatively small, making the speed of data mining much higher than that of other algorithms, which can greatly reduce the use of data mining time, resulting in very high work efficiency. (3) Discrete and continuous data can be processed. There are many different types of students’ academic performance, including grades and teachers’ evaluation of students’ grades. Therefore, there are not only continuous data but also discrete data in grades. At the same time, discrete data constitutes the majority. For example, many courses adopt discrete data variables of high, medium, and low levels, and decision tree type can process both kinds of data. (4) It shows the importance of attributes. The expression way of decision tree is very intuitive, which can directly reflect the importance of the zodiac, which is directly represented by the current level of decision tree node. For example, a more important attribute will be at a higher level, and a less important attribute will be at a lower level [17–20].

- (4) The structure of the decision tree mining algorithm steps: The decision tree mining algorithm to construct the decision tree is divided into two steps, respectively, decision tree generation and pruning; The decision tree is a root node that is initially generated and has undergone a gradual process generation from top to bottom. The process of data segmentation and formation completes the structure of the decision tree. Using algorithms in the decision tree mining process, the decision tree can be tested from the root node, and the decision result will determine the next one until the last point. The pruning of decision tree is to prune the decision tree and remove the redundant branches through testing, so as to get the decision tree with the least expected error rate.
- (5) ID3 algorithm: The currently most commonly used decision tree mining algorithm is the ID3 algorithm. The basic idea of the method is recursive downward search on the training sample set. It is a kind of typical greedy algorithm. The decision tree of every node can test each attribute and the use of information gain as the attribute selection criteria and select the maximum attribute of information gain as the node of decision tree. To build a decision tree, the concept of information will be used in the algorithm to complete. (1) Information entropy. Information entropy is the expectation of all kinds of information, which can measure the uncertainty of the whole information element  $X$ . In the data set with  $X$  as the sample, all possible signals in the signal source are represented by the symbolic number  $N$ . Let the

possible value be AI, and the probability when the value is AI is  $P(AI)$ . The relationship between information entropy is obvious. (2) Conditional entropy: it is the expected value of information entropy under different conditions, where the signal source is  $b_j$  corresponding to  $Y$ , and the signal source is  $a_i$  corresponding to  $X$ , the probability of which is  $p(a_i|b_j)$ . (3) Average information gain: in the process of operation, the difference between two amounts of information is generally expressed by information gain. In the process of selecting classification attributes, the larger information gain of gyroscope is generally regarded as classification attributes.

The decision diagram of ID3 algorithm is shown in Figure 2.

### 3.2. Features and Stages of Decision Tree Mining Algorithm.

Decision trees are shaped like trees, so the characteristics include the following: (1) a decision tree consists of a series of nodes and branches. (2) Branches are formed between nodes and subnodes. Nodes represent attributes considered in the decision-making process, and different attribute values form different branches. Based on the idea of CLS algorithm and Quinlan's ID3 algorithm, the improved decision tree learning algorithm is as follows: (1) Generate an empty decision tree and a training sample table. (2) If all samples in the training sample set  $T$  belong to the same class, node  $T$  is generated and the learning algorithm is terminated. (3) According to the principle of maximum information gain, first select the attribute with the largest information gain from the training sample attributes, and generate testability node, namely, root node 8. (4) If the value of  $A$  is  $1, \dots, m$ , then according to the different value of  $A$ ,  $T$  is divided into  $m$  subsets  $T_1, T_2, \dots, T_m$ . (5) For each  $T_i$ , turn to Step (2). The specific process is shown in Figure 3:

To analyze students' performance data, it is necessary to integrate the data of students' various basic subjects, relevant professional courses, and participation in various school activities and then show the comprehensive quality of students. At the same time, for the students who are about to graduate, all the results including the graduation design results can be analyzed, so as to make more clear the interaction between various disciplines, and teachers and educational administrators can arrange courses more scientifically in the future.

- (1) Data collection stage: there are many students' scores in the educational administration system. This stage is to collect these scores and conduct a preliminary analysis of the scores. In the process of collection, the integrity of the students' scores will be checked. Secondly, the questionnaire survey can also be used in the form of statistics of specific data, for the comprehensive analysis of the data of students. In the process of collection, the data can be relatively simple statistical collation.
- (2) Data preprocessing stage: among the data, there may be some incomplete data and some data containing

noise, which have a great impact on the final data analysis results. The data can be used in project risk assessment to analyze the feasibility of the project. In the preprocessing of the data, there is a need to eliminate the incomplete data and noise data, find the defects in the data, and reasonably complete the data to ensure the integrity and stability of the data. In addition, for some special data, conversion is needed to ensure the processing of data, improve the efficiency of data mining, and reduce the time of data processing.

- (3) Data conversion: the decision tree mining algorithm mainly carries out the calculation of discrete data, and it is difficult to give full play to the superiority of the decision tree mining algorithm when dealing with continuous data such as students' grades. The work of data transformation is to process continuous data into discrete data. For example, students' grades are divided into grades: 85–100 is an A; a score of 84 to 70 is B; a score below 69 is CO.
- (4) Data mining: in data mining, information entropy, conditional entropy, and average gain information can be used for mining according to the actual situation of the data. In actual work, the relevant personnel can determine which method to use according to the situation.
- (5) Function of the achievement analysis system: (1) management function of the system. The management module of the system needs to be able to record the information of users. At the same time, in order to avoid information leakage of students, the identity of users must be verified before entering the system, and the identity information that meets the requirements of the system can enter the system; otherwise, it is not allowed to enter. At the same time, the identity of people logging into the system will also be different, which requires the classification of the permissions of different people. (2) Basic data management function. The main objective of basic data management is to maintain the basic data used by the system, such as teachers' information, students' grades, course information, and relationship between teaching and so on. This function includes modifying, adding, deleting and querying all kinds of information. For example, for the teacher, you can give the administrator's authority; after the teacher logs in, the system needs to give the relevant authority. 3. Managing achievement data. The management of score data includes querying score, printing score, and modifying program. Students should only have the right to query their scores, so they can query their scores of each subject, while teachers can manage their grades. 4. Performance early-warning management. This function mainly realizes early warning of students' grades and encourages students to strengthen their learning of a certain subject. In order to realize the function of warning students, students' performance data can be

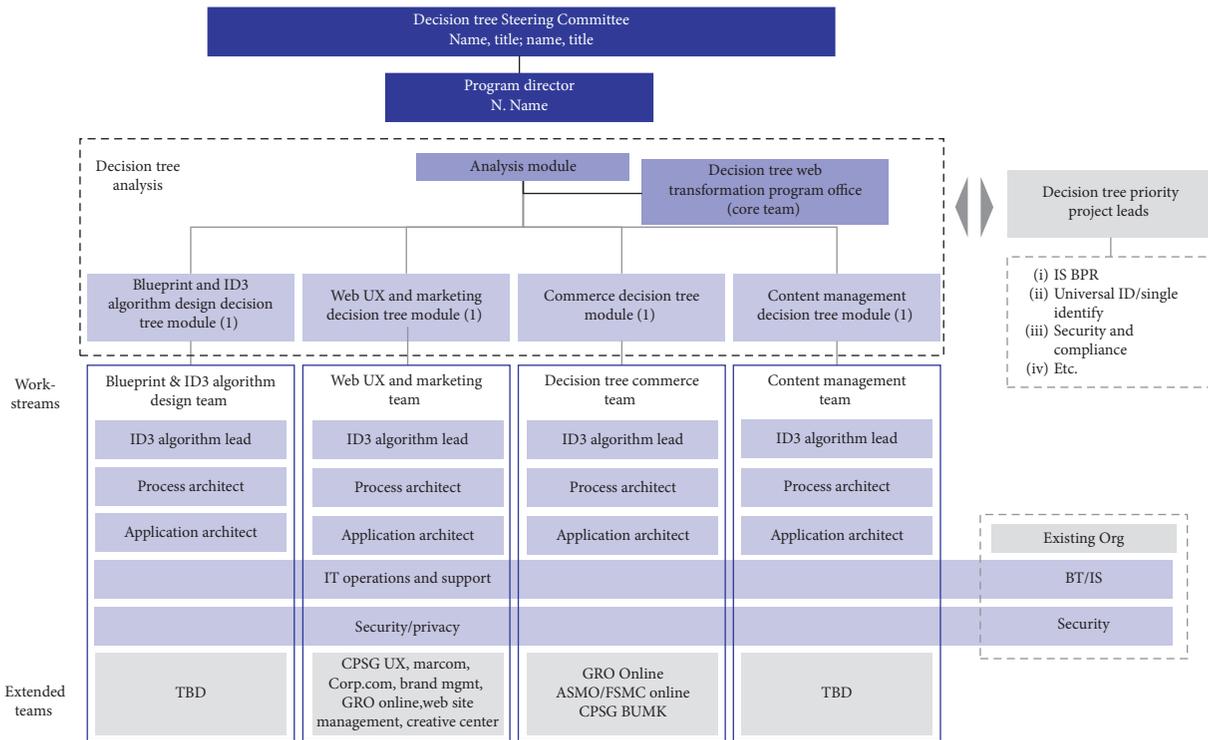


FIGURE 2: The decision diagram of ID3 algorithm.

Subthemes	Potential future signpost events	Actions	Class	Themes										
				AB	DC	DE	FG	HI	JK	LM	NO	PQ		
Alpha	Genetically-modified bio-system produces methane from organic waste at a steady state cost of \$5 per 1000 cubic feet of gas produced.	1	SB	●	●			●						●
	Publicized commercial use of metallic organic framework (MOF) in semiconductor manufacture.	0	EK	●		●								●
Beta	Reported lab demonstration of MOF gas storage device coupled with fuel cell to produce electricity.	4	EK	●				●	●					
	Public report of sensor network used in detecting and repairing faults in a dike or levee system, preventing flooding.	0	EK	●					●	●				

Recommended action key 0 recommend increased frequency of observation 1 recommend funding exploratory research 2 recommend early (in-house) experiments to test feasibility	3 recommend adoption of new standards 4 recommend funding of new development in this area 5 recommend trial development of new technology 6 recommend deployment of new technology	Classification key SB scientific break through P/E political or economic event EK enhancement of known technology	Theme key AB alpha DC beta DE gamma FG delta HI epsilon JK zeta LM sigma NO chi
--	---	--	---

FIGURE 3: The specific process.

analyzed to determine whether students' performance has met the warning conditions. It also includes the setting of the warning object and the way of alarm. The early-warning operation management sends early-warning information to the early-warning object, and the early-warning level can be divided according to the performance level, including first-level early warning and second-level

early warning. The achievement evaluation model is shown in Figure 4.

#### 4. Online Education Course Score Evaluation Based on Decision Tree Mining Algorithm

4.1. Establishment of Online Education Course Achievement Evaluation Model Based on Decision Tree Mining Algorithm.

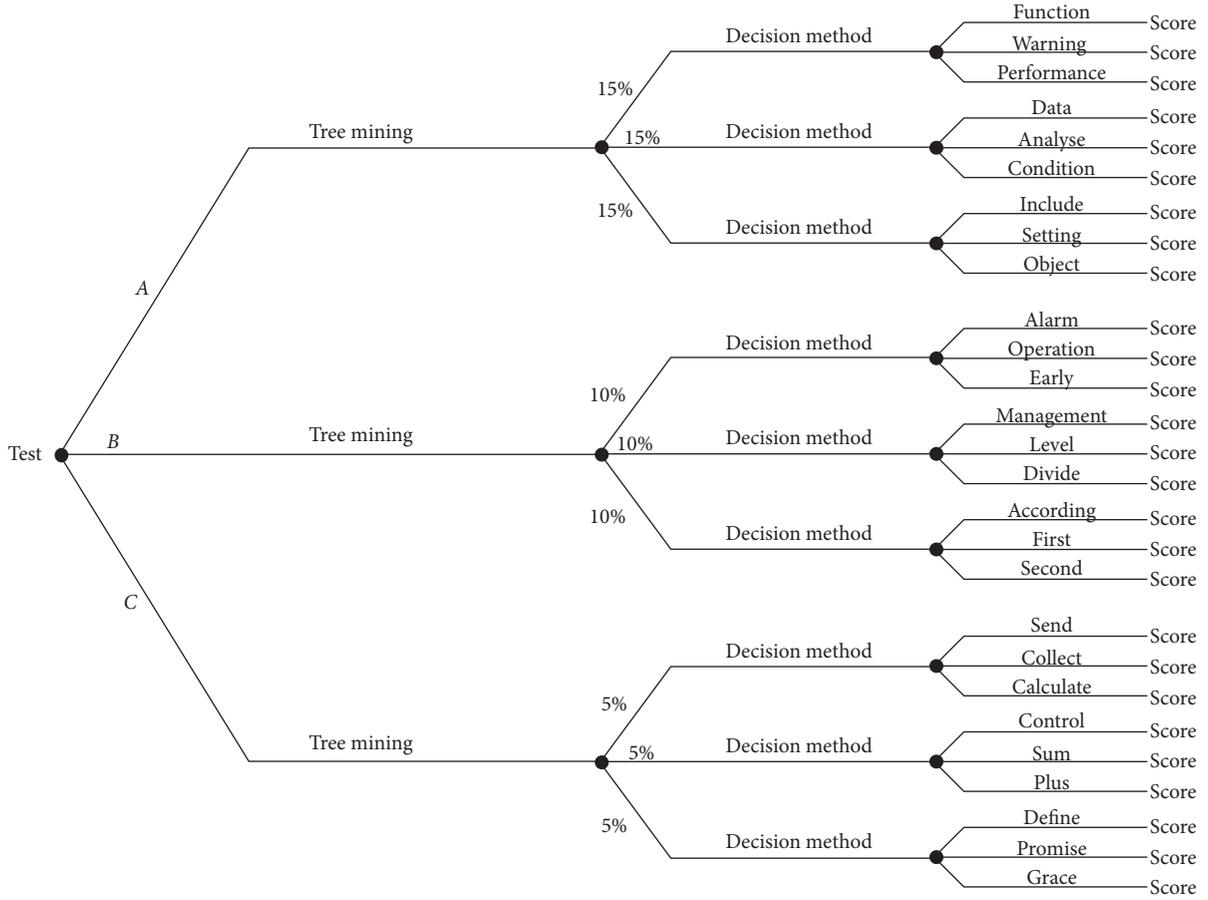


FIGURE 4: The achievement evaluation model.

This paper takes the online learning status of financial management in the first semester of the 2019–2020 academic year of a university as the research object. The classes learning this course include Taxation Classes 1 and 2 and Accounting Class 1, among which 168 students have valid data sets. In this paper, the learning behavior data of students is collected through the Superstar Learning link learning platform. The data set includes students' basic information, learning behavior characteristics, and academic performance. In order to facilitate the subsequent performance evaluation and analysis, the learning behavior data is encoded. The meanings and codes of students' learning behavior characteristics and achievements are shown in Table 1.

Data preprocessing is related to the mining of the adequacy of behavioral characteristics affecting performance. Because the achievement is continuous attributes, before data discretization, it is divided into four grade intervals: A (good), B (good), C (in general), D (poor), representing the score range for [85, 100], [75, 85], [60, 75], [0, 60). Other learning behavior characteristics using entropy are divided into less or more intervals, and the calculation formulas are

$$J(E) = J(e_1, \dots, e_n) = - \sum_{m=1}^n Q_m * \log_2(Q_m) + \text{ff}m,$$

$$F(I) = \sum_{n=1}^m \frac{|E_n|}{e} J(e_{1n}, \dots, e_{in}) + \text{ff}in,$$

$$J(e_{1n}, \dots, e_{in}) = \sum_{m=1}^m Q_{mn}(e_{1n}, \dots, e_{in}) + \log_2(Q_{mn}).$$

(1)

Some kind of cross validation is also used in the experiment, such as the cross of experimental data models to prove that the decision model is fully available. The Pearson correlation coefficient analysis method is commonly used to analyze the linear correlation degree of two variables. This paper uses this analysis method to analyze the correlation between learning behavior characteristics and learning performance. In addition, this article also uses relevant evaluation criteria: 0.00–0.30 means not relevant, 0.30–0.50 is weakly relevant, 0.50–0.80 is moderately relevant, and 0.80–1 is strongly relevant. This paper analyzes the correlation coefficient matrix of learning behavior characteristics and performance and visualizes the results through heat maps. The results are shown in Table 2.

TABLE 1: The meanings and codes of students' learning behavior characteristics and achievements.

Code	Characteristic of learning behavior	Meaning
SP1	Video task point completion	Completing the number of videos to watch
SP2	Video viewing time	Total time spent watching the video
HM1	Number of chapter quizzes completed	Number of chapter tests completed
HM2	Job completion number	The number of assignments completed
HM3	Study visits	Viewing the learning resources and courseware of the learning platform
CL1	Number of sign-in times	The number of times a student attends class on time
CL2	Quantity of course interaction	Number of interactions between teachers and students
TS1	Section test average score	The average score for all section tests
TS2	Job average	The average grade for all assignments
TS3	The final result	Total score on the final exam

TABLE 2: Learning behavior characteristic correlation.

	SP1 (%)	SP2 (%)	HM1 (%)	HM2 (%)	HM3 (%)	CL1 (%)	CL2	TS1	TS2 (%)	TS3 (%)
SP1	100	45	70	60	45	45	50%	25%	45	80
SP2	45	100	25	35	35	30	40%	2%	20	35
HM1	70	25	100	70	65	55	30%	60%	50	70
HM2	60	35	70	100	55	55	30%	40%	45	80
HM3	45	35	65	55	100	75	55%	20%	70	60
CL1	45	30	55	55	75	100	45%	20%	55	50
CL2	50	40	30	30	55	45	100%	-10%	40	40
TS1	25	2	60	40	20	20	-10%	100%	20	30
TS2	45	20	50	45	70	55	40%	20%	100	50
TS3	80	35	70	80	60	50	40%	30%	50	100

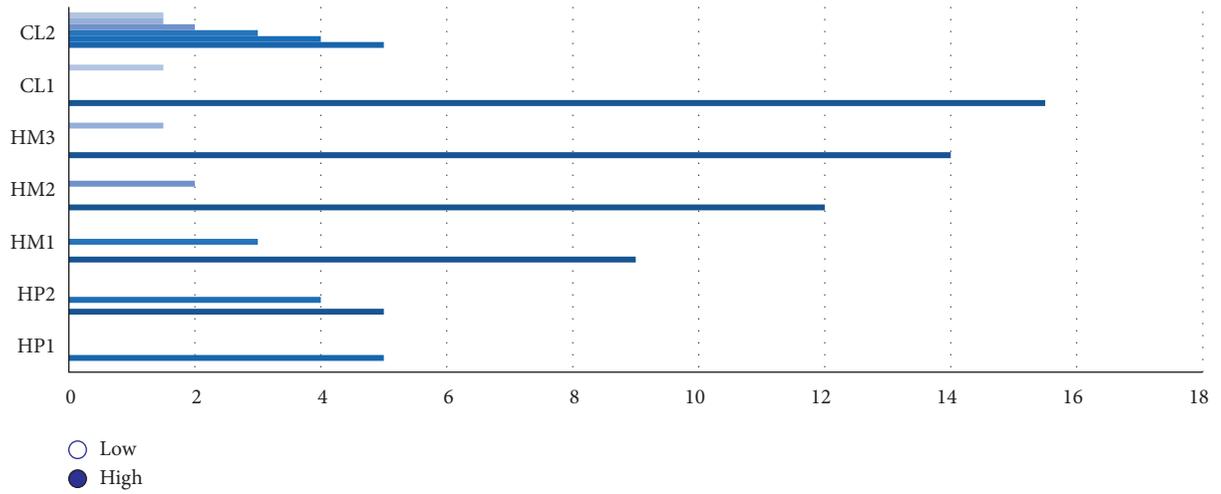
The  $t$ -test method is also feasible in this paper. The  $t$ -test is divided into single-population test and double-population test. Single-population test is to test whether the difference between the mean of a sample and a known population mean is significant. When the population distribution is normally distributed, for example, the population standard deviation is unknown and the sample size is less than 30, then the deviation statistics between the sample mean and the population mean show a T-distribution. Double-population sample test includes independent sample  $t$ -test and paired sample  $t$ -test. Independent sample  $t$ -test is used to test whether the difference of population mean represented by two independent samples is significant. Applicable conditions: (1) both samples are from the normal population; (2) the two samples are independent of each other; (3) the samples satisfy homogeneity of variance (i.e., pass homogeneity test of variance). Homogeneity test of variance is a method to check whether the population variance of different samples is the same in mathematical statistics. The basic principle is to make a certain assumption about the characteristics of the population and then make inferences about whether the assumption should be rejected or accepted through statistical reasoning of the sampling study. Commonly used methods are Hartley test, Bartlett test, and modified Bartlett test. The learning behavior characteristic justification is shown in Figure 5.

As can be seen from Figure 5 and Table 2, there is a moderate correlation between the completion of the video task points, the completion of the chapter test, the average score of the chapter test, the completion rate of the homework, the average score of the homework, and the

grade. Video viewing time, study visits, and course interactions were weakly correlated with grades. There is no correlation between academic completion rate and academic achievement. Therefore, from the results of the correlation coefficient analysis, it can be seen that the sign-in completion rate is not considered in the construction of the subsequent performance evaluation model. By comparing the above calculation results, it can be seen that B1 has the highest information gain rate. Therefore, select it as the root node. Then, use the branch of the root node to calculate the information gain rate of the remaining features, and recursively call the above calculation process to select the best feature until the training set is empty.

*4.2. Evaluation of Online Education Course Achievement Evaluation Model Based on Decision Tree Mining Algorithm.* Model evaluation uses the test data set to test and evaluate the performance evaluation model through the confusion matrix, and evaluates the classification performance of the performance evaluation model based on the accuracy, precision, recall, and F1 value indexes of the model, respectively.

In this paper, decision tree C4.5 algorithm is used to evaluate students' performance, and its evaluation accuracy is about 88%. The higher the recall rate and F value are, the better the model performance is. It can be clearly seen from Table 2 that the recall rate in each interval is greater than 85%, achieving the expected effect. Therefore, this paper builds a performance evaluation model based on decision tree C4.5 algorithm, as shown in Figure 6 in the form of a tree.



	HP1	HP2	HM1	HM2	HM3	CL1	CL2
Breadth and depth of functionality	○	○	○	○	○	○	○
Strength of architecture	◐	◐	◐	◐	◐	◐	◐
Ability to support	◑	◑	◑	◑	◑	◑	◑
Ease of implementation	◒	◒	◒	◒	◒	◒	◒
Overall strength of solution	●	●	●	●	●	●	●

FIGURE 5: The learning behavior characteristic justification.

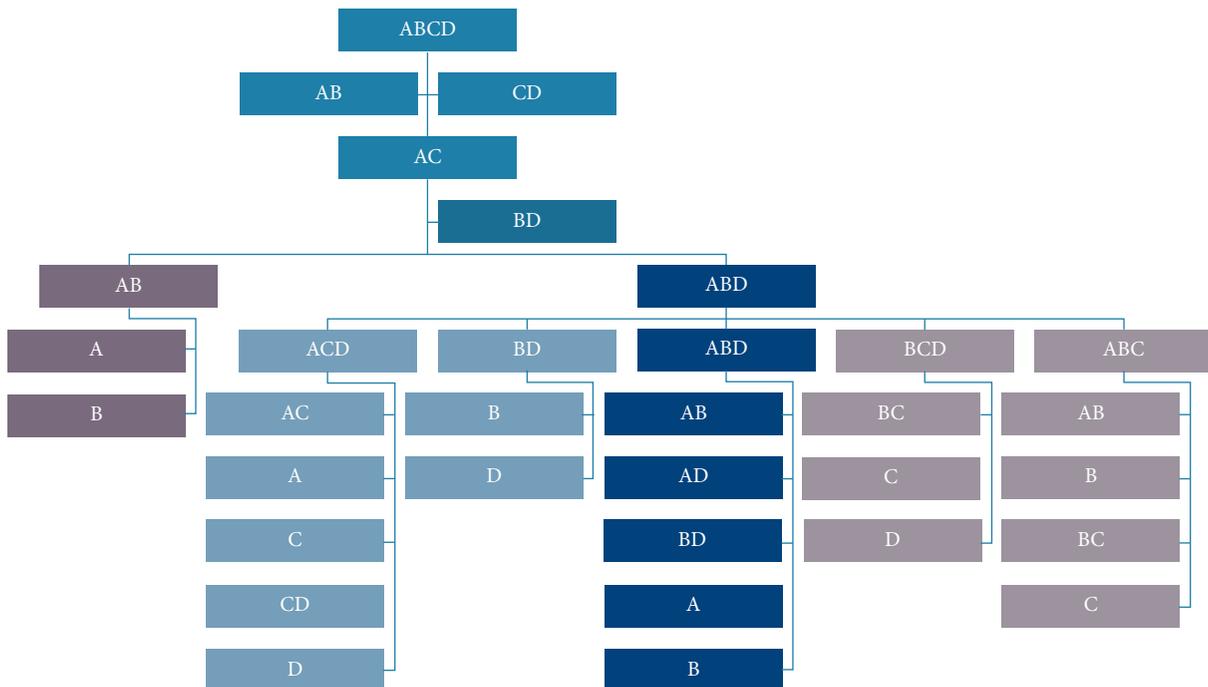


FIGURE 6: A performance evaluation model based on decision tree C4.5 algorithm.

The classification rules generated by the decision tree model are the path from the root node of the decision tree to each leaf node, and the classification rules are described in the form of if-then, so as to extract the main factors affecting the performance. According to Figure 6, the classification rule of decision tree generation is as follows: If the number of video task points completed is 2 “much,” the average score of chapter test is 2 “A,” the number of chapter tests completed is 2 “much,” and the number of learning visits is 2 “much,” then the score is 2 “C.” If the number of video task points completed is 2 “A,” the average score of chapter test is 2 “A,” the average score of chapter test is 2 “A,” the average score of video viewing time is 2 “A,” and the average score of homework is 2 “C,” the formula is

$$F(I_n) = \sum_{n=1}^m \frac{|E_n|}{e} J(e_{1n}, \dots, e_{in}) + \iiint im. \quad (2)$$

According to the analysis of the model results, the video task point completion quantity is the root node of the tree; the chapter test completion number and the chapter test average separation root node are the closest; and the video viewing time, homework average score, and course interaction quantity are closer to the root node. Therefore, the most significant impact on students’ performance is that of the number of video task points completed, followed by the number of chapter test completed and chapter test average score, and the last is the amount of course interaction and homework average score. Based on this, this paper suggests the following in the actual online learning and teaching process: (1) students should strengthen the self-study ability and conscientiously complete the course video task. (2) In addition to learning the knowledge points of each chapter by heart, students should also pay attention to the examination content of each chapter to test their mastery of each chapter’s knowledge points. (3) Teachers should guide students to actively participate in the interaction of course content and questions, deepen students’ understanding of knowledge points, and effectively intervene in students’ online learning.

**4.3. Empirical Analysis.** The use case of this study is to use the data of students’ final math scores in a certain university as the mining object, and the data comes from the network. Through digging and analyzing, find out the main factors that affect students’ achievement. In this paper, an improved algorithm based on ID3 algorithm is used to build decision tree. To establish the decision tree, the following attributes should be considered: (1) the degree of interest of students in math class; (2) learning attitude; (3) attendance; (4) completing homework independently.

After data processing, the training set of data score information divides the test results into four categories: specifically, A (excellent), B (good), C (pass), and D (fail). Our output is A, B, C, D, 30 records in total. There are 17 records with A value, 9 records with a value of B, 2 records with a value of C, and 2 records with a value of D.

The following is to calculate the information gain: “degree of interest in math class,” “learning attitude,” and

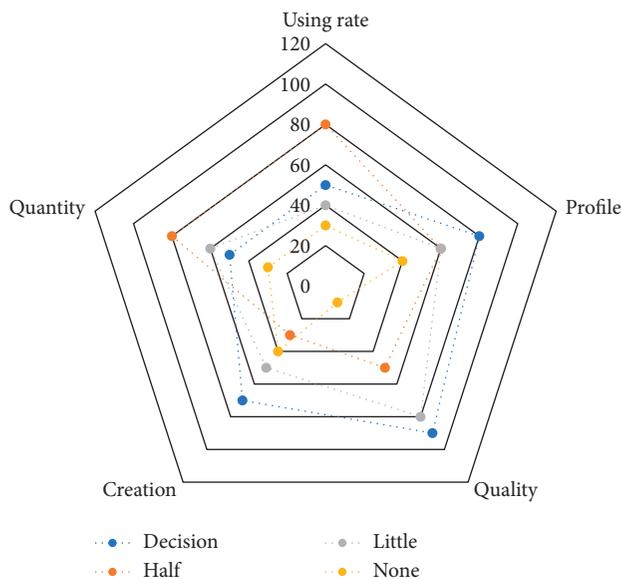


FIGURE 7: The difference between the decision result and the nondecision tree mining algorithm.

“class attendance” are taken as root nodes, respectively, to calculate the information gain. Take the attribute “degree of interest in math” as the root node. There are 17 records with values of interest, including 14 A, 2 B, 1 C, and 0 D values. There are 11 records with average values, including 3 A’s, 7 B’s, 0 C’s, and 1 D. There are 2 records with the value of “no interest,” including 1 D, 1 C, 0 A, and 0 B values.

Calculate its corresponding entropy:  $\text{Info}(\text{interest}) = 0.83435$ ;  $\text{Info}(\text{common}) = 1.24067$ ;  $\text{Info}(\text{uninterest}) = 1$ . Similarly,  $\text{Gain}(\text{interest in mathematics}) = 0.511972$ ,  $\text{Gain}(\text{learning attitude}) = 0.708688$ ,  $\text{Gain}(\text{class attendance}) = 0.395689$ , and  $\text{Gain}(\text{independent completion of homework}) = 0.774990$  are obtained for the corresponding information appreciation of the above four attributes. The difference between this result and the nondecision tree mining algorithm is shown in Figure 7.

Finally, according to the principle of maximizing information gain, “independently completed work” was selected as the root node, and the sample was divided into three parts. Then, each subtree was calculated in a recursive way and pruned. The results of data analysis and comparison are shown in Figure 8.

In the confusion matrix, there is a large proportion of negative cases “NO,” and a false negative case “YES” is judged to be cancer-free “NO.” The further calculation of the sensitivity of the classifier is  $90/300 = 30.00\%$ ,  $90/300 = 30.00\%$ , and the special effect is  $9650/9700 = 98.56\%$ . It can be seen that the classifier is very sensitive, simple, and effective. The decision tree obtained from the research shows that students who get an A (excellent) do a good job on their own when completing their homework. Students who are interested in mathematics tend to perform well in exams, most of which get A (excellent) or B (good). In contrast, students who performed poorly or were not interested in mathematics also tended to perform poorly in exams. In addition, the student’s learning attitude is also a factor that cannot be ignored. The impact of student attitudes on the decision tree model is shown in Figure 9.

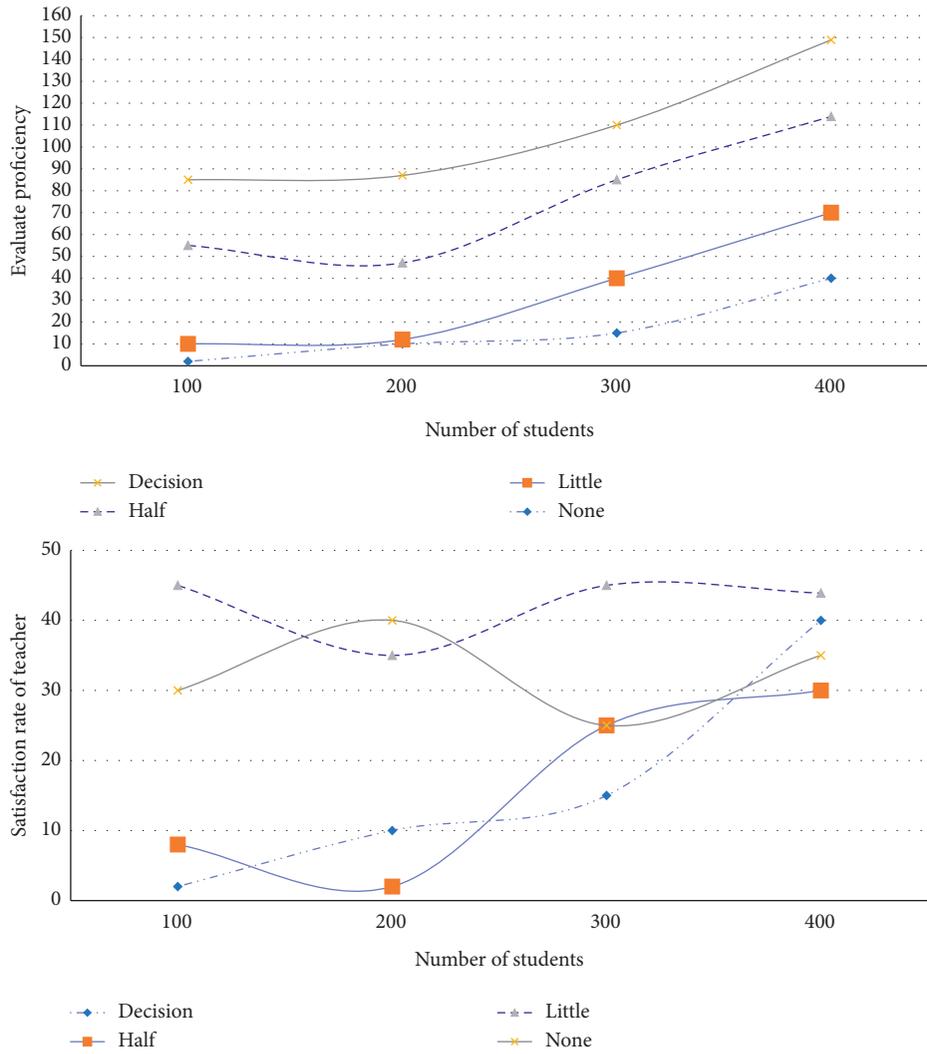


FIGURE 8: The results of data analysis and comparison.

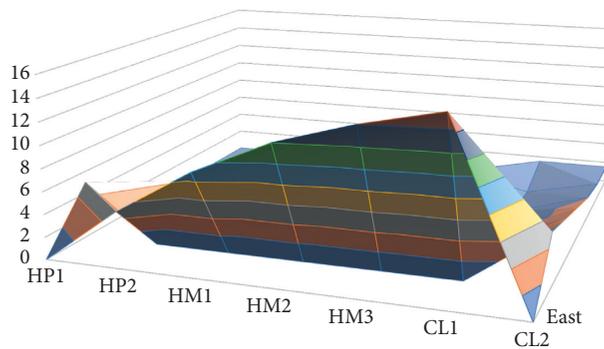


FIGURE 9: The influence of students' attitude on the decision tree model.

### 5. Conclusion

Through decision tree analysis method, we can analyze the implicit relationship between different subjects and construct decision analysis tree, so as to analyze students'

learning situation. Not only can we dig the implicit relationship between the buddies course, but we can also timely prompt students in the current problems existing in learning and improve the learning efficiency of students. In this paper, the decision tree C4.5 algorithm is used to construct

the student achievement evaluation model for the research, and the evaluation accuracy is about 88%, which has reached the expected effect of the research. The results of model classification rules generated by the score evaluation model based on decision tree C4.5 algorithm show that the learning behavior characteristics such as video task point completion amount, chapter test completion number, and chapter test average score are the most critical characteristics in students' online learning and have the greatest impact on students' academic performance. Therefore, both students and teachers should pay enough attention to the above key learning behavior characteristics in order to achieve ideal results in online learning or teaching. However, this paper still needs to be further improved. Firstly, this paper is based on the data generated by the teaching process of Superstar Learning online learning platform. The learning behavior characteristics involved in the data need to be supplemented, for example, the teacher's class arrangement, the difficulty of homework, and the teaching method of the course. Secondly, the obtained research results have not been applied to the actual online learning and teaching process, and the effect of the research results on the online learning and teaching process is not known. In the following work, we will continue to discuss the above deficiencies. On the one hand, the online learning platform is constantly used to improve the learning behavior data, so as to obtain more comprehensive key factors affecting students' academic performance, and make the research results more effective for online learning or teaching. On the other hand, we will work with students and course teachers to apply the research results of this paper to practical teaching, so as to test the effect of the research results on online learning and teaching process.

### Data Availability

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

### Conflicts of Interest

The author declares no conflicts of interest.

### Acknowledgments

The study was supported by the Financing of First-Class Subject (Education) Construction project of Ningxia University (Grant no. YLXKZD1928).

### References

- [1] K. Ramya, Y. Teekaraman, and K. A. Ramesh Kumar, "Fuzzy-based energy management system with decision tree algorithm for power security system," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1173–1178, 2019.
- [2] A. B. Raut and M. A. A. Nichat, "Students performance prediction using decision tree," *International Journal of Computational Intelligence Research*, vol. 13, no. 7, pp. 1735–1741, 2017.
- [3] R. M. Rani and M. Pushpalatha, "Generation of Frequent sensor epochs using efficient Parallel Distributed mining algorithm in large IOT," *Computer Communications*, vol. 148, no. 1, pp. 107–114, 2019.
- [4] A. K. Hamoud, A. S. Hashim, and W. A. Awadh, "Predicting student performance in higher education institutions using decision tree analysis," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 2, pp. 26–31, 2018.
- [5] J. Di and Y. Xu, "Decision tree improvement algorithm and its application," *International Core Journal of Engineering*, vol. 5, no. 9, pp. 151–158, 2019.
- [6] M. Jin, H. Wang, Q. Zhang et al., "Financial management and decision based on decision tree algorithm," *Wireless Personal Communications*, vol. 7, no. 2, pp. 118–147, 2018.
- [7] A. I. Weinberg and M. Last, "Interpretable decision-tree induction in a big data parallel framework," *International Journal of Applied Mathematics & Computer Science*, vol. 27, no. 4, pp. 737–748, 2018.
- [8] Q. Wu, "MOOC learning behavior analysis and teaching intelligent decision support method based on improved decision tree C4.5 algorithm," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 14, no. 12, p. 29, 2019.
- [9] W. F. W. Yaacob, S. A. M. Nasir, W. F. W. Yaacob, and N. M. Sobri, "Supervised data mining approach for predicting student performance," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 16, no. 3, pp. 1584–1592, 2019.
- [10] C. Fiarni, E. M. Sipayung, and P. B. T. Tumundo, "Academic decision support system for choosing information systems sub majors programs using decision tree algorithm," *Journal of Information Systems Engineering and Business Intelligence*, vol. 5, no. 1, pp. 57–66, 2019.
- [11] H. Y. Jeong and T. S. Yoon, "Analysis of hepatitis C virus using data mining algorithm-apriori, decision tree," *International Journal on Bioinformatics & Biosciences*, vol. 7, no. 3, pp. 01–11, 2017.
- [12] D. Y. Y. Sim, C. S. Teh, and A. I. Ismail, "Improved boosted decision tree algorithms by adaptive apriori and post-pruning for predicting obstructive sleep apnea," *Advanced Science Letters*, vol. 24, no. 3, pp. 1680–1684, 2018.
- [13] S. Hussain, N. Abdulaziz Dahan, F. M. Ba-Alwi, and N. Ribata, "Educational data mining and analysis of students' academic performance using WEKA," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 9, no. 2, pp. 447–459, 2018.
- [14] H. Kaur, "A literature review from 2011 TO 2014 on student's academic performance prediction and analysis using decision tree algorithm," *Journal of Global Research in Computer Science*, vol. 9, no. 5, pp. 10–15, 2018.
- [15] J.-J. Sheu and K.-T. Chu, "Mining association rules between positive word-of-mouth on social network sites and consumer acceptance: a study for derivative product of animations, comics, and games," *Telematics and Informatics*, vol. 34, no. 4, pp. 22–33, 2017.
- [16] M. Chen, J. Yang, J. Zhou, Y. Hao, J. Zhang, and C.-H. Youn, "5G-Smart diabetes: toward personalized diabetes diagnosis with healthcare big data clouds," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 16–23, 2018.
- [17] J. Wang, A. Liu, T. Yan, and Z. Zeng, "A resource allocation model based on double-sided combinational auctions for transparent computing," *Peer-to-Peer Networking and Applications*, vol. 11, no. 4, pp. 679–696, 2018.

- [18] S. Amaran, N. V. Sahinidis, B. Sharda, and S. J. Bury, "Simulation optimization: a review of algorithms and applications," *Annals of Operations Research*, vol. 240, no. 1, pp. 351–380, 2016.
- [19] Q. Feng, D. He, S. Zeadally, M. K. Khan, and N. Kumar, "A survey on privacy protection in blockchain system," *Journal of Network and Computer Applications*, vol. 126, pp. 45–58, 2019.
- [20] H. Li, L. Pei, D. Liao, G. Sun, and D. Xu, "Blockchain meets VANET: an architecture for identity and location privacy protection in VANET," *Peer-to-Peer Networking and Applications*, vol. 12, no. 5, pp. 1178–1193, 2019.