Research on college English teaching based on data mining technology

Jinhui Duan[1*] and Rui Gao[2]

*Correspondence:
djhui98@163.com
[1] Kunming University
of Science and Technology
Oxbridge College,
Kunming 650106, Yunnan,
China
Full list of author information
is available at the end of the
article

Abstract

To improve the efficiency and quality of college English teaching, we analyzed the feasibility and application process of data mining technology in college English teaching. The entire process of data classification mining was fully realized. A new teaching program was proposed. The object and target of data mining were determined. Online surveys were used to collect data. Data integration, data cleaning, data conversion, data reduction and other pre-processing technologies were adopted. The decision tree was generated by using the C4.5 algorithm, and the pruning was carried out. The result analysis decision tree model was completed. A detailed survey of the students' English learning in University was made in detail. The results showed that the qualified rate of students' English performance was increased from 20–30% to 50–60%. Therefore, the classification rules provide theoretical support for the school teaching decision. This method can improve the quality of English teaching.

Keywords: Data mining, Classification, Decision tree, Classification rules

1 Introduction

Due to the popularization of mass higher education, the increase in the number of students has brought many new problems to the teaching work of colleges and universities. Some students have not developed good study habits [1–3]. The students' English foundation is very poor. Self-control ability is very weak. The learning goal is not clear. Most students lack the spirit of hard work [4–6]. Therefore, teachers urgently need to study a large amount of data and information in each link of university teaching, obtain knowledge from it, and guide teaching scientifically. Over the years, a large amount of data has been accumulated in the teaching and management of universities. However, these data have not been effectively used [7, 8]. Rich information resources are used. Use data mining classification technology. Acquire knowledge to aid decision-making. It can guide teaching and further improve the quality of teaching.

Data mining is the process of extracting implicit and potentially useful information and knowledge from a large amount of incomplete, noisy, fuzzy and random data. It is the core of knowledge discovery in the database [9, 10]. Compared with foreign countries, the domestic research on data mining is relatively late. There is no overall strength. At present, people engaged in data mining research in China are mainly concentrated

in universities, and some are in research institutions or companies. There are many research fields involved, usually focusing on algorithm research, the practical application of data mining and the research of data mining theory. Currently, organizations that research data mining emerge in endlessly, such as the Data Mining Centre of Renmin University of China [11]. In recent years, research work on data mining has been carried out. Representative experts engaged in data mining research include Professor Xin Banghyang. Although China has not yet developed data mining products worth promoting, the development of data mining disciplines is very rapid. It is now at the advanced level in the world [12].

In order to improve the efficiency and quality of college English teaching, the feasibility and application process of data mining technology in college English teaching are analyzed. A new teaching plan was proposed. Determine the purpose and goal of data mining. Online surveys are used to collect data. Data integration, data cleaning, data conversion, data reduction and other pre-processing techniques are used. This method can improve the quality of English teaching.

2 Related work

In China, data mining technology is used in online English learning platforms, college English listening, teaching quality, test results, course score analysis and management, etc., providing data references for teaching decision-making and improvement. In the online English learning platform, cluster analysis is used to cluster the students' mastery of English level, association rule algorithm is used to analyze the connection between exercises and exercises, and genetic algorithm is used to develop an automatic organization of English learning content system [13, 14]. In terms of the quality of college English teaching, data mining and analysis are carried out based on the teaching factors that affect the quality of students' learning, and the importance of the factors affecting the quality of learning is explored, and based on the analysis results, the teaching quality is improved in a targeted manner [15]. In terms of college English listening According to college English listening scores, Internet research data, etc., using data such as gender, teaching methods, learning goals, test difficulty, teaching language, etc., dig out the main factors affecting listening test scores [16]. In the college English test, according to English Data such as foundation, effort level, learning method, interest, and English atmosphere, use decision tree algorithm to mine the rules to predict whether the test will pass [17].

Data mining can be classified according to database types, mining objects, mining tasks, mining methods and techniques [18]. Data mining tasks include correlation analysis, cluster analysis, classification, prediction, timing model, and bias analysis [19]. The objects of data mining are mainly relational databases, but it is very difficult to find the necessary data and information from millions of multimedia data. Data mining methods are developed from artificial intelligence and machine learning methods. Combining traditional statistical analysis methods, fuzzy mathematics methods and visualization techniques, the database is taken as the research object, and data mining methods and techniques are formed. Classification is an important data mining technology. The commonly used classification methods include decision tree classification, Bayesian classification, neural network classification, and genetic algorithm. Different algorithms are

suitable for different types of data [20]. The comparison of the three algorithms is shown in Table 1.

The decision tree can be easily used for discrete attribute data. However, when the value of the property is very complex, the effect may be worse. If the number of branches is limited, the effect of the decision tree is good. The neural network needs to transform the discrete attributes into numerical attributes. For cases where numerical properties are dominant, the neural network converts all inputs to 0–1. The objects of data mining are mainly relational databases, but it is very difficult to find the necessary data and information from millions of multimedia data. Data mining methods are developed from artificial intelligence and machine learning methods [21, 22]. Combining traditional statistical analysis methods, fuzzy mathematics methods and visualization techniques, the database is taken as the research object, and data mining methods and techniques are formed. Decision trees can be handled by dividing the values. If the number of attributes in the record is large, the neural network will be affected by it. The degree of influence of the decision tree is relatively small [23, 24]. If there are multiple dependent variables, neural network is the best choice. For the chronological data, the neural network has a better ability to process the chronological data. Decision trees can also handle the time sequence, but the required data preparation is relatively high. The Bayesian classification algorithm is quick and easy, and it takes only one traversal of data to end and get the result. However, this algorithm does not output direct classification rules. Therefore, it is usually used in processing studies to obtain a preliminary result [25].

3 Methods

Data mining is a decision support process. It is a deep-level data information analysis method. The application of data mining technology to teaching evaluation is undoubtedly very beneficial. Each value of the attribute produces a branch. The corresponding sample subset of the branch attribute values is moved to the newly generated child node. This algorithm is applied recursively to each child node until all samples of the node are partitioned into a class. Each path of the leaf node that reaches the decision tree represents a classification rule [26, 27]. It can fully analyze the hidden relationship between test results CRS and various factors, which is not a traditional evaluation method. Through data mining analysis, its evaluation results can bring unprecedented gains and surprises to teaching [28]. The structure diagram of the decision tree model is shown in Fig. 1.

An online survey system was designed. The data of 360 students were collected. The first is the student achievement table, including student number, name, and English score. This information is obtained through the student achievement management

Table 1 Comparison of several common classification algorithms

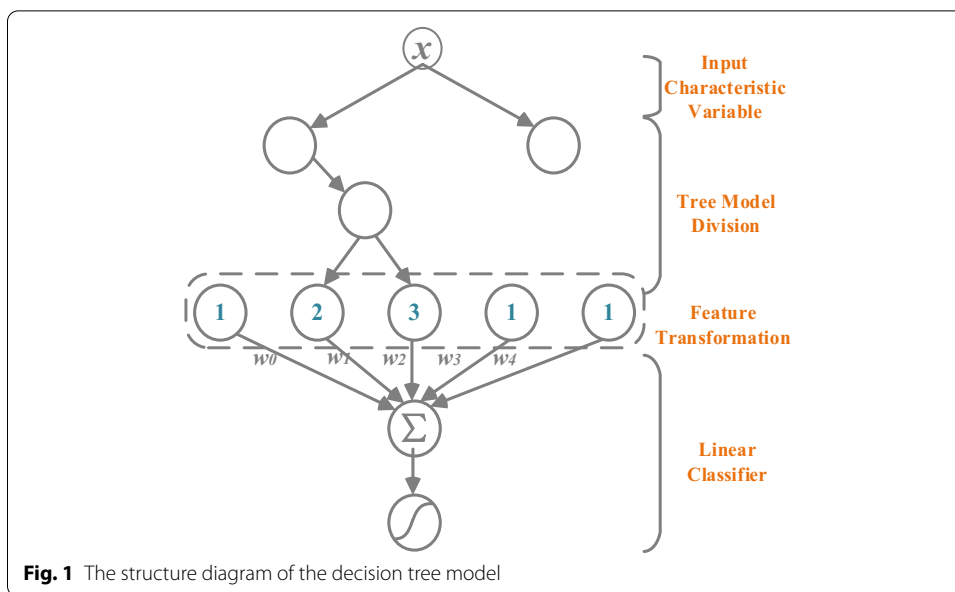|	Understandability	Easy training	Feasibility	Generality	Usefulness	Degree of product availability
Bayesian network	B	B	B	B	B	B
Decision tree	A+	B	A+	A	A	B+
Neural network	C—	B—	A—	A	A	A

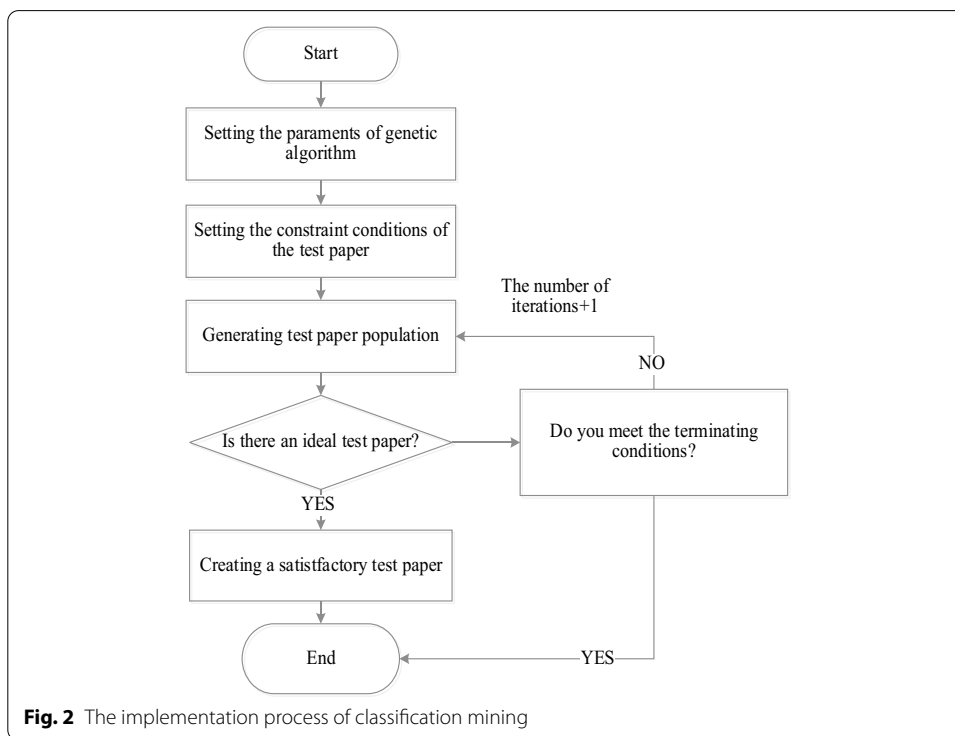**Fig. 1** The structure diagram of the decision tree model

Fig. 2 The implementation process of classification mining

system of the school administration office. The second is students' interest in English, the effects of classroom learning, the knowledge of the course, students' abilities of listening, speaking, reading and writing, the vocabulary they master, the teaching methods of teachers, and the time of extracurricular study [29, 30]. The implementation process of classification mining is shown in Fig. 2.

This part is generated through an online survey and completed by the student. Through data compilation, a total of 356 valid data were obtained, accounting for 98.9% of the survey participants. The valid data are then merged into a student performance analysis database. After data pre-processing, the structure of the student performance analysis database is as follows: "The degree of interest in the course," "The effect of the classroom learning," "The vocabulary that the students have mastered," and "The teaching method of the teacher." "The time of extracurricular learning" is a non-category attribute. "Whether it is good" is a category attribute [31, 32].

The decision tree C4.5 algorithm is used in this paper. The algorithm starts from all training samples at the root node of the tree and selects one attribute to distinguish the samples. Each value of the attribute produces a branch. The corresponding sample subset of the branch attribute values is moved to the newly generated child node. This algorithm is applied recursively to each child node until all samples of the node are partitioned into a class [33, 34]. Each path of the leaf node that reaches the decision tree represents a classification rule. The key decision of the top-down decision tree generation algorithm is the selection of the node attribute values. Different attribute values make the subsets of recorded records different. It affects the growth of the decision tree and the quality of the decision tree structure, thus affecting the rule information [35]. According to the above principles, the following steps are used to construct the decision tree. The first step is to calculate the information entropy required for a given sample classification. The second step calculates the rate of information gain for each attribute. The third step determines the test properties. The fourth step is to further divide the branch nodes with the above method.

$$\text{net}_{pj} = \text{net}_j = \sum_{i=1}^{M} \omega_{ij} o_i \tag{1}$$

The output of the i-th node is as follows:

$$o_j = f(\text{net}_j) \tag{2}$$

The Sigmoid function is selected as the activation function.

$$f(\text{net}_j) = \frac{1}{1 + \exp[-(\text{net}_j - \theta_j)]} \tag{3}$$

The differential function of Eq. (3) is transformed into the following equation.

$$f'(\text{net}_j) = f(\text{net}_j)[1 - f(\text{net}_j)] \tag{4}$$

By adjusting the value of the weighting coefficient of the hidden layer, the information of the first node obtained by the hidden layer is transmitted to the corresponding node of the next layer. The degree of influence of the decision tree is relatively small. If there are multiple dependent variables, neural network is the best choice. For the chronological data, the neural network has a better ability to process the chronological data. Decision trees can also handle the time sequence, but the required data preparation is relatively high.

$$\text{net}_k = \sum_{j=1}^{q} \omega_{jk} o_j \tag{5}$$

In the principle of the decision tree C4.5 algorithm, the information entropy formula of the sample space is as follows:

$$I(p, n) = -\frac{p}{p+n} \log\left(\frac{p}{p+n}\right) - \frac{n}{p+n} \log\left(\frac{n}{p+n}\right) \tag{6}$$

The expected information entropy with X_i as the test attribute is as follows:

$$E(X_i) = \sum_{j=1}^{k} \frac{p_j}{p_j + n_j} I(p_j, n_j) \tag{7}$$

The information gain with X_i as the root node is as follows:

$$\text{Gain}(X_i) = I(p, n) - E(X_i) \tag{8}$$

The rate of information gain is as follows:

$$\text{Gain Ratio}(X_i) = \frac{\text{Gain}(X_i)}{\text{Spliti}(X_i)} \tag{9}$$

Among them:

$$\text{Spliti}(X_i) = -\sum_{j=1}^{k} \frac{p_j}{m} \log\left(\frac{p_j}{m}\right) \tag{10}$$

The tree obtained by the above algorithm tends to grow too large to produce "excessive" phenomena to the training data, which in turn reduces the comprehensibility and usability of the tree. In other words, this decision tree may be very accurate for this historical data. Once it is applied to new data, the accuracy drops sharply. To prevent over-training and reduce training time, methods that can stop the tree from growing at the right time need to be established. Post-pruning is used. The basic idea is to make the decision tree fully grow first and then use the pruning technique to remove the non-general foliage. Generally, it is determined whether to reserve this branch by measuring the improvement of the classification performance of a certain branch, such as the specified classification error rate, and the degree of complexity of the entire tree.

4 Experiment

The first step is to calculate the information entropy required for a given sample classification. 356 samples were obtained for analysis and calculation. Using formula (1), I (S1, S2) = 0.960 is calculated. The second step is to calculate the information gain rate for each attribute. The Bayesian classification algorithm is quick and easy, and it takes only one traversal of data to end and get the result. However, this algorithm does not output direct classification rules. Therefore, it is usually used in processing studies to obtain a preliminary result. The information gain rate of a certain attribute is calculated according to the formula (1), respectively. According to the formula (2),

Table 2 Experimental result

	Extracurricular learning time	Effect of classroom learning	Teaching methods	Vocabulary mastered by students
E	0.583	0.433	0.648	0.547
Gain	0.377	0.527	0.312	0.413
Spliti	1.380	1.475	1.290	1.475
Gain Ratio	0.273	0.357	0.242	0.280

Fig. 3 The classification process of decision tree
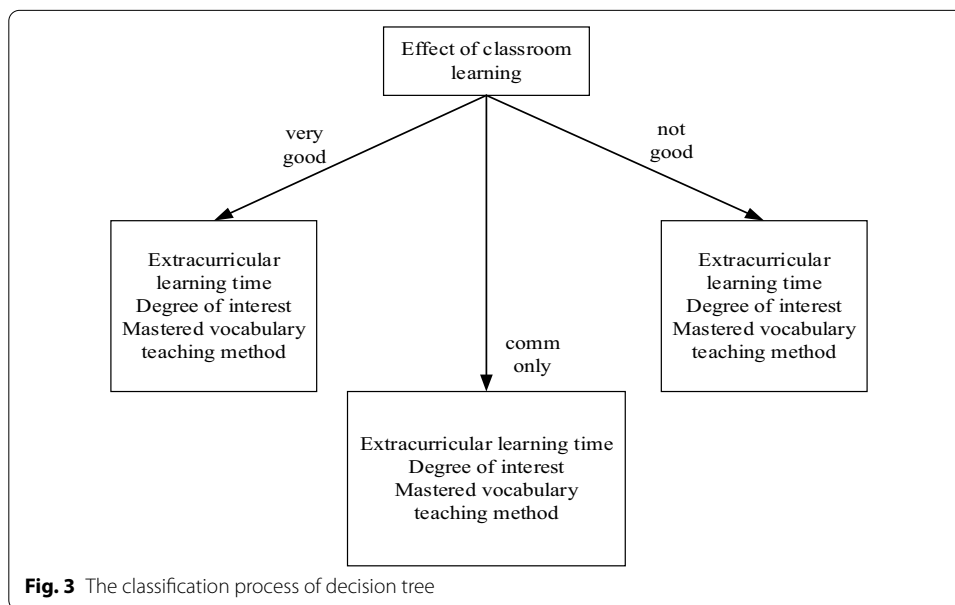
the desired information entropy is calculated for this attribute to be divided into a given sample. According to the formula (3), the information gain of this attribute is calculated. According to the formula (5), the splitting information of this attribute is calculated. According to the formula (4), the information gain rate of this attribute is calculated. The specific results are shown in Table 2.

The third step is to determine the test properties. As the "classroom learning effect" attribute has the highest information gain rate, it is selected as a test attribute. The key decision of the top-down decision tree generation algorithm is the selection of the node attribute values. Different attribute values make the subsets of recorded records different. It affects the growth of the decision tree and the quality of the decision tree structure, thus affecting the rule information. According to the above principles, the following steps are used to construct the decision tree. The first step is to calculate the information entropy required for a given sample classification. A node is created, which is marked with the "classroom learning effect." For each attribute value, a branch is drawn up, which is divided by the sample. The result is shown in Fig. 3.

The fourth step is to use the above method to further divide the branch nodes and use the post-pruning method to prune the final decision tree. The aforementioned

rules extracted from the decision tree are applied to the reality. Teachers have learned a lot of knowledge from them and thus improved their teaching methods. The students' English learning performance will be improved from the current 20%-30% rate to 50%-60%. The result of the pruning of a decision tree is shown in Fig. 4.

Once it is applied to new data, the accuracy drops sharply. To prevent overtraining and reduce training time, methods that can stop the tree from growing at the right time need to be established. Post-pruning is used. The basic idea is to make the decision tree fully grow first and then use the pruning technique to remove the non-general foliage. Generally, it is determined whether to reserve this branch by measuring the improvement of the classification performance of a certain branch, such as the specified classification error rate, and the degree of complexity of the entire tree.

5 Case analysis and testing

Through the research in the above chapters, we analyzed the feasibility and application process of data mining technology in college English teaching. The whole process of data classification and mining has been fully realized. In this chapter, we use clustering experiments on two data sets to verify the application effect of the decision tree model in college English learning. After comparing the results of different clustering algorithms, we found that no matter which kind of data set, the effect of decision tree model is higher. The results can prove that the decision tree model algorithm has good global optimization capabilities. The number of iterations and the results of convergence are shown in Fig. 5.

Determine the purpose and goal of data mining. Online surveys are used to collect data. Data integration, data cleaning, data conversion, data reduction and other preprocessing techniques are used. The decision tree is generated by using the C4.5 algorithm and pruned. The whole process of data classification and mining has been fully realized. In this chapter, we use clustering experiments on two data sets to verify the
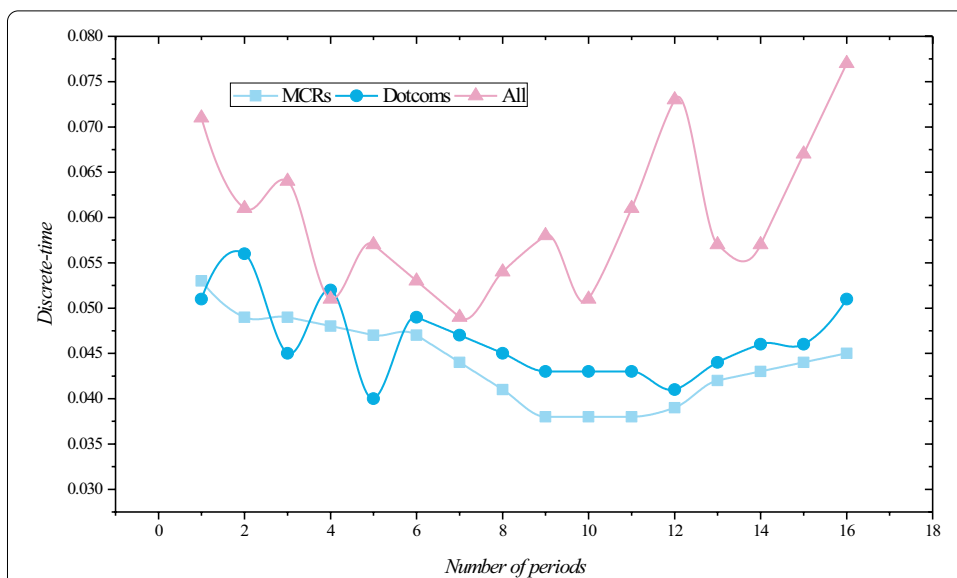

Fig. 4 The result of the pruning of a decision tree
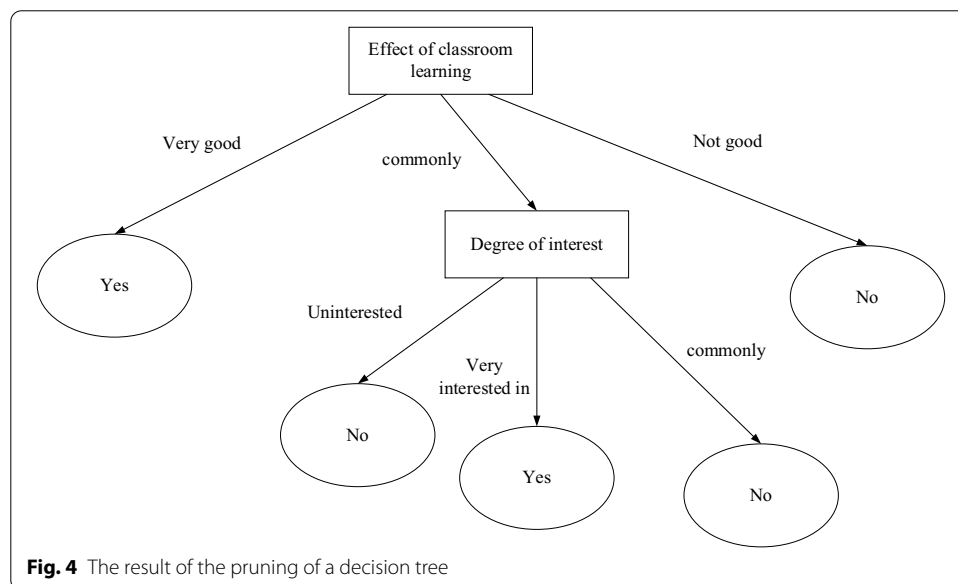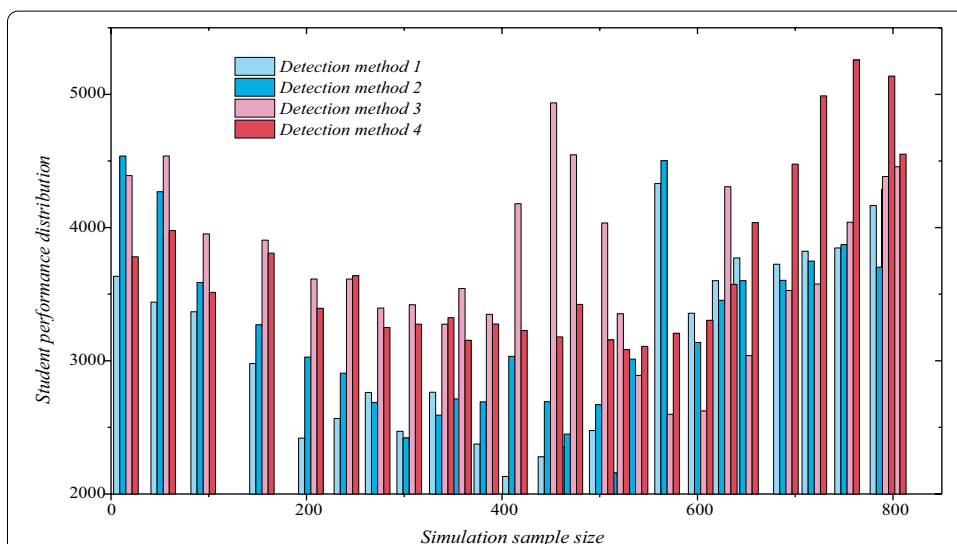
Fig. 5 Test generation times and convergence

Fig. 6 Distribution of the number of students in English learning

application effect of the decision tree model in college English learning. After comparing the results of different clustering algorithms, we found that no matter which kind of data set, the effect of decision tree model is higher. The result analysis decision tree model is completed. A detailed survey was conducted on the students' English learning in the university. The results show that the passing rate of students' English performance has increased from 20–30% to 50–60%. Figure 6 shows the distribution of students' English learning scores.

The data mining realized by this system mainly consists of the user selecting the data source, setting the mining parameters, and then mining the corresponding association rules from the selected data source. First generate frequent item sets, and then generate

association rules to find out the influencing factors of those teachers whose characteristics are excellent or good. The first step is to screen the database and obtain a total of 64 records with a total score of no less than 85, and a total of 29 records with a score between 70 and 84. Set the minimum support degree of 10% and the minimum confidence degree of 5%, and use the improved decision tree algorithm to mine the frequent item sets of the entire database. The next step is to generate association rules. For any set frequent K item sets, first get the corresponding candidate set, and treat it as an association rule A prerequisite for, and then calculate the corresponding confidence. Figure 7 shows the results of mining teaching evaluation data.

A better teaching attitude and teaching methods also bring a better level of teaching quality. Therefore, in class, teachers should consider using teaching methods that are easier for students to accept knowledge. It is important to improve business ability, but while working hard to do it, we must also strive to improve one's own professional accomplishment. In order to comprehensively evaluate teachers' abilities from multiple aspects, a teaching quality evaluation system came into being. It can more accurately reflect whether teachers have undertaken a moderate amount of teaching and research work, so that the school's teaching management department can base on the system's evaluation results. Make appropriate adjustments to more objectively present the advantages and disadvantages of teachers in teaching.

6 Results and discussion

By using the decision tree classification algorithm, the analysis of students' achievement in the teaching research is realized. The application of data mining classification technology in college teaching research is put forward. An online survey system is designed. Then, the basic information base of the students' learning situation is established. The decision tree algorithm is used to establish the decision tree model of student achievement analysis. The pruning of decision tree is realized by the post-pruning method. By using the final decision tree, the classification rules for improving the quality of college English teaching are extracted.

Fig. 7 Partial results of mining teaching evaluation data
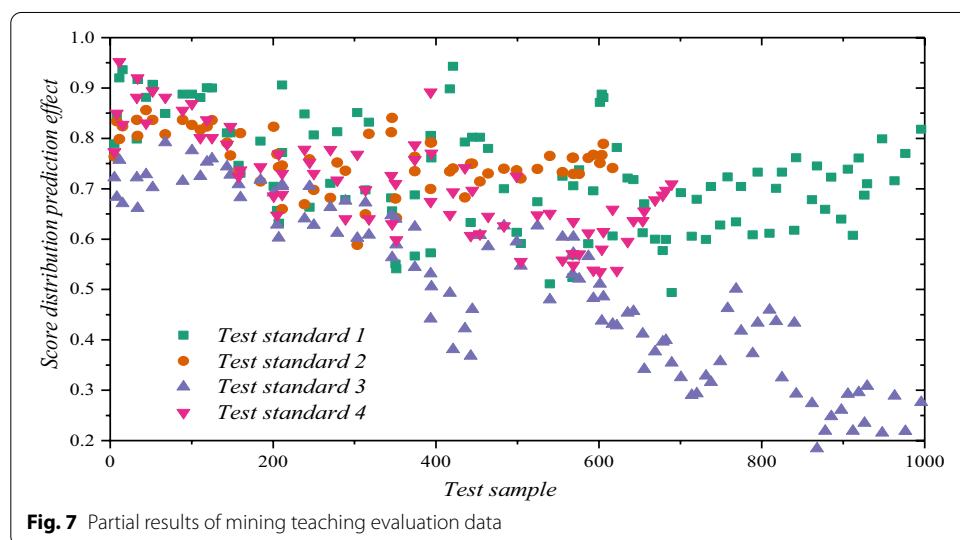
Abbreviation
CRS: Computer reservation system.

Authors' contributions
RG is responsible for the collection of experimental data, and JD is responsible for the writing of the paper. Both authors read and approved the final manuscript.

Availability of data and materials
Data sharing not applicable to this article as no datasets are generated or analyzed during the current study.

Declarations

Ethical approval
This article does not contain any studies with human participants or animals performed by any of the authors. All authors agree to submit this version and claim that no part of this manuscript has been published or submitted elsewhere.

Competing interests
All authors declare that they have no conflict of interest.

Author details
[1]Kunming University of Science and Technology Oxbridge College, Kunming 650106, Yunnan, China. [2]Digital Hollywood University Graduate School, Tokyo 113-0022, Japan.

References

1. W. Zhu, Y. Hou, E. Wang, Y. Wang, Design of geographic information visualization system for marine tourism based on data mining. J. Coast. Res. **103**(sp1), 1034 (2020)
2. Y. Ye, T. Li, D. Adjeroh, S.S. Iyengar, A survey on malware detection using data mining techniques. ACM Comput. Surv. **50**(3), 1–40 (2017)
3. J.B. Varley, A. Miglio, V.A. Ha, M.J.V. Setten, G.M. Rignanese, G. Hautier, High-throughput design of non-oxide p-type transparent conducting materials: data mining, search strategy, and identification of boron phosphide. Chem. Mater. **29**, 2568–2573 (2017)
4. D. Tien Bui, T.C. Ho, B. Pradhan, B.T. Pham, V.H. Nhu, I. Revhaug, GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks. Environ. Earth Sci. **75**(14), 1–22 (2016)
5. K.L. Thompson, A.N. Kuchera, J.N. Yukich, Teaching college writing from a physicist's perspective. Am. J. Phys. **89**(1), 61–66 (2021)
6. K.M. Stawiarski, G.P. Jeyashanmugaraja, G. Bindelglass, G. Lancaster, Utility of a "limited code" status in an inner city community teaching hospital. J. Am. Coll. Cardiol. **75**(11), 3566 (2020)
7. H. Song, N. Gunkelmann, G. Po, S. Sandfeld, Data-mining of dislocation microstructures: concepts for coarse-graining of internal energies. Model. Simul. Mater. Sci. Eng. **29**, 035005 (2021)
8. W. Shi, X. Ke, E.R. Meshot, D.L. Plata, The carbon nanotube formation parameter space: data mining and mechanistic understanding for efficient resource use. Green Chem. **19**(16), 3787–3800 (2017)
9. X. Ren, J. Cui, The development and application of multimedia technology in college gymnastics teaching. J. Test. Eval. **49**(4), 20200196 (2021)
10 Y. Qiang, N.S.-N. Lam, The impact of Hurricane Katrina on urban growth in Louisiana: an analysis using data mining and simulation approaches. Int. J. Geograph. Inf. Sci. **30**, 1832–1852 (2016)
11. S. Qamar, A. Khalique, M.A. Grzegorczyk, On the Bayesian network based data mining framework for the choice of appropriate time scale for regional analysis of drought Hazard. Theor. Appl. Climatol. **143**, 1–19 (2021)
12. L.I. Ming-Fei, L.X. Chen, Application of Literature Analysis Software HistCite in the Teaching for Innovation Training Program of College Students (2016).
13. M.L. Merani, D. Croce, I. Tinnirello, Rings for privacy: an architecture for large scale privacy-preserving data mining. IEEE Trans. Parallel Distrib. Syst. **PP**(99), 1 (2021)
14 S. Madrakhimov, G. Rozikhodjaeva, K. Makharov, The use of data mining methods for estimating of vascular aging. Atherosclerosis **315**, e135 (2020)
15. A.O. Luna, M.A. Simmons, S. Abraham, R. Karnik, Beyond see one, do one, teach on, the effect of targeted echo teaching labs on the quality of first year pediatric cardiology echoes. J. Am. Coll. Cardiol. **75**(11), 3653 (2020)

16. T. Lorberbaum, K.J. Sampson, J.B. Chang, V. Iyer, R.L. Woosley, R.S. Kass, N.P. Tatonetti, Coupling data mining and laboratory experiments to discover drug interactions causing QT prolongation. J. Am. Coll. Cardiol. **68**(16), 1756–1764 (2016)

17. H. Li, X. Bu, X. Liu, X. Li, Q. Lyu, Evaluation and prediction of blast furnace status based on big data platform of ironmaking and data mining. ISIJ Int. **61**, 108–118 (2020)

18. W. Kang, E.-K. Jang, C.-Y. Yang, P.Y. Julien, Geospatial analysis and model development for specific degradation in South Korea using model tree data mining. CATENA **200**, 105142 (2021)

19. S.R. Joseph, H. Hlomani, K. Letsholo, Data mining algorithms: an overview. Neuroscience **12**(3), 719–743 (2016)

20. A. Jain, G. Hautier, S.P. Ong, K. Persson, New opportunities for materials informatics: resources and data mining techniques for uncovering hidden relationships. J. Mater. Res. **31**(8), 977–994 (2016)

21 H. Hong, H.R. Pourghasemi, Z.S. Pourtaghi, Landslide susceptibility assessment in Lianhua County (China): a comparison between a random forest data mining technique and bivariate and multivariate statistical models. Geomorphology **259**(Apr. 15), 105–118 (2016)

22 M. Giles, Foody, Uncertainty, knowledge discovery and data mining in GIS. Prog. Phys. Geogr. **27**(1), 113–121 (2016)

23. R.M. Geilhufe, A. Bouhon, S.S. Borysov, A.V. Balatsky, Three-dimensional organic Dirac-line materials due to nonsymmorphic symmetry: a data mining approach. Phys. Rev. B **95**(4), 041103 (2017)

24 S. Garcia, J. Luengo, F. Herrera, Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. Knowl. Based Syst. **98**(Apr. 15), 1–29 (2016)

25 A.E.A. Elrazek, M. Amer, B. Hawary, A. Salah, A.S. Bhagavathula, M. Al-Boraie, S. Saab, Prediction of HCV vertical transmission: what factors should be optimized using data mining computational analysis. Liver Int. **37**, 529–533 (2016)

26 X. Du, H. Xu, F. Zhu, A data mining method for structure design with uncertainty in design variables. Comput. Struct. **244**, 106457 (2021)

27 B. Chabalenge, S. Korde, A.L. Kelly, D. Neagu, A. Paradkar, Understanding matrix assisted continuous cocrystallisation using data mining approach in quality by design (QbD). Cryst. Growth Des. **20**, 4540–4549 (2020)

28. L. Carmichael, S. Stalla-Bourdillon, S. Staab, Data mining and automated discrimination: a mixed legal/technical perspective. IEEE Intell. Syst. **31**(6), 51–55 (2016)

29 Y. Cadavid, C. Echeverri-Uribe, C.C. Mejía, A. Amell, J.A.M. Ospina, Analysis of potential energy savings in a rotary dryer for clay drying using data mining techniques. Dry. Technol. (2021). https://doi.org/10.1080/07373937.2021.1872610

30. Y. Yang, N. Xiong, N.Y. Chong, X. Défago, A decentralized and adaptive flocking algorithm for autonomous mobile robots, in *The 3rd International Conference on Grid and Pervasive Computing*, (2008).

31. A. Shahzad, M. Lee, Y.K. Lee, S. Kim, N. Xiong, J.Y. Choi, Y. Cho, Real time MODBUS transmissions and cryptography security designs and enhancements of protocol sensitive information. Symmetry **7**(3), 1176–1210 (2015)

32. Q. Zhang, C. Zhou, N. Xiong, Y. Qin, X. Li, S. Huang, Multimodel-based incident prediction and risk assessment in dynamic cybersecurity protection for industrial control systems. IEEE Trans. Syst. Man Cybern.: Syst. **46**(10), 1429–1444 (2015)

33. K. Huang, Q. Zhang, C. Zhou, N. Xiong, Y. Qin, An efficient intrusion detection approach for visual sensor networks based on traffic pattern learning. IEEE Trans. Syst. Man Cybern. Syst. **47**(10), 2704–2713 (2017)

34. W. Wu, N. Xiong, C. Wu, Improved clustering algorithm based on energy consumption in wireless sensor networks. IET Netw. **6**(3), 47–53 (2017)

35. J. Sun, X. Wang, N. Xiong, J. Shao, Learning sparse representation with variational auto-encoder for anomaly detection. IEEE Access **6**, 33353–33361 (2018)

Publisher's Note