Journal of Computers Vol. 28, No. 1, 2017, pp. 1-12
doi:10.3966/199115592017022801001

Application of Fuzzy Association Rules in the Analysis on Higher Vocational College Students' Performance

Yan-Rui Lei[1], Li Lei[2], and Lai-Quan Liu[3*]

[1] Department of Network Engineering, Hainan College of Software Technology,
No. 128, Fuhai Road Qionghai, Hainan, China
leiyanrui@139.com

[2] Department of Software Engineering, Chongqing University of Arts and Sciences,
No. 598, Huachuang Road, Yongchuan, Chongqing, China
657546253@qq.com

[3] Department of Digital Design, Hainan College of Software Technology,
No. 128, Fuhai Road, Qionghai, Hainan, China
liulaiquan@139.com

Received 21 May 2015; Revised 14 July 2015; Accepted 6 April 2016

Abstract. Data mining technology can be used to discover useful knowledge from a large number of seemingly unrelated data. It can enhance the simple data query function to mining knowledge from data and thus provide a high level of decision-making reference for the decision makers. The data mining techniques provide an opportunity for the analysis of the results. By using the data mining techniques, we can conduct effective mining to large amount of achievement data and thus extract valuable information to provide guidance for teachers' targeted teaching. Mining algorithm of fuzzy association rules is to introduce fuzzy set theory to associate rule mining to improve the disadvantage of edge data loss caused by traditional association rule analysis, and to expand the membership degree of factors in set to characteristic function from original {0,1} to [0,1] through membership function, then to mine out the hidden "knowledge". This paper regards it as reference, then applies it to performance analysis, and employs the obtained association rule to guide the teaching of teachers and the learning of students, then to serve for the employment competitiveness improvement of students.

Keywords: application, association rules, data mining, research

1 Introduction

Improving the quality of education has been paid attention by all colleges and universities. Whereas, one of the important basis for the evaluation of teaching quality is the achievement of student, which is a crucial symbol to evaluate the knowledge and skills acquired by students as well as basis and assessment of teachers' teaching effectiveness check. But with ascension of informatization management in colleges and universities, a large number of student achievements have been accumulated, and the rise of data mining can work out the discovery of valuable knowledge from these massive data [1]. It is of great significance to apply data mining knowledge to students' achievement assessment: the real cause of affecting fluctuation trend in students achievement can be analyzed from students achievements, and valuable information including the implicit relationship between curriculum and curriculum, students' learning interest, *etc.*, can also be analyzed, and these meaningful information exhumed can be applied in the service of education reform in colleges and universities, and be taken as the reference to further enhance the quality of teaching, the competitiveness and so on.

* Corresponding Author

On October 28, 2005, the State Council issued The State Council's Decision on Digorously Developing the Vocational Education, and on November 7th and 8th the same year, the National Vocational Education Working Conference was convened under the State Council. At the meeting, the State Council put forward decisions of "paying strategic importance to vocational education" and "vigorously developing vocational education". On June 23, 2014, the State Council convened the National Vocational Education Working Conference, proposing to "accelerate the development of modern vocational education", thus higher vocational education became a hit concerned by educational circles once again. Under this opportunity and circumstance, how to keep pace with the times and to take effective teaching measures to ensure the training objective of higher vocational education, which became one of the most important tasks of higher vocational education.

Subsequently, numerous higher vocational colleges have implemented teaching reform, applying Germany's action-oriented teaching into various aspects of education. During the implementation process of action-oriented teaching, students' performances recorded for each course are no longer a simple collection of grades on final exam, daily performance and attendance, but instead, grades in each experiment and practical training of the course are all recorded, thus abundant data of grades has been accumulated. At present, the processing of these data of grades only includes storage, reference and back-up, and hidden information has not been unearthed from a large amount of data, and transformed into usable "knowledge". Action-oriented teaching process highlights the learning in the field of career action. Teaching of learning field has completely shredded the discipline-based teaching system in the teaching of vocational training. In this teaching process, knowledge are combined as required by areas of activity to develop a teaching plan in the learning field, and conduct learning according to the requirements proposed by teaching of learning field [2]. In this process, the results of students in a course come from the sum of the results of learning field in this course, and any analysis has not been made whether the learning between fields of study will influence each other. Numerous analyses of performance based on data mining have also been made [3-5], but few studies have been made in linkages between experiments of the course, total scores of experiment and the course, and between grades of practical trainings. In view of this, this paper, based on the technology of fuzzy association rules in data mining, has made in-depth mining and study on the linkages among experimental results of the course, total score and grades in practical training to dig out a potential relationship, and to put the research results into practice, which provides more helpful information for administrators and implementor of teaching.

2 Data Mining

Data mining (DM) [6] is a process of automatically discovering useful knowledge in the existing mass of data. Data mining technique is used to discover some useful models which are previously unknown for the people in large-scale databases, also an important step in the process of knowledge discovery in databases. DM also has the ability to predict future observed results.

Currently, relevant researches have been conducted in many fields by data mining, including Security Company, bank system, education and teaching. Data mining has indeed become the hot issue of computer science, which will be researched and applied by many scientists for a long time in the future. The appearance of data mining is not an accident, but a long-term process. At the early period, scientists pay all attention to the learning of computer. Computer learning is also called machine learning whose process is to input the known or solved problems into computer as samples and generate the rules adapted to the samples after learning the samples and concluding, using these rules to solve similar problems. Later, with the appearance of computer being applied to neural network technique, people gradually research the knowledge discovery. The so-called knowledge discovery is to input coding rules into computer, through which to solve the corresponding problems. This method is abandoned due to the huge input and non-ideal effect. After then, under the common guidance of computer knowledge and neural network theory, machine learning is adopted again, which is also applied to business competition successfully. At the end of 1980s, a new terminology- KDD appeared which was discussed quickly and gained common recognition. Since then, KDD has been used to express the process of data mining, and the sub-process of data mining is expressed by mining algorithm.

To date, research on data mining and knowledge discovery has made some progress, and important algorithms are mainly association rule, clustering algorithm, decision tree, and so on. Association analysis discovers association rules. These rules demonstrate potential connections among data items of set data-set and have been widely applied in shopping basket or transaction data analysis; while clustering is to divide data object into various classifications or clusters. The objects within the same cluster have relatively high similarity while those in different clusters have more differences with unclear cluster partition. Decision tree algorithm is a discrete function value approximation method [3]. It is a typical classification method and firstly makes data processing to generate readable rules and decision trees through induction algorithm and then to use decision to conduct new data analysis. Some international companies engaged in related research have begun developing data mining systems used in commercial fields, like BehaviorScan, MDT, Stock Selector, AI (Automated Investor), Clonedetector Falcon, and Fais [7].

Data mining technology starts relatively late in China, and interdisciplinary interaction has not been formed. Research projects of data mining are mainly funded by the government, namely National Natural Science Fund and “the 12th Five-Year Plan” program. And most of them are concentrated in universities and research institutes. China’s success in business decisions by using data mining technology is relatively rare, so the research of data mining technology and development have great potential and broad prospects in China.

3 The Application of the Data Mining Technology in the Field of Education

Data mining technology has achieved abundant results so far and been successfully applied to various fields. NBA once applied data mining analysis offered by a company and won the game. Moreover, domestic telecommunications industry implements data mining technology to dig out the potential customers to increase sales, *etc*. In conclusion, data mining technology can be used as decision-making guidance in all walks of life and it profoundly contributes to the economic development and decision-making of enterprises [8].

In recent years, Chinese higher education has been developing rapidly, and the vocational colleges are on the rise as well. Under the context of rapid-developing higher vocational education, how to ensure the teaching quality and cultivate qualified practical senior talents for the society is one of the important subjects and the fundamental starting point for the teaching activities in higher vocational colleges. So far, there are various evaluation methods towards the education and teaching quality. For example, data mining is applied to academic score analysis. And then, through certain analysis, the factors that have close relationship with students’ examination scores are found out from many factors, including educational management, curriculum setting, teaching management, etc. With such analysis and feedback, it can provide important information feedback for education and teaching activity; at the same time, it can provide basic information and support for the decision-makers of the management department in higher vocational colleges. In the field of education, the main function of data mining is to extract data in the database, transform, and then dig out the potential rules, which can be used as the education decision-making reference. Data mining offers valuable reference information for the teaching decision and improves students’ ability. Data mining technology is gradually becoming mature, and its application range is expanding. It is introduced to teaching and service for teaching and management by many colleges and universities. For example, apply the data mining to the analysis as shown below (see Fig. 1), which play a good role in improving teaching management level of higher education, the ability of students and other aspects [9].

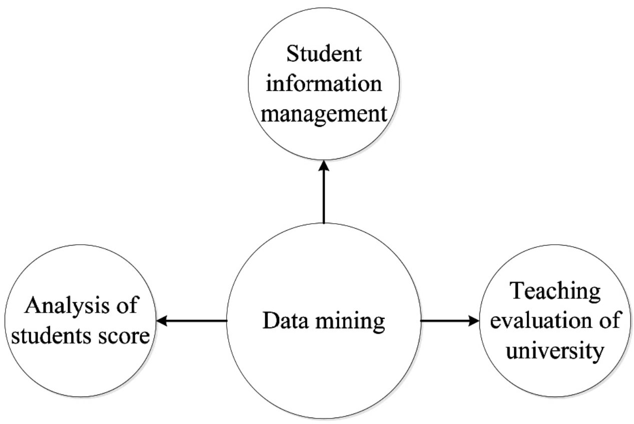

Fig. 1. Application of data mining

At present, applications of data mining in education are embodied in the following aspects (see Table 1) [10].

Table 1. Applications of data mining in education

Application areas	Contents	Objectives
Achievement analysis	Find out the main factor influencing students' test scores by decision tree algorithm of student's performance and research data mining, such as interest, extra-curricular learning practice or teaching methods, *etc.*	Make reference to students' learning.
	Utilize the attribute importance analysis method of rough set theory in the score analysis, and find out which part of the examination largely impacts on the overall score.	Offer reference for the course teaching.
	Association rules algorithm is used to mine the course grade analysis and find out the courses which made students repeat easily, have no degree and drop out.	Forewarn students.
Student management [11]	Look into students' personal information, course information, grades, employment situation after graduation and activities information such as of web-surfing and book-borrowing and find out the potential relation.	Benefit student management in colleges.
Question bank	Use clustering method to research data in question bank and find out the information or association hidden behind data.	Contribute to the reasonable organization of question bank and overall evaluation of the comprehensive quality of teachers' teaching.
Teaching quality evaluation	Based on the evaluation of students, teachers and supervisors, using ID3, C4.5 algorithm to construct the decision-making tree, analyze the influencing factors of teaching quality evaluation, and find out the characteristics of the popular teachers, *etc.*	Promote teaching.
Distance learning system	Analyze log information in remote education by data mining technology, and find out how to improve the education quality and service level.	Make reference to teachers, students, and administrator in remote teaching.

4 Fuzzy Association Rules

In 1895, the German mathematician G. Cantor founded set theory; the crucial idea of set theory is the principle of comprehension: a property P is given, and for any objects able to satisfy property P, which are pooled together by the objects only with the property P to form a set, expressed as:

$$A=\{a|P(a)\} \tag{1}$$

Where, a represents any element in A, and P (a) indicates a with property P [12]. It shall not be ambiguous but differentiated for property P, that is any object has no property P or has property P. In this theory, the description of the properties of things is usually expressed by "True" and "False". If an element has a certain property of P, it can be referred to as a "1" or "True", otherwise referred to as "0" or "False". However, in the objective world, not all things can be expressed by the "True or False". For example, for people at the age of 33, they may belong to young people in terms of some features, i.e. about 55%; while in terms of other characteristics, they may belong to the middle-aged, i.e. about 45%.

On the basis of the above theory, Zadeh proposed the concept of fuzzy sets in 1965. Using the membership function to represent the degree of the object subordinate to the set, he indicated the membership degree of elements of the set to characteristic function from the previous {0,1} expanded to [0,1], thus the concept of fuzzy set theory was formed. For a long time after that, he has been engaged in research on fuzzy sets, and has raised the possibility theory of being complementary with fuzzy sets. Since then, the disciplinary research on fuzzy phenomenon has identified a certain status in research community, and been applied and promoted in many dimensions.

4.1 Fuzzy Set Theory

To make comparison between the classical theory and fuzzy theory, the distinction between fuzzy membership function and ordinary sets is illustrated through comparison as follows [12].

Definition 1: characteristic function of set A: $A \in P(U)$, and defining mapping: μA: $U \to \{0,1\}$, among it, $\mu_A(U) = \begin{cases} 1(u \in A) \\ 0(u \notin A) \end{cases}$. Then, the function μ_A becomes the characteristic function of set A.

Fuzzy sets are described by their membership function.

Definition 2: membership function of fuzzy sets: mapping setting on domain U: A: μA: $U \to [0,1]]$, namely $u \to \mu A(u)$. Then mapping μA has defined a fuzzy set A on U, and the function μA (u) is the membership function of A; μA (u) is the membership of A, i.e. membership degree of u to A.

Specification by example:

Assuming that domain U (see Fig. 2) takes the line segment with unit length, denoting fuzzy sets on U as A, if the element x (segment) is located inside the A (circle), it is denoted by 1; if the element x is located outside A, it is denoted by 0; if an element is in both interior and exterior of A, then it is in the "intermediary state" of membership, and the membership degree of this element to A is its length inside A.

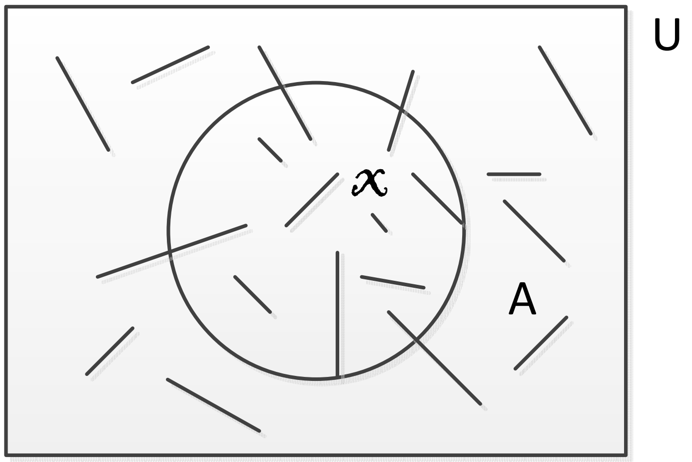

Fig. 2. Membership function instance figure

When μA takes [0, 1] in {0, 1}, fuzzy set A becomes an ordinary set A, and therefore, the ordinary set is a special case of fuzzy sets. This article abbreviates the degree of membership as A (u).

Traditional association rule algorithm is to mine plenty of association rules in the analysis of classical "shopping basket", then use certain concept to express the connection among things. The rules are generated through the analysis on these data, which also provides decision-making support to managers. Association rule mining model is shown below (see Fig. 3).

But in the real world, not all things can be described as the state of "either this or that", which not only contains plenty of Boolean or Category data, but also contains the continuous discrete data and discrete data with larger numerical range.

During the mining process of traditional association rules, it will transfer this type of data into Boolean, to rigidly divide the data, which is simple and fast, but the divided attribute set association rules can not express the fuzzy phenomenon in the reality. Also, the rigid division of data may cause the loss of edge data, so the mined association rules do not have great significance.

For example, during the analysis of "shopping basket", besides the items, the database also reserves the numerical data related to these items, such as the quantity and price of commodities. At the early period when association rule was put forward, it only paid attention to the membership information rather than numerical value. But some knowledge and information may be hidden in these numeric data, which will provide assistance to association discovery. Because the perspectives that people understand thing are general uncertain, it adopts the fuzzy set in attribute field to soften the periphery of attribute value. In the N divided clusters, it can provide smooth transition interval to the factors which may belong to or not

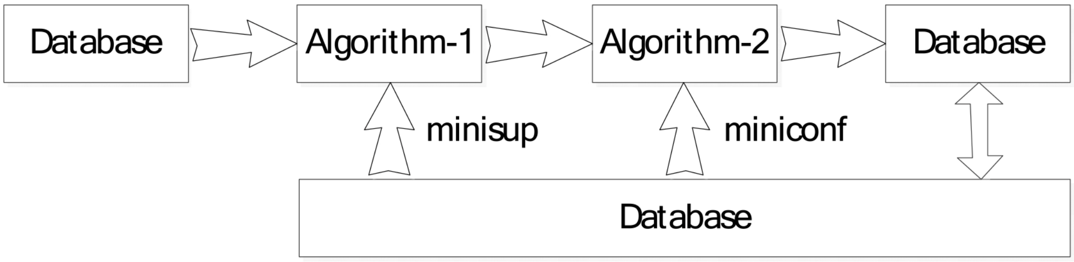

Fig. 3. Model of association rules mining

belong to a certain category. Then it mines the association rules that people are interested in from database, in order to describe by the semantic expression understood by people, which more conforms to the inference and thinking habit of human. The fuzziness of fuzzy association rule not only requires the fuzziness of fuzzy concept, but also requires the fuzziness of membership function. During the conversion process, different results will be obtained by using different membership functions, which will greatly impact the mining results.

4.2 Related Concepts of Fuzzy Association Rules

Some basic concepts of fuzzy association rules [13]:

Assumed I= $\{I_1, I_2, \ldots I_m\}$ is an attribute set of database D, and for each property I_i ($1 \le i \le m$), they can be divided into q_j fuzzy attributes with fuzzy subjection function. After the original numerical attributes are divided into fuzzy attributes, database D is transformed into fuzzy database D_f, whose attribute set is $I_f=\{ I_1^1, I_1^2, \ldots I_1^{q_1}, I_2^1, I_2^2, \ldots, \ldots, I_m^1, I_m^2, \ldots I_m^{q_m} \}$, and range extensions for all new properties are [0,1].

Definition 3: support degree of record i in fuzzy item-set $X=\{x_1, x_1, \ldots, x_p\}I_f$ to fuzzy item-set x is defined as follows:

$$\mathrm{SupT}_i(X)=x_{1i} \wedge x_{2i} \wedge \ldots \wedge x_{pi} \text{ or } \mathrm{SupT}_i(X)=x_{1i} \times x_{2i} \times \ldots \times x_{pi}\ . \tag{2}$$

Where, x_{ji} is the recorded value of i in fuzzy item x_{ji}, $x_{ji}[0,1]$ (i=1, 2, ..., n; j=1, 2, ..., p).

Definition 4: support degree of entire data set of fuzzy set $X=\{x_1, x_1, \ldots, x_p\}I_f$ to X is defined as follows:

$$\sup(X)=\frac{\sum_{i=1}^{n} \sup T_i(X)}{|D_f|} \tag{3}$$

In the above formula, $|D_f|$ is the number of affair in the database, if support degree of fuzzy item-set is not less than the given minimum fuzzy support degree minsup, X is the fuzzy frequent item-set.

Definition 3: For implicative expression $X \Rightarrow Y$in fuzzy association rules, similar to the Boolean association rules, X is called the antecedent of fuzzy association rules, while Y is called the consequent of fuzzy association rules. Similarly, in $X \subseteq I_f$, $Y \subseteq I_f$, and $X \neq \phi$, $Y \neq \phi$, $X \cap Y \neq \phi$, $I=X \cup Y$, there is no continuous item from the same property.

Definition 5: Implicative expression $X \Rightarrow Y$, in fuzzy association rules, sup ($X \Rightarrow Y$) and conf ($X \Rightarrow Y$) are defined as:

$$\sup(X \Rightarrow Y)=\sup(X \cup Y) \tag{4}$$

$$\mathrm{conf}(X \Rightarrow Y)=\frac{\sup(X \cup Y)}{\sup(X)} \tag{5}$$

As in the above case, in $X \subseteq I_f$, $Y \subseteq I_f$, and $X \neq \phi$, $Y \neq \phi$, $X \cap Y \neq \phi$, $I=X \cup Y$, there is no continuous item from the same property.

Similar to the Boolean association rules, policymakers set minimum support degree minsup and minimum confidence minconf in advance, and fuzzy Association rules are mined.

4.3 Mining Steps of Fuzzy Association Rules

Mining process [14, 15] of fuzzy association rules is as shown in Fig. 4.

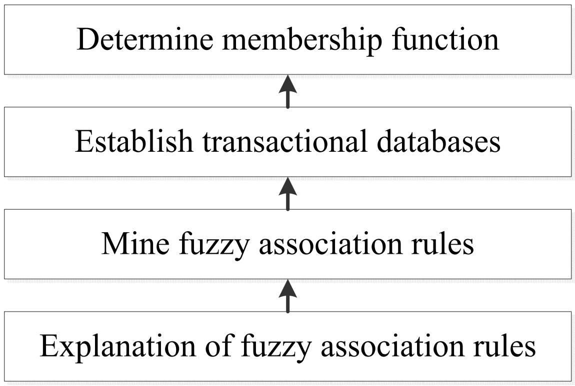

Fig. 4. Mining process of fuzzy association rules

4.4 Principles and Flaws of the Fuzzy C-Means Clustering Algorithm

The fuzzy C-means algorithm performs data clustering by optimizing the objective function [16]. In its clustering results, each sampled data has a membership level with respect to a cluster center, which is in the range [0, 1]. The algorithm uses the MSE approximation theory to construct a constrained nonlinear programming function, and mostly uses the objective function as an approach to clustering. WGSS and J_1 are the general forms of the cluster's objective function.

The fuzzy C-means algorithm is able to automatically classify sampled data, and assign the samples to classes by optimizing the standard function J . Then, the membership level of the sample data with respect to the center of its classes can be obtained based on the optimized standard function. And J denotes the quadratic sum of the error between the sampled data and the center of the class it belongs to.

It is a partitioning algorithm, the ultimate goal of which is to minimize the quadratic sum of the distance between the samples of each class and the centers of the clusters.

The membership U_{ij} and the cluster center V_i are defined as:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{dkj}\right)^{\frac{2}{m-1}}} \tag{6}$$

$$v_i = \frac{\sum_{j=1}^{n} (u_{ij})^m x_j}{\sum_{j=1}^{n} (u_{ij})^m} \tag{7}$$

The fuzzy C-means clustering algorithm is a process of simple iterations and dynamic clustering. Its flow chart is shown in Fig. 5.

Execution process of fuzzy C-mean value clustering algorithm:

1. Initialization: vector matrix of clustering center $V^{(0)}$, the fuzzy weighted index number m=2, clustering class number is $c(2 \le c \le n)$, iteration termination threshold $\varepsilon =$ e, initial value of iteration b=0, the maximum iterations is $b_{max.}$

2. Update divided matrix $U^{(b)}$, form the new divided matrix through calculating the membership degree value in formula (6).

As for formula (6), $\forall i, j$, if $d_{ij}^{(b)} > 0$, calculating the value of $u_{ij}^{(b)}$ through formula (6);

If $d_{ij}^{(b)} = 0$, then $u_{ij}^{(b)} = 1$, while $j \ne k$, $u_{ik}^{(b)} = 0$.

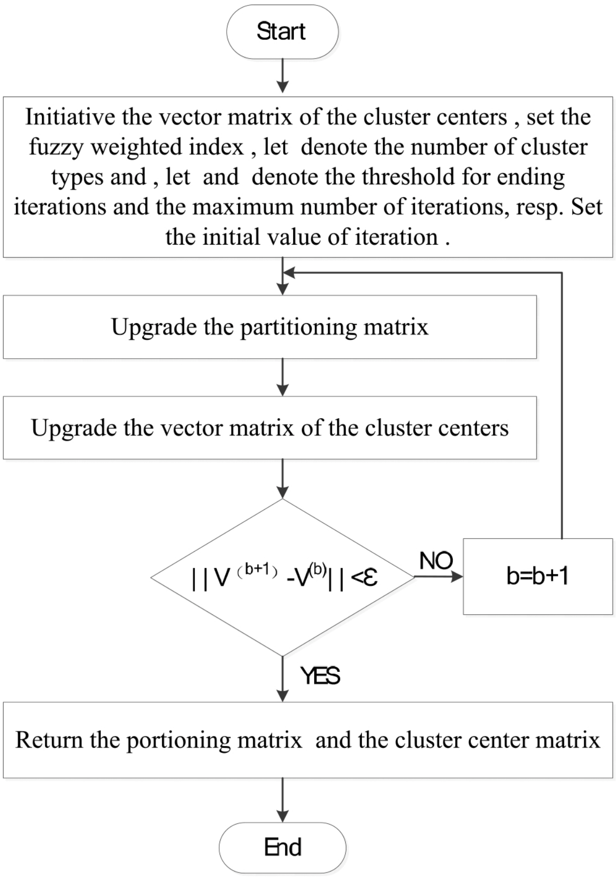

Fig. 5. Flow charts of the fuzzy C-means clustering algorithm

3. Update the clustering center $V^{(b+1)}$. $\forall i, j$, if $\exists d_{ij}^{(b)} > 0$, then:

$$v_i^{(b+1)} = \frac{\sum_{j=1}^{n} (u_{ij}{}^{m})^{(b+1)} x_j}{\sum_{j=1}^{n} (u_{ij}{}^{m})^{(b+1)}} \tag{8}$$

4. if $\left\| V^{(b+1)} - V^{(b)} \right\| < \varepsilon$ establishes, then the algorithm ends and outputs divided matrix U and clustering center matrix V; if $\left\| V^{(b+1)} - V^{(b)} \right\| < \varepsilon$ does not establish, make b=b+1 and return to the step 2.

The above execution process is a dynamic clustering process.

5 Application of Fuzzy Association Rules in the Analysis on Higher Vocational College Students' Performance

5.1 Data Acquisition and Preprocessing

This application research, based on experiment scores, total points of courses and grades in practical training of students in 2011 in a certain vocational technical institute in Hainan (4 classes, a total of 152 students), has mined out potential effects among all experiments of courses and potential effects of all experiments on integrated practical training of courses, and has provided references for teachers' teaching using mining results, thus improving students' practical ability. Data in the study process is derived from the database of teaching management system of a certain vocational technical institute in Hainan. Detailed information of the course experiment and data on the performance is as shown in Table 2 and Table 3.

Table 2. Detailed information of students' performance

Serial number	Experimental contents	Percentage
Experiment 1	Linux system installation, configuration, and installation of common software	10%
Experiment 2	LAN construction	15%
Experiment 3	Software package management system, user management	20%
Experiment 4	Maintenance and troubleshooting of Grub and password	10%
Experiment 5	Data backup and configuration of automatic login environment of KDE	20%
Experiment 6	System trimming, start testing and debugging	25%

Table 3. Detailed information of students' performance

Serial number	Experiment 1 (S1)	Experiment 2 (S2)	Experiment 3 (S3)	Experiment 4 (S4)	Experiment 5 (S5)	Experiment 6 (S6)	Total grades of course (SUM)	Training perform-ance (TRA)
1	7	12	15	9	17	15	75	85
2	10	13	20	10	20	20	93	95
3	3	10	0	8	18	5	44	60
4	3	10	12	10	16	22	42	80
5	8	11	13	10	15	14	73	75
…	…	…	…	…	…	…	…	…
149	8	10	15	8	12	22	75	80
150	10	15	20	10	20	25	100	100
151	8	10	14	9	12	24	77	77
152	8	9	13	8	10	20	68	68

The data format of information of students' performance is stringent in the original system, with no need to clean up, but for some special records, such as records of students' retake classes; they need to be converted [17].

In the mining process, method of data cleansing adopted is as follows:

(1) The duplicate records are removed, and for results of student's retake classes, only the grade in the first exam of this course is kept.

(2) If student's performance is missing due to absence, it can be processed according to performance of make-up examination.

(3) If student neither has performance in initial exam nor in make-up examination, the student's performance is not included in the process of mining.

The purpose of this paper is to apply fuzzy association rules in data mining of students' performance in process evaluation of action-oriented teaching. As to data conversion, the results are sorted by using optimal fuzzy c-means clustering algorithm, and then they are transformed into fuzzy membership values.

5.2 Mining Process and Discovery of Association Rules

1. Similarly, the minimum supporting number of minsup=0.4, minconf=0.7 is set.

2. Cluster centers (see Table 4) and fuzzy database are calculated based on fuzzy c-means clustering algorithm [18].

Table 4. Cluster centers of all properties

Property	Cluster centers		
	1	2	3
S1	2	6	10
S2	9	12	15
S3	2	13	20
S4	5	8	10
S5	11	16	20
S6	4	16	22
SUM	43	73	96
TRA	65	80	96

3. Potential Count of all fuzzy items is calculated, with results shown in Table 5.

Table 5. Fuzzy database after conversion of experimental grades

No.		1	2	3	4	5	…	149	150	151	152	Count
	Low	0	0	1	0.76	0	…	0	0	0	0	29.6
S1	Mid-dle	0.75	0	0	0.24	0.5	…	1	0	0.5	0.26	54.6
	High	0.25	1	0	0	0.5	…	0	1	0.5	0.74	67
	Low	0	0	1	1	0.35	…	0.71	0	0.86	1	82.7
S2	Mid-dle	1	0.67	0	0	0.65	…	0.39	0	0.14	0	47.88
	High	0	0.33	0	0	0	…	0	1	0	0	22.3
	Low	0	0	1	0.18	0	…	0	0	0.09	0	21.3
S3	Mid-dle	0.71	0	0	0.82	1	…	0.71	0	0.91	1	86.5
	High	0.29	1	0		0	…	0.29	1	0	0	43.3
	Low	0	0	0	1	0	…	1	0	0	0	33.6
S4	Mid-dle	0.5	0	1	0	0	…	0	0	0	1	42
	High	0.5	1	0	0	1	…	0	1	1	0	75.6
	Low	0	0	0	0.8	0.2	…	0	0	0	1	33.6
S5	Mid-dle	0.75	0	0.5	0.2	0.8	…	0.5	0	1	0	63
	High	0.25	1	0.5	0	0	…	0.5	1	0	0	54.6
	Low	0.08	0	0.92	1	0.17	…	0	0	0	0	36.5
S6	Mid-dle	0.92	0.33	0.08	0	0.83		0.17	0	0.17	0.67	53.3
	High	0	0.67	0	0	0	…	0.83	1	0.83	0.33	61.5
	Low	0	0	0.97	1	0.03	…	0	0	0	0.17	36.5
SUM	Mid-dle	0.91	0.13	0.03	0	0.97	…	0.91	0	0.03	0.83	64
	High	0.09	0.87	0	0	0	…	0.09	1	0.97	0	50.7
	Low	0	0	1	0	0.33	…	0	0	0.2	0.8	39.1
TRA	Middle	0.69	0.06	0	1	0.67	…	1	0	0.8	0.2	74.3
	High	0.31	0.94	0	0	0	…	0	1	0	0	37.8

4. The fuzzy items mentioned above with potential no less than 40 are classified into frequent 1-item-set L_1 , L_1 = {S1. high S2. low, S3.middle, S4.high, S5.middle, S6. high, SUN. middle, TRA. middle}.

5. C_2 is generated when the above L_1 is connected, and the fuzzy item of the same property is not connected. Thus it can be obtained ={{S1. high, S2. low},{S1. high, S3. middle},{S1. high, S4. high},{S1. high, S5. middle},{S1. high, S6. high},{S1. high, SUM. middle},{S1. high, TRA. middle},{S2. low, S3. middle}, {S2. low, S4. high},{S2. low, S5. middle},{S2. low, S6. high},{S2. low, SUM. middle},{S2. low, TRA. middle}, {S3. middle, S4. high}, {S3. middle, S5. middle}, {S3. middle, S6. high}, {S3. middle, SUM. middle}, {S3.middle, TRA. middle}, {S4. high, e. middle},{S4. high, S6. high},{S4. high, SUM. middle},{S4. high, TRA. middle},{S5. middle, S6. high},{S5. middle, SUN. middle},{S5. middle, TRA. middle},{S6. high, SUM. middle},{ S6. high, TRA. middle},{SUM. middle, TRA. middle}}.

6. As in above C_2, such as {S1. low, S2. high}, it is explained that if the experiment S1 is low, then the experiment S2 is high, and such a conclusion apparently cannot meet reality. Therefore after removal of these unpractical cases, it can be got that C_2={{S1. high, S4. high}, {S1. high, S6. high},{S1. high, SUM. middle},{S1. high, TRA. middle}, {S3. middle, S5. middle}, {S3. middle, SUM. middle}, {S3. middle, TRA. middle},{ S4. high S5. high},{S4. high, SUM. middle},{S4. high, TRA. mid-dle},{S5.middle, SUM. middle},{S5.middle, TRA. middle},{S6. high, SUM. middle},{S6.high, TRA. middle},{SUM. middle, TRA. middle}}.

7. Potentials of candidate fuzzy items are calculated, such as the Count ({A. tall, D. high}) =0.25*0.5+1*1+0.5*1+...+1*1+0.5*1+0.74*0=52.77. Similarly, Count ({S1. high, S6. high}) =39.33, Count ({S1. high, SUN. middle}) =24.86, Count ({S1. high, TRA. middle}) =18.8, Count ({S3. middle, S5. middle}) =46.6, Count ({S3. mid-dle, SUM. middle}) =52.7, Count ({S3. middle, TRA. middle}) =61.1, Count

({ S4. high, S6. high}) =42.2, Count ({S4. high, SUM. middle}) =26.8, Count ({S4. high, TRA. middle}) =31.7, Count ({S5. middle, SUM. middle}) =33.07, Count ({S5. middle, TRA. middle}) =43.13, Count ({S6. high, SUM. middle}) =19.27, Count ({S6. high, TRA. middle}) = 27.03, Count ({SUM. middle, TRA. middle}) = 40.29.

8. C_2 is generated from the above C_2, and fuzzy items less than the minimum support degree is deleted. It can be got that L_2 = {{S1. high, S4. high}, {S3.middle, S4.middle}, {S3.middle, SUN. middle}, {S3.middle, TRA. middle}, {S4. high, S6. high}, {S5.middle, TRA. middle}, {SUN. middle, TRA. middle}}.

9. It can be obtained by connecting L_2 that C_3 = {{S3.middle, S5.middle, SUM. middle}, {S3.middle, S5.middle, TRA. middle}}. Count ({S3.middle, S5.middle, SUN. middle}) =43.6>40; Count ({S3.middle, S5.middle, TRA. middle}) =41.32 >40. Excavation work ends here.

10. By frequent item-set L_3 and L_2, and minimum confidence of minconf=0.8, fuzzy association rules are generated:

$\{\mathrm{S3}.middle，\mathrm{S5}.middle\} \Rightarrow \mathrm{SUM}.middle$, SUP=0.42, conf=0.8;

$\{\mathrm{S3}.middle，\mathrm{S5}.middle\} \Rightarrow \mathrm{TRA}.middle$, SUP=0.43, conf=0.8

5.3 Application of the Conclusion

From the above rules, it can be drawn that for student whose total scores and the training results in the experiment are in the middle, their results in the experimental S3 and S5 are all in the middle, thus we can strengthen training of experiments S3 and S5, so as to improve students' achievements. In addition, though some rules cannot meet the requirements of our minimum support and confidence, we can also refer to its use in teaching to elevate students' achievements, thus in the service of cultivation of professional capabilities for students.

The results of above association rules are reported to teachers, and according to the rules, when teachers of this course are teaching the same course in the second semester, they shall focus on experiments of S3 and S5. Also, S2 experiment with generally low results is also intensely trained. The average score of this course in the second semester has been risen 9.59 percent, and the number of students whose scores are more than 85 points has increased from 11.18 percent to 20.39 percent, which is a significant increase.

6 Conclusion

Association rule mining is extensively applied to many fields. Mining the potential association among data items from the large-scale database also provides referable suggestion for the decision-making in the industry. Fuzzy association rules mining algorithm takes advantage of the ideas of dealing with fuzzy and uncertain problems in fuzzy set theory to keep the edge data, and makes good achievement in automatic control, machine learning and other fields by the handling method which more conforms to the semantic expression and thinking mode of human.

Based on the performance data of the course of Linux application for students in 2011 of a certain major of a vocational technical institute in Hainan, this paper has realized the data analysis of students' performance by using fuzzy association rules and has drawn some conclusions, which has provided decision support for teachers with certain effects achieved. In the next phase of work, a further excavation study can be made by combining students' strengths and interests, and subjective and objective factors for teachers, *etc.* Data mining technique can be fully applied to the analysis of performance to help teachers and administrators gain valuable "knowledge", so as to provide services for teaching, which is a great help to the improvement of teaching quality in higher vocational colleges.

Acknowledgement

This work is supported by Natural Science Foundation of Hainan Province of China (20156232).

References

[1] B. Thomas, G. Raju, A novel unsupervised fuzzy clustering method for preprocessing of quantitative attributes in association rule mining, Information Technology & Management 15(2014) 9-17.

[2] Z Liaghat, A.H. Rasekh, A. Mahdavi, Application of data mining methods for link prediction in social networks, Social Network Analysis and Mining 3(2013) 143-150.

[3] H.M. Liu, Application of data mining in the analysis of higher vocational colleges' achievement, [disseration] Hefei: Anhui University, 2011.

[4] F. Zhang, Analysis of vocational college student achievement based on association rules, Journal of Chuzhou University 2011.

[5] H.F. Jiang, J.F. Gu, Y.J. Liu, F.Y. Ye, H.X. Xi, M.F. Zhu, Study of clustering algorithm based on fuzzy c-means and immunological Partheno genetic, Journal of Software 8(1)(2013) 134-141.

[6] J. Liu, G.S. Xu, D. Xiao, L.Z. Gu, X.X. Niu, A semi-supervised ensemble approach for mining data streams, Journal of Computers (Taiwan) 8(11)(2013) 2873-2879.

[7] J. Li, Research on the application of data mining technology in the analysis of student achievement, [disseration] Xi'an: Xi'an Shiyou University, 2010.

[8] B. Wu, Research and evaluation of data mining in student performance analysis, [disseration] Nanchang: Nanchang University, 2009.

[9] M. Ramaswami, R. Bhaskaran, A CHAID based performance prediction model in educational data mining, International Journal of Computer Science Issues 1(2010) 10-18.

[10] J. Liu, Research on decision tree classification algorithm and its application in teaching analysis, [disseration] Nanjing: Hohai University, 2006.

[11] H. M. Li, The application of data mining in students management in university, [disseration] Beijing: Beijing Jiaotong University, 2010.

[12] C. H. Chen,.T. P. Hong, Y. Li, Intelligent Information and Database Systems, 2015.

[13] J. Cranshaw, E. Toch, J. Hong, A. Kittur, N. Sadeh, Bridging the gap between physical location and online social networks, in: Proc. UBICOMP'10, 2010.

[14] P. Yan, G.Q. Chen, Discovery implication-based fuzzy association rules, Fuzzy Systems and Mathematics 18(2004) 279-283.

[15] W.D. Wang, Research and application of fuzzy text clustering algorithm, [disseration] Xi'an: Xidian University, 2012.

[16] D.M. Li, T. Li, T. Zhao, A new clustering method based on type-2 fuzzy similarity and inclusion measures, Journal of Computers (Taiwan) 9(11)(2014) 2559-2569.

[17] W.H. Liu, The Research and Application of Weight Fuzzy Association Rules Mining Algorithm, [disseration] Suzhou: Suzhou University, 2010.

[18] W.H. Shou, W.H. Fan, Knowledge map mining of financial data, Tsinghua Science and Technology 18(1)(2013) 68-76.