Research on Teacher Evaluation Based on Artificial Intelligence and Big Data

Hongbin Gu*, Feng Jin[1], Kongpeng Wei[2]

{guhongbin@pjvtc.edu.cn*, 0332@pjvtc.edu.cn[1], kpwei@pjvtc.edu.cn[2]}

Panjin Vocational & Technical College, No.999 Zhengbang Road, Liaodong Bay New District, Panjin City, Liaoning Province 124000

Abstract. Big data focuses on multi-dimensional, in-depth mining and scientific analysis of a large amount of data, discovering the hidden relationship and value behind the data, which helps the evaluation of teaching quality shift from speculation based on small-sample data or fragmented information to evidence-based decision-making based on a full range of full-scale data. The use of artificial intelligence technology to analyze the evaluation results helps to improve the credibility and validity of teacher evaluation and reduce the tension and conflict in the evaluation process. In this paper, based on the characteristics of vocational education, based on artificial intelligence and big data technology, we analyze the data of each relevant factor of vocational education teacher evaluation and put forward relevant suggestions.

Keywords: Artificial Intelligence, Big Data, Teacher Evaluation.

1 Introduction

Teaching quality evaluation is an important part of teaching management. Retrievable Chinese research results on teaching quality evaluation can be traced back as early as Li Hongxing's application of fuzzy mathematics in teaching quality evaluation in 1983 [1], research on teaching quality evaluation in higher vocational colleges and universities began in 2000 [2], in 2001, research on multivariate analysis of teaching quality evaluation began [3], research on teaching quality evaluation based on big data appeared in 2015 [4] , the research on teaching quality evaluation of higher vocational colleges and universities based on big data is late, starting in 2016 [5], and there are few research results.

Big data and artificial intelligence technologies are also gradually applied to the field of teaching quality evaluation. Algorithms such as neural network [6], factor analysis [7], Markov chain [8], principal component analysis [9], cluster analysis [10], association rule analysis [11], support vector machine [12], and analysis of variance [13] have been applied to the research of teaching quality evaluation.

2 Methods

By collecting teaching quality evaluation information for two semesters of an academic year (2018-2019 academic year) in a higher education institution, 289 teaching quality evaluation

ICMEIM 2023, September 08-10, Wuhan, People's Republic of China

DOI 10.4108/eai.8-9-2023.2340179

data were obtained (teachers who had taken more than one course in two semesters were calculated by average score). Teaching quality evaluation data includes: teacher's name, faculty number, age, title, education, and evaluation score.

In order to find the correlation between teachers' age, title, education and evaluation scores, the correlation of these attributes was calculated using python's seaborn library (with values ranging from 0-1, the larger the correlation between the two attributes, the greater the correlation), and it was found that only the correlation between the age and title was larger, which reflects the fact that the evaluation of titles in colleges and universities can only be realized after many years of hard work. As shown in Figure 1.

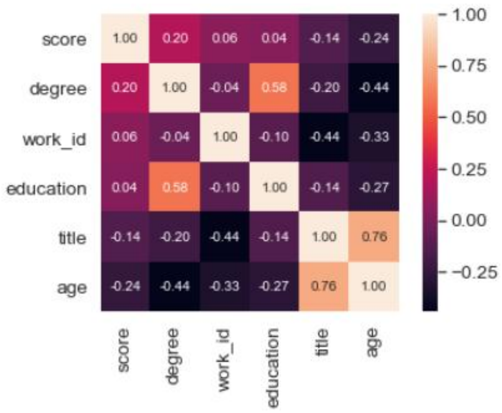

Figure 1 Correlations between factors associated with teacher ratings

It was also found that there was some negative correlation between teacher ratings and age (-0.24 points), and the correlation obtained by analyzing the relationship between teacher ratings and age using python's seaborn library is shown below:

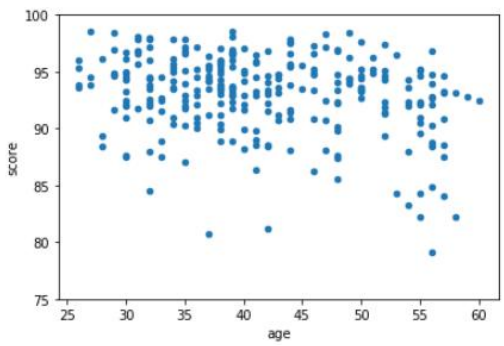

Figure 2 Correlation between Teacher Ratings and Age

Various machine learning algorithms were tried to analyze the relationship between these attributes and the results were not obvious, so an attempt was made to mine the relationship between these data using the association rule Apriori algorithm.

Association rule mining is a commonly used method in data mining. The concept of association rules was proposed by Agrawal in 1993, association rules are used to find frequent patterns, associations, correlations that exist between sets of items or sets of objects in relational data, and to get the interdependencies between data objects by analyzing them. The association rules are described below in a formalized language:

Definition 1 The dataset for association rule mining is denoted as D. D is a general transaction database, D={t_1,t_2,... ,t_n},t_k is a transaction and the elements in t_k are items.

Definition 2 Let I={i_1,i_2,...,i_m} be a set consisting of all items in D. Any subset X of I becomes a set of items in D, |X|=k, then the set X is said to be a k-item set.

Definition 3 The number of transactions in a dataset D containing an itemset X is called the support number of the itemset, and the support of an itemset X is extremely left support(X). If support(X) is not less than minsupport, then X is said to be a frequent itemset, otherwise X is said to be an infrequent itemset.

Definition 4 If X,Y are itemsets and $X \cap Y = \varphi$, the implication X => Y is called an association rule. The support of the itemset ($X \cup Y$) is called the support of the association rule X => Y. support(X => Y) = support($X \cup Y$). Confidence of association rule X=>Y is the percentage of transactions in D that contain X while also containing Y.

$$\text{Confidence}(X=Y)=\text{support}(X \cup Y)/\text{support}(X)\times 100\% \tag{1}$$

Definition 5 An association rule X=>Y is said to be strong if support(X=>Y)≥ minsupport and confidence(X=>Y) ≥minconfidence, otherwise it is said to be weak. minconfidence is the minimum confidence.

Apriori algorithm is the core of the association rule mining algorithm and the specific algorithm is described as follows:

Input: transactional dataset D, minimum support

Output: set of all frequent items in D

Method:

```
L1 = All frequent data itemsets;
for(k=2;Lk-1≠φ;k++)
{
    Ck = apriori-gen(Lk-1,minsupport);
    For all t∈D do
    {
        Ct=Subset(Ck,t);
        For all c∈Ct do c.count++
    }
}
```

```
Lk={ c∈Ct |c.count≥minsupport};
Return L={All Lk}
```

The steps for using association rules for teacher evaluation are as follows. First of all, it is necessary to convert the teaching quality evaluation data into a dataset represented by symbols. The division of the dataset is related to the support and confidence. The computation of confidence is shown as equation (1). Too much division of the dataset will lead to too low support and confidence, and the persuasive power of the rules is too low. Too little division of the dataset will easily cause some infrequent feature items implied in the data to be ignored, and it is difficult to discover sufficiently useful rules. As a rule of thumb, the datasets for titles, degrees, ages, and assessment scores are divided into four parts.

The division of the dataset for faculty titles and degrees is done according to their natural attributes:

The dataset representation of titles is as follows:

P1 - Professor; P2 - Associate Professor; P3 - Lecturer; P4 - Assistant Professor and below.

The dataset of degrees is represented as follows:

D1 - Doctoral; D2 - Master's; D3 - Bachelor's; D4 - No degree.

The age dataset can be divided into different age groups of young and middle-aged people, such as over 49 years old, 35-49 years old, 30-35 years old, and under 30 years old [5]. The age of teachers in the dataset was used because the age of teachers in the stage of 35-49 years old accounted for 49% of all teachers, which led to the fact that the regions with the highest confidence level for rating scores excellent and good in the results of the association rule algorithm were all in this age group. Improvement was made by dividing the teachers in the dataset into four groups equally according to their number, and the age distribution of the teachers after the average grouping was as follows: more than 47 years old, 37 - 47 years old, 34 - 37 years old, and less than 34 years old. They are denoted by A1, A2, A3, and A4, respectively.

Different attempts were made to quantify the rules for quantifying the teaching effectiveness rating scores, and the use of a fixed range of score bands began to be considered: above 85, 70-85, 60-70, and below 60 [5]. Comparison of the distribution of scores in the adopted dataset revealed that most of the scoring data were concentrated above 90 points as shown in Figure 3, which was not suitable for the fixed score band approach. The classification method in reference [14] was categorized into four classes of excellent to poor according to the standard deviation of the normal distribution: 15.8%, 34.2%, 34.2%, and 15.8% of the ranking by grades. The support and confidence of the association rule for obtaining excellent with this classification is too low because the sample proportion is too low when counting 15.8% as excellent according to a normal distribution. The final approach used was that the scores were categorized into 4 classes according to 25%, 25%, 25%, 25%.

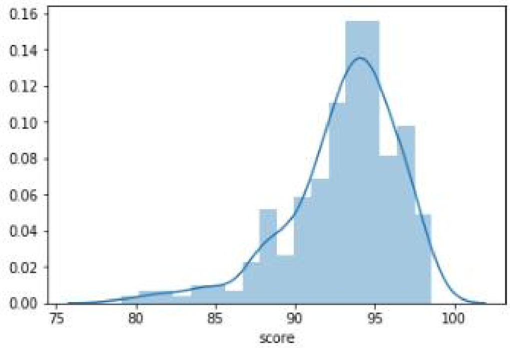

Figure 3 Distribution of scores

3 Results

The association rule Apriori algorithm [15] is used to find the frequent item set and the association rule is obtained according to the given minimum confidence level. Setting the minimum support as 0.07 and the minimum confidence level as 0.3, the association rules with excellent classroom teaching effect are obtained as shown in Table 1, and the association rules with good classroom teaching effect are shown in Table 2.

Table 1 Association rules for excellent classroom teaching effectiveness

Association Rules	Support	Confidence
A4 -> S1	0.080	0.319
A3 -> S1	0.080	0.319
D2 -> S1	0.156	0.326
[A4,D2] -> S1	0.031	0.265
[A3,D2] -> S1	0.073	0.368

From the data analyzed, it was found that students' satisfaction with teachers is mainly related to age, education, and title. Teachers who are younger or have higher qualifications are more likely to have excellent classroom teaching results. Teachers who are older or have higher titles are more likely to have good classroom teaching.

Teachers in the age groups of less than 34 (A4) and 34 to 37 (A3) are easy to be popular with students because they are close to their students' age and have a smaller generation gap, or have higher education and higher professional tiers, resulting in a high support and confidence level for high rating scores (S1). Teachers with a master's degree (D2) have high support and confidence for high rating scores (S1), indicating that teachers with higher levels of expertise teach better. The correlation between title and rating scores is not significant. The correlation between the attributes of younger age and higher education can be further analyzed. The lower support and confidence ([A4,D2] -> S1) for higher rated scores for teachers under the age of 34 (A4) and with a Master's Degree (D2) suggests that the main reason for the popularity of teachers in this age group is the proximity of their age to the students, the smaller generational gap, and the ease of interaction with the students. The higher support and confidence levels for the higher rating scores of teachers aged 34-37 (A3) with a master's degree (D2) indicate that the reasons for the popularity of teachers in this age group include the educational factor in addition to age.

Table 2 Association rules for good classroom teaching effectiveness

Association Rules	Support	Confidence
A1 -> S2	0.076	0.306
P2 -> S2	0.090	0.309
[A1,P2] - > S2	0.055	0.615

Teachers who are 47 years of age or older (A1), or whose title is Associate Professor (P2) have higher support and confidence levels for their rating scores at the good level. This indicates that teachers with longer teaching experience or higher titles are more likely to teach well. The analysis reveals that there is a relatively strong correlation ([A1,P2] - > S2) between the two, age above 47 and Associate Professor, in the rank of good rating scores, which is also in line with the fact that more teachers with longer teaching experience are rated with higher titles.

4 Discussion

The above used are the conclusions obtained by using the correlation rule for different subjects and teachers in the same school year. However, the teaching ability and level of teachers are dynamically changing, and in order to make comparisons on the same basis, we analyzed the data on student evaluation scores for the same course (Computer Fundamentals) taken from the full data set over a span of 5 years. The change in the titles of some of the teachers during these 5 years has produced a change that is reflected in the dataset. Data correlation analysis through python's seaborn library revealed that the correlation between the various factors associated with different teacher ratings for the same course was similar to the correlation of data from different courses in the same academic year.

It was also found that there was some negative correlation between teacher ratings and age (-0.27 points) for the same course, and the relationship between teacher ratings and age was analyzed using python's seaborn library to obtain a correlation similar to that in Figure 2.

5 Conclusion

The conclusion obtained for the same course across multiple academic years is similar to the results of the ratings analysis of different courses in the same academic year, i.e., younger teachers get higher student ratings than older teachers. Through interviews with teachers and students as well as questionnaires, we found that this phenomenon is due to the fact that young teachers are more in tune with the development of the times by teaching more new courses because of the rapid updating of knowledge nowadays, and on the other hand, it is due to the fact that the age difference between the young teachers and the students is smaller, which makes it easier for them to communicate with the students. Because of this, our suggestion is that older teachers should also update their knowledge and culturally study more about the psychology of younger students, so as to improve students' acceptance from both sides.

Acknowledgments. Panjin Vocational & Technology College Project (PJZYKYKT202307)

References

[1] LI Hongxing,LIU Shuqi. Application of fuzzy mathematics in teaching quality evaluation[J]. Research on Teaching,1983(01):9-11.

[2] HAN Wenzhong,XU Ming,XIAO Lizhong,YU Lijun. Research on Teaching Quality Assurance System and Teaching Evaluation System in Higher Vocational Schools[J]. Journal of North China Institute of Aerospace Industry,2000(04):1-8.

[3] Niu Yuqi, He Ping. Multivariate statistical analysis of teaching quality evaluation data[J]. Journal of Shaanxi Institute of Technology,2001(04):67-70.

[4] YE Zhiwei,YAN Lingyu,LIU Wei,CHEN Hongwei. A preliminary study of teaching quality evaluation in colleges and universities based on big data[J]. Science and Technology Information,2015,13(26):190-191.

[5] Zhang Weiguo,Chen Haiyan. Research and analysis of big data in the multivariate evaluation system of teachers' teaching quality in higher vocational colleges and universities[J]. Modern Computer(Professional Edition),2016(18):45-49.

[6] WANG Xuhui,HUANG Feihua. Teaching quality evaluation model and application based on BP neural network[J]. Research on Higher Engineering Education,2007(05):78-81.

[7] WANG Ji-Quan, QIU Li-Chun, WANG Fulin, SO RuiXia. The application of factor analysis method in the evaluation of teaching quality of college teachers[J]. Practice and Understanding of Mathematics,2011,41(15):28-34.

[8] FENG Hong,ZOU Hua,WEI Wenyuan. Application of Markov chain in teaching quality evaluation[J]. Journal of Tianjin Normal University(Natural Science Edition),1999(01):6-10.

[9] Zhang Sumei. Comprehensive evaluation of teachers' teaching quality based on principal component analysis[J]. Science and Technology Information (Science Teaching and Research),2008(22):541-542.

[10] Feng Mei. Evaluation of teachers' classroom teaching quality based on fuzzy cluster analysis[J]. Practice and Understanding of Mathematics,2008(02):12-15.

[11] Xie Qiuli. Data mining for teaching quality evaluation based on association rules[J]. Modern Computer (Professional Edition),2008(06):25-27.

[12] LIU Wei,SUN Lin. Classroom teaching quality evaluation based on support vector machine[J]. Journal of Hefei University of Technology (Natural Science Edition),2010,33(07):968-971.

[13] Feng Mei. One-way ANOVA of teachers' influence on teaching quality[J]. Mathematics Practice and Understanding,2005(11):61-66.

[14] SONG Fuqing,PANG Shanqi,LIU Wanli. Application of Ma's normal process in education quality evaluation[J]. Journal of Henan Normal University(Natural Science Edition),2007(03):8-10.

[15] Lina Lu,Yaping Chen,Hengyi Wei,Maishun Yang. Research on Apriori algorithm in mining association rules[J]. Small Microcomputer Systems,2000(09):940-943.