The Interaction of College Students' Sores in Different Courses Based on Association Rule

Lingyue Gao[1, a], Xiufang Hao[2, b, *]

[1]Shanxi University of Finance and Economics, Shanxi 030000, China

[2]Shanxi University, Shanxi 030000, China

[a]dczxglyxhjl@126.com, [b]tydzhhxf@163.com

*Corresponding author

Keywords: Linkages of courses, Improve level of education, Association rule, Apriori algorithm

Abstract: We transform scores data into discrete attributes, using Association rule to analyze the students' scores of Economic Statistics major to find the linkages of courses. This paper provides the basics for improving level of education and reminding students to study hard. The mining results can provide targeted guidance for students, such as cultivating integrating knowledge ability, and provide effective information for the adjustment and improvement of teaching methods.

1. Introduction

Many students find that their grades are getting worse after they enter the university. Compared with compulsory education, the different teaching modes in higher education make many freshmen unable to fit in it, together with poor possessiveness and so on, these students must cram to rescue grades. But this may have bad influence on learning other follow-up professional courses because of high linkages to the backward and forward. Besides, universities attach importance to enhancing the atmosphere of study to construct the “Double First-Class”. Universities expect every student could take each course seriously.

2. Model Theory

2.1 Introduction to Association Rule

Association rule is a classical rule-based machine learning method for discovering relations among variables in the large databases. So, using Apriori algorithm to analyze our issue is reasonable. This method originated from shopping basket analysis and it is employed today in many application areas including Web usage mining, network intrusion detection [1], and biomedicine [2]

It was proposed by Agrawal and Srikant in 1994 [3]. The monotonicity of frequent item-sets is emphasized, stating if item-set k is a frequent item-set, its subsets are also frequent item-sets. Apriori algorithm is scanning and taking recursion layer-by-layer to generate strong associated rules by the frequent item-sets k, and construct candidate k+1 item-sets. Then check the rules with a given minimum support degree and confidence threshold to decide whether it is reasonable. When there is no new candidate itemset generated, the Apriori algorithm stops.

2.2 Evaluation Criteria

The commonly used evaluation criteria for frequent item-sets are support, confidence and lift.

Support is the proportion of the number of occurrences of several associated data in the data set to the total data set. Or the probability of the occurrence of several data associations. Confidence reflects the probability of one data appearing. To select a frequent data set, we need to customize the evaluation criteria in different problems. Lift represents the ratio of probability of containing X after containing Y to probability of X population occurrence. That is:

$$\text{Lift}(X \rightarrow Y)=P(Y|X)/P(Y)=\text{Confidence}(X \rightarrow Y)/P(Y) \quad (1)$$

The degree of lift reflects the relationship between Y and X. X→Y is an effective strong association rule with a lift degree greater than 1, and X→Y is an invalid strong association rule with a lift degree less than or equal to 1.

3. Association Rule Analysis on Scores

3.1 Preprocess Data

Considering that there are great different courses in different majors, and the scores of different teachers in the same course are also different, through comparison, I used the data of Economic Statistics major during two years in Shanxi University of Finance and Economics as training set. In all the courses of two years, I choose only 15 courses with a high degree of relevance to the major of economic statistics. Each course is represented by "A~O" separately, shown in the Table 1.

Table 1 Course Name, Classification and Symbol of Courses

Classification	Name	Symbol
General Foundation Course	Advanced Mathematics (Volume One)	A
	Advanced Mathematics (Volume Two)	B
	Fundumentals of University Computer	C
	Experiment of University Computer Foundation	D
	Computer Programming (VB)	E
	Linear Algebra	F
Subject Foundation Course	Political Economics	G
	Microeconomics	H
	Macroeconomics	I
	Econometrics	J
	Econometric Experiments	K
Professional Compulsory Course	Economic Optimization Method	L
	Introduction to Statistics	M
	Probability Theory	N
	Mathematical Statistics	O

To transform continuous data into discrete attributes, the scores are divided into three grades, and "1, 2, 3" is used to represent excellence, good and poor respectively. In the reference, the general ranking is 80-100 points for excellence, 60-80 points for good and below 60 points for poor [4].

Through overall data, it is found that the scores of some subjects are highly concentrated due to the influence of external factors and the overall distribution is skewed. In Table 2 I take the scores of professional compulsory course as example. The standard deviation of Economic Optimization is on the low side, indicating that students' scores are concentrated in 73 points. For another example, students get an average score of 77.07 in Introduction to Statistics, and being a left skewness distribution, indicating that students' scores tend to high scores. While the general classification criteria for high scores above 77 points is only 80 points.

Table 2 Descriptive Statistics of Professional Compulsory Course Achievements

	Min	Max	Mean		Standard Deviation	Skewness	
			Statistic	S.E.		Statistic	S.E.
Economic Optimization Method	60	96	73.71	1.032	9.787	.368	.254
Introduction to Statistics	41	98	77.07	1.257	11.922	-1.005	.254
Probability Theory	60	100	81.68	1.184	11.230	-.435	.254
Mathematical Statistics	60	100	80.78	1.223	11.606	-.311	.254

If the general classification criteria are adopted, the stratification may be not obvious, and there may be deviations, so it is impossible to objectively analyze the relationship among them. Combining with the idea of "3σ", this paper redesigns a set of grading standards listed in Table 3.

Table 3 Discretization Standard of Scores

Grade	Standard
1	$Xi \geq \bar{X}+3*S(X)$
2	$\bar{X}+3*S(X) \geq Xi \geq \bar{X}-3*S(X)$
3	$Xi < \bar{X}-3*S(X)$

3.2 Establish Association Rules

Considering that the time of course learning varies, it is necessary to specify a variable as the target value and then to study it with SPSS Modeler 18.0.

General curriculum and professional compulsory curriculum are closely related, which is mainly reflected in two aspects. First, general curriculum is the basis of professional curriculum. By learning the basic, comprehensive and extensive knowledge, students broaden their horizons and lay a solid foundation for the deepening process of knowledge. Second, professional knowledge and general knowledge are cross-integrated, sometimes there is no clear boundary between them [5]. Professional education aims to cultivate students' professional skills in a certain field of knowledge. As the purpose of higher education is to train qualified person, professional curriculum is particularly critical.

Therefore, I take the scores of the compulsory courses in the second semester as the objective value to analyze. Because the courses studied at the same time also influence each other, these variables do not be deleted in the correlation analysis.

When the thresholds of support and confidence are set to 0.15 and 0.85 respectively, the mining results are shown in the Table 4.

Table 4 Association Rules Between Professional Course and Other Courses

No.	Consequent	Antecedent	Rule Support	Rule Confidence
1	M = 3	H = 3 and G = 3 and N = 3	16.67	86.67
2	M = 1	J = 1 and O = 1	15.56	85.71
3	M = 3	E = 3 and H = 3 and G = 3 and N = 3	15.56	85.71
4	L = 3	J = 3	13.33	91.67
5	L = 1	J = 1 and G = 1 and B = 1	12.22	81.82

It can be found that each rule contains subject foundation course. The score of common course determines the scores of professional courses. Therefore, I further study the factors affecting common course. Since Introduction to Statistics and Economic Optimization Method study after the common course, two variables are deleted in the filter, and then the correlation analysis is carried out for each common course.

The thresholds of support and confidence are still set to 0.15 and 0.85 respectively, the mining results are shown in the Table 5.

Table 5 Association Rules Between Common Course and Other Courses

No.	Consequent	Antecedent	Rule Support	Rule Confidence
1	G = 1	H = 1 and O = 1 and N = 1	15.56	92.86
2	G = 3	E = 3 and H = 3	22.22	90.00
3	H = 3	E = 3 and N = 3	18.89	94.12
4	H = 1	G = 1 and E = 1 and B = 1	17.78	93.75
5	H = 3	E = 3 and G = 3	22.22	90.00
6	I = 1	G = 1 and H = 1 and O = 1	16.67	93.33
7	I = 1	G = 1 and O = 1 and E = 1	15.56	92.86
8	I = 1	H = 1 and O = 1 and B = 1	15.56	92.86
9	I = 1	G = 1 and H = 1 and O = 1 and B = 1	15.56	92.86
10	J = 2	B = 2	18.89	82.35
11	J = 2	K = 1 and F = 2	16.67	80.00

3.3 Test the Association Rules

However, in strong association rules meet the given support and confidence, there are also effective strong association rules and ineffective strong association rules. According to the formula (1), I test these rules in Excel and get the lift of each rule. After testing, the promotion degree of each rule is greater than 1. The rules are effective.

4. Conclusions

For professional compulsory courses, general foundation course and subject common courses will affect performance.

Many professional courses will repeat the basic contents of general education courses in learning, and on this basis continue to deepen. As in Table 4, Rule 2, students learn Mathematical Statistics excellently, it is easy to get excellent scores in Introduction to Statistics.

General foundation course and subject common courses enable students to understand the original theory and thought. But the theory of professional compulsory course is closer to the reality. That is the application of ideal theory. So, if students have poor foundation, it is not easy to learn professional compulsory courses well.

Some courses seem to be no correlation outwardly, but the basic course is helpful for students to expand their thinking, enhance their logical reasoning ability, and help them to learn professional knowledge better.

Analysis also shows that no matter what the internal linkage between the courses is, the better students' scores are, the easier students will get good grades, and vice versa.

5. Suggestions

a) Schools should manage to help freshmen get through the adaptation period. For example, offering lectures to introduce various effective learning methods, guide students to overcome psychological problems. To avoid losing interest in learning in the early stage, and ultimately affect the follow-up course learning.

b) Students should take the initiative to enhance their awareness, set clear goals from freshman year, and study actively. Every course should be taken seriously, no course should be despised because of its attribute.

c) Students should not only pay attention to grades, but also pay attention to the learning process, like cultivating self-learning ability and integrating knowledge. Teachers should also train students' diversified abilities in routine teaching. Only in this way can students make continuous progress in learning and better adapt to social development.

References

[1] Liping Yang. Research and Application of Apriori Association Rules Algorithm in Network Intrusion Detection [J]. *Digital Technology and Application*,2018,36(08):107-108.

[2] Hui Teng, Lan He, Yunna Song. Research on Biomedical Data Mining Method[J]. *China Continuing Medical Education*,2017,9(32):22-24.

[3] Agrawal R, Srikant R. Fast Algorithm for Mining Association Rules[J]. *Journal of Computer Science & Technology*, 1994, 15(6):619-624.

[4] Hua Wang, Ping Liu, Application of improved association rule algorithm in early warning of student performance [J]. *Computer Engineering and Design*,2015,36(03):679-682+752.

[5] Jianmin Zhang, Jian Zhu. Problems and Countermeasures: Probe and Analysis of the Curriculum Reform of the Universities' General Education [J]. *Forum on Contemporary Education*,2017(05):107-113.