

# Analysis of Influencing Factors of Learning Efficiency for Undergraduates

Yan Chen

College of Information System and Management  
National University of Defense Technology  
Changsha, China  
15575186160@163.com

Ping Jiang

College of Information System and Management  
National University of Defense Technology  
Changsha, China  
jiangping@nudt.edu

**Abstract**—For undergraduate student, the learning efficiency for a course study often determines the final academic record. Thus improving efficiency is one of the most concerns for undergraduates during their university study. This paper aims at researching the influencing factors of learning efficiency, including students' subjective factors and objective factors such as course arrangement, especially, establishes a type of method in studying the influential factors and let students find the causes influencing their learning efficiency. A survey is designed to help collecting information of undergraduates' course learning efficiency and related factors. Then, the survey data is analyzed by data mining method to identify some interesting relationships between learning efficiency and its influencing factors.

**Keywords**—learning efficiency; data mining; system engineering

## I. INTRODUCTION

High score for a course is always regarded as the measurement of success in university study. As high learning efficiency often helps to improve the comprehensive quality for undergraduates, learning efficiency is generally accepted as the necessity to improve academic performance, which may result in high score[1]. However, due to various reasons, many college students suffer from insufficient learning efficiency. Having spent a lot of efforts in courses study, many students still could not receive their ideal score[2]. Therefore, this research's object is to analyze the impacts of various factors on undergraduates' learning efficiency, and to identify relationships among them. The results can be used as references to improve the learning efficiency of undergraduates' course study and help improve course arrangements for teachers and colleges.

The study of learning efficiency is interdisciplinary. It's an issue of common concern in many fields. In recent years, a lot of researches have been made in areas such as management and psychology[3]. Yet researchers often concentrate on studying efficiency from the psychological factors, and most of them are qualitative[4]. There are few quantitative analysis of the impact on learning[5]. So, this study focuses on undergraduate students, which makes it more desirable by them, to analyze efficiency and factors quantitatively, also to provide suggestions for improvement.

To analyze the factors comprehensively, literature learning in this area is necessary, which costs a lot of time. Then, an

eligible questionnaire is designed and sent to undergraduates in the National University of Defense Technology (NUDT). Data mining and system engineering methods are both used to analyze the survey data. Finally, some associated relationships are identified. The conclusion summarized our research.

## II. QUESTIONNAIRES AND DATA PREPROCESSING

### A. Questionnaire design

For research purposes and realities, a course is chosen and the undergraduates participating in the course are the research objects in the survey. The questionnaire is comprised of 69 questions. The survey is carried out in two steps, which are completed by the same batch of students during middle and final of the course respectively. Through the elaborate design of the questionnaire, and 350 questionnaires were handed out, among which 302 valid responses were fed back. the survey data was analyzed through correlation analysis and data mining.

### B. Data preprocessing

Data preprocessing or pretreatment includes survey data quantification and removal of invalid data. The software of Excel is used to store the survey data. In each column of the sheet in Excel, respondent's ID number is the key word, to be helpful to number the surveys. Every respondent was recorded to one column in the sheet, so each column records the students' answers to all questions.

## III. DATA MINING

### A. Introduction

Data mining is to reveal unknown meaningful relationships or trends and patterns through analysis of large amounts of data[6]. Our research is to reveal possible relationships between the learning efficiency and its influencing factors, through the data mining of the survey.

Associated rule mining is one of the activist data mining methods. It can find relationships between events[7]. At first, it was used to found connection in supermarket trading database: through the analysis of sales data, Wal-Mart supermarket founded that male consumers who buy beers usually buy diapers together. In further survey, they found these parts of men are mainly young fathers, who don't forget to buy diapers for their children when buying beers for themselves. Eventually, the diaper shelves of the supermarkets have been moved to beer shelves. Wal-Mart received praises unanimously.

## B. Basic concepts of Apriori algorithm [8]

### 1) The definition of support degree

$\text{support}(X \rightarrow Y) = |X \cap Y|$  = The appearance time of items in sets X and Y at the same time in each record.

For example:  $\text{support}(\{\text{option 1 question 3}\} \rightarrow \{\text{option 4 question 34}\})$  = the appearance time of option 1 question 3 and option 4 question 34 in the same records = 3.

### 2) The definition of confidence

$\text{confidence}(X \rightarrow Y) = |X \cap Y| / |X|$  = [The appearance time of items in sets X and at the same time in each record] / [the appearance time of x].

For example:  $\text{confidence}(\{\text{option 1 question 3}\} \rightarrow \{\text{option 4 question 34}\})$  = The appearance time of option 1 question 3 and option 4 question 34 in the same records / The appearance time of option 4 question 34 =  $3/3 = 100\%$ .

### 3) The definition of itemset

If event A has k elements, then call this event a k-itemsets.

When one rule meet the minimum support thresholds (min\_sup) and minimum confidence threshold (min\_conf) at the same time, this rule was called strong a rule. If one itemset can meet the minimum level of support degree, it was called a frequent itemset. Where the minimum support degree and minimum confidence degree was given by data analyst appropriately according to subjective experience.

## C. The process of association rules mining

### 1) Overview

#### a) Find all frequent itemsets

According to definition, the frequency of the appearance of these itemsets is equal to the predefined minimum support.

#### b) Produce strong association rules through the frequent itemsets

According to definition, these rules must meet the minimum support degree and minimum confidence degree.

### 2) Apriori law

#### a) Law 1

If one set is a frequent itemset, all subsets of it are frequent itemsets.

Example: suppose a set {A, B} is a frequent itemset, that is, the appearance time of A and B in one record is greater than or equal to minimum support degree, the min\_support, then the appearance time of its subsets {A}, {B} must be greater than or equal to min\_support, that is, it's subsets are frequent itemsets.

#### b) Law 2

If a set is not a frequent itemset, none of its supersets are frequent itemsets.

Example: suppose that the set {A} is not a frequent itemset, That is, the appearance number of it is less than min\_support, then the appearance number of any superset of it such as {A, B} must be less than the number of min\_support. So, its supersets cannot be frequent itemsets.

### 3) The mining process of Association Rules(specifically)

#### a) Find ordinary itemset

Process:

- i. Scan
- ii. Count
- iii. Prepare
- iv. Figure out ordinary itemsets
- v. Connect and cut to form candidate itemset

Repeat the above steps until cannot discovered more frequent itemsets.

#### b) Form association rules

Process:

According to the aforementioned definition of confidence degree, association rules form as follows:

- i. For each ordinary itemset L, form all nonempty subsets of it;
- ii. For each nonempty set S of L, if where:

$$P(L)/P(S) \geq \text{min\_conf}$$

then output regulation:

$$s \rightarrow L-S$$

**Example: the analysis of question 30, 40 and examination performance (a virtual question 70) from 4 students (Minimum support degree was set to be 2)**

Followed by the process of (2) to various candidate itemsets, which mined out relationships meet the minimum confidence degree.

From:

$$P(\{301, 401, 702\}) / P(\{301, 401\}) = 1 \geq \text{min\_conf}$$

Output regulation:

$$\{301, 401\} \rightarrow \{702\}$$

Obtain Association rules:

[I think the class is relatively fast and difficult], [This course is useful]  $\rightarrow$  [get scores in 80-90]

Matlab was used to imply Apriori algorithm. In the mining process, each student is equivalent to one "instance" (column), the digit in each row of data records the students' answers to one specific question (They are arrayed in the order of question number). For these answers, Association rules mining have been used.

## IV. RESULT ANALYSIS

### A. Analysis of learning subject

By studying the data from students who scored in 80-90 and the resulted association (support degree is 10, confidence degree is 0.8) (there are too few students score above 90), the following rule has been found:

*[I preview, and review less for the course content], [The efforts to achieving goals is relatively positive for me, attitude is not strong, nor confused], [I think if this course is took in the 1, 2 classes in the morning will give me a good spirit]  $\rightarrow$  [Get score in 80-90]*

It is found that students who score higher rarely review or preview. This appears to be a contrary to our conventional impression: among students who are learning well, except a few natively wise guys, most of them are hard-working people.

Looking the following respondents, the doubt would remove. These parts of students think they are driven and think that morning time make them more efficient. It seems that they are able to strive for their goals, and they are more energetic. College students are prone to be tired in today's Chinese colleges, which is particularly evident in the morning. But students scored higher have better efficiency in the morning, which suggested that their everyday statuses are good. Their inner class learning productivity is more efficient than average person. So, it's no doubt that they can learn well without too much previewing or reviewing.

#### *B. Analysis of learning object*

By studying the data from students who scored in 80-90 and the resulted association (support degree is 10, confidence degree is 0.8) (there are too few students score above 90), the following rule has been found:

*[I think the course movement is relatively fast and difficult], [I think the course is useful] → [Get score in 80-90]*

It seems that students who get better scores usually think their specialized courses are relatively fast and difficult. And this is not an isolated phenomenon during this survey. After investing some of the students, it is suggested that the students who think the courses are useful are prone to work harder to complete their tasks. And more difficult the course is, the stronger sense of urgency will arise. Thus they can achieve good results is not surprising at all.

#### *C. Analysis of learning environment*

By studying the data from students who scored in 80-90 and the resulted association (support degree is 10, confidence degree is 0.8) (there are too few students score above 90), the following rule has been found:

*[My main extracurricular activities are community events], [Review and finding relevant information is required for me to finish the homework after class], [The time invested in learning is not productive for me] → [Get score in 80-90]*

Through further analysis of association rules, it is found that students who scored higher tend to invest amount of time in various community activities, which will correspondingly reduce the learning time. But rather than giving them backwards in learning, the reduction of study time proposed their score level.

This phenomenon can be explained from two aspects. First, Chairman Mao said: students are for studying, also learn others. That is, in college, whatever the activity may be, we should be mainly academic, but study should not be the whole of college[9] mentioned that the students who are not good at truancy are not good students. This sentence is in fact to advice that it is necessary to join community events and develop hobbies in the case of not neglecting students' studies. Just as the Germany educator Johann Gottlieb Fichte said, way to inspire students' pure love to learning, is to encourage students to active spontaneously, which should become the base of access to knowledge, so as to enable students to learn through their own activities no matter what to learn. This shows that participation in community activities enhances the activity of learning and "pure love", which in particular plays an

irreplaceable role in developing self-learning ability, also can promote the improvement of students' comprehensive ability and innovation spirit. Especially in communities such as the kinds of electronics, machinery and computers, have a direct role to upgrade the professional course record. On the other hand, the reduction of learning time requires them to learn more efficiency. They have to access to knowledge better in a much shorter time. College should also arrange a variety of Club activities to provide opportunities for exercising and practicing.

#### *D. Analysis of study manners*

The strong rules of 20 support degree of and 0.8 confidence degree:

*[I often develop learning plan], [I learn sincerely in class all the time] → [Review and finding relevant information is required for me to finish the homework after class]*

It is apparent that the "clusters" phenomenon of good habits easily upon good students. Who learn sincerely in class always develop their own study plan regularly and initially. Most of them are able to complete assignments independently. It also reflected that most good students have similar habits of study.

### V. CONCLUSION

What result in students' low productivity after large investment of learning time is caused by the problems of learning attitudes. The common characteristics of good students have been summarized as following: goals, a sense of urgency. Unsatisfactory students usually learn without allocating their time appropriately or fail to find a reasonable position, which caused "not playing well, nor learning well". This can prove the vital status of learning attitudes, followed by study manners and objective environments, because only the ones who have a clear goal and determine to work hard for it can achieve the final victory.

The results also show that appropriate participation in extracurricular activities[10] is conducive to improving the efficiency of learning. What more surprising are those teachers' faster speeds also helps to stimulate students' enthusiasm for class. This indicated that appropriate pressure is conducive to improving the productivity, just the same as appropriate relaxation.

### REFERENCES

- [1] Wang, Z.S. 2007. College students role perceiving and countermeasures of educators in college. *China Youth Research* 8: 11-13.
- [2] Du, Y.F. 2008. Survey and analysis on study motivation of college students. *Materials of Humanities & Education* 30: 56-58.
- [3] Ruo, D.Q. 2010. College students learning efficiency problems causes and countermeasures. *Realistic Magazine* 2: 13-19.
- [4] Ryan, R.M. & Deci, E.L. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55(1):68-78.
- [5] Zhao, L.J. 2011. Research on factors affecting learning efficiency of college students and improving the learning efficiency. *China Market* 39: 77-80.
- [6] Wang, G.H. & Jiang, P. 2004. Review on data mining. *Journal of Tongji University* 2: 51-53.

- [7] Wang A.P. et al. 2010. Association rule mining approaches in data mining. *Computer Technologies and Development* 4:89-91.
- [8] Qian, X.Z. & Kong, F. 2008. Research on Apriori approach in association rule mining. *Computer Engineering and Application* 17: 35-37.
- [9] Qin, B. X. 2012. Go to university, what to study? Guang Zhou: Nanfang Daily.
- [10] Zheng,Z.J. 2014. Improving engineering practical ability of undergraduates by combining extracurricular activities for innovation. *Experimental Technology and Management* 31: 25-27.