

Analysis of College Students' Practical Teaching Effect Based on Machine Learning Correlation Analysis Algorithm: Take the Software Technology Course as an Example

Liu Xingxing^{1,a,*} and Yang Chenggang^{2,b}

¹Department of Information Engineering, Private Sichuan Tianyi College, Cheng Du, China

²School of Marxism, Wuhan University of Technology, Wuhan, China

^a 342270413@qq.com, ^b 1274297207@qq.com

*corresponding author

Keywords: Training courses, Auxiliary management, Abnormal early warning, Evaluation reference, Machine learning, Correlation analysis

Abstract. In view of the difficulty in managing the training class in Colleges and universities, this paper tries to assist the training teachers to strengthen management from the point of machine learning algorithm. After analyzing the characteristics of the training class, we get 7 characteristic attributes including the training performance, the click of the mouse, the number of keyboards and the number of programs, and then we decide to use the association analysis method in machine learning to analyze the association rules between the characteristic attributes. . Finally, by using three steps of the association analysis algorithm - collecting data, preparing data and training algorithms, the frequent item sets of students with good training results {mouse, keyboard and program} are obtained. It can be found out that if the characteristic attribute values of a student are deviate from the trait attribute values of the students with good results during the training process. Too large indicates problems, which can help teachers warn students' abnormal situations and provide reference for evaluating training results.

Research Background and Significance

In college classrooms, especially in applied undergraduate and higher vocational colleges, training is very important nowadays. So in this background, Offering training class teacher want to several method, especially when the teacher demonstrates the content of the training course. A teacher to face fifty or sixty students, it is impossible to do a small class teaching one to one tutoring, so the effect of their own practical training can not be realized in real time, so there are many problems: students are lazy, do not do practical training, do things like playing cell phones or chatting, and have nothing to do with the classroom.

Students are in trouble when they encounter problems, but they are embarrassed to ask the teacher.

Students with ability can do fast and poor students do things slowly and not uniformly.

The above problems have puzzled teachers in practical training, and they have also thought of several ways:

Walking method

The so-called walking method, that is the teacher walked in the classroom after the end of teaching content of training, if found to have the above three problems in a timely manner to solve. But the problem is that the teacher can not keep moving. He may be solving some of the students' problems, when other students are not aware of the situation, and the teacher is not small, as long as the students do small movements will not be found by the teacher.

Grouping method

In order to solve the above 1.2 1.3 problem, the teacher usually divides the students into groups and divides the good and the learning students into one group so that they can form a mutual help group to solve the problem in part, but there will be a problem in the group: for example, the

students in the group are not balanced, some are strong, and there are some groups. Some groups were weak, and the students' psychological feelings were affected because the scores were calculated by the scores of groups; and some members of the group became more inactive after the division, thinking that they were always hoping for the team leader and the good members of the group, making their training worse.

Strong band weak method

In order to solve the above 1.1 problem, most teachers will assign some good students to the poor students to help them with one to one, but most of the situations are contrary to their wishes. According to statistics, the students who are not good at learning are unwilling to communicate with the students who have learned well. This may be from the perspective of psychology. The cause of the heart.

Based on the above analysis, training teachers are divided into skills, there is always no place to take care of. So from the point of view of computer thinking, when teachers are separated from surgery, we can turn to computer software technology to assist our teachers in teaching. This article is from the teacher training course of auxiliary software technology point of view, using the analysis algorithm to find some common training students association algorithm in machine learning based on, which can make an evaluation and monitoring mechanism, evaluation of students training effect and to monitor the training process of the students.

Concept of Machine Learning Association Analysis

Association analysis is a task to find interesting relationships in large data sets. These relationships can be of two types: frequent itemsets or association rules. Frequent itemsets often appear in a collection of items. Association rules imply that there may be strong relationships between the two items. Here is an example to illustrate two concepts. Table 1 gives a list of transactions in a grocery store

Table 1 .A simple transaction list from Hole Foods natural food store

Trading number	commodity
0	Soy milk, lettuce
1	Lettuce, diaper, wine, beet
2	Soy milk, diaper, wine, orange juice
3	Lettuce, soy milk, diaper, wine
4	Lettuce, soy milk, diaper, orange juice

Frequent itemsets are collections of items that often appear together. The set of Table 1, wine, diapers, soybean milk} is an example of frequent itemsets. From the above dataset, you can also find association rules such as diapers > wine. This means that if someone buys a diaper, he will probably buy wine. With frequent itemsets and association rules, businesses can better understand their customers. Although most of the instances of association rules analysis are from retail, the technology can also be used in other industries, such as web traffic analysis and the pharmaceutical industry.

When looking for frequent itemsets, there are two important concepts in the frequent definition: support and credibility.

The degree of support for a frequent itemset is defined as the percentage of records containing the frequent itemsets in the dataset. As you can see from table 1, the support of {soy milk} is 4/5. Of the 5 transaction records, 3 contained {soybean milk and diaper}, so the support of {soybean milk and diaper} was 3/5. Support is for frequent itemsets, because a minimum support degree can be defined, while only frequent items that satisfy minimum support are reserved.

Confidence or confidence is defined for an association rule such as {diaper} - {wine}. The credibility of this rule is defined as "support ({diapers, wine}) / support ({diaper})". As you can see from table 1, the reliability of "diaper - > wine" is 3/4=0.75 because the support of the diaper, wine} is 3/5 and the diaper is supported by 4/5. This means that for all records containing "diapers", our rules apply to 75% of the records.

Support and credibility are used to quantify the success of association analysis. In this study, these two important indicators will be used to analyze the teaching effect of training courses.

Domestic University Student Training Courses Use Relevance Analysis Teaching Effect Research and Analysis

There are few researches on machine learning relevance analysis in teaching effectiveness in China, especially in practical training teaching. Taking the knowledge network as an example, with "association analysis" + "teaching effect" as the keyword search article, only 37 records, but the use of "grey relational analysis", this is not the same method of correlation analysis in this paper, if add the "training" keyword, only the next 5 records, if added "machine learning" "It is the 0 record. This shows that there are few researches on the machine learning relevance analysis algorithm applied to the teaching effect analysis of training courses. This paper applies association analysis to the education industry, mainly analyzing the frequent itemsets and correlation of the teaching effect of training courses.

Research on Algorithm Model of Machine Learning Association Analysis

Suppose we are running a grocery store with a small variety of goods, and we are very interested in those products that are often purchased together. We have only 4 commodities: commodity 1, commodity 2, commodity 3 and commodity 4. So what are the combinations of goods that might be purchased together? There may be only one commodity, such as commodity 1, or two, three or all four commodities. We don't care if someone buys two goods 1 and four goods 3. We only care about him buying one or more commodities. The general process of association analysis algorithms is like table 2.

Table 2 general process of association analysis algorithm

General process of association analysis algorithm	
1	Collecting data: using any method
2	Preparation data: any data type can be
3	Analysis data: using any method
4	Training algorithm: using association analysis algorithm to find frequent itemsets
5	Test algorithm: do not need the test process
6	Use algorithm: used to discover frequent itemsets and association rules between items.

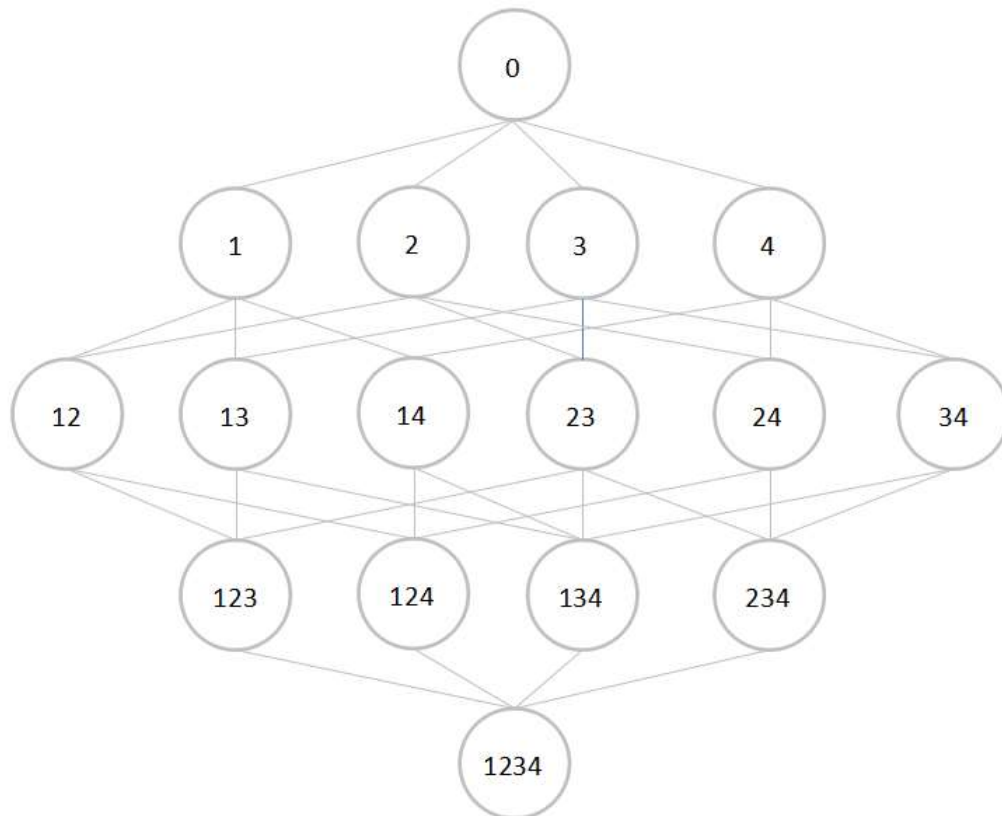


Figure 1 all possible frequent itemset combinations in set {1,2,3,4}

Figure 1 shows all possible combinations of items. The item number 1 carries the package item 1 itself. 0 in the figure is an empty set. The connection between items indicates that two or more sets can be combined to form a larger set. The goal of the case is to find a collection of items that are often purchased together. The above has been used to measure the frequency of occurrence with support. The degree of support of a set is how many transactions records contain the set. How to make a calculator support for a given set, such as {1,4}, should be traversed by each record and check that the record contains 1 and 4, and if the record does contain the two items, then the total number is added. After scanning all the data, the total number of statistics is divided by the total number of transaction records, and the support can be obtained. The above process and results are only aimed at a single set of {1,4}. To achieve the support of each possible set, we need to repeat the above process many times.

The above is machine learning association analysis algorithm model. Next, we will use this model to study the teaching effect of Java training course.

Take Java as an Example, Use Apriori Algorithm to Study Teaching Effect.

The goal of association analysis includes two aspects: Discovering Frequent Itemsets and discovering association rules. First, frequent itemsets should be found before obtaining association rules. The whole model is like Figure 2 below.

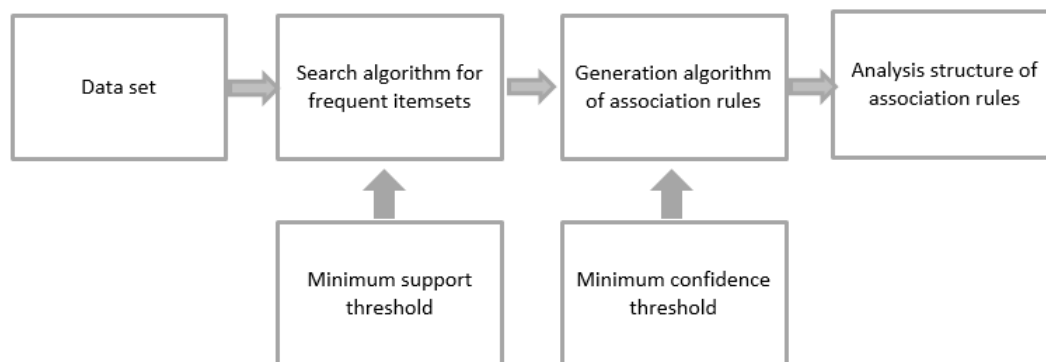


Figure 2 Diagram of association rule algorithm model

Generating Frequent Itemset

Usually generates frequent itemsets

Generating frequent itemsets has two input parameters, namely minimum support and data set. First, collect the observation data of the Java training course; then generate a list of frequent item sets for a single observation data; then scan the dataset to see those frequent itemsets to meet the minimum support and remove all those that are not satisfied; then combine the remaining data sets to generate frequent items containing two elements. Set; next, scan the observation data records again and remove the frequent itemsets that do not meet the minimum support. The process is repeated, knowing that all the combinations have been scanned.

Generation frequent itemsets for specific elements

Sometimes, we do not want to find all frequent itemsets but are interested only in itemsets containing a particular element. In this article, we will find some public features of the students who have good training results, and use these features to warn the students who do not take the training seriously, and can also provide reference for the final evaluation of the training results. Referring to the general steps in Table 2, the specific steps of this article are as follows:

Collection of data

Through the school training center to collect the data on the computer (student) in the classroom of the training room, there are 5 types of data: the click frequency of the mouse, the input frequency of the keyboard, the opening length (proportion) of the development tool, the frequency of the program and the network traffic. The meaning of data type is explained below, as shown in Table 3.

Table 3 the meaning of the data type

1 mouse click frequency	Average mouse click times per minute
2input frequency of the keyboard	The number of input times per minute of the keyboard
3 the opening length of the development tool (proportion)	Time length / lesson length of a development tool
4 The number of programs	The number of times a student uses a development tool to run the program
5 network traffic	Network traffic used by students during class time
6 rxaises the question	The number of questions raised by the students during the class time

As an example of the Android training course, the data collected, as shown in Table 4, lists only part of the data for example. The table header explains the click times of the mouse for the mouse;

the number of keys that the keyboard is the keyboard; the development tool is the opening time of the development tool; the number of programs representative of the program; network representative network The traffic flow; the question represents the number of questions.

Table 4 data collected

Serial number	mouse	keyboard	development tool	program	network	problem
1	1395	8235	45	22	217.35MB	2
2	1800	7245	44	20	105.75MB	0
3	2385	7785	45	15	72MB	1
4	254	134	45	3	456.15MB	0

Preparation of data

The data collected can not be used directly. It needs standardization and unification. Whether the results are excellent is digitalized, 1 represents excellence, and 0 represents no excellence. Other headers in the table: the click frequency of the mouse representing the mouse (the average number of clicks per minute); the keyboard is the input frequency of the keyboard (the average number of input times per minute); the development tool represents the opening time ratio of the development tool (opening time / time time); the number of times the program represents the program; the network Represents network traffic (average traffic per minute); questions represent the number of questions raised. The results are shown in Table 5.

Table 5 data after finishing

Serial number	mouse	keyboard	development tool	program	network	problem	Serial number
1	1	31	183	1	22	4.68MB	2
2	1	40	161	0.97	20	2.35MB	0
3	1	53	173	1	15	1.76MB	1
4	0	5	3	1	3	10,13MB	0

Training algorithm

Through the Apriori algorithm, we get the frequent itemsets containing 1 good training results, such as table 6.

Table 6 contains frequent item sets with good training results

'1'	mouse
'1'	keyboard
'1'	program

Conclusion and Prospect

Conclusion and result analysis

According to the results of our association rules algorithm, we can see what common students have in good training results. They have common features in three aspects of mouse, keyboard and program running times. If some students' training monitoring and monitoring data are too far away from these data, they can provide early warning to the teacher. Maybe these students are doing something irrelevant to the classroom.

Prospect and deficiency analysis

In addition to the observation data types analyzed in this article, there may be other feature types related to the good feature value 1 of the training performance, hoping to continue to find and join the algorithm model in the follow-up teaching process. In addition, the training courses considered in this paper have some limitations, which are basically suitable for the professional department of the Department of science and engineering. However, other departments such as the humanities and construction departments are not within the scope of consideration. The following can continue to study the professional training of other departments.

Reference

- [1] Li Si. Application and innovation of machine learning in intelligent education [J]. information and computer.2017, (20): 231-232.
- [2] Yu Minghua, Feng Xiang, Zhu Zhi ting. The application and innovation of machine learning in the field of AI ([J].JOURNAL OF DISTANCE EDUCATION.2017), (3).
- [3] Liu Shiyun, Sheng Zhiwei, Guo Benjun. Application of data mining in education informatization in higher education [J]. education and occupation.2009, (2).
- [4] Yuan Wanlian, Cheng Cheng. Application of data mining technology in classroom teaching evaluation in Colleges and universities [J]. computer technology and development, 2008, (11).
- [5] Chai Yan Mei, Lei Chen Fang. Summary of research on online learning behavior based on data mining technology [J]. computer application research, 2018, (5).