The Influence of Ideological and Political Education on Employment Quality of College Students based on Association Rule Analysis

Ying Zhang*
College of Marxism, Jilin Engineering Normal University, Changchun, China

*Corresponding author e-mail: zhangying@jlenu.edu.cn

Abstract. A large number of employment data of college graduates are stored in the database. It is necessary to mine a large number of unknown and useful information hidden in these employment data. This paper aims to analyze the impact of Ideological and political education on the quality of employment by using association rules. Firstly, this paper studies the calculation formula of association rules. Firstly, the concept of association rules is described. Secondly, the association rules of employment data are analyzed and designed. Secondly, the employment data of college students and the learning data of Ideological and political education of college students are mined. Finally, the employment data is analyzed and studied in detail by using software. The experimental results show that the credits of Ideological and political education are in the range of [3,4]. 38.69% of the graduates whose achievements in Ideological and political education are within the range of [3,4] belong to class II. Among the graduates whose score of Ideological and political education in 2019 is [1,2], 35.29% are not satisfied with their current situation, that is to say, their confidence in the rules is 0.3529%. It can be seen that the credit system of Ideological and political education of college students has affected the employment quality of graduates after half a year to some extent.

Keywords: Association Rules, Data Mining, Ideological and Political Education, Employment Quality

1 Introduction
Association rule is a research direction in the field of computer science and technology, and it is an important subject of knowledge discovery [1, 2]. In our life, we accumulate a lot of data every day. In such an environment, the research of data mining has made great progress, making association rules attracted the attention of various industries [3, 4]. If we can use association rules to mine the potential knowledge we need from these data, and provide targeted guidance for students' learning and life, association rules can make significant contributions to the field of education. Ideological and political education of college students is an important data to measure the quality of employment, but at present, using this data is still in the primary stage of simple statistical analysis [5]. This paper makes an in-depth analysis of the ideological and political education of college students by using the technology of...
association rules. Students can understand their own ideological and political education according to these laws, and provide reference for later learning [6-7]. Teachers can understand students’ ideological and political education according to these laws and provide suggestions for students. School administrators can find the employment quality of ideological and political education according to these laws, and provide the basis for the later adjustment of teaching arrangement [8]. It is of great practical significance to analyze the influence of ideological and political education on the employment quality of college students [9-11].

Won, dongwoo proposed an effective method to discover knowledge from transaction data by using association rules. First, we use ontology to solve this problem. The hierarchy of ontologies defines the generalization of concepts at different levels of abstraction to minimize search space. Next, we develop an effective algorithm, hierarchical Association Rule Classification (HARC), which uses a new measure called association degree to classify association rules. Therefore, users can now efficiently find the required rules by searching the compact generalization rules first, then the specific rules belonging to them, rather than scanning the entire rule list [12]. In order to improve the efficiency and interpretability of temporal association rules mining, L. Wang proposed a temporal association rules mining algorithm based on frequent itemsets tree. Based on the discretization of the reduced dimension time series data, the frequent term set is generated by vector operation, which improves the efficiency. In view of the advantages of the tree structure and the time interval relationship between items, the frequent item set tree is constructed while mining frequent item sets, which improves the efficiency of rule mining without generating candidate item sets. Experimental results show that compared with other algorithms, the algorithm has better efficiency and interpretability in mining time rules, and has a good application prospect [13].

Firstly, this paper analyzes the background and practical significance of the topic, and introduces the current situation of association rules and related applications. It also introduces the concept, method and process of association rules in detail. Secondly, the classical algorithms of association rules are summarized, and they are analyzed and discussed in detail. It focuses on the analysis of strong association rules. And use SPSS modeler data mining software to build a model of student employment data [14]. On the basis of setting the important indicators of data mining, namely support and confidence, analyze and calculate through SPSS modeler data mining software.

2 Proposed Method

Association Rules

For association rules, we first need to clarify some concepts of association rules. The rule is “if, then...”, the former is the condition, the latter is the result. To measure whether a rule is a “good rule”, we need to understand the following definitions.

(1) Item, item set and dimension of item set

Item: a field in the database. For the record of education cloud platform, it can refer to a resource in a website transaction. For example, when the customer selects “single choice question” and “ppt” when making courseware, then “single choice question” and “ppt” represent two different items. Expressed as (M = 1,2,...) N)

Itemset: a collection of several items. In the above example, {“single choice question”, “ppt”}, {“single choice question”} and {“ppt”} can be item sets. That is, the item set t is:

$$T = \{T_1, T_2, ..., T_n\}$$  \hspace{1cm} (1)

K-item set is used to represent the set with K items, and the number of items, K, is used to define the length of item set. On this basis, all parts of the content with support equal to or greater than the minimum support are called frequent itemsets.

Dimension: the number of items in an item set is called the dimension of the item set, which is recorded as |t|. A set of terms whose dimension is k is called a set of K terms.

(2) Confidence level
In order to measure the accuracy of an association rule, the concept of confidence is proposed. Confidence degree is the degree of credibility and reliability of a conclusion according to a certain condition. That is, when a set of supporting items in the set has a certain proportion of transactions supporting a set of items at the same time, the percentage ratio of the proportion is associated with the corresponding set of items, which is the confidence of this group of association rules.

(3) Support level
In order to express the probability that the union of two item sets appears in the whole transaction. The ratio of the number of transactions in a transaction set to the total number of transactions in a transaction set is called the support degree of the transaction set (for the convenience of shorthand, sup (T) is used to represent the support degree of the transaction set t).

(4) Minimum support and confidence
In order to quantify and evaluate the importance of association rules concerned by users, the concepts of support and confidence are defined for item sets, and min sup and min conf are used to represent their minimum values respectively. When k is used to represent the length of item set, an item set whose support is not below the minimum support can be called k-frequent set. The setting of the above minimum value is very important in Apriori algorithm. Wrong setting will not only make the selection rules meaningless, but also reduce the overall efficiency.

(5) Association rules
Association rules are implicit: R: XY where XT, YT, and X ∩ y =. It means that if item set X appears in a transaction, y will appear with a certain probability, and X and y are independent of each other and cannot be interchanged.

(6) Confidence and support of association rules
For an association rule, its value needs to be evaluated to reflect which association rule is more interested by users. In order to accomplish this task, support and confidence need to be defined. In association rules, the ratio between the number of transactions containing two items and the original database dimension is defined as the support degree of association rules, namely:

\[
support(X \Rightarrow Y) = \frac{\text{count}(X \cup Y)}{|T|}
\]

(2)

To a certain extent, the value of the association rules can be determined by examining the support and confidence. Generally speaking, users are concerned about the association rules with higher values in these two aspects.

(7) Strong association rules:
Finding strong association rules is the ultimate goal of association rules mining, and it can also know the root cause of business decision. Since it is called strong association rules, it naturally represents those association rules that are valuable and can help to understand the characteristics of users. After the above definition, it is easy to link the performance of association rules with the support and confidence. When the support and confidence of an association rule can be above the minimum value of both, it can be considered that it has a considerable degree of value, which is the so-called strong association rule itself.

In addition, there are two key features that can help you learn more about association rules.

1) For frequent itemsets, because frequent itemsets represent an itemset with a support degree above the minimum support degree, for an itemset, the number of itemsets of non empty true subsets must not be less than the number of itemsets, that is, the confidence of the true subsets must not be less than the confidence of the itemset, so the true subsets must also be a confidence of the minimum Set of frequent items on.

2) Similarly, for an item set, the occurrence times of the item set (i.e. the superset of the set) that takes it as the true subset must not be greater than that of the item set itself, that is to say, the confidence degree of its superset must be below that of itself. Because of this, for an infrequent item set, its superset can’t have the property of the aforementioned frequent item set.
According to the above six definitions and two properties, we can make clear the basic data needed by the association rule algorithm and what the calculation goal of the algorithm is. The strong association rules obtained by the algorithm can guide the education cloud platform to make plans to a certain extent, understand the actual teaching situation of teachers and the use of students, and facilitate the further promotion of the website [15].

3 Experiments

3.1 Data Acquisition
According to the research and design scheme of students' employment data, this paper uses the principle of relevance rules to process the employment data, and forms a database of major, gender, ideological and political education learning achievement and employment destination of the employed students, as shown in Table 1.

Association rules study the implementation of students' employment, and discretize the data in each database. Specifically, divide the data in the database into several data segments, and then click the OK button on the discreteness interface. Under normal circumstances, all the information in the discreteness data segments will be displayed [16]. If there is no display, you need to readjust the data segment information until it is displayed normally, and then click again Run the button, and the system will automatically jump out of the window after running. The upper part of the window after running is the frequent item set of support threshold [17].

Table 1. Basic information of employed students

| Essential information | Data attribute | length | Primary key | Not empty |
|-----------------------|----------------|--------|-------------|-----------|
| Full name             | char           | 8      | Y           | Y         |
| Credit points of Ideological and political education for College Students | char           | 2      | Y           | Y         |
| Employment satisfaction | char           | 10     | Y           | Y         |
| Employment income     | char           | 20     | Y           | Y         |

3.2 Test Steps
Based on the SPSS modeler software, the data of College Students' employment is mined, and two related analyses are carried out. The first correlation analysis selects the satisfaction of college students with the current situation as the next item, and takes the score of Ideological and political education credits of college students as the first item; the second correlation analysis takes the monthly income of college students after graduation as the second item, and takes the satisfaction of college students with the current situation as the second item and takes the credit point of Ideological and political education of college students as the front item.

Specific steps: import the employment database composed of students' academic performance, psychological quality and employment destination into the source file of SPSS modeler data mining software for data preprocessing, select the Apriori algorithm in SPSS modeler data mining software, set all parameters, and then carry out operation and analysis.

4 Discussion

4.1 Association Rules between Current Situation Satisfaction and Credit Points of Ideological and Political Education
First, the first association analysis is carried out. The minimum confidence threshold is 30%, and the minimum support threshold is 0.4%. After redundancy processing, the remaining three association rules are shown in Figure 1.
Figure 1. Rules for the connection between credit score points of Ideological and political education and their satisfaction with their current situation

The promotion degree of the three association rules is greater than 1, that is, when students have ideological and political education credit scores, compared with other scores, their satisfaction with their own status is higher, and the higher the promotion degree, the more valuable the association rules are.

The first association rule: Ideological and political education credit = [3,4] → whether satisfied with the status quo = dissatisfied support degree is 0.1633, that is, when the average score of 2019 graduates is in the range of [3,4], 16.33% are dissatisfied with the status quo; its confidence degree is 0.3668, that is to say, 36.68% of the graduates whose ideological and political education achievements are in the range of [3,4] are dissatisfied with my status quo.

The reason why the credit score of Ideological and political education does not meet the current situation in the range of [3,4] may be that the requirements are too high. People with good scores hope to find a good job or a good way out, but their personal feelings are not very good, so they are not satisfied with the current situation. Therefore, in the process of employment guidance for graduates, students of different levels can be cultivated in batches according to their performance, and guidance, persuasion and guidance for students whose ideological and political education credits are in the range of [1,2] and [3,4] can also be properly strengthened to enhance their satisfaction with the current situation.

4.2 Association Rules between Monthly Income and Credit of Ideological and Political Education

When the minimum support threshold is 0.004 and the minimum confidence threshold is 0.1, the promotion of seven association rules is greater than 1. After redundant data processing, association rules are not reduced. There are 10 association rules before and after redundancy, indicating that there is no redundancy rule for this data.

On the whole, the overall support is high. It can be seen that the credit score of Ideological and political education of graduates has a certain impact on the monthly income of graduates after half a year. It can be seen that the lower the credit score of Ideological and political education of graduates, the lower the monthly income of graduates after half a year. On the other hand, it also shows that our school should provide students with necessary ideological work when carrying out student work, and properly communicate with students with poor achievements in Ideological and political education, so that they can cope with the more severe employment situation.

Table 2. Association rules between average grade point and monthly income half a year after graduation

| rule | The aforesaid                              | Support degree | Confidence degree | Promotion degree |
|------|-------------------------------------------|----------------|-------------------|------------------|
| 1    | Score point of Ideological and political education = [1,2) | 0.0049         | 0.1765            | 1.4292           |
Score point of Ideological and political education = [1, 2)
Score point of Ideological and political education = [2, 3)
Credit point of Ideological and political education = [3, 4)
Credit point of Ideological and political education = [3, 4)

5 Conclusions
This paper introduces the concept, method and process of association rules in detail. Secondly, the classical algorithms of association rules are summarized, and they are analyzed and discussed in detail. On the basis of a questionnaire survey of graduates from an independent college, through relevant analysis, we draw the following conclusions: graduates with average credits of [3,4] are generally dissatisfied with the current situation; the lower the average credits, the lower the monthly income of graduates six months later. When providing employment guidance for graduates, we can train students of different levels in batches according to their performance, properly strengthen the guidance, persuasion and guidance for students with average credits between [1,2] and [3,4], improve their satisfaction with the current situation, and make them better cope with the severe employment situation.

Acknowledgements
This work was supported by Scientific research project of Jilin Institute of Higher Education (number: JGJX2019D199).

References
[1] Wen-Yang Lin, You-En Wei, Chun-Hao Chen. GIAMS: A generic approach for mining indirect association rules in data streams[J]. Intelligent Data Analysis, 2017, 21(1):S177-S194.
[2] Zabri Z, Nohuddin P N E. Analyzing Network Intrusion Behavior of Packet Capture Using Association Rules Technique: An Initial Framework[J]. Journal of Computational and Theoretical Nanoscience, 2018, 24(3):1628-1632.
[3] M. Özgür Cingiz, G. Biricik, B. Diri. ARNetMiT R Package: Association rules based gene co-expression networks of miRNA targets[J]. Cellular and molecular biology (Noisy-le-Grand, France), 2017, 63(3):18.
[4] Sachi Nandan Mohanty, E. Laxmi Lydia, Mohamed Elhoseny, Majid M. Gethami Al Otaibi, K. Shankar, Deep learning with LSTM based distributed data mining model for energy efficient wireless sensor networks, Physical Communication, 2020, In Press.
[5] Addi Ait-Mlouk, Tarik Agouti, Fatima Ghrami. Mining and prioritization of association rules for big data: multi-criteria decision analysis approach[J]. Journal of Big Data, 2017, 109(2):272-278.
[6] Frederick Parente, John-Christopher Finley. Using association rules to measure Subjective Organization after Acquired Brain Injury[J]. Neurorehabilitation, 2018, 42(1):9-15.
[7] Zhang Y, Xie Y M, Chen C, et al. [Association rules analysis of Fufang Kushen injection in combination with traditional Chinese medicine or modern medications in treating malignant tumor: real-world retrospective study].[J]. China Journal of Chinese Materia Medica, 2017, 42(2):378.
[8] Z. Xu, C. Zheng, Y. Zhang. Operation Optimization of Flue Gas Desulfurization System in
Power Plant Based on an Improved Association Rules Algorithm[J]. Proceedings of the Csee, 2017, 37(15):4408-4414.

[9] Zhang Z, Guo C. Association Rules Evaluation by a Hybrid Multiple Criteria Decision Method[J]. International Journal of Knowledge & Systems Science, 2017, 2(3):14-25.

[10] Dong-Pil Shin, Young-Jun Park, Jongwon Seo. Association Rules Mined from Construction Accident Data[J]. Ksce Journal of Civil Engineering, 2017, 22(1):1-13.

[11] Anh Tran, Tin Truong, Bac Le. Efficiently mining association rules based on maximum single constraints[J]. Vietnam Journal of Computer Science, 2017, 4(2):1-17.

[12] Won D, Mcleod D. An efficient approach to categorising association rules.[J]. International Journal of Data Mining, 2017, 4(4):309-333.

[13] L. Wang, S.-L. Li, P.-P. Xu. Temporal association rules mining algorithm based on frequent item sets tree[J]. Control & Decision, 2018, 33(4):591-599.

[14] Ahmed M. Eassa, Mohamed Elhoseny, Hazem M. El-Bakry, Ahmed S. Salama, NoSQL Injection Attack Detection in Web Applications Using RESTful Service, Programming and Computer Software, 44(6), pp 435–444.

[15] Krishnaraj, N, Elhoseny, M, Lydia, EL, Shankar, K, ALDabbas, O. An efficient radix trie - based semantic visual indexing model for large - scale image retrieval in cloud environment. Software Practice and Experience, 2020; In Press.

[16] E. Laxmi Lydia, Joshua Samuel Raj, R. Pandi Selvam, Mohamed Elhoseny, K. Shankar, Application of discrete transforms with selective coefficients for blind image watermarking. Transactions on Emerging Telecommunications Technologies, 2019, In press.

[17] Xiaohui Yuan, Daniel Li, (2018) “Deepankar Mohapatra, Mohamed Elhoseny, Automatic Removal of Complex Shadows from Indoor Videos Using Transfer Learning and Dynamic Thresholding", Computers and Electrical Engineering, 70, pp. 813-825."