Journal of Information Technology Review

Print ISSN: 0976-3511
Online ISSN: 0976-2922

DLINE JOURNALS

JITR 2024; 15 (1)
https://doi.org/10.6025/jitr/2024/15/1/7-13

The Learning Platform System of Theory and Education Based on Association Rule Learning Algorithm

Xiaojuan Chen
School of Marxism, Chongqing Vocational College of Transportation
Chongqing, 402247,China
{cxj3408@163.com}

ABSTRACT

Integrating the traditional advantages of promoting theory and education courses and information technology is necessary for educational development. Achieving true "teaching based on individual differences" requires the scientific application of educational big data. This paper uses association rule mining technology to start from the demand for student task point learning management in the current personalized learning platform system. It analyzes the current status and research trends of personalized learning and association rules. Then, it discusses the principles, advantages and disadvantages of the classic Apriori algorithm and proposes an improved algorithm. Finally, the feasibility and actual effects of the algorithm are verified through simulations on the learning platform system, providing certain assistance for future research on theory and educational courses.

Keywords: Association Rule Learning Algorithm, Data Mining, Education, Marxist Theory, Learning Platform System

Received: 26 September 2023
Revised: 4 December 2023
Accepted: 14 December 2023

1. Introduction

With the popularization of big data, we are in a critical period of transformation. Therefore, educators in universities must actively embrace this transformation and combine it with existing traditional advantages to better meet the needs of current social development. Based on this, a more creative big data thinking mode conducive to social development should be constructed. This can help us better understand the essence of education and more effectively carry out related preaching and training. With the development of science and technology, data mining has become an important tool that can help us discover valuable information hidden behind massive amounts of information. It uses advanced mathematical algorithms and computer science techniques to support our decision-making. However, due to the existence of massive information, traditional data mining techniques still have shortcomings. The Apriori algorithm is considered to be a very important tool that can help us find relationships between different events to anticipate future developments better. It decomposes the information content into multiple subsets through pruning and other cleaning operations for quick discovery and analysis. With the continuous improvement of science and technology, some new calculations have been developed to greatly reduce the search range and shorten *the search process, thus greatly improving search efficiency. The development of association rules has become a hot topic in data analysis today and has been deeply integrated into daily human life. By using association rule mining technology, we can extract valuable content from a large amount of data sets and better understand future trends, such as students' learning background and achievements at graduation. This information can be applied to college graduation employment planning and management to achieve better results.*

2. Related Work

Agrawal and other researchers initially regarded association rule mining as a significant technology aimed at gaining in-depth insights into consumer needs by exploring the relationships between different types of products. This helps companies better manage and monitor purchasing efficiency, leading to more efficient sales [4]. In recent years, an increasing number of algorithms have been developed to enhance the efficiency, flexibility, real-time performance, and universality of rule mining algorithms. These include hierarchical mining, incremental updates, distributed and parallel mining, multi-level, multi-value, concept lattice mining, and more [5]. In the book "Reimagining the Role of Technology in Education" of the U.S. National Educational Technology Plan in 2017, it was proposed that mobile data collection tools and network collaborative platforms bring personalized learning approaches to students, providing them with more opportunities. Furthermore, the book highlighted that personalized learning will be one of the core areas of effective leadership in future development, which can help students achieve better teaching results through collecting and integrating formative and summative assessment statistics during individualized digital teaching processes [6]. With the introduction of the "New Generation Artificial Intelligence Development Plan," China is vigorously developing an education system based on intelligent learning. Through teaching interventions and decisions, it aims to promote innovation and development in teaching methods and advance the sustainable and healthy development of higher education [7]. With the advancement of artificial intelligence technology, collecting and utilizing data analysis from different sources have become necessary steps for personalized learning, and maximizing the value of these data analyses is crucial for achieving this goal. Currently, research on personalized learning has evolved from traditional dimensions, such as personal information collection, preferences, and interests, to more comprehensive levels [8], and continues to explore and improve. By using artificial intelligence systems in schools and combining them with the latest research in the education field, such as big data analysis and deep neural networks, it is possible to better track students' behavior and accurately predict their academic performance. By combining the characteristics of association rule algorithms, we can better apply them to education courses at universities and more effectively utilize knowledge in mathematics and computer science [9]. Furthermore, we can enhance educational outcomes through data mining, such as presenting information better and conducting more accurate behavior predictions. By combining traditional education with the latest big data technology, we can make better predictions, comprehensively grasp the needs of different periods, achieve overall educational effectiveness, accurately understand the relationship between them, and conduct practical operations more effectively [10].

3. Association Rule Mining Technology

3.1. Generation of Association Rule Learning Algorithms

In a frequent itemset, if itemset I contains multiple items, such as tennis racket, tennis ball, sports shoes, and shuttlecock, the rule can be used to describe the interactions between these items. We can use a more complex rule to describe such interactions, for example, in an itemset, if the XY of an item is 3, then the D of this rule is 6, and its XY will result in a higher confidence. If α*=0.5,* β*=0.6, we can assert that the selection of a tennis racket and the selection of tennis service have a close mutual influence.*

$$Confidence\ (X \Rightarrow Y) = (P \mid X) = \frac{Support\ (X \cup Y)}{Support\ (X)} \times 100\% \qquad (1)$$

By applying the above methods, we can formulate a series of corresponding association rules:

Step 1: Select appropriate frequencies from the availability levels of D.

Journal of Information Technology Review Volume 15 Number 1 February 2024

Step 2: Use the similarity between these frequencies and availability levels to construct corresponding relationships.

3.2. Apriori Algorithm

The Apriori algorithm is considered the first association rule mining algorithm and is regarded as a very effective method. This algorithm can extract valuable content from a large number of association rules and effectively mine the associations between these rules. To achieve this goal, the Apriori algorithm follows two main steps: firstly, extracting valuable content from the candidate set, and secondly, conducting in-depth evaluation of this content to determine the effectiveness of the association rules. The Apriori algorithm is typically seen as a one-dimensional, multi-level classification algorithm that uses multilevel iterations to achieve effective recognition of multiple sub-objectives. As shown in Figure 1, it is named Apriori. The core idea of this algorithm lies in discovering the similarity between multiple sub-objectives and combining them to achieve effective recognition of multiple sub-objectives.

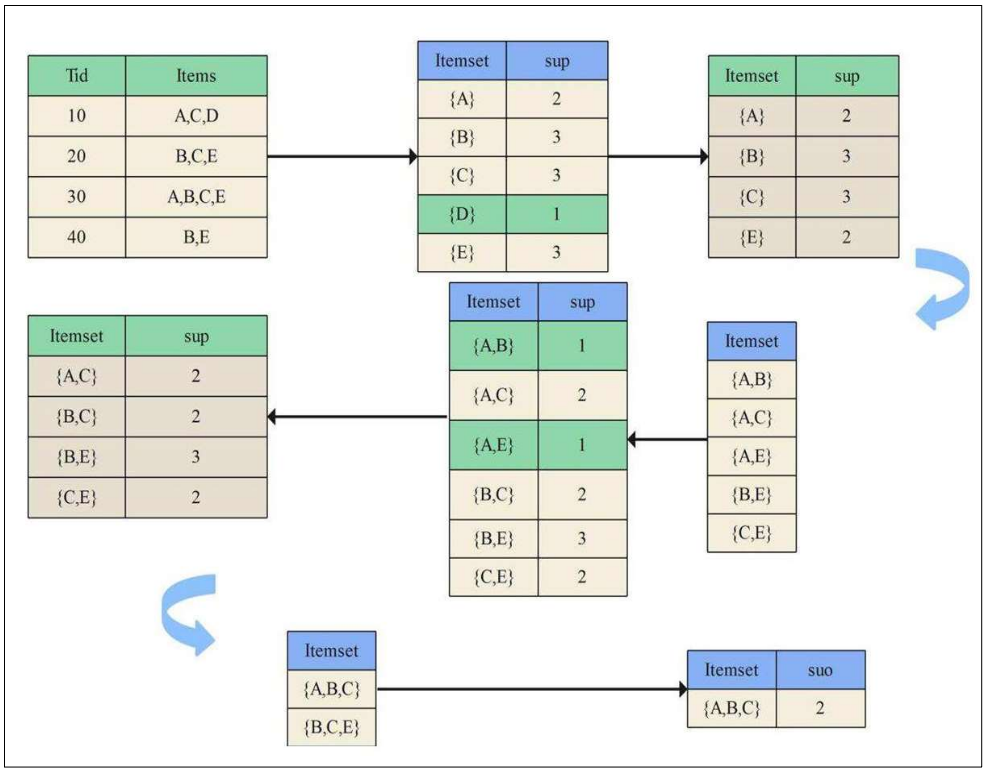

Figure 1. Apriori Algorithm Example

In some cases, we can search for frequent itemsets through database scanning. First, we can find L, then use it to find two Ls, and finally, we can continue the search until we find k frequent itemsets. This way, we can determine that the support of each Lx exceeds the minimum threshold. The Apriori algorithm has been widely applied due to its simple and understandable operations. Its association rules are based on multiple repetitions of item combinations, ensuring that their support meets the expected standards and has generality and credibility. Although this approach has many advantages, such as significantly reducing the number of database system scans, reducing the formation of intermediate itemsets, improving the reliability of individual items, extending the algorithm's scalability, and enhancing its operability, it still faces challenges such as heavy I/O load, inability to meet large-scale data processing requirements, and lack of high reliability.

3.3. Improvement Methods

Several optimization techniques have been developed for the Apriori algorithm. For example, dynamic counting of itemsets involves checking, evaluating, and filtering the candidate sets of *each itemset to determine which itemsets meet the requirements and which ones do not, thus achieving fast and accurate evaluation of Apriori. The specific approach is as follows:*

(1) *Statistically determine the minimum support itemsets among those containing a single element to construct a one-dimensional frequent itemset Lt.*

(2) *Repeat this process until no more complex frequency itemsets are generated. For example, in step k, generate k-dimensional candidate itemsets based on k-1 frequent itemsets generated in step k-1. By performing p partial classifications, we can obtain a classifier containing p rules. When the test set and the training set data are identical, the misidentification rate can be calculated as follows: [Note: The content afterÿ" is missing and needs to be provided for a complete translation.]*

$$error\ (r_i) = 1 - Conf\ (r_i) \qquad (2)$$

The Apriori algorithm can help us determine which k-dimensional frequent itemsets have been accurately collected, thus avoiding duplicate searches and improving search efficiency. It allows us to filter out those k-dimensional frequent itemsets that do not meet the requirements, thereby enhancing search efficiency. The number of instances covered by rule riis ri.Sup x |Di| / ri.Conf, and the number of misidentified instances is (1 - ri.Conf) x ri.Sup x |Di| / ri.Conf. Therefore, the overall classifier misidentification rate is: [Note: The content after " Therefore, the overall misclassification rate of the classifier is missing and needs to be provided for a complete translation.]

$$error = \sum_{i=1}^{p} \left[(1/Conf(ri)-1) \text{ x } sup(r_i) \text{ x } |D_i|\right] \qquad (3)$$

4. Design of Experiments and Result Analysis

4.1. Design of Experiments and Data Collection

This article uses a student learning behaviour record database of education courses for 19-level computer science students in vocational and technical colleges to extract possible correlations between chapters and then obtain functional correlations between knowledge points and exam scores. Finally, association rule mining algorithms find the relationship between students' knowledge points and exam scores. There are several main processes. (1) Collect students' learning records and convert them into a searchable database of learning outcomes. (2) Extract data mining objects from the learning results database, compile code, and convert the relationship list into a business processing database. (3) For business databases, generate frequent item sets based on a given minimum support level, and then combine them with a given minimum confidence level to generate association rules. Data mining is based on the provided objects in the code table and simplified relationship table in the learning log database, including the number of students, school hours, teaching situation, exam scores, and learning method data. Convert the relationship table into a corresponding transaction library and write the code. For example, selecting a portion of students' "Database Principles" as a transactional database, and using a coding table to convert corresponding items to good-graded chapters. After generating transactional database D, | D |=9, K={K01, K02, K03, K04, K05, K06, K07, K08, K09}. Assuming a given minimum support of 0.25, use the Apriori algorithm to obtain all frequency term sets in D. After generating a frequency itemset, for any K types of frequent itemsets, search for all possible true subsets and calculate the corresponding rule reliability. When it exceeds the given minimum value, the rule is output until it is finally connected to the k-1 itemized set.

4.2. Result Analysis

As shown in Figure 2, when the minimum confidence threshold is set to 0.75, the rule is K01 $\wedge$ K08 $\rightarrow$ K03. Specifically, when studying Marxist Leninist education courses, if students achieve good grades in the "Introduction" and "Database Programming" chapters, the grades in the "Relational Database Standard Language SQL" chapter will also be better. Therefore, in personalized learning systems, it can be considered to place "Database Programming" before "Relational Database Standard Language SQL". Alternatively, when students choose to learn the "Standard Language for Relational Database SQL", the personalized learning system may suggest that they first review and consolidate the content of the "Introduction" and "Database Programming" chapters. When *the union value gradually increases to 100%, the confidence threshold will increase and decrease. Therefore, when students with "academic" Learning styles choose learning content, the personalized learning system will give priority to recommending theoretical text-based learning materials for students. When students with "operational" Learning styles choose learning content, the system will prioritise presenting video animation-based learning materials for students.*

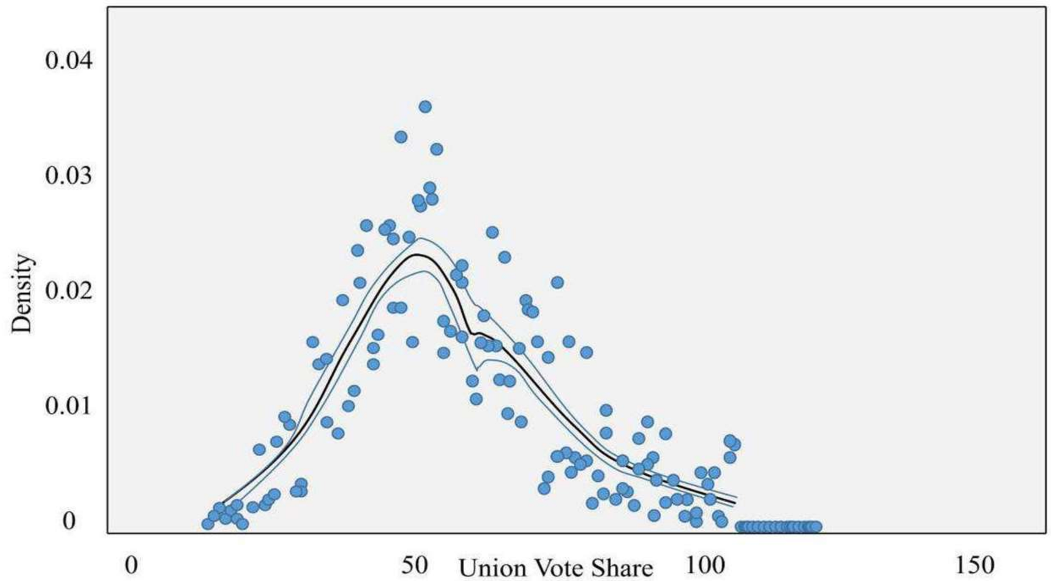

Figure 2. Confidence of Association Rules Generated from Transaction Database

Through the Apriori algorithm, we can explore the associations between course contents in depth and discover how students with different learning styles utilize the best learning strategies to improve learning effectiveness. The experiments have demonstrated that optimising the Apriori algorithm can significantly enhance computational efficiency, and the provided rules can effectively assist teachers in learning supervision and guidance. By studying the vertical layout dataset based on bit mapping, this paper aims to investigate the learning efficiency of the CAAR (Classification Model for Association Rules) and distributed computing. To obtain more accurate results, this paper used 33.2 seconds, instead of the original 33.2 seconds. By leveraging the method of bit vectors, we can partition and efficiently transmit the data to other locations. A new algorithm can be employed to construct a more effective classification system, which automatically adjusts algorithm parameters according to different requirements, thereby achieving more efficient completion of classification tasks. Using Master/Slave technology, we can synchronize the processing of multiple datasets in a single Agent, thereby constructing an efficient multi- Agent ensemble. To achieve this, we simulated the working states of different Agents to explore the additional time they require to extract and process similar information, including communication and information aggregation. In this experiment, we set minsup and minconf as evaluation criteria, with minsup being 0.01 and minconf being 0.50.

Based on the experimental results in Figure 3, we found that if the number of agents is sufficient, the processing efficiency of classification association rules will be significantly improved. For example, if the Agents reach 6, the processing efficiency will decrease to 188 seconds. However, if the number of agents exceeds 6, even if processed more frequently, it will result in additional processing costs, which will not improve processing efficiency. With the continuous expansion of agents, even in the case of 6, the data mining efficiency of this dataset cannot be significantly improved.

5. Conclusions

In recent years, due to the rapid popularization of artificial intelligence technology, many new data have come from various types of models and algorithms. Among them, the Apriori algorithm is one of the best, as it can provide accurate and reliable models and help us quickly identify and predict future trends. Therefore, the Apriori algorithm has proven to be a suitable model for Marxist Leninist courses. It can help us quickly identify and predict future trends, and its reliability and manoeuvrability are also very strong. Introducing the Apriori algorithmperforms well in supervised image content analysis and large-scale data mining. Its characteristics include fast computation, excellent analysis accuracy, clear rules, compact rule specifications, and convenient model operation, thus enriching future research fields.

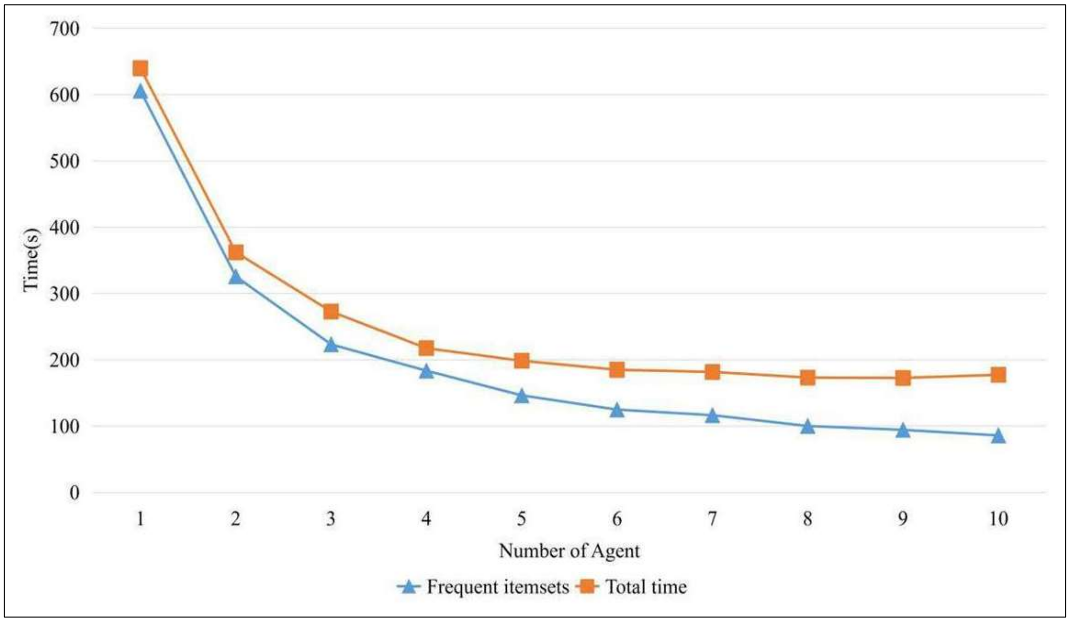

Figure 3. Number of Agents and Mining Time

References

[1] Mao, C., Wang, Q., Feng, K., et al. (2021). Research on the Mechanism of Multi-resource Mapping Matching and Transmission of Road System Based on Meme Theory. *Journal of Physics: Conference Series*, 1838(1), 01-20.

[2] Zhang, H., Padua, S. A., Li, Y. (2021). Research on the Design of Preschool Education Management Information System Based on Computer Technology. *Journal of Physics Conference Series*, 1915(2), 02-03.

[3] Xu, J. (2022). Research of Online Courses Construction Based on the Theory of Interaction. *Journal of Higher Education Research*, 3(1), 26-28.

[4] Jianhua, W., Yuanrui, Z.,Jianbin, Z. (2021). Research on Ideological and Political Education of E-commerce Course for Marketing Major. *Education Study*, 3(1), 50-54.

[5] Nakabayashi, K., Morimoto, Y. (2023). Design and Implementation of a Distributed Multi-Platform Learning Environment Based on an Extensible Learning Support System Architecture. *Transactions of Japanese Society for Information and Systems in Education*, 40(1), 28-41.

[6] Li, H., Dong, X. (2022). Construction of Online Ideological and Political Education Platform Based on Artificial Intelligence Technology. In International Conference on E-Learning, EEducation, and Online Training (pp. 129-144). Cham: Springer Nature Switzerland.

[7] Zhang, Y. (2021). Research on Ideological and Political Education of College Students from the Perspective of System Theory. *Frontier of Higher Education*, 2(2), 1-17.

[8] An, N., Wang, J., Wang, H. (2018). Research and Practice on Innovative Methods of Ideological and Political Education for College Students Based on Internet of Things+ Technologies. *Educational Sciences: Theory & Practice*, 18(5).

[9] Wang, Y. (2020). Analysis on the Construction of Ideological and Political Education System for College Students Based on Mobile Artificial Intelligence Terminal. *Soft Computing*, 24(11), 8365-8375.

[10] Wu, X. (2022). Research on the Reform of Ideological and Political Teaching Evaluation Method of College English Course Based on "Online and Offline" Teaching. *Journal of Higher Education Research*, 3(1), 87-90.