



## Using Association Rules based Apriori Algorithm for Student's Learning, Behavior, and Activity

<sup>1</sup>Andino Maselena, <sup>2</sup>Miftachul Huda, <sup>1</sup>Chotirat Ann Ratanamahatana\*

<sup>1</sup>Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Thailand

<sup>2</sup>Faculty of Human Sciences, Universiti Pendidikan Sultan Idris, Malaysia

\*Corresponding author email: chotirat.r@chula.ac.th

**Abstract:** This research aims to determine the tendency of a student to answer questions related to learning, behavior and activity. Data from the results of this study can be developed to assess a student's creativity. In this study, we used a questionnaire containing 30 questions to find out about students' learning, behavior, and activity. The questions in this questionnaire are related to multicultural education, with scores between 0 - 10 for the answer option. The questionnaire answers from 10 respondents was used to identify several types of association rules related to the answers data from the questionnaires given to the students, namely the values of support and confidence, knowing the combination of items that students often answer. Calculations were performed using association rules based on the Apriori algorithm. From the results of this research, a set of rules is obtained that describes the tendency of a student to answer questions related to learning, behavior and activity.

**Keywords:** student, learning, behavior, activity, association rules, apriori algorithm

**Received:** 16 March 2024

**Revised:** 12May 2024

**Accepted:** 12 June 2024

### 1. Introduction

The worldwide range of the prediction process on the learning behavior is important to recognize the potential value to transmit along with the entire process of driving the significant attribution about synthesizing different types of information [1]. It is necessary to highlight the wide range of efforts to give insights into improving knowledge and skills among the students. This relates to the learning styles, which play a pivotal value in determining the outstanding role in directing the development of acquiring knowledge. The particular stages derive from modifying and reinforcing the existing knowledge, behaviors, skills and values [2]. In the attempts to assess the ability to learn related to concern about the core elements in the learning process, the effort to organize the instruction in the classroom basis and borderless space should be served professionally along with traditional basis and electronic devices in supporting the effectiveness and efficiency in the learning process [3]. Incorporated with an organized learning style, the wide range of maintaining self-regulation is supposed to foster learning enhancement with an effective way of performing the planning prepared to meet the challenges [4]. Set out amongst the students to make them get well in the attempt to see the potentials of the classroom dynamics, creating a conducive classroom atmosphere would potentially help this process point out actual time on settling tasks in the learning activities through an effort to restore the motivational substance in maintaining the learning procedure effectively.

Regarding dealing with the necessary tasks to achieve progress and advancement, the motivation basis in learning needs to establish the required attainment in giving insightful value along with emphasizing the psychological and biological basis into the integrated ranges in directing the quality behavior [5]. In

particular, the higher level of expertise the number of users achieves will increasingly enhance their motivational substance. Thus, their learning enhancement also seems more advanced in understanding the learning process and complying with the fundamental components [6]. However, those with a low motivational score are less likely to be interested and need more practice and training to ensure they can obtain the best result. To see this attainment, the assessment model is required to address the achievement based on the process. In this view, the assessment model has several comparisons with the previous section on cognitive learning styles. There is an opportunity to combine or re-use the results from that section to enhance the findings here. Unlike the Creativity scale, it is more plausible that a user may rate his/herself as a 0 or a 10 on this scale, although this is still highly unlikely. A user who rates themselves as a perfect 0 will not benefit from the course and indicate no interest in a career. It is doubtful that a user would rate themselves as a 10 could be an entrepreneur running a company focused on a professional job.

## **2. Student's Learning, Behavior, and Activity**

Enhancing learning through acquiring knowledge, critical thinking, and pragmatic skills is part of continuing experiential life expertise [7]. When attempting an overall process to approach learning, the variety of such attempts must commit to possessing outstanding human characteristics engaged with the advancement and complete skill sets [8]. Continuously activated into the learning process, the extent of individual life experiences refers to enhancing the significant way with the ability to analyze what happens in the learning process to respond to environmental life circumstances [9]. It enables them to achieve individual transformations into various aspects, such as physical and intellectual bases, to transmit into the behavioral substance [10]. In terms of transforming the particular expertise to ensure skillful engagement, strengthening the learning experience by giving them the knowledge and understanding with acquiring skills and attitudes should be considered to improve the way to solve problems widely [11]. The emphasis on the training basis could also be determined in the learning process with its particular principles to improve their skillful potential and maximize the utilization of devices designed within the learning environments [12].

Along with enhancing the potential of learning behavior and activity within the learning environment, this attempt could enable the learners to understand their way, leading to building their attitude on a learning basis. Regarding further inquiry about the influential basis to improve problem-solving skills and mindset [13], a comprehensive approach to learning should be incorporated internationally with the guidelines to emphasize the instructions process set out in the program and training [14]. The concern on how to learn properly integrated into fundamental principles to obtain the opportunity should be well cooperated with carrying out the wise approach with an adept task in gaining information and knowledge inquiry. This enhancement should be carried out with networking influence to support the learning progress in creating the circumstance within a supportive setting [15]. With the significant role in supporting the learning process, learning activities incorporated into an environmental basis to build cultural recognition and awareness must be combined with carrying out the tasks carefully and observing the lectures related to the subject content. These potential stages need to work with constituting personal development to enhance self-awareness with in-depth understanding towards improving the quality process with further attention to achieve the ability to manage the ethical and professional attribution to give insights into handling challenges [16]. Moreover, the quality of classroom experiences is the core point of extensive behavior settling within schooling and training, which is fundamental to improving the holistic learning process in acquiring knowledge, behaviors, skills, and preferences [17]. Learning activities with the processes on successive stages involved in fitting the learning environment in supporting the quality of self-advocacy and self-regulation should be complied with, emphasizing paying attention to the fundamental values within educational principles.

In transmitting the incorporated skills in facilitating the training and development programs, this attempt, combined with cultural recognition and spiritual self, needs to appropriately manage in a particular way with the principles of mechanical aids and rules [18]. The worldwide variety of skills and

subjects customized and personalized into learning basis has to be entirely engaged with the underlying major components of resolving complex problems to settle on the particular guidelines as an essential enhancement of moral purpose [19]. Through a holistic approach, strengthening the rightful intention with ethical engagement should be applied in sustainable learning with comprehensive perseverance in sustaining the purification of self-regulation in transmitting inward and outward elements of the learning process [20]. The achievement of the way to play a key role in strengthening such processes like physical transformation and building the spiritual in the contextual basis to support the self-discipline needs to obtain working hard through a wide range of planning basis, observing, and evaluating the phenomenon derived amidst their circumstance with creativity and self-control [21]. Both ethical and technical aspects, along with the social context to help increase the quality of self-discipline, should be emphasized to guide the learning process in a particular way [22]. As a result, this process would lead to the potential behavior of obtaining self-discipline in sustaining the extent of self-motivation and willingness to have such value of responsibility awareness. This encouragement refers to adopting the potency of independent thinking associated with the spiritual fortitude entirely engaged in self-reliance and self-assurance to help the learners possess skillful and educated individuals with the level of self-criticism assigned to commit to carrying out duties comprehensively in the teaching and learning process.

### 3. Association Rules

Association rules are a specific or one of the existing techniques in data mining that states a correlation relationship between the emergence level of multiple attributes in a database. As a data mining technique for finding associative rules between a combination of items, association rules can be derived from various data sources, including data transactions, data warehouses, and other information stores [23]. In general, processed data is homogeneous. The first study of association rules search is obtained from the items that often appear together. One of the most commonly used algorithms for association rules search is Apriori. The importance of an association rules can be identified by two parameters, namely *support*, i.e., the percentage of item combination events or the *support count* of the number of *items* that appear in an aggregate of transactions, and *confidence*, in example the strength of relationship between items in the association rules [24]. An association analysis defines a process for finding all associative rules that meet minimum requirements for *support* (*minimum support*) and minimum requirements for *confidence* (*minimum confidence*). *Support* is a measure that shows the level of dominance of an item or itemset over the entire transaction. *Support* is the percent of records that contain one or more combinations of items compared to the total number of records. *Confidence* is a measure that shows the relationship between two items conditional. *Confidence* is a percentage of records that contains a combination of items compared to the total number of transactions that select the item [25].

### 4. Association Rules for Student's Learning, Behavior, and Activity

The first step is to make the required adjustments to the input data. In this case, since all answers can be the same for all questions, adjustments should be made to differentiate if the same answer occurs for each question. The adjustment is adding the question index number before each answer. As an example: for question 1 with answer 0, then the data value is 100; for question 1 with answer 10, then the data value is 110; for question 2 with answer 10, then the data value is 210; for question 10 with answer 10, then the data value is 1010. The next step is determining the parameter percentage value of *support* as a determinant of an answer that can be said to be 'often answered'. In this case, the *support* percentage value is 0.3. Later, this support value will be multiplied by the number of respondents, which is worth 10. So  $0.3 * 10 = 3$ , which means an answer is said to be answered frequently if answered thrice in all questionnaire answer data. Then, a minimum and maximum limit of answers will be set to be counted. In this case, the minimum and maximum values are 2 and 4. It means that the computed answer is the set of answers with two answers, three answers, and four answers. The next process is to determine the itemset that occurs in all the answers to the questionnaire. The process is as follows: Calculate the minimum value of the answer data that is often answered. The formula used is the percentage of *support* value

multiplied by the number of respondents. In this case, it means that  $0.3 * 10 = 3$  times. This variable will be used in the calculation of number 3. Calculate the number of each answer on all questions. All answers to the questionnaire that have been adjusted will be counted as the number of answers that appear in all the questionnaires. As an example, data 100 appears as many as three times, that is in the second, seventh, and ninth respondents; data 101 appears as many as two times, which is the fifth and eighth respondents; data 102 appears only one time, which is in the first respondent. And so on for each questionnaire's answer data.

Based on the calculation, find all data that fulfill the minimum support value calculated at number 1, i.e., three times. From each of these data, arrange the itemset for all combinations that can be found. In this case, the combination size used is within the range of 2 to 4 combinations. For each itemset with the combination, also calculate the number of occurrences in the questionnaire answer, and if the number of occurrences of the combination meets the minimum *support* limit, enter it as the candidate response item. For example, the calculation starts by combining data from two answers. Concerning the election results, the first combination is 100 and 204. If we look at the questionnaire data table of responses to the questionnaire, then this combination is only found in the seventh respondent. Since the number is only one occurrence, this combination does not meet the minimum value of *support* worth three times, so this combination is not a candidate answer itemset. The second combination is 100 and 304. If we look at the questionnaire data table adjustment, this combination is only found in the ninth respondent. Since the number is only one occurrence, this combination does not meet the minimum value of *support* worth three times, so this combination is not an itemized answer. Thus, the calculations performed on all combinations of 2 data can be found. The first successful combination of 2 data is a combination of 705 and 1203. This combination is found in the third, sixth, and eighth respondents. Since this combination appears three times, it is included in the itemset answer.

After the combination of 2 data is complete, then proceed to the combination of 3 data and the combination of 4 data. Record all combinations included in the itemset answer. The next step is to determine the rules for each item set that has been found. However, when entering the rule calculation, the next step is to determine the parameters of the percentage value of *confidence* as the determinant factor of the truth level of the relationships between the answers in itemset. In this case, the percentage value of *confidence* is 0.8. This means the relation will be included as an answer if its truth level on all questionnaires is above 80%.

## 5. Results and Discussion

The process of determining the rules is as follows: divide each item set into two parts, for example, the cause clause and the resultant clause of the candidate rule. Find all possible combinations of the cause and effect clauses that can be found from each item set. For each candidate rule, calculate the value of the candidate's rule *confidence* in the data table of questionnaire answer adjustment (Table 3). Note the comparison between the exact number of rules and all the answers to the questionnaires that have the data on the cause clause. As an example: "If 705 then 1203", by looking at the data table of questionnaire answer adjustment, then the data 705 is on the third, sixth, and eighth respondents. After looking at the three respondents, all three also enter data 1203.

So the value of its *confidence* is for the first itemset (705, 1203); the first candidate rule is obtained from the number of exact rules divided by all answers to the questionnaire that have the data in the cause clause, which is three divided by three equal to 1. The second future rule is obtained for the first item set (705, 1203): "If the 1203 is 705". By looking at the data table of questionnaire adjustment (Table 3), 1203 data are in the second, third, sixth, and eighth respondents. After looking at the four respondents, only the third, sixth, and eighth respondents entered data 1203, while the second respondent entered data 706. So, the value of its *confidence* is the number of exact rules divided by all answers to the questionnaire that have the data on the cause clause, which is three divided by four, equal to 0.75. Then, calculate the value of *confidence* in all candidates. Compare all *confidence* values to the future rule with a minimum value of a

predefined percentage of *confidence*, which is 0.8. The final answer is all the rules that meet the minimum confidence value. The final step is converting this rule into a sentence the reader understands.

For example, the first rule, "If 705 then 1203," will be converted to "If answering question 7 with number 5 then it will answer question 12 with number 3".

[IF (705) Then (1203)]; [IF (705) THEN (2402)]; [IF (705) THEN (2609)]; [IF (2402) THEN (2609)]; [IF (2609) THEN (2402)]; [IF (705) THEN (1203,2402)]; [IF (705,1203) THEN (2402)]; [IF (705,2402) THEN (1203)]; [IF (1203,2402) THEN (705)]; [IF (705) THEN (1203,2609)]; [IF (705,1203) THEN (2609)]; [IF (705,2609) THEN (1203)]; [IF (1203,2609) THEN (705)]; [IF (705) THEN (2402,2609)]; [IF (705,2402) THEN (2609)]; [IF (705,2609) THEN (2402)]; [IF (1203,2402) THEN (2609)]; [IF (1203,2609) THEN (2402)]; [IF (705) THEN (1203,2402,2609)]; [IF (705,1203) THEN (2402,2609)]; [IF (705,2402) THEN (1203,2609)]; [IF (705,2609) THEN (1203,2402)]; [IF (1203,2402) THEN (705,2609)]; [IF (1203,2609) THEN (705,2402)]; [IF (705,1203,2402) THEN (2609)]; [IF (705,1203,2609) THEN (2402)]; [IF (705,2402,2609) THEN (1203)]; [IF (1203,2402,2609) THEN (705)].

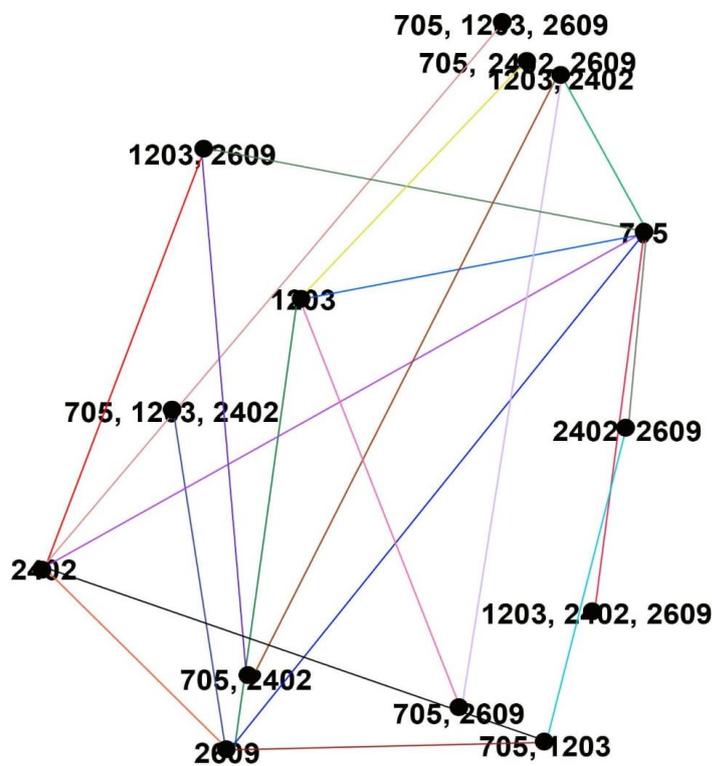


Figure 1. Mapping Questionnaire Answer

If question 7 is answered with number 5, it will answer question 12 with number 3. If question 7 is answered with number 5, then question 24 will be answered with number 2. If question 7 is answered with number 5, then question 26 will be answered with number 9. If question 24 is answered with number 2, it will answer question 26 with number 9. If question 26 is answered with the number 9, then question 24 will be answered with the number 2. If answering question 7 with the number 5, it will answer question 12 with the number 3 and answer question 24 with the number 2. If answering question 7 with the number 5 and question 12 with the number 3, it will answer question 24 with the number 2. If answering question 7 with the number 5 and answering question 24 with the number 2, then it will answer question 12 with the number 3. If answering question 12 with the number 3 and answering question 24 with the number 2, it will answer question 7 with the number 5. If answering question 7 with the number 5, it will answer question 12 with the number 3 and answer question 26 with the number 9. If answering question 7 with the number 5 and answering question 12 with the number 3, then it will

answer question 26 with the number 9. If answering question 7 with the number 5 and answering question 26 with the number 9, it will answer question 12 with the number 3. If answering question 12 with the number 3 and answering question 26 with the number 9, it will answer question 7 with the number 5. If answering question 7 with the number 5, then it will answer question 24 with the number 2 and answer question 26 with the number 9. If answering question 7 with the number 5 and answering question 24 with the number 2, it will answer question 26 with the number 9. If answering question 7 with the number 5 and answering question 26 with the number 9, it will answer question 24 with the number 2. If answering question 12 with the number 3 and answering question 24 with the number 2, it will answer question 26 with the number 9. If answering question 12 with the number 3 and answering question 26 with the number 9, it will answer question 24 with the number 2. If answering question 7 with the number 5, it will answer question 12 with the number 3, answer question 24 with the number 2, and answer question 26 with the number 9. If answering question 7 with the number 5 and question 12 with the number 3, it will answer question 24 with the number 2 and answer question 26 with the number 9. If answering question 7 with the number 5 and answering question 24 with the number 2, then it will answer question 12 with the number 3 and answer question 26 with the number 9. If answering question 7 with the number 5 and answering question 26 with the number 9, then it will answer question 12 with the number 3 and answer question 24 with the number 2. If answering question 12 with the number 3 and answering question 24 with the number 2, then it will answer question 7 with the number 5 and answer question 26 with the number 9. If answering question 12 with the number 3 and answering question 26 with the number 9, then it will answer question 7 with the number 5 and answer question 24 with the number 2. If answering question 7 with the number 5, question 12 with the number 3, and question 24 with the number 2, then it will answer question 26 with the number 9. If answering question 7 with the number 5, question 12 with the number 3, and answer question 26 with the number 9, then it will answer question 24 with the number 2. If answering question 7 with the number 5, question 24 with the number 2, and question 26 with the number 9, then it will answer question 12 with the number 3. If answering question 12 with the number 3, question 24 with the number 2, and question 26 with the number 9, then it will answer question 7 with the number 5.

Association rules, as one of the methods in data mining used in this research, is used to see the relation between some answers to student questionnaires. Data mining will extract information from the answers to student questionnaires and then look for association rules by finding the relationship between the questionnaire answers in a set of data using an Apriori algorithm to produce the association rules. Given the knowledge of students' answer patterns in answering questionnaires about learning, behavior, and activity, it can be further developed to find students' creativity based on theories proposed by Amabile [26].

## 6. Conclusion

From the results of this research, the Apriori algorithm can find the pattern of the interrelationship between students' learning, behavior, and activity variables with high *support* and trust values determined by students. With the low *confidence* value of the input (minimum *confidence*), the rules produced more and more. Conversely, the higher the confidence value, the better the pattern of associations. The result of the experiment combines minimum *confidence* and threshold *support* between 0.1 and 1 (10-100%); the highest value *supports* 0.95 or 95%. Experimental data show that 90% *support* and 90% minimum *confidence* will produce a few best rules, with association pattern with itemset combination showing difference and equation. The required processing time is influenced by the amount of data processed and the minimum *support* value specified. The more data is processed, the longer the student's learning process, behavior, and activity data will be. The smaller the minimum *support* value specified, the longer the processing time for that data will be. This is due to the small value of minutes, which means the probability of establishing association rules between items is more significant, which requires more checking of students' learning, behavior, and activity data. For further development, data from the results of calculations in this study can be used to determine a student's creativity.

## Acknowledgements

This research project is supported by the Second Century Fund (C2F), Chulalongkorn University, Thailand. We gratefully appreciate this support.

## References

- [1]. Xia, X., & Wang, T. (2024). Multi objective evaluation between learning behavior and learning achievement. *The Asia-Pacific Education Researcher*, 33(1), 1-15.
- [2]. Lunde, I. M. (2024). From generic skills to behaviour monitoring: exploring materialisations of the key skills framework in public-private relationships. *Journal of Educational Administration and History*, 56(1), 39-53.
- [3]. Borham, A. H., Huda, M., Rasid, M. S. A., Rahim, M. M. A., & Hamid, N. Z. A. (2024). Teaching approach for indigenous people: an empirical study from Pahang, Malaysia. *Journal of Education and Learning (EduLearn)*, 18(3), 773-782.
- [4]. Maseleno, A., Huda, M., Jasmi, K. A., Basiron, B., Mustari, I., Don, A. G., & bin Ahmad, R. (2019). Hau-Kashyap approach for student's level of expertise. *Egyptian Informatics Journal*, 20(1), 27-32.
- [5]. Schweder, S., & Raufelder, D. (2024). Does changing learning environments affect student motivation?. *Learning and Instruction*, 89, 101829.
- [6]. Pincus, J. D. (2024). Theoretical and empirical foundations for a unified pyramid of human motivation. *Integrative Psychological and Behavioral Science*, 58(2), 731-756.
- [7]. Hanson, C. L., Magnusson, B. M., Crandall, A. A., Barnes, M. D., McFarland, E., & Smith, M. (2024). Life experience pathways to college student emotional and mental health: A structural equation model. *Journal of American College Health*, 72(3), 826-833.
- [8]. Soto, C. J., Napolitano, C. M., Sewell, M. N., Yoon, H. J., & Roberts, B. W. (2024). Going beyond traits: Social, emotional, and behavioral skills matter for adolescents' success. *Social Psychological and Personality Science*, 15(1), 33-45.
- [9]. Sunnemark, F., Lundqvist Westin, W., Al Saad, T., & Assmo, P. (2024). Exploring barriers and facilitators to knowledge transfer and learning processes through a cross-departmental collaborative project in a municipal organization. *The Learning Organization*, 31(3), 358-374.
- [10]. Nguyen, A., Lämsä, J., Dwiarie, A., & Järvelä, S. (2024). Lifelong learner needs for human-centered self-regulated learning analytics. *Information and Learning Sciences*, 125(1/2), 68-108.
- [11]. Schachner, M. K., Hölscher, S., Moscardino, U., Ceccon, C., Juang, L., & Pastore, M. (2024). Adolescent cultural identity development in context: The dynamic interplay of the Identity Project with classroom cultural diversity climate in Italy and Germany. *Journal of Youth and Adolescence*, 1-19.
- [12]. Nan Cenka, B. A., Santoso, H. B., & Junus, K. (2024). Using the personal learning environment to support self-regulated learning strategies: a systematic literature review. *Interactive Learning Environments*, 32(4), 1368-1384.
- [13]. Chen, P., & Teo, Q. K. (2024). A Strategic Mindset for Learning and Life. In *Applying the Science of Learning to Education: An Insight into the Mechanisms that Shape Learning* (pp. 99-112). Singapore: Springer Nature Singapore.
- [14]. Romiszowski, A. J. (2024). *Producing instructional systems: Lesson planning for individualized and group learning activities*. Taylor & Francis.
- [15]. Christensen, L., & Aldridge, J. (2012). *Critical pedagogy for early childhood and elementary educators*. Dordrecht, Netherlands: Springer.

- [16]. Dugan, K. A., Vogt, R. L., Zheng, A., Gillath, O., Deboeck, P. R., Fraley, R. C., & Briley, D. A. (2024). Life events sometimes alter the trajectory of personality development: Effect sizes for 25 life events estimated using a large, frequently assessed sample. *Journal of Personality*, 92(1), 130-146.
- [17]. Keskin, M. T., Alagül, Ö., & Gürsel, F. (2024). Self-discovery and life skill building through holistic learning in high school students. *Journal of Adventure Education and Outdoor Learning*, 24(2), 175-186.
- [18]. Kuknor, S., & Kumar, V. R. (2024). Impact of training and development interventions for diversity & inclusion: proposing an organizational schema. *Development and Learning in Organizations: An International Journal*, 38(1), 16-19.
- [19]. Han, H. (2024). Exploring the relationship between purpose and moral psychological indicators. *Ethics & Behavior*, 34(1), 28-39.
- [20]. Mukhlis, H., Kristianingsih, A., Fitrianti, F., Pribadi, T., Kumalasari, D., Febriyanti, H., & Maselena, A. (2020). The effect of expressive writing technique to stress level decrease of new student at Al-Falah Putri Islamic Boarding School, Margodadi, Tanggamus. *Annals of Tropical Medicine and Public Health*, 23(06), 192-200.
- [21]. Huda, M., Hashim, A., Teh, K. S. M., Shankar, K., Ayshwarya, B., Nguyen, P. T., Hashim, W., & Maselena, A. (2019). Learning quality innovation through integration of pedagogical skill and adaptive technology. *International Journal of Innovative Technology and Exploring Engineering*, 8(9), 1538-1541.
- [22]. Jandra, M., Sentono, T., Huda, M., & Maselena, A. (2019). The effects of emotional intelligence, family environment and learning styles on social-science learning outcomes: an empirical analysis. *TEST Eng. Manag.*, 81, 4374-4386.
- [23]. Fernandez-Basso, C., Ruiz, M. D., & Martin-Bautista, M. J. (2024). New spark solutions for distributed frequent itemset and association rule mining algorithms. *Cluster Computing*, 27(2), 1217-1234.
- [24]. Du, L., Huang, F. S., Chen, S., Xiong, Y., Cai, Q., & Guo, Q. An Association Rule Mining-Based Modeling Framework for Characterizing and Predicting Urban Road Traffic Accidents. *Available at SSRN 4873399*.
- [25]. Kaushik, M., Sharma, R., Kõiva, P., Fister, I., & Draheim, D. (2024). An Exhaustive Multi-Aspect Analysis of Swarm Intelligence Algorithms in Numerical Association Rule Mining. *IEEE Access*.
- [26]. Amabile, T. M. (1997). Motivating creativity in organizations: On doing what you love and loving what you do. *California management review*, 40(1), 39-58.
- [27]. Maselena, A., Hardaker, G., Sabani, N., & Suhaili, N. (2016). Data on multicultural education and diagnostic information profiling: Culture, learning styles and creativity. *Data in brief*, 9, 1048-1051.

## Appendix A. Questionnaire

**Table 1.** Questionnaire [27]

No	Questions
1	Do you feel that the people you learn with on a regular basis are interesting?
2	Do you consider the learning you undertake on a regular basis to be a satisfying experience?
3	How often is there clear communications between you and the people you regularly learn with?
4	How often are your highest achievements in the learning environment recognised and rewarded?
5	How often do you feel the deadlines that your supervisor/tutor sets for you are realistic?

6	Can the people you learn with effectively manage failure in the course of their daily activities?
7	Are the people you regularly learn with outstanding at the learning activities assigned?
8	Do you feel your work pushes you to undertake activities that are achievable but challenging?
9	When you are learning in a group, are your colleagues helpful and respectful in the collaborative activities?
10	How often can you set yourself reasonably challenging goals in your learning and get feedback on your results?
11	Do you feel that most of the work/learning you undertake relates to your own personal interests?
12	Do you feel there is a sufficient amount of natural light in the space where you learn?
13	How often do you feel you become absorbed in the activities of learning?
14	Is the place where you learn a friendly environment?
15	Do you feel that your family or close friends are supportive of your learning development?
16	Are the people you learn and collaborate with on a regular basis competent at what they do?
17	How often do you expect to receive constructive criticism for your work?
18	How often do you feel your learning activities are appropriately monitored and evaluated?
19	Are you given sufficient resources to adequately undertake you learning activities?
20	How often are you in cooperation with the other people you work with?
21	How often are you encouraged by your colleagues in your efforts to be innovative in your learning?
22	Do you feel that the people you learn and interact with on a regular basis are cooperative?
23	Do you consider your achievements in learning important?
24	Is the place where you learn effectively managed?
25	Would you describe your learning experiences as challenging?
26	Are your colleagues constructive when they comment on and challenge your ideas?
27	How often do you feel you have influence over your learning activities in terms of making your own decisions and choices in direction?
28	Do you feel that learning has been a rewarding experience for you?
29	Do you feel your general activities are valuable to others?
30	How often do you feel that your learning is enabled through flexible procedures?

### Appendix B. Data Processing using Association Rules

Each question has answers from 0 to 10. The data of respondents who filled out the questionnaire is 10, with questionnaire answer data as shown in Table 2.

**Table 2.** Questionnaire Answer Data

Question	Student									
	1	2	3	4	5	6	7	8	9	10
1	02	00	07	03	01	03	00	01	00	09
2	06	05	04	02	04	09	04	10	06	01
3	03	02	07	04	00	04	10	06	04	06
4	06	06	09	08	10	00	07	00	05	09
5	09	05	00	02	10	08	05	05	10	06
6	10	08	10	06	06	00	00	09	10	05
7	08	06	05	08	06	05	02	05	10	02
8	08	00	03	07	08	09	04	05	03	03
9	10	00	01	03	03	09	09	07	04	02
10	05	09	04	02	09	02	04	03	05	02
11	06	01	01	06	00	08	10	00	10	01
12	06	03	03	02	05	03	01	03	05	04
13	02	07	02	00	10	07	01	00	01	00
14	05	04	05	02	04	02	01	03	07	00
15	02	06	06	01	10	01	01	04	09	05
16	02	10	05	01	04	09	02	04	10	03
17	09	10	10	05	10	07	02	05	05	07
18	10	04	07	00	00	10	08	07	00	10
19	07	02	05	00	01	06	02	01	08	04
20	01	06	10	08	09	04	10	01	05	00
21	06	08	05	10	01	07	01	08	10	09
22	00	05	07	09	03	01	01	03	08	06
23	04	08	07	07	02	00	02	06	00	07
24	05	08	02	01	02	02	08	02	00	08
25	07	06	08	05	07	04	09	06	10	03
26	08	08	09	03	09	09	04	09	07	01
27	04	03	06	06	01	06	05	05	08	01
28	10	02	05	03	09	00	10	00	02	05
29	08	04	08	05	04	07	04	05	00	09
30	03	10	01	05	02	07	06	04	00	10

The first step is to make the required adjustments to the input data. In this case, since all answers can be the same for all questions, adjustments should be made to differentiate if the same answer occurs for each

question. The adjustment is adding the question index number before each answer. As an example:

- for question 1 with answer 0, then the data value is 100
- for question 1 with answer 10, then the data value is 110
- for question 2 with answer 10, then the data value is 210
- for question 10 with answer 10, then the data value is 1010

So, the results of the questionnaire answer after the adjustment are shown in Table 3.

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	2	6	3	6	9	0	8	8	0	05	06	06	02	05	02	02	09	10	07	001	06	00	04	05	07	08	04	10	08	03
2	0	5	2	6	5	8	6	0	0	09	01	03	07	04	06	10	10	04	02	006	08	05	08	08	06	08	03	02	04	10
3	7	4	7	9	0	0	5	3	1	04	01	03	02	05	06	05	10	07	05	010	05	07	07	02	08	09	06	05	08	01
4	3	2	4	8	2	6	8	7	3	02	06	02	00	02	01	01	05	00	00	008	10	09	07	01	05	03	06	03	05	05
5	1	4	0	0	0	6	6	8	3	09	00	05	10	04	10	04	10	00	01	009	01	03	02	02	07	09	01	09	04	02
6	3	9	4	0	8	0	5	9	9	02	08	03	07	02	01	09	07	10	06	004	07	01	00	02	04	09	06	00	07	07
7	0	4	0	7	5	0	2	4	9	04	10	01	01	01	01	02	02	08	02	010	01	01	02	08	09	04	05	10	04	06
8	1	0	6	0	5	9	5	5	7	03	00	03	00	03	04	04	05	07	01	001	08	03	06	02	06	09	05	00	05	04
9	0	6	4	5	0	0	0	3	4	05	10	05	01	07	09	10	05	00	08	005	10	08	00	00	10	07	08	02	00	00
0	9	1	6	9	6	5	2	3	2	02	01	04	00	00	05	03	07	10	04	000	09	06	07	08	03	01	01	05	09	10

**Table 3.** Adjustment Answer Data Questionnaire

The next step is to determine the parameter percentage value of support as a determinant of an answer that can be said to be 'often answered.' In this case, the support percentage value is 0.3. Later, this support value will be multiplied by the number of respondents, which in this case is worth 10. So  $0.3 \times 10 = 3$ . This means an answer is said to be answered frequently if answered thrice in all questionnaire answer data.

Then, set a minimum and maximum number of answers to be counted. In this case, the minimum and maximum values are 2 and 4. This means that the computed answer is the set of answers with two answers, three answers, and four answers. Table 4 shows the number of each questionnaire answer.



Data	Number								
------	--------	------	--------	------	--------	------	--------	------	--------

**Table 4.** The Number of Each Questionnaire Answer

	Appear		Appear		Appear		Appear		Appear
100	3	702	2	1310	1	2000	1	2509	1
101	2	705	3	1400	1	2001	2	2510	1
102	1	706	2	1401	1	2004	1	2601	1
103	2	708	2	1402	2	2005	1	2603	1
107	1	710	1	1403	1	2006	1	2604	1
109	1	800	1	1404	2	2008	1	2607	1
201	1	803	3	1405	2	2009	1	2608	2
202	1	804	1	1407	1	2010	2	2609	4
204	3	805	1	1501	3	2101	2	2701	2
205	1	807	1	1502	1	2105	1	2703	1
206	2	808	2	1504	1	2106	1	2704	1
209	1	809	1	1505	1	2107	1	2705	2
210	1	900	1	1506	2	2108	2	2706	3
300	1	901	1	1509	1	2109	1	2708	1
302	1	902	1	1510	1	2110	2	2800	2
303	1	903	2	1601	1	2200	1	2802	2
304	3	904	1	1602	2	2201	2	2803	1
306	2	907	1	1603	1	2203	2	2805	2
307	1	909	2	1604	2	2205	1	2809	1
310	1	910	1	1605	1	2206	1	2810	2
400	2	1002	3	1609	1	2207	1	2900	1
405	1	1003	1	1610	2	2208	1	2904	3
406	2	1004	2	1702	1	2209	1	2905	2
407	1	1005	2	1705	3	2300	2	2907	1
408	1	1009	2	1707	2	2302	2	2908	2
409	2	1100	2	1709	1	2304	1	2909	1
410	1	1101	3	1710	3	2306	1	3000	1
500	1	1106	2	1800	3	2307	3	3001	1
502	1	1108	1	1804	1	2308	1	3002	1
505	3	1110	2	1807	2	2400	1	3003	1
506	1	1201	1	1808	1	2401	1	3004	1
508	1	1202	1	1810	3	2402	4	3005	1
509	1	1203	4	1900	1	2405	1	3006	1
510	2	1204	1	1901	2	2408	3	3007	1

600	2	1205	2	1902	2	2503	1	3010	2
605	1	1206	1	1904	1	2504	1		
606	2	1300	3	1905	1	2505	1		
608	1	1301	2	1906	1	2506	2		
609	1	1302	2	1907	1	2507	2		
610	3	1307	2	1908	1	2508	1		

The following process is to determine the itemset that occurs in all the answers to the questionnaire. The process is as follows:

1. Calculate the minimum value of the answer data that is often answered. The formula used is the percentage of support value multiplied by the number of respondents. In this case, it means  $0.3 * 10 = 3$  times. This variable will be used in calculating number 3.
2. Calculate the number of each answer on all questions. All answers to the questionnaire that have been adjusted will be counted as the number of answers that appear in all the questionnaires. As an example:
  - Data 100 appears as many as three times, that is, in the second, seventh, and ninth respondent
  - Data 101 appears as many as two times, which is the fifth and eighth respondent
  - Data 102 appears only one time, which is in the first respondent

And so on for each questionnaire's answer data. The final result of the calculation is as follows:

1. Based on the calculation in Table 4, find all data that fulfil the minimum value of support calculated at number 1, in this example, three times. The result of the data selection is shown in Table 5.

**Table 5. Minimum Support Selection Result**

Data	Number appear
100	3
204	3
304	3
505	3
610	3
705	3
803	3
1002	3
1101	3
1203	4
1300	3
1501	3
1705	3
1710	3
1800	3
1810	3
2307	3

2402	4
2408	3
2609	4
2706	3
2904	3

2. From each of these data, arrange the itemset for all combinations that can be found. In this case, the combination size used is within the range of 2 to 4 combinations. For each itemset with the combination, also calculate the number of occurrences in the questionnaire answer (same as process number 3), and if the number of occurrences of the combination meets the minimum support limit, enter the candidate response item (same as process number 4). As an example:

3. The calculation starts from a combination of 2 answers data. Concerning the selection results (Table 5), then:

- The first combination is 100 and 204. If we look at the questionnaire data table of responses to the questionnaire (Table 3), this combination is only found in the seventh respondent. Since the number is only one occurrence, this combination does not meet the minimum value of support worth three times, so this combination is not a candidate answer itemset.
- The second combination is 100 and 304. If we look at the questionnaire data table adjustment (Table 3), then this combination is only found in the ninth respondent. Since the number is only one occurrence, this combination does not meet the minimum value of support worth three times, so this combination is not an itemized answer.

Thus, the calculations performed on all combinations of 2 data can be found. The first successful combination of 2 data is 705 and 1203. This combination is found in the third, sixth, and eighth respondents. Since this combination appears three times, it is included in the itemset answer.

After the combination of 2 data is complete, then proceed to the combination of 3 data and the combination of 4 data. Record all combinations included in the itemset answer. The final result of the item answer is shown in Table 6.

**Table 6.** Itemset Answer

Itemset	Number Appear
705, 1203	3
705, 2402	3
705, 2609	3
1203, 2402	3
1203, 2609	3
2402, 2609	4
705, 1203, 2402	3
705, 1203, 2609	3
705, 2402, 2609	3
1203, 2402, 2609	3
705, 1203, 2402, 2609	3

The next step is to determine the rules for each item set that has been found. However, when entering the rule calculation, the next step is to determine the parameters of the percentage value of confidence as the determinant factor of the truth level of the relationships between the answers in itemset. In this case, the

percentage value of confidence is 0.8. This means that the relation will be included as an answer if its truth level on all questionnaires is above 80%

The process of determining the rules is as follows:

1. Divide each item set into two parts, for example, the cause clause and the resultant clause of the candidate rule. Find all possible combinations of the cause and effect clauses that can be found from each item set. The outcome of the candidate rule is shown in Table 7.

**Table 7.** Future Rule

Itemset	Cause Clause	Resultant Clause
705, 1203	705	1203
705, 2402	1203	705
705, 2609	705	2402
1203, 2402	2402	705
1203, 2609	705	2609
2402, 2609	2609	705
705, 1203, 2402	1203	2402
	2402	1203
	1203	2609
	2609	1203
	2402	2609
705, 1203, 2609	2609	2402
	705	1203, 2402
	1203	705, 2402
	2402	705, 1203
	705, 1203	2402
705, 2402, 2609	705, 2402	1203
	1203, 2402	705
	705	1203, 2609
	1203	705, 2609
	2609	705, 1203
1203, 2402, 2609	705, 1203	2609
	705, 2609	1203
	1203, 2609	705
	705	2402, 2609
	2402	705, 2609
705, 1203, 2402, 2609	2609	705, 2402
	705, 2402	2609

	705, 2609	2402
	2402, 2609	705
	1203	2402, 2609
	2402	1203, 2609
	2609	1203, 2402
	1203, 2402	2609
	1203, 2609	2402
	2402, 2609	1203
	705	1203, 2402, 2609
	1203	705, 2402, 2609
	2402	705, 1203, 2609
	2609	705, 1203, 2402
	705, 1203	2402, 2609
	705, 2402	1203, 2609
	705, 2609	1203, 2402
	1203, 2402	705, 2609
	1203, 2609	705, 2402
	2402, 2609	705, 1203
	705, 1203, 2402	2609
	705, 1203, 2609	2402
	705, 2402, 2609	1203
	1203, 2402, 2609	705

2. For each candidate rule, calculate the value of the candidate's rule confidence in the data table of questionnaire answer adjustment (Table 3). Note the comparison between the exact number of rules and all the answers to the questionnaires that have the data on the cause clause. As an example: "If 705 then 1203."

Looking at the data table of questionnaire answer adjustment (Table 3), data 705 is for the third, sixth, and eighth respondents. After looking at the three respondents, it turns out all three also entered data 1203.

So the value of his confidence are:

- For the first item (705, 1203), the first candidate rule is obtained

= The number of exact rules / all answers to the questionnaire that have the data in the cause clause

= 3 / 3

= 1

- For the first item (705, 1203), the second future rule is obtained

"If the 1203 is 705."

The data table of questionnaire adjustment (Table 3) shows that the second, third, sixth, and eighth respondents entered 1203 data. After looking at the four respondents, only the third, sixth, and eighth respondents entered data 1203, while the second respondent entered data 706.

So, the value of its confidence is

= The number of exact rules / all answers to the questionnaire that have the data on the cause clause

= 3 / 4

= 0.75

Calculate the value of confidence on all candidate rules found in Table 7. The final result of the calculation is shown in Table 8.

**Table 8.** Confidence Value at Future Rule

Itemset	Cause Clause	Resultant Clause	Confidence value
705, 1203	705	1203	1
	1203	705	0.75
705, 2402	705	2402	1
	2402	705	0.75
705, 2609	705	2609	1
	2609	705	0.75
1203, 2402	1203	2402	0.75
	2402	1203	0.75
1203, 2609	1203	2609	0.75
	2609	1203	0.75
2402, 2609	2402	2609	1
	2609	2402	1
705, 1203, 2402	705	1203, 2402	1
	1203	705, 2402	0.75
705, 1203, 2609	2402	705, 1203	0.75
	705, 1203	2402	1
705, 1203, 2609	705, 2402	1203	1
	1203, 2402	705	1
705, 2402, 2609	705	1203, 2609	1
	1203	705, 2609	0.75
705, 2402, 2609	2609	705, 1203	0.75
	705, 1203	2609	1
705, 2402, 2609	705, 2609	1203	1
	1203, 2609	705	1
1203, 2402, 2609	705	2402, 2609	1
	2402	705, 2609	0.75
1203, 2402, 2609	2609	705, 2402	0.75
	705, 2402	2609	1

705, 1203, 2402, 2609	705, 2609	2402	1
	2402, 2609	705	0.75
	1203	2402, 2609	0.75
	2402	1203, 2609	0.75
	2609	1203, 2402	0.75
	1203, 2402	2609	1
	1203, 2609	2402	1
	2402, 2609	1203	0.75
	705	1203, 2402, 2609	1
	1203	705, 2402, 2609	0.75
	2402	705, 1203, 2609	0.75
	2609	705, 1203, 2402	0.75
	705, 1203	2402, 2609	1
	705, 2402	1203, 2609	1
	705, 2609	1203, 2402	1
	1203, 2402	705, 2609	1
	1203, 2609	705, 2402	1
	2402, 2609	705, 1203	0.75
	705, 1203, 2402	2609	1
	705, 1203, 2609	2402	1
705, 2402, 2609	1203	1	
1203, 2402, 2609	705	1	

3. Compare all confidence values to the future rule with a minimum value of predefined percentage of confidence, which is 0.8. The final answer is all the rules that meet the minimum value of confidence. The final result is shown in Table 9.

**Table 9.** Last Rule Answer

Itemset	Cause Clause	Resultant Clause	Confidence Value
705, 1203	705	1203	1
705, 2402	705	2402	1
705, 2609	705	2609	1
2402, 2609	2402	2609	1
705, 1203, 2402	2609	2402	1
	705	1203, 2402	1
	705, 1203	2402	1
705, 1203, 2609	705, 2402	1203	1
	1203, 2402	705	1
	705	1203, 2609	1

705, 2402, 2609	705, 1203	2609	1
	705, 2609	1203	1
1203, 2402, 2609	1203, 2609	705	1
705, 1203, 2402, 2609	705	2402, 2609	1
	705, 2402	2609	1
	705, 2609	2402	1
	1203, 2402	2609	1
	1203, 2609	2402	1
	705	1203, 2402, 2609	1
	705, 1203	2402, 2609	1
	705, 2402	1203, 2609	1
	705, 2609	1203, 2402	1
	1203, 2402	705, 2609	1
	1203, 2609	705, 2402	1
	705, 1203, 2402	2609	1
	705, 1203, 2609	2402	1
	705, 2402, 2609	1203	1
	1203, 2402, 2609	705	1

4. The final step is to convert this rule into a sentence to be understood by the reader. As an example:

For the first rule, "If 705 then 1203", it will be converted to

"If answering question 7 with number 5 then it will answer question 12 with number 3."

This rule is always appropriate in all respondent data, with a confidence score of 1. Table 10 shows some samples of translation rules into words.

**Table 10.** Translating rule into word

Cause Clause	Resultant Clause	Rule Conversion	Confidence
--------------	------------------	-----------------	------------

			Value
705	1203	If answering question 7 with the number 5, then it will answer question 12 with the number 3	1
705	2402	If answering question 7 with the number 5, then it will answer question 24 with the number 2	1
705	2609	If answering question 7 with the number 5, then it will answer question 26 with the number 9	1
2402	2609	If answering question 24 with the number 2, then it will answer question 26 with the number 9	1
2609	2402	If answering question 26 with the number 9, then it will answer question 24 with the number 2	1
705	1203, 2402	If answering question 7 with the number 5, then it will answer question 12 with the number 3 and answer question 24 with the number 2	1
705, 1203	2402	If answering question 7 with the number 5 and answering question 12 with the number 3, then it will answer question 24 with the number 2	1
705, 2402	1203	If answering question 12 with the number 3 and answering question 24 with the number 2, then it will answer question 7 with the number 5	1
1203, 2402	705	If answering question 7 with the number 5, then it will answer question 12 with the number 3 and answer question 26 with the number 9	1
705	1203, 2609	If answering question 7 with the number 5 and answering question 12 with the number 3, then it will answer question 26 with the number 9	1
705, 1203	2609	If answering question 7 with the number 5 and answering question 26 with the number 9, then it will answer question 12 with the number 3	1
705, 2609	1203	If answering question 7 with the number 5, then it will answer question 24 with the number 2 and answer question 26 with the number 9	1
1203, 2609	705	If answering question 7 with the number 5 and answering question 24 with the number 2, then it will answer question 26 with the number 9	1
705	2402, 2609	If answering question 7 with the number 5 and answering question 26 with the number 9, then it will answer question 12 with the number 3 and answering question 24 with the number 2, then it will answer question 26 with the number 9	1
		If answering question 12 with the number 3 and answering question 26 with the number 9, then it will	

705, 2402	2609	answer question 24 with the number 2	1
705, 2609	2402	If answering question 7 with the number 5, then it will answer question 12 with the number 3, answer question 24 with the number 2, and answer question 26 with the number 9	1
1203, 2402	2609	If answering question 7 with the number 5 and answering question 12 with the number 3, then it will answer question 24 with the number 2 and answer question 26 with the number 9	1
1203, 2609	2402	If answering question 7 with the number 5 and answering question 24 with the number 2, then it will answer question 12 with the number 3 and answer question 26 with the number 9	1
705	1203, 2402, 2609	If answering question 7 with the number 5 and answering question 26 with the number 9, then it will answer question 12 with the number 3 and answer question 24 with the number 2	1
705, 1203	2402, 2609	If answering question 12 with the number 3 and answering question 24 with the number 2, then it will answer question 7 with the number 5 and answer question 26 with the number 9	1
705, 2402	1203, 2609	If answering question 12 with the number 3 and answering question 24 with the number 2, then it will answer question 26 with the number 9	1
		If answering question 7 with the number 5, answering question 12 with the number 3, and answering question 26 with the number 9, then it will answer question 24 with the number 2	
		If answering question 7 with number 5 and answering question 24 with number 2 and answering question 26 with number 9 then it will answer question 12 with number 3	
		If answering question 12 with number 3 and answering question 24 with number 2 and answering question 26 with number 9 then it will answer question 7 with number 5	