# Analysis of Kinematics Graph Interpretation Skills Using RapidMiner

Kanokporn Intakaew[a] and Pornrat Wattanakasiwich[a]

Corresponding author: pornrat.w@cmu.ac.th
[a]Department of Physics and Materials Science, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

## Abstract

This study explores the application of data mining techniques in physics education research, focusing on the analysis of kinematics graph interpretation skills. The research had two main objectives: to demonstrate the utility of RapidMiner, a data mining tool, in analyzing educational data, and to compare its effectiveness with traditional item analysis methods. Fifty-nine Grade-10 students at Chiang Mai University Demonstration School completed the Test of Understanding Graphs in Kinematics (TUG-K) before and after participating in a problem-based learning module integrating high-speed video analysis. Traditional statistical analysis revealed significant improvement in student performance ($p<0.001$, effect size 0.76). Association rule mining, conducted using RapidMiner, uncovered key relationships between test items that were not apparent through traditional analysis. These relationships provided insights into common student misconceptions and areas requiring targeted instruction. The study demonstrates the potential of advanced data mining techniques to reveal deeper patterns in educational data compared to conventional item analysis methods. This novel application of RapidMiner in physics education research offers a promising approach for more detailed analysis of student understanding, potentially informing more effective teaching strategies and curriculum design in physics education.

## Introduction

Understanding the specific areas where students struggle is key to improving teaching strategies and designing effective curricula (Wattanakasiwich, Taleab, Sharma & Johnston, 2013). In physics education, decades of research have revealed a wide range of alternative conceptions that students hold, which often hinder their understanding of key concepts. Various methods have been developed to identify these misconceptions, with multiple-choice surveys being one of the most effective tools (Ding & Beichner, 2009). These surveys, such as the Test of Understanding Graphs in Kinematics (TUG-K) (Beichner, 1994), have been instrumental in exploring students' understanding of complex physics topics like force and motion (Hestenes & Wells, 1992; Thornton & Sokoloff, 1998), energy and momentum (Singh & Rosengrant, 2003), and mechanical waves (Tongchai, Sharma, Johnston, Arayathanitkul & Soankwan, 2009).

In analyzing the data from such assessments, traditional methods like classical test theory, factor analysis, and item response theory have been widely employed in physics education research (Ding & Beichner, 2009). These methods provide important metrics such as item difficulty and discrimination indices, but they are limited in their ability to uncover deeper

relationships between test items. For instance, classical analysis techniques typically focus on individual item performance, overlooking potential associations between items that could reveal how students perceive connections between related concepts.

To address these limitations, this study introduces the use of association rule mining—a data mining technique that can uncover hidden patterns and relationships between test items. Data mining, particularly in the field of education, has emerged as a powerful tool for analyzing large datasets and extracting meaningful insights into student behavior and learning outcomes (Romero & Ventura, 2010). In the context of physics education, data mining techniques have been successfully applied to predict student performance and to identify misconceptions (Zabriskie, Yang, DeVore & Stewart, 2019).

## Background

### Education data mining
Data mining is the process of extracting patterns, trends, and insights from big databases, often via the use of statistical and computer tools (Bachhal, Ahuja & Gargrish, 2021). It is the process of analyzing and extracting meaningful information from massive datasets, including structured and unstructured data, using sophisticated algorithms and statistical models. Educational data mining (EDM) is a growing field that applies data mining techniques to educational data to extract insights and improve educational outcomes (Romero & Ventura, 2010). Educational data can come from a variety of sources, including student performance data, demographic data, assessment data, and learning management system data.

EDM has become increasingly important in recent years because of the growing amount of data generated in educational settings (Salloum, Alshurideh, Elnagar & Shaalan, 2020). Student behavior patterns, such as attendance, academic achievement, and learning preferences, may be found through data mining in education (Márquez-Vera, Morales & Soto, 2013). Overall, educational data mining is a critical tool for improving teaching and learning outcomes. By leveraging the power of data, educators and researchers can gain insights into student behavior, performance, and learning needs, and use this information to develop effective interventions and personalized learning experiences (Romero & Ventura, 2010; Bachhal et al., 2021).

Data mining techniques can be useful in identifying student misconceptions in physics, predicting student performance (Yang et al., 2020; Zabriskie et al., 2019), improving curriculum design and personalizing instruction. Previous physics education research had been focused in using data mining to predict student performance in physics courses (Zabriskie et al., 2019).

### Kinematics graph interpretation skills
Research on graphical interpretation skills in physics has focused on understanding how students learn to read, interpret, and use graphs to understand physical phenomena (Beichner, 1994; Bollen, De Cock, Zuza, Guisasola & van Kampen, 2016). Previous physics education research has found that students often struggle with graphical interpretation in physics, particularly when interpreting non-linear graphs or graphs with complex axes. In addition, studies have shown that students often have difficulty identifying and interpreting key features of graphs, such as slopes, intercepts, and areas under the curve (Susac, Bubic, Kazotti, Planinic & Palmovic, 2018).

The test of understanding graphs in kinematics (TUG-K) is one of the most widely used tests to assess student understanding and interpretation skills in kinematic graphs (Beichner, 1994; Klein, Becker, Küchemann & Kuhn, 2021). Physics education researchers use TUG-K to study various aspects of students' understanding of kinematics graphs, including the specific misconceptions that students have and how those misconceptions can be addressed in the classroom, such as the slope and the area under the curve (Susac et al., 2018). The test is also used to compare the performance of students from different backgrounds or educational settings, as well as to evaluate the effectiveness of instructional interventions or teaching strategies (Araujo, Veit & Moreira, 2008).

Extensive use of the TUG-K test in physics education research has resulted in numerous significant findings. However, the quantitative method used to analyze items in TUG-K is still limited to classical test analysis (Klein et al., 2021). Large-scale assessment data is usually reduced to distractor frequencies and item response accuracy rates, overlooking potential item associations that could directly compare how students perceive differences or similarities between items. To address these limitations, data mining techniques can be employed to analyze student responses from the test.

**RapidMiner**
The RapidMiner software package is designed for data mining and modeling. Its user-friendly graphical interface allows for the rapid and intuitive implementation and execution of data mining processes (Fernández & Luján-Mora, 2017). Two significant trends in its use with educational data are analyzing large datasets and developing predictive models which can assist educators in identifying at-risk students and providing targeted interventions to enhance their outcomes (Márquez-Vera et al., 2013). Another trend is analyzing text data, such as essays or forum posts, to identify patterns and themes in the data. RapidMiner's text mining capabilities enable researchers to evaluate large amounts of text data. Additionally, it has been used to analyze social network data extracted from online learning platforms to identify significant trends in student interactions and how these affect learning outcomes (Gao, Li & Wu, 2021).

RapidMiner is a versatile data mining tool that can assist educational researchers in extracting insights and improving educational outcomes (Slater, Joksimović, Kovanovic, Baker & Gasevic, 2017). RapidMiner can be used to analyze multiple-choice tests in terms of descriptive statistics, association rule mining and predictive models. In this paper, RapidMiner was used to perform association rule mining to identify relationships between different test items in the Test of Understanding Graphs in Kinematics (TUG-K) (Beichner, 1994). This can help to identify which test items are strongly correlated, which can be useful for identifying areas where students need more instruction.

In comparison, traditional spreadsheet software like Excel is typically limited to basic statistical analysis, such as calculating item difficulty and discrimination. While these metrics are useful, they lack the ability to reveal complex relationships between test items. RapidMiner, by contrast, allows for a deeper analysis by uncovering patterns and correlations between items, providing insights into how student responses are interconnected.

## Purpose of the study

The aim of this study is twofold. First, it seeks to demonstrate the utility of RapidMiner in analyzing educational data and gaining insights into students' understanding of kinematic graphs. By applying association rule mining to responses from the Test of Understanding Graphs in Kinematics (TUG-K), the study identifies relationships between test items. Second, the study compares the effectiveness of data analysis using RapidMiner with traditional item analysis using Excel.

## Methodology

### Participants

Participants in this study consisted of 60 students in grade 11 at Science Classrooms in University - Affiliated School Project (SCiUS Project) at Chiang Mai University Demonstration School. This project is to establish science classrooms in schools for talented students in science and technology under the supervision of university. All 60 students took the test before starting the STEM learning module "Sport science" and after finishing the module. The Sports Science module is part of the Integrated Science and Mathematics 8 curriculum, where students explore the application of physics principles, such as kinematics and dynamics, to analyze motion in sports. This hands-on approach helps students understand the relationship between forces, energy, and movement in physical activities. The module emphasizes both theoretical learning and practical data analysis, such as measuring and correlating speed, acceleration, and other variables in sports scenarios. However, only 59 students completed the post-test because one student went abroad for an exchange program. Therefore, we analyzed the data collected from the 59 students who completed both pre-test and post-test.

### Instrument

The Test of Understanding Graphs in Kinematics (TUG-K) is a tool that is commonly used to assess a student's ability to interpret and analyze graphs in the context of kinematics, which is the study of motion (Beichner, 1994). The modified version of TUG-K has been made to achieve parallelism of the objectives of the test and proven to be better in terms of item difficulty, item discrimination and reliability (Zavala, Tejeda, Barniol & Beichner, 2017). However, we used the original version of the TUG-K test in this study because we would like to later compare findings with student responses over the past 12 years. The test consists of 21 multiple-choice questions that depict various aspects of kinematic graphs, such as position, velocity, and acceleration, and students are asked to answer questions related to these graphs. The TUG-K is designed to assess a student's graphical interpretation skills, as well as their understanding of kinematic concepts. In order to perform well on the test, students must be able to read and interpret the information presented in the graphs, including the axes, units, scales, and key features of the graphs. In this study, the Thai version of TUG-K was used and administered to students. An example of question 2 in the test is displayed in Figure 1.

When is the acceleration the most negative?
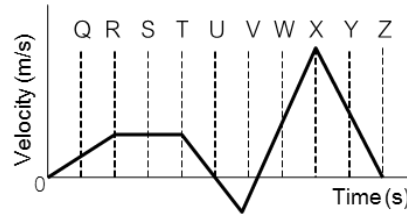
(A) R to T
(B) T to V
(C) V
(D) X
(E) X to Z

**Figure 1** Question 2 of the TUG-K test

## Data Mining Process with RapidMiner

For the data analysis process using data mining, a model of the Cross-Industry Standard Process for Data Mining (CRISP-DM) process is utilized, which consists of six steps as illustrated in Figure 2 (Wirth & Hipp, 2000).
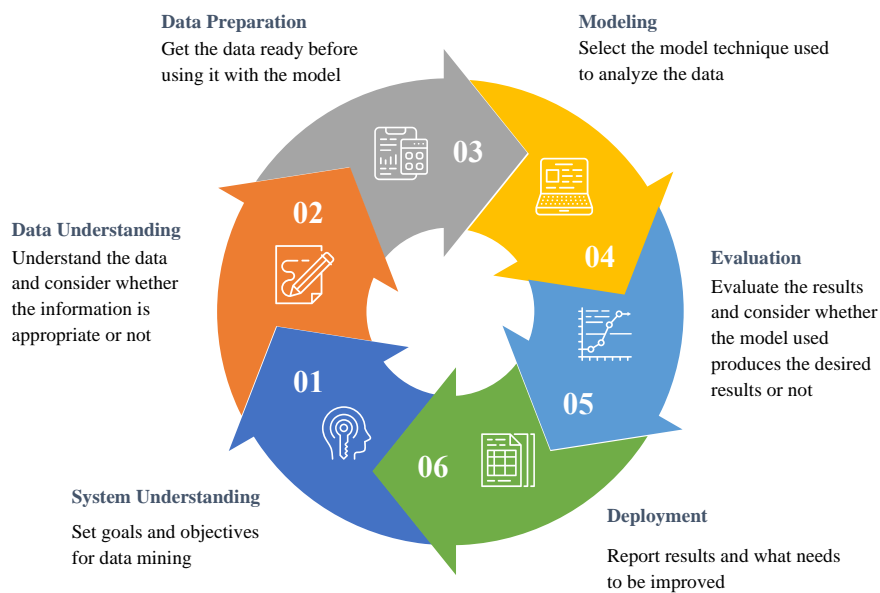


**Figure 2** The CRISP-DM process is used to provide a framework for the data mining process.

In this study, the data mining analysis process is as follows:

**1. Systems Understanding**— The system in our study is the Test of Understanding Graph in Kinematics or TUG-K, consisting of 21 multiple-choice questions. TUG-K is a well-known conceptual test in assessing students' understanding and interpretation of kinematic graphs.

**2. Data Understanding**—TUG-K pre-test and post-test responses were collected from 59 students in grade 10 at Chiang Mai University Demonstration School.

**3. Data Preparation**—The students' correct answer to each question was represented with a numerical value of "1" and the incorrect answer with a numerical value of "0."

**4. Modeling**— The Frequency Pattern (FP) Growth operator is used to determine the frequency of item sets and uncover any relationships among items in the dataset. This operator analyzes the transaction database by constructing an FP-tree data structure to calculate all frequent item

sets from an example set. The frequency of each item set is computed using the FP-tree data structure. Figure 3 displays a process of association rules with multiple operators in RapidMiner.
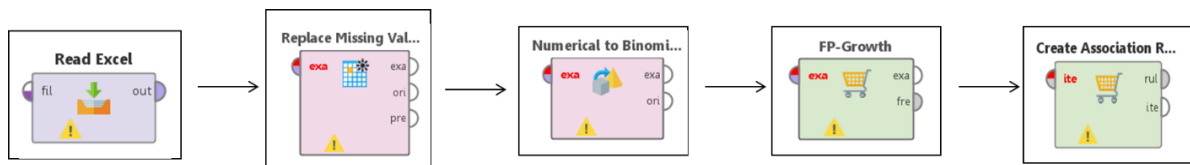


**Figure 3** Process of creating association rules with multiple operators in RapidMiner**,** including Read Excel, Replace Missing Value, Numerical to Binominal, FP-Growth and Create Association Rules

- "Read Excel" operator is for importing data from Excel spreadsheet into RapidMiner.
- "Replace Missing Value" operator is for replacing the missing data with 0.
- "Numerical to binominal" operator is for changing the type of the selected numeric attributes to a binominal type.
- "FP-Growth" operator is shorted for "The Frequent Pattern Growth" operator. The FP-Growth is a method used to identify frequent patterns by scanning the Transaction ID (TID) just twice and constructing a tree to locate support. To identify frequent item sets in RapidMiner, we can set the minimum support value to 0.5 and examine which minimum support ranges occur frequently with this item set. After identifying these ranges, we can adjust the minimum support value to 0.78 ; Support refers to the proportion of transactions that contain both item X and item Y ; the equation for support is

$$\text{Support} = \frac{\text{No. of students who answered correctly}}{\text{Total No. of students}}$$

- "Create Association Rules" operator is for creating association rules, setting the minimum confidence level at 0.80; confidence measures how often item in Y appear in transactions that contain X; the equation for confidence is

$$\text{Confidence} = \frac{\text{Support (premises,conclusion)}}{\text{Support (premises)}}$$

While premises are the antecedent conditions that determine the probability of a consequent, conclusions refer to the predicted item sets or conditions that follow from the premises. For example, consider the TUG-K test, where answering Question 5 correctly might represent a premise indicating that the student is likely to answer Question 20 correctly.

**5. Evaluation**—Consider how each question relates to other questions on TUG-K.

**6. Deployment**—The relationship rules obtained from the modeling step will be used to analyze other data sets of students' answers to TUG-K exams from previous academic years.

## Results and discussion

The statistical results from both pre and post-tests showed that the mean scores significantly differed at the 0.001 level, indicating that the mean scores after the instruction (mean = 14.3,

SD = 3.7) were significantly higher than the pre-test mean scores (mean = 12.4, SD = 3.2). The effect size was 0.76, which suggests that the post-test scores were substantially different from the pre-test scores. The classical test theory was used to analyze each item based on item difficulty, item discrimination index and point-biserial coefficient.

**Item Analysis**

*Item Difficulty*

The Item Difficulty Index (P) measures the level of difficulty of a question, calculated by the ratio of correct responses to total responses. The average difficulty ratings of 59 students are shown in Figure 4, where a P value of 0.0 indicates no correct answers and 1.0 indicates all correct answers. A question is considered to be easy if P > 0.9, and it is considered to be difficult if P < 0.3 (Wuttiprom, Sharma, Johnston, Chitaree, & Soankwan, 2009). As illustrated in Figure 4, most questions in TUG-K have a difficulty index between 0.3 and 0.8, with a few items slightly below 0.3 or above 0.9. The average P values for the pre-test and post-test were 0.59 and 0.67, respectively, which are close to the ideal value of 0.5.
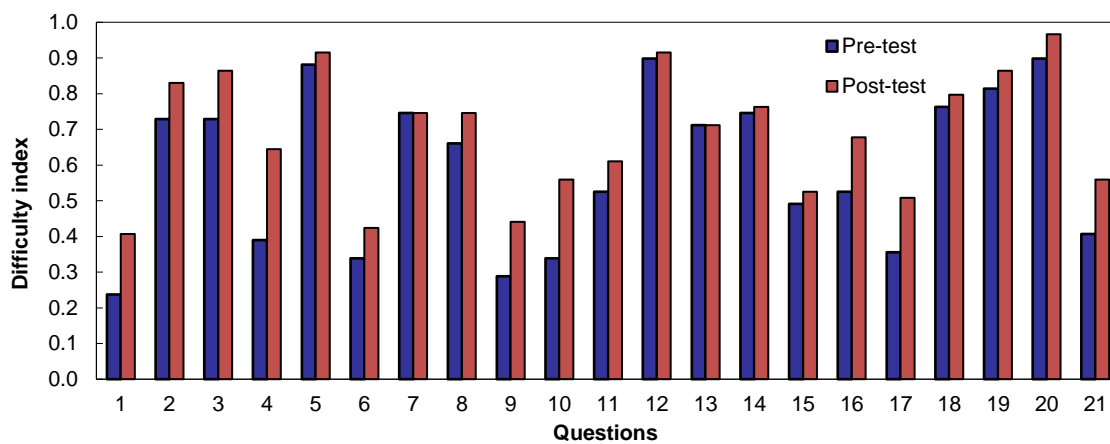


**Figure 4** Figure 4 Item difficulty on the pre-test and post-test using the TUG-K (Thai version).

*Item Discrimination*

The discrimination index (D) measures how well a question distinguishes between competent and less competent students. Specifically, it distinguishes those who performed well on the survey from those who did not. A satisfactory question discrimination index is D ≥ 0.3 (Wuttiprom et al., 2009). Figure 5 shows the discrimination index for each question. More than 15 questions have a D value greater than 0.3 for both the pre-test and post-test, and the average discrimination index is 0.46. Although 6 other questions fall slightly below the standard range, they are still acceptable, as noted by Adams and Wieman (2011). These researchers state that the acceptable range of values for item analysis was established for single constructs and summative tests designed to differentiate individual students. However, instruments used for formative assessments "may have statistics that fall outside of the "standard" range" (Adams & Wieman, 2011).
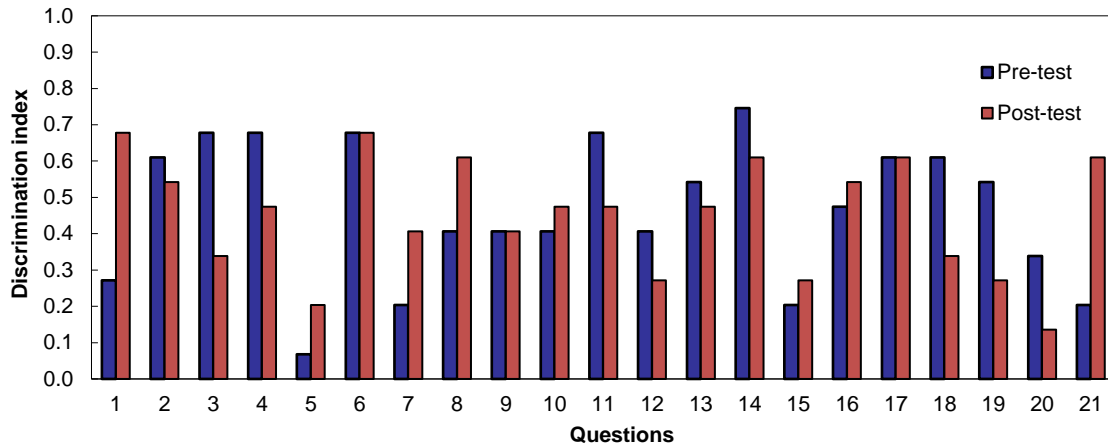
**Figure 5** Discrimination index of all 21 TUG-K questions for both pre-test and post-test

*Point-biserial coefficient*

The point-biserial coefficient (PBI) measures the correlation of each question with the total score of the TUG-K test. A high PBI value indicates that students who answered the question correctly are more likely to have a high overall score. The desired value of the coefficient is greater than 0.2, as suggested (Wuttiprom et al. 2009). Figure 6 illustrates the PBI for each question. Most questions have a coefficient higher than the criterion value, except for questions 5, 7, and 21 in the pre-test. The average coefficient for the pre-test is 0.34, and for the post-test, it is 0.42, indicating that the questions are internally consistent with the whole TUG-K test.
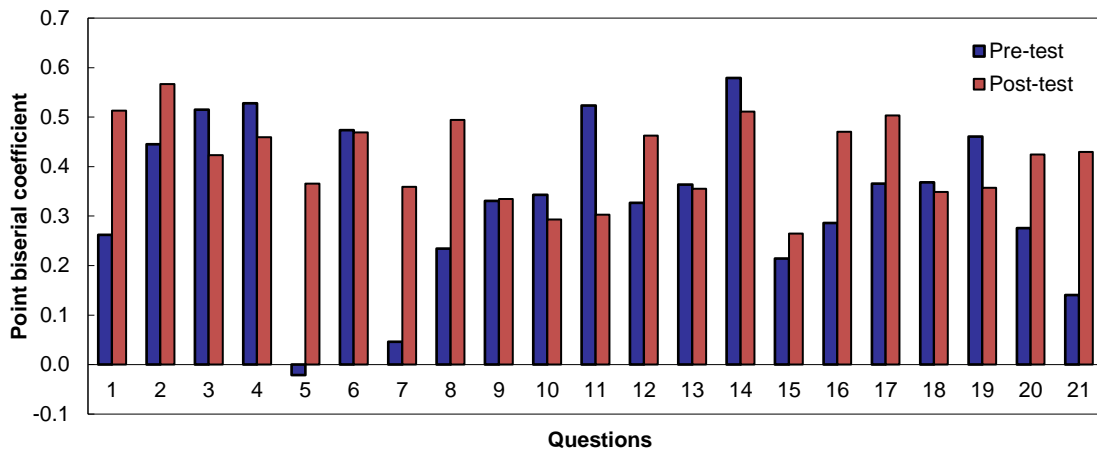


**Figure 6** Point biserial coefficient (PBI) of all 21 TUG-K questions for both pre-test and post-test

While the item analysis performed using Excel provided valuable insights into the difficulty and discrimination indices of each test item, it did not offer information about the relationships between the items. Traditional item analysis is limited to evaluating individual item performance metrics, but it lacks the capability to identify how students' responses to different items may be interconnected. To address this limitation, we employed association rule mining with RapidMiner to uncover relationships between test items, offering a deeper understanding of how students' responses to one question may influence their performance on others.

**Association between items**

In analyzing the correlation of questions in the TUG-K test from the post-learning scores of students using RapidMiner, it was found that 4 questions passed with a minimum support of 0.78 and a minimum confidence of 0.80, which was specified in Questions 5, 12, 19, and 20 as in Table 1.

**Table 1:** Displays the correlation pairs among questions in the TUG-K test that meet both the minimum support and minimum confidence requirements.

| Premises | Conclusion | | Support | Confidence |
|:---:|:---:|:---:|:---:|:---:|
| **T19** | T12 | – | 0.787 | 0.960 |
| **T05** | T20 | - | 0.803 | 0.907 |
| **T05** | T12 | | 0.803 | 0.907 |
| **T20** | T05 | – | 0.803 | 0.891 |
| **T12** | T05 | | 0.803 | 0.891 |
| **T20** | T12 | | 0.803 | 0.891 |
| **T12** | T20 | | 0.803 | 0.891 |
| **T19** | T12 | | 0.787 | 0.873 |

An analysis conducted using RapidMiner indicated an interesting relationship between Questions 5 and 20, and between Questions 12 and 19, as each of these pairs exhibited the highest effective confidence. Table 1 displays the relationships between the questions, and it can be observed that row 2 and row 9 (indicated by the red frame) correspond to the relationship between Questions 12 and 19, while row 3 and row 5 (indicated by the blue frame) correspond to the relationship between Questions 5 and 20.

To understand how to interpret the correlation chart in RapidMiner, as illustrated in Figure 7(A), note that the values in parentheses on the left indicate support, while the values on the right represent confidence values obtained by constructing correlation rules in RapidMiner.
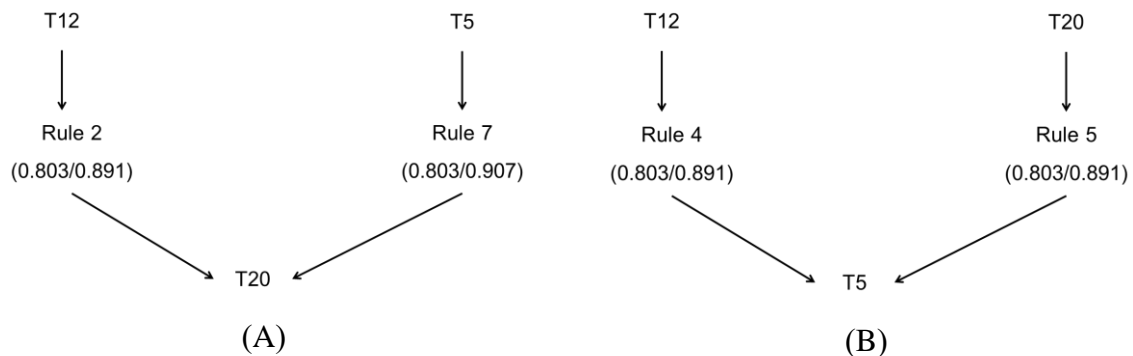


**Figure 7** (A) question pairs that are associated with Question 20
(B) question pairs that are associated with Question **5**

According to the analysis of Questions 5 and 20 in Figure 7, it was discovered that a correct answer to Question 5 would also result in Question 20 being correct (shows in Figure 7 (A)). The confidence value for this relationship was determined to be 0.907. This means that if 100 students took the TUG-K test and all of them answered Question 5 correctly, then 98 out of the 100 students would also get Question 20 correct. But if Question 20 is answered correctly, then it is likely that Question 5 will also be answered correctly (shows in Figure 7 (B)). The

confidence value for this relationship was found to be 0.891, which indicates a slight decrease compared to the previous relationship. In other words, if all 100 students answered Question 20 correctly on the TUG-K test, then it is estimated that 89 out of the 100 students would also get Question 5 correct.

When taking these two questions into consideration in more detail, it was found that Question 5 (shown in Figure 8) is a consideration of finding the speed. Students must understand the slope as well, but for Question 20 (shown in Figure 9) they can count the area under the graph to find the answer. So, Question 20 seems simpler than Question 5. However, these two questions ask similarly about how to find the slope and area under the graph. As a result, the correlation between these two questions is high.

5. The velocity at the 2 second point is:
      (A) 0.4 m/s
      (B) 2.0 m/s
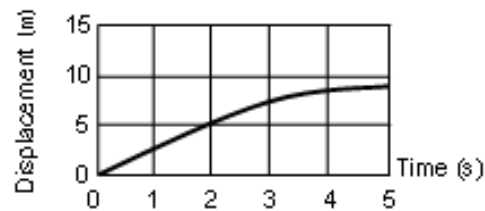      (C) 2.5 m/s
      (D) 5.0 m/s
      (E) 10.0m/s

**Figure 8** Question 5 in TUG-K

20. An object moves according to the graph below. How far does it move during the interval from t = 4 s to t = 8 s?
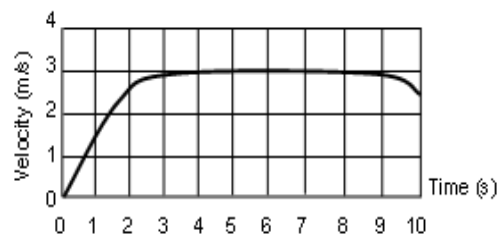(A) 0.75 m    (B) 8.0 m
(C) 3.0 m    (D) 12.0 m
(E) 4.0 m

**Figure 9** Question 20 in TUG-K

T19 → Rule 8 (0.787/0.960)
T5 → Rule 6 (0.803/0.907)
T20 → Rule 3 (0.803/0.891)
Rule 8, Rule 6, Rule 3 → T12

(A)
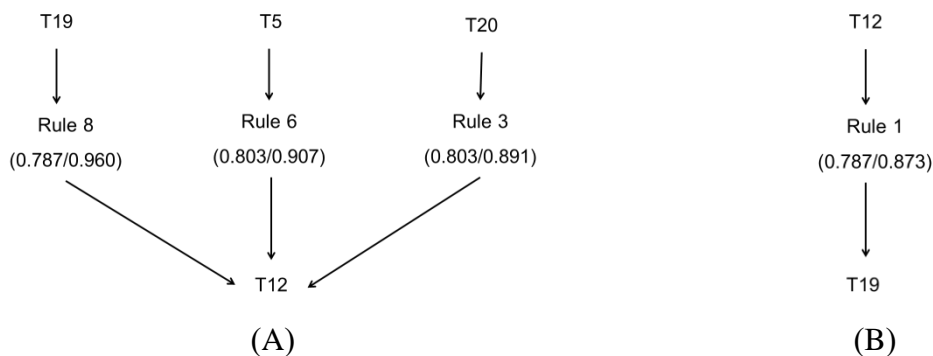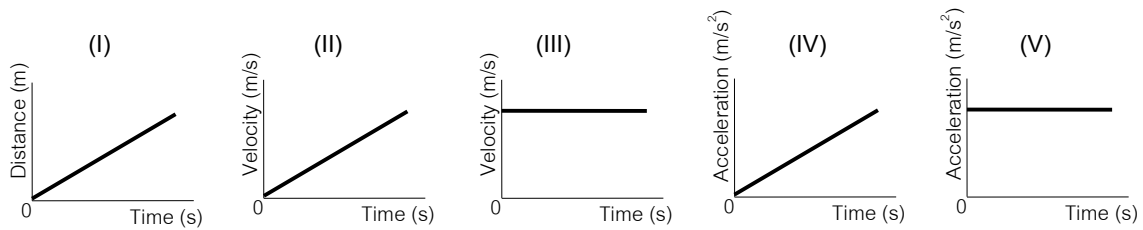
T12 → Rule 1 (0.787/0.873) → T19

(B)

**Figure 10**  (A) question pairs that are associated with Question 12
(B) question pairs that are associated with Question 19

From the analysis of Questions 12 and 19, it was found that if Question 19 were done correctly, it would result in Question 12 being correct (shows in Figure 10 (A)). with a confidence value of 0.960, but if Question 12 were done correctly, it would result in Question 19 being correct (shows in Figure 10 (B)).  with a confidence value of 0.873.

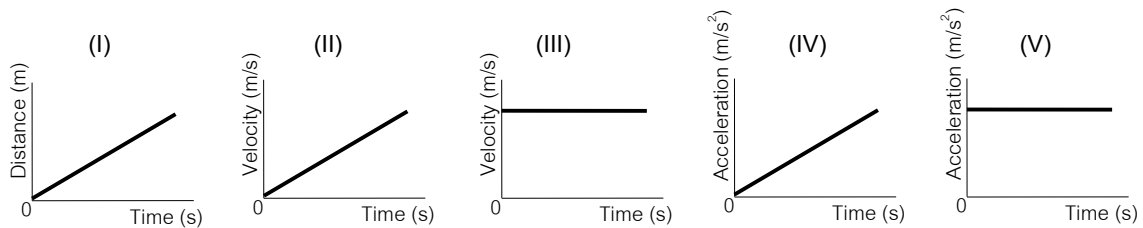12. Consider the following graphs, noting the different axes:



Which of these represent motion at constant velocity?
(A) I,II and IV    (B) I and III    (C) II and V    (D) IV only    (E) V only

**Figure 11** Question 12 in TUG-K

19. Consider the following graphs, noting the different axes:



Which of these represent motion at constant, non-zero acceleration?
(A) I,II and IV    (B) I and III    (C) II and V    (D) IV only    (E) V only

**Figure 12** Question 19 in TUG-K

When taking these two questions into consideration in more detail, it was discovered that both Question 12 (shown in Figure 11) and Question 19 (shown in Figure 12) share the same graph. Question 12 inquired about identifying a graph with constant velocity, while Question 19 asked about identifying a graph with constant and non-zero acceleration. The analysis revealed that Question 19 is more complex than Question 12, as understanding the graph regarding velocity is a prerequisite for correctly answering Question 19, which, if done correctly, can also result in a more accurate answer to Question 12. Due to their interconnectedness, the correlation between these two questions is high.

The analysis of item relationships using RapidMiner has provided valuable insights into how certain questions on the TUG-K test are interconnected, particularly between Questions 5 and 20 and between Questions 12 and 19. These associations highlight underlying conceptual links in students' understanding, such as the connections between interpreting the slope and area under a graph. This type of insight, which traditional item analysis methods in Excel cannot reveal, demonstrates the enhanced capability of RapidMiner in uncovering deeper patterns in educational data. By identifying these correlations, the study fulfills its second purpose—comparing the effectiveness of data analysis using RapidMiner with traditional item analysis in Excel—and underscores the advantage of employing advanced data mining techniques to gain richer insights into student learning and test item relationships.

## Conclusion

This study aimed to achieve two main objectives: first, to demonstrate the utility of advanced data analysis techniques using RapidMiner in educational research, and second, to compare

these techniques with traditional item analysis methods, such as those performed using Excel. Initially, the classical approach to item analysis was applied to the Test of Understanding Graphs in Kinematics (TUG-K), providing insights into the difficulty, discrimination, and point-biserial correlations of individual test items. While useful, these traditional methods are limited in their ability to reveal relationships between items or identify patterns in how students approach different concepts.

To overcome these limitations, association rule mining was used to analyze post-test responses with RapidMiner. This technique uncovered significant relationships between certain test items, such as the strong correlations between Questions 5 and 20 (velocity and the area under the curve) and between Questions 12 and 19 (motion graph interpretation). These associations provided deeper insights into the conceptual connections and common misunderstandings among students—insights that traditional item analysis techniques cannot reveal.

The comparison between these two techniques highlights the value of complementing traditional item analysis with more advanced data mining approaches. While classical methods offer foundational metrics for evaluating individual test items, techniques like association rule mining offer a richer understanding of how students engage with related concepts, providing deeper insights for improving both assessment and instruction.

## Implication for practice

For educators, these findings suggest that using advanced data analysis techniques, such as association rule mining, can reveal patterns of student misconceptions and highlight connections between different concepts. This allows for more targeted interventions and instructional strategies aimed at addressing specific areas where students struggle. Additionally, insights from this analysis can be used to refine assessments, ensuring they capture not only isolated knowledge but also the relationships between key concepts.

For researchers, this study demonstrates the benefit of combining traditional item analysis with advanced data mining techniques. Doing so can uncover more complex patterns in student responses and provide a more comprehensive view of how students interact with assessment items. Future research could apply these methods to larger datasets to further explore learning behaviors in various educational contexts.

In conclusion, by employing more advanced analysis techniques such as association rule mining, educators and researchers can gain a deeper understanding of student learning. These techniques offer the potential to move beyond surface-level metrics and provide more detailed insights into conceptual understanding, ultimately enhancing both educational research and teaching practices.

## ACKNOWLEDGEMENTS

## REFERENCES

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education, 33*, 1289-1312.

Araujo, I. S., Veit, E. A., & Moreira, M. A. (2008). Physics students' performance using computational modelling activities to improve kinematics graphs interpretation. *Computers & Education, 50*(4), 1128-1140.

Bachhal, P., Ahuja, S., & Gargrish, S. (2021). Educational data mining: A review. In *Journal of Physics: Conference Series* (Vol. 1950, No. 1, p. 012022). IOP Publishing.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American journal of Physics*, 62(8), 750-762.

Bollen, L., De Cock, M., Zuza, K., Guisasola, J., & van Kampen, P. (2016). Generalizing a categorization of students' interpretations of linear kinematics graphs. Physical Review Physics Education Research, 12(1), 010108.

Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions. *Physical Review Special Topics-Physics Education Research, 5*(2), 020103.

Fernández, D. B., & Luján-Mora, S. (2017, March). Comparison of applications for educational data mining in Engineering Education. In *2017 IEEE World Engineering Education Conference (EDUNINE)* (pp. 81-85). IEEE.

Gao, S., Li, L., & Wu, J. (2021). Application Analysis of Data Mining Technology in the Field of Learning Analysis in China-From the Perspective of Content Analysis. *The Educational Review, USA, 5*(6), 155-163.

Hestenes, D., & Wells, M. (1992). A mechanics baseline test. *The Physics Teacher, 30*, 159-166.

Klein, P., Becker, S., Küchemann, S., & Kuhn, J. (2021). Test of understanding graphs in kinematics: Item objectives confirmed by clustering eye movement transitions. *Physical Review Physics Education Research*, 17(1), 013102.

Márquez-Vera, C., Morales, C. R., & Soto, S. V. (2013). Predicting school failure and dropout by using data mining techniques. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 8*(1), 7-14.

Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (applications and reviews), 40*(6), 601-618.

Salloum, S. A., Alshurideh, M., Elnagar, A., & Shaalan, K. (2020). Mining in educational data: review and future directions. In *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020)* (pp. 92-102). Springer International Publishing.

Singh, C., & Rosengrant, D. (2003). Multiple-choice test of energy and momentum concepts. *American Journal of Physics, 71*, 607-617.

Slater, S., Joksimović, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics, 42*(1), 85-106.

Susac, A., Bubic, A., Kazotti, E., Planinic, M., & Palmovic, M. (2018). Student understanding of graph slope and area under a graph: A comparison of physics and nonphysics students. *Physical Review Physics Education Research, 14*(2), 020109.

Thornton, R., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics, 66*, 338-352.

Tongchai, A., Sharma, M. D., Johnston, I. D., Arayathanitkul, K., & Soankwan, C. (2009). Developing, evaluating and demonstrating the use of a conceptual survey in mechanical waves. *International Journal of Science Education*, *31*, 2437-2457.

Wattanakasiwich, P., Taleab, P., Sharma, M. D., & Johnston, I. D. (2013). Construction and implementation of a conceptual survey in thermodynamics. *International Journal of Innovation in Science and Mathematics Education*, 21(1).

Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (Vol. 1, pp. 29-39).

Wuttiprom, S., Sharma, M. D., Johnston, I. D., Chitaree, R., & Soankwan, C. (2009). Development and use of a conceptual survey in introductory quantum physics. *International Journal of Science Education, 31*, 631-654.

Yang, J., DeVore, S., Hewagallage, D., Miller, P., Ryan, Q. X., & Stewart, J. (2020). Using machine learning to identify the most at-risk students in physics classes. *Physical Review Physics Education Research, 16*(2), 020130.

Zabriskie, C., Yang, J., DeVore, S., & Stewart, J. (2019). Using machine learning to predict physics course outcomes. *Physical Review Physics Education Research, 15*(2), 020120.

Zavala, G., Tejeda, S., Barniol, P., & Beichner, R. J. (2017). Modifying the test of understanding graphs in kinematics. *Physical Review Physics Education Research, 13*(2), 020111.