ACTA DIDACTICA NAPOCENSIA
Volume 16, Number 2, 2023 - DOI: 10.24193/adn.16.2.9

DETERMINATION OF ASSOCIATION RULES FOR MATHEMATICAL SKILLS

Hanife Merve ERDOĞAN, Nazan SEZEN YÜKSEL

Abstract: The aim of this study is to classify the subjects and skills of middle school mathematics course in the context of MATH Taxonomy and to determine their relations. For this purpose, the questions and answers related to the mathematics subtest of a national exam were analyzed over the answers of 20154 students. The study continued with the analysis of the data set containing the answer patterns of the students according to the rules of association of data mining methods by using RapidMiner program and FP-Growth algorithm. The results show that students with any level of skill in MATH taxonomy related to mathematical skills could perform skills in sub-categories at a high rate. Students who have attainment in the learning areas provided later in the mathematics program had a high rate of knowledge in the previously provided learning areas. However, this result is not valid for Probability of Simple Events.

Key words: mathematical skills, data mining, association rules, FP-Growth algorithm, RapidMiner.

1. Introduction

Data is produced in almost every moment of daily life with technological developments. The data produced is often stored unconsciously and this data may contain various information that is not apparent at first glance. Different and confidential information can only be accessed after processing, analyzing, and interpreting the big data obtained. The process of accessing information from this data is defined as data mining. The data mining process reveals information that is hidden in large data sets and presents meaningful and useful patterns to users (Han, Kanber & Pei, 2012). Processed, analyzed, and interpreted data is used in different fields, such as health (Yıldırım, Uludağ, &Görür, 2008), engineering (Eker, 2016), insurance (Söylemez, Doğan & Özcan, 2016) and education (Kurt-Pehlivanoğlu & Duru, 2015; Şevgin 2020). This process in education is expressed as educational data mining.

As seen in Figure 1, the cycle of educational data mining is illustrated by Romero and Ventura (2007, p.136). The data processed in educational data mining may be related to students, teachers, administrators, or parents. The outputs obtained at the end of the data mining process may be used by one of these target groups who produce this data. It is expected that the outputs will provide meaningful, qualified, useful information to people who can be expressed as stakeholders of the education system (Bilen et al., 2014).

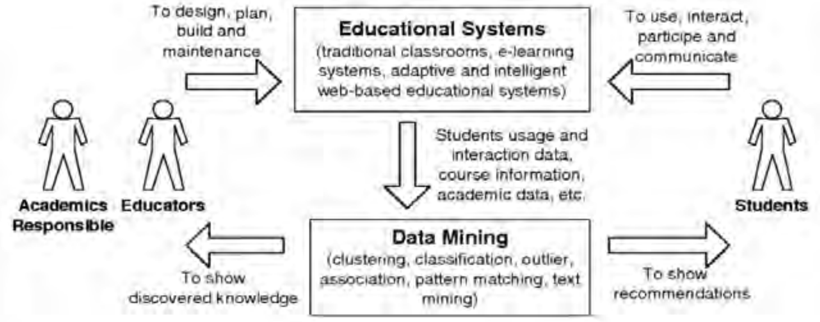

Figure 1. *The application of data mining in education (Romero & Ventura, 2007, p.136)*

Received March 2023.
Cite as: Erdoğan, H. M. & Sezen Yüksel, N. (2023). Determination of association rules for mathematical skills. *Acta Didactica Napocensia,* 16(2), 113-128, https://doi.org/10.24193/adn.16.2.9

Similarly, as demostrated in Figure 2, García, Romero, Ventura and Castro (2011), who defined educational data mining as the process of converting raw data obtained from educational systems into information that educational software, developers, teachers and researchers can use, described this process in a more cyclical way:

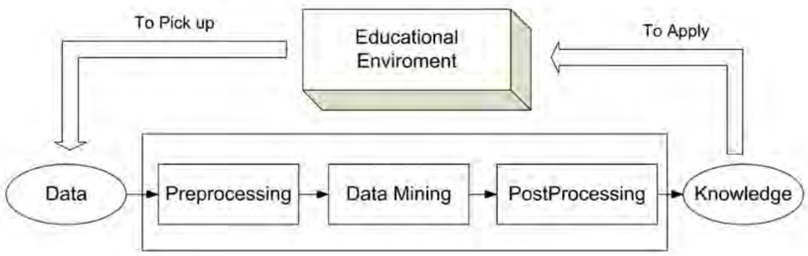

Figure 2. *Educational data mining process (García vd., 2011)*

Students produce educational data in classrooms with online or written exams such as Basic Education to Secondary Education Transition (TEOG) exam. TEOG exam, which was a central examination conducted throughout the country, could be considered an important data provider in the Turkish education system. TEOG exam, which focused on the simultaneous implementation of educational programs, differed in this sense from many measurement and evaluation applications (PISA, TIMSS, PIRLS) that are applied internationally and focus on skills while dealing with the acquisition of achievements in educational programs. TEOG also disseminated the evaluation of achievement throughout the process and ensured the simultaneous implementation of the curriculum across the country. TEOG had been implemented since 2013-2014 academic year and terminated as of 2017-2018 academic year. Within the scope of this exam, over a million 8th grade students across Türkiye took mathematics exams as well as the other lessons every year, which were applied in first term and second term. Following the evaluation of the data of these exams, students had been placed in high school education.

In the literature, studies related to the TEOG exam have been examined in various areas; such as the relationship between mathematics course success and mathematics test success applied in the exam (Çağlar, 2015; Çelikel and Karakuş, 2017; Delioğlu, 2017; Karakoç and Köse, 2018), the suitability of TEOG mathematics exam questions for mathematics curriculum (Bağcı, 2016; Başol et al. , 2016) and the comparison of TEOG mathematics questions according to the updated Bloom Taxonomy (Altun, 2016; Başol, Balgamış, Karlı and Öz, 2016; Dalak, 2015; Karaman, 2016; Karaman and Bindak, 2016; Yakalı, 2016). In addition, there is a study conducted by Can (2017) that examines the achievements of the TEOG exam with data mining. The results provided in this study, were only related to learning outcomes in Math curriculum. As seen in the literature, a variety of studies on students' TEOG exam conducted during the years when the TEOG exams were applied. But a sufficient number of studies using data mining were not conducted at the stage of evaluating TEOG exams data.

TEOG was prepared to measure the level of acquisition in the curriculum. The acquisitions include the basic math skills that students are expected to know. MATH (Mathematical Assessment Task Hierarchy) taxonomy, developed from BLOOM's taxonomy and used to measure mathematical knowledge, skills and abilities, can be used to evaluate a course where cognitive processes such as mathematics are at the forefront. MATH taxonomy was developed specifically for mathematics in order to make an accurate assessment in mathematics as an exchange of BLOOM taxonomy (Wood, Smith, Petocz, & Reid, 2002). There are 3 groups in MATH taxonomy as A, B and C. Group A consists of three categories, Group B consists of two categories, and Group C consists of three categories (D'Souza and Wood, 2003; Smith, Wood, Coupland, Stephenson, Crawford & Ball, 1996; Wood and Smith, 2007). MATH Taxonomy groups and categories that belong to each group are shown in Figure 3.

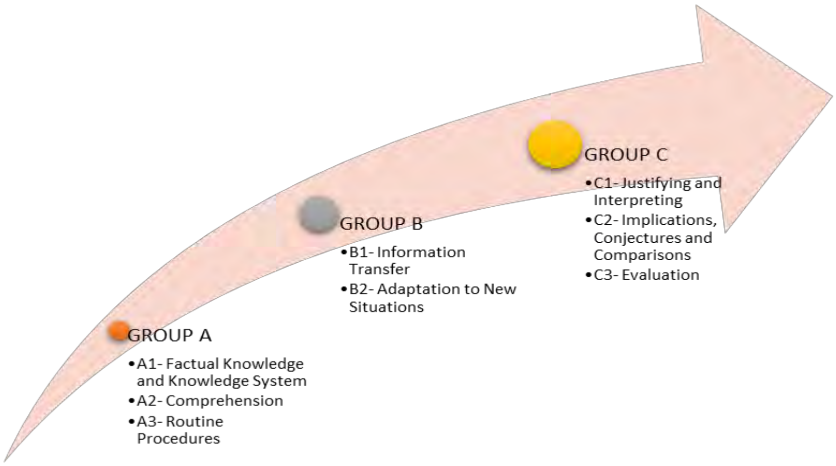

Figure 3. *MATH Taxonomy groups and categories that belong to each group*

In group A; categories such as factual information and information system, comprehension, routine operations, respectively, are included according to the situations requiring high-level skills. In group B, which requires higher-level skills than in group A, there are categories of knowledge transfer and adaptation to new situations, again according to situations requiring higher-level skills. In group C, which requires a higher level of skill compared to groups A and B, the justifying and interpreting category is followed by Implications, Conjectures and Comparisons category, which require higher level of skill, and the taxonomy, which is the last category, is completed (D'Souza and Wood, 2003).

Smith et al. (1996) explained the categories in MATH taxonomy as follows:

The “Factual Knowledge and Knowledge System” requires remembering a formula, definition or theorem, and the skills in this category consist only of re-imagining previously learned knowledge. “Comprehension” is a category in which it is somewhat possible to reproduce information without understanding. It requires understanding the importance of symbols in a formula, recognizing examples and counter examples of a mathematical concept or goal, and making a simple definition. “Use of Routine Processes” requires proper use of information beyond simple factual recall.

“Information Transfer” requires being able to convert information from one form to another form such as verbal to digital, numerical to graphic etc. “Adapting to New Situations” requires the ability to select and apply appropriate methods or information in new situations.

“Justifying and interpreting” requires verification and justification of a result given or reached by the student. "Implications, Conjectures and Comparisons" requires students to make inferences and predictions from a given or achieved result or situation, and to verify or prove it. “Evaluation” is the ability to verify the value of information for a given purpose based on certain criteria. This category includes the ability of students to provide or determine criteria, to judge, and to consistently discuss the principles of an algorithm. It also includes building something beyond what is given, reconstructing information as a new whole, and perceiving an unprecedented application regarding the application of information.

MATH taxonomy can be used to decipher the hierarchy between mathematical skills. When the studies on MATH Taxonomy are analyzed, it is seen that both the teacher-prepared exam questions (Kesgin, 2011; Aygün et al., 2016) and the questions presented in large-scale exams (Uğurel, Moral, and Kesgin, 2012; Moralı, Karaduman and Uğurel, 2014) are categorised according to taxonomy. For example, Uğurel et al. (2012) made a comparative analysis of the mathematics questions in the Student Selection Examination for Secondary Education Institutions (OKS) and Placement Exam (SBS) which were two of the ones utilized in our nation and also TIMSS exams within the framework of 'MATH Taxonomy'.

In this study, the focus is to discover important information that is considered to be hidden in the data collected within the scope of the TEOG exam, and the level of math skills included in the TEOG exam and the relationship between the math skills. In this study, the following research questions have been addressed:

• How is the distribution of the exam questions in terms of the categories of MATH Taxonomy and learning areas in mathematic curriculum?

• While examining the result of data, are there any associations observed between the mathematical skills of the students in MATH Taxonomy?

• While examining the result of data, are there any associations observed between the learning areas in the curriculum?

2. Method

The research model used in the study, the population and sampling, data collection and data analysis are provided in the following sections.

2.1. Research Model

In this study, which was carried out with the aim of determining the association between mathematical skills, the process and the visual element related to the process are shown in Figure 4.

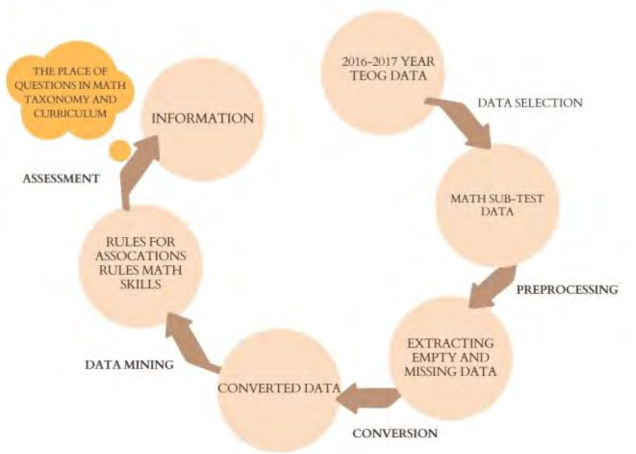

Figure 4. *The processes of the research*

Firstly, 2016-2017 TEOG exam data has been provided by an official letter from Ministry of Education Directorate General for Measurement, Assessment and Examination Services (ODSHGM). Mathematics subtest results were then selected from all data. Next, empty and missing data had been extracted from the data. Then, the correct answers of the students were coded as 1 and the wrong answers and left blanks as 0. Finally, converted data was mining by using RapidMiner program and FP-Growth algorithm. After these processes, the hidden information was obtained by evaluating both the MATH taxonomy and math curriculum.

2.2. The Population of the Study

According to the statistics of the Ministry of National Education for the academic year of 2016-2017, the total number of students expected to take the exam was 1.179.803, consists of 610.010 male and 569.793 female students. However, the number of students could not take the exam for various reasons. Therefore, 1.174.247 students in the first term and 1.153.551 students in the second term participated in the exams in the 2016-2017 academic year. However, the study group consists of the the maximum number of data were obtained from ODSHGM, was allowed to use by researchers.

2.3. Data Collection

Within the framework of TEOG, data from two time periods were collected with mathematics subtests. On November 23, 2016, and on April 27, 2017, respectively. The mathematics subtest had 20 questions, which means that there were 20 variables for each student. The data of 10,000 students for the first term and 10,154 students for the second term were chosen at random from the student exam data. In this manner, a total of two matrices for the 2016–2017 academic year were obtained: one matrix of size 20x10000 for the first term and one of size 20x10154 for the second term.

2.4. Data Collection Tools

The data collection tools were the mathematics subtests. Each mathematics subtest used as a data collection tool contains 20 questions. The participants were given multiple choice questions with four options as part of the data collection tools.

In Table 1, the test statistics for the mathematics subtest used in the first and second terms of the 2016-2017 academic year are provided.

Table 1. *2016-2017 Mathematics Subtest Test Statistics*

Subtest	Reliability Coefficient	Average Difficulty Index	Average Discrimination Index (Double Series)
Mathematics (1st Term)	0.85	0.49	0.66
Mathematics (2nd Term)	0.86	0.55	0.68

Table 1 shows that both terms of the mathematics subtest had reliability coefficients of 0.85 and 0.86, respectively, indicating that both terms were reliable. Both terms' mathematics subtests had medium difficulty, with mean difficulty indices of 0.49 for the first term and 0.55 for the second. Additionally, the first and second periods' respective average discrimination indices were 0.66 and 0.68, respectively. They were both different subtests.

2.5. Data Analysis

In this study, which was carried out with the aim of determining the association rules regarding mathematical skills, the following steps were followed in the analysis phase of the data, respectively:

a) The mathematics subtest items in booklet A were matched to the learnings outcomes.
b) The questions in the subtest were classified by the researcher in accordance with the MATH Taxonomy.
c) For the study's connection of the questions with the accomplishments and classification of them in accordance with the MATH Taxonomy, the opinions of three specialists were consulted.
d) Two separate data sets obtained from Ministry of National Education (MoNE) ODSHGM were analyzed with the RapidMiner program developed by Dortmund Technical University.

The FP Growth algorithm, which is frequently preferred in the determination of the association rules, was used in the analysis of data performed with the RapidMiner program.

3. Results

3.1. Findings Regarding MATH Taxonomy Category and Learning Areas in Mathematic Curriculum Distributions

The process of categorizing the questions in the TEOG mathematics subtest into MATH Taxonomy categories was initially made by the researcher and presented to the opinion of three experts. The taxonomy catagory and sub-learning area to which each questions belongs were listed in the expert opinion form as "1: Appropriate, 2: Not Appropriate."

Krippendorff's Alpha coefficient was calculated for the analysis of expert opinions (3 experts and 2 categories). As a result of the expert opinion analysis, the Krippendorff's Alpha coefficient was calculated as 0.71 for the first term data and 0.68 for the second term data. Accordingly, it is possible to say that expert opinions were moderately compatible. This phase was followed by the evaluation of the opinions of the experts received as a whole and the examination of the items that were disagreed individually. By re-evaluating the questions with the experts, the sub-learning area and taxonomy category that the disagreed items belong to were defined, and consensus was achieved.

In the first and second terms of the 2016-2017 academic year, Table 2 demonstrates how the questions on math subtests connected to the MATH Taxonomy categories and groups were distributed. In addition, the number of questions included in the categories and the distribution percentages of the questions are also given in Table 2.

Table 2. *Distribution of Questions in Mathematics Subtests According to Math Taxonomy Categories*

		1ST TERM							**2ND TERM**							
		QUESTION NUMBER					**TOTAL (f)**	**%**	**QUESTION NUMBER**						**TOTAL (f)**	**%**
GROUP A	Factual Knowledge and Knowledge System	1	9	18	-	-	3	15	-	-	-	-	-	-	0	0
	Comprehension	11	-	-	-	-	1	5	1	8	18	-	-	-	3	15
	Routine Procedures	3	6	15	17	20	5	25	6	10	11	14	16	19	6	30
GROUP B	Information Transfer	2	10	12	-	-	3	15	2	4	5	7	17	-	5	25
	Adaptation to New Situations	4	5	7	13	19	5	25	1	9	-	-	-	-	2	5
GROUP C	Justifying and Interpreting	8	14	-	-	-	2	10	2	3	12	-	-	-	3	10
	Implications, Conjectures and Comparisons	16	-	-	-	-	1	5	3	13	15	20	-	-	4	15
	Evaluation	-	-	-	-	-	0	0	-	-	-	-	-	-	0	0
TOTAL							**20**	**100**							**20**	**100**

As indicated in Table 2, the number of the questions on the mathematics subtest for the first term were divided into mostly Group A, followed by Group B, and then Group C. Additionally, the questions in the second term's mathematics subtest were designed to fall into Group A and resemble those from the first term more closely. In Group C, there will be more questions in the second term than there were in the first.

In Figure 5, the weighted distribution of the questions in mathematics subtests in MATH Taxonomy categories and groups are shown.

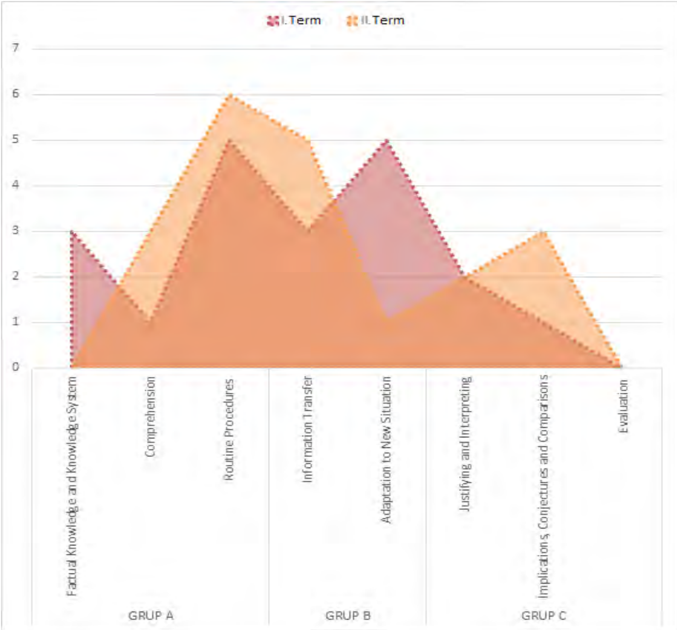

Figure 5. *Distribution of the questions in TEOG mathematics subtests to MATH Taxonomy categories*

Figure 6 for the mathematics subtests shows the distribution of the questions throughout the learning areas and sub-learning areas of the mathematics curriculum.

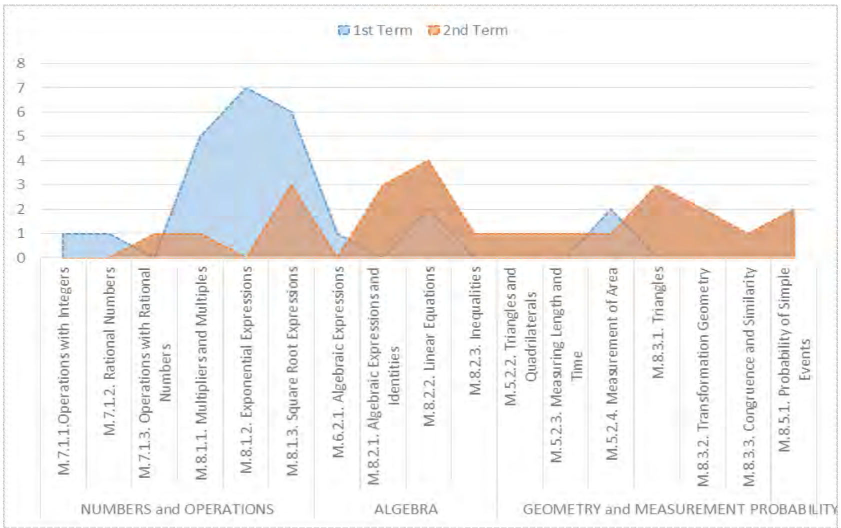

Figure 6. *Distribution of the questions in mathematics curriculum*

The "Numbers and Operations" study area is where the majority of the first term's questions are located, as seen in Figure 6. The questions are not equally distributed across the different sub-learning areas, and they are mostly collected from the "Exponential Expressions" sub-learning area of the "Numbers and Operations" learning area. Additionally, no questions from the "Probability" study area are provided for the first term. The second term's test questions are mostly focused on the subject of "Algebra." All four learning areas were used to prepare questions for the second term. Compared to the first term, there is a more balanced distribution of the questions.

3.2. Findings on Association Rules for the First Term

Applying the FP Growth algorithm on the 10.000 data from the first term in the RapidMiner program generated the association rules. The rules obtained from first term mathematics subtest data with the confidence value of above 0.97 through RapidMiner program and the association of questions are given in Figure 7.

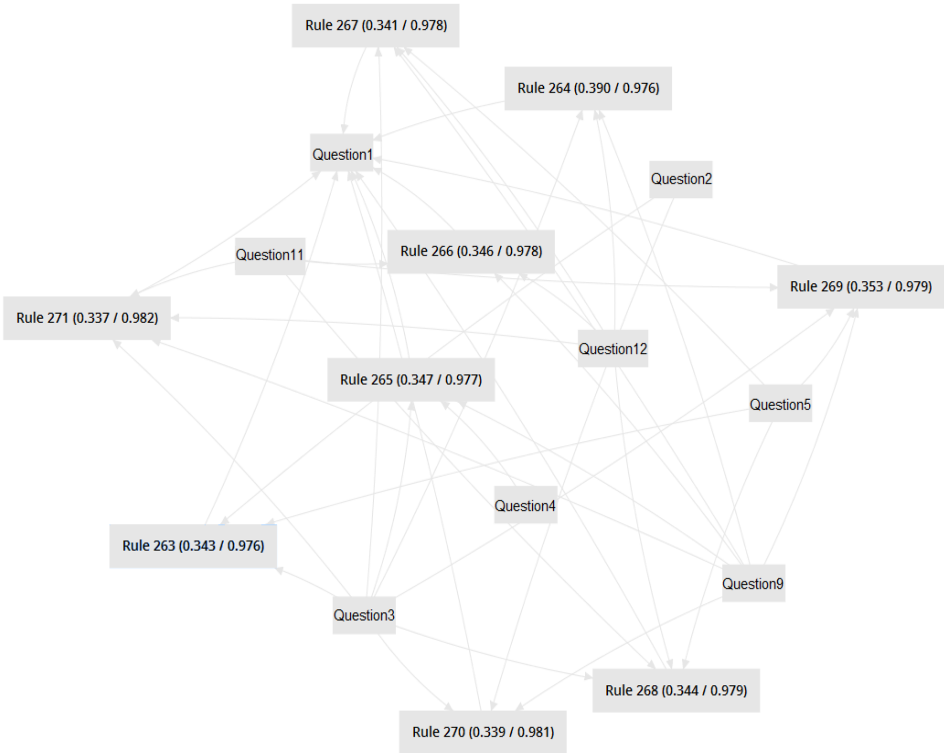

Figure 7. *The rules obtained through the RapidMiner regarding the first term mathematics subtest data*

The associations between the questions are also given in Figure 7. The first of the numerical expressions given in parentheses regarding the association rules between the questions shows the support value, and the second one shows the confidence value. The support value expresses the availability of the obtained rules, while the confidence value indicates their correctness. In other words, the support value explains how often the two questions appear together in the entire data set. When one is answered correctly, the confidence value explains the situation when the other is answered correctly. Looking at this network of relationships mentioned in Figure 7, it seems that the first question most often shows a relationship with other questions.

The association rules, which are obtained through the RapidMiner program with the first term mathematics subtest data and have a support value greater than 0.5, are given in Table 3.

Table 3. *Association Rules Regarding the First Term*

Rule No	Prior	Result	Support Value	Confidence Value
1	Question3	Question1	0.69	0.92
2	Question5	Question1	0.61	0.91
3	Question11	Question1	0.57	0.93
4	Question5	Question3	0.56	0.85
5	Question9	Question1	0.54	0.93
6	Question1, Question5	Question3	0.53	0.88
7	Question3, Question5	Question1	0.53	0.95
8	Question11	Question3	0.53	0.87
9	Question12	Question1	0.53	0.92
10	Question9	Question3	0.51	0.89
11	Question11	Question1, Question3	0.51	0.84
12	Question1, Question11	Question3	0.51	0.90

13	Question3, Question11	Question1	0.51	0.96
14	Question12	Question3	0.50	0.88

The association rules given in Table 3 can be explained as follows:

We can give the Rule 1 as shown below as an example of the interpretation of the association rules,

Rule 1: 69% of the students answered the 3rd and 1st questions correctly together. 92% of the students who answered the 3rd question correctly, answered the 1st question correctly. In other words, students who have the skills specified in the “Routine Procedures” category also have the skills given in the "Factual Knowledge and Knowledge System" category with a 92% probability. This association also reveals that 92% of the students who answered the question in the "Multipliers and Multiples" sub-learning area correctly, answered the question in the "Operations with Integers" sub-learning area correctly.

Rule 13: 51% of the students answered the 3rd, 11th and 1st questions correctly together. 96% of the students who answered the 3rd and 11th questions correctly answered the 1st question correctly. This rule shows that students who have skills specified in the “Routine Procedures” category and “Comprehension” category also have the skills given in the "Factual Knowledge and Knowledge category with a 96% probability. At the same time, a unity is also revealed in terms of learning areas. 96% of the students who answered the questions in the “Exponential Expressions” and “Operations on Integers” sub-learning areas correctly, answered the question in the “Multipliers and Multiples” sub-learning area correctly.

The rules related to the MATH Taxonomy obtained according to the first term mathematics subtest data can be summarized as in Figure 8.

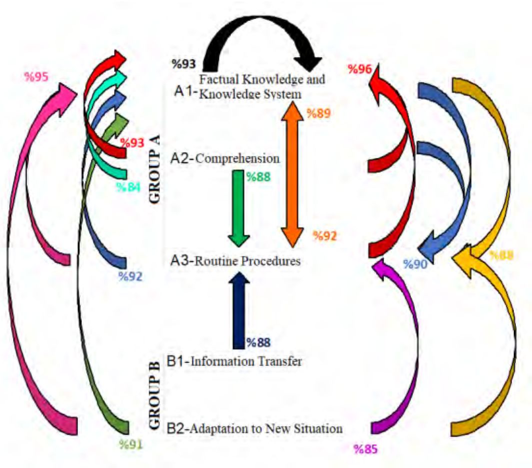

Figure 8. *The rules related to the MATH Taxonomy obtained according to the first term mathematics subtest data*

The rules related to the sub-learning areas obtained according to the first term mathematics subtest data can be summarized as in Figure 9.

Figure 9. *The rules related to the sub-learning areas obtained according to the first term mathematics subtest data*

3.3. Findings on Association Rules for the Second Term

Applying the FP Growth algorithm on the 10.154 data from the second term in the RapidMiner program generated the association rules. The rules with the confidence value of above 0.98 obtained through RapidMiner program from second term mathematics subtest data and the association of items are given in Figure 10.

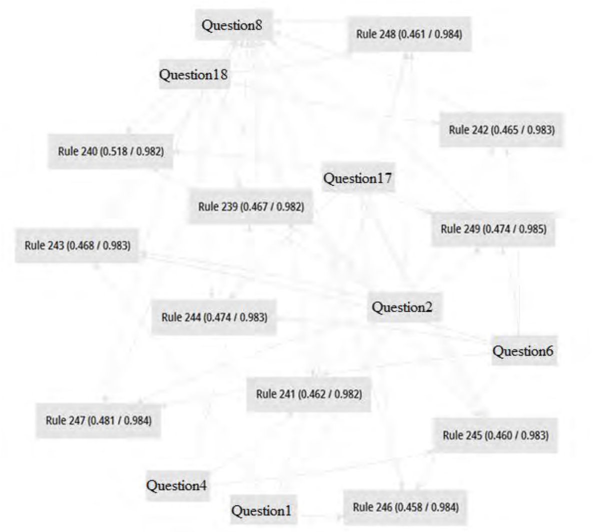

Figure 10. *The rules obtained through the RapidMiner regarding the second term mathematics subtest data*

The association rules, which are obtained through the RapidMiner program with the second term mathematics subtest data and have a support value greater than 0.6, are given in Table 4.

Table 4. *Association Rules Regarding the Second Term*

Rule No	Prior	Result	Support Value	Confidence Value
1	Question17	Question8	0.71	0.96
2	Question1	Question8	0.71	0.95
3	Question18	Question8	0.70	0.95
4	Question2	Question8	0.69	0.95
5	Question6	Question8	0.63	0.96
6	Question4	Question8	0.61	0.95
7	Question18	Question2	0.61	0.83
8	Question2	Question18	0.61	0.83

9	Question18	Question17	0.61	0.83
10	Question2	Question1	0.60	0.83
11	Question2	Question17	0.60	0.82
12	Question18	Question1	0.60	0.82

We can give the Rule 1 as shown below as an example to the association rules given in Table 4.

Rule1: 71% of the students answered the 17th and 8th questions correctly together. 96% of the students who answered the 17th question correctly, answered the 8th question correctly. In other words, students who have the skills specified in the "Information Transfer" category also have the skills given in the "Comprehension" category with a 96% probability. This association also reveals that 96% of the students who answered the question in the "Triangles" sub-learning area correctly, answered the question in the "Probability of Simple Events" sub-learning area correctly. The association between these two questions, which are not directly related in terms of learning areas, becomes more significant when evaluated from a taxonomic point of view.

The rules related to the MATH Taxonomy obtained according to the second term mathematics subtest data can be summarized as in Figure 11.

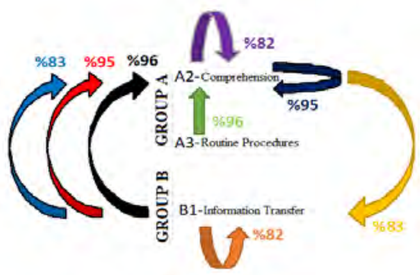

Figure 11. *The rules related to the MATH Taxonomy obtained according to the second term mathematics subtest data*

The rules related to the second term mathematics subtest data and the sub-learning areas in the mathematics curriculum can be summarized as in Figure 12.

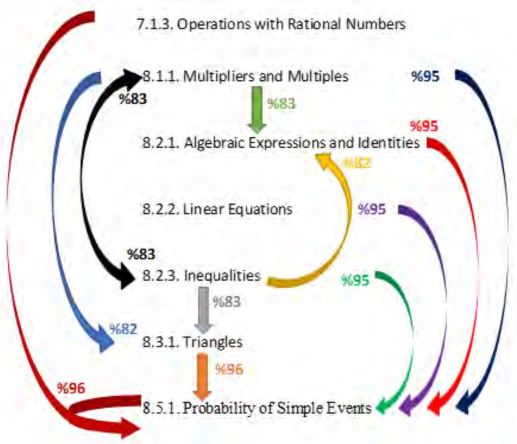

Figure 12. *The rules related to the sub-learning areas obtained according to the second term mathematics subtest data*

4. Discussion and Conclusion

The primary purpose of this study was to determine how the math subtest's questions were distributed throughout the MATH Taxonomy categories and curriculum learning areas. The second purpose was to use student answer patterns to show how categories in the MATH Taxonomy relate to one another. The final purpose was to show the connection between the different learning areas in math curriculum.

Firstly, the association of the questions in the mathematics subtest according to the MATH Taxonomy category was revealed in order to describe the exam. The majority of the questions in the mathematics subtest of the TEOG exam were in the "Routine Procedures" category. Similarly, Uğurel et al. (2012) find that the majority of the questions in the SBS 7 and SBS 8 national exams were related to "Routine Procedures". Aygün et al. (2016) also drew attention to the fact that two out of every three questions prepared by teachers in mathematics exams were questions about the routine procedures. The results show that the majority of the questions on both the large-scale exam and the teacher-prepared exams are designed to assess Routine Procedures category in the MATH Taxonomy. The questions in the large scale exams are more likely to assess the skills on the lower steps of the taxonomy due to the age level of the group and the structure of the test technique. Overall, the results show that the majority of questions were prepared in the Group A category, while the minority of questions were prepared in the Group B and Group C categories in line with the studies of Moralı et al. (2014), Kesgin (2011), Aygün et al. (2016) and Mullis et al. (2020). It can be said that the structure and purpose of the TEOG exam are also not conducive to assess high-level mathematical skills.

The TEOG exam's questions were developed in accordance with the curriculum's objectives. As a consequence, the distribution of the questions in relation to the mathematical curriculum's learning areas is provided. The first term was mostly based on "Numbers and Operations''and the second term was mostly based on "Geometry and Measurement" learning areas. This result is similar to the results of the study conducted by Bağcı (2016). In her study, 2013-2014 TEOG mathematics sub-test question distribution was examined and it was concluded that the questions in the first term mainly related to the square root numbers (Numbers and Operations), and in the second term, questions are related to triangles and equations (Geometry and Measurement-Algebra). This result regarding the inclusion of more questions in the "Numbers and Operations" learning area and the "Algebra" learning area is also similar to the TIMSS 2019 results. The weight of these two areas in the exam in TIMSS 2019 is 60% (Mullis et al. 2020).

As a consequence, it was discovered that "Numbers and Operations" was the most frequently utilized learning area to assess these skills, and that "Using Regular Operations" skills were the most frequently measured. It is also noteworthy that there are variety of learning areas that are widely used in the measurement of certain skills. Therefore, those who prepare large-scale exams and teachers can create different types of questions for the skills they want to measure when preparing exam questions, the connection of the questions to everyday life can be increased and the relationship with different learning areas can be ensured.

Following the results of the distribution of mathematical skills and learning areas in the TEOG exam, association rules were reached by the analysis of student response patterns. The relationships between the mathematical skills that students have and these skills are as follows dec:

In the mathematics subtest in the first term, the categories of the five questions answered most correctly in the MATH taxonomy are respectively listed as "Knowledge and Knowledge System", "Routine Procedures", "Adaptation to New Situations" and "Comprehension". The least correctly answered question is in the category of "Implications, Conjectures and Comparisons". In other words, it is seen that the skills in Group A and Group B categories are acquired more; the acquisition rate of the higher-level skills in Group C is low. In the second term, the categories of the five questions answered most correctly in the MATH taxonomy are respectively listed as "Comprehension" and "Information Transfer", ". The least correctly answered question is in the category of "Justifying and Interpreting". These results showed that the most correctly answered questions are mainly in category A; the least correctly answered questions are in Category C. The results of a study conducted by Kesgin (2011) with pre-service mathematics teachers are in line with this result. According to his study, pre-service mathematics teachers prepared the exam questions mainly in Group A in the MATH Taxonomy compared to the Group B and C.

In addition, students with any level of skills in MATH taxonomy are able to perform skills in subcategories at a high rate. And this result confirms Smith et al. (1996)'s work. Smith et al. (1996)

stated that there is a hierarchy between the groups and the categories in the groups. He noted that Group A contains lower level of mathematical skills than Group B.

In the study, the group with the most correct answers was Group A and the group with the least correct answer was Group C. Smith et al. (1996) is an indicator of the hierarchy between skills. However, in this study, it has been determined that students with skills required for the “Knowledge and Knowledge System” category can fulfill the skills in “Routine Procedures”, that students with skills required for the “Comprehension” category can fulfill the skills in “Routine Procedures”, and that students with skills required for the “Knowledge and Knowledge System” and "Comprehension" categories together can fulfill the skills in “Routine Procedures”. These results show that students with the skills required for the sub-categories in MATH taxonomy can correctly answer questions in categories that require higher skills in the same group. This is a sign that the mathematical skills included in the “Routine Procedures” can go beyond the hierarchical structure. This finding does not overlap with the opinion stated by Smith et al. (1996) that there is a hierarchy between the categories in the same group.

The relationship of the most correctly answered questions with the learning areas in the curriculum can be summarized as follows: students who have educational acquisition in the learning areas provided later in the mathematics curriculum have a high rate of knowledge in the previously provided learning areas. The "Operation with Integrers" sub-learning area has prerequisites for many other learning areas, which shows the value of fundamental knowledge when the facts from the first term are evaluated broadly. This result has shown that it is extremely important for students to have basic operation skills before focusing on intensive learning areas.

The most striking situation in the association analysis of the second term data is that the "Probability of Simple Events” sub-learning area is associated with a large number of sub-learning areas. Accordingly, it is inevitable for students to encounter failure even if they are deprived of pre-learning even on a subject that has its unique rules and algorithms such as probability.

In addition, the questions answered correctly by the students in the first term showed more correlation with the different groups and levels of the MATH taxonomy. On the other hand, the questions answered correctly in the second term were mostly related to the steps of Group A and the first step of Group B. This shows that questions in the first term exam were able to measure different and high-level skills. The fact that the questions were not prepared in the first term to cover all the learning areas in the curriculum is a sign that a higher level of question was prepared when measuring mathematical skills compared to the second term. The content validity of the test was increased by the fact that the questions asked in the second term were more balanced than the curriculum. However, the steps by which the questions were related to the skills in MATH taxonomy and the mathematical skills that were desired to be measured remained at lower levels. The mean difficulty index of the first-dec test (0.49) and the mean difficulty index of the second-dec test (0.55) also confirmed this difference between the two periods regarding the levels of mathematical skills measured.

5. Suggestions

Within the scope of the research, important results have been obtained regarding the relationship between mathematical skills and mathematics sub-learning areas. It should be taken into consideration by mathematics educators that the hierarchical structure of MATH Taxonomy appears even in exam conducted with large masses. In this context, it is recommended that mathematics educators and, in particular, mathematics teachers should consider this taxonomy when choosing the skills or learning areas, they expect from their students to exhibit. Because important information about which skills or learning areas should be acquired primarily is included in the taxonomy.

In the mathematic education, the importance of pre-learning is constantly emphasized. However, the full learning of a subject may not always occur even if the pre-learning about the subject is completed. In the light of the results, the clear determination of other topics related to a subject may be important for the full learning of a subject. For this reason, it is recommended for teachers to advance their lessons by establishing various associations based on the association relations of mathematics subjects.

In addition to this, during the preparation of mathematics curriculum, the time and place of the subjects can be re-evaluated in the curriculum by taking into consideration the associations regarding learning and sub-learning areas.

Furthermore, the following suggestions are recommended:

(a) The study can be expanded with associations with higher support value by expanding the data set used in the research.

(b) Associations regarding mathematical skills can be revealed by using the exam data applied within the scope of TEOG system in different years.

(c) Associations regarding mathematical skills can be evaluated together by including different types of exams in the research.

(d) The study can be detailed by comparing the results through different data mining analysis programs.

References

Altun, H. (2016). *TEOG sınavı matematik soruları hakkında öğretmen görüşlerinin incelenmesi ve yenilenmiş Bloom taksonomisine göre sınıflandırılması* [*Analysis of Teachers' Opinion about Maths Questions of TEOG Exam and Classification of the Questions According to the Renewed Bloom Taxonomy*] (Master thesis). Ondokuz Mayıs University, Samsun.

Aygün, B., Baran-Bulut, D., & İpek, A. S. (2016). İlköğretim matematik dersi sınav sorularının MATH taksonomisine göre analizi [The analysis of the primary school mathematics exam questions according to the MATH taxonomy]. *Turkish Journal of Computer and Mathematics Education*, 7(1), 62-88.

Bağcı, E. (2016). *TEOG sınavı matematik sorularının matematik öğretim programına uygunluğunun ve TEOG sisteminin hedeflerine ulaşma düzeyinin belirlenmesi* [*Alignment of TEOG mathematics questions to the mathematics curriculum and the level of attainment the goals of TEOG system*] (Master thesis). Ankara University, Ankara.

Başol, G., Balgamış, E., Karlı M.G. & Öz, F.B. (2016). Content analysis of TEOG mathematics items based on MONE attainments, TIMSS levels, and reformed Bloom Taxonomy [Content analysis of TEOG mathematics items based on MONE attainments, TIMSS levels, and reformed Bloom Taxonomy]. *Journal of Human Sience*, 13, 5945-5967.

Can, E. (2017). *Temel eğitimden ortaöğretime geçiş sınavı kazanımlarının veri madenciliği yöntemleri ile değerlendirilmesi* [*Evaluation of acquisitions of the transition from primary education to secondary education exam with data mining methods*] (Master thesis). Afyonkarahisar University, Afyon.

Çağlar, M. (2015). *Matematik dersi TEOG sınavları ile öğretmen sınavlarının bazı değişkenler açısından karşılaştırılması* [*The comparison of match lesson TEOG examination with teacher exams according to some variables*] (Master thesis). Düzce University, Düzce.

Çelikel, F. & Karakuş, M. (2017). TEOG sınavının matematik dersindeki akademik başarıyla ilişkisinin ve matematik dersi öğretim süreci üzerindeki etkilerinin incelenmesi [The Analyzing the Relevance of TEOG Exam to Academic Achievement and the Effects of TEOG Exam on Teaching Process of Math Class]. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education* Vol. 11, Issue 2, December 2017, pp. 1-18, 11, 1-18. DOI: 10.17522/balikesirnef.373133

D'Souza, S. M., & Wood, L. N. (2003). Designing assessment using the MATH taxonomy. In L. Bragg, C. Campbell, G. Herbert, & J. Mousely (Eds.), *Mathematics Education Research: Innovation, Networking, Opportunity: Proceedings of the 26th Annual Conference of MERGA Inc.* (pp. 294-301). Deakin University, Geelong, Australia.

Dalak, O. (2015). *TEOG sınav soruları ile 8.Sınıf öğretim programlarındaki ilgili kazanımların yenilenmiş Bloom taksonomisine göre incelenmesi* [*The study of the objectives in curriculum of 8th* *grade with the questions of TEOG in accordance with revised bloom taxonomy*] (Master thesis). Gaziantep University, Gaziantep.

Delioğlu H.N. (2017). *Sekizinci sınıf öğrencilerinin matematik başarısı ile sınav ve matematik kaygısı, matematiğe yönelik özyeterlik algısı arasındaki ilişki* [*8th grade students of success of mathhematics and test and mathematics anxiety, mathematics self efficacy of investigation*] (Master thesis). Adnan Menderes University, Aydın.

Eker, M.E. (2016). *Veri madenciliği yöntemleri ile uçuş biletleme analizi* [*Apriori algorithm implementation on exam data in data mining*] (Master thesis). Ondokuz Mayıs University, Samsun.

García, E., Romero, C., Ventura, S., & de Castro, C. (2011). A collaborative educational association rule mining tool. *The Internet and Higher Education*, 14, 77-88. DOI: 10.1016/j.iheduc.2010.07.006

Han, J., Kanber, M. & Pei, J. (2012). *Data Mining: Concepts and Techniques*, Third Edition. Waltham, USA: Morgan Kaufmann Publishers.

Karakoç, G. & Köse, İ.A. (2018). İlköğretim akademik başarı ölçüleri ile temel eğitimden ortaöğretime geçiş sınav puanları arasındaki ilişki [The Relationship between Elementary School Academic Achievement Levels and High School Entrance Exam Scores]. *Cumhuriyet International Journal of Education*, 7(2), 121–142.

Karaman, M. (2016). *İlköğretim matematik öğretmenlerinin sınav soruları ile TEOG matematik sorularının yenilenmiş Bloom taksonomisine göre analizi* [*Analysis of elementary mathematics teachers' examination questions and TEOG mathematics questions according to revised Bloom's Taxonomy*] (Master thesis). Gaziantep University, Gaziantep.

Karaman, M. & Bindak, R. (2016). İlköğretim matematik öğretmenlerinin sınav soruları ile TEOG matematik sorularının yenilenmiş Bloom taksonomisine göre analizi [Primary Mathematics Teachers' Exam Questions and Teog Mathematics Questions Analysis According To the Updated Bloom Taksonomy]. *Current Research Education*, 3, 51-65.

Kesgin, Ş. (2011). *Matematik öğretmen adaylarının soyut matematik dersindeki bilgilerinin MATH taksonomi çerçevesinde analizi* [*Analysis of abstract mathematics knowledge of mathematics teacher trainees through MATH taxonomy*] (Master thesis). Dokuz Eylül University, İzmir.

Kurt-Pehlivanoğlu, M. & Duru, N. (2015). Veri madenciliği teknikleri kullanılarak ortaokul öğrencilerinin sosyal ağ kullanım analizi: Kocaeli ili örneği [Use of Social Network Analysis of Middle School Students Using Data Mining Techniques: Kocaeli Sample]. *Duzce University Journal of Science and Technology*, 3, 508-517.

Mullis, I.V.S., Martin, M. O., Foy, P., Kelly, D.L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. United States: IEA, https://www.iea.nl/publications.

Moralı, H.S., Karaduman, H. & Uğurel, I. (2014). Matematik öğretmenliği alan bilgisi sınavlarındaki soruların MATH taksonomi çerçevesinde analizi [Analysis of questions of subject-related exams in mathematics education in terms of MATH Taxonomy]. *International conference on education in mathematics, science & technology*, 16-18 Mayıs 2014, Konya, Türkiye.

Özcan, A. (2021). Büyük veri: fırsatlar ve tehditler [Big Data: Opportunities and threats]. *TRT Akademi*, 6(11).

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, 33, 135-146.

Smith, G., Wood, L., Coupland, M., Stephenson, B., Crawford, K. & Ball, G. (1996). Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal of Mathematical Education in Science and Technology*, 27, 65-77, DOI: 10.1080/0020739960270109. Accessed from https://www.tandfonline.com/loi/tmes20

Söylemez, İ., Doğan, A., & Özcan, U. (2016). Trafik kazalarında birliktelik kuralı analizi: Ankara ili örneği [Association Rules on Traffic Accident: Case of Ankara]. *Ege academic review*, 16 (Special issue), 11-20. DOI: 10.21121/eab.2016OZEL24423.

Şevgin, H. (2020). *ABİDE 2016 fen başarısının yordanmasında MARS ve BRT veri madenciliği yöntemlerinin karşılaştırılması* [*Predicting the ABİDE 2016 science achievement: the comparison of MARS and BRT data mining methods*] (Doctoral dissertation). Gazi University, Ankara.

Uğurel, I., Moralı, H. S., & Kesgin, Ş. (2012). OKS, SBS ve TIMSS matematik sorularının 'MATH Taksonomi' çerçevesinde karşılaştırmalı analizi [A comparative analysis on the mathematics questions in OKS, SBS and TIMSS under the lens of MATH Taxonomy]. *Gaziantep University Journal of Social Sciences*, 423-444.

Wood, L.N., Smith, G.H., Petocz, P. & Reid, A. (2002). Correlation between student performance in linear algebra and categories of a taxonomy. In M. Boezi (Ed.), *2nd International Conference on the Teaching of Mathematics (At the Undergraduate Level)*, Crete, John Wiley.

Wood, l.N., & Smith, G.H. (2007). Perceptions of difficulty. *Second International Conference on the Teaching of Mathematics, 1-6 July 2002*, Hersonissos, Greece.

Yakalı, D. (2016). *TEOG sınavlarındaki matematik sorularının yenilenmiş Bloom taksonomisi ve öğretim programına göre değerlendirilmesi* [*Evaluation of math questions in TEOG exams according to renovated Bloom taxonomy and curriculum*] (Master thesis). Adnan Menderes University, Aydın.

Yıldırım, P., Uludağ, M. & Görür, A. (2008). Hastane bilgi sistemlerinde veri madenciliği [Data mining in hospital information systems]. *Academic Informatics 2008*, 429-434. Retrieved from https://ab.org.tr/ab08/bildiri/149.pdf

Authors

Hanife Merve ERDOĞAN, National education specialist in Ministry of National Education, Ankara (Türkiye). E-mail: h.merverdogan@gmail.com

Nazan SEZEN YÜKSEL, Associate professor in Hacettepe Univesity, Ankara (Türkiye). E-mail: nsezen@hacettepe.edu.tr