
The Eurasia Proceedings of Educational & Social Sciences (EPESS), 2017

Volume 8, Pages 19-29

IConSE 2017: International Conference on Science and Education (IConSE)

ACHIEVEMENT ANALYSIS OF A DATASET OF DISTANCE EDUCATION SAMPLES WITH APRIORI ALGORITHM

Tolga Demirhan
Trakya University

Ilker Hacıoğlu
Trakya University

Abstract: The purpose of this study is to analyze the academic success of distance education students using the apriori algorithm which is one of the association rule extraction algorithms in data mining. The students enrolled in Trakya University, Tunca Vocational High School (Distance Education) in 2016-2017 education year constitute the data set of the study. A questionnaire with 51 items, participation in which was on a voluntary basis, was used for the creation of the data set. On the data set, Apriori experiments were carried out separately for the attributes with course grades and the obtained rules were interpreted.

Keywords: Apriori, weka, success analysis, data mining, distance education

Introduction

Data mining is the process of discovering information by analyzing large amounts of data in various ways to reach useful information. (Pandeewari, L., Rajeswari, K., 2014) Data mining is used successfully in the academic field as well as disciplines such as health or engineering. The data mining area, which aims to reveal useful patterns previously unknown to the data set to understand the students better and for the evaluation of learning processes, is referred to as educational data mining. (Chan, A.Y.K., Chow, K.O., Cheung, K.S., 2008; Chandra, E., Nandhini, K., 2010) Association rule mining in educational data mining is a traditional and convenient research method used to determine the relationship among attributes in large databases. (Sebastian, A number of research studies are present in the literature, in which apriori algorithm and association rule mining are performed on educational data sets.

In the research study which is entitled "Student Success Analysis with Apriori Algorithm", conducted by Murat Karabatak and Melih Cevdet İnce, a success analysis was performed using the course grades of 250 students. In that study, missing, noisy and inconvenient data were cleaned by means of queries. 26 rules were discovered and the rules which were found interesting by the researchers were investigated. (Karabatak, M., İnce, M.C., 2004)

In Ahmet Selman Bozkır, Bilge Gök and Ebru Sezer's work, entitled "Identification of Factors Affecting Students' Use of the Internet for Educational Purposes by Data Mining Methods", it was attempted to discover the use of internet for educational purposes among undergraduate students. Data was collected by means of a questionnaire which was given to 380 randomly sampled students and the important rules were identified. In the study, it was seen that it was necessary to provide training for the students regarding effective online research methods and to improve the technological facilities of universities. (Bozkır, A.S., Gök, B., Sezer, E., 2008)

In a study by Nur Sena Tanrıverdi, which is entitled "A Data Mining Application to Improve the Graduation Success of ITU Mathematical Engineering Students", the relationship between the graduation success status of alumni and the success status in their departmental courses was attempted to be revealed by means of association rule applications with the help of the data obtained from the registrar of the university. (Tanrıverdi, N.S., 2013)

In a study by Nesibe Yağın, Emre Güngör and Nilüfer Yurtay, "Technical Elective Course Selection Analysis with Apriori Algorithm", the underlying reasons behind students' technical elective course selections were

- This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

- Selection and peer-review under responsibility of the Organizing Committee of the conference

*Corresponding author: Tolga Demirhan- E-mail: tolgademirhan@trakya.edu.tr

attempted to be found out by using the Apriori algorithm. As a result of the study, it was revealed that the algorithm could be used to discover the associations among exchange programs such as Erasmus / Farabi / Mevlana, summer school, practical training and distance education courses, which are all of particular concern to students. (Yalçın, N., Güngör, N., Yurtay, N. 2013)

In Mehmet Taş, M. Fatih Adak and Nilüfer Yurtay's study entitled "The Extraction of Association Rules from the Practical Training Data of Students and the Identification of their Practical Training Tendencies", Apriori Algorithm was applied on practical training records taken between 1999-2012 in order to identify students' practical training preferences. The results indicated 989 rules and it was observed that the students preferred software-related companies most frequently while the rest of them completed their training at their own university. (Taş, M., Adak, M.F., Yurtay, N., 2013)

A 51-item questionnaire was developed and given to the distance education students in the present study. The results obtained through these questionnaires were analyzed by means of the Apriori Algorithm, which is one of the association rule algorithms within Weka. With this regard, it was aimed within the context of this study to find out the rules and relationships among the academic success of the students and their demographic information, means of reaching academic resources and the frequency of using those resources.

Methods

The educational instances used in the study were created through the data collected from the students of the distance education computer programming program who volunteered to partake in the study. The students were enrolled in the computer programming program in 2016-2017 academic year. Within the context of the study, a questionnaire was given to the participants in order to collect data regarding their demographic information, study habits and academic success. Google Forms application was used to collect the questionnaire data. The responses of the participants were transferred to Excel via Google Forms for analyses. The errors in the open-ended questions were corrected during the preprocessing stage and the four indicators of the academic success variable, "Poor", "Mediocre", "Good", "Very Good" were reduced to two classes as "Success" and "Failure". The records which were formed were then converted into .arff format, which could be understood by Weka. The questionnaire items, codes and the arff file is given below.

Items	Code
What is your gender?	S1
How old are you?	S2
Please state what year of studies you are in.	S3
Please state your marital status.	S4
Please state the number of dependents as children you have.	S5
What is your most recently received qualification?	S6
Please state the city you live in.	S7
Are you employed?	S8
Which sector does your job belong to?	S9
For how many years have you been working?	S10
What is your average monthly income?	S11
What is your high school diploma grade?	S12
What is your average grade in Turkish, History and Foreign Language courses?	S13
What is your average grade in the Mathematics course?	S14
What is your average grade in the software based courses? (Intro. To Programming, Visual Programming, Internet Programming, Object Oriented Programming etc...)	S15
What is your average grade in theoretical vocational courses? (Server OS, Computer Architecture, Database Management Systems etc...)	S16
What is the type of company you completed your practical training?	S17
In which province is that company located?	S18
In which professional field does the company operate?	S19
In which unit did you complete your practical training?	S20

What was the weekly duration of work? (5-6-7 days...)	S21
How would you evaluate the duration of practical training?	S22
I believe that the practical training was beneficial in terms of knowledge and it contributed to my receiving information related to my profession.	S23
Were you able to use the knowledge you acquired in your department during your practical training?	S24
I believe that practical training helped me gain experience.	S25
Practical training encouraged me to learn more.	S26
I would recommend the company where I completed my practical training to my friends.	S27
The administration of practical training company were concerned with the trainees.	S28
Practical training contributed to my teamwork skills.	S29
Practical training contributed to the improvement of my written and oral communication skills.	S30
Practical training increased my sense of duty and responsibility.	S31
What is your level of computer use?	S32
How frequent do you use computers and the internet?	S33
What is your frequency of using the LMS?	S34
For how many hours in a day do you use the internet for learning purposes?	S35
Please select the learning material or materials you prefer while studying.	S36
The content of the courses are sufficient theoretically.	S37
The content of the courses are sufficient in terms of applications.	S38
The videos related to the courses are sufficient.	S39
I was able to communicate my requests and suggestions regarding the courses.	S40
The language of the courses are clear and understandable.	S41
Course content fits the aims of the program.	S42
Transport to the examination center and physical facilities are sufficient.	S43
Exam results are announced timely.	S44
I participate in the virtual class communication hours.	S45
I am content with getting in contact with the teaching staff during the virtual class communication hours.	S46
I am content with the Learning Management System (LMS) that is used.	S47
I know how to use the LMS.	S48
I can easily access the content visa LMS and the web site of the university.	S49
I receive informative e-mails and SMS regarding my distance education courses and exams.	S50
I find the announcements and notifications insufficient.	S51

Figure1. Table of questionnaire items – codes and responses

```

1 @relation 'Uzaktan_Egitim-weka.filters.unsupervised.attribute.Remove-R17-31-weka.filters.unsupervised.attrib
2
3 @attribute s3 {1sınıf,2sınıf}
4 @attribute s6 {lise_mezunu,universite_mezunu,yuksekokul_mezunu,yukse_lisans_doktora}
5 @attribute s12 {3_3.99,4_5,2_2.99}
6 @attribute s13 {basarili,basarisiz}
7 @attribute s14 {basarili,basarisiz}
8 @attribute s16 {basarili,basarisiz}
9 @attribute s32 {ileri_duzey,orta_duzey,temel_duzey}
10 @attribute s33 {cok_sik,sik,cok_az,ara_sira}
11 @attribute s34 {gunluk_takip_ederim,hafiraladikca_katilirim,haftalik_takip_ederim,sadece_sinavlara_katilirim}
12 @attribute s35 {2_3_saat,4_5_saat,0_1_saat,6_saat_ve_uzeri}
13 @attribute s36 {pdf_icerik_video,pdf,video,pdf_icerik,pdf_video,icerik_video,icerik}
14 @attribute s37 {3,4,2,5,1}
15 @attribute s38 {3,4,2,5,1}
16 @attribute s39 {3,2,5,1,4}
17 @attribute s40 {3,2,1,5,4}
18 @attribute s41 {3,4,5,2,1}
19 @attribute s42 {3,4,5,1,2}
20 @attribute s43 {3,4,2,5,1}
21 @attribute s44 {3,5,4,1,2}
22 @attribute s45 {3,2,1,4,5}
23 @attribute s46 {3,4,2,1,5}
24 @attribute s47 {3,4,5,2,1}
25 @attribute s48 {3,5,4,1,2}
26 @attribute s49 {3,5,4,2,1}
27 @attribute s50 {3,4,5,1,2}
28 @attribute s51 {3,1,2,5,4}
29 @attribute s15 {basarili,basarisiz}
30
31 @data
32 1sınıf,lise_mezunu,3_3.99,basarili,basarili,basarili,ileri_duzey,cok_sik,gunluk_takip_ederim,2_3_saat,pdf_icerik
33 2sınıf,lise_mezunu,3_3.99,basarili,basarisiz,basarili,ileri_duzey,cok_sik,gunluk_takip_ederim,4_5_saat,pdf_icerik
34 1sınıf,lise_mezunu,3_3.99,basarisiz,basarisiz,basarisiz,ileri_duzey,cok_sik,gunluk_takip_ederim,4_5_saat,pdf_icerik
35 1sınıf,lise_mezunu,3_3.99,basarisiz,basarisiz,basarili,ileri_duzey,cok_sik,hafiraladikca_katilirim,2_3_saat,pdf_icerik
36 2sınıf,lise_mezunu,4_5,basarisiz,basarili,basarisiz,ileri_duzey,cok_sik,haftalik_takip_ederim,4_5_saat,pdf_icerik
    
```

Figure2. Arff data set file

The view of an attribute whose success status had four classes is presented below including its reduction to four classes.

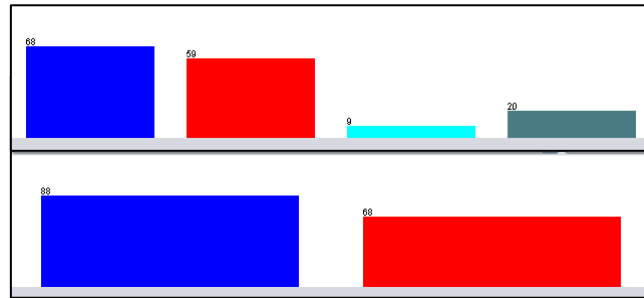


Figure3. Classify value with two classes

Association Rule and Apriori Algorithm

Data mining techniques are generally divided into two models, predictive and descriptive. The predictive model is based on prediction, forecasting and classification steps. The descriptive model reveals the hidden relationships within the data. Association rule is a model that defines certain types of data relationships. In this respect, it is a descriptive model. (Kurt Pehlivanoglu, M., Nevcihan Duru, N.,2015)

Association rules define the dependency rules between two or more attributes from databases where different support and confidence criteria are used. Here, support and confidence are two criteria related to the rule. They reflect the actual usefulness and certainty of the rule that is discovered. (Sebastian, S., Performance, 2016) These rules, which meet the minimum support and confidence thresholds are called strong rules since they are obtained as a result of user-defined values. (Man, M., Abu Bakar, W.A.W., Abdullah, Z., Jalil, Z.A, Herawan, T., 2016)

The Apriori algorithm was developed by Agrawal and Srikant in 1994. This algorithm is the most commonly used and known association rule extraction algorithm. Each rule obtained as a result of the apriori algorithm is expressed in terms of support and confidence. The support criterion refers to the frequency of association between items, while the trust criterion refers to the correctness of these associations. (Yalçın, N., Güngör, E., Yurtay, N.,2013)

The rules have an $A \Rightarrow B$ notation. If A is part of an event, X% of B is part of the event, too. A is called the premise and B is the conclusion. $A \Rightarrow B$ is interpreted as fulfilling the conditions of A and B at the same time. (Gour S., Jaloree, S., Gour, M., 2016)

Support and confidence are calculated as follows.

Where X and Y are different products, support for product X is the ratio of product X in all purchases. $|X|$ shows the number of purchases containing X products, $|D|$ shows the number of all purchases made;

$$\Rightarrow \text{Support}(X) = |X| / |D|$$

Where support for products X and Y is the number of purchases which include products X and Y;

$$\Rightarrow \text{Support}(X \Rightarrow Y) = |X.Y| / |D|$$

And confidence for X and Y is as follows:

$$\Rightarrow \text{Confidence}(X \Rightarrow Y) = \text{support}(X.Y) / \text{support}(X)$$

For example, if a customer who purchases an X product also buys Y product, the expression of this situation with the association rule is $X \Rightarrow Y$ [Support = 30%, confidence = 70%]. Here, support and confidence are the novelty criteria of the rule. They respectively show the usefulness and correctness of the rule discovered.

The large support value in the algorithm reduces the apriori steps and the number of results obtained. The result set obtained in this case is not useful. What is important is that the confidence value is large. Because confidence values indicate the correctness of the rule. (Özseven,T., Düğenci, M., 2011).

Within the scope of the study, the Apriori algorithm settings used on Weka were set through the algorithm GenericObjectEditor dialog window. The adjustments made in this context are given below;

- In the dataset containing success information for more than one course, the CAR line was set to True to create a separate rule for each feature that contained course success, instead of general association rules. This feature is a preferred feature for creating class association rules. Where car = False, the algorithm extracts a general association rule.
- The ClassIndex line is the section in which the classes to take as a basis while creating the association rules are identified. In this section, S13, S14, S15 and S16 attributes were assigned to the Classindex value, and the association rules were created.
- LowerBoundMinSupport is the minimum support value with a default value of 0.1. Within the scope of the study, rules were created with the values 0.2, 0.3, 0.4, 0.5 and 0.6 for the target classes.
- MinMetric is used as the minimum confidence interval. The default value is 0.9. It was used as 0.8 in this study.
- NumRules is the number of rules to be displayed on the screen. The default value is 10. Even if a very large value is written in this section, the generated rules will be displayed according to the determined criteria (for 5, for example, only 5 in the result screen). In this study, a value of 90.000.000 was input in this section since the aim was to identify the number of rules which fitted the criteria.

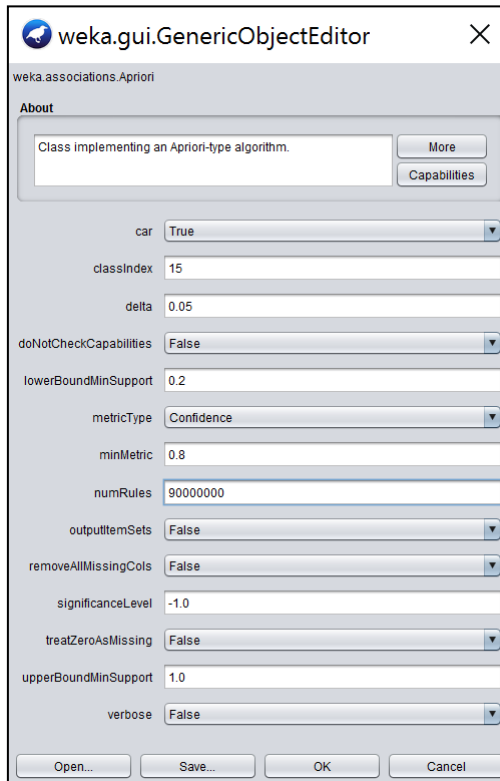


Figure 4. Apriori genericobjecteditör

Results and Findings

The data set with 51 different features of 156 students was analyzed through the Apriori algorithm for the classindex value of S13. 5 rules with the highest number of rules and confidence values obtained when the Apriori algorithm was run with different (0.2, ..0.6) support values and a minimum of 80% confidence are given in the table below.

Table1. Car=True - ClassIndex=13 | average Turkish, history and foreign language scores

Rank	MinMetric	LowerBound MinSupport	Total RulesCount	First Five Rules
1	0.8	0.2	49307	1. s5=cocugum_yok s16=basarili s44=5 37 ==> s13=basarili 33

				conf:(0.89) 2. s8=evet_calisiyorum s14=basarili 41 ==> s13=basarili 36 conf:(0.88) 3. s5=cocugum_yok s15=basarili s16=basarili s33=cok_sik 40 ==> s13=basarili 35 conf:(0.88) 4. s5=cocugum_yok s15=basarili s16=basarili s32=ileri_duzey s33=cok_sik 39 ==> s13=basarili 34 conf:(0.87) 5. s14=basarili s16=basarili 38 ==> s13=basarili 33 conf:(0.87)
2	0.8	0.3	3	1. s5=cocugum_yok s16=basarili 61 ==> s13=basarili 49 conf:(0.8) 2. s16=basarili s44=5 60 ==> s13=basarili 48 conf:(0.8) 3. s7=istanbul s8=evet_calisiyorum s16=basarili 60 ==> s13=basarili 48 conf:(0.8)
3	0.8	0.4	0	-
4	0.8	0.5	0	-
5	0.8	0.6	0	-

Apriori algorithm created a sum of 49307 rules in the experiments with a classindex of S13, minimum support of 0.2 and minimum confidence of 80%. When the first 5 of these rules are investigated:

1. According to the rule, 89% of those who do not have children and are successful in vocational theoretical courses are also successful in Turkish Language, History and English courses.
2. According to the rule, 88% of those who work and are successful in mathematics are also successful in Turkish Language, History and English courses.
3. According to the rule, 88% of those who are successful in the theoretical and practical vocational courses and who use computer-internet very frequently are also successful in Turkish Language, History and English lessons.
4. According to the rule, 87% of those who do not have children, are successful in vocational theoretical and practical courses, use the computer very well and use computers-internet very frequently are also successful in Turkish Language, History and English lessons.
5. According to the rule, 87% of those who are successful in mathematics and theoretical vocational courses are also successful in Turkish Language, History and English.

Apriori algorithm created a sum of 3 rules in the experiments with a classindex of S13, minimum support of 0.3 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 80% of those who do not have children and are successful in the theoretical vocational courses are also successful in Turkish Language, History and English courses.
2. According to the rule, 80% of those who are successful in the theoretical vocational courses and whose response for the item "Exam results are announced timely." is "Yes" are also successful in Turkish Language, History and English courses.
3. According to the rule, 80% of those who live and work in İstanbul and are successful in the theoretical vocational courses are also successful in Turkish Language History and English courses.

For the minimum support values of 0.4, 0.5 and 0.6 in the S13 attribute, no rule was generated.

In Table 2, S14 classindex value was analyzed using Apriori algorithm. In the table, 5 rules with the highest number of rules and confidence values obtained when the Apriori algorithm was run with different (0.2, ..0.6) support values and a minimum confidence value of 80% are given.

Table2. Car=True - ClassIndex=14 | Mathematics Score

Rank	MinMetric	LowerBound MinSupport	Total RulesCount	First Five Rules
1	0.8	0.2		-
2	0.8	0.3	214125	1. s6=lise_mezunu s8=evet_calisiyorum s13=basarisiz 53 ==> s14=basarisiz 52 conf:(0.98) 2. s1=bay s5=cocugum_yok s6=lise_mezunu s7=istanbul s8=evet_calisiyorum 52 ==> s14=basarisiz 49 conf:(0.94) 3. s1=bay s6=lise_mezunu s13=basarisiz 50 ==> s14=basarisiz 47 conf:(0.94) 4. s5=cocugum_yok s6=lise_mezunu s13=basarisiz 50 ==>

				s14=basarisiz 47 conf:(0.94) 5. s1=bay s5=cocugum_yok s6=lise_mezunu s7=istanbul 55 ==> s14=basarisiz 51 conf:(0.93)
3	0.8	0.4	2082	1. s6=lise_mezunu s7=istanbul s8=evet_calisiyorum 77 ==> s14=basarisiz 66 conf:(0.86) 2. s1=bay s6=lise_mezunu s8=evet_calisiyorum 89 ==> s14=basarisiz 76 conf:(0.85) 3. s6=lise_mezunu s24=* 77 ==> s14=basarisiz 65 conf:(0.84) 4. s6=lise_mezunu s29=* 77 ==> s14=basarisiz 65 conf:(0.84) 5. s6=lise_mezunu s31=* 77 ==> s14=basarisiz 65 conf:(0.84)
4	0.8	0.5	2	1. s6=lise_mezunu s8=evet_calisiyorum 103 ==> s14=basarisiz 85 conf:(0.83) 2. s1=bay s6=lise_mezunu 96 ==> s14=basarisiz 79 conf:(0.82)
5	0.8	0.6	0	-

Apriori algorithm created a sum of 214125 rules in the experiments with a classindex of S14, minimum support of 0.3 and minimum confidence of 80%. When the first 5 of these rules are investigated:

1. According to the rule, 98% of those who hold a high school diploma, work and are unsuccessful in Turkish Language, History and English courses are also unsuccessful in Mathematics.
2. According to the rule, 94% of those who are male without children, hold a high school diploma, live and work in İstanbul are unsuccessful.
3. According to the rule, 94% of those who are male, unsuccessful in Turkish Language, History and English courses and hold a high school diploma are unsuccessful.
4. According to the rule, 94% of those who do not have children, hold a high school diploma and are unsuccessful in Turkish Language, History and English courses are also unsuccessful in Mathematics.
5. According to the rule, 93% of those who are male without children, hold a high school diploma and live in istanbul are also unsuccessful in Mathematics.

Apriori algorithm created a sum of 2082 rules in the experiments with a classindex of S14, minimum support of 0.4 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 86% of those who hold a high school diploma, live and work in İstanbul are unsuccessful in Mathematics.
2. According to the rule, 85% of those who are male, hold a high school diploma and work are unsuccessful in Mathematics.
3. According to the rule, 84% of those who hold a high school diploma are unsuccessful in Mathematics.

Apriori algorithm created a sum of 2 rules in the experiments with a classindex of S14, minimum support of 0.5 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 85% of those who hold a high school diploma and work are unsuccessful in Mathematics.
2. According to the rule, 82% of those who are male and hold a high school diploma are unsuccessful in Mathematics.

For the minimum support values of 0.2 and 0.6 in the S14 attribute, no rule was generated.

In Table 3, S15 classindex value was analyzed using Apriori algorithm. In the table, 5 rules with the highest number of rules and confidence values obtained when the Apriori algorithm was run with different (0.2, ..0.6) support values and a minimum confidence value of 80% are given in rank order.

Table3. Car=True - ClassIndex=15 | Average success in practical vocational courses

Rank	MinMetric	LowerBound MinSupport	Total RulesCount	First Five Rules
1	0.8	0.2	107896	1. s13=basarili s16=basarili s48=5 49 ==> s15=basarili 43 conf:(0.88) 2. s13=basarili s16=basarili s32=ileri_duzyey s33=cok_sik 52 ==> s15=basarili 45 conf:(0.87) 3. s7=istanbul s13=basarili s16=basarili 49 ==> s15=basarili 42

				conf:(0.86) 4. s13=basarili s16=basarili s32=ileri_duzey 55 ==> s15=basarili 47 conf:(0.85) 5. s13=basarili s16=basarili s33=cok_sik 55 ==> s15=basarili 47 conf:(0.85)
2	0.8	0.3	33	1. s13=basarili s16=basarili s32=ileri_duzey 55 ==> s15=basarili 47 conf:(0.85) 2. s13=basarili s16=basarili s33=cok_sik 55 ==> s15=basarili 47 conf:(0.85) 3. s16=basarili s33=cok_sik s48=5 55 ==> s15=basarili 47 conf:(0.85) 4. s1=bay s16=basarili s48=5 61 ==> s15=basarili 52 conf:(0.85) 5. s16=basarili s48=5 67 ==> s15=basarili 57 conf:(0.85)
3	0.8	0.4	3	1. s16=basarili 88 ==> s15=basarili 72 conf:(0.82) 2. s1=bay s16=basarili 81 ==> s15=basarili 66 conf:(0.81) 3. s8=evet_calisiyorum s16=basarili 80 ==> s15=basarili 64 conf:(0.8)
4	0.8	0.5	0	-
5	0.8	0.6	0	-

Apriori algorithm created a sum of 107896 rules in the experiments with a classindex of S15, minimum support of 0.2 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 88% of those who are successful in Turkish Language, History and English courses along with the theoretical vocational courses and who also know how to use the LMS very well are also successful in the the practical vocational courses.
2. According to the rule, 87% of those who are successful in Turkish Language, History and English courses along with the theoretical vocational courses, who can use computers in an advanced level and and who use computers and internet very frequently are also successful in the practical vocational courses.
3. According to the rule, 86% of those who live in İstanbul, are successful in Turkish Language, History and English courses along with the theoretical vocational courses are also successful in the practical vocational courses.
4. According to the rule, 85% of those who are successful in Turkish Language, History and English courses along with the theoretical vocational courses and who can use computers in an advanced level are also successful in the practical vocational courses.
5. According to the rule, 85% of those who are successful in Turkish Language, History and English courses along with the theoretical vocational courses and who use computers and internet very frequently are also successful in the practical vocational courses.

Apriori algorithm created a sum of 33 rules in the experiments with a classindex of S15, minimum support of 0.3 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 85% of those who are successful in Turkish Language, History and English courses along with the theoretical vocational courses and who can use computers in an advanced level are also successful in the practical vocational courses.
2. According to the rule, 85% of those who are successful in Turkish Language, History and English courses along with the theoretical vocational courses and who use computers and internet very frequently are also successful in the practical vocational courses.
3. According to the rule, 85% of those who are successful in the theoretical vocational courses, use computers and internet very frequently and the LMS very well are also successful in the practical vocational courses.
4. According to the rule, 85% of those who are male, successful in the theoretical vocational courses and can use the LMS very well are also successful in the practical vocational courses.
5. According to the rule, 85% of those who are successful in the theoretical vocational courses and can use the LMS very well are also successful in the practical vocational courses.

Apriori algorithm created a sum of 3 rules in the experiments with a classindex of S15, minimum support of 0.4 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 82% of those who are successful in the theoretical vocational courses are also successful in the practical vocational courses.
2. According to the rule, 81% of those who are male and successful in the theoretical vocational courses are also successful in the practical vocational courses.
3. According to the rule, 80% of those who work and are successful in the theoretical vocational courses are also successful in the practical vocational courses.

For the minimum support values of 0.5 and 0.6 in the S15 attribute, no rule was generated.

In Table 4, S16 classindex value was analyzed using Apriori algorithm. In the table, 5 rules with the highest number of rules and confidence values obtained when the Apriori algorithm was run with different (0.2, ..0.6) support values and a minimum confidence value of 80% are given in rank order.

Table4. Car=True - ClassIndex=16 | Average success in theoretical vocational courses

Sıra	MinMetric	LowerBound MinSupport	Total RulesCount	First Five Rules
1	0.8	0.2	116901	1. s7=istanbul s8=evet_calisiyorum s13=basarili s15=basarili s32=ileri_duzey 36 ==> s16=basarili 35 conf:(0.97) 2. s7=istanbul s8=evet_calisiyorum s13=basarili s15=basarili s32=ileri_duzey s33=cok_sik 36 ==> s16=basarili 35 conf:(0.97) 3. s1=bay s12=3_3.99 s15=basarili s33=cok_sik 33 ==> s16=basarili 32 conf:(0.97) 4. s1=bay s7=istanbul s13=basarili s15=basarili s32=ileri_duzey 33 ==> s16=basarili 32 conf:(0.97) 5. s1=bay s12=3_3.99 s15=basarili s32=ileri_duzey s33=cok_sik 33 ==> s16=basarili 32 conf:(0.97)
2	0.8	0.3	47	1. s1=bay s15=basarili s32=ileri_duzey 64 ==> s16=basarili 58 conf:(0.91) 2. s1=bay s15=basarili s33=cok_sik 62 ==> s16=basarili 56 conf:(0.9) 3. s13=basarili s32=ileri_duzey 61 ==> s16=basarili 55 conf:(0.9) 4. s15=basarili s32=ileri_duzey s33=cok_sik 65 ==> s16=basarili 58 conf:(0.89) 5. s13=basarili s15=basarili 63 ==> s16=basarili 56 conf:(0.89)
3	0.8	0.4	3	1. s1=bay s15=basarili 75 ==> s16=basarili 66 conf:(0.88) 2. s8=evet_calisiyorum s15=basarili 76 ==> s16=basarili 64 conf:(0.84) 3. s15=basarili 87 ==> s16=basarili 72 conf:(0.83)
4	0.8	0.5	0	-
5	0.8	0.6	0	-

Apriori algorithm created a sum of 116901 rules in the experiments with a classindex of S16, minimum support of 0.2 and minimum confidence of 80%. When these rules are investigated:

1. According to the rule, 97% of those who live and work in İstanbul, are successful in Turkish Language, History and English courses along with the practical vocational courses and who use computers in an advanced level are successful in the theoretical vocational courses.
2. According to the rule, 97% of those who live and work in İstanbul, are successful in Turkish Language, History and English courses along with the practical vocational courses and who use computers in an advanced level and who use computers and internet very frequently are successful in the theoretical vocational courses.
3. According to the rule, 97% of those who are male, have a high school diploma grade between 3 and 3.99, are successful in the practical vocational courses and use computers and internet very frequently are also successful in the theoretical vocational courses.
4. According to the rule, 97% of those who are male, are successful in the Turkish Language, History and English courses and use computers in an advanced level are also successful in the theoretical vocational courses.
5. According to the rule, 97% of those who are male, have a high school diploma grade between 3 and 3.00, who use computers in an advanced level and very frequently are also successful in the theoretical vocational courses.

Apriori algorithm created a sum of 47 rules in the experiments with a classindex of S16, minimum support of 0.3 and minimum confidence of 80%. When the first 5 of these rules are investigated:

1. According to the rule, 91% of those who are male, are successful in the practical vocational courses and use computers in an advanced level are also successful in the theoretical vocational courses.
2. According to the rule, 90% of those who are male, are successful in the practical vocational courses and use computers and internet very frequently are also successful in the theoretical vocational courses.
3. According to the rule, 90% of those who are successful in Turkish Language, History, English courses and use computers in an advanced level are also successful in the theoretical vocational courses.
4. According to the rule, 89% of those who are successful in the practical vocational courses and use computers and internet very frequently are also successful in the theoretical vocational courses.
5. According to the rule, 89% of those who are successful in Turkish Language, History, English courses along with the practical vocational courses are also successful in the theoretical vocational courses.

Apriori algorithm created a sum of 3 rules in the experiments with a classindex of S16, minimum support of 0.4 and minimum confidence of 80%. When the first 5 of these rules are investigated:

1. According to the rule, 88% of those who are male and successful in the practical vocational courses are also successful in the theoretical vocational courses.
2. According to the rule, 84% of those who work and are successful in the practical vocational courses are also successful in the theoretical vocational courses.
3. According to the rule, 83% of those who are successful in the practical vocational courses are also successful in the theoretical vocational courses.

For the minimum support values of 0.5 and 0.6 in the S16 attribute, no rule was generated.

Conclusion

The present study was conducted with the participation of 156 volunteer students enrolled in Trakya University Tunca Vocational School (Distance Education) Computer Programming Program in 2016-2017 academic year. The data set, which was created with questionnaire data, was analyzed by means of the Apriori algorithm, which is included in WEKA as an association rule extraction algorithm that is widely known and used.

In the Apriori experiments made within the context of the study, association rules are generated for each of the S13 (Average Grade for Turkish Language, History, English Courses), S14 (Mathematics Course Grade), S15 (Average Grade for Practical Vocational Courses) and S16 (Average Grade for Theoretical Vocational Courses) attributes instead of generic association rules. In this regard, the following generalization is reached for the rules focusing on the S13, S15 and S16 attributes in the data set: The participants who spend time with computers and internet, have advanced computer skills, can use very well th LMS where course content is published and mid-term exams are done have a high rate of success in the courses coded as S13, S15 and S16.

In addition, a strong link can also be seen between the success status of vocational S15 (practical) and S16 (theoretical) courses. According to the rule with the highest confidence level obtained in the classindex S13, S15 and S16 experiments, 97% of the participants who live and work in İstanbul, are successful in Turkish Language, History and English courses along with the practical vocational courses and can use computers in an advanced level are also successful in the theoretical vocational courses.

When the attributes S13, S14, S15 and S16 are examined in terms of class distribution, it is seen that the class distributions are balanced except for S14 (Mathematics course). In this respect, association rules are based on failure taking into account the fact that the class distribution in the S14 attribute indicate that 68% of the participants are unsuccessful in this particular attribute while 32% are successful. An investigation taking into account the instance distribution in the classes which the attribute has can be proposed as a further study which would necessitate the use of the Apriori Association Algorithm. According to the rule with the highest level of confidence in the classindex S14 experiments, 98% of the participants who hold a high school diploma, work and are unsuccessful in Turkish Language, History and English courses are also unsuccessful in the Mathematics course.

References

- Bozkır, A.S., & Gök, B., & Sezer, E. (2008) Üniversite Öğrencilerinin İnterneti Eğitimsel Amaçlar İçin Kullanmalarını Etkileyen Faktörlerin Veri Madenciliği Yöntemleriyle Tespiti. *Bilimde Modern Yöntemler Sempozyumu*, Eskişehir.
- Chan, A.Y.K., & Chow, K.O., & Cheung, K.S. (2008). Online Course Refinement through Association Rule Mining. *Journal of Educational Technology Systems*, ISSN 1450-216X 36(4), (pp. 43–44).
- Chandra,E., & Nandhini, K. (2010). Knowledge Mining from Student Data. *European Journal of Scientific Research*, 47(1), (pp. 156-163).
- Gour S., & Jaloree, S., & Gour, M. (2016). Water Quality Assessment using Association Rule Mining for River Narmada. *Indian Journal of Science and Technology*, 9(10), DOI: 10.17485/ijst/2016/v9i10/82155.
- Karabatak, M., & İnce, M.C., (2004). Apriori Algoritması İle Öğrenci Başarısı Analizi. *Elektrik Elektronik Bilgisayar Mühendisliği Sempozyumu*, Bursa.
- Kurt Pehlivanoğlu, M., & Duru, N., (2015). Veri Madenciliği Teknikleri Kullanılarak Ortaokul Öğrencilerinin Sosyal Ağ Kullanım Analizi: Kocaeli İli Örneği, *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 3, (pp. 508-517).
- Man, M.B., & Abu Bakar, W.A.W., & Abdullah, Z., & Jalil, Z.A, &Herawan, T., (2016). Mining Association Rules: A Case Study on Benchmark Dense Data. *Indonesian Journal of Electrical Engineering and Computer Science*, 3(3), (pp. 546-553), DOI: 10.11591/ijeecs.v3.i2.pp546-553.
- Özseven,T., & Düğenci, M., (2011). LOG Analiz: Erişim Kayıt Dosyaları Analiz Yazılımı ve GOP Üniversitesi Uygulaması. *Bilişim Teknolojileri Dergisi*, 4(2), (pp. 55-66).
- Pandeeswari, L., & Rajeswari, K., (2014). Student Academic Performance Using Data Mining Techniques. *International Journal of Computer Science and Mobile Computing, IJCSMC*, 3(10), (pp. 726–731).
- Sebastian, S., (2016). Performance, Evaluation By Artificial Neural Network Using WEKA. *International Research Journal of Engineering and Technology (IRJET)*, 3(3), e-ISSN: 2395 -0056 (pp. 1459-1464).
- Tanrıverdi, N.S., (2013). *İTÜ Matematik Mühendisliği Öğrencilerinin Mezuniyet Başarılarını Yükseltmeyi Hedefleyen Bir Veri Madenciliği Uygulaması*. (Unpublished Dissertation), İstanbul Üniversitesi Fen Edebiyat Fakültesi Matematik Mühendisliği Programı, İstanbul-Turkey.
- Taş, M., & Adak, M.F., & Yurtay, N., (2013). Öğrencilerin Staj Verileri Üzerine Uygulanan Apriori Algoritması ile Birliktelik Kurallarının Çıkarılması ve Staj Eğiliminin Belirlenmesi. *1st International Symposium on Innivative Technologies in Engineering and Science*, (pp. 1086-1095).
- Yalçın, N., & Güngör, E., & Yurtay N., (2013). Apriori Algoritması ile Teknik Seçmeli Ders Seçim Analizi, *Ulusal Uzaktan Eğitim ve Teknolojileri Sempozyumu*, Konya.