Available online at www.sciencedirect.com

ScienceDirect

Procedia Social and Behavioral Sciences 2 (2010) 5251–5259

Procedia
Social and Behavioral Sciences

WCES-2010

Data mining application on students' data

Ali Buldu[a] *, Kerem Üçgün[a]

[a]*Technical Education of Marmara University, Istanbul, 34722, Turkey*

Received November 15, 2009; revised December 3, 2009; accepted January 25, 2010

Abstract

In today's world, due to the rapid development of technology, the amount of data stored has been constantly increasing in every field. It is intended to obtain meaningful, valuable information that is not previously known from these data by applying data mining techniques.

In data mining techniques, association rules are one of the most preferred techniques. Apriori algorithm is the most used one in these association rules. In this work, by being carried out apriori algorithm upon the data of students of Istanbul Eyup I.M.K.B. Vocational Commerce High School, the rules have been produced and from the results obtained the relation between the courses that the students failed have been revealed.

Keywords: Data mining; education; association-relation rule; apriori algorithm; data reducing.

1. Introduction

The data is for their sizes covers much space in pages but the merit of their usage is little. However, if we sum it up by putting numbers in an order, if we convert into meaningful sentences by arranging letters, and if we produce a melody by putting notes in a raw and if we produce a graphic or a picture of a tree by combining data on computer screen it is only at this point that we convert these data into information. Info covers less space as for its size in contrast to data but is more powerful in terms of usage worth (Gursakal, 2001).

Because of this reason, data mining comes in the first line in the process of information exploration on databases. From these data in hand, data mining is extracting potentially handy information which is not so clear, unknown before and up closed (Alatas, Akin, 2004). Data mining is not a solution itself at this point, instead it is a tool which supports decision making process and which tries to ensure required knowledge in order to reach the solution of the problem.

* Ali Buldu. Tel.: +90-216-336-5770; fax: +90-216-337-8987
E-mail address: alibuldu@marmara.edu.tr

1877-0428 © 2010 Published by Elsevier Ltd.
doi:10.1016/j.sbspro.2010.03.855

1.1. The procedure of application

In this study, in order to realize data mining with apriori algorithm, the students' data has been entered into the school automation program which has been developed by us. The frequent items sets are fixed by regarding all the points that students have got. Then, the rules are revealed which can be obtained from databases.

As a data set which apriori algorithm will be applied, the data of the students of Eyup İ.M.K.B. Vocational Commerce High School 11-A class is used. There are 28 students in 11-A class. When the lessons are dealt separately as a first and second term, apriori algorithm is applied on the points that the students got from 74 courses in total. Which means 30 courses in 9th, 26 courses in 10th and 18 courses in 11th classes. As minimum support rate 9 and as minimum confidence rate %85 is entered. Each step of algorithm can be seen below.

1st Step: In the first step of algorithm, algorithm scans the table of course lists in figure 1 and the information about how many operations each objects have passed through is obtained. These obtained data represents C_1 candidate item set of apriori algorithm.

SCHOOL NUMBER	THE COURSES HE/SHE FAILED
413	9AFİZ - 9AKİM - 9AMAT - 9AMATBHR - 10AMAT
414	9AFİZ - 9AKİM - 9AMATBHR - 9AAUB - 10AMAT
420	9AMAT - 10AMAT
423	9AMAT - 9AFİZ
436	9ABİT- 9ADİH - 9AFİZ - 9AKİM - 9AAUB - 9AİNGBHR - 10AMAT - 10AMATBHR - 10ATARBHR - 10ATEÖ - 10ATEÖBHR - 10ABTT - 11AİTA
441	9AMAT - 9ATEDBHR - 9AAUB - 10AMATBHR - 10ATEÖ - 10ATEÖBHR - 10ABTTBHR - 11AGÖPBHR
443	9ADVABHR - 9AFİZ - 9AKİM - 9AMATBHR - 10AMAT - 10ATEÖ
464	9AKİM - 9AAUB - 10AMAT
468	9AKİM
469	9ACOGBHR - 9ADVABHR - 9AKİM - 9AKİMBHR - 9AMATBHR - 9AAUB - 10AMAT - 10ATAR - 10ATEÖ
471	9AFİZ - 9AKİM - 10AMAT
472	9AFİZ - 9AMAT - 9ATARBHR - 9ATED - 9AAUB - 9İNGBHR - 9ABİT - 9ABİTBHR - 9ADVA - 10ACOG - 10ADVA - 10AMAT - 10AMES - 10APPR - 10ATARBHR - 10ATED - 10ATEÖ - 11AFEL - 11AGÖPBHR - 11AİTA
480	9AFİZ - 9AKİM - 9AMAT
491	11AGÖPBHR
494	9ABİT - 9AFİZ - 9AKİM - 9AAUB - 9AİNGBHR - 10AMAT - 10ATEÖBHR - 11AGÖPBHR - 11AİTA
498	9ABYLBHR - 9AFİZ - 9AKİM - 9AAUB - 9AİNG - 10AMAT - 10AMATBHR - 10ATEÖ - 10ATEÖBHR - 10ABTT
500	10AMAT - 10AİNGBHR
506	9ADVA - 9AFİZ - 9AKİM - 9AMAT - 9AMATBHR - 9ATED - 9AAUB - 9AİNG - 9AİNGBHR - 10AMAT - 10AMATBHR - 10ATEÖ - 10ATEÖBHR - 10ABTT - 11AGÖPBHR
516	9ABİT - 9AFİZ - 9AMAT - 9AAUB - 10AMAT - 10ATED - 10ATEÖ - 11AİTA
520	9ABİT - 9ACOG - 9ACOGBHR - 9ADVA - 9AFİZ - 9AAUB - 10AMATBHR - 10APPRBHR - 10ATARBHR - 10ATED - 10ATEÖ - 11AİTA
522	9ABİT - 10ABTT - 10ABTTBHR
529	9ADVABHR - 9AFİZ - 9AMAT - 10AMAT - 10ATAR - 10ATED - 10ATEÖ - 10ATEÖBHR - 11AFEL - 11AFELBHR - 11AGÖPBHR
539	9ABİT - 9AFİZ - 9AKİM - 9AMAT - 9AMATBHR - 9AAUB - 9AAUBBHR - 10AMAT - 10ATED - 10ATEÖ - 10ATEÖBHR - 10ABTT - 11AGÖPBHR - 11AİTA
541	9ABYLBHR - 9ACOG - 9ADVA - 9ADVABHR - 9ADİH - 10ADVA - 9AFİZ - 9AKİM - 9AMAT - 9AMATBHR - 9ATEDBHR - 9AAUB - 9AİNGBHR - 10AMAT - 10APPR - 10ATAR - 10ATEÖ - 10AİNG - 11AİTA
553	9ABİT - 9ABİTBHR - 9AAUB - 10AMAT - 10ATEÖ

Figure 1. Course list table

2nd Step: In the second step, with the help of minimum support which is determined at previous L_1 item set is obtained which includes the most frequent single items. That is to say, they form L_1 set equal to the minimum support rate or having higher support rate.

C_1 candidate item set

COURSE CODE	COURSE NAME	SUPPORT
10AMAT	MATH	18
9AFİZ	PHYSICS	16
9AAUB	ASTRONOMY AND SPACE SCIENCES	14
9AKİM	CHEMISTRY	14
10ATEÖ	BASIC ELECTRONICS AND MEASUREMENT	13
9AMAT	MATH	11
9ABİT	INFORMATION AND COMMUNICATION TECHNOLOGIES	8
10ATEÖBHR	BASIC ELECTRONICS AND MEASUREMENT	7
11AGÖPBHR	VISUAL PROGRAMMING	7
9AMATBHR	MATH	7
11AİTA	HISTORY	7
10AMATBHR	MATH	5
9AİNGBHR	ENGLISH	5
10ABTT	INFORMATION TECHNOLOGY	5
10ATED	TURKISH LITERATURE	5
9ADVA	LANGUAGE AND EXPRESSION	4
9ADVABHR	LANGUAGE AND EXPRESSION	4
10ATARBHR	HISTORY	3
10ATAR	HISTORY	3
10APPR	PACKAGE PROGRAMS	2
9ATEDBHR	TURKISH LITERATURE	2
11AFEL	PHILOSOPHY	2
10ABTTBHR	INFORMATION TECHNOLOGY	2
9ABYLBHR	BIOLOGY	2
9ABİBHR	INFORMATION AND COMMUNICATION TECHNOLOGIES	2
9AİG	ENGLISH	2
9ADİH	DEMOCRACY AND HUMAN RIGHTS	2
9ACOG	GEOGRAPHY	2
10ADVA	LANGUAGE AND EXPRESSION	2
9ATED	TURKISH LITERATURE	2
9ACOGBHR	GEOGRAPHY	2
10AMES	PROFESSIONAL DEVELOPMENT	1
10İNGBHR	ENGLISH	1
9AAUBBHR	ASTRONOMY AND SPACE SCIENCES	1
10AİNG	ENGLISH	1
10APPRBHR	PACKAGE PROGRAMS	1
9ATARBHR	HISTORY	1
11AFELBHR	PHILOSOPHY	1
10ACOG	GEOGRAPHY	1
9AKİMBHR	CHEMISTRY	1

=> Minimum support (9)

L_1 frequent item set

COURSE CODE	COURSE NAME	SUPPORT
10AMAT	MATH	18
9AFİZ	PHYSICS	16
9AAUB	ASTRONOMY AND SPACE SCIENCES	14
9AKİM	CHEMISTRY	14
10ATEÖ	BASIC ELECTRONICS AND MEASUREMENT	13
9AMAT	MATH	11

Figure 2. Apriori algorithm $C_1 \propto L_1$ transformation

3rd Step: In this step, in order to determine L_2 frequent item set L_1 as a new candidate item set is formed from the combination of $L_1 \propto L_1$. This process is called as joining. The forming of C_2 candidate item set can be seen figure 3 with sample data.

L_1 frequent item set

COURSE CODE	COURSE NAME	SUPPORT
10AMAT	MATH	18
9AFİZ	PHYSICS	16
9AAUB	ASTRONOMY AND SPACE SCIENCES	14
9AKİM	CHEMISTRY	14
10ATEÖ	BASIC ELECTRONICS AND MEASUREMENT	13
9AMAT	MATH	11

$L_1 \infty L_1$

Join process

C_2 candidate item set

L1 X L1	L1_1	L1_2
10AMAT X 10ATEÖ	10AMAT	10ATEÖ
10AMAT X 9AAUB	10AMAT	9AAUB
10AMAT X 9AMAT	10AMAT	9AMAT
10AMAT X 9AKİM	10AMAT	9AKİM
10AMAT X 9AFİZ	10AMAT	9AFİZ
10ATEÖ X 9AAUB	10ATEÖ	9AAUB
10ATEÖ X 9AMAT	10ATEÖ	9AMAT
10ATEÖ X 9AKİM	10ATEÖ	9AKİM
10ATEÖ X 9AFİZ	10ATEÖ	9AFİZ
9AAUB X 9AMAT	9AAUB	9AMAT
9AAUB X 9AKİM	9AAUB	9AKİM
9AAUB X 9AFİZ	9AAUB	9AFİZ
9AFİZ X 9AMAT	9AFİZ	9AMAT
9AFİZ X 9AKİM	9AFİZ	9AKİM
9AKİM X 9AMAT	9AKİM	9AMAT

Figure 3. C_2 candidate item set formed by $L_1 \infty L_1$ transformation

4th Step: According to the summary code of C_2 pruning procedure is necessary to be done after candidate set is formed. In pruning, the existence of subsets of C_2 candidate item sets in L_1 set is controlled and if a sub set of any item doesn't take place in L_1, the related item is excluded from evaluation and it is deleted from C_2 candidate item set. When this procedure is necessary to be done in C_2 and later candidate sets, this procedure isn't done C_2. the reason of this is that the subsets of all the items C_2 consists of single item and they have come to C_2 candidate item set after being exposed to joining from L_1 set. This means that all the sub sets of C_2 take place L_1 set. Thus, in this step C_2 candidate item set is looked and L_2 candidate item set is formed from the rates which are equal to minimum support rate or high rate. The forming of L_2 frequent is indicated in Figure 4 with sample data.

C_2 candidate item set

C2 COURSE CODE	SUPPORT
10AMAT X 10ATEÖ	11
10AMAT X 9AAUB	12
10AMAT X 9AFİZ	13
10AMAT X 9AKİM	12
10AMAT X 9AMAT	8
10ATEÖ X 9AAUB	11
10ATEÖ X 9AFİZ	10
10ATEÖ X 9AKİM	7
10ATEÖ X 9AMAT	7
9AAUB X 9AFİZ	10
9AAUB X 9AKİM	9
9AAUB X 9AMAT	6
9AFİZ X 9AKİM	11
9AFİZ X 9AMAT	9
9AKİM X 9AMAT	5

=> Minimum support (9)

L_2 frequent item set

L2 COURSE CODE	SUPPORT
10AMAT X 10ATEÖ	11
10AMAT X 9AAUB	12
10AMAT X 9AFİZ	13
10AMAT X 9AKİM	12
10ATEÖ X 9AAUB	11
10ATEÖ X 9AFİZ	10
9AAUB X 9AFİZ	10
9AAUB X 9AKİM	9
9AFİZ X 9AKİM	11
9AFİZ X 9AMAT	9

Figure 4. L_2 frequent item set formed by $C_2 \infty L_2$ transformation.

5th Step: In 5th step of application, in order determine L_3 frequent item set, a new item set C_3 is formed from the dual combination of L_2 set items. C_3 candidate item set's forming is indicated in Figure 5 with sample data.

L_2 frequent item set

L2 COURSE CODE	SUPPORT
10AMAT X 10ATEÖ	11
10AMAT X 9AAUB	12
10AMAT X 9AFİZ	13
10AMAT X 9AKİM	12
10ATEÖ X 9AAUB	11
10ATEÖ X 9AFİZ	10
9AAUB X 9AFİZ	10
9AAUB X 9AKİM	9
9AFİZ X 9AKİM	11
9AFİZ X 9AMAT	9

$L_2 \infty L_2$ Join

C_3 candidate item set

L2 X L2	L2_1	L2_2	L2_3
10AMAT X 10ATEÖ X 9AAUB	10AMAT	10ATEÖ	9AAUB
10AMAT X 9AFİZ X 9AMAT	10AMAT	9AFİZ	9AMAT
10AMAT X 9AKİM X 9AMAT	10AMAT	9AKİM	9AMAT
10AMAT X 9AAUB X 9AMAT	10AMAT	9AAUB	9AMAT
10AMAT X 10ATEÖ X 9AMAT	10AMAT	10ATEÖ	9AMAT
10AMAT X 9AFİZ X 9AKİM	10AMAT	9AFİZ	9AKİM
10AMAT X 9AAUB X 9AKİM	10AMAT	9AAUB	9AKİM
10AMAT X 10ATEÖ X 9AKİM	10AMAT	10ATEÖ	9AKİM
10AMAT X 9AAUB X 9AFİZ	10AMAT	9AAUB	9AFİZ
10AMAT X 10ATEÖ X 9AFİZ	10AMAT	10ATEÖ	9AFİZ
10ATEÖ X 9AFİZ X 9AMAT	10ATEÖ	9AFİZ	9AMAT
10ATEÖ X 9AKİM X 9AMAT	10ATEÖ	9AKİM	9AMAT
10ATEÖ X 9AAUB X 9AMAT	10ATEÖ	9AAUB	9AMAT
10ATEÖ X 9AFİZ X 9AKİM	10ATEÖ	9AFİZ	9AKİM
10ATEÖ X 9AAUB X 9AKİM	10ATEÖ	9AAUB	9AKİM
10ATEÖ X 9AAUB X 9AFİZ	10ATEÖ	9AAUB	9AFİZ
9AAUB X 9AFİZ X 9AMAT	9AAUB	9AFİZ	9AMAT
9AAUB X 9AKİM X 9AMAT	9AAUB	9AKİM	9AMAT
9AAUB X 9AFİZ X 9AKİM	9AAUB	9AFİZ	9AKİM
9AFİZ X 9AKİM X 9AMAT	9AFİZ	9AKİM	9AMAT

Figure 5. C_3 candidate item set formed by $L_2 \infty L_2$

6th Step: According to the summary code of apriori algorithm, pruning procedure is required after C_3 candidate item set is formed. In pruning, whether the subsets of C_3 candidate item set exists in L_2 set is controlled. If a subset of any item doesn't take place in L_2 set, the related item is excluded from the evaluation process and the algorithm can be ended. At the end of the pruning C_3 candidate set when the algorithm is ended, the items L_{k-1} frequent item sets are used in order to form association rules.

C_3 candidate item set (before pruning)

L2 X L2	L2_1	L2_2	L2_3
10AMAT X 10ATEÖ X 9AAUB	10AMAT	10ATEÖ	9AAUB
~~10AMAT X 9AFİZ X 9AMAT~~	10AMAT	9AFİZ	9AMAT
~~10AMAT X 9AKİM X 9AMAT~~	10AMAT	9AKİM	9AMAT
~~10AMAT X 9AAUB X 9AMAT~~	10AMAT	9AAUB	9AMAT
~~10AMAT X 10ATEÖ X 9AMAT~~	10AMAT	10ATEÖ	9AMAT
10AMAT X 9AFİZ X 9AKİM	10AMAT	9AFİZ	9AKİM
10AMAT X 9AAUB X 9AKİM	10AMAT	9AAUB	9AKİM
~~10AMAT X 10ATEÖ X 9AKİM~~	10AMAT	10ATEÖ	9AKİM
10AMAT X 9AAUB X 9AFİZ	10AMAT	9AAUB	9AFİZ
10AMAT X 10ATEÖ X 9AFİZ	10AMAT	10ATEÖ	9AFİZ
~~10ATEÖ X 9AFİZ X 9AMAT~~	10ATEÖ	9AFİZ	9AMAT
~~10ATEÖ X 9AKİM X 9AMAT~~	10ATEÖ	9AKİM	9AMAT
~~10ATEÖ X 9AAUB X 9AMAT~~	10ATEÖ	9AAUB	9AMAT
~~10ATEÖ X 9AFİZ X 9AKİM~~	10ATEÖ	9AFİZ	9AKİM
~~10ATEÖ X 9AAUB X 9AKİM~~	10ATEÖ	9AAUB	9AKİM
10ATEÖ X 9AAUB X 9AFİZ	10ATEÖ	9AAUB	9AFİZ
~~9AAUB X 9AFİZ X 9AMAT~~	9AAUB	9AFİZ	9AMAT
~~9AAUB X 9AKİM X 9AMAT~~	9AAUB	9AKİM	9AMAT
9AAUB X 9AFİZ X 9AKİM	9AAUB	9AFİZ	9AKİM
~~9AFİZ X 9AKİM X 9AMAT~~	9AFİZ	9AKİM	9AMAT

Forming subset k-1

Dual subsets of C_3 candidate item set

COURSE CODE
(10AMAT X 10ATEÖ) (10AMAT X 9AAUB) (10ATEÖ X 9AAUB)
(10AMAT X 9AFİZ) (~~10AMAT X 9AMAT~~) (9AFİZ X 9AMAT)
(10AMAT X 9AKİM) (~~10AMAT X 9AMAT~~) (~~9AKİM X 9AMAT~~)
(10AMAT X 9AAUB) (~~10AMAT X 9AMAT~~) (~~9AAUB X 9AMAT~~)
(10AMAT X 10ATEÖ) (~~10AMAT X 9AMAT~~) (~~10ATEÖ X 9AMAT~~)
(10AMAT X 9AFİZ) (10AMAT X 9AKİM) (9AFİZ X 9AKİM)
(10AMAT X 9AAUB) (10AMAT X 9AKİM) (9AAUB X 9AKİM)
(10AMAT X 10ATEÖ) (10AMAT X 9AKİM) (~~10ATEÖ X 9AKİM~~)
(10AMAT X 9AAUB) (10AMAT X 9AFİZ) (9AAUB X 9AFİZ)
(10AMAT X 10ATEÖ) (10AMAT X 9AFİZ) (10ATEÖ X 9AFİZ)
(10ATEÖ X 9AFİZ) (~~10ATEÖ X 9AMAT~~) (9AFİZ X 9AMAT)
(~~10ATEÖ X 9AKİM~~) (~~10ATEÖ X 9AMAT~~) (9AKİM X 9AMAT)
(10ATEÖ X 9AAUB) (~~10ATEÖ X 9AMAT~~) (~~9AAUB X 9AMAT~~)
(10ATEÖ X 9AFİZ) (~~10ATEÖ X 9AKİM~~) (9AFİZ X 9AKİM)
(10ATEÖ X 9AAUB) (~~10ATEÖ X 9AKİM~~) (9AAUB X 9AKİM)
(10ATEÖ X 9AAUB) (10ATEÖ X 9AFİZ) (9AAUB X 9AFİZ)
(9AAUB X 9AFİZ) (~~9AAUB X 9AMAT~~) (9AFİZ X 9AMAT)
(9AAUB X 9AKİM) (~~9AAUB X 9AMAT~~) (~~9AKİM X 9AMAT~~)
(9AAUB X 9AFİZ) (9AAUB X 9AKİM) (9AFİZ X 9AKİM)
(9AFİZ X 9AKİM) (9AFİZ X 9AMAT) (~~9AKİM X 9AMAT~~)

C_3 candidate item set (after pruning)

L2 X L2	L2_1	L2_2	L2_3
10AMAT X 10ATEÖ X 9AAUB	10AMAT	10ATEÖ	9AAUB
10AMAT X 9AFİZ X 9AKİM	10AMAT	9AFİZ	9AKİM
10AMAT X 9AAUB X 9AKİM	10AMAT	9AAUB	9AKİM
10AMAT X 9AAUB X 9AFİZ	10AMAT	9AAUB	9AFİZ
10AMAT X 10ATEÖ X 9AFİZ	10AMAT	10ATEÖ	9AFİZ
10ATEÖ X 9AAUB X 9AFİZ	10ATEÖ	9AAUB	9AFİZ
9AAUB X 9AFİZ X 9AKİM	9AAUB	9AFİZ	9AKİM

The existence of k-1 subsets of C_k in L_{k-1} is being controlled

L2 COURSE CODE
10AMAT X 10ATEÖ
10AMAT X 9AAUB
10AMAT X 9AFİZ
10AMAT X 9AKİM
10ATEÖ X 9AAUB
10ATEÖ X 9AFİZ
9AAUB X 9AFİZ
9AAUB X 9AKİM
9AFİZ X 9AKİM
9AFİZ X 9AMAT

Figure 6. Pruning in C_3 candidate item set in apriori algorithm

7th Step: In this step C_3 candidate item set is looked and L_3 frequent item set is obtained from the rates that are equal to minimum support rate or from higher rates. The consisting of L_3 frequent item set is seen at Figure 7 with sample data.

C_3 candidate item set

C3 COURSE CODE	SUPPORT
10AMAT X 10ATEÖ X 9AAUB	9
10AMAT X 10ATEÖ X 9AFİZ	9
10AMAT X 9AAUB X 9AFİZ	9
10AMAT X 9AAUB X 9AKİM	9
10AMAT X 9AF□ X 9AKİM	10
10ATEÖ X 9AAUB X 9AFİZ	~~8~~
9AAUB X 9AFİZ X 9AKİM	~~7~~

= > Minimum support (9)

L_3 candidate item set

L3 COURSE CODE	SUPPORT
10AMAT X 10ATEÖ X 9AAUB	9
10AMAT X 10ATEÖ X 9AFİZ	9
10AMAT X 9AAUB X 9AFİZ	9
10AMAT X 9AAUB X 9AKİM	9
10AMAT X 9AFİZ X 9AKİM	10

Figure 7. L_3 frequent item set formed by $C_3 \infty L_3$ transformation

8th Step: In this step, in order to determine L_4 frequent item set q new candidate item set C_4 is formed from the join process of L_3 set items. The forming of C_4 candidate item set is seen in figure 8 with sample data.

L_3 candidate item set

L3 COURSE CODE	SUPPORT
10AMAT X 10ATEÖ X 9AAUB	9
10AMAT X 10ATEÖ X 9AFİZ	9
10AMAT X 9AAUB X 9AFİZ	9
10AMAT X 9AAUB X 9AKİM	9
10AMAT X 9AFİZ X 9AKİM	10

$L_3 \infty L_3$ Join process

C_4 candidate item set

L3 X L3 COURSE CODE	L3_1	L3_2	L3_3	L3_4
10AMAT X 9AAUB X 9AFİZ X 9AKİM	10AMAT	9AAUB	9AFİZ	9AKİM
10AMAT X 10ATEÖ X 9AAUB X 9AFİZ	10AMAT	10ATEÖ	9AAUB	9AFİZ

Figure 8. C_4 candidate item set formed by $L_3 \infty L_3$ transformation

9th Step: According to the summary code of apriori algorithm, pruning procedure is required after C_4 candidate item set has been formed. If a subset of any item doesn't take place L_3 set, the related item is excluded from the evaluation and is deleted from C_4 candidate item set. All items at the end of pruning process can be excluded from the evaluation and the algorithm can be ended. At the end of the pruning made in C_k candidate item set when the algorithm is ended, the items in L_{k-1} frequent item sets are used in order to form association rules. Pruning process in C_4 candidate item set is seen with sample data in Figure 9.

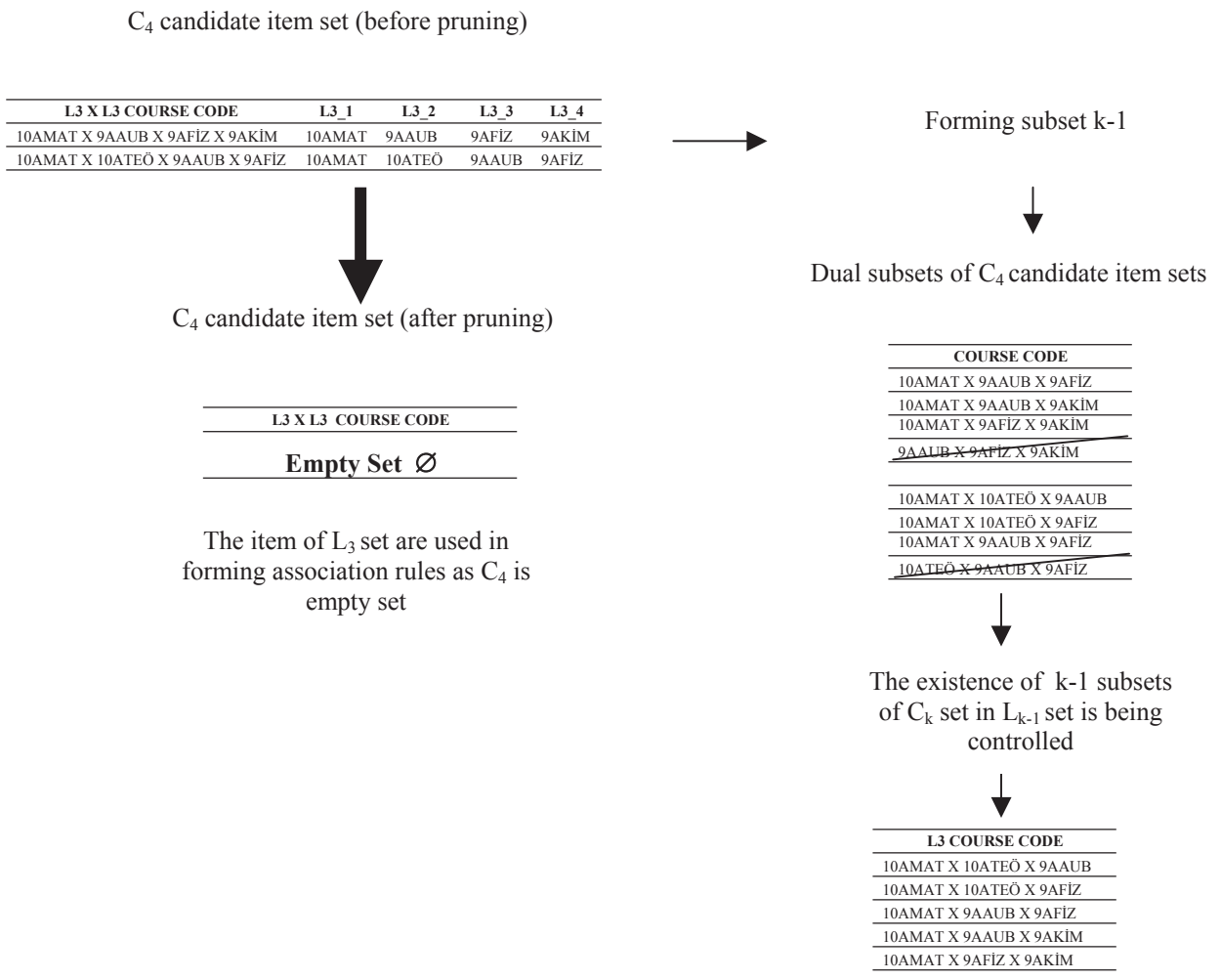

Figure 9. Prune process in C_4 candidate item set of apriori algorithm

10th Step: In tenth step, C_4 candidate item set is looked and from the rates which are equal to minimum support rate or from high rates L_4 frequent item set is obtained and the data is kept in veri_L4 table. During obtaining L_k frequent item set in C_k candidate item set, no item in C_k candidate item set can go beyond minimum support rate. In this situation the algorithm is finished and the items of L_{k-1} which are equal to determined minimum support rate or having high rate are used in forming association rules. These steps can be repeated according to the situation of data set or can be ended before they have reached at this step.

The Rules: In this part, association rules are formed from the item of the most frequent item set which has been obtained as the last and the ones which are equal to or higher than confidence rate determined in first step are showed. The rules are seen with data on Figure 10.

RULES	CONFIDENCE
9AFİZ X 9AKİM --> 10AMAT	91%
9AAUB X 9AKİM --> 10AMAT	100%
9AAUB X 9AFİZ --> 10AMAT	90%
10ATEÖ X 9AFİZ --> 10AMAT	90%

Figure 10. Association rules which are formed at the end of algorithm

2. Results (Findings)

When these rules are observed, we can got the result that the students who are unsuccessful in numeral courses become unsuccessful again in numeral courses of one year after. It is observed that the students who are unsuccessful in especially 9^{th} class numeral courses become unsuccessful in the course of math of 10^{th} class as well. According to rules; from the students of 11-A class it has been observed that:

- 91 percent of the ones who are unsuccessful in physic and chemistry of 9^{th} class are also unsuccessful in math of 10^{th} class,
- 100 percent of the ones who are unsuccessful in space science and astronomy are unsuccessful in math of 10^{th} class,
- 90 percent of the ones who are unsuccessful in astronomy, space science and physic of 9^{th} class are unsuccessful in math of 10^{th} class,
- 90 percent of the ones who are unsuccessful in basic electronics and measuring and physic of 9^{th} class are unsuccessful in math of 10^{th} class.

At last, the ways of students' being successful in these courses can be found by seeking the reasons of being unsuccessful. The students can be helped in choosing profession for themselves by revealing the data relations between their social activities, their concern fields and the lessons in which they are unsuccessful.

In this study, a sample study has been conducted to show how the apriori algorithm can be used in educational field and the results have been observed. Some rules about students' course points have been made out. Apriori algorithm is the most know algorithm in exploring the frequent item in data mining. In this algorithm, the database is required to be scanned many times to find frequent item set. It is used especially in analyzing market basket.

A school automation software has been prepared for algorithm's studying on data. This software is general purpose and prepared in order to reveal association rules from databases. With this reason, a database in general purpose has been prepared. Thanks to this software, all the information about the students and the school has been transferred on database. Then, association rules have been obtained by these data being used.

Data mining can be used effectively in educational institutes for leading education activities in an effective way, for watching students' performances continuously and directing students in course and profession choosing. Thus, the level of students' success can be raised. Being evaluated previously of the associations in which the students are unsuccessful with associations being observed, different strategies can be determined to make this situation away. Besides, it can be used as a helping tool in profession choosing of students according to their aptitudes and characteristics

The results we have made out from the data will allow us to guess truly for the future.

References

Gürsakal, N. (2001). Sosyal Bilimlerde Araştırma Yöntemleri, VİPAŞ, Bursa, 189.

Alataş, B., & Akın, E. (2004). Madenciliğinde Yeni Yaklaşımlar, YA/EM'2004 – Yöneylem Araştırması/Endüstri Mühendisliği – XXIV Ulusal Kongresi, Gaziantep – Adana, 15-18 Haziran.