



Knowledge Discovery from Library Automation via Bibliomining using the Apriori Algorithm

Ishak FIRAT¹ , Fulya ASLAY^{2,*} 

¹ *Erzincan Binali Yildirim University, Electrical Electronics Engineering Department, 24060, Erzincan, Turkey*

² *Erzincan Binali Yildirim University, Computer Engineering Department, 24060, Erzincan, Turkey*

Highlights

- It can be benefit for university library information system's managers in decision making.
- Detailed information can be obtained on the selection of library resources.
- Patterns of the publication borrowing process are presented.
- Member profiles are determined through association rules.
- It is ensured that publications can be purchased in the future time planning.

Article Info

Received: 30 Aug 2021
Accepted: 07 Jan 2022

Keywords

Association analysis
Apriori algorithm
Bibliomining
Library information system
Data mining

Abstract

Today which is called as the digital age with the considerably developing information systems, the constant increase in the data amount being recorded has revealed the concept of big data. Obtaining the strategic information which is crucial for decision-makers especially in managerial terms is only possible through processing these big data with accurate techniques. Data mining techniques have frequently been used in recent years in order to reach meaningful and useful knowledge among data stacks. In this study, the Apriori Algorithm was used for the managers of university library information systems, which provide data-oriented service, to make investment decisions in the future effectively and create user profiles. Within the scope of the study, an application was performed on the basis of an information system comprising of real data of Erzincan Binali Yildirim University Central Library. By means of association rules, which are one of the descriptive models of data mining, ten different association rules regarding the joint borrowing of publications were applied and results were obtained in the confidence intervals of 57.1% and 95.8%. In addition, information such as library inventory, member profile, and publication borrowing habits were obtained and evaluations were made in line with this information at the end of the study.

1. INTRODUCTION

In today's enterprises where most operations from production to management have changed and developed owing to Industry 4.0, techniques such as artificial intelligence, machine learning and data mining are used for processing the data recorded in all processes. Among these techniques, data mining which has been heard frequently in recent years and emerged as an interdisciplinary discipline, originates from statistical methods such as the Bayes' Theorem in the 1700s and the Regression Analysis in the 1800s. Data mining is associated with many techniques primarily such as statistics and also data science, database theory and machine learning. Data mining became a whole of techniques used by most enterprises owing to its contribution to the profitability of enterprises especially at the end of the 1990s. Analyzing the big data sets from different points of view, making the information acquired useful in line with a purpose and summarizing them and thus revealing the secret patterns within the data can only be possible through data mining. Today, data mining applications can be used in the business world, public institutions, universities and in every area of science. It is important to store big data stacks via technology which is obtained from all information technologies services and which now enables the knowledge, objects and humans to connect

to each other in the same environment via the internet of objects. Also it is important to create real-time decision-making strategies among these stacks.

As in most organizations desiring to gain a competitive advantage in line with strategic goals, universities too aim to provide a better service to their students and employees in order to provide a competitive advantage and a quality education. Therefore, they have to produce useful information for themselves among the data that are stored multidimensionally in the present information systems.

In the study, it was aimed to define the patterns that would be meaningful and useful for decision-makers by processing the data within the information system of Erzincan Binali Yıldırım University Central Library via the Apriori Algorithm, which was developed for the association rules analysis that is one of the descriptive data mining techniques. In this context, useful outcomes such as book borrowing habits, library inventory information, library member profile and publications borrowed together were produced. With the analysis of the aforementioned information, it was aimed to plan the library more efficiently in physical terms, renew the resources in such a way to meet the needs and increase the rate of using books by recommending new books to the members.

In the first section of the study, the concepts of data mining, bibliomining and association rules were discussed and the literature abstracts of studies conducted in this field were included. In the second section the acquirement, prior processing and cleaning process of the data used for data mining was explained and the association rules were analyzed. All the findings obtained as a result of the application were presented in the third section. Finally in the fourth section the findings obtained from the application were interpreted, the study was compared with similar studies and recommendations were made.

1.1. Literature Review

The literature related to association analysis used for numerous studies in different fields [1-5]. However, the studies on bibliomining which use library data are not adequate in number. In the study in proposed in [6], using statistical methods, the web diaries of the internet page of a library were analyzed via data mining techniques and it was tried to find the access patterns of library users. In the study in proposed in [7], the two-year book borrowing records of a university library in Taiwan were analyzed and they suggested that it was possible to explore different association classes within dispersed subject hierarchies via the circulation of the resource preference patterns of similar users. In the study in proposed in [8], it is examined the effectiveness and efficiency of the service provided by the library of Bielefeld University in Germany and the user intelligence applications in academic libraries. By this way, it was aimed to detect the patterns within the resource quest of users and thus enhance the service quality.

In the study in proposed in [9], the behaviors of Akdeniz University library users were examined on the basis of borrowing information using the Apriori and TwoStep algorithms. In the study in proposed in [10], it is examined whether the public funding transferred to a big public library in the Czech Republic via the K-Means algorithm is used efficiently or not and whether the users are able to reach the books and documents in the library in the most efficient way or not. In the study in proposed in [11], concerning the data mining application in library systems, the association analysis, clustering and ordered pattern analyses were used in revealing the frequency and indices of the books borrowed from libraries. Accordingly, it is possible to recommend books to users and also it becomes easier for users to find the books they may borrow in book arrangement on the basis of book use frequency.

In the study in proposed in [12], it is conducted a study to develop the library services by processing categorical information related to library users such as their university department, age range, and gender. In the study, the Apriori Algorithm, which is an association analysis application, was used for the purpose of detecting patterns such as publications borrowed and borrowing durations from the data of library users. In the study in proposed in [13], it is conducted a study to create a profile in academic libraries on the basis of fictional characters representing the purposes, benefits, and values of real persons. In the study which chose the K-Means application that is among clustering analysis methods, a simple chat reference dump was coded on the basis of chats made in the virtual environment in order to use the data services produced in libraries. The aforementioned dump was drawn on an X/Y diagram. While the X line represents the

quality of the user's information need in the diagram, the Y line represents the quality of the user's motivation.

In the study proposed in [14], association rule mining algorithm is applied to find out borrowing rules of readers according to their borrowing records and to recommend other booklists for them in a personalized way, so as to increase utilization rate of data resources at library. In the study proposed in [15], a markov logic network method was improved for reconstruct association rule-mining tasks for library recommendation and compared the method to traditional Apriori, FP-Growth, Inverse, Sporadic and UserBasedCF algorithms on two history library data sets and the Chess and Accident data sets.

2. DATA MINING

In the study [16], it is defined data mining as a discipline which is interactive with statistics, database technology, pattern recognition and machine learning and a secondary analysis of correlations which cannot be predicted in large databases. Data mining is the process of extracting self-knowledge from large-scale data [17]. Scott Nicholson denominated the data mining studies which he conducted with library data as "Bibliomining" and added a new term to the literature. Bibliomining was derived from the terms "bibliometry" and "data mining" [18]. Library data mining examines the parameters concerning by what users in which category the resources in libraries are used and how often and in what order. By examining the aforementioned parameters and detecting the patterns between them, it is aimed to enable decision-makers to procure the resources which address to relevant users, offer a higher productivity and are probable to be demanded afterwards. By this way, the data mining studies to be conducted especially in university libraries enable to determine the resource utilization of students and examine the effect of this situation on success level [18].

Data mining models are divided into two parts as predictive and descriptive. In predictive models, future behaviors can be determined by using predetermined data. On the other hand, descriptive models allow for finding the patterns to be used in decision-making [19]. While classification and regression models are predictive models, clustering, association rules, and sequential-time pattern models are descriptive models [20].

2.1. Association Rules

Association rules which are one of the descriptive models of data mining aim to explore interesting association patterns among great amounts of data and contribute to marketing, decision-making and business management. Therefore, the subject of association rules usually plays a key role in studies conducted in knowledge discovery in databases [21]. The purpose of this method is to reveal rules measuring the correlation between two or more properties [22]. Different algorithms have been used in the literature in analyzing the association rules. Some of these algorithms are AIS, SETM, Apriori, and AprioriTid. Of these algorithms, the most known one is the Apriori algorithm [23].

The Apriori algorithm differs from the AIS and SETM algorithms in the way in which candidate object sets are counted and these candidate sets are produced in each pass. In both the AIS algorithm and the SETM algorithm, candidate object sets are generated during scanning while reading the data. After a transaction is read, it is also checked whether large object sets are in these transactions. The generation of new candidate object sets is produced by combining large object sets obtained with other objects in the processes. Of course, this unnecessarily results in generating and counting many candidate object sets, which are actually the small set of objects, as if they were the large set of objects. This increases the time complexity of the algorithm. The Apriori algorithm, on the other hand, creates candidate objects by using only the object sets that were determined to be large in the previous scan, without involving any operations in the database. The Apriori algorithm is based on the assumption that any subset of a large set of objects will also be large. Thus, an object set consisting of k objects can be obtained by combining large object sets with $k-1$ objects and deleting those whose subsets are not large. At the end of this merge and deletion process, fewer candidate object sets are formed [24]. Due to all these advantages, the apriori algorithm was preferred in the study.

The problem of mining association rules notations and definitions were shown below [24]:

- $I = \{i_1, i_2, \dots, i_m\}$ – set of items;
- D = set of transactions; each transaction t is included in I ;
- X = set of items from I , t contains X .
- An association rule is a pair $X \rightarrow Y$, where $X \subseteq I$, $Y \subseteq I$, $X \cap Y = \emptyset$.
- Confidence of the rule $X \rightarrow Y$ is c , if $c\%$ of the transactions in D that contain the set X , contain also the set Y .
- Support of rule $X \rightarrow Y$ is s , if $s\%$ of the transactions in D contains the set $X \cup Y$.

2.2. The Apriori Algorithm

The Apriori algorithm was developed by Agrawal and Srikant in 1994 for the association rules. The association rules check whether an element exists or not. In contradistinction to others, this algorithm was designed for databases containing a procedure and time stamp. Every procedure is a cluster of elements [25]. The best known method for reducing the calculation complexity of object-clusters mining is the Apriori algorithm. The Apriori algorithm offers an efficient solution for the elimination of certain candidate object-clusters without counting their support values. The algorithm was named after the use of prior knowledge in the mining of frequent object-clusters [23]. If an object-cluster is frequent, all subsets of this cluster should also be a frequent object-cluster.

In the Apriori algorithm, the correlation between elements is calculated for certain support and confidence values. As these two values increase, the importance of the association of the two elements also increases. The confidence value is a value demonstrating the possibility for an element to coexist with another element. The support value is a value demonstrating how often the two elements coexist [23]. The Apriori algorithm uses a repetitive approach which is known as the level logic quest. In this approach the k -element object-clusters are created by combining the $(k-1)$ -element object-clusters. First of all, the support values of objects are calculated for creating the 1-element object-clusters. Frequent elements in the $(k-1)$ transition are used for creating the frequent candidate object-clusters (C_k) in the (L_{k-1}) , k transition. The database is screened, the support values of the C_k s are calculated and the L_k s are created. By this way, the algorithm continues until it is no more able to find frequent object-clusters [23]. In order to understand how the L_{k-1} is used in obtaining the L_k , the steps comprising the Merging and Pruning procedures of the algorithm are as follows [26].

In the process of Merging step, the candidate k -element object-clusters (C_k) are created by the combination of the L_{k-1} among themselves in order to find the L_k . Let's say the I_1 and I_2 are object-clusters in the L_{k-1} . The $I_i[j]$ demonstration represents a j object in the I_i (for example, $I_1[k-2]$ represents the second last object in the I_1). The Apriori assumes that the objects or object-clusters in a procedure are arranged in an alphabetical order. For the $k-1$ -element object-clusters the objects are arranged as $I_i[1] < I_i[2] < \dots < I_i[k-1]$. If the first $(k-2)$ objects of the L_{k-1} are common, the $L_{k-1} \times L_{k-1}$ combination may occur [23].

In the process of Pruning step, the elements of the C_k may or may not be frequent. However, it contains all frequent object-clusters and it is the superset of the L_k . The results of the screening that is performed for calculating the support value of each candidate in the C_k determine the L_k (for example, all candidates with a value greater than and equal to the minimum support value are defined as frequent object-cluster and are included in the L_k). The C_k may be enormous and may require considerably heavy calculations. In order to reduce the size of the C_k the Apriori property is used. Any non-frequent $(k-1)$ -object-cluster will not be the subset of a frequent k -object-cluster. Therefore, unless the great $(k-1)$ -element subsets of a candidate k -object-cluster are available in the L_{k-1} , this candidate will not be a frequent object-cluster and will be excluded from the C_k [23].

3. MATERIAL AND METHODS

In this study, the analyses were performed using an inventory comprising of the data of Erzincan Binali Yıldırım University Central Library via the IBM SPSS Modeler which is among data mining tools. The data were prepared within the application and made available in order for the algorithm to function. The

data preparation process comprised of the phases of data selection, data prior processing and data reduction [19]. The present data undergoing the prior processing were primarily subjected to the cleaning and transformation processes. Then the Apriori algorithm, which is an association rules analysis, was used on the data and the results were interpreted. Figure 1 shows the flowchart of the study.

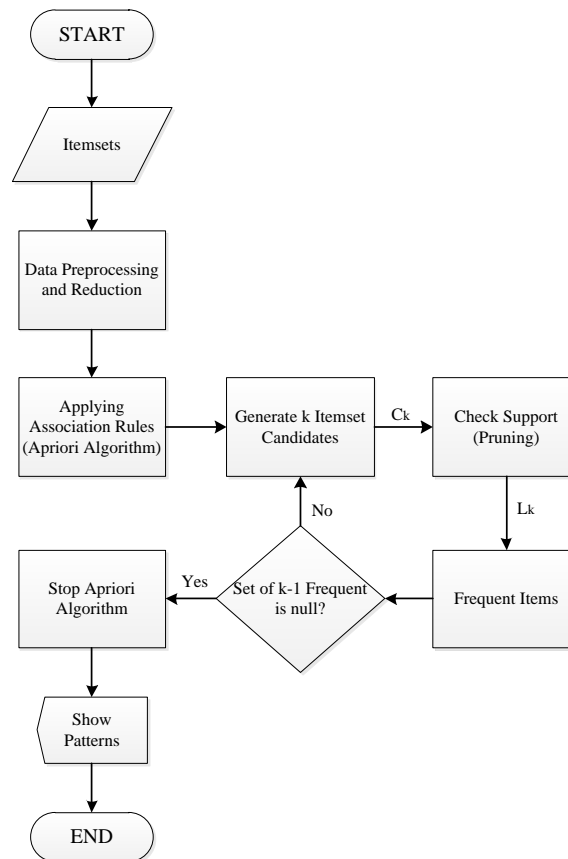


Figure 1. Flowchart of the study

3.1. Data Selection

The library data comprises of attributes such as book asset id, Dewey class, book publish date, number of book borrowings, number of borrowing days, number of members, faculty of members, title of members and sex of members. The Dewey Decimal Classification system is a classification system enabling all books which contain knowledge on the same topic to be found under the same class, in order to access the information sources in a library easily. In this system the books are grouped according to their topics and these topics are identified with numeric expressions. The system has 10 main categories. These main categories are also divided into sub-categories and topics within themselves again with numeric expressions. For example, in the Dewey Classification System the classification number “400” defines the works in the area of “Language and Linguistics” in general. When another number is added to 4, subsets from general to detail will emerge. For example, there are works related to the “Turkish Language” within the subset represented by the number “410” and there are works related to “Turkish” within the subset represented by the number “413”.

3.2. Data Prior Processing and Reduction

In the study there were three different files containing the records of the library database. These files were the “Catalogue” file containing core data such as the publish dates and authors of publications available in the library; the “Borrowing” file containing data such as the publication information, member information and borrowing dates of borrowing processes in the library; and the “Member” file containing data such as the department, title, address of the library members. Examining all the data in the files; irregular and

missing records were encountered and all these records were primarily cleaned via the SQL queries. Areas with nominal values such as the sex, faculty and title of the members were transformed into categorical values and were made suitable for the IBM SPSS Modeler program. In addition since the members table contained all members borrowing and not borrowing from the library, the borrowing table was rearranged only according to the members borrowing and was made suitable for the analyses.

4. MODELING VIA THE APRIORI ALGORITHM FROM THE ASSOCIATION RULES

Prior to creating the association rules, the data were visualized with the help of a network diagram in the IBM SPSS Modeler program. A general view is created for the data before obtaining rules with the help of data visualization, which is as important as all other techniques in data mining.

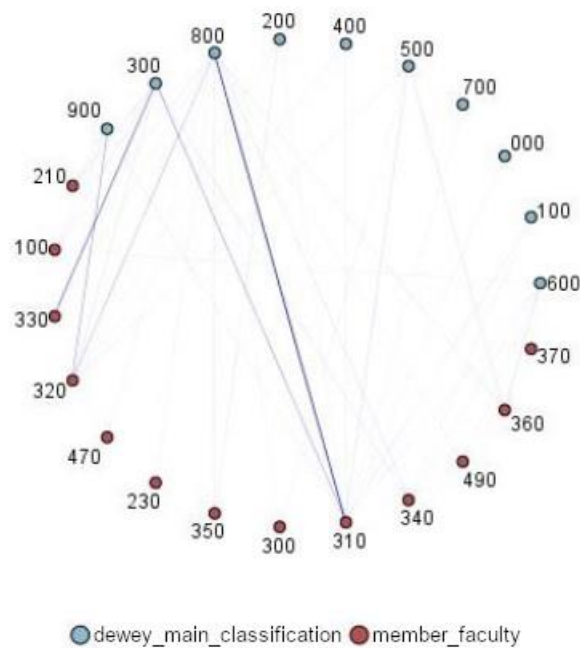


Figure 2. Correlation between the faculties of the members and the codes of the publications borrowed

In Figure 2, the faculties of the library members and the main class codes of the books they borrowed were visualized with the help of the network diagram. The figure has lines demonstrating how often the main class of publications and the faculties of the members were together. The dark lines demonstrate that the associations were higher in number, while the thin lines demonstrate that the associations were lower in number. Here, the most frequent association was observed in the borrowing process conducted by the members registered to the Faculty of Education coded 310 from the Literature and Rhetoric main publication group with the Dewey code of 800. Also it was observed that the borrowing process conducted by the members registered to the Faculty of Science and Letters coded 320 from the Geography and History main publication group with the Dewey code of 900 was a frequently repeated association.

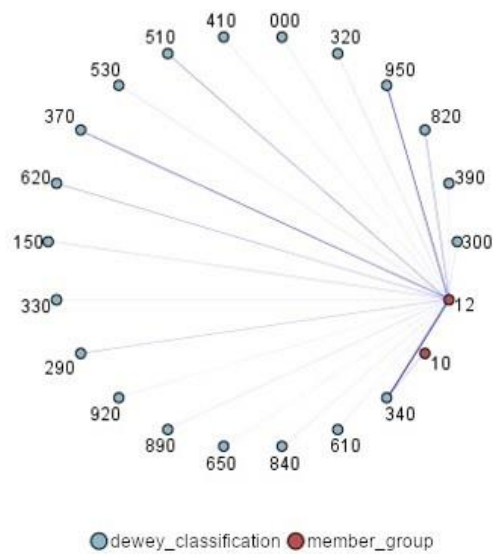


Figure 3. *Correlation between the titles of the members and the codes of the publications borrowed*

Figure 3 demonstrates the occurrence frequency of associations between the member groups and dewey publication classes during borrowing processes. According to the figure, the students with a member code of 12 frequently borrowed the following publications: law coded 340, Education coded 370, Maths coded 510 and works on the General Asian History including the Turkish History coded 950. It was observed that the academicians with member code of 10 borrowed publications mostly from the law group coded 340. Figure 4, which was obtained with the help of the network diagram, demonstrates which publication groups the members chose more according to their sex. According to the figure it was observed that the female members coded 02 borrowed from the Turkish Literature publication group coded 810 very frequently. Another important matter here is that the members coded 03 whose sex information was not available had very frequent borrowing process records, as well.

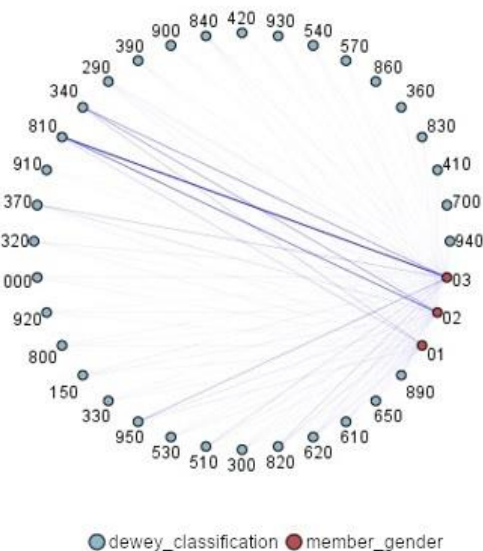


Figure 4. *Correlation between the sex of the members and the codes of the publications borrowed*

Figure 5 demonstrates in what time the members returned the publications according to their groups. It was observed that the students’ association of returning the publications within the first 15 days was higher than all other associations.

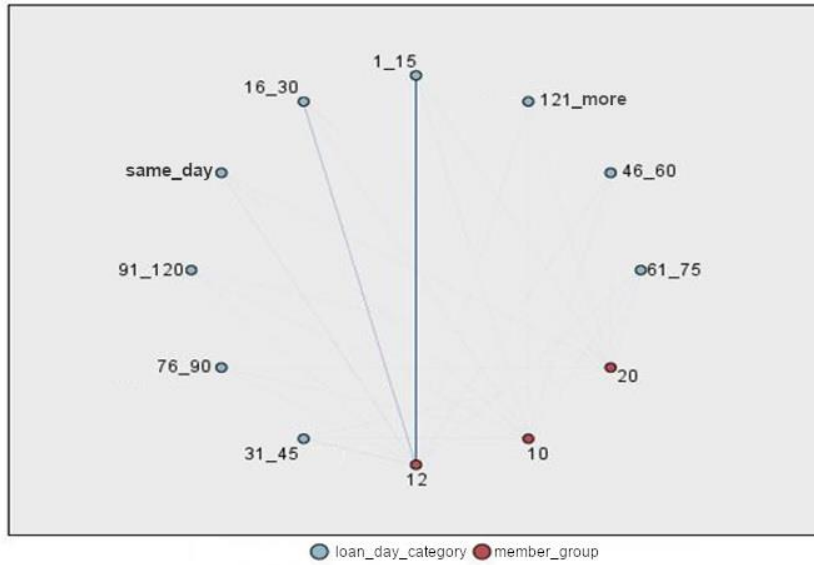


Figure 5. *Correlation between the titles of the members and the borrowing durations*

In the visualization processes conducted with the help of the network diagrams, since the student group had too many members, this repressed the data of the academican group and made it difficult to visualize it. Therefore, the two network diagrams in Figures 6 and 7 were obtained by filtering the data of only the academican member group. The diagrams acquired demonstrate the most intense associations of the academicians coded 10 according to their sex, educational institution and time of returning the publications they borrowed.

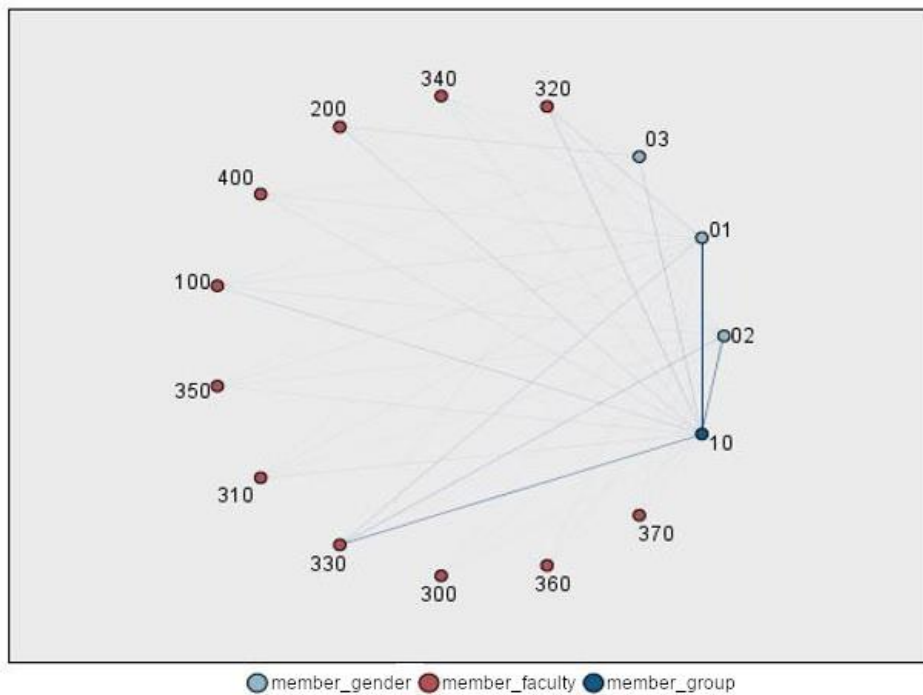


Figure 6. *Correlation between the faculties and sex of the academican members*

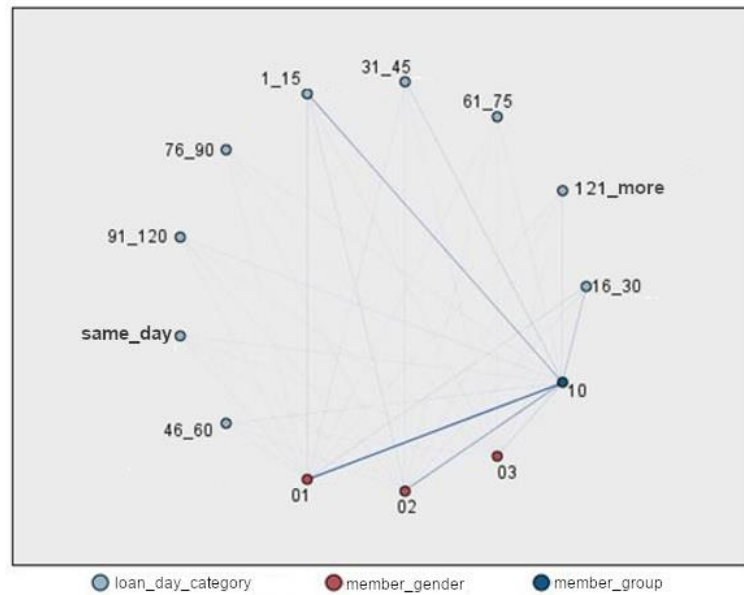


Figure 7. Correlation between the sex of the academicians and the borrowing durations

Following the phase of visualization with the network diagrams, the stream in Figure 8 which was created for performing association analysis via the IBM SPSS Modeler program was operated and a node listing the association rules was created. Figure 9 demonstrates the output acquired. According to the support and confidence threshold value entered into the system, the Apriori algorithm had produced 4470 rules. These rules can be arranged individually according to their support, confidence and lever values. The support value indicates at what percentage the initial and successor publication groups are included in all processes; the confidence value indicates with what percentage of possibility a member borrowing a publication from the initial publication group borrows a publication from the successor publication group and finally the leverage value indicates the interestingness of the rules. The threshold value was determined to be 1% for minimum value and 10% for minimum confidence. The maximum initial element number was determined to be three.

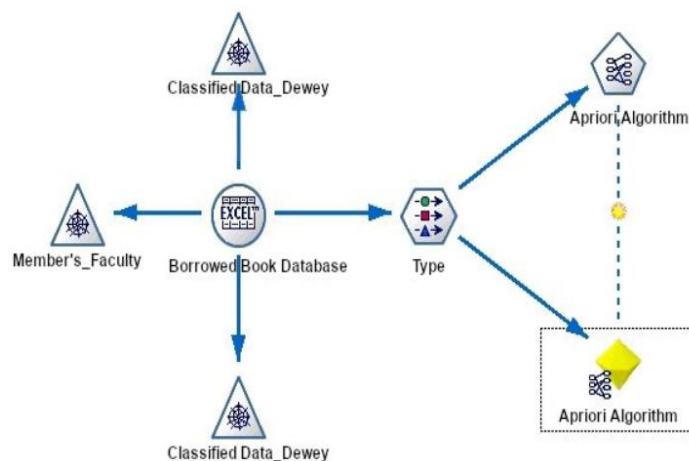


Figure 8. The stream created for the association analysis

Consequent	Antecedent	Support %	Confidence %	Lift
810 = T	150 = T 890 = T 820 = T	1,243	95,876	2,288
810 = T	840 = T 150 = T 820 = T	1,051	95,122	2,27
950 = T	910 = T 320 = T	1,013	94,937	7,102
810 = T	830 = T 890 = T 820 = T	1,038	93,827	2,239
810 = T	920 = T 820 = T	1,23	93,75	2,237
810 = T	890 = T 370 = T 820 = T	1,307	93,137	2,223
810 = T	830 = T 840 = T 820 = T	1,089	92,941	2,218
810 = T	860 = T 840 = T	1,064	92,771	2,214
810 = T	800 = T 840 = T	1,038	92,593	2,21
810 = T	860 = T 820 = T	1,525	92,437	2,206

Figure 9. Association rules arranged according to the confidence value

5. RESULTS AND DISCUSSION

Table 1 shows a part of the rules obtained as a result of applying the Apriori algorithm for the association analysis. In order to see what kind of publications were borrowed together, the rules were arranged according to the confidence criteria and the results were interpreted.

Table 1. Arrangement of the association rules in different topic titles

Premise	Successor	Support %	Confidence%	Lever
Turkish Literature	Psychology–Other Languages’ Literature–English Literature	1,243	95,876	2,288
Asian History	Geography–Political Sciences	1,013	94,937	7,102
English Literature	French Literature–Social Sciences–Turkish Literature Biography – Non-Christian	1,141	76,404	4,276
Political Sciences	Religions–Asian History	1,025	65,000	8,699
Physics	Engineering–Maths-Turkish Literature	1,115	62,069	9,666
Psychology	Medical Sciences–Social Sciences	1,013	62,025	6,220
Education	Life Sciences–Turkish Literature	1,295	61,386	3,994
Maths	Physics–English Literature–Turkish Literature	1,192	60,215	5,947
Political Sciences	Law–Asian History	1,871	58,904	7,883
Non-Christian Religions	Biography–Political Sciences–Turkish Literature	1,077	57,143	5,184

For the association rule of the Turkish Literature → Psychology – Other Languages’ Literature – English Literature publication groups; the possibility of the Turkish Literature and Psychology – Other Languages’ Literature – English Literature publication groups to be encountered together in the total borrowing movements was 1.243%. In addition, it can be asserted that the library members borrowing from the Turkish

Literature publication group also borrowed from the Psychology – Other Languages' Literature – English Literature publication groups at a probability of 95.876%.

For the association rule of the Asian History → Geography – Political Sciences publication groups; the possibility of the Asian History and Geography – Political Sciences publication groups to be encountered together in the total borrowing movements was 1.013%. In addition it can be asserted that the library members borrowing from the Asian History publication group also borrowed from the Geography – Political Sciences publication groups at a probability of 94.937%.

For the association rule of the English Literature → French Literature – Social Sciences – Turkish Literature publication groups; the possibility of the English Literature – French Literature – Social Sciences – Turkish Literature publication groups to be encountered together in the total borrowing movements was 1.141%. In addition it can be asserted that the library members borrowing from the English Literature publication group also borrowed from the French Literature – Social Sciences – Turkish Literature publication groups at a probability of 76.404%.

For the association rule of the Political Sciences → Biography – Non-Christian Religions – Asian History publication groups; the possibility of the Political Sciences – Biography – Non-Christian Religions – Asian History publication groups to be encountered together in the total borrowing movements was 1.025%. In addition it can be said that the library members borrowing from the Social Sciences publication group also borrowed from the Biography – Non-Christian Religions – Asian History publication groups at a probability of 65%.

For the association rule of the Physics → Engineering – Maths – Turkish Literature publication groups; the possibility of the Physics – Engineering – Maths – Turkish Literature publication groups to be encountered together in the total borrowing movements was 1.115%. In addition it can be stated that the library members borrowing from the Physics publication group also borrowed from the Engineering – Maths – Turkish Literature publication groups at a probability of 62.069%.

For the association rule of the Psychology → Medical Sciences – Social Sciences publication groups; the possibility of the Psychology – Medical Sciences – Social Sciences publication groups to be encountered together in the total borrowing movements was 1.013%. In addition it is possible to state that the library members borrowing from the Psychology publication group also borrowed from the Medical Sciences – Social Sciences publication groups at a probability of 62.025%.

For the association rule of the Education → Life Sciences – Turkish Literature publication groups; the possibility of the Education – Life Sciences – Turkish Literature publication groups to be encountered together in the total borrowing movements was 1.295%. In addition it can be asserted that the library members borrowing from the Education publication group also borrowed from the Life Sciences – Turkish Literature publication groups at a probability of 61.386%.

For the association rule of the Maths → Physics – English Literature – Turkish Literature publication groups; the possibility of the Maths – Physics – English Literature – Turkish Literature publication groups to be encountered together in the total borrowing movements was 1.1192%. In addition it is possible to state that the library members borrowing from the Math publication group also borrowed from the Physics – English Literature – Turkish Literature publication groups at a probability of 60.215%.

For the association rule of the Political Sciences → Law – Asian History publication groups; the possibility of the Political Sciences – Law – Asian History publication groups to be encountered together in the total borrowing movements was 1.871%. In addition it can be asserted that the library members borrowing from the Political Sciences publication group also borrowed from the Law – Asian History publication groups at a probability of 58.904%.

For the association rule of the Non-Christian Religions → Biography – Political Sciences – Turkish Literature publication groups; the possibility of the Non-Christian Religions – Biography – Political

Sciences – Turkish Literature publication groups to be encountered together in the total borrowing movements was 1.077%. In addition it is possible to state that the library members borrowing from the Non-Christian Religions publication group also borrowed from the Biography – Political Sciences – Turkish Literature publication groups at a probability of 57.143%.

6. CONCLUSION

In this study the bibliomining process was applied to the 2010-2017 data of Erzincan Binali Yıldırım University Central Library and useful information such as borrowing habits, inventory information, library member profile and publications borrowed together, was obtained for the publications in the library. In the study, the data to be analyzed were primarily acquired and were subjected to a prior processing. Then the cleaning and transformation processes were performed on the data via the SQL queries. In analysis of the data, the Apriori algorithm which is among the association analysis algorithms also known as the market basket analysis, was used with the help of the IBM SPSS Modeler software. 4470 association rules were produced and thus patterns concerning what kind of publications the members borrowed together were obtained. The support and confidence values needed by the Apriori algorithm were shaped according to the dimension of the data.

All the patterns obtained guide library decision-makers and useful outcomes such as library inventory information, library member profile, borrowing habits and publications borrowed together, were produced. In the light of these information; taking the book circulation into account, it will be possible to recreate the settlement orders, have a more efficient physical planning for the library, renew the resources in such a way to meet the needs and increase the publication borrowing rate by recommending new books to the members.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

- [1] Sözen, E., Bardak, T., Peker, H., Bardak, S., “Analysis of Factors Effecting Furniture Selection Using Apriori Algorithm”, *Journal of Advanced Technology Sciences*, 6(3): 679-684, (2017).
- [2] Bardak, S., Bardak, T., “Investigation of the Most Problems in Furniture Products With Data Mining”, *Euroasia Journal of Mathematics, Engineering, Natural & Medical Sciences International Indexed and Refereed*, 7(13): 285-292, (2020).
- [3] Keleş, A. E., Kaya, M., “The Analysis of the Factors Affecting the Productivity in the Wall Construction of the Using Apriori Data Mining Method”, *Academic Informatics Conference*, 831-836, (2014).
- [4] Anwar, M.A., Ahmed, S.S., Khan, M.A.U., “Application of Apriori Algorithm on Examination Scores, In: Patnaik”, *Advances in Machine Learning and Computational Intelligence*, Springer, Singapore, (2021).
- [5] Wang, C., Zheng, X., “Application of improved time series Apriori algorithm by frequent itemsets in association rule data mining based on temporal constraint”, *Evolutionary Intelligence*, 13: 39–49, (2020).
- [6] Takci, H., Sogukpinar, I., “Discovery of Library Users' Access Patterns”, *Information World*, 3(1): 12-26, (2002).
- [7] Pu, H., Yang, C., “Enriching user-oriented class associations for library classification schemes”, *The Electronic Library*, 21(2): 130-141, (2003).

- [8] Decker, R., Höppner, M., "Strategic planning and customer intelligence in academic libraries", *Library Hi Tech*, 24(4): 504-514, (2006).
- [9] Ucan, O., "Data Mining Applications in Digital Libraries: The Example of Akdeniz University Central Library", Master Thesis, Akdeniz University, Institute of Social Sciences, Antalya, 1-3, (2010).
- [10] Hajek, P., Stejskal, J., "Advances in Enviroment Computational Chemistry and Bioscience", *Wseas Publisher, Montreux*, 339-344, (2012).
- [11] Uppal, V., Chandwani, G., "An Empirical Study of Application of Data Mining Techniques in Library System", *International Journal of Computer Applications*, 74(11): 42-46, (2013).
- [12] Zhang, Q. S., Wang, X. Y., "Research of Personalized Information Service Based on Association Rules", *Advanced Materials Research*, 760(1): 1800-1803, (2013).
- [13] Tempelman-Kluit, N., Pearce, A., "Invoking the User from Data to Design", *College and Research Libraries*, 75(5): 616-640, (2014).
- [14] Yi, K., Chen, T., Cong, G., "Library personalized recommendation service method based on improved association rules", *Library Hi Tech*, 36(3): 443-457, (2018).
- [15] Wang, S., Xu, J., Feng, Y., Peng, M., Ma, K., "A Markov logic network method for reconstructing association rule-mining tasks in library book recommendation", *Information Discovery and Delivery*, (2021).
- [16] Hand, D. J., "Data Mining: Statistics and More?", *The American Statistician*, 52(2): 112-118, (1998).
- [17] Ganesh, S., "Data Mining: Should It be Included in the 'Statistics' Cirriculum", *The 6th International Conference on Teaching Statistics*, South Africa, 1-4, (2002).
- [18] Nicholson, S., "The bibliomining process: Data Warehousing and Data Mining for Library Decision Making", *Information Technology and Libraries*, 22(4): 146-156, (2003).
- [19] Fayyad, U., Shapiro, G. P., Smyth, P., "Knowledge Discovery and Data Mining: Towards a unifying framework", *International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon, 82-88, (1996).
- [20] Schroeder, A. T., "Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support", *Journal of the American Society for Information Science*, 48(9): 862-863, (1997).
- [21] Han, J., Fu, Y., "Discovery of Multiple-Level Association Rules from Large Databases", *21th International Conference on Very Large Data Bases*, Zurich, 1-12, (1995).
- [22] Larose, D. T., Larose, C. D., "Discovering knowledge in data: an introduction to data mining", 2nd edition, *A John Wiley & Sons Publisher*, New Jersey, 249-255, (2004).
- [23] Agrawal, R., Srikant, R., "Fast Algorithms For Mining Association Rules", *20th International Conference on Very Large Data Bases*, Santiago, 1-32, (1994).
- [24] Agrawal, R., Imielinski, T., Swami, A., "Mining association rules between sets of items in large databases", In *Proceeding: International Conference Management Data*, (1993).

- [25] Liu, H., Wang, B., “An association rule mining algorithm based on boolean matrix”, *Data Science Journal*, 6(9): 559-565, (2007).
- [26] Han, J., Kamber, M., “Data Mining: Concepts & Techniques”, 3rd Edition, Morgan Kauffmann Publishers, San Francisco, 155-182, (2006).