

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

Fall 11-17-2021

Analysis of Library Book Borrower Patterns Using Apriori Association Data Mining Techniques

Lucky Zamzami

Universitas Andalas

Ari Agung Prastowo

Universitas Padjadjaran

R Rulinawaty

Universitas Terbuka

Robbi Rahim

Sekolah Tinggi Ilmu Manajemen Sukma, usurobbi85@zoho.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>

 Part of the Computer and Systems Architecture Commons, Educational Technology Commons, Instructional Media Design Commons, and the Other Education Commons

Zamzami, Lucky; Prastowo, Ari Agung; Rulinawaty, R; and Rahim, Robbi, "Analysis of Library Book Borrower Patterns Using Apriori Association Data Mining Techniques" (2021). *Library Philosophy and Practice (e-journal)*. 6632.

<https://digitalcommons.unl.edu/libphilprac/6632>

Analysis of Library Book Borrower Patterns Using Apriori Association Data Mining Techniques

Lucky Zamzami¹, Ari Agung Prastowo², Rulinawaty³, Robbi Rahim^{4*}

¹Universitas Andalas, Padang, Indonesia. Email: luckyzamzami@soc.unand.ac.id

²Universitas Padjadjaran, Indonesia. Email: ari.agung@unpad.ac.id

³Universitas Terbuka, Indonesia. Email: ruly@ecampus.ut.ac.id

⁴Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia. Email: usurobbi85@zoho.com

Corresponding Email: usurobbi85@zoho.com

Abstract. The library is one of the most important facilities because it manages collections of written works, printed works, and recorded works and can provide information resources as well as be a driving force for the advancement of an educational institution. Conventional libraries will have piles of book borrowing transaction data recorded in the agenda book, which is only an archive, and the placement of books far apart, which causes members to take longer to find books when borrowing books of different types, is an issue that must be addressed. To overcome these two issues, a recommendation for an intelligent system is required. This study was carried out in one of Indonesia's vocational schools, with data collected through observation and interviews with librarians. The goal of this study is to examine the borrowing pattern of books using association data mining techniques. The association method used is a priori, and it will result in recommendations for association rules. The result of the association rule with reference to the 2-itemset with the highest value is a combination of Religion book and Physical Education book with 8 percent support and 100 percent confidence, whereas the association rule with 3-itemset reference resulted in 4 rules with 6 percent support and 100 percent confidence. The result is an application that can generate association rules for book recommendations and book placement recommendations.

Keywords: Data mining, apriori method, library, book borrowing, rule association, book placement.

1. Introduction

The library is one of the most important facilities for students in educational institutions. In addition to managing collections of written works, printed works, and recorded works, the library is now regarded as an information resource that drives the advancement of an institution [1]–[4], including secondary and vocational school levels. Borrowing data is traditionally processed in such a way that transaction data for borrowing books is simply stored without further processing. Book lending transaction data recorded in the loan agenda book could not be reprocessed optimally. As a result, the data only becomes a pile of archives, despite the fact that the data can be used for book recommendations by locating the data sets that appear the most frequently in a data set. In addition, one of the other initiatives is to improve book placement. The placement of books tends to be far apart, so when borrowing books of various types, the bookshelves will be far apart, causing users to take longer to find books. As a result, there is a requirement for a method to provide book recommendations to library visitors or users as well as book placement recommendations to librarians.

Data mining is the process of extracting useful information from large amounts of data by employing statistical methods, mathematics, and artificial intelligence technology [5]–[10]. The process of discovering hidden patterns in the data of each previously unknown book lending transaction is known as data mining [11]. Several commonly used data mining methods are a priori methods [12]–[15]. The a priori method was chosen because it takes less time than other methods in the association rule method with actual data if the specified minsup is large [16], [17]. This study is similar to several previous studies, namely those conducted by (Waliyansyah, Novita, and Saputro 2021) [18] about book placement in libraries using the Apriori method, and (Prehanto et al. 2019) [19] about modeling library books using the association rule method with the a priori algorithm in determining book placement and book lending analysis. However, this study differs from the others in one important way: the method of processing it prior to mining. The dataset and the test application used in the research (Waliyansyah, Novita, and Saputro 2021) differ. It is in the dataset and the percentage of support values and confidence values from the prediction results while in research (Prehanto et al. 2019). These factors distinguish the author's research from previously conducted studies.

2. Research Methodology

The apriori method is used in this study to analyze transaction history data on library book lending. Data collection, data pre-processing, analysis of the highest frequency pattern using the a priori method, formation of association rule patterns, and testing of experimental results are the stages of research. The use of association rules with the a priori method has advantages in terms of simplicity and data handling capability. The Rapidminer application is used in this study. For more details, see Figure 1 below.

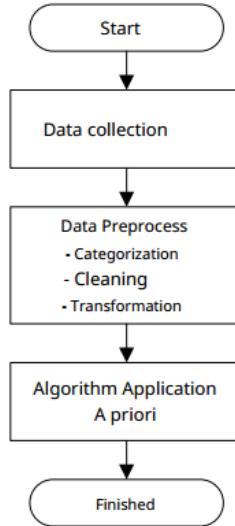


Figure 1. Data research stages

From March 2021 to June 2021, the data collected is transaction data on borrowing books from a library at one of the vocational schools. Three activities are carried out during the data pre-processing stage: (1) book categorization, (2) data cleaning, and (3) data transformation. The Apriori Algorithm is implemented using the RapidMiner software tools. The minimum antecedent support and confidence used are 5% and 10%, respectively. Data cleaning is performed during the data pre-processing stage by deleting book lending transactions that only contain one book item. While the data transformation is carried out by converting loan transaction data into a boolean table.

Figure 2 depicts the steps of the Apriori method, which is a solution method for analysing the pattern of borrowing library books.

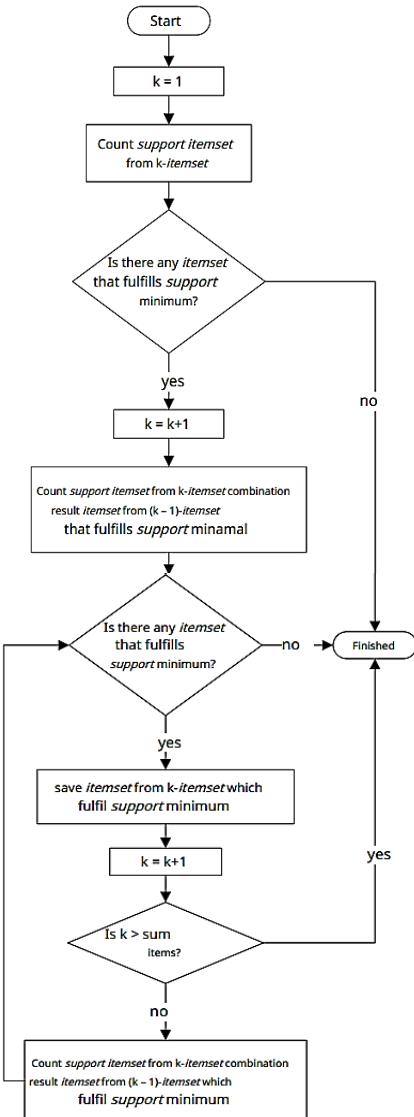


Figure 2. Apriori Algorithm Steps

The process of calculating support is depicted in Figure 2. Formula 1 is used to calculate the support.

$$\text{Support (A)} = \frac{\text{Number of transaction in which A appears}}{\text{Total number of transactions}}$$

$$(1)$$

The Apriori method, in addition to calculating support, also computes confidence. Formula 2 is used to calculate confidence.

$$\text{Confidence (A} \rightarrow \text{B)} = \frac{\text{Support(AUB)}}{\text{Support(A)}}$$

$$(2)$$

3. Results and Discussion

There were 97 transactions for borrowing library books at one of the vocational schools that were collected. The data becomes 50 transactions after preprocessing. Books are divided into the following categories: (1) ICT textbooks, (2) Religion textbooks, (3) Indonesian language textbooks, (4) English textbooks, (5) Chemistry textbooks, (6) Mathematics textbooks, (7) Physical Education textbooks, and (8) PKN textbooks. Association analysis, also known as association rule mining, is a data mining technique used to discover associative rules between item combination rules.

The data processed in determining the borrowing pattern for this book is visitor data. For library visitors at one of the vocational schools, the data variables are as follows: (1) name; (2) day/date of borrowing; (3) gender; (4) class; (5) the title of the first book; (6) the title of the second book; and (7) the title of the third book. Each visitor will be divided into groups based on the title of the book borrowed from the variables. This is done so that the a priori method can be used to easily analyze the data. Table 1 contains a recapitulation of library visitor data.

Table 1. Library Visitor Data

Visitor Name	Date	Class	First book title	Second book title	Third book title
Person1	23 March 2021	XII KCR ²	ICT textbooks	PKN textbooks	Physical Education textbooks
Person2	23 March 2021	XI KCR ²	PKN textbooks	Mathematics textbooks	-
Person3	23 March 2021	XII KCR ¹	Chemistry textbooks	Physical Education textbooks	-
Person4	23 March 2021	XII KK	Religion textbooks	Chemistry textbooks	Physical Education textbooks
Person5	04 April 2021	XI BUSAN A ²	Religion textbooks	PKN textbooks	ICT textbooks
Person6	04 April 2021	XII BOGA	ICT textbooks	-	-
Person7	04 April 2021	XII BUSAN A ²	Chemistry textbooks	Mathematics textbooks	ICT textbooks
Person8	04 April 2021	XII KCR ²	ICT textbooks	Religion textbooks	Physical Education textbooks
Person9	05 April 2021	XII BOGA	Chemistry textbooks	ICT textbooks	Religion textbooks
Person10	05 April 2021	X KCR ²	ICT textbooks	Religion textbooks	-
Person11	13 Mei 2021	XII PERHO TELAN	Physical Education textbooks	-	-
Person12	13 Mei 2021	XII KCR ²	English textbooks	Mathematics textbooks	Chemistry textbooks
Person13	13 Mei 2021	XI RPL	Physical Education textbooks	-	-
Person14	14 Mei 2021	XII TKJ	Mathematics textbooks	Indonesian language textbooks	-
Person15	21 Mei 2021	XI RPL	Chemistry textbooks	-	-
Person16	21 Mei 2021	XI BUSAN A ¹	Mathematics textbooks	English textbooks	Indonesian language textbooks
Person17	21 Mei 2021	X BOGA	Mathematics textbooks	Indonesian language textbooks	English textbooks
Person18	21 Mei 2021	X KK	Chemistry textbooks	English textbooks	ICT textbooks
Person19	21 Mei 2021	XI KK	Physical Education textbooks	ICT textbooks	-
Person20	21 Mei 2021	XI PERHO TELAN	Religion textbooks	ICT textbooks	-

Following the collection of visitor data, visitors are classified based on the title of the book borrowed. However, before doing so, a book guide table based on the title of each book is created, as shown in Table 2 below.

Table 2. Shows the book lending transactions.

No	Visitor	Borrowed book
1	Person1	ICT textbooks, PKN textbooks, Physical Education textbooks
2	Person2	PKN textbooks, Mathematics textbooks
3	Person3	Chemistry textbooks, Physical Education textbooks
4	Person4	Religion textbooks, Chemistry textbooks, Physical Education textbooks
5	Person5	Religion textbooks, PKN textbooks, ICT textbooks
6	Person6	ICT textbooks
7	Person7	Chemistry textbooks, Mathematics textbooks, ICT textbooks
8	Person8	ICT textbooks, Religion textbooks, Physical, Education textbooks
9	Person9	Chemistry textbooks, ICT textbooks, Religion textbooks
10	Person10	ICT textbooks, Religion textbooks

No	Visitor	Borrowed book
11	Person11	Physical Education textbooks
12	Person12	English textbooks
13	Person13	Physical Education textbooks
14	Person14	Mathematics textbooks, Indonesian language textbooks
15	Person15	Chemistry textbooks
16	Person16	Mathematics textbooks, English textbooks, Indonesian language textbooks
17	Person17	Mathematics textbooks, Indonesian language textbooks, English textbooks
18	Person18	Chemistry textbooks, English textbooks, ICT textbooks
19	Person19	Physical Education textbooks, ICT textbooks
20	Person20	Religion textbooks, ICT textbooks

The transaction data is then represented as a list table, as shown in Table 3:

Table 3. Data Representation of Book Borrowing

Visitors	Borrowed book
1	ICT textbooks
1	PKN textbooks
1	Physical Education textbooks
2	PKN textbooks
2	Mathematics textbooks
3	Chemistry textbooks
3	Physical Education textbooks
4	Religion textbooks
4	PKN textbooks
4	ICT textbooks
5	Chemistry textbooks
5	ICT textbooks
5	Religion textbooks
6	ICT textbooks
7	Chemistry textbooks
7	ICT textbooks
7	Mathematics textbooks
8	ICT textbooks
8	Religion textbooks
8	Physical Education textbooks
9	Chemistry textbooks
9	ICT textbooks
9	Religion textbooks
10	ICT textbooks
10	Religion textbooks
11	Physical Education textbooks
12	English textbooks
12	Mathematics textbooks
12	Chemistry textbooks
13	Physical Education textbooks
14	Mathematics textbooks
15	Indonesian language textbooks
16	Mathematics textbooks
16	English textbooks
16	Indonesian language textbooks
17	Mathematics textbooks
17	Indonesian language textbooks
17	English textbooks
18	Chemistry textbooks
18	English textbooks
18	ICT textbooks
19	ICT textbooks
20	Religion textbooks
20	ICT textbooks

The minimum support value in this test is set at 5%, and the minimum confidence is set at 10%. Table 4 displays the 2-itemset candidates from the book lending data who meet the minimum support and minimum confidence requirements outlined in Table 3.

Table 4. Candidate 2-itemset

Combination	Amount
PKN textbooks, Mathematics textbooks	1
Chemistry textbooks, Physical Education textbooks	1
ICT textbooks, Religion textbooks	2
Mathematics textbooks, Indonesian language textbooks	1
Physical Education textbooks, ICT textbooks	1

From table 4 it can be concluded that if the value of the threshold is set where $\Phi = 2$, then: $F2 = \{(PKN \text{ textbooks}, Mathematics \text{ textbooks}), (Chemistry \text{ textbooks}, Physical \text{ Education \text{ textbooks}}), (ICT \text{ textbooks}, Religion \text{ textbooks}), (Mathematics \text{ textbooks}, Indonesian \text{ language \text{ textbooks}}), (Physical \text{ Education \text{ textbooks}}, ICT \text{ textbooks})\}$.

Table 5. Candidate 3-itemset

Combination	Amount
Religion textbooks, ICT textbooks, Physical Education textbooks	3
ICT textbooks, Chemistry textbooks, Religion textbooks	3
ICT textbooks, Religion textbooks, Physical Education textbooks	3
Chemistry textbooks, ICT textbooks, Religion textbooks	3

The combination of itemset in $F2$ can be combined into a candidate 3-itemset. The itemsets from $F2$ that can be combined are itemsets that have similarities in the first $k-1$ items. The 3-itemset candidates that can be formed from $F2$ are shown in table 7. Thus $F3 = \{(Religion \text{ textbooks}, ICT \text{ textbooks}, Physical \text{ Education \text{ textbooks}}), (ICT \text{ textbooks}, Chemistry \text{ textbooks}, Religion \text{ textbooks}), (ICT \text{ textbooks}, Religion \text{ textbooks}, Physical \text{ Education \text{ textbooks}}), (Chemistry \text{ textbooks}, ICT \text{ textbooks}, Religion \text{ textbooks})\}$.

After you've discovered all of the high-frequency patterns, look for association rules that meet the minimum requirements for confidence by calculating the confidence of association rules $A \rightarrow B$. The $A \rightarrow B$ rule's confidence value is calculated using a formula (2). If you borrow a religion book and an ICT book, you will borrow a physical education book. To get the confidence value of this rule, enter the total value of loan transactions containing religious books and ICT books divided by the total loan transactions containing religious books and ICT books. The reference of this formula is in Table 5 with 3-itemset and Table 4 with 2-itemset. Meanwhile, to get a support value, if you borrow a Math book, you will borrow a science book. The results of the calculation of the total transactions containing the borrowing of Mathematics and Science books divided by the total loan transactions. Association rules that are formed based on a minimum support of 5% and a minimum confidence of 10% can be seen in Table 6 below.

Table 6. Association Rules of 2-itemset

Rules	Support	Confidence
If you borrow a math book, you will borrow a chemistry book	7%	50%
If you borrow a religion book, you will borrow a physical education book	8%	100%
If you borrow an ICT book, you will borrow a Religion book	11%	75%
If you borrow ICT books, you will borrow Chemistry books	11%	37%
If you borrow an ICT book, you will borrow a physical education book	11%	37%

Rules	Support	Confidence
If you borrow a physical education book, you will borrow a PKN book	7%	50%
If you borrow a chemistry book, you will borrow a religion book	8%	33%
If you borrow Chemistry books, you will borrow ICT books	11%	55%
If you borrow English books, you will borrow ICT books	8%	50%

Table 6 shows that the combination of the Religion book and the Physical Education book, with a total support x confidence of 80 percent, has the highest value of support and confidence with a 2-itemset reference. In the meantime, Table 7 shows the association rules with 3-itemsets that match the minimum support and confidence criteria.

Table 7. Association Rules of 3-itemset

Rules	Support	Confidence
If you borrow a religion book and an ICT book, you will borrow a physical education book	6%	100%
If you borrow ICT books and Chemistry books, you will borrow Religion books	6%	100%
If you borrow ICT books and Religion books, you will borrow Physical Health books	6%	100%
If you borrow Chemistry books and ICT books, you will borrow Religion books	6%	100%

Table 7 yielded four association rules that meet the minimal support and confidence criteria. It can be explained by the fact that all combinations of borrowing religious books, ICT books, and physical education books from a vocational high school's library have the same high likelihood of borrowing at the same time.

4. Conclusion

Based on the findings and discussions, it is possible to conclude that using the association method can aid in determining the pattern of library visitors in one of the majoring schools, which is currently done manually. The Apriori method, with its association rules, can provide effective information to describe the process related to the pattern of library visitors, and the resulting rules can be used as a reference in the library's acquisition of collection books. Furthermore, the variables chosen have a significant impact on the resulting rule or knowledge. This research can be expanded in the future by employing a hybrid method or data mining patterns other than association rules.

References

- [1] O. Manurung and P. S. Hasugian, "Analisa Algoritma Apriori Untuk Peminjaman Buku Pada Perpustakaan SMA 1 Silima Pungga- Pungga Paronggil," *Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 4, no. 1, pp. 1–8, 2019.
- [2] S. Saefudin and D. Fernando, "Penerapan Data Mining Rekomendasi Buku Menggunakan Algoritma Apriori," *JSiI (Jurnal Sist. Informasi)*, vol. 7, no. 1, p. 50, 2020.
- [3] E. Srikanti, R. F. Yansi, Norhavina, I. Permana, and F. N. Salisah, "Penerapan Algoritma Apriori untuk Mencari Aturan Asosiasi pada Data Peminjaman Buku di Perpustakaan," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 4, no. 1, pp. 77–80, 2018.
- [4] I. Afifah, "Implementation of environmental care character for elementary school students through verticultural culture techniques," *J. Pengabdi. dan Pemberdaya. Masy. Indones.*, vol. 1, no. 2, 2021.
- [5] M. Widayastuti, A. G. Fepdiani Simanjuntak, D. Hartama, A. P. Windarto, and A. Wanto, "Classification Model C.45 on Determining the Quality of Customer Service in Bank BTN Pematangsiantar Branch," *J. Phys. Conf. Ser.*, vol. 1255, no. 012002, pp. 1–6, 2019.
- [6] B. Supriyadi, A. P. Windarto, T. Soemartono, and Mungad, "Classification of natural disaster prone areas in Indonesia using K-means," *Int. J. Grid Distrib. Comput.*, vol. 11, no. 8, pp. 87–98, 2018.
- [7] A. P. Windarto, U. Indriani, M. R. Raharjo, and L. S. Dewi, "Bagian 1: Kombinasi Metode Klustering dan Klasifikasi (Kasus Pandemi Covid-19 di Indonesia)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 855, 2020.
- [8] A. P. Windarto, J. Na, and A. Wanto, "Bagian 2 : Model Arsitektur Neural Network dengan Kombinasi K-Medoids dan Backpropagation pada kasus Pandemi COVID-19 di Indonesia," vol. 4, pp. 1175–1180, 2020.
- [9] A. Waluyo, H. Jatnika, M. R. S. Permatasari, T. Tuslaela, I. Purnamasari, and A. P. Windarto, "Data Mining

- Optimization uses C4.5 Classification and Particle Swarm Optimization (PSO) in the location selection of Student Boardinghouses," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, pp. 1–9, 2020.
- [10] W. Katrina, H. J. Damanik, F. Parhusip, D. Hartama, A. P. Windarto, and A. Wanto, "C.45 Classification Rules Model for Determining Students Level of Understanding of the Subject," *J. Phys. Conf. Ser.*, vol. 1255, no. 012005, pp. 1–7, 2019.
- [11] N. Wandini, R. A. Hendrawan, and A. Mukhlason, "Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalian Association Rule Menggunakan Algoritma Apriori," *J. Tek. ITS*, vol. 1, pp. 1–5, 2012.
- [12] M. Bhargava and A. Selwal, "Association Rule mining using Apriori Algorithm: A Review," *Int. J. Adv. Res. Comput. Sci.*, vol. 4, no. 2, p. 2013, 2013.
- [13] R. Ruswati, A. I. Gufroni, and R. Rianto, "Associative Analysis Data Mining Pattern Against Traffic Accidents Using Apriori Algorithm," *Sci. J. Informatics*, vol. 5, no. 2, pp. 91–104, 2018.
- [14] I. N. Jha and S. Borah, "Efficient Association Rule Mining Using Improved Apriori Algorithm," *Int. J. Sci. Eng. Res.*, vol. 3, no. 11, pp. 1–4, 2012.
- [15] L. Hanguang and N. Yu, "Intrusion Detection Technology Research Based on Apriori Algorithm," *Phys. Procedia*, vol. 24, pp. 1615–1620, 2012.
- [16] A. Lewis, M. Zarlis, and Z. Situmorang, "Penerapan Data Mining Menggunakan Task Market Basket Analysis Pada Transaksi Penjualan Barang di AB Mart dengan Algoritma," *J. Media Inform. Budidarma*, vol. 5, no. April, pp. 676–681, 2021.
- [17] I. Djamarudin and A. Nursikuwagus, "Analisis Pola Pembelian Konsumen Pada Transaksi Penjualan Menggunakan Algoritma Apriori," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 2, p. 671, 2017.
- [18] R. R. Waliyansyah, M. Novita, and N. D. Saputro, "Utilization of data mining for placement of books in the library using the Apriori method," *J. Phys. Conf. Ser.*, vol. 1869, no. 1, 2021.
- [19] D. R. Prehanto, A. D. Indriyanti, G. S. Permadi, T. Z. Vitadiar, and F. D. Jayanti, "Library book modeling data using the association rule method with apriori algorithm in determining book placement and analysis of book loans," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 5, pp. 1244–1250, 2019.