

# Implementation of the Apriori Algorithm for Library Book Borrowing Patterns

Ida Farida, Yunita Sartika Sari, Fauzi Nur Iman

Universitas Mercu Buana, Jl. South Meruya No.1

Email address: dae.farida@mercubuana.ac.id, yunita.sartika@mercubuana.ac.id, fauzi@mercubuana.ac.id

**Abstract**—Because libraries have a large collection of books from a wide range of sciences, data mining is necessary in order to explore added value in the form of knowledge that has not been known manually from a data collection. Association rule techniques and a priori algorithms are used to create candidate combinations of items based on certain rules. Then, the combination of items meets the minimum support requirements, which is then used to create rules that meet the minimum requirements. *percaya diri*. Information in the form of book borrowing patterns and suggestions for libraries to arrange book layouts according to the itemsets formed are the outcomes of this research.

**Keywords**— Borrowing; Data Mining; Association Rules; Apriori Algorithm.

## I. INTRODUCTION

The demand for quick, accurate, and pertinent information is growing along with the advancement of information technology. In order for information to play a significant role in the current and future growth of society, it is imperative that it be accurate in all aspects of daily life. However, sometimes there isn't a sufficient balance between the strong demand for information and the acquisition of new knowledge. Frequently, more work is required to extract this information from data including very large populations.

It is not sufficient to rely solely on operational data when using existing data in information systems to support decision-making; data analysis is needed to fully realize the potential of current data. Because decision makers attempt to use the data warehouse they currently have, a new field of study known as data mining has emerged to address the challenge of identifying significant and intriguing patterns or information hidden in massive amounts of data. It is anticipated that the application of data mining will reveal information that was previously concealed in the data warehouse, turning it into useful knowledge.

The term "data mining" refers to the process of searching through enormous databases for relevant and valuable information. The Association guideline—a guideline for identifying high frequency connections between a set of itemsets—is the most crucial component of data mining approaches. The AIS Algorithm, DHP Algorithm, Partition Algorithm, and Apriori Algorithm are a few of the algorithms that are incorporated in the association rules. Nevertheless, one of these algorithms—the a priori algorithm—is frequently employed in data mining to examine lending trends. An example of an a priori algorithm is seen in data mining, where rules that indicate relationships between several qualities are

referred to as affinity analysis or market basket analysis. Where this system can later work by analyzing and finding patterns associated with borrowed books. This technique is usually called association analysis or association rules, where the study is concerned with 'what and what'.

This study is named "Implementation of the Apriori Algorithm in Data Mining for Book Lending Patterns" based on the background information provided above.

### 1.1. Literature review

#### 1.1.1. Definition of Association Rule

In order to support the recommendation system by identifying similarities between items in the transactions that occur, association rules is an approach that looks for patterns that frequently arise among numerous transactions, where each transaction consists of multiple items. As stated by Fadlina (2014)

#### 1.1.2. Understanding Apriori Algorithms

Agrawal and Srikant introduced the a priori technique in 1994 as a fundamental method for identifying frequent itemsets for Boolean association rules. One of the communication rules used in data mining is the Apriori algorithm. Market basket analysis and affinity analysis are terms used to describe rules that express the relationship between multiple variables. A data mining technique used to determine rules for combining elements is relational analysis, also known as relational rule mining. High frequency pattern analysis, sometimes known as frequent pattern mining, is one of the stages of association analysis that has drawn the interest of numerous academics in an effort to create effective algorithms. Two benchmarks—support and confidence—can be used to assess the significance of an association. Support (supporting value) is a percentage a combination of these items in the database, while confidence (certainty value) is the strength of the relationship between items in the association rules.

The a priori algorithm is broken down into multiple phases known as passes or narratives:

#### 1. Formation of candidate itemsets.

The candidate k-itemset is formed from a combination of (k-1)- itemsets obtained from the previous iteration. An alternative to the previous algorithm is to prune k-element array candidates whose subsets contain k-1 elements that are not included in the highest frequency pattern of length k-1.

#### 2. Calculation of support for each k-itemset candidate.

The number of transactions that contain every item in a candidate k-itemset is counted when scanning the database to

determine the support of each candidate k-itemset. This is another feature of the a priori algorithm, which calls for finding the longest k-itemset over the entire database.

3. Establish a high frequency pattern

PHigh frequency patterns containing k items or kitemsets are determined from k-itemset candidates whose support is greater than the minimum support.

4. If no new high frequency pattern is obtained then the entire process is stopped.

To better understand the Apriori algorithm process, the following will provide an illustration of the use of the Apriori algorithm. Using the database in assuming minimum support is 2 transactions.

II. RESEARCH METHODOLOGY

The research method used in this research is a qualitative approach. The qualitative method is called a new method, because it has not been popular for a long time, it is called the postpositivism method. Qualitative research method is a research method based on the philosophy of postpositivism, used to research natural objects (initially as experiments) where the researcher is the key instrument, data collection techniques are carried out in a triangulated (combined) manner, the analysis is inductive/qualitative and the research results are qualitative. emphasizes eating rather than generalization (Sugiyono, 2014;7). Several characteristics of research using qualitative methods are as follows:

1. Qualitative research has a general design, is flexible, develops and emerges in the research process.
2. Qualitative research is carried out to find interactive relationship patterns, describe complex realities and gain an understanding of meaning.
3. Qualitative research uses data collection techniques in the form of participant observation, in-depth interviews, documentation and triangulation.
4. Qualitative research research instruments in the form of notebooks, tape recorders, cameras, handycams and others.
5. Qualitative research data used in the form of personal documents, field notes, respondents' words and actions, documents and others (Sugiyono, 2014; 14).

This research is qualitative in nature by analyzing book borrowing patterns in libraries using data mining methods and a priori algorithm calculations. The a priori algorithm itself has minimum support and confidence values.

The stages that will be taken are:

1. Preparation. This stage is the subject (population) stage of the library. The object taken is a book. Limitations and developing a research plan.
2. Literature review. In the literature review, a review of library slims and literature studies regarding book borrowing and related matters was carried out.
3. Data collection. Data collection was carried out by means of interviews with library staff, observations and documents.
4. Data mining processing. The data collected is processed according to stages Knowledge Discovery in Databases (KDD).

5. Results and Discussion. The discussion at this stage explains the results of the data mining process which was carried out using an a priori algorithm.

6. Conclusions and suggestions. Draw conclusions from the research results and make suggestions for the library to improve.

III. RESULTS AND DISCUSSION

The following is a classification of books in the library:

TABLE 1 Book Classification

Classification	Title
100	Statistic
200	Algorithm and Structure Data
300	Software Engineering
400	Data Science
500	Artificial Intelligence
600	Architecture Enterprise
700	Audit Information System
800	Network
900	Machine Learning

The following are the steps for implementing the apriori algorithm:

1. Itemset formation

The process of forming C1 or what is called 1 itemset with minimum amount of support = 40% with the following formula:

$$\text{Support}(A) = \frac{\sum \text{Number of transactions containing } A}{\sum \text{Total transactions}} \times 100\%$$

2. Establishment of Association Rules

After a high frequency pattern is found, then we look for an association rule that meets the minimum requirements for confidence by calculating the confidence of the associative rule A→B. Minimum Confidence = 70%. The confidence value is calculated using the following formula:

$$\text{Confidence}(A|B) = \frac{\sum \text{Number of transactions containing } A \text{ and } B}{\sum \text{Total transactions containing } A} \times 100$$

TABLE 2 Association rules

Association Rules	Support	Confidence
If you borrow a book type 100 then borrow 400 book types	42.5%	77.27%
If you borrow book type 100 then borrow book type 200	42.5%	77.27%
If you borrow book type 200 then borrow book type 400	42.5%	77.27%
If you borrow book type 300 then borrow book type 500	40%	72.72%
If you borrow 600 types of books then borrow 700 types of books	55%	78.57%
If you borrow a 600 type of book then borrow a 800 type book	55%	84.62%
If you borrow 300 types of books then borrow 700 types of books	47.5%	79.17%
If you borrow a 500 type book then borrow a 900 type book	40%	80%
If you borrow a 400 type of book then borrow a 900 type of book	45%	75%

Based on Table 2, the types of books or book classifications that are often borrowed by library members are 100, 200, 300, and 400.

### 3. RapidMiner Implementation

In this section we will discuss the RapidMiner process in finding book borrowing patterns as support for the results of the a priori algorithm discussed previously. The data that will be used is book borrowing data in Excel form. The result of this data mining is information in the form of book borrowing patterns/rules that support research. Below are the work steps for implementing data mining using RapidMiner, which are as follows:

#### 3.1. Creating the Tabular Format

At this stage, what is done is to prepare the data to be processed, namely data on library book borrowing transactions in 2023. To create tabular format data using Microsoft Excel, where the data used can be seen in the attachment. Then import the Microsoft Excel table that has been created into the Local Repository.

#### 3.2. Arrangement of Apriori Algorithm Operators

At this stage, what is done is drag and drop the 2023 data table into the process. So that database operators appear in the main process. The Numerical to Binominal operator is needed to convert different attribute values in the borrowing table into binominal form. Then connect the data table 2023 1 with Numerical to Binominal operators. This process will convert the values from the lending transaction table into Binominal Attributes. Next, connect the Numerical to Binominal operator with the W-Apriori operator and fill in the W-Apriori parameter C with a minimum confidence value of 70% or 0.7. after that connect W-Apriori to the results. So, the whole thing is like in the image below:



Figure 1. Arrangement of Apriori Algorithm Operators

#### 3.3 Apriori Algorithm Results

This stage is the final stage of the Rapidminer data mining process, after all operators are connected, then click the play icon, the F11 button, a new Apriori tab will appear, which contains a description of all itemsets that meet the W-Apriori parameters, as in the following image:



Figure 2 Results of the Apriori Algorithm for Book Borrowing Data in 2023

From the results of data mining analysis with RapidMiner, the number of rules produced in 2023 is 10 rules.

From the results of data mining analysis with RapidMiner, the number of rules produced from 2023 is 10 rules. Based on the results from RapidMiner, one rule can be taken to be used as information that supports book layout settings using a data mining process obtained from an a priori algorithm to make finding books easier. Of course, the one with the highest Support and Confidance values, namely, if you borrow a 100 type of book, borrow a 400 type book with a support value of 55% and a confidence value of 84.62%.

## IV. CONCLUSION

This research comes to the following conclusion based on the conducted research: 10 rules are formed annually based on the outcomes of data mining calculations made using an a priori algorithm, book borrowing transaction data in libraries with a minimum support limit of 40% and minimum confidence of 70%. With a support value of 55% and a confidence value of 84.62%, one of the rules that is generated states that if you borrow a book of type 100, you must also borrow a book of type 400. This information can help the library make recommendations for how to arrange the bookshelves.

## REFERENCES

- [1.] Nuraisana, "Decision Support System For Determining Recipients Of The Hope Family Program (PKH) Aid Using The Analytical Hierarchy Process Method," J. Mantik Penusa, vol. 3, no. 1, pp. 129–137, 2019.
- [2.] Jadianan Parhusip, "Application of the Analytical Hierarchy Process (AHP) Method in the Design of a Decision Support System for Selection of Non-Cash Food Assistance (BPNT) Recipient Candidates in Palangka Raya City," J. Teknol. Inf. J. Science and Applications. Bid. Tech. Inform., vol. 13, no. 2, pp. 18–29, 2019, doi: 10.47111/jti.v13i2.251.
- [3.] A. Sopandi Solihin, D. Gustian, F. Sembiring, M. Muslih, and N. Arianti Destria, "Decision Support System For Cash Social Assistance Recipients Using The Technique For Order Preference Method By Similarity To Ideal Solution," J. Rekayasa Technol. Nusa Putra, vol. 8, no. 1, pp. 1–9, 2021.
- [4.] E. Martyani and S. Santoso, "Decision Support System for Selection of Hanging Out Places Using the Analytical Hierarchy Process Method," J. V-Tech (Vision Technol., vol. 2, no. 1, pp. 14–23, 2019, doi: 10.35141/jvt.v2i1.455.
- [5.] B. Suprpto and A. Sujoni, "Decision Support System For Prospective Recipients Of The Hope Family Program (PKH) Aid Using The Analytical Hierarchy Process (Ahp) Method," J. Inf. And Comput., vol. 7, no. 2, pp. 47–56, 2019.

- [6.] MA Fauzi, "Decision Support Systems," 2016.
- [7.] A. R Laisouw, S. Lutfi, and F. Tempola, "Decision Support System for Providing Family Hope Program (Pkh) Assistance to Poor People in Ternate City Using the AHP Method," *JIKO (Journal of Information and Computers)*, vol. 2, no. 1, pp. 34–60, 2019, doi: 10.33387/jiko.v2i1.973.
- [8.] Ministry of Social Affairs, "Get to Know the Non-Cash Food Assistance Program Closer," Ministry of Social Affairs.go.id, 2021, [Online]. Available: <https://kemensos.go.id/uploads/topics/15664651387355.pdf>.
- [9.] N. Serrano, G. Gallardo, and J. Hernantes, "Infrastructure as a Service and Cloud Technologies," *IEEE Softw.*, vol. 32, pp. 30–36, 2015.
- [10.] Beer, M.I., & Hassan, M.F. (2017). Adaptive security architecture for protecting RESTful web services in enterprise computing environments. *Service Oriented Computing and Applications*, 12(2), 111–121. doi:10.1007/s11761-017-0221-1
- [11.] Ardiansyah, HI (2017). Development of the REST API for the Indoexplore.id Tourism Marketplace Information System.
- [12.] Kusuma, I. (2017). Quick Way to Create API Documentation with APIDOCjs. Retrieved from <https://indrakusuma.web.id/2017/11/07/Cara-Mudah-Mabuat-Dokumentasi-APIdengan-apidocjs>
- [13.] Perkasa, MI, & Setiawan, EB (2018). Development of Community Data Web Services Using REST API with Access Token, X(1), 19–26.
- [14.] Pratama, YEB (2017). Implementation of REST Web Service for Mobile Based Reseller Orders, 16–37
- [15.] Rulloh et al. (2017). Implementation of the REST API in the Android-Based Fitness Guide Application. *Teknikom : Vol. 1 No. 2, 1(2)*, 85–89