Application of Apriori Algorithm in One State University's Library Book Borrower Records for Efficient Library Shelving

Jerson D. Cecilio*, Gene Marck B. Catedrilla, Jonardo R. Asor
Laguna State Polytechnic University, Los Baños, Laguna, Philippines.

* Corresponding author.; Email: ceciliojd@gmail.com (J.D.C.)
Manuscript submitted March 16, 2023; revised June 20, 2023; accepted September 10, 2023.
doi: 10.17706/jsw.18.4.172-184

Abstract: Association rule mining is a technique for discovering patterns, associations, and relationships in large data sets or in a variety of databases such as relational, transactional, and other archives or repositories. It is significantly used in libraries to provide a data-driven approach in management of books, reports, theses, manuscripts, and other literature. This article was conducted to examine book borrowing patterns using the Apriori algorithm for efficient book shelving to assist Laguna State Polytechnic University's library in effectively managing resources, and services. The three year book borrower records of Laguna State Polytechnic University were used as the dataset in this article. Hence, rapidminer was used as a data mining tool in implementing apriori algorithm in the latter and for association discovery. Through the use of apriori algorithm, it was discovered that histories, and consumer preferences books give a high relationship rating therefore, the library may consider rearranging the shelves and place the latter closer with each other. Moreover, all the combinations of two item sets or books with a confidence value greater than 60% as shown in this article were strongly advised to be placed or grouped together for a more effective shelving and efficient searching of books.

Key words: library, books, patrons, data mining, LSPU, Los Baños, association, apriori, library shelving, books placement

1. Introduction

Libraries are valuable since they play a significant part in society by providing all community members, regardless of status or location, with access to information and resources. Hence, most libraries used a library system to record and track all borrowed books in a more efficient manner. Due to its functions, it generates a large transactional database. Different innovative approach is applied to library management to enhance its service which includes a data-driven approach in book management. Commercial businesses, in specific, are interested in determining association rules that identify frequently purchased items. M. Kavitha and S. Subbaiah, discovered and discussed in their study that the association rule shows a promising result in determining relationships with two entities. The findings of market basket analysis are able to suggest package deals such as special offerings, item combinations, to design a better store display, and to give understanding into other brand recognition [1]. Allowing to conduct a study on the association of books to identify commonalities or associated books in the book borrower's records.

S. J. Cunningham and E. Frank stated in their article that the apriori algorithm is employed to identify classifications that exist in transactional data of borrowed books in University Library. This data helps

visitors find extra parts of gathering which may have records relating to visitor's needed data, as well as to determine the physical layout of a library [2]. While D. R. Prehanto, *et al*. [2] stated that the presence of everyday transactions in the library can lead to an increasing number of books borrowed. An apriori algorithm with association rule method as model classifier on book borrower's records at the Asy'ari Hasyim University Library used to make book recommendations and evaluate borrowing of books as a result of its application and define the relationship between one book to another. The calculation process is then done based on the resulting modeling, so that it is known that the same book is borrowed every day, with the support and confidence value percentage obtained from 600 data of books borrower's record, with support averaging 6% and confidence values averaging 67-100% [3].

The Laguna State Polytechnic University Los Baños (LSPU-LB) Library uses an information system, namely LSPU Library Borrowing System. The system aims to manage all the functions and transactions of the library. It assists the librarians in maintaining the database of books as well as books borrowed by patrons and their due dates contributing to a transaction-oriented database. The association data mining works best to the transactional database to identify the frequent patterns of item sets. In this article, association data mining is used in discovering common patterns, associations, connections, correlations, and frameworks in a library's database. The primary application of association rule mining is the evaluation of transaction history. Hence, this article aims to determine book borrowing patterns to identify association of books using the apriori algorithm to provide librarians recommendations for effective book shelving.

2. Related Works

Association rule mining is a technique for identifying frequent item sets, relationships, connections in sets of data collected in different databases, such as operational and transaction-oriented, relational, as well as other forms of database archives. The primary objective of association rule mining is to identify the rules that permit us to assess the presence of a particular item relying on the occurrences of the other items in the data [4].

Apriori algorithm was used to analyze supermarket datasets, the manager can learn which items are frequently bought and which items are bought together by the customer. It will be used to make decisions and promote earnings. They can also offer special promotions and product combinations. His one example was Aisle shopping, which increased sales by arranging items on the same shelf and made it easier for customers to choose items. Street vegetable vendors sell potatoes and onions together. Observing the behavior of the customers who bought the items together. Sales increased substantially. Best market basket analysis which assists customers in purchasing their items with greater ease, thereby increasing market sales [5].

In [6], stated that traditional libraries will have mountains of transaction records of borrowed books that will only serve as a repository, and book spacing, which induces patrons to take more time to see different kinds of books, is an issue need to be solved. The study employed Apriori algorithm with an association rule method, and it will result in generating association rules for book placement recommendations. While in [7], authors realized that all association mining techniques achieve their objectives perfectly, this technique has unique characteristics and specifications that demonstrate its accuracy, proficiency, and preference. It can be accomplished by using Association data mining in the library database. Also, in [8], authors determined the sequence of books arrangement inside the library by integrating the Apriori calculations with the Association data mining and analyzing the results of patrons' proclivity to borrow books relying on the combination of two itemsets. Making it more convenient for library patrons to take books and program system may place near the location of the library materials.

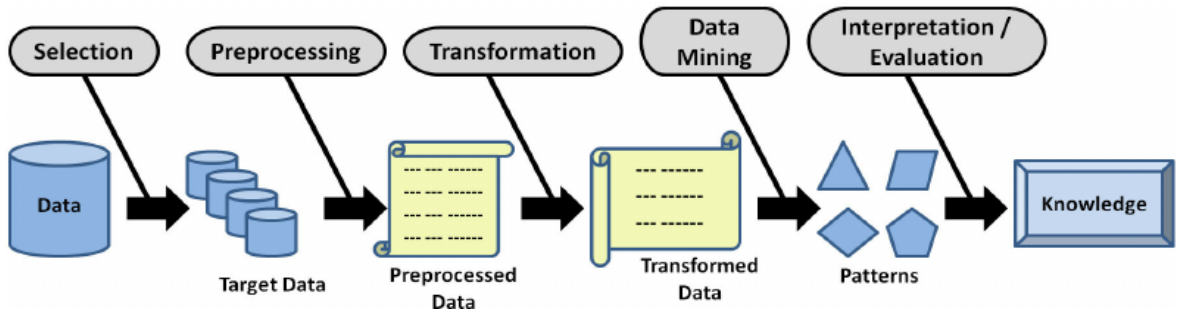
In [9], investigated the correlation between borrowed books in Dewan Bahasa dan Pustaka's (DBP)

libraries in order to determine common borrowing patterns in Brunei libraries. Association rule mining was used in the analysis, which is a method to identify frequent patterns, correlations, or connections in datasets from various types of databases, including relational, transactional, and other repositories. The Apriori Algorithm, a common technique for frequent itemset searches in association rule mining, was used to transform data from book borrowing transactions collected by the library into useful information. According to the study findings, Brunei patrons typically request Malay-titled books and academic past exam papers with a support value of about 0.5% and as a result, the investigation gained an average confidence value of around 70%.

Overall, the Apriori algorithm is a strong instrument for examining customer purchasing patterns and can be used to generate relevant book recommendations for customers as well as textbook recommendations for students depending on their fields of study. The use of methodologies for data mining in university libraries, such as the Apriori algorithm, can assist libraries in grouping similar borrowing behaviors, improving preventative measures, and timely detection of high population density, limiting fatalities.

3. Methodology

Knowledge Discovery in Database is a method of identifying hidden patterns in data, rules, as well as other components inherent in large volumes of data, as shown in Fig. 1, [10].



Source: https://www.researchgate.net/publication/274425359/figure/fig1/AS:294725178413056@1447279366136/The-Knowledge-Discovery-in-Databases-KDD-process.png

Fig. 1. Knowledge discovery in database.

3.1. Data Collection

The datasets collected are the transaction history of borrowed books from LSPU Los Baños from 2019 to 2021. It was recorded using the Library System, a computerized system that saves all library activities and transactions. The collected data is suitable for mining association rules.

3.2. Data Preprocessing

In the analysis, the data extracted from the system needs some conversions, modifications, and groupings relevant to the Apriori algorithm association rule method without compromising the data's integrity.

Table 1. Attributes of book borrower's records data set

Attributes	Description
Date	Date of transactions
Title	Title of Borrowed Book
BibType	Category of Materials
PatronType	Group of Borrowers

Belongs to data preprocessing, it identifies the required attributes necessary for Apriori algorithm calculations.

3.3. Model Development

3.3.1. Market basket analysis

Market-Basket Analysis is identified as analysis of affinity, and is a data mining technique which is applied in a variety of areas such as education, nuclear science, marketing, and bioinformatics [11]. In most libraries, data is collected by recording borrowed books using the Library Management System. It consists of a large number of transaction histories of borrowed books. Each transaction contains book borrowing, lending, and due dates in one transaction. Using Market Basket Analysis, it will identify commonalities or associated books in the book-borrowing records.

3.3.2. Association rules

Data analysis for frequently occurring patterns and classifying the most significant associations using support and confidence criteria yields association rules. The frequency by which the items appear in the data is indicated by the level of support. Confidence represents the number of times the if-then statements are proven to be true. Lift is a third metric for comparing confidence to expected confidence or the couple of times an if-then statement is expected to be true. To create results on association rule, two or more itemsets are employed [12].

For this association rule mining, rules are extracted from a given transactional database. This algorithm primarily employs three (3) measures: support, confidence, and lift.

Support is a metric that indicates how regularly an individual item seems in a given dataset. It is measured by:

$$Support1 = \frac{NumberofTransactionsofOneItemSet}{TotalNumberofTransactions}$$

or

$$Support2 = \frac{NumberofTransactionsWithBothAandB}{TotalNumberofTransactions}$$

In the context of association rule mining, confidence is a statistic that measures how frequently a specific relationship or association between two itemsets occurs in a dataset. It is commonly stated as a percentage or as a ratio. More specifically, the confidence of two itemsets A and B signifies the probability that itemset B will also be present in a transaction or record when itemset A is present. Confidence implies how frequently the rule has been proven true. It is measured by:

$$Confidence = \frac{NumberofTransactionsWithBothAandB}{TotalNumberofTransactionsWithA}$$

$$Confidence = \frac{P\ (A\ \Omega\ B\ )}{P\ (A)}$$

Lift assists researchers in determining whether association rules are useful for decision-making. A lift value larger than one shows a significant association, making it an effective indicator for activities such as market basket analysis and recommendation systems. Lift is measured by these two parameters. Confidence and Expected Confidence Values. It is measured by:

$$Lift = \frac{Confidence}{ExpectedConfidence}$$

$$Lift = \frac{P\ (\ A\ \Omega\ B\ )}{P\ (A).P\ (B)}$$

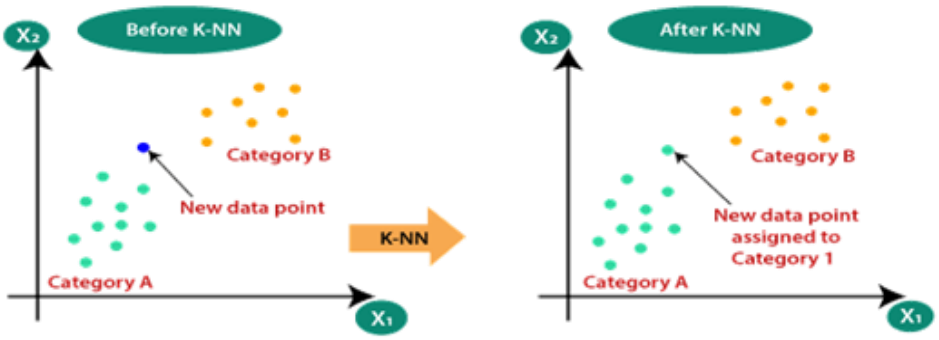
3.3.3. FP Growth algorithm

For this association rule mining, rules are extracted from a given transactional database. This algorithm primarily employs three (3) measures: support, confidence, and lift. Without the need for candidate generation, a frequent pattern is generated. The data is represented by the Frequent Pattern - growth algorithm as a tree known as a FP tree. This tree structures will keep the connection in between item sets. The data is fragmented by a common item which is referred to as a fragment pattern. These fragmented pattern item sets were evaluated. The analysis of frequent item sets is then extensively discussed using this method [13].

3.3.4. Apriori algorithm

The Apriori algorithm is a data mining method often used to discover frequently occurring item sets and relevant association rules. This module defines association techniques and Apriori calculations, as well as how to use an Apriori algorithm. It is intended to operate on a database containing transaction information. The Apriori algorithm discusses each item as an item set and counts the support based on their frequency found in the dataset before capturing those with support equal to or greater than the minimum support [14].

3.3.5. K-Nearest Neighbors Algorithm (KNN)



https://static.javatpoint.com/tutorial/machine-learning/images/k-nearest-neighbor-algorithm-for-machine-learning2.png

Fig. 2. K-Nearest neighbors algorithm.

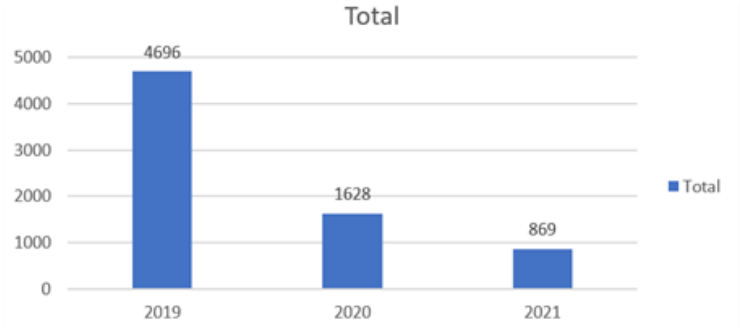
Because it's non-parametric technique, it makes no assumptions about the fundamental data. A model uses distance to classify or predicts the grouping of a single data point. Consider the A and B categories, as well as the new data point, x1. This data point will fall into which of the following categories? To address these problems, a K-Nearest Neighbors method is needed. By utilizing the KNN method, we were able to identify quickly the categories on a specific dataset [16]. The KNN algorithm's aim is to determine new classified elements from the given set.

4. Results and Discussion

The specific books borrowed by the patrons per day. The restatement of borrower's records groups the most frequently borrowed books in order to analyze the pattern of borrowing specific book from given history data, as described in Table II. It also provides the number of borrowed books, that will used to perform the Apriori algorithm.

Table 2. Sample book borrower's records

Date	Book Titles	Types	Patron Types
1/9/2019	Applied nutrition and food technology	Books	Faculty Reg
1/9/2019	Food technology	Books	Faculty Reg
1/9/2019	Practical food preservation and processing: a practical guide for small businesses, hobbyists and enthusiasts	Books	Faculty Reg
1/9/2019	Practical food preservation and processing: a practical guide for small businesses, hobbyists and enthusiasts	Books	Faculty Reg
1/10/2019	Introduction to criminology and psychology of crime	Books	Faculty Reg
1/10/2019	Introduction to criminology and psychology of crime	Books	Faculty Reg
1/10/2019	Criminology 101: a reference manual for introduction to criminology and psychology of crimes	Books	Faculty Reg
1/11/2019	Introduction to criminology and psychology of crime	Books	Faculty Reg
1/11/2019	Police administration, organization and planning: an instructional book	Books	Faculty Reg
1/11/2019	Police administration, organization and planning: an instructional book	Books	Faculty Reg
1/14/2019	Introduction to humanities arts for fine living	Books	BSED
1/14/2019	Introduction to art appreciation a textbook in humanities	Books	BSED
1/14/2019	Introduction to art appreciation: a textbook in humanities	Books	BSED



Source: Laguna State Polytechnic University Library System

Fig. 3. Total borrower's record per year.

The borrower's record or transaction history over a year. There were 4,696 transactions recorded in 2019, 1,628 transactions recorded in 2020, and 869 transactions recorded in 2021. The statistics indicate that the covid-19 disease outbreak had a massive effect in library, as represent in Fig. 3.

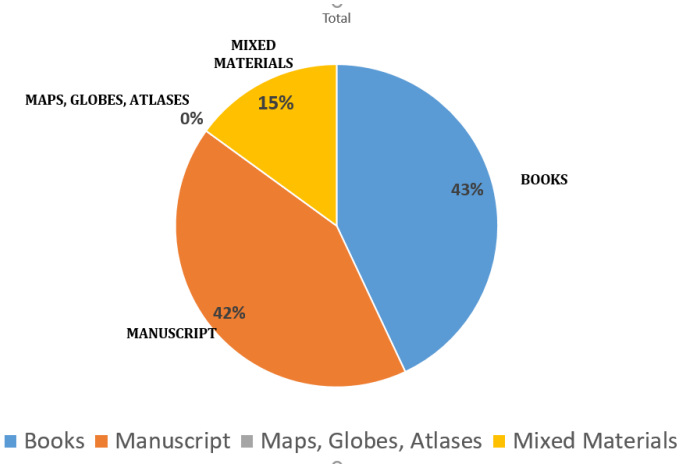


Fig. 4. Percentage of borrowed materials.

The percentage of all borrowed materials for the past three (3) years from 2019-2021, with books

accounting for 43%, manuscripts accounting for 45%, and mixed materials accounting for 15%. According to preliminary data analysis, books and manuscripts are the most frequently borrowed items by patrons, and the researcher regards them as patrons' interests, as illustrated in Fig. 4.

As part of the data preprocessing, the dataset was grouped by the same or similar date and all the numbers of the same books borrowed on the same day were combined, resulting in 239 transactions without compromising the data's integrity.

In data mining, the support of an individual itemset represents the proportion or frequency of that itemset's occurrences within a dataset. It is often stated as a percentage or fraction of the total number of transactions or records in the dataset containing the relevant item. High support for an itemset indicates that it is common or popular in the dataset, whereas low support shows that the item is uncommon. Analyzing the support of each itemset is essential in determining the individual relevance and occurrence patterns of items, which is important for various data mining activities such as market analysis and recommendation systems. In the subsequent stage, we will provide assistance for each individual itemset that satisfies the minimum support criteria. To better illustrate this process, an example calculation is provided below.

$$Support(Technical\ Research) = \frac{82}{239} = 34.20\%$$

As a result, all book titles with support of one itemset.

Table 3. support of each one (1) item set

Book Titles	Total	Support
Title = Technical Research	82	34.20%
Title = The law on obligations and contracts: for business students	28	11.70%
Title = Jose Rizal: buhay, mga ginawa at mga sinulat ng isang henyo, manunulat, siyentipiko, at pambansang bayani	23	9.60%
Title = Jose Rizal: life, works, and writings of a genius, writer, scientist, and national hero	21	8.70%
Title = Philippine history & constitution	21	8.70%
Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna	20	8.30%
Title = Effectiveness of reading enhancement program in the reading comprehension of the grade 9 students	17	7.10%
Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna	17	7.10%
Title = Costumers level satisfaction in selected restaurant in Los Baños Laguna	16	6.70%
Title = Human resource management: from the practitioner's point of view	16	6.70%
Title = Philippine history	16	6.70%
Title = foundations of education	16	6.70%
Title = Effectiveness of localized materials in developing the reading comprehension level of grade 9 students	15	6.20%
Title = Jose P. Rizal his life, works, and role in the Philippine revolution	15	6.20%
Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan	15	6.20%
Title = Brief history of the Filipino people	11	4.60%
Title = The Philippines a unique nation	10	4.20%

The list of frequently borrowed books from the library. Within the dataset's 239 transactions and 2,360 total books borrowed, *Title = Technical Research,* has been borrowed 82 times with a support value of 34.20%, *Title = The law of obligations and contracts: for business students,* has been borrowed 28 times with a support value of 11.70%, *Title = Jose Rizal: buhay, mga ginawa at mga sinulat ng isang henyo, manunulat,* siyentipiko*, at pambansang bayani,* has been borrowed 23 times with a support value of 9.60% and so on, as

shown in Table 3.

Understanding the support of two itemsets enables researchers to detect relationships and trends in data, which is necessary for activities such as market basket analysis and recommendation systems. The result obtained is a direct outcome of the combined support provided by two distinct sets of items. Let me illustrate this concept through an explanatory example calculation.

$$Support(ConsumerPrefandConsumerDining) = \frac{14}{239} = 5.80\%$$

Table 4. support of two (2) item set

Book Titles	Total	Support
Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna	14	5.80%
Title = Technical Research, Title = Jose Rizal : buhay, mga ginawa at mga sinulat ng isang henyo, manunulat, siyentipiko, at pambansang bayani	13	5.40%
Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan	13	5.40%
Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan	11	4.60%
Title = Jose Rizal : buhay, mga ginawa at mga sinulat ng isang henyo, manunulat, siyentipiko, at pambansang bayani, Title = Jose Rizal : life, works, and writings of a genius, writer, scientist, and national hero	10	4.20%
Title = Philippine history & constitution, Title = Philippine history	10	4.20%
Title = Philippine history & constitution, Title = Brief history of the Filipino people	10	4.20%
Title = Philippine history & constitution, Title = The Philippines a unique nation	10	4.20%

The combination of two (2) itemsets that meet the minimum support, examples are as follows: *Title = Factors affecting* customer's *preference in choosing casual restaurant in Pila, Laguna | Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna*, has been borrowed 14 times and with a support value 5.80%, as shown in Table 4.

After identifying and locating its borrowing frequency patterns, we need to find a truly particular rule that meets the confidence for all intents and purposes minimum criteria by computing the confidence of the associative rule, which is very significant for all intents and purposes. The following is an example calculation for confidence value, which is highly significant.

$$Confidence(PhilNation\ and\ PhilHistory) = \frac{10}{10} = 100\%$$

Table 5. Confidence of two (2) item set

Book Titles	Confidence
Title = The Philippines a unique nation, Title = Philippine history & constitution	100.00%
Title = Brief history of the Filipino people, Title = Philippine history & constitution	90.91%
Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan, Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna	90.91%
Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan, Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna	86.67%
Title = Consumer perception on fast food outlet in Los Baños Laguna, Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna	83.33%
Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna, Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna	82.35%
Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna	76.92%
Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna	73.33%
Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan	71.43%
Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna	70.00%
Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna	66.67%

Combination of two item set meets the minimum support based on the computation of the Apriori algorithm with the Association data mining technique from Table 5. The minimum value of confidence for an Association Rule calculation was set to 60% during the evaluation stage. Setting the confidence value to 60% results in a more precise and accurate estimate. All of the book titles in Table V were defined with a confidence value greater than 60%.

Table 6. Associated books

Association Rules	Confidence
[Title = The Philippines a unique nation] --> [Title = Philippine history & constitution]	1.000 or 100%
[Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan] --> [Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna]	0.909 or 90.9%
[Title = Brief history of the Filipino people] --> [Title = Philippine history & constitution]	0.909 or 90.9%
[Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan] --> [Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna]	0.867 or 86.7%
[Title = Consumer perception on fast food outlet in Los Baños Laguna] --> [Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna]	0.833 or 83.3%
[Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna] --> [Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna]	0.824 or 82.4%
[Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan] --> [Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna]	0.769 or 76.9%
[Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan] --> [Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna]	0.733 or 73.3%
[Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna] --> [Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan]	0.714 or 71.4%
[Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna] --> [Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna]	0.700 or 70.0%
[Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan] --> [Title = Factors affecting customer's preference in choosing casual restaurant in Pila, Laguna, Title = Factors that affect consumer dining decision of selected restaurant in Los Baños, Laguna]	0.667 or 67.7%

The results of datasets run in RapidMiner to justify the support and confidence in the processed data. A confidence of 60% means that the 60% of the library patrons, who borrowed book *Title = The Philippines a unique nation* also borrowed *Title = Philippine history & constitution*, as shown in table 6. If patrons borrow books during that period, the placement of specific books will be positioned nearest to or in front of the library accommodation area in the following months. This information can be given to a librarian as a recommendation for effective and efficient book shelving.

5. Comparison

The performance of the FP-Growth, Apriori, and K-Nearest Neighbors (KNN) algorithms in book borrowing records varies depending on the task and the size of the dataset. Below is a particular comparison of these algorithms' performance:

FP-Growth is well-known for its ability to extract frequent patterns in large datasets. In contrast to Apriori, which generates a large number of candidate itemsets, FP-Growth obtains frequent patterns using a compressed data structure called a frequent pattern tree (FP-tree). This can result in shorter processing times, especially with massive datasets. However, when the dataset contains a large number of unique items, FP-Growth may not perform well because the structure of the FP-tree can become computationally expensive.

Apriori is a well-known algorithm for revealing frequent patterns in transactional data. The method produces a great deal of candidate itemsets and extracting them according to their support. Apriori's performance is affected by the data set's size as well as the minimum support threshold used to extract candidate itemsets. For massive data with low support thresholds, Apriori may be slower than FP-Growth, but it may perform effectively for smaller datasets with higher support thresholds.

K-Nearest Neighbor algorithm is a technique which may be employed to solve classification and regression problems. KNN can be used in book borrowing records to classify patrons depending on their borrowing behavior and predict their future borrowing patterns. The performance of KNN is determined by the number of neighbors used for classification and the distance metric used to measure user similarity. For large datasets, KNN can be slower than Apriori and FP-Growth, especially if the number of neighbors used is huge.

Finally, the performance of FP-Growth, Apriori, and K-Nearest Neighbor algorithms in book borrowing records depends on the specific task and the size of the dataset. FP-Growth and Apriori are valuable algorithms for extracting frequent patterns in transactional datasets, while K-Nearest Neighbor can be used for classification and regression problems in book borrowing records. Overall, Apriori Algorithm was used as the main data mining technique, since our dataset contains all transactions and need to identify all associated books by generating candidate itemsets based on support and confidence values.

6. Conclusion

Associated books were identified by examining the given results based on the combination of two (2) item sets that meet the minimum support and confidence from the Apriori algorithm calculations with association technique on book borrower's records. As shown in Table 5, all books with a confidence value greater than 60% were strongly advised to be placed together such as group of *Title = The Philippines a unique nation | Title = Philippine history & constitution*, with 100% confidence value; *Title = Factors that affect consumer dining decision of selected restaurant in Los Baños Laguna, | Title = Students satisfaction on the food service quality of the university cafeterias a basis for the creation of quality management plan | Title = Factors affecting customer's preference in choosing casual restaurants in Pila, Laguna*, with 91% confidence value etc. All the results will be extremely beneficial to the librarian in terms of organizing books and managing library resources. The researcher determined which books will be grouped together or arranged in the most convenient order and discovered that related books are more focused on histories and consumer preferences.

It is highly recommended that all of the books listed in table 6 Associated Books should be placed on the same bookshelf and near the front of the library service or lobby so that library patrons can easily see and find the books they need based on book borrowing patterns. Following that, the proposed calculation can be performed on a yearly transaction or monthly basis, if possible, for future research to generate better and

more accurate recommendations for book placement and shelving. This will be followed by more progressive service and greater customer or library patron satisfaction.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Mr. Gene Marck B. Catedrilla created the research title, built the model and conceptual framework, and supervised the research. Mr. Jerson D. Cecilio gathered and preprocessed the data, carried out the implementation, and performed the numerical calculations. Mr. Jonardo R. Asor verified and validated the data analysis techniques utilized in the research, assisted in the interpretation of the findings, and worked on the manuscript. All authors contributed significantly, provided criticism, and helped develop the research, analysis, and manuscript.

Acknowledgment

The authors would like to extend their heartfelt gratitude to the Laguna State Polytechnic University Los Baños library for extending and entrusting their book borrowers' record to be used for this article.

References

[1] Kavitha, M., & Subbaiah, S. (2020). Association rule mining using apriori algorithm for extracting product sales patterns in groceries. *International Journal of Engineering Research & Technology (IJERT)*, 2278–0181.

[2] Cunningham, S. J., & Frank, E. (2002). Market basket analysis of library circulation data. *Proceedings of the 6th International Conference on Neural Information Processing*.

[3] Prehanto, D. R., *et al*. (2020). Library book modeling data using the association rule method with apriori. *Algorithm in determining Book Placement and Analysis of Book Loans*, *29(05)*, 1244–1250.

[4] Techopedia. (n.d.). Association rule mining. Retrieved from: https://www.techopedia.com/definition/30306/association-rule-mining.

[5] Sharma, V. (2020, April 13). Supermarkets as an example – Apriori algorithm. LinkedIn. Retrieved from: https://www.linkedin.com/pulse/supermarkets-example-apriori-algorithm-vaibhav-sharma

[6] Zamzami, L., *et al*. (2021). Analysis of library book borrower patterns using apriori association data mining techniques. *Library Philosophy and Practice*, 6632.

[7] Moses, V. A., *et al*. (2011). Library management system using association rule mining. *International Journal* of *Biotech Trends and Technology*, 2249–0183.

[8] Husain, T., & Hidayati, N. (2021). The optimize of association rule method for the best book placement patterns in library*:* A monthly trial. *TEKNOKOM: Jurnal Teknologi dan Rekayasa Sistem Komputer*, *4(2)*, 53–59.

[9] Lai, D. (2022). Analyzing book borrowing patterns in dewan bahasa dan pustaka brunei using apriori algorithm. Retrieved from: https://ubdsds.substack.com/p/sufi-presents-at-myce-2022

[10] Devedzic, V. (2001). Knowledge discovery and data mining in databases. *Handbook of Software Engineering and Knowledge Engineering*, Yugoslavia.

[11] Kaur, M., & Kang, S. (2016). Market basket analysis: Identify the changing trends of market data using association rule mining. *Procedia Computer Science*, *85*, 78–85.

[12] Lutkevich, B. *Association Rules*. TechTarget. Retrieved from: https://www.techtarget.com/searchbusinessanalytics/definition/association-rules-in-data-mining

[13] Softwaretestinghelp. (2022). Frequent pattern (FP) growth algorithm in data mining. Retrieved from: https://www.softwaretestinghelp.com/fp-growth-algorithm-data-mining/
[14] Intellipaat. (2022). Data science – Apriori algorithm in python, market basket analysis. Retrieved from https://intellipaat.com/blog/data-science-apriori-algorithm
[15] Lubis, Z., Sihombing, P., & Mawengkang, H. (2020). Optimization of K Value at the K-NN algorithm in clustering using the expectation maximization algorithm. *Sci. Eng*., *725*.
[16] Javatpoint. K-Nearest neighbor (KNN) algorithm for machine learning. Retrieved from: https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning
[17] Verma, Y. (2021, September 21). Apriori vs. FP growth in market basket analysis – A comparative guide. Retrieved from: https://www.analyticsvidhya.com/blog/2021/09/apriori-vs-fp-growth-in-market-basket-analysis-a-comparative-guide/
[18] IBM. K-nearest neighbors algorithm. Retrieved from: https://www.ibm.com/ph-en/topics/knn