



Enhancing library book recommendations using clustering and association rule mining

Komgrit Rumdon, Siriwan Kajornkasirat, and Nathaphon Boonnam

DOI: <https://doi.org/10.47989/ir31363451>

Abstract

Introduction. This study investigates library user clusters to improve book recommendations. Data on daily borrowing behaviours from 2020 to 2025 were collected from the Prince of Songkla University, Pattani Campus library.

Method. Users were clustered using K-means. The optimal number of clusters was selected via the elbow method combined with silhouette scores. Borrowing-pattern relationships were analysed using association rule mining with the FP-Growth algorithm (minimum support 15%, minimum confidence 90%).

Analysis. The study examined the distribution of borrowed book categories and evaluated clustering performance using precision, recall, F-measure, and hit rate. Association rules were extracted to identify co-borrowing trends among categories.

Results. *Languages* was the most borrowed category, followed by *History & Geography* and *Science*. The best user clustering solution was $k = 3$, achieving precision 82.98%, recall 98.55%, F-measure 0.8815, and hit rate 97.57%. The strongest rule showed that users borrowing *Economics* [330], *Social Sciences* [300], and *History of Europe* [940] were likely to also borrow *History of Asia* [950]. Associations were also found among life sciences, biology, physics, and chemistry.

Conclusions. The study results indicate patterns of interest and knowledge connections derived from borrowing behaviour, which can be used to improve library services and serve as a model for data-driven library management.

Introduction

With the rapid development of the digital age, advances in information technology have greatly expanded the ways in which people obtain and process information. Especially in the field of library services, digital transformation has become the main trend. Libraries are no longer just traditional physical collections of books but have evolved into comprehensive centers for information processing and knowledge sharing. The main challenge of the digital transformation of libraries, as the key providers of knowledge services, is efficiently handling massive, diverse text resources and achieving accurate classification to support real-time responses, such as intelligent retrieval and personalised recommendations (Isnain et al., 2021; Wang et al., 2025; Yang et al., 2023). With a vast selection of books available in libraries, it can be time-consuming for some users to find the book they want. Additionally, many books are poorly recommended, leading to inefficient use of books (Safira & Mohammed, 2026).

Analysis of resource usage behaviour of users is a way to make executives and those involved, such as procurement and strategic planning departments, aware of the behaviours in using information resources at the library, and the quantity and type of resources used. This information can be used to improve the management and procurement of resources to meet real needs (Zhang et al., 2020). Previous studies on behaviour analysis in the use of information resources by service users have shown that such analysis can be useful in planning and improving the quality of operations in many dimensions, including procuring information resources that meet the needs, budget allocation, and quality assurance of the agency (Zhu & Xie, 2022). In addition, the Behaviour analysis can also be used to improve service quality and to provide information resources in line with the needs of service users in a cost-effective manner. Phoemsangwuwan et al. (2022) found that such analysis can be used in educational quality assurance. Therefore, to effectively manage the complexity and vastness of user behaviour data, we must analyse how users interact with these information resources.

However, the study of resource use behaviours primarily relies on observations and statistical data. Additionally, it was discovered that as data sets become larger, and more complex, the increasing power of computer technology has significantly increased data collection, storage, and manipulation capabilities (Khademizadeh et al., 2022). Universities and other educational institutions' libraries have gathered information on books borrowed over the years, but the useful information contained in this data has not yet been fully utilised. In many cases, students cannot find the necessary books in the library because other students have already borrowed them. On the other hand, many books remain unread by the students. From this perspective, library management faces challenges in deciding whether to provide books that would benefit users and how to place them on shelves (Smith, 2023). There is an urgent need for systems that can help library management make intelligent choices about how to handle these problems (Khademizadeh et al., 2022).

Data mining techniques should be applied to address these important problems. In this regard, association rules derived using the FP-Growth algorithm have been used to find connections among the data on book borrowing history in the database, indicating what lecturers and students borrow, with consideration of levels of support and confidence (FIRAT & Aslay, 2022). Applying association rules to analyse user behaviour makes it possible to identify trends when deciding to use various products and services. Such rules can be used in planning or determining strategies for designing services that respond to the interests of users or customers. Prior research shows that the application of relationship rules can be useful in organising promotional activities, designing services that are consistent with user needs, making marketing plans, and sourcing products and services efficiently (Baka et al., 2022; Jiwitan, 2022). Personalised recommendation systems optimise library resources by reducing wastage, improving user search processes, and enhancing marketing efforts by effectively targeting user needs (Makawana & Mehta, 2023). These

systems are designed to predict the likelihood of a book being borrowed by a user based on their interests, behaviour, and other relevant data, thus encouraging more reading among all users of the university library (Amin et al., 2025; Safira & Mohammed, 2026). The implementation of personalisation into recommendation systems can provide significant benefits to libraries.

The results from the analysis of library users' behaviour using an association rule mining approach can be used to support executives' policy decisions to set management strategies, allocate budgets, and provide services. Directions regarding purchasing the library's information resources to meet the needs lead to decisions about information use that are consistent with the current situation, such as improving service quality, quality assurance, and trends in the local context. The analysis of library users' behaviour will provide useful information to librarians, staff, or library operators to recommend resources that are consistent with the analysis of book borrowing behaviour, design user experiences that can support different types of users appropriately, and sustainably adjust information resource services. Policy and planning analysts can use the analysis results in risk management regarding budget allocation and organisational analysis, to be consistent with the main mission of the library. Finally, it can lead to innovations in library services, such as a book tracking system, a book recommendation system, re-borrowing notifications through Line Official, or creating membership renewal campaigns for users.

Literature review

Data mining can be used to analyse user needs from the library database, by identifying association rules (Guo, 2021; Singh et al., 2021) "To provide a comprehensive overview of the current research landscape, this literature review is organized into the following key themes".

K-Means clustering

Cluster analysis is a data mining technique that maximises object similarity within each group and minimises object similarity between groups to explore patterns and structures in datasets, aiding in the classification or grouping of data and facilitating decision-making (Zhang & Zhang, 2023).

K-Means clustering is an iterative algorithm (Mao et al., 2021). The main concept is to separate the data into k groups by first choosing k objects at random to serve as the initial centroids for clustering. Next, each object is assigned to the closest cluster based on the distance to each seed cluster centre. The cluster centre will be recalculated using the objects that are currently in the cluster each time a sample is assigned.

Assume that the initial dataset is $D = \{X_1, X_2, \dots, X_n\}$, where X_i, X_j are the data objects, and X_{il} is the " l " characteristic attribute of X_i (Fan et al., 2022). The definition of the Euclidean distance is shown in Equation 1.

$$d(X_i, X_j) = \sqrt{\sum_{l=1}^n (X_{il} - X_{jl})^2} \quad (1)$$

The choice of k has a significant impact on the results from K-means clustering. In this study, the K-means using the silhouette score and elbow method are employed to select the k value based on the sum of squares of errors (SSE) (Fan et al., 2022). The idea is that SSE describes each cluster's compactness to a certain extent. Therefore, if the SSE result is small, it will lead to better clustering, which is determined by Equation 2.

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} (C_i - x)^2 \quad (2)$$

Regarding the application of the K-means results in library work to organise reading promotion activities or library services, Wang et al. (2022) classified users into four categories. They identified which classification easily utilises the information resources and which one hardly does, along with which clusters have longer loan terms. Guzman et al. (2025) classified the library's collection of books to demonstrate which ones are frequently borrowed, sometimes borrowed, and seldom borrowed.

Association rule mining

There are several types of algorithms for recommending books to customers of a library, and the most used approach is association rules-based recommendations (Monika et al., 2025). The recommendation algorithm-based, on association rules, focuses on mining association rules between borrowed books from the library database. The general mining process can be divided into two steps: searching for frequent sets in the database and searching for strict rules within these frequent sets (Duchen, 2022).

Association rules are logical rules developed by data mining. Agrawal et al. (1993) proposed them to reveal potential relationships between data items. The concept of support is fundamental to understanding association rules. In the context of book borrowing, using literature books and computer books as examples, the support of a rule connecting these two indicates the probability that literature books and computer books will appear simultaneously throughout the entire data item set, while the confidence indicates the probability that literature books will be borrowed after computer books have been borrowed (Zhou, 2022). The basic steps of mining the association rule between borrowed books and the association rule-based recommendation algorithm are shown in Figure 1.

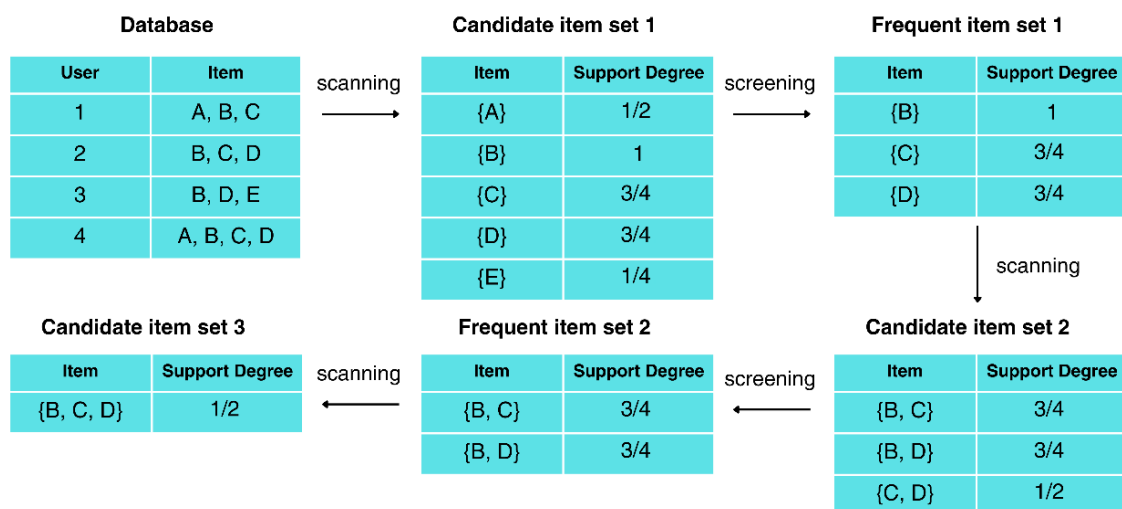


Figure 1. Basic diagram of the association rule mining step.

In Figure 1, the database of book borrowing records is simplified for illustrative purposes. The book titles are represented by letters, and the record number is reduced to four users and five books.

The flow steps are as follows:

(1) The database was scanned. Each book is treated as an item. Therefore, there are five items in the set of candidate items obtained from the scan (Duchen, 2022). The support degree, of each item was also calculated following Equation 3.

$$Support = \frac{n}{N} \quad (3)$$

where Support is the support level of the item, N is the total number of records in the database, here the number of users, and n is the number of records that the item contains.

(2) The pruning operation is performed on the candidate item set to remove items with support degree below a pre-set threshold, and the remaining item set constitutes the frequent item set and is recorded.

(3) The items in the frequent item set obtained in the previous step are combined in to pairs in a new candidate item set, and the support degree of each item in the new candidate item set is calculated.

(4) Return to step 2 to obtain the new set of frequent items set and record it.

(5) Repeat steps 2, 3, and 4 until no new candidate items sets can be generated. The association item set is a rule that may generate a strong association between a non-empty true subset of all the complete item set and the set of the remaining elements in the item set that occur frequently.

For example, take the item set {B, C} in Figure 1. The possible strong rules are $\{B\} \Rightarrow \{C\}$ and $\{C\} \Rightarrow \{B\}$ (Duchen, 2022). The confidence degree for calculating a strong rule $X \Rightarrow Y$ is calculated according to Equation 4.

$$Confidence = \frac{|X \cap Y|}{|X|} \quad (4)$$

Where Confidence is the confidence degree of the strong rule, $|X \cap Y|$ is the number of records that have both X and Y, and $|X|$ is the number of records that have X. When the confidence degree exceeds a certain threshold, the association rule is considered a strong rule. The recommendation algorithm provides book recommendations based on the user's book borrowing records and strong rules (Duchen, 2022).

Library management decision-making is supported by association rules. Liu (2018) developed a library management system based on association rules and described the modules required for mining, and used association rule mining to identify relationships among the most borrowed books each semester. They also used association rule mining to examine the books borrowed by users from the same group.

Apriori algorithm

The Apriori algorithm was the first association rule mining method to utilize support-based pruning to manage the exponential increase in candidate item sets (Mariana et al., 2017). Support-based pruning involves removing item sets that are infrequent in the dataset, specifically eliminating any $(k + 1)$ -length item set if its k -length subset is infrequent. This method also introduces the apriori principle (Mohanty et al., 2023), which states that every non-empty subset of a frequent item set must also be frequent. The algorithm helps find all frequent item sets that commonly appear in a database. Data processing was performed across various minimum support thresholds until frequent k -item sets with k of at least two were identified. The apriori algorithm is widely used to analyse customer behaviours for product and service recommendations across different domains. For example, it is applied in business to analyse social media and online shopping habits (Liao et al., 2021) and to understand the relationship between customers and products (Qisman et al., 2021). This data is then used to predict trends and guide marketing strategies. In education, the algorithm has been employed to analyse borrowing behaviours to recommend resources aligned with user interests (Zhou, 2020). It is also utilised in health science applications to analyse patterns within health-related data and to inform recommendations for the prevention of dengue (Kajornkasirat et al., 2025).

FP-Growth algorithm

The Frequent Pattern Growth (FP-Growth) algorithm is an improvement over the apriori algorithm in terms of execution speed, thereby addressing the shortcomings of the apriori algorithm. FP-Growth is one alternative that can be used to identify the most frequently occurring patterns (frequent item sets) within a dataset (Zhang & Zhang, 2023). The apriori algorithm needs to generate candidates to obtain a frequent item set. However, for the FP-Growth algorithm, generating candidates is not done because FP-Growth uses the concept of a development tree in search of a frequent item set. Characteristics of the FP-Growth algorithm are that the data structure used is a tree called the FP-Tree (Lakshika & Caldera, 2022). The FP-Growth algorithm can extract a frequent item set directly from the FP-Tree. The extraction of frequent item sets using the FP-Growth algorithm is performed by constructing a data structure called an FP-Tree (Lael & Pramudito, 2023).

Automated library system

Library automation employs information and communications technologies (ICT) to reduce manual processing in libraries, saving librarians significant time and effort. Integrating recommendation systems into these automated environments improves user satisfaction while delivering scalable efficiencies for large user bases. Researchers and librarians in academic settings have examined such systems, particularly their application to the OPAC module, which serves as the primary public interface for searching and retrieving resources (Puritat & Intawong, 2020; Puritat et al., 2021; Ikwuanusi et al et al., 2024). Open-source platforms such as Koha, Evergreen, and OpenBiblio support modular integration of advanced algorithms directly into bibliographic databases (Mamoriya, 2021). This flexibility enables the use of collaborative filtering and content-based techniques to deliver tailored discovery services aligned with academic research patterns (Alaoui et al., 2024). Numerous studies apply association rule mining to circulation logs, identifying recurring borrowing behaviours to predict and recommend relevant titles (Bhajantri et al., 2024; Xie & Liu, 2025). These approaches collectively advance personalised library services.

Recent studies indicate that K-Means Clustering is effective in classifying users according to their service usage behaviour, borrowing frequency, and specific interests. Meanwhile, association rule mining, along with the FP-Growth algorithm, has proven to be a valuable approach for identifying patterns of relationships among books and co-occurring borrowing behaviours. Integrating techniques with data from an automated library system can significantly enhance the accuracy of book recommendation systems. This advancement enables users to more efficiently discover books that align with their interests;, improves access to pertinent information resources;, and enhances the overall quality of library services.

Methodology

This study adopted a multi-step approach that involves data collection, data preprocessing, user clustering using RFM analysis, and association rule mining using the FP-Growth algorithm, as shown in Figure 2.

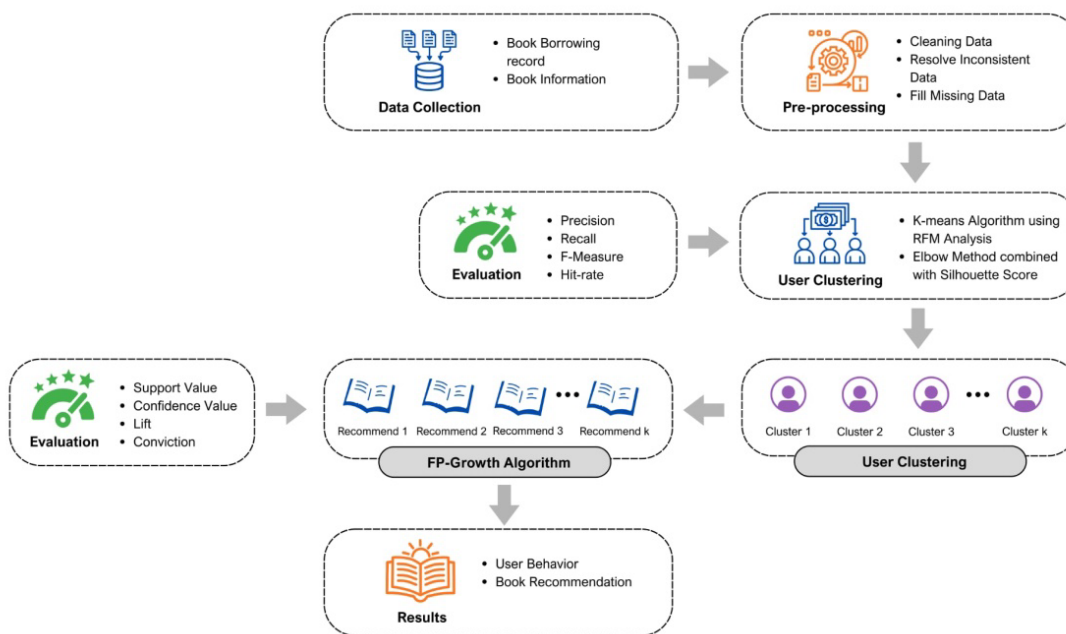


Figure 2. Research methodology.

Data collection

The dataset used in this study was obtained from the daily borrowing behaviour records of the Library of Prince of Songkla University, Pattani Campus, through the Automated Library System for Thai Higher Education Institutes (ALIST). This dataset covers the period from January 1, 2020, to June 30, 2025. The data include as variables daily borrowing, number of book borrowing, book data, book title, call number, book category, book collection, and number of copies.

This study followed established ethical principles and data governance frameworks by employing thorough anonymisation and de-identification procedures. All personally identifiable information was permanently removed before data processing and analysis. Additionally, techniques such as pseudonymisation, data masking, and tokenisation were utilised to reduce the risk of re-identification. The data were protected through encryption, strict access controls, and secure storage systems. There was no intention to identify individuals or infringe on library users' rights. These measures ensure the protection of participant confidentiality and compliance with relevant data protection regulations and institutional review requirements.

Data pre-processing

After collecting the data, the next step was data pre-processing, which involved data cleaning aimed at removing unused or duplicate data during the process of building a recommendation model. Additionally, to address data inconsistencies, various issues were identified that required correction, such as inconsistent data types and missing data, which were filled in with *null* values.

Then we identified the association in the book data, which includes a book collection and book classification, based on the book borrowing records. This stage produced four variables that serve as inputs for K-Means clustering.

User clustering

Cluster analysis is defined as the process of grouping a set of similar objects into clusters (Bhatia, 2019). The main idea of clustering is to try to divide instances into clusters so that instances within a cluster are highly similar but with high dissimilarity between clusters (Bhatia, 2019; Silwattananusarn et al., 2020). The K-Means clustering algorithm was used to analyse user

characteristics. In the K-Means clustering algorithm, objects are grouped into k clusters based on their attributes or features.

After data was pre-processed, we checked for recent transactions, frequency and the amount spent by the customers. To create a recency variable, decide the reference date, that is, one day prior to the last transaction. RFM (recency, frequency, monetary) analysis is a very popular customer segmentation and identification technique in database marketing (Christy et al., 2018; Patil & Anita, 2022). RFM has been applied to analyse the borrowing behaviour of users as follows: Recency (R) implies the number of days since the last borrowing records; Frequency (F) refers to the number of times or records that users make when borrowing books; Monetary (M) is defined as the total quantity of books borrowed.

K-Means algorithm is applied using the Euclidean distance metric to partition the users for RFM values. K-Means is used twice to analyse the amount obtained for Recent and Frequent transactions as mentioned below: i) To partition the customers based on the amount generated with recent transactions, and ii) To group the customers by the amount generated with frequent transactions. Clusters obtained in RFM analysis are evaluated using the silhouette score and elbow method (Cui, 2020; W3Schools, 2025), which analyses show well the resulting clusters are separated based on their book borrowing records. The elbow method analysis results combined with silhouette score values are compared for k values to identify the optimal number of clusters in this study. For each cluster, a book recommendation model is developed using the association rules from the FP-Growth algorithm.

Association rule mining

Association rule mining is a commonly used data mining technique to explore the relationships between items in large data sets. The rule has the form ‘cause’ $\Rightarrow$ ‘consequence’, which means “if the cause is satisfied, the consequence is likely to also occur”. In this study, association rules are used to recommend books. For example, the association rule ‘book 1’, ‘book 2’ $\Rightarrow$ ‘book 3’ means that if a reader reads book 1 and book 2, they are likely to read book 3 as well. Rules like this can be used for book recommendations (Zhang et al., 2024). The FP-Growth algorithm is used in order to convert the raw borrowing logs into useful association rules; by setting strict thresholds for support, confidence, and lift, the system eliminates insignificant patterns so as to focus on the most important associations, thus enabling the identification of new titles that are tailored to individual user behaviours. Books are recommended based on a set of highly prioritised criteria, ensuring that users are introduced to new titles that they have not previously read.

Evaluation

A confusion matrix is used to summarise the performance of the classification model. The Confusion matrix holds Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) counts, which will be used to calculate precision and recall scores, as shown in Table 1.

	Positive	Negative
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

Table 1. Confusion matrix

The metrics used to assess the performance of the research method includes Precision, Recall, and F-measure (Sundari & Subaji, 2022). The F-measure is defined as the harmonic average of precision (P) and recall (R).

(1) Precision (P) is the ratio of true positive predictions to the total predicted positives is defined by Equation 5.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

(2) Recall (R) is the number of true positive predictions divided by all relevant test values defined by Equation 6.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

(3) F-measure is defined as the harmonic average of Precision (P) and Recall (R) by Equation 7.

$$\text{F-Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

In this study, the hit rate is a binary metric that indicates whether relevant items are included in a user's recommendation list (Annisa et al., 2024). The higher the hit rate, the better the quality of the recommendations generated. The hit rate indicates that developing a high-quality recommendation system requires each recommendation target to receive at least one relevant item aligned with their interests (Annisa et al., 2024). The hit rate is defined by Equation 8.

$$\text{Hit Rate} = \frac{n \text{ hits}}{n \text{ users}} \quad (8)$$

Results

This section presents the results from K-means clustering and association rule mining. The research evaluation used a confusion matrix along with hit-rate, precision, recall, and F-measure.

General information analysis

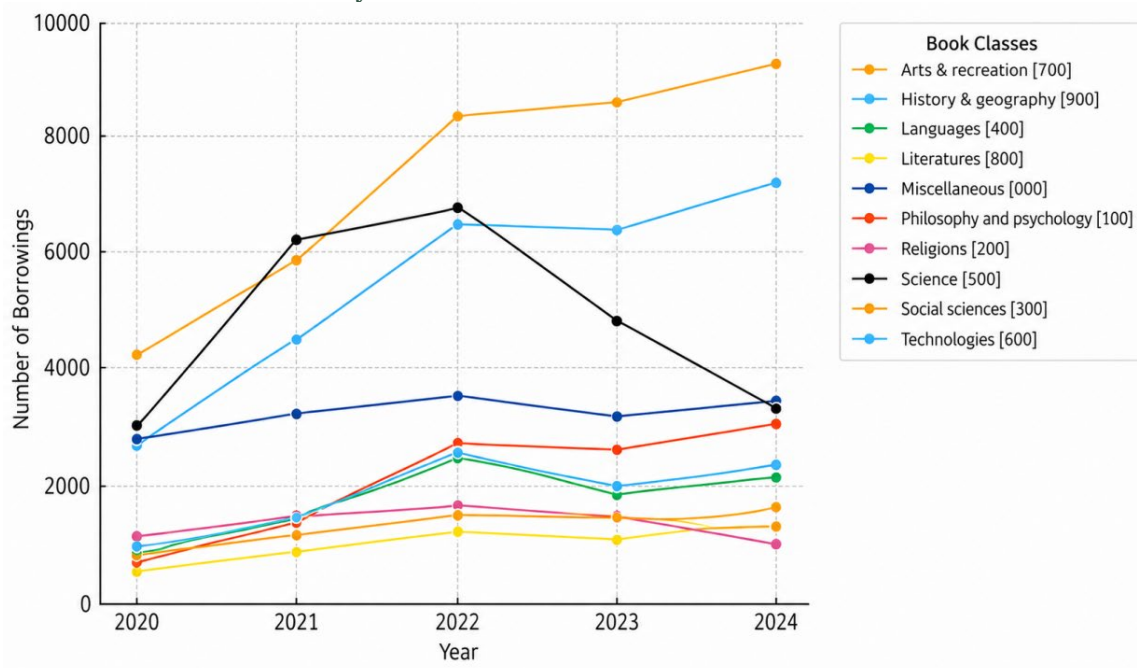


Figure 3. Trends of book borrowing by category 2020-2024

Figure 3 shows the distribution of book borrowing events across Dewey Decimal Classification (DDC) classes from 2020 to 2024, showing both annual totals and the proportion of each category. Overall, book borrowing fluctuated slightly over the five years. The class with the highest number of borrowing was Languages [400], followed by History & Geography [900] and Science [500], reflecting the popularity of media in linguistics, history, and science among users. In contrast, the categories Philosophy and Psychology [100] and Religions [200] saw low and relatively stable borrowing, while the category Technologies [600] showed an increasing trend after 2022, indicating a growing interest in technology and its applications. The results from this graph highlight trends and differences in book borrowing behaviour, which can be used to effectively plan the development of information resources to meet future user needs.

Clustering analysis

The results of the trained algorithm are then used to predict clusters in the test data. To find the most appropriate number of clusters from the results of K-means, the variance of the best cluster is searched using the silhouette score and the elbow method.

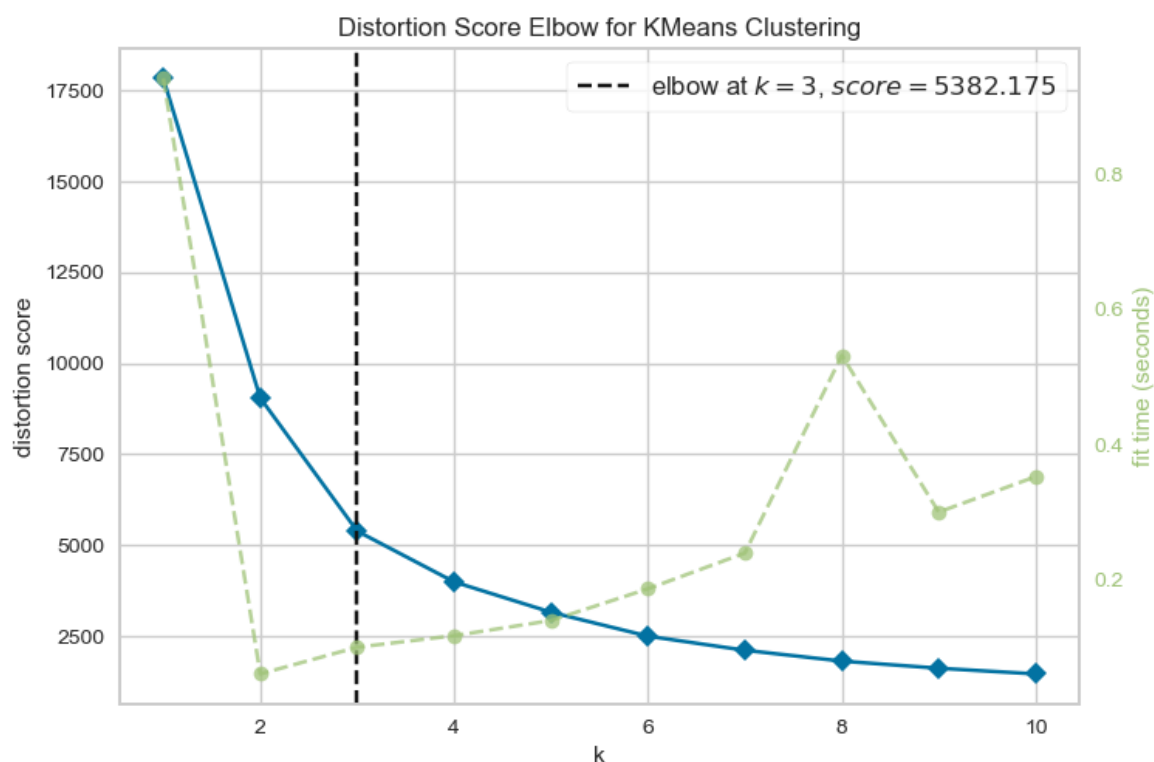


Figure 4. The optimal number of clusters of library user groups from K-Means analysis.

Figure 4 shows the relationship between the number of clusters (k) and the distortion score obtained from the K-Means clustering process. The left y-axis shows the distortion score, which is the sum of the squares of the distances between each data point and the centre of that cluster. The distortion score decreases rapidly as the number of clusters increases from 1 to 3, and after that, the rate of decrease begins to slow down. This turning point, or "elbow point," indicates that the optimal number of clusters is $k = 3$, which gives a distortion score of approximately 5,382.175. This suggests that dividing the library user data into three groups is the most appropriate grouping in terms of model accuracy and simplicity.

In addition, the right y-axis shows the model's fit time, which changes slightly as the number of clusters increases, reflecting the variation in the computation time of each clustering round. The

elbow method analysis results combined with silhouette score values confirm that $k = 3$ is the most appropriate number of clusters for library user behaviour.

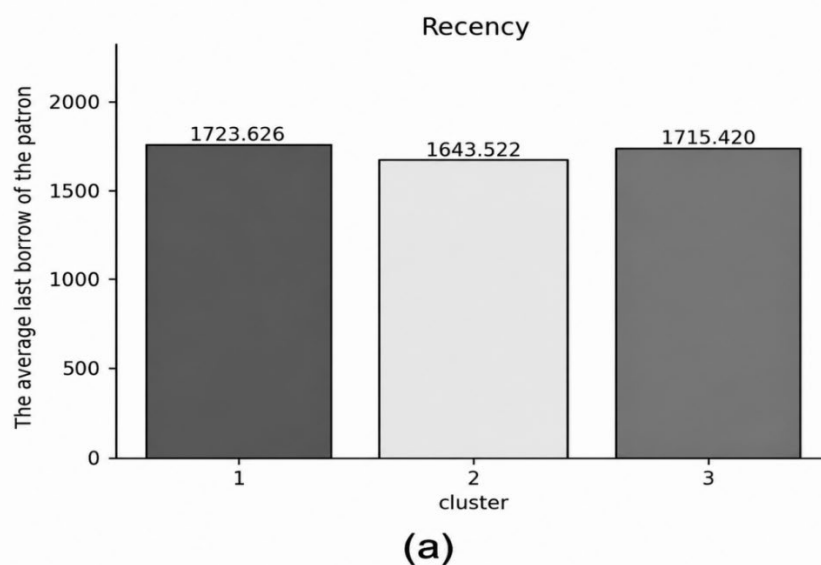
Number of clusters	Precision	Recall	F-Measure	Hit Rate
3	0.8298	0.9855	0.8815	0.9757
6	0.8241	0.9704	0.8865	0.9229
10	0.8059	0.9687	0.8745	0.9065

Table 2. Performance metrics of the recommendations across, three, six or ten clusters using the same dataset.

From Table 2 on the performance metrics of library user groupings using K-Means with the number of clusters equal to, three, six or ten, it is seen that all four metrics, namely *precision*, *recall*, *F-measure* and *hit rate*, were at a high level, indicating accuracy and stability of the model used to cluster the data. However, the model with three clusters was the most appropriate choice, with *precision* at 82.98%, *recall* at 98.55%, *F-measure* at 0.8815, and *hit rate* at 97.57%.

Based on the results of the user clustering, the researchers defined the behaviour of the three user groups using RFM, (Recency, Frequency, and Monetary) analysis, as shown in Figure 5.

Figure 5 (a), 5(b) and 5(c) illustrates the behaviour of each cluster as follows: Cluster 1 is defined as dormant, low-activity users. This group reflects users with the lowest level of library engagement across all dimensions, including both borrowing frequency and volume. They also have the longest gap between their last borrowing visits compared to other groups, indicating a tendency towards churn or being one-time users. ; Cluster 2 is defined as loyal high-value users. This group represents the highest quality target group, exhibiting significantly higher borrowing frequency and volume of resources borrowed compared to other groups (High Frequency & Monetary). Furthermore, they have the lowest Recency, indicating consistent return visits, reflecting their engagement and satisfaction with library services. Cluster 3 is defined as drifting moderate users. This group has a moderate borrowing history, neither too high nor too low, but their Recency is similar to Cluster 1, indicating that although they used the service moderately in the past, recent usage has become less frequent, and they may be heading towards cessation of service if not stimulated.



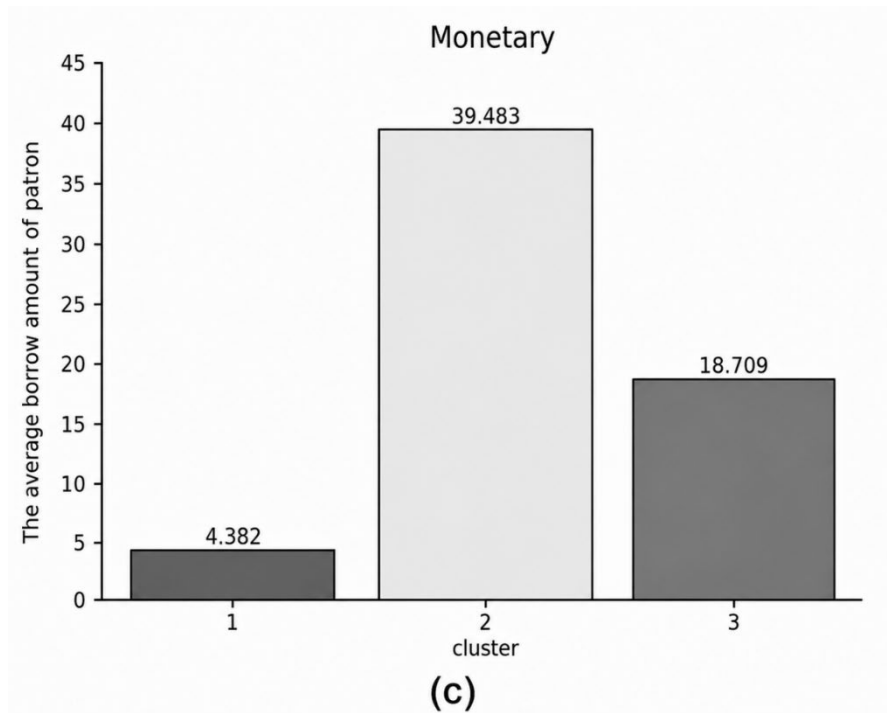
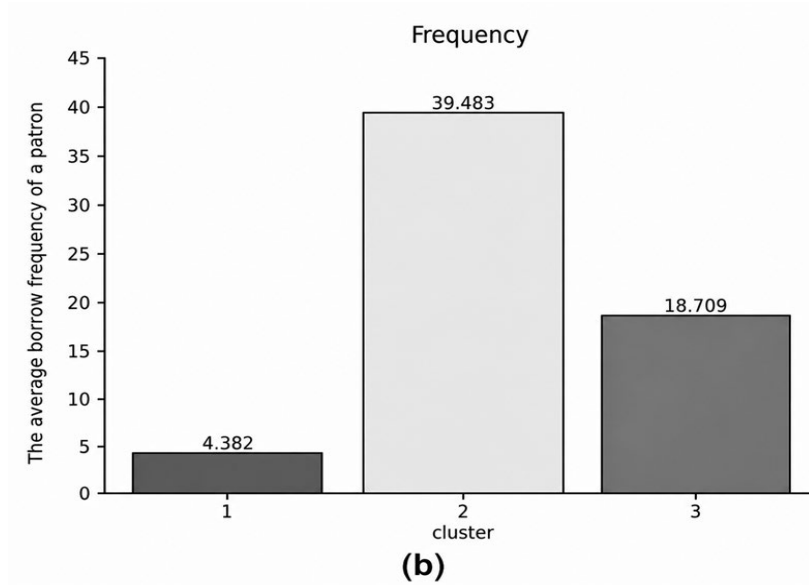


Figure 5. RFM analysis of the user groups was analysed using K-Means clustering.

Association rule mining results

Association rule mining was used to identify patterns in book borrowing within each user cluster, with a minimum support value of 0.15 and a minimum confidence value of 0.90. The results from all three clusters were similar, demonstrating a highly consistent pattern of book borrowing behaviour across the clusters, as shown in Table 3.

Rule	Premises	Conclusion	Support (%)	Confidence (%)	Lift	Conviction
1	Economics [330], Social Sciences [300], History of Europe [940]	History of Asia [950]	1.6	100.0	6.7	∞
2	Life Sciences; Biology [570], Languages [420], Physics [530]	Chemistry [540]	2.0	96.0	6.4	21.2
3	Political Science [320], Economics [330], History of Europe [940]	History of Asia [950]	1.9	95.8	6.4	20.2
4	Mathematics [510], Physics [530]	Chemistry [540]	1.9	95.7	6.4	19.8
5	Political Science [320], History of Europe [940], Social Sciences [300]	History of Asia [950]	1.6	95.0	6.3	17.0
6	Plants [580], Chemistry [540]	Life Sciences; Biology [570]	1.6	95.0	6.3	17.0
7	Animals [590], Physics [530]	Life Sciences; Biology [570]	1.5	94.7	6.3	16.0
8	Economics [330], History of Europe [940]	History of Asia [950]	2.4	93.5	6.2	13.1
9	Political Science [320], History of Europe [940]	History of Asia [950]	2.8	91.7	6.1	10.2
10	Economics [330], Law [340], History of Asia [950]	Political Science [320]	1.9	91.7	6.1	10.2
11	Animals [590], Medicine and Health [610]	Life sciences; Biology [570]	1.8	91.3	6.1	9.8
12	Law [340], Public Administration [350], Social Sciences [300]	Political Science [320]	1.6	90.5	6.0	8.9
13	Economics [330], Public Administration [350], History of Asia [950]	Political Science [320]	1.5	90.0	6.0	8.5

Table 3. Association rule mining results.

The analysis demonstrates significant interdisciplinary borrowing patterns among library users (Table 3). The association rules identified thirteen significant rules meeting the criteria for Support, Confidence, Lift, and Conviction.

Strongest relationship. Rule 1: Economics [330], Social Sciences [300], History of Europe [940] $\Rightarrow$ History of Asia [950] has a confidence of 100%. This means that whenever a book from each of the premises has been borrowed, the user also borrows one from History of Asia. Although the Support is only 1.6%, this perfect Confidence indicates a strong and consistent connection in historical data. This rule may reflect a group of researchers or students studying social science and economics topics in a comparative context between Europe and Asia.

Relationship between disciplines. Economics & History: There is a strong relationship with Asian History (see Rules 1, 3, 8, and 13 in Table 3), indicating that those interested in Economics are also likely to be interested in the History of Asia. Science-related subjects (Life Sciences, Physics, and Chemistry) in rules 2, 6, and 7 demonstrate the relationship between these subjects. For example, those who borrow Life Sciences and Biology books are likely to also borrow Chemistry, which is very reasonable because these subjects are related. Political Science, Public Administration, and Social Sciences show strong connections with History of Europe, History of Asia, and Law (see Rules 5, 9, 10, and 12), indicating that those interested in politics and government tend to study history and law simultaneously.

The association rule results in Table 3, showing the relationships among library users' book borrowings, are presented as a knowledge graph, as shown in Figure 6.

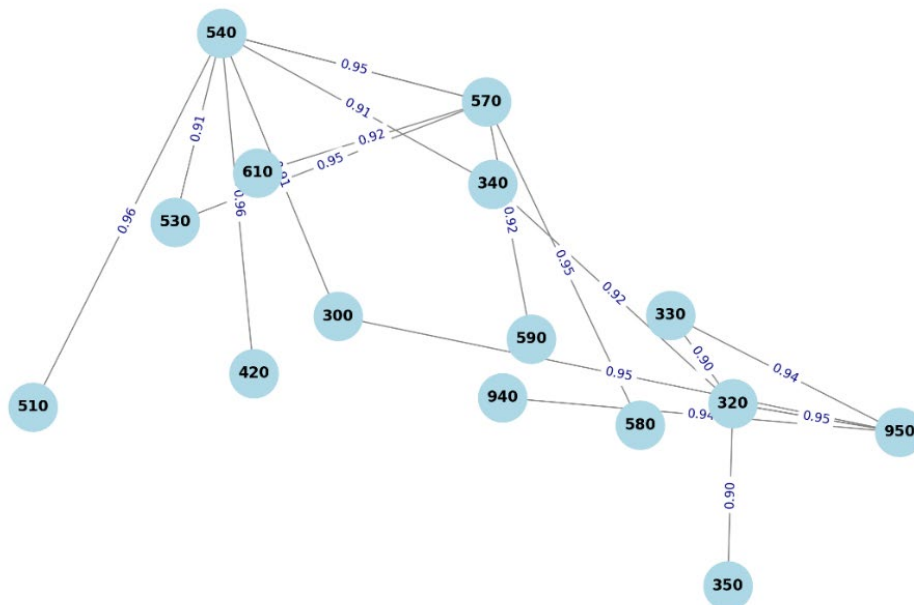


Figure 6. Knowledge graph of library book borrowing patterns.

Discussion

From the analysis of the RFM model for grouping users by K-means clustering, using the elbow method combined with silhouette scores, the most appropriate number of clusters was $k = 3$. This is because the borrowing behaviour of users tends to fall into a few main groups, with each group having different characteristics of interest (for example, a group of users who focus on specialised academic books, a group interested in general books, or a group with diverse borrowing behaviour. This is consistent with the quantitative evaluation using the four performance indicators *precision*, *recall*, *f-measure* and *hit rate*, which are all at a high level when compared to the models with the number of clusters equal to six and ten.

Furthermore, the analysis results show that the model with a smaller number of clusters ($k = 3$) outperforms the models with larger numbers of clusters ($k = 6$ and $k = 10$). This reflects that adding too many clusters may cause the model to over-segment. This may be because users have different book borrowing behaviour, resulting in too many clusters, which affects the performance evaluation of indicators such as Precision, Recall, F-Measure, and Hit Rate. Furthermore, setting the number of clusters at an appropriate level results in the loss of important behavioural relationships between users in the same group and reduces the ability to learn aggregate patterns in the data. Annisa et al. (2024) analysed performance with the number of clusters set at two, six and sixteen, and they found that the most appropriate number of clusters was $k = 2$. Setting the number of clusters at an appropriate level improves the accuracy of the recommendation system because it can maintain the relationship within the clusters more efficiently. Furthermore, this is consistent with the idea of Occam's Razor, which states that less complex models that produce results that are as efficient as or better than more complex models are often the most appropriate models in practice (Bargagli Stoffi et al., 2022). Therefore, user grouping by K-means clustering with three clusters was not only statistically appropriate but also reflected the actual behavioural structure of the library users, which can be used as an important basis for the development of the book recommendation module in the next step of the study.

An analysis of association rules using the FP-Growth algorithm was used to identify borrowing patterns within each user cluster. Analysis of all three clusters revealed highly similar borrowing behaviour, demonstrating a significant consistency across clusters. This indicates that borrowing behaviour does not vary based on user clustering but rather reflects the common behaviour of most library users. It was found that the rule with the highest strength is Economics [330], Social Sciences [300], European History [940] $\Rightarrow$ Asian History [950], with a Confidence of 100% and Support of 1.6%. This may be because the study of history helps to reflect on the social, cultural, economic and political background of each continent, which most humanities and social sciences refer to the regional context in the form of comparative analysis as a case study to create reflection (van Eeden, 2024). This relationship, despite its relatively low support, has a perfect confidence, reflecting a stable and consistent academic connection among the user group, which may be researchers or students studying economic and social issues in a comparative dimension between Europe and Asia to see social and economic distinctions. Rules 1, 3, 8, and 13, indicate a structural pattern of book borrowing: users interested in economics also tend to be interested in history. This is consistent with the idea of interdisciplinary knowledge discovery, which suggests that contemporary learners tend to connect knowledge across disciplines to gain deeper and more comprehensive perspectives (Wu et al., 2024).

On the other hand, the rules related to the natural sciences (rules 2, 6 and 7) show a relationship between Life Sciences, Biology, Physics and Chemistry, reflecting the nature of these disciplines sharing common ground and continuously sharing knowledge resources. This relationship is in line with the trend of science integration learning systems (STEM Integration) in which learners often connect knowledge from many fields, especially in general education or foundational subjects. Tinpun et al. (2025) stated that faculties that have integrated learning in education with science and technology will borrow books that are relevant and necessary for knowledge integration.

In summary, the discovery of association rules using the FP-Growth algorithm in this study not only reveals behavioural patterns in book borrowing, but also reflects interdisciplinary knowledge dynamics that demonstrate the linkage of knowledge from multiple disciplines to solve complex problems that cannot be solved by a single discipline. This concept emphasises the application of knowledge, skills, and ideas from multiple disciplines to real-world situations (Wu et al., 2024). The results can be used to develop a user-centric recommendation system based on borrowing history, which can significantly improve the efficiency of information and learning services in educational

institutions (Cox & Mazumdar, 2022). This may be because university libraries have a large budget for purchasing books to be consistent with the courses offered. Analysing the borrowing behaviour of users will increase efficiency in deciding how to purchase books, which will help reduce the budget for purchasing books that meet the needs of users. Khademizadeh et al., (2022) stated that libraries are the source of all knowledge and learning, and provide a large amount of information; therefore, association rule mining is required to dynamically analyse library databases and make strategic decisions for effective library management.

Conclusion

Users were grouped using K-means clustering. The results from the elbow method and silhouette scores indicated that the most suitable number of clusters was $k = 3$. The model in this case gave the highest accuracy measures when compared to clustering done with $k = 6$ and $k = 10$. This shows that the model is accurate and stable in classifying users with similar book borrowing behaviour. The fact that the model with fewer clusters gave better results than the model with more clusters suggests that the data structure reflects only a few main user groups with specific behavioural characteristics. This appropriate grouping reduces the problem of over-segmentation and supports a deeper understanding of the service usage patterns of each user group.

Analysis of the relationships in book borrowing using the FP-Growth algorithm revealed significant association rules consistent with academic characteristics across disciplines. The results reflect strong relationships among user groups interested in comparative economic and social issues between Europe and Asia. Other important relationship clusters were also found, such as those between the natural sciences (Biology, Physics, and Chemistry), indicating interdisciplinary borrowing patterns, a key trend among users in the current era.

The findings indicate that the integration of K-means clustering and the FP-Growth algorithm significantly enhances data-driven analyses of user behaviour. The clustering model effectively distinguishes between distinct user segments based on shared behavioural characteristics, while association rule mining reveals underlying patterns in book borrowing activities. This synergy of analytical techniques facilitates the development of a personalised book recommendation system, resulting in improved accuracy and operational efficiency.

While the theoretical foundations of behavioural analysis for decision-making support are well-documented in existing literature, this study contributes valuable empirical insights by extracting hidden patterns from real-world service usage data. Rather than restricting recommendations to broad subject categories for collection development, this analysis uncovers deeper interdisciplinary connections that may not be readily apparent through traditional curriculum structure examinations. The methodological framework employs a combination of data analytics techniques, utilising the RFM model to measure user engagement and assess borrowing intensity, K-means clustering to categorise users based on shared behaviours and the FP-Growth algorithm to identify frequently co-borrowed book sets through association rule mining. The outcomes of this analysis are then translated into a personalised recommendation system capable of offering precise and context-sensitive suggestions.

We note that while curriculum-based analysis serves as a valuable mechanism for guiding resource acquisition, an exploration of actual user behaviour provides a more accurate representation of user needs. By aligning collection management decisions with empirical usage patterns, libraries can effectively reduce the incidence of underutilised materials and enhance budget allocation efficiency. Overall, the proposed approach has significant implications for improving the quality of information services in digital-era libraries and presents a scalable model for advancing data-driven library management within educational institutions.

This study highlights the transformative potential of intelligent library systems through user clustering and association rule mining for book recommendations. By shifting from reactive to active services, these systems deliver targeted suggestions aligned with users' borrowing patterns and cluster interests. Consequently, libraries enhance resource utilisation efficiency, minimise unnecessary acquisitions of underused materials, and enable data-driven decisions for collection development. Additional benefits include elevated user satisfaction, reduced search times, and wider access to relevant knowledge. Overall, the findings advance service quality, optimise resource management, and support practical applications of recommendation technologies in digital library environments, fostering more responsive and efficient information systems.

Acknowledgement

The authors thank Seppo J. Karrila for constructive comments on this manuscript. This work was supported by the Graduate School, Prince of Songkla University, through the Talent Utilisation Type 2 under Grant No. TU2-04/2564. We also thank the Prince of Songkla University, Pattani Campus library, for the data in this study.

About the authors

Komgrit Rumdon received a Master of Information Science degree from Khon Kaen University, Thailand. He is currently a PhD candidate in the Philosophy Program in Data Science at the College of Digital Science, Prince of Songkla University, Thailand. He can be contacted at email: 6410035002@email.psu.ac.th.

Siriwan Kajornkasirat received the PhD degree in Computational Science from Walailak University, Thailand, in 2011. She has participated in PhD research experience at Deakin University, Australia, funded by the Royal Golden Jubilee PhD Program (RGJ-Ph.D. Program). In 2014, she was invited to a STEM education workshop under the International Visitor Leadership Program (IVLP). This is a U.S. Department of State program funded by the U.S. government. She is currently an associate professor at the Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani, Thailand. She serves as an assessor for the Education Criteria for Performance Excellence (EdPEX) and the Thailand Quality Award (TQA). Her research interests include Data Science, Artificial Intelligence, Computing Science, Advanced Analytic Online, STEM Education, Smart Farming, Internet of Things (IoT), Smart Health, Digital Marketing, E-Marketing for Tourism, AR/VR Technology. She can be contacted at email: siriwan.wo@psu.ac.th.

Nathaphon Boonnam received the Sc.D. degree in mathematics from Tokai University, Kanagawa, Japan, in 2014, and the M.Sc. and B.Sc. degrees in applied mathematics from King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, in 2011 and 2017, respectively. He was awarded the Senior Fellowship of the Higher Education Academy (SFHEA) under the UK Professional Standards Framework (UK-PSF). His research interests include STEM education, Mathematics education, Riemannian geometry, Computational geometry, Sphere theorem, Data Science, Agricultural Data Analytics, Smart Farming, wireless sensor networks.

References

- Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record*, 22(2), 207–216.
<https://doi.org/10.1145/170036.170072>.

- Alaoui, D., Riffi, J., Sabri, M. A., Aghoutane, B., Yahyaouy, A., & Tairi, H. (2024). Comparative study of filtering methods for scientific research article recommendations. *Big Data and Cognitive Computing*, 8(12), 190–190. <https://doi.org/10.3390/bdcc8120190>.
- Amin, F. M., Rusydiyah, E. F., & Azizah, A. N. (2025). Personalized library book recommendations using K-Means clustering and association rules. *Journal of Scientometric Research*, 14, 32–45. <https://doi.org/10.5530/jscires.20251005>.
- Anitha, P., & Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 1785–1792. <https://doi.org/10.1016/j.jksuci.2019.12.011>
- Annisa, S., Rini, D. P., & Abdiansah, A. (2024). Collaborative filtering recommendation system using a combination of clustering and association rule mining. *Journal of Information Systems and Informatics*, 6(3), 1499–1516. <https://doi.org/10.51519/journalisi.v6i3.802>.
- Baka, A., Persoh, S., Latekeh, I., Abdulyasat, A., Pangtip, P., & Deumong, F. (2022). Data mining techniques by using association rules for purchasing behavior analysis of retail customers: A case study of D2Shop, Pattani province. *Journal of Technology Management Rajabhat Maha Sarakham University*, 9(2), 18–29. <https://ph02.tci-thaijo.org/index.php/itm-journal/article/view/247578>.
- Bargagli Stoffi, F. J., Cevolani, G., & Gnecco, G. (2022). Simple models in complex worlds: Occam's Razor and statistical learning theory. *Minds & Machines*, 32, 13–42. <https://doi.org/10.1007/s11023-022-09592-z>.
- Bhajantri, A., Nagesh, K., Goudar, R. H., Dhananjaya, G. M., Kaliwal, R. B., Rathod, V. N., Kulkarni, A., & Govindaraja, K. (2024). Personalized book recommendations: A hybrid approach leveraging collaborative filtering, association rule mining, and content-based filtering. *EAI Endorsed Transactions on Internet of Things*, 10. <https://doi.org/10.4108/eetiot.6996>
- Bhatia, P. (2019). *Data mining and data warehousing: Principles and practical techniques*. Cambridge University Press. <https://doi.org/10.1017/9781108635592>.
- Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking: An effective approach to customer segmentation. *Journal of King Saud University - Computer and Information Sciences*, 33(10), 1251–1257. <https://doi.org/10.1016/j.jksuci.2018.09.004>.
- Cox, A. M., & Mazumdar, S. (2022). Defining artificial intelligence for librarians. *Journal of Librarianship and Information Science*, 56(2), 330–340. <https://doi.org/10.1177/09610006221142029>
- Cui, M. (2020). Introduction to the K-Means clustering algorithm based on the elbow method. *Geoscience and Remote Sensing*, 3, 9–16. <https://doi.org/10.23977/accaf.2020.010102>.
- Duchen, H. (2022). A comparative study of various book recommendation algorithms for public libraries. *Technical Services Quarterly*, 39(4), 369–380. <https://doi.org/10.1080/07317131.2022.2125676>.
- Fan, S., Xu, Y., Zhu, B., et al. (2022). Student behavior analysis research of based on data mining. *The 2nd International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR)*, Xi'an, China, November 25–27, 2022 (pp. 130–134). IEEE. <https://doi.org/10.1109/ICBAR58199.2022.00032>.

- FIRAT, İ., & Aslay, F. (2022). Knowledge Discovery from Library Automation via Bibliomining using the Apriori Algorithm. *Gazi University Journal of Science*, 35(4), 1344–1357. <https://doi.org/10.35378/gujs.983642>.
- Guo, Y. (2021). CNS: interactive intelligent analysis of financial management software based on apriori data mining algorithm. *International Journal of Cooperative Information Systems*, 30(01n04), 2150008. <https://doi.org/10.1142/S0218843021500088>.
- Guzman, E. E. C., Tano, I. M., Piad, K., Lagman, A. C., Espino, J., Mababa, J., & Victoriano, J. M. (2025). Unlocking insights from academic library data using clustering and recommender dashboard analytics for enhanced book collection management. *World Journal of Advanced Research and Reviews*, 25(3), 645–657. <https://doi.org/10.30574/wjarr.2025.25.3.0779>.
- Ikwuanusi, U. F., Onunka, O., Owoade, S., & Uzoka, A. C. (2024). Revolutionizing library systems with advanced automation: A blueprint for efficiency in academic resource management. *International Journal of Scholarly Research in Multidisciplinary Studies*, 5(2), 19–40. <https://doi.org/10.56781/ijsrms.2024.5.2.0045>.
- Isnain, A. R., Supriyanto, J., & Kharisma, M. P. (2021). Implementation of the K-Nearest Neighbor (K-NN) algorithm for public sentiment analysis of online learning. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(2), 121–130. <https://doi.org/10.22146/ijccs.65176>.
- Jiwitan, N. (2022). The discovery of association rules and data analysis to optimize Japanese food sales. *Journal of Information Science and Technology*, 12(1), 1–12. <https://doi.org/10.14456/jist.2022.1>.
- Kajornkasirat, S., Hnusuwan, B., Puttinaovarat, s., Puangsuwan, K., & Kaewsuwan, N. (2025). Recommender system for dengue prevention using machine learning. *IAES International Journal of Artificial Intelligence*, 14(2), 1106–1115. <https://doi.org/10.11591/ijai.v14.i2.pp1106-1115>.
- Khademizadeh, S., Nematollahi, Z., & Danesh, F. (2022). Analysis of book circulation data and a book recommendation system in academic libraries using data mining techniques. *Library & Information Science Research*, 44(4), 101191. <https://doi.org/10.1016/j.lisr.2022.101191>.
- Kortemeyer, G., & Dröschler, S. (2021). A user-transaction- based recommendation strategy for an educational digital library. *International Journal on Digital Libraries*, 22(2), 147–157. <https://doi.org/10.1007/s00799-021-00298-8>.
- Lael, T. A. D., & Pramudito, D. A. (2023). Use of Data Mining for the analysis of consumer purchase patterns with the Fp growth algorithm on motor spare part sales transactions data. *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, 4(2), 128–136. <https://doi.org/10.34306/itsdi.v4i2.582>.
- Lakshika, T., & Caldera, A. (2022). Association Rules for knowledge discovery from e-news articles: A review of Apriori and FP-Growth algorithms. *Advances in Science Technology and Engineering Systems Journal*, 7(5), 178–192. <https://doi.org/10.25046/aj070519>.
- Liao, S., Widowati, R., & Hsieh, Y. (2021). Investigating online social media users' behaviors for social commerce recommendations. *Technology in Society*, 66(2021), 1–14. <https://doi.org/10.1016/j.techsoc.2021.101655>.

- Liu, Y. (2018). Data mining of university library management based on improved collaborative filtering association rules algorithm. *Wireless Personal Communications*, 102(4), 3781–3790. <https://doi.org/10.1007/s11277-018-5409-y>.
- Makawana, M., & Mehta, R. G. (2023). Discovering search space using M-distance clustering of semantic relatedness based weighted network for the content-based recommender system. *Journal of Scientometric Research*, 12(2), 243–253. <https://doi.org/10.5530/jscires.12.2.024>.
- Mamoriya, A. (2021). Open-source software and proprietary software for library automation. *Journal of Advanced Research in Library and Information Science*, 8(2), 1–3. <https://doi.org/10.24321/2395.2288.202101>.
- Mariana, S., Surjandari, I., Dhini, A., Rosyidah, A., & Prameswari, P. (2017). Association rule mining for building book recommendation system in online public access catalog. 2017 3rd International Conference on Science in Information Technology (ICSITech), 246–250. <https://doi.org/10.1109/ICSITech.2017.8257119>.
- Mao, Z., Wu, J., Qiao, Y., & Yao, H. (2021). Government data governance framework based on a data middle platform. *Aslib Journal of Information Management*, 74(2), 289–310. <https://doi.org/10.1108/AJIM-03-2021-0068>.
- Mohanty, B., Tripathy, M., & Champati, S. (2023). Performance Analysis of Association Rule Mining Algorithms: Evidence from the Retailing Industry. *Journal of Engineering Science and Technology Review*, 16(5), 108–122. <https://doi.org/10.25103/jestr.165.14>.
- Monika, W., Nasution, A. H., Syam, F. A., & Wijesundara, C. (2025). Clustering Of Library's Patron behavior Using Machine Learning. *Digital Zone Jurnal Teknolog iInformasi DanKomunikasi*, 16(1), 1–11. <https://doi.org/10.31849/digitalzone.v16i1.19680>.
- Phoemsangwuan, P., Keawmen, K., & Muenpo, S. (2022). Assessment of information technology services and analyse of the behavior of using information technology resources of the library for educational quality assurance: a case study of Surasak Srisuk Library, Mahidol University. *TLA Bulletin*, 66(1), 104–118. https://so06.tci-thaijo.org/index.php/tla_bulletin/article/view/255954.
- Puritat, K., & Intawong, K. (2020). Development of an open source automated library system with book recommendation system for small libraries. 2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), Pattaya, Thailand, March 11–214, 2020 (pp. 128–132). IEEE. <https://doi.org/10.1109/ECTIDAMTNCN48261.2020.9090753>.
- Puritat, K., Julrode, P., Ariya, P., Sangamuang, S., & Intawong, K. (2021). Book recommendation for library automation use in school libraries by multi features of support vector machine. *International Journal of Advanced Computer Science and Applications*, 12(4). <https://doi.org/10.14569/ijacsa.2021.0120426>.
- Qisman, M. Rosadi, R., & Abdullah, A. S. (2021). Market basket analysis using apriori algorithm to find consumer patterns in buying goods through transaction data (case study of Mizan computer retail stores). *Journal of Physics: Conference Series*, 1722(2021). <https://doi.org/10.1088/1742-6596/1722/1/012020>.
- Safira, W. E., & Mohammed, S. (2026). Supporting academic library collection decisions using K-Means-based book recommendation. *Artificial Intelligence in Educational Decision Sciences*, 1(2), 22–33. <https://doi.org/10.66053/aieds.v1i2.25>.

- Silwattananusarn, T., Kulkanjanapiban, P., & Athakravisunthorn, K. L. (2020). Mining and analyzing patron's book-loan data and university data to understand library use patterns. *International Journal of Information Science and Management*, 18(2), 151–172. <https://doi.org/10.48550/arXiv.2008.03545>.
- Singh, P. K., Othman, E., Ahmed, R., Mahmood, A., Dhahri, H., & Choudhury, P. (2021). Optimized recommendations by user profiling using apriori algorithm. *Applied Soft Computing*, 106, 107272. <https://doi.org/10.1016/j.asoc.2021.107272>.
- Smith, T. (2023, May 4). Library funding becomes the “nuclear option” as the battle over books escalates. NPR. <https://www.npr.org/2023/05/04/1173274834/book-bans-library-funding-missouri-texas-ashcroft> (Archived by the Internet Archive at <https://web.archive.org/web/20230621222422/https://www.npr.org/2023/05/04/1173274834/book-bans-library-funding-missouri-texas-ashcroft>)
- Sundari, P. S., & Subaji, M. (2022). An improved hidden behavioral pattern mining approach to enhance the performance of recommendation system in a big data environment. *Journal of King Saud University - Computer and Information Sciences*, 34(10, Part A), 8390–8400. <https://doi.org/10.1016/j.jksuci.2020.09.010>.
- Tinpun, N., Rumdon, K., Kaewsuwan, N., Theppaya, T., Matcha, W., Kammanee, K., & Khammanee, W. (2025). Using the association rule to analyze the book borrowing behavior of Prince of Songkla University students with data mining techniques. *Journal of Information Science Research and Practice*, 43(3), 23–41. <https://doi.org/10.14456/jiskku.2025.16>.
- Kajornkasirat, S., Hnusuwan, B., Puttinaovarat, s., Puangsuwan, K., & Kaewsuwan, N. (2025). Recommender system for dengue prevention using machine learning. *IAES International Journal of Artificial Intelligence*, 14(2), 1106–1115.
- Van Eeden, E. S. (2024). Considering the many variations of regional history. *Histories of People and Place*, 1–22. <https://doi.org/10.1080/28334299.2024.2433919>.
- W3Schools. (2025). *Machine learning: K-means*. W3Schools. https://www.w3schools.com/python/python_ml_k-means.asp (Archived by the Internet Archive at https://web.archive.org/web/20250617131157/https://www.w3schools.com/python/python_ml_k-means.asp)
- Wang, C., You, F., & Wang, Y. (2025). Text intelligent classification model based on improved KNN algorithm in library service transformation. *Systems and Soft Computing*, 7, 200313. <https://doi.org/10.1016/j.sasc.2025.200313>.
- Wang, J., Alroobaea, R., Baqasah, A. M., Althobaiti, A., & Kansal, L. (2022). Study on library management system based on data mining and clustering algorithm. *Informatica (Slovenia)*, 46(9), 17–24. <https://doi.org/10.31449/inf.v46i9.3858>.
- Wu, S., Lin, M., Ji, M., & Wang, T. (2024). Exploring core knowledge in interdisciplinary research: insights from topic modeling analysis. *Applied Sciences*, 14(21), 10054. <https://doi.org/10.3390/app142110054>.
- Xie, S., & Liu, Z. (2025). Library resource recommendation integrating collaborative filtering and association rules. *Journal of Computational Methods in Sciences and Engineering*. <https://doi.org/10.1177/14727978251364442>.

- Yang, K., Tian, C., Xian, H., Tian, W., & Zhang, Y. (2023). Query on the cloud: improved privacy-preserving k-nearest neighbor classification over the outsourced database. *World Wide Web*, 26(4), 1747–1774. <https://doi.org/10.1007/s11280-022-01093-4>.
- Zhang, W., Ji, M., Wang, H., Zhang, C., Xu, S., Wu, Y., & Song, F. (2020). An exploration of how the resourcing of staff and book purchasing affects the development of China's public school libraries in Guangzhou. *Journal of the Australian Library and Information Association*, 69(1), 70–92. <https://doi.org/10.1080/24750158.2019.1702140>
- Zhang, X., & Zhang, J. (2023). Analysis and research on library user behavior based on apriori algorithm. *Measurement: Sensors*, 27, 100802. <https://doi.org/10.1016/j.measen.2023.100802>.
- Zhang, X., Zang, L., Bai, W., & Liu, H. (2024). Big data analysis and application of library information resources. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns-2024-1212>.
- Zhou, X. (2022). Personalized book recommendation of the university library coupled with collaborative filtering and FP-Growth algorithm. *Proceedings of the 4th International Conference on Machine Learning and Machine Intelligence*, Hangzhou, China, September 17–19, 2021 (pp. 172–176). ACM Digital Library. <https://doi.org/10.1145/3490725.3490751>.
- Zhu, Q., & Xie, X. (2022). Users' needs and expectations of immersive learning spaces in an academic library: A survey. *Journal of Librarianship and Information Science*, 55(2), 371–382. <https://doi.org/10.1177/09610006221081844>.
- Zhou, Y. (2020). Design and implementation of book recommendation management system based on improved Apriori algorithm. *Intelligent Information Management*, 12(3), 75–87. <https://doi.org/10.4236/iim.2020.123006>.

Copyright

Authors contributing to *Information Research* agree to publish their articles under a [Creative Commons CC BY-NC 4.0 license](https://creativecommons.org/licenses/by-nc/4.0/), which gives third parties the right to copy and redistribute the material in any medium or format. It also gives third parties the right to remix, transform and build upon the material for any purpose, except commercial, on the condition that clear acknowledgment is given to the author(s) of the work, that a link to the license is provided and that it is made clear if changes have been made to the work. This must be done in a reasonable manner, and must not imply that the licensor endorses the use of the work by third parties. The author(s) retain copyright to the work. You can also read more at: <https://publicera.kb.se/ir/openaccess>