Supporting Academic Library Collection Decisions Using K-Means–Based Book Recommendation

Wahyuni Edsa Safira ✉
Universitas Negeri Makassar, Makassar, Indonesia
Saif Mohammed
Bilkent University, Turkey

ABSTRACT

Purpose - This study aims to develop a data-driven book recommendation system to support academic library collection management using the K-Means clustering method.

Methods - The study utilized book borrowing data from the Library of the Department of Informatics and Computer Engineering at Makassar State University collected over a 22-month period. Borrowing records were grouped by book categories and monthly borrowing frequencies, then processed into numerical variables. The K-Means algorithm was applied to identify borrowing pattern clusters, and cluster quality was evaluated using the Silhouette Coefficient to assess cohesion and separation.

Findings - The analysis produced three distinct clusters representing different borrowing behaviors. Programming and information technology books formed the most frequently borrowed cluster, research methodology books showed increased demand during specific academic periods, and education and learning methods books exhibited relatively lower borrowing intensity. The average Silhouette Coefficient value of 0.35 indicates a moderate yet acceptable clustering structure for recommendation and managerial purposes.

Research limitations - This study is limited to historical transaction data from a single departmental library and does not incorporate user profiles or qualitative preference data, which may restrict generalizability to other academic library contexts.

Originality - This study contributes empirical evidence on the use of K-Means clustering for book recommendation and decision support in academic libraries, demonstrating how borrowing pattern analysis can inform data-driven collection management and improve the relevance of library services. The findings also highlight the practical role of clustering analytics in supporting efficient resource allocation and evidence-based planning within higher education libraries and departmental level strategic decisions.

To cite this article : Wahyuni Edsa Safira & Saif Mohammed (2026). Supporting Academic Library Collection Decisions Using K-Means–Based Book Recommendation. Artificial Intelligence in Educational Decision Sciences, 1(2), 22–33.

INTRODUCTION

In the digital age, access to information has become increasingly rapid and abundant, including in the form of textbooks, journals, and academic literature (Maesaroh et al., 2025). This condition is characterised by the increasing availability of information sources that can be accessed through various media and learning support tools. Many universities in Indonesia still rely on conventional book search systems in libraries, such as manual catalogues or searches based on title and author (Hermawan & Risparyanto, 2025; Mutiarani et al., 2022). This search pattern generally requires direct user involvement in browsing the available collections. This is often inefficient, especially when users need book recommendations that suit their interests or level of understanding (Elmustian & Firdaus, 2024; Jannah et al., 2025; Mugnia & Mutoffar, 2024). In such situations, the search process can require relatively more time and effort. Meanwhile, libraries must be able to manage their collections of books or library materials in line with user needs, including providing book recommendations that can be used as library reference materials (Abdurrachman & Chahyati, n.d.; Nurcahyani, 2023).Proper collection management is an important part of supporting the optimal use of information sources.

Technological developments also mean that libraries and librarians must face increasingly significant and diverse challenges, both in managing print and digital collections. This situation requires adjustments to be made to the information management process. One of the main challenges is the organisation of collections, as mentioned above regarding book recommendations (Putri & Rahmah, 2019; Zabidi, 2024). This challenge is related to efforts to manage collections so that they can be used more effectively. One alternative for organising library collections is the implementation of data mining using the K-Means Clustering method to help librarians analyse and group books based on borrowing data. The results of this grouping can be used to support decisions regarding frequently borrowed books so that they can be used as book recommendations, especially in university libraries such as the Makassar State University Library, particularly in the Department of Informatics and Computer Engineering (Hasanah & Purnomo, 2022; Iryani, 2020).

Based on a survey conducted at the Makassar State University Library, Department of Information Technology and Computer Science, approximately 37.72% of respondents stated that the library collection was still in the uncertain category. These results illustrate that some students are not yet fully convinced that the available collection meets their academic needs. This perception is related to students' experiences in utilising the library collection. Students find it difficult to find books that suit their academic needs, resulting in low utilisation of the available collection (Pandita, 2017). This condition indicates that there are obstacles in the process of searching for relevant library materials. As the number of books in the collection increases each year, the complexity of the search process also tends to increase. Students often need more time to find books that suit their needs quickly and efficiently (Ulfah & Irwaty, 2022).This situation indicates that the increase in the collection has not been fully offset by ease of access and information retrieval.

In planned, sustainable library collection development research that meets user needs, it is important to improve services, support learning and research, and strengthen the role of libraries as information centres. The application of the K-Means Clustering algorithm enables the grouping of collections based on borrowing data, so that the system can provide relevant book recommendations. The results show that the clusters formed according to borrowing patterns and popular book categories (Asminah, 2022; Nasir, 2020).Another study combining TF-IDF in synopses and Cosine Similarity formed three user interest clusters, although there were differences in results between manual calculations and the WEKA application due to sensitivity to centroid initialisation and the suboptimal determination of the number of clusters and integration of historical data (Fadlina et al., 2024).A study by Zabidi at the Indonesian National Defence University Library also proved the effectiveness of K-Means in collection and service management, but also highlighted the importance of data quality, system evaluation, and technical training for users (Zabidi, 2024).

The use of the K-Means Clustering method can be adapted to assess book collections based on user borrowing data and reading interests (Dacwanda & Nataliani, 2021). This approach provides an overview of the patterns of collection utilisation in libraries. On the other hand, this method also reflects the application of K-Means in the context of academic data evaluation and management in libraries (Fathuroh, 2023).The findings show that K-Means can be used as a data analysis tool in library information management. In addition, K-Means is also used to identify library visitor profiles, which is relevant to efforts to group library users and collections more accurately (Asriningtias et al., 2023).The results of this study show that the K-Means method has broad application potential in supporting library data and collection management. Furthermore, Gustin, S., et al in their study "Data Mining Techniques Using the K-Means Method for Clustering and Searching Books in the Asahan Regency Library" directly discusses the implementation of K-Means for grouping library book collections, focusing on improving the efficiency of book searches (Gustin et al., 2022). Research in the same year also explored the application of clustering algorithms in visual search models in libraries, which is relevant to the development of more interactive and efficient library information systems (Zeng et al., 2022). Finally, another study also discussed exploring the efficiency of K-Means in data clustering, but more specifically in the context of health, proving that this method is very suitable for research aimed at clustering data (Abdurrachman & Chahyati, n.d.; Bin Samer & Darujati, 2023; Sutanto et al., n.d.).

Previous studies have shown that the application of the K-Means Clustering algorithm in library systems has yielded good results in grouping books and making recommendations based on borrowing patterns and user interests. These results indicate that a historical data-based clustering approach can help libraries understand trends in the use of their collections by users. However, previous research on library book clustering has generally been limited to the use of general variables such as works, philosophy, religion, and language, so that the resulting clusters do not fully consider the characteristics of specific scientific fields. This condition means that there are no specific variables designed specifically for the context of the Information Technology and Computer Science department library. Due to these limitations, the implementation of K-Means Clustering in previous studies has not thoroughly utilised variables relevant to the discipline of Information Technology and Computer Science, both in terms of collection content and the academic needs of its users. The resulting book groupings tend to still use general categories, making them less capable of capturing the nuances of interest, focus of study, and more specific information needs of students and teaching staff in these departments. This situation indicates that there is room for development in the application of clustering methods to make them more contextual and appropriate to the characteristics of specific academic environments.

This study aims to develop a book recommendation system that can group collections based on user borrowing patterns so that it can provide more relevant, efficient, and appropriate book recommendations for students in the Department of Informatics and Computer Engineering. This approach is intended to help make the collection search process more focused and tailored to the academic needs of users. The K-Means Clustering algorithm is used to develop a book recommendation system tailored to the needs of users in the Department of Informatics and Computer Engineering at Makassar State University. This technique has a good ability to classify data into groups with similar characteristics and has been proven effective in various previous studies related to historical data-based recommendation systems. Thus, the application of this method is expected to support more systematic collection management. This research is expected to contribute to the development of data-based book recommendation systems in higher education institutions and serve as a reference for other libraries in improving services and formulating more effective collection management policies that meet user needs.

METHOD

Research Methodology

The research method is a technique for searching for, obtaining, and collecting data, both primary and secondary, which will later be used for the purposes of compiling scientific work and analysing factors related to the main issues so that the data obtained will be accurate (Andini et al., 2022).The method used in this study is the data mining method with K-Means Clustering. This research was conducted to implement K-Means Clustering to recommend books in the field of Information Technology and Computer Science at Makassar State University. From the results obtained, we can identify the patterns used as a reference in determining the books recommended to facilitate visitors. There are several methods of data collection used for this research, including:

1. Library Research, which utilises libraries, books, proceedings, or journals as reference media in determining the factors, parameters, and labels used in the research.

2. Observational Research, which is a data collection method by conducting direct field observations to find data used in this research, such as data collection in libraries.

Research Design

The following is the research design that will be used in the study of Book Recommendations Using K-Means Clustering in the Library of the Department of Informatics and Computer Engineering, Makassar State University, as shown in Figure 1 below:
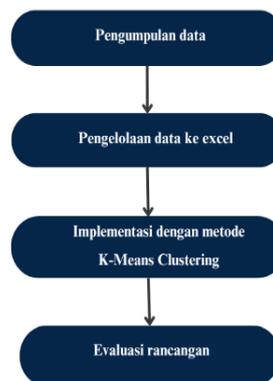

Figure 1. Research Design

The research design flowchart process in Figure 1 is as follows:

Data Collection.

The borrowing data collected in this study is book borrowing data from the Library of the Department of Informatics and Computer Engineering, Makassar State University, for 22 months, from September 2022 to April 2025. The data was obtained through direct observation at the library and is in the form of tables containing information such as book numbers, borrower names, book titles, membership numbers, classes, borrowing dates, and return dates. This data was then focused on book categories and borrowing months, with five book categories processed, namely Programming, Research Methodology, Learning Methods, Education, and Information Technology.

During the processing, the data was cleaned by removing irrelevant information such as borrower names and telephone numbers. Next, the data was grouped based on month and book category, calculating the number of books borrowed in each category every month. This processed data became the input for the clustering process using the K-Means method, with the aim of grouping

book borrowing patterns so that more relevant book recommendations could be provided to library users.

Data Processing

A total of 131 book borrowing records were collected from the Library of the Department of Informatics and Computer Engineering, Makassar State University. The initial data included attributes such as book number, borrower name, book title, membership number, class, date of borrowing, date of return, telephone number, and librarian name. For the purposes of K-Means Clustering analysis, data cleaning was carried out by removing irrelevant attributes, so that the data focused on book titles, dates of borrowing, and book categories. Each book title was then classified into five categories, namely Programming, Research Methodology, Learning Methods, Education, and Information Technology. Next, the number of books borrowed in each category was calculated per month, forming five numerical variables for each month as K-Means input format.

In the experimental stage, each month is represented as one instance with five numerical attributes. The clustering process is carried out by initialising three initial centroids, calculating the Euclidean distance between the data and the centroids, and iteratively updating the cluster membership. The process is stopped at the fourth iteration because there is no change in cluster membership. The final results show the formation of three clusters, namely Cluster 1 with 12 months of borrowing, Cluster 2 with 3 months of borrowing, and Cluster 3 with 7 months of borrowing. The stages of cleaning, category grouping, and conversion to numerical form are necessary because the raw library data does not meet the format required by the K-Means algorithm.

Method implementation

The steps for processing data in K-Means Clustering in this study are as follows (Fakhri et al., 2021).

1. Enter the data to be clustered.

2. Determine the number of clusters.

3. Randomly select data equal to the number of clusters as cluster centres (centroids).

    Formula for calculating the centroid in the K-Means algorithm:

    $$v = \frac{\sum_{i=1}^{n} x}{n}, \text{ for } i = 1,2,3, \ldots, n$$

    Description :

    - v = cluster centroid
    - x_i = object to - i
    - n = the number of objects

4. Calculate the distance between the data and the cluster centre using the Euclidean Distance equation.

    Formula for Calculating Euclidean Distance

    $$d(x, y) = ||x - y|| = \sqrt{\sum_{i=0}^{n}(x_i - y_i)^2}, i = 1, 2, 3, \ldots, n$$

    Description:

    - x_i = object x to- i
    - y_i = object y to- i

- n = the number of objects

5. Recalculate the cluster centres with the new cluster membership. 6. If there is no change in the cluster centres, the clustering process is complete. If not, repeat step 4 until the clusters no longer change.

Design Evaluation

The evaluation of clustering results in this study was conducted using the Silhouette Coefficient method. This method is used to measure how well each object has been grouped into its respective cluster. The Silhouette value indicates how close an object is to other clusters compared to its own cluster. The Silhouette value ranges from -1 to 1, where a value close to 1 indicates that the object fits very well with its cluster and less well with other clusters, while a value close to -1 indicates poor clustering.

Rumus dari *Silhouette Coefficient* (s (ì)) adalah sebagai berikut:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Description:
- a(ì) = the average distance between data i and all other data in the same cluster.
- b(ì)= the average distance between data i and all data from different nearest clusters.

If s (ì) approaching 1, then the data is in the correct cluster; if it is close to 0, the data is on the border between two clusters; and if it is negative, the data may be misclassified. By using the Silhouette Coefficient, the quality of cluster formation can be evaluated objectively to ensure that the clustering results from the K-Means algorithm are optimal and can be used in recommendation systems.

RESULTS AND DISCUSSION

In this study, the book borrowing data at the Library of the Department of Informatics and Computer Engineering, Makassar State University, was clustered using the K-Means Clustering method. The analysed data came from a recapitulation of book borrowing during a specific observation period, which had been transformed into numerical data based on five book categories, namely Programming (Pm), Learning Methods (Mp), Other Learning Methods (Mpl), Education (Ed), and Information Technology (Ti). Each category is represented as a numerical variable that shows the number of book loans each month. The transformation of data into numerical form aims to ensure that all attributes used meet the input requirements of the K-Means algorithm, which works based on the calculation of the distance between objects. By representing borrowing patterns in monthly units, each month is treated as a single unit of analysis that reflects borrowing trends for specific book categories. This approach allows for the identification of relatively homogeneous borrowing patterns within a cluster and more measurable differences in characteristics between clusters.

Collecting Loan Data

Book borrowing data was obtained from the Library of the Department of Informatics and Computer Engineering, Makassar State University, for 22 months (September 2022–April 2025). The data was processed by removing irrelevant attributes (such as the borrower's name) and grouped into five categories: The initial borrowing data can be seen in Table 1.

Table 1. Initial Loan Data

Book Number	Name	Book Title	Class	Loan Date	Return Date
ED-185	indah fadila r	the phisicology of money	PTIK E	3/4/2025	3/6/2025
PM-215	muh zaki mubarak	belajar phyton	PTIK E	3/4/2025	3/7/2025
JK-058	muh haidir nugrah	jaringan komputer	Tekom A	2/3/2025	2/10/2025
HST-034	nurul khiarul hisam	sepotong hati yang baru	PTIK E	10/22/2024	10/20/2024
HST016	nurul khiarul hisam	great expect action	PTIK E	10/23/2024	10/24/2024
HST-106	nurul khiarul hisam	great expect action	PTIK E	10/25/2024	10/24/2024
JK-004	rehan	mikro ruter os	PTIK D	-	-
MP-032	Sasmita	metode peneltian kuantitatif,kulitatif, dan r&d	SC1	8/25/2023	9/5/2023
MPL-024	alfiya pratiwi f	e-learning dan blendd learning	PTIK G	8/29/2023	9/10/2023
MPL-024	alfiya pratiwi f	e-learning dan blendd learning	PTIK G	9/4/2023	9/6/2023

Based on Table 1, the initial data on book borrowing at the Library of the Department of Informatics and Computer Engineering, Makassar State University, is still in the form of raw administrative data that includes attributes such as book number, borrower name, book title, class, and borrowing and return dates, so it cannot be used directly as input for the K-Means Clustering algorithm. Therefore, the initial data was normalized by removing irrelevant attributes, such as borrower names and other administrative information, so that the analysis could focus on book titles and borrowing times. Next, each book title was grouped into five main categories, namely Programming, Learning Methods, Other Learning Methods, Education, and Information Technology, as shown in Table 2. The grouped data was then summarized based on the month of borrowing and book category, resulting in aggregated numerical data suitable for the clustering process and enabling a more structured identification of book borrowing patterns.

Table 2. Data Borrowing After Normalization

Month	Pm	Mp	Mpl	Ed	Ti
September 2022	3	1	1	1	1
August 2023	4	1	2	1	1
September 2023	1	1	2	1	1
October 2023	3	1	1	1	1
November 2023	1	3	1	2	1
December 2023	1	3	1	1	1
January 2024	2	1	1	1	1
February 2024	1	2	1	1	1
March 2024	1	2	1	1	3
April 2024	1	1	4	1	1

Based on the table, the borrowing data for 22 months from 2022 to 2025 at the Library of the Department of Informatics and Computer Engineering, Makassar State University, was processed by

dividing the frequency of book borrowing each month into each book category. Each row in the table represents one month of borrowing, while each column shows the number of books borrowed in a particular category in that month. Thus, the borrowing data for 22 months was divided into the number of times each month that books were borrowed in each book category, forming aggregated numerical data that describes the variation in borrowing intensity between months and between categories. This data format allows each month to be treated as a single unit of analysis with several numerical variables, which are then used as input in the clustering process to systematically identify book borrowing patterns. Clustering Process with K-Means

1. In this study, the data was clustered into 3 clusters (K=3). These clusters will later be categorized as less desirable, desirable, and highly desirable.
2. Initialize Initial Centroid
   In determining the initial centroid value, it can be taken randomly from the data of the sample used in the study. This initial centroid is the center point of the first cluster. The centroid for C1 is taken from the second month's data sample (K2), C2 is taken from the fifth month's data sample (K5), and C3 is taken from the fifteenth month's data sample (K15). The following is the centroid form that has been translated into Table 3.

Table 3. Initial Centroid

Initial Centroid	Pm	Mp	Mpl	Ed	Ti
C1	4	1	2	1	1
C2	1	3	1	2	1
C3	3	1	1	3	3

3. Clustering Iteration Results and Borrowing Pattern Analysis

The iteration process was stopped at the 4th iteration because the centroid was stable. The results of the 4th iteration can be seen in Table 3 as follows:

Table 4. Iteration Results to-4

C1	C2	C3
1,073674894	3,958114029	2,792117184
1,775606313	4,654746681	3,350296972
1,285604052	3,559026084	2,867837129
1,073674894	3,958114029	2,792117184
2,546784465	1,632993162	2,965791356
2,342244887	1,527525231	3,26390276
0,807946643	3,559026084	2,606898228
1,67912014	2,449489743	2,867837129
2,611406603	2,708012801	1,911298309
2,579297923	4,546060565	3,883507777
0,807946643	4,082482904	2,867837129
1,52297224	2,645751311	2,941608612
6,336621953	3	5,982969026
0,3908679801	3,696845502	2,687841103
2,674467756	3,651483716	1,169336111
2,037836543	3,872983346	1,628822036
0,807946643	4,082482904	2,867837129
3,186342382	3,958114029	1,22890361
3,02535581	4,434711565	1,22890361
3,977366521	3,785938897	2,020305089
4,544532735	6,027713773	3,26390276
1,073674894	3,958114029	2,792117184

Based on the results of the clustering process using the K-Means algorithm, distance calculations and cluster membership updates were performed iteratively until convergence was achieved. From the results of the fourth iteration, the final centroid can be seen in Table 5, which shows that there were no further changes in the cluster center values in subsequent iterations. Each column in the table represents the centroid of each cluster, namely C1, C2, and C3, while the numerical values in each

row indicate the position of the centroid in each variable used in the analysis. These final centroids describe the average characteristics of the book borrowing data in each cluster, so that the differences in values between centroids reflect the variation in borrowing patterns between clusters that are formed and form the basis for determining the results of the book borrowing data clustering.

Table 5. Final Centroid

C1	C2	C3
2,25	1	2
1,166666667	4,333333333	1,285714286
1,75	1	1,285714286
1	1,333333333	2,428571429
1	1,666666667	3,142857143

Based on the results of the fourth iteration, analysis of borrowing patterns from the application of the K-Means Clustering algorithm shows the formation of three main clusters from book borrowing data for approximately 22 months at the Library of the Department of Informatics and Computer Engineering, Makassar State University. The table shows the average borrowing value for each book category for each cluster, represented by C1, C2, and C3. Each column shows the characteristics of the borrowing pattern in one cluster, while each row represents the book category analyzed. The differences in average values between clusters reflect variations in the intensity of book borrowing in each category, so these results illustrate the existence of relatively homogeneous borrowing patterns within the same cluster and differences between clusters, which are then used as a basis for analysis and determining book recommendations.

- Cluster 1 Programming Books (Highly sought after: Cluster 3 is dominated by programming and computer science books with fairly high and stable borrowing rates in September 2022, August 2023, January 2024, and February 2024. This illustrates the routine needs of students majoring in Information Technology and Computer Science for materials related to their field of study.
- Cluster 3 (Research Methodology): Cluster 3 is dominated by books on Research Methodology, with a significant spike in borrowing in March and April 2024, coinciding with the period of thesis or final project preparation.
- Cluster 2 (Less popular): Cluster 2 includes books on Education and Learning Methods, which are only popular during certain months, such as November and December 2023, indicating low interest in non-technical and pedagogical books.

Cluster Evaluation

The quality of clustering is measured using the Silhouette Coefficient to assess the degree of suitability of each data object to the formed clusters. Most of the data has a positive Silhouette value with an average of 0.35, indicating that the resulting clusters have a fairly good degree of separation and that most data objects are in the appropriate clusters. This value indicates that the distance between clusters is relatively greater than the distance between data within the same cluster, so that the clustering structure can be identified quite clearly. However, some data, such as K6 ($S(i) = -0.05$) and K8 ($S(i) = -0.47$), indicate the potential for less than optimal clustering. This condition shows that these data objects have almost the same proximity or are even closer to other clusters, which is likely due to the selection of initial centroids that are not yet fully ideal. Nevertheless, overall, the results of this evaluation still show that the K-Means Clustering method is capable of producing a fairly representative grouping and reinforces the validity of its use in identifying book borrowing patterns in libraries.

CONCLUSION

This study aims to develop a book recommendation system based on borrowing patterns using the K-Means Clustering algorithm to improve the relevance and efficiency of collection searches at the Library of the Department of Informatics and Computer Engineering, Makassar State University.

Based on the results and discussion presented, the application of the K-Means algorithm to book borrowing data for approximately 22 months successfully grouped the data into three main clusters with different borrowing pattern characteristics. The clusters formed were able to represent variations in borrowing intensity based on book categories, thus providing a more structured picture of library users' trends and needs. The evaluation of the clustering quality using the Silhouette Coefficient showed a positive average value, indicating that the resulting cluster structure was quite good and in line with the initial research objectives. Thus, the results of this study show the suitability between the issues raised in the introduction and the findings obtained in the results and discussion sections. Furthermore, these findings open up opportunities for further development, both in refining the book recommendation system based on borrowing data and in applying more adaptive clustering methods in further research, in order to support more effective and sustainable library collection management.

ACKNOWLEDGMENT

AUTHOR CONTRIBUTION STATEMENT

WES designed the study and supervised data collection. WES conducted data preprocessing, clustering implementation, and result interpretation. SM contributed to methodological review and analytical validation. WES drafted the manuscript. All authors reviewed and approved the final manuscript.

AI DISCLOSURE STATEMENT

The authors declare that AI tools were utilized only for language polishing and editorial support. AI tools did not affect the research methodology, data analysis, or conclusions. Responsibility for the integrity of the work lies entirely with the authors.

REFERENCES

Abdurrachman, N., & Chahyati, D. (n.d.). *Peningkatan kualitas citra bawah air menggunakan gan dengan mekanisme residual dan attention*.

Andini, Y., Hardinata, J. T., & Purba, Y. P. (2022). Penerapan Data Mining pada Tata Letak Buku Di Perpustakaan Sintong Bingei Pematangsiantar dengan Metode Apriori. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, *7*, 13–18. http://dx.doi.org/10.30645/jurasik.v7i1.410.g387

Asminah, A. (2022). Sistem Penentuan Penambahan Koleksi Buku di Perpustakaan Menggunakan Metode K-Means Clustering. *Journal of Information System Research (JOSH)*, *4*(1), 330–338. https://doi.org/10.47065/josh.v4i1.2383

Asriningtias, S. R., Wulandari, E. R. N., Persijn, M. B., Rosyida, N., & Sutawijaya, B. (2023). Identification of Public Library Visitor Profiles using K-means Algorithm based on The Cluster Validity Index. *Sinkron*, *8*(4), 2615–2626. https://doi.org/10.33395/sinkron.v8i4.12901

Bin Samer, T., & Darujati, C. (2023). Acceleration and Clustering of Liver Disorder Using K-Means Clustering Method with Mahout's Library. *Journal of Systems Engineering and Information Technology (JOSEIT)*, *2*(2), 37–44. https://doi.org/10.29207/joseit.v2i2.5334

Dacwanda, D. O., & Nataliani, Y. (2021). Implementasi k-Means Clustering untuk Analisis Nilai Akademik Siswa Berdasarkan Nilai Pengetahuan dan Keterampilan. *AITI: Jurnal Teknologi Informasi*, *18*(2), 125–138. https://doi.org/10.24246/aiti.v18i2.125-138

Elmustian, & Firdaus, M. (2024). Filologi, Transformasi Teks, dan Filsafat Pendidikan: Strategi Pelestarian Budaya dalam Konteks Pendidikan Kontemporer. *Indonesian Research Journal on Education*, *4*(4), 1073–1081. https://doi.org/10.31004/irje.v4i4.1213

Fadlina, J., Utami, R., & Sitinjak, N. M. (2024). Pengelompokkan Buku dan Rekomendasi Buku Menggunakan K-Means Clustering pada Dinas Perpustakaan dan Kearsipan Kota Medan. *Jurnal Widya*, *5*(2), 1045–1058. https://doi.org/10.54593/awl.v5i2.288

Fakhri, D. A., Defit, S., & Sumijan. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi dan Teknologi*, 160–166. https://doi.org/10.37034/jidt.v3i3.137

Fathuroh, S. (2023). Metode K-Means Clustering Dalam Optimalisasi Kinerja Dosen Pendamping Akademik Pada Program Kampus Merdeka. *Jurnal Sistim Informasi Dan Teknologi*, *5*(2), 55–60. https://doi.org/10.37034/jsisfotek.v5i2.172

Gustin, S., Ramdhan, W., & Kifti, W. M. (2022). Teknik Data Mining Menggunakan Metode K-Means Untuk Mengcluster Dan Pencarian Buku Di Perpustakaan Daerah Kabupaten Asahan. *JUTSI (Jurnal Teknologi Dan Sistem Informasi)*, *2*(3), 195–204. https://doi.org/10.33330/jutsi.v2i3.1901

Hasanah, N. N., & Purnomo, A. S. (2022). Implementasi Data Mining Untuk Pengelompokan Buku Menggunakan Algoritma K-Means Clustering (Studi Kasus: Perpustakaan Politeknik LPP Yogyakarta). *Jurnal Teknologi Dan Sistem Informasi Bisnis*, *4*(2), 300–311. https://doi.org/10.47233/jteksis.v4i2.499

Hermawan, B., & Risparyanto, A. (2025). *Transformasi Perpustakaan Perguruan Tinggi dan Peran Pustakawan di Era Artificial Intelli- gence (AI): Penelusuran Data Sekunder Tahun 2018–2024*.

Iryani, L. (2020). Penerapan Datamining Menentukan Minat Baca Mahasiswa Di Perpustakaan Universitas Bina Darma Palembang Menggunakan Metode Clustering. *INTECOMS: Journal of Information Technology and Computer Science*, *3*(1), 82–89. https://doi.org/10.31539/intecoms.v3i1.1251

Jannah, M., Yumami, E., Julianto, A., & Rahmi, E. (2025). Sistem Rekomendasi Buku di Perpustakaan Menggunakan Machine Learning dan Algoritma Apriori. *Jurnal Sains dan Informatika*, *11*(1), 21–29. https://doi.org/10.34128/jsi.v11i1.1868

Maesaroh, I., Mujib, A., & Kholis, N. (2025). Navigating the Digital Shift: Assessing Skills, Training, and Resources for Virtual Reference Librarians. *Khizanah Al-Hikmah : Jurnal Ilmu Perpustakaan, Informasi, Dan Kearsipan*, *13*(1), 54–68. https://doi.org/10.24252/v13i1a5

Mugnia, A., & Mutoffar, M. M. (2024). *Implementasi Algoritma Apriori Untuk Sistem Rekomendasi Buku Pada Perpustakaan Digital*. *11*(1).

Mutiarani, R., Yoanda, S., & Gunaidi, A. (2022). Analisis Kendala Penerapan Otomasi Perpustakaan Di Perpustakaan Perguruan Tinggi Bina Sriwijaya Palembang. *JIPI (Jurnal Ilmu Perpustakaan dan Informasi)*, *7*(2), 271. https://doi.org/10.30829/jipi.v7i2.12824

Nasir, J. (2020). Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means. *Jurnal SIMETRIS*, *11*(2), 1–13.

Nurcahyani, H. (2023). Penelitian Strategi Pengembangan Koleksi Di Perpustakaan Pada Google Scholar: Sebuah Narrative Literature Review. *Jurnal Pustaka Budaya*, *10*(1), 2442–7799. https://doi.org/10.31849/pb.v10i1.11275

Pandita, A. (2017). *Pengaruh Kualitas Pelayanan Terhadap Kepuasan Pemustaka Di Upt Perpustakaan Universitas Negeri Makassar*. Universitas Islam Negeri Alauddin Makassar.

Putri, A. D., & Rahmah, E. (2019). Persepsi Mahasiswa terhadap Instrumen Musik di Perpustakaan Universitas Bung Hatta dalam Kenyamanan Membaca. *Ilmu Informasi Perpustakaan dan Kearsipan*, *8*(1), 27. https://doi.org/10.24036/107294-0934

Sutanto, Y., Amin, B. A., Setyadi, H. A., & Purnama, B. E. (n.d.). *Prediksi harga perumahan menggunakan metode principal component analysis dan random forest regresi*.

Ulfah, M., & Irwaty, A. S. (2022). Penerapan Data Mining Clustering Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan. *Fidelity : Jurnal Teknik Elektro*, *4*(3), 62–68.

Zabidi, A. F. (2024). Penerapan Algoritma K-Means untuk Pengelompokan Koleksi Perpustakaan dengan Data Mining. *Media Jurnal Informatika*, *16*(2), 233–242. https://doi.org/10.35194/mji.v16i2.4814

Zeng, Z., Sun, S., Li, T., Yin, J., & Shen, Y. (2022). Mobile visual search model for Dunhuang murals in the smart library. *Library Hi Tech*, *40*(6), 1796–1818. https://doi.org/10.1108/LHT-03-2021-0079