

Data Mining in Academic Libraries: A Systematic Review

Shahnaz khademizadeh

Associate Prof., Department of Knowledge and Information Science, Faculty of Educational Sciences and Psychology, Shahid Chamran University of Ahvaz; Ahvaz, Iran.

Corresponding Author: s.khademi@scu.ac.ir
ORCID iD: <https://orcid.org/0000-0003-4494-7709>

Fatemeh Rafieinasab

1. PhD. Candidate, Department of Knowledge and Information Science, Faculty of Educational Sciences and Psychology, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

2. Librarian Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.

fmhrafiei@gmail.com
ORCID iD: <https://orcid.org/0000-0000-0000-0000>

Natarajan Radhakrishnan

Professor, Department of Library and Information Science, Periyar University, Salem, India.
mkuradha@gmail.com ORCID iD: <https://orcid.org/0000-0002-2046-094X>

Received: 26 June 2022

Accepted: 01 October 2022

Abstract

Libraries, as organizations that typically deal with significant amounts of data, can use data mining techniques to make informed decisions based on discovered knowledge to optimize their services. Finding valuable hidden information or patterns in large amounts of data is highly important in library management and services. Academic libraries, as centers that continuously serve the scientific community, have been able to provide better services to their users by utilizing results from data-based studies for better resource and budget management or by using the information within their organizations. This study was conducted to systematically collect and analyze data mining studies in university libraries with a focus on applications, subject areas, techniques, and software used. Articles were retrieved from two major scientific databases (Scopus, Web of Science, and Google Scholar) following PRISMA's preferred guidelines for literature review, and were screened and selected based on their relevance, techniques, software used, and subject areas.

Keywords: Data mining, Knowledge Discovering, Data Management, Academic Libraries.

Introduction

Information about the user and behavior patterns is an organization's most valuable asset. Librarians can provide services more systematically by improving awareness in this regard. Data mining can help library managers make more accurate and successful decisions using documentation and decision-making. Data mining can extract explicit and implicit knowledge from large databases for various purposes, such as decision-making, process control, planning, and service promotion. Data mining can also quickly review and store knowledge while facilitating research (Han, Kamber & Pei, 2011). Data mining can be defined as extracting knowledge from the volume of many data (Berry & Linoff, 1997; Duan & Wang, 2021). According to Han, Kamber and Pei (2011), data mining is the crucial phase of knowledge

discovery to extract knowledge from a large amount of data. It aims to analyze exploratory data, discover patterns, rules, and algorithms, predict modeling, and find deviations. As some authors, like Chatfield, have argued, data mining is related to "discovering knowledge" from databases. (Chatfield, 1995); in this case, data mining means extracting valuable and potentially useful knowledge from the data (Ansari, Vakili Mofrad, Mansoorizadeh & Amiri, 2019). This process uses statistical approaches, machine learning, and set training to extract specific information for management in the future (Bussaban & Kularbphettong, 2014).

Data was considered a data process after specifying the objectives and requirements, editing data, and an initial strategy for achieving the goals (Hendrickx, Cule, Meysam, Naulaerts, Laukens & Goethals, 2015). After analyzing them, the initial approach is discovered, and the initial quality of the data can be quickly evaluated. After the data is prepared, the primary data set is extracted, and appropriate variables are analyzed. Then, the results are created as appropriate patterns or models based on the applied techniques (Chapman & *et al.*, 2000). Data mining can generally be defined as using analytical tools to discover valuable and unknown relationships and patterns in a large data set. Thus, data mining involves data collection, management, analysis, and prediction.

Long and Wu (2012) point out that academic libraries, the central pillar of educational and research activities, use database technologies for facilitating services. These libraries are the centers that constantly deal with many data. Data mining technology in the digital information age, due to its functional impact on libraries in data research, how to efficiently extract data and features and provide quality personal services to users, is one of the crucial issues in academic libraries (Duan & Wang, 2021). Libraries and information centers increase their services statistically and qualitatively due to the massive development of new information technology. On the other hand, established tools and techniques may not respond rapidly and adequately to users aside from time and money when access to a considerable volume of information increases. Tools and procedures like data mining are becoming increasingly important to access data and extract valuable knowledge efficiently. Their most significant is gaining knowledge from massive amounts of data to enhance the quality of management choices (Uppal & Chandwani, 2013); within this volume of data, patterns, and relationships between different parameters are unseen. Therefore, modern data mining technology is essential for information centers and library management systems (Zhang, 2014).

When data mining is used to manage and understand users' information needs, helpful information can be uncovered and provided (Suresh, Anand, Vianesh & Mohammad, 2018). As a suitable tool for preceding users' behavior for knowledge-based decisions, this cutting-edge and dependable technology may make libraries' and information centers' jobs easier (Mishra & Mishra, 2013). Data mining can provide users with relevant books by analyzing the lending process, leading to improved library resource quality and facilitation of library management (Yi, Chen & Cong, 2018). Many such studies (Chen & Chen, 2007; Tsuji & *et al.*, 2012; Jomsri, 2014; Liu, 2018; Yi, Chen & Cong, 2018), and data mining techniques have been used in recommender and intelligent systems in libraries. In some cases, research data can be combined with library data, and the results can be examined more accurately. O'Leif believed it would be difficult to show value until foreign data was combined with library data, creating partnerships. Researchers will benefit from other experiences to effectively use data mining (Oakleaf, 2010).

Academic libraries have used different methods and techniques for various uses and thematic areas of data mining. Libraries faced with a large volume of data, using data mining

techniques and relying on the results of such studies can be considered proper support for library management decisions. It is possible to predict the values of other variables and use them in decision-making. Therefore, it is inevitable to refine and compare large amounts of data using data mining techniques and use them to produce new added value. There is no investigation conducted on this subject for systematic review. The current study focused on analyzing the status of research conducted in this field, uses, and techniques using data in academic libraries about all the uses and thematic areas of data and value-generating. Therefore, this paper seeks to answer the following review questions:

- What is the use of data mining in academic libraries?
- What are the subject areas of data mining covered in academic libraries?
- What is the data mining software used in academic libraries?
- What are the data mining techniques used in academic libraries?

Material and Methods

This article examines scholarly data mining in academic libraries using PRISMA guidelines. PRISMA helps researchers to report a helpful set of minimum evidence-based cases published for critical evaluation (Moher, Liberati, Tetzlaff, Altman & the PRISMA group, 2009). Initially, PRISMA was created for health and healthcare-related studies. Systematic reviews of scientific research guidelines have been successfully demonstrated in recent library information management and information science studies. (Ashiq, Usmani & Naeem, 2022; Safdar, Batool & Mahmood, 2020).

We start by specifying the research questions being addressed, and then we detect relevant literature to augment our understanding of these questions in a structured way. In this section, we present the details of the following procedure.

In this systematic review, four stages (phases) have been identified using the PRISMA diagram: identification, screening, eligibility, and inclusion of studies; the first stage or identification of all studies related to data mining in academic libraries was collected from Scopus and ISI databases. Then, in the second stage, only original research papers in English were used for the present study. Other types of documents, such as review articles, reports, book chapters, and congress papers, were excluded from this study. If there was no access to the full text of the article, they were excluded from the analysis process. In the third stage of eligibility, the titles and abstracts of the articles were reviewed. Finally, by reviewing the full text of the article and final confirmation, the articles which deserved to be present in the systematic review in the final stage were evaluated.

A comprehensive search strategy was designed to extract this field's most relevant studies. Two databases (Scopus, Web of Science) were selected for this study. By August 1, 2021, we searched two databases using a search strategy. We searched keywords for titles, abstracts, and keywords in the Scopus database and titles and topics in the WOS database. A total of 114 publications were found articles without date of publication limitation. We searched the WOS database on August 1, 2021, and located 109 published articles. All three databases' relevant data were downloaded and entered into the Mendeley software. Data were retrieved from Scopus and Web of Science. This process was repeated to ensure that the data was accurate. The review process (screening, eligibility, and inclusion of studies) was conducted two times. Data extraction was completed then, followed by a re-examination of the articles. All of these

steps were performed and repeated at each stage to ensure the accuracy and validity of the inclusion. The following keywords and search strategies were used to search and retrieve articles from the databases.

("Academic library" OR "university library" OR "campus library" OR "school library" OR "faculty library" OR "faculty-librar*" OR "scholar library*") AND ("datamining" OR "data mining" OR "data-mining OR bibliomining OR bibliomin* OR Pattern recognition OR Artificial intelligence OR Data analysis" OR "Data visualization" OR "Digital forensics" OR "Knowledge discovery" OR "Association rules" OR "Text analysis" OR "Text mining" OR "Web mining").

In this article, to evaluate the quality of the analyzed articles (Batten, Oakes & Alexander, 2014; Ashiq, Usmani & Naeem, 2022; Safdar, Batool & Mahmood, 2020), the checklist has been used, and this checklist has 45 questions, and the different parts of the documents are evaluated, and scored in three spectrums of "yes there" (2 points), somewhat (1 score) and "unreported" (zero points) are scored, and the overall percentage is extracted. In this study, qualitative analysis was performed by two coders, and the agreement of evaluators based on the Miles and Haberman formula (1994) was 89% (Miles & Huberman, 1994).

Results

From the total of research articles related to the subject and according to the criteria, 223 abstract articles have been reviewed. And 109 articles were analyzed from Web of Science, and 114 from Scopus. In this study, 53 articles have been excluded from the review of the title and abstract due to non-compliance with the specified criteria, and 50 sources have been entered into the process to review the full text. The criteria and relevance to the subject of 32 documents were taken for systematic analysis. (Figure 1). Two experts were consulted about identifying, searching, and retrieving resources related to the research subject and the criteria for including systematic review, and the expert's suggestions were applied. It is necessary to explain that the search process was conducted separately To validate the research stages. A systematic search output based on the PRISMA review process approach for entering final resources is reviewed below.

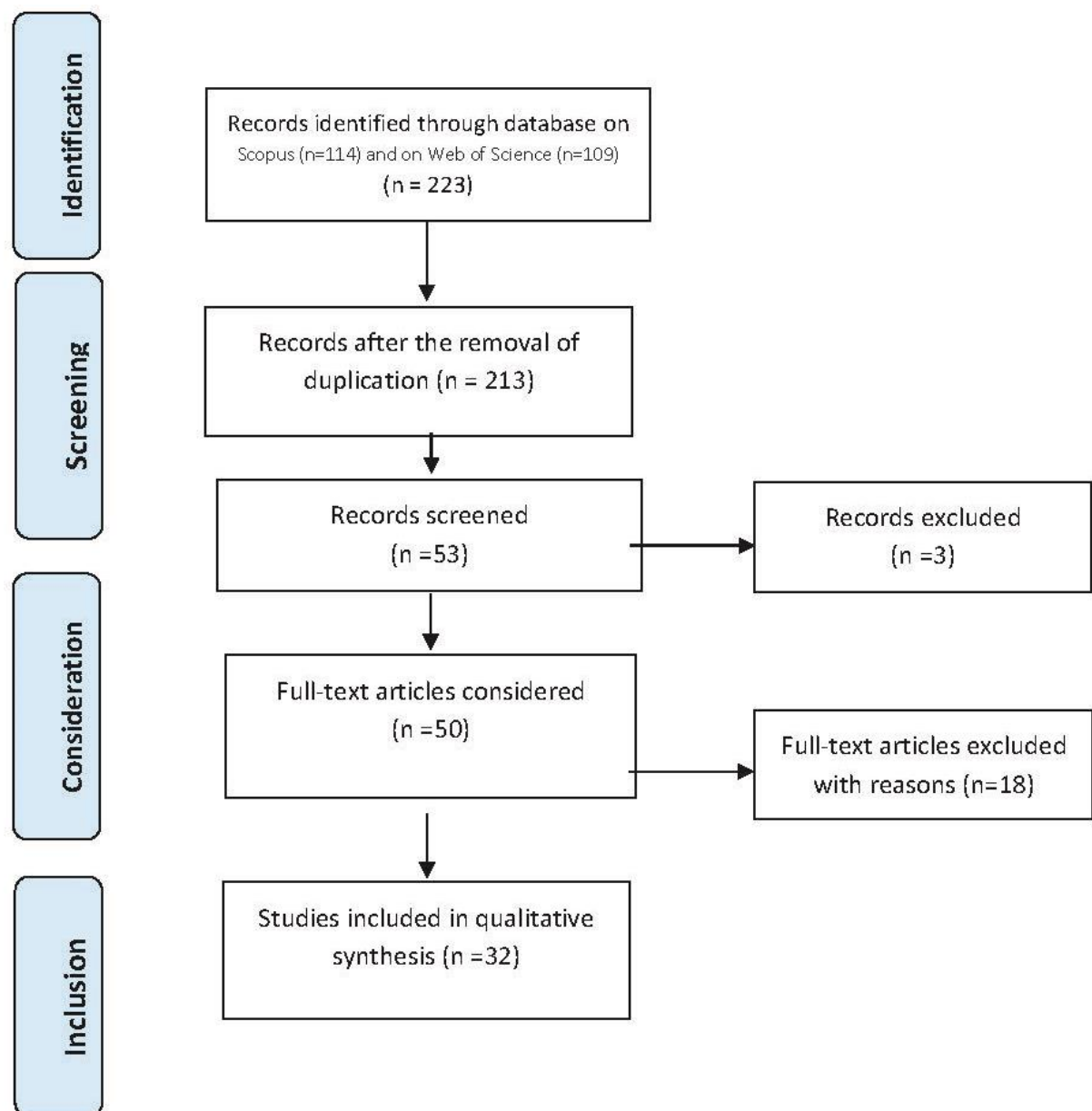


Figure 1: Steps Involved in Selecting and Excluding Articles for the Research (Moher, Liberati, Tetzlaff, Altman & the PRISMA group, 2009)

Figure 2 presents the number of articles in this study according to the year of publication. Based on the results, the highest number of articles entered into the analysis in this study is related to 2018, and the lowest are 2006, 2008, and 2015.

The specifications of the articles included in this study are shown in Table 1.

Table 1

The studied research specifications

Row	Code	Title	Writers	Year of publication
1	K3	Research on the Service Mode of the University Library Based on Data Mining	Duan & Wang	2021
2	K7	Mining and Analyzing Patron's Book-Loan Data and University Data to Understand Library Use Patterns	Silwattananusarn & Kulkanjanapiban	2020
3	K8	Toward Effective Planning and Management Using Predictive Analytics Based on Rental Book Data of Academic Libraries	Iqbal, Jamil, Ahmad & Kim	2020
4	K10	An Analysis of Research Trends on Data Mining in Chinese Academic Libraries	Huancheng, Tingting & Rocha	2019
5	K11	Library book modeling data using the association rule method with apriori algorithm in determining book placement and analysis of book loans	Prehanto, Indriyanti, Permadi, Vitadiar & Jayanti	2019
6	K13	Mining reference chat transcripts to analyze noise complaints	Vance	2018
7	K14	Understanding occupancy and user behavior through Wi-Fi-based indoor positioning	Wang & Shao	2017
8	K15	Data Mining of University Library Management Based on Improved Collaborative Filtering Association Rules Algorithm	Liu	2018
9	K16	Research on Library Recommendation Reading Service System Based on Adaptive Algorithm	Huang, Li & Xiao	2018
10	K17	Discovering Research topics from library electronic references using latent Dirichlet allocation	Fang, Yang, Gao & Li	2018
11	K18	Technical services and the virtual reference desk: Mining chat transcripts for improved e-resource management	Kimbrough	2018
12	K19	Analysis of Academic Libraries' Facebook Posts: Text and Data Analytics	Al-Daihani & Abrahams	2018
13	K42	Investigating Factors Affecting Library Visits by University Students Using Data Mining	Puarungroj, Pongpatrakant, Boonsirisumpun & Phromkhot	2018
14	K21	Text mining in business libraries	Anderson & Craiglow	2017
15	K22	Understanding occupancy patterns and improving building energy efficiency through Wi-Fi-based indoor positioning	Wang & Shao	2017
16	K23	Linking special collections to classrooms: A curriculum-to-collection crosswalk	Yaco, Brown & Konrad	2016

Row	Code	Title	Writers	Year of publication
17	K24	A Text Mining Analysis of Academic Libraries' Tweets	Al-Daihani & Abrahams	2016
18	K29	Usage of E-resources: VIRTUAL Value of Demographics	Samson	2014
19	K31	The E-book Power User in Academic and Research Libraries: Deep Log Analysis and User Customisation	Ahmad, Brogan & Johnstone	2014
20	K37	Data-mining: Improvement of university library services	Will	2006
21	K38	Bibliomining for automated collection development in a digital library setting: Using data mining to discover web-based scholarly Research works	Nicholson	2003
23	K39	Decision support for the academic library acquisition budget allocation via circulation database mining	Kao, Chang & Lin	2003
23	G1	E-book data mining: real information behavior of university academic community	Rafique, Ameen & Arshad	2021
24	B6	Mining library and university data to understand library use patterns	Renaud, Britton, Wang & Ogihara	2015
25	B12	Reference question data mining a systematic approach to library outreach	Finnell & Fontane	2010
26	G2	Using data mining techniques to predict user's behavior and create recommender systems in the libraries and information centers	Ansari, Vakili Mofrad, Mansoorizadeh & Amiri	2020
27	G16	Classifying and predicting students' performance using improved decision tree C4.5 in higher education institutes	Sadiq & Ahmed	2019
28	A1	Using data mining to improve digital library services	Kovacevic	2010
29	B9	Negotiating a Text Mining License for Faculty Researchers	Williams	2014
30	B13	Using adaptive resonance theory and data-mining techniques for materials recommendation based	Tsai & Chen	2008
31	S7	ASSESSMENT OF LIBRARY USERS' FEEDBACK USING MODIFIED MULTILAYER PERCEPTRON NEURAL NETWORKS	Nandha Kumar & Christopher	2017
32	B2	Four Categories of Academic Libraries: A Cluster Analysis Based on Collections, Expenditures, and Circulation per Student Data	Lund	2017

Use of Data Mining in Academic Libraries

Six categories of using data mining in academic libraries were identified; (discovering hidden patterns and discovering the rules, predicting and presenting models, classification, clustering, identifying research topics and trends, and analyzing social networks) and explained in the following. See Figure. 4.

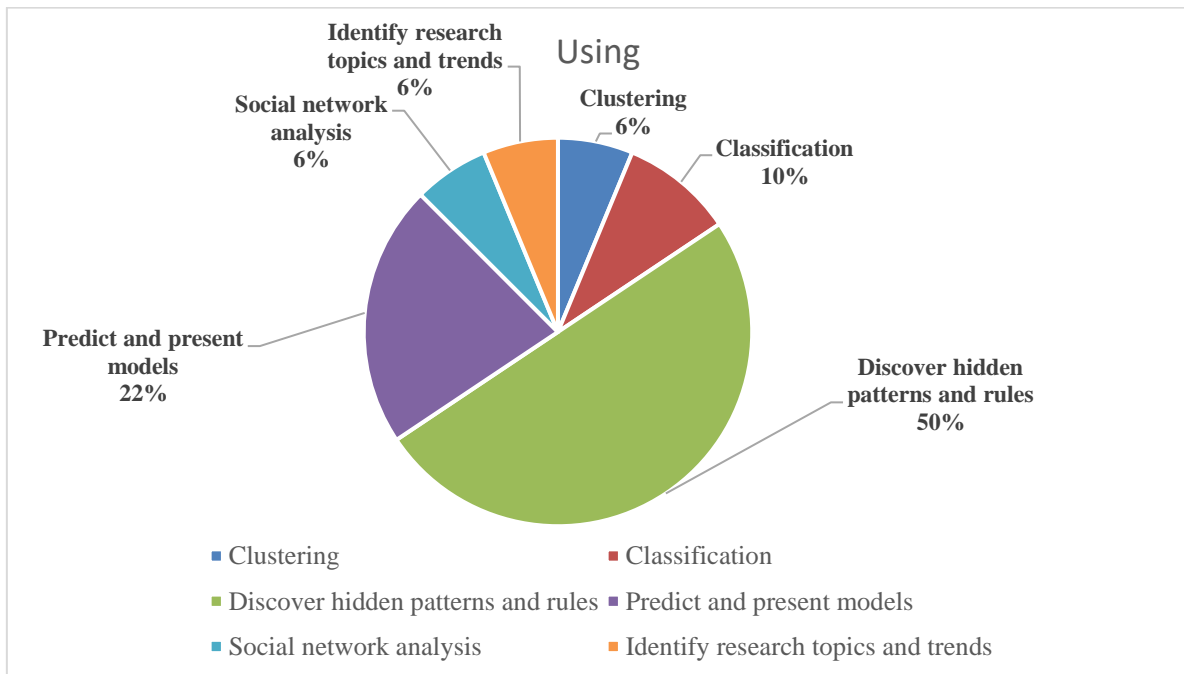


Figure 4: uses of Data Mining

Discover hidden patterns and rules account for 50% of studies in this field; one of the attractive uses of data is discovering hidden patterns between the data. In academic libraries, it is possible to discover the pattern of using electronic resources through the data extracted from library software about the circulation of library resources, to discover the pattern and provide a model to determine users' priorities. Due to the limited resources and facilities of academic libraries, data mining helps library managers to prioritize more important criteria in their programs and improve their clients' satisfaction by improving the quality of management decisions on time and at a lower cost. A set of library materials being recommended by users is the foundation of recommendation systems, a data filtering process and a personal service provider. Managers and policymakers may coordinate their choices, provide information resources, and effectively employ them in libraries and information centers by using the patterns that have been found. (Bracke, 2004; Hoffman-Apitius, Younesi & Kasam, 2009; Scarnò, 2010; Samson, 2014; Renaud, Britton, Wang & Ogihara, 2015; Puarungroj, Pongpatrakant, Boonsirisumpun & Phromkhot, 2018; Rafique, Ameen & Arshad, 2021)

Predicting and presenting the model-which includes 22% of the articles in this study- is one of the other functions of data that use this technique in academic libraries. It can be estimated by using existing data, users' behavior and performance, and the process of using electronic and library resources to a large extent. Such studies can lead to effective services and improve the quality of interaction between librarians and users.

Another use of data mining in academic libraries is classifying members, libraries, or other items based on similar variables and characteristics, such as behavior. Identifying fixed client's

and dating clients according to their similar behavior in searching, lending information resources, and analyzing the client's credibility, will help identify the main factors in providing services. Studies revealed that the categorization method could be an excellent reference for planning in academic libraries (Nandha Kumar & Christopher, 2017; Sadiq & Ahmed, 2019; Tsai & Chen, 2008). Research findings have brought that academic libraries and their management cannot be considered homogeneous. These institutions can be pretty distinct in terms of budget and use. The efficiency of cluster analysis of academic libraries is to assess trends in a wide range of library data to help plan and manage libraries better (Lund, 2017).

The importance of using data mining and text mining methods in understanding the social information set of academic libraries is to help make decisions and strategic planning processes. This function helps academic libraries to use social media strategies to increase participation and strengthen marketing (Al-Daihani & Abrahams, 2016; Al-Daihani & Abrahams, 2018). Studies showed that the topics discovered by Data mining are consistent with topics outlined by human experts. Therefore, this data mining provides new knowledge and demonstrates a method for discovering the subject at a lower cost and more efficiency than current methods (Fang, Yang, Gao & Li, 2018).

The subject areas of data mining in academic libraries

Data mining studies have been conducted in various subject areas in academic libraries. The most common domains used in this field were analyzing reference questions, recommending systems (to suggest relevant or interesting resources), predicting user behavior and performance, analyzing library trust data, and discovering the pattern of occupying the library space. See Figure. 5.

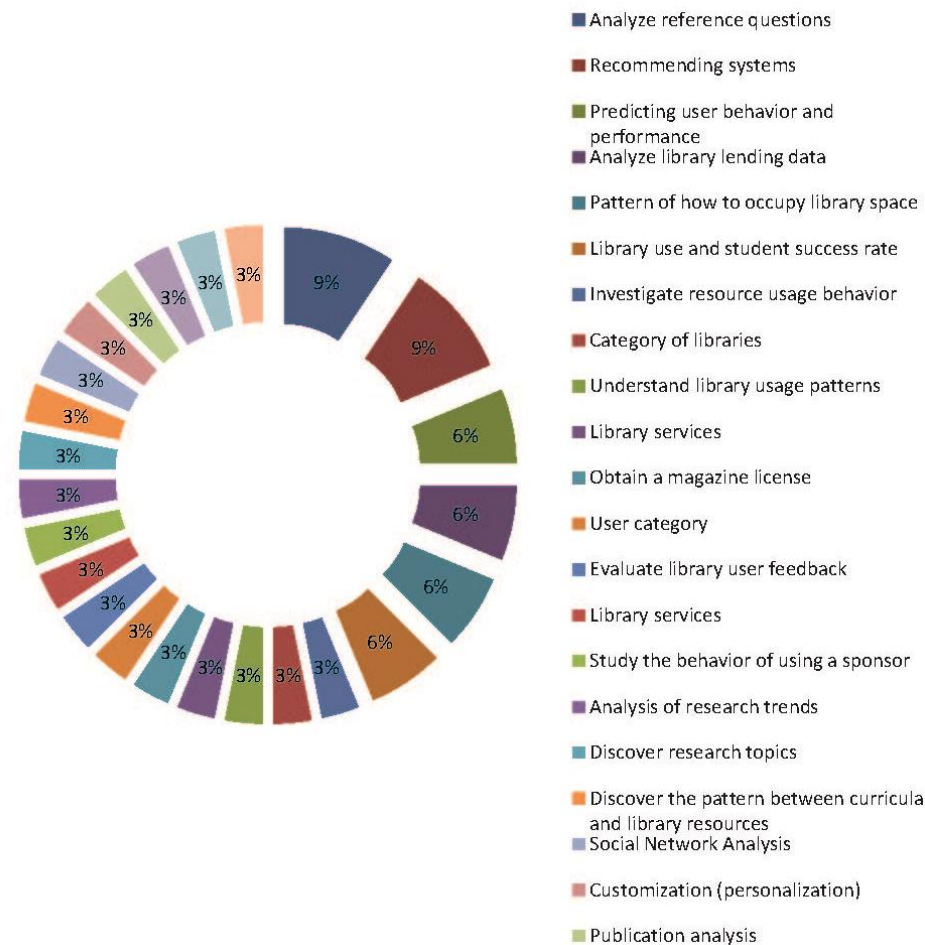


Figure 5: Subject Areas of Data mining in Academic Libraries

Data mining developments used in Academic libraries

To perform data mining on the data of academic libraries, a range of data mining techniques can be used. In these articles, communication rules, text mining, clustering, and neural network were the most commonly used techniques. Communication rules identify the correlation between two or more items by identifying a hidden pattern in the data set. Hence, it is also called relationship analysis. This technique has been used to discover usage, pattern, and modeling in academic libraries research. Clustering is widely used in discovery, especially in analyzing Patron's Book-Loan Data transactions. Model the relationship between the input and output, the neural network technique used. This method is used for classification, regression analysis, data processing, etc. (Han, Kamber & Pei, 2011). The neural networks are more suitable for situations without identifying how things happen, and only the outcome and output are essential.

Data mining software used in Academic libraries

Interestingly, 21 different software have been used for 32 studies (Figure 6), among which R software, Matlab, with 5 cases, and Excel and Weka software, with four and three cases, had the highest usage among different data fields. It can be inferred that in typical use, this software's features are not unique.

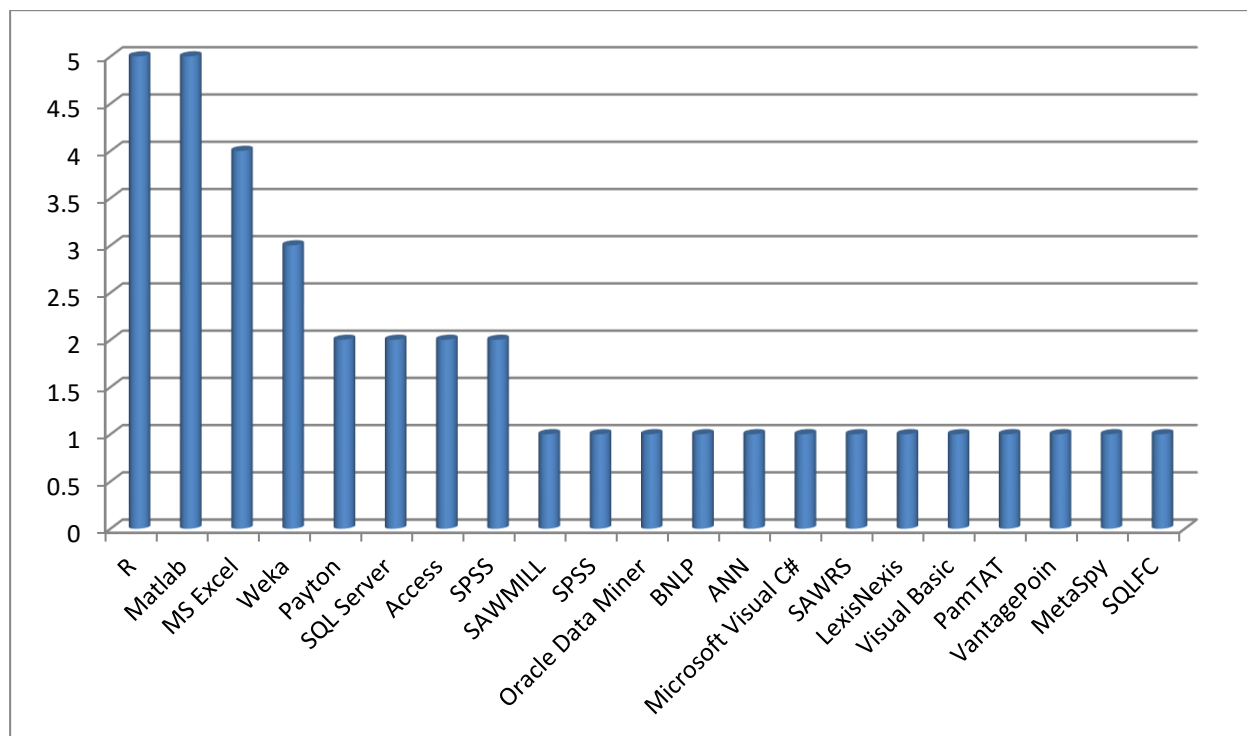


Figure 6: University library Data Mining Software

Discussion

This study is a systematic review of data mining in academic libraries, which was conducted to investigate using techniques, software, and thematic areas based on 32 select studies. Articles were studied from different aspects, such as using technique, software, and thematic areas. This study revealed that data mining in academic libraries uses various uses, such as discovering patterns and the relationship between data sharing of trust and resource circulation (Renaud, Britton, Wang & Ogihara, 2015; Iqbal, Jamil, Ahmad & Kim, 2020), user Behavior Prediction (Sadiq & Ahmed, 2019), such as discovering the pattern of how to occupy library space (Wang & Shao, 2017), analysis of reference questions (Finnell & Fontane, 2015; Vance, 2018), new services such as proposing practical resources and recommended systems (Tsai & Chen, 2008; Huang, Li & Xiao, 2018; Liu, 2018), the Budget allocation has been used to provide the resources needed by different groups (Kao, Chang & Lin, 2003), social network analysis (Al-Daihani & Abrahams, 2016; Al-Daihani & Abrahams, 2018) and more. Libraries can improve the quality of interaction between themselves and users through data mining. The collected data contain valuable information which can be shared with the library strategy and is fruitful for improving management decisions (Chang & Chen, 2006). Data mining has a cycle-like nature because more questions will arise after discovering patterns that shape the next process round. Utilizing advanced technologies such as data mining will constantly challenge librarians and educational institution managers as they demand creativity and strive for innovation (Patkar, 2005/2011).

Limitations

This study faced limitations regarding database selection, language, search strategy, and quality assessment of selected studies. The study was limited to original Research papers indexed from the Scopus database, ISI, which published research papers in English. The present

work does not include any research articles which were published other than in the English language.

Conclusion

Any new technique in organizations to be implemented faces challenges and obstacles. At first, data mining in libraries, especially university libraries, constantly faces issues and problems that must be studied. Until the fundamental issues and problems are identified, time, cost, and human resources will be wasted. Considering the importance of research, the high cost and time spent on it, and the impact that this research has on the process of managerial decisions, a closer look at research topics for effectiveness and use is essential. For example, one of the cases in which data mining can be used in academic libraries is to study the behavior of users of university library websites. Unfortunately, most university library websites and portals extract data sets to study behavior. Users face technical and managerial limitations. It is necessary to review this issue to be able to use this technique in analyzing the behavior of users of university library websites to correct the policies of universities.

References

- Ahmad, P., Brogan, M. & Johnstone, M. N. (2014). The e-book power user in academic and research libraries: Deep log analysis and user customisation. *Australian Academic & Research Libraries*, 45(1), 35-47. <https://doi.org/10.1080/00048623.2014.885374>
- Al-Daihani, S. M. & Abrahams, A. (2016). A text mining analysis of academic libraries' tweets. *The Journal of Academic Librarianship*, 42(2), 135–143. <https://doi.org/10.1016/j.acalib.2015.12.014>
- Al-Daihani, S. M. & Abrahams, A. (2018). Analysis of academic libraries' Facebook posts: text and data analytics. *The Journal of Academic Librarianship*, 44(2), 216–225. <https://doi.org/10.1016/j.acalib.2018.02.004>
- Anderson, C. B. & Craiglow, H. A. (2017). Text mining in business libraries. *Journal of Business & Finance Librarianship*, 22(2), 149-165. <https://doi.org/10.1080/08963568.2017.1285749>
- Ansari, N., Vakili Mofrad, H., Mansoorizadeh, M. & Amiri, M. R. (2019). Discovery of user access pattern and loan transactions analysis of information resources using association rules technique in data mining (case study: Libraries and information centers of Hamedan University of medical sciences). *Iranian Journal of Information Processing & Management (Information Science & Technology)*, 34(3), 1155-1186. <https://doi.org/10.35050/JIPM010.2019.035> [in Persian]
- Ashiq, M., Usmani, M. H. & Naeem, M. (2022). A systematic literature review on research data management practices and services. *Global Knowledge, Memory and Communication*. 71(8-9), 649-671. <https://doi.org/10.1108/GKMC-07-2020-0103>
- Batten, G., Oakes, P. M. & Alexander, T. (2014). Factors associated with social interactions between deaf children and their hearing peers: A systematic literature review. *Journal of Deaf Studies and Deaf Education*, 19(3), 285-302. <http://doi.org/10.1093/deafed/ent052>
- Berry, M. J. A. & Linoff, G. (1997). *Data mining techniques: for marketing, sales and customer support*. New York, John Wiley & Sons publication.

- Bracke, P. J. (2004). Web usage mining at an academic health sciences library: an exploratory study. *Journal of the Medical Library Association*, 92(4), 421-428. Retrieved from https://www.researchgate.net/publication/8221845_Web_usage_mining_at_an_academic_health_sciences_library_An_exploratory_study
- Bussaban, K. & Kularbphetong, K. (2014). Analysis of users' behavior on book loan log based on association rule mining. *World Academy of Science, Engineering and Technology*, 8(1), 18-20. Retrieved from <file:///C:/Users/Reza/Downloads/9997109.pdf>
- Chang, C. C. & Chen, R. S. (2006). Using data mining technology to solve classification problems: A case study of campus digital library. *The Electronic Library*, 24(3), 307-321. <https://doi.org/10.1108/02640470610671178>
- Chapman, P., Clinton, J., Kerber, R., Khabza, T., Reinartz., Shearer, C. & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step Data Mining Guide*. Copenhagen, SPSS. Retrieved from <http://www.statoo.com/CRISP-DM.pdf>
- Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 158(3), 419-444. <https://doi.org/10.2307/2983440>
- Chen, C. C. & Chen, A. P. (2007). Using data mining technology to provide a recommendation service in the digital library. *The Electronic Library*, 25(6), 711-724. <https://doi.org/10.1108/02640470710837137>
- Duan, S. & Wang, Z. (2021). Research on the service mode of the university library based on data mining. *Scientific Programming*, 2021, 1-9. <http://doi.org/10.1155/2021/5564326>
- Fang, D., Yang, H., Gao, B. & Li, X. (2018). Discovering research topics from library electronic references using latent Dirichlet allocation. *Library Hi Tech*, 36(3), 400-410. <https://doi.org/10.1108/LHT-06-2017-0132>
- Finnell, J. & Fontane, W. (2010). Reference question data mining: A systematic approach to library outreach. *Reference & User Services Quarterly*, 49(3), 278-286. Retrieved from <http://www.jstor.org/stable/20865263>
- Han, J., Kamber, M. & Pei, J. (2011). *Data mining: concepts and techniques*. 3rd edition. Burlington: Morgan Kaufman publishers. Retrieved from <https://www.sku.ac.ir/Datafiles/BookLibrary/43/Data-Mining-Concepts-and-Techniques-Han.pdf>
- Hendrickx, T., Cule, B., Meysam, B., Naulaerts, S., Laukens, K. & Goethals, B. (2015). Mining association rules in graphs based on frequent cohesive itemsets. Part of the *Lecture Notes in Computer Science book series* (LNAL), 9078. Springer, Cham. https://doi.org/10.1007/978-3-319-18032-8_50
- Hoffman-Apitius, M., Younesi, E. & Kasam, V. (2009). Direct use of information extraction from scientific text for modeling and simulation in the life sciences. *Library Hi Tech*, 27(4), 505-519. <https://doi.org/10.1108/07378830911007637>
- Huancheng, L., Tingting, W. & Rocha, Á. (2019). An analysis of research trends on data mining in Chinese academic libraries. *Journal of Grid Computing*, 17, 591-601. <https://doi.org/10.1007/s10723-018-9461-3>
- Huang, Z., Li, T. & Xiao, S. (2018). Research on library recommendation reading service system based on adaptive algorithm. *Wireless Personal Communications*, 102, 1963-1977. <https://doi.org/10.1007/s11277-018-5249-9>

- Iqbal, N., Jamil, F., Ahmad, S. & Kim, D. (2020). Toward effective planning and management using predictive analytics based on rental book data of academic libraries. *IEEE Access*, 8, 81978-81996. <http://doi.org/10.1109/ACCESS.2020.2990765>
- Jomsri, P. (2014). Book recommendation system for digital library based on user profiles by using association rule. In *Fourth Edition of the International Conference on the Innovative Computing Technology (INTECH 2014)*, (pp.130-134). Luton, Uk. <https://doi.org/10.1109/INTECH.2014.6927766>
- Kao, S. C., Chang, H. C. & Lin, C. H. (2003). Decision support for the academic library acquisition budget allocation via circulation database mining. *Information Processing & Management*, 39(1), 133-147. [https://doi.org/10.1016/S0306-4573\(02\)00019-5](https://doi.org/10.1016/S0306-4573(02)00019-5)
- Kimbrough, J. (2018). Technical services and the virtual reference desk: Mining chat transcripts for improved e-resource management. *The Serials Librarian*, 74(1-4), 212-216. <https://doi.org/10.1080/0361526X.2018.1428482>
- Kovacevic, A., Devedzic, V. & Pocajt, V. (2010). Using data mining to improve digital library services. *The Electronic Library*, 28(6), 829-843. <https://doi.org/10.1108/02640471011093525>
- Liu, Y. (2018). Data mining of university library management based on improved collaborative filtering association rules algorithm. *Wireless Personal Communication: An International Journal*, 102(4), 3781-3790. <https://doi.org/10.1007/s11277-018-5409-y>
- Long, X. & Wu, Y. (2012). Borrowing data mining based on association rules. In *2012 International Conference on Computer Science and Electronics Engineering*, (pp. 239–242). Hangzhou, China. <http://doi.org/10.1109/ICCSEE.2012.179>
- Lund, B. D. (2017). Four categories of academic libraries: A cluster analysis based on collections, expenditures, and circulation per student data. *Library Collections, Acquisitions, & Technical Services*, 40(3-4), 96-105. <https://doi.org/10.1080/14649055.2020.1794748>
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd edition). Sage Publications.
- Mishra, R. N. & Mishra, A. (2013). Relevance of data mining in digital library. *International Journal of Future Computer and Communication*, 2(1), 10-14. <http://doi.org/10.7763/IJFCC.2013.V2.110>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & the PRISMA group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology*, 62(10), 1006-1012. <http://doi.org/10.1016/j.jclinepi.2009.06.005>
- Nandha Kumar, K. G. & Christopher, T. (2017). Assessment of library users' feedback using modified multilayer perceptron neural networks. *ICTACT Journal on Soft Computing*, 7(4), 1505-1509. <http://doi.org/10.21917/ijsc.2017.0209>
- Nicholson, S. (2003). Bibliomining for automated collection development in a digital library setting: Using data mining to discover Web-based scholarly research works. *Journal of the American Society for Information Science and Technology*, 54(12), 1081-1090. <https://doi.org/10.1002/asi.10313>
- Oakleaf, M. (2010). *The value of academic libraries: A comprehensive research review and report*. American Library Association. Retrieved from https://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_report.pdf

- Patkar, V. N. (2005). Data mining applications in library and academic institutions. *Informations Studies*, 11(3), 145-156. (M. Sarafzadeh, Trans.) Retrieved from <https://www.unp.ir/article/university/sources-scientific/1582> [in Persian]
- Prehanto, D. R., Indriyanti, A. D., Permadi, G. S., Vitadiar, T. Z. & Jayanti, F. D. (2020). Library book modeling data using the association rule method with apriori algorithm in determining book placement and analysis of book loans. *International Journal of Advanced Science and Technology*, 29(05), 1244 -1250. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/9786>
- Puarungroj, W., Pongpatrakant, P., Boonsirisumpun, N. & Phromkhot, S. (2018). Investigating factors affecting library visits by University Students using data mining. *Libres*, 28(1), 25-33. <http://doi.org/10.32655/LIBRES.2018.1.3>
- Rafique, A., Ameen, K. & Arshad, A. (2021). E-book data mining: real information behavior of university academic community. *Library Hi Tech*, 41(2), 413-431. <https://doi.org/10.1108/LHT-07-2020-0176>
- Renaud, J., Britton, S., Wang, D. & Ogihara, M. (2015). Mining library and university data to understand library use patterns. *The Electronic Library*, 33(3), 355-372. <https://doi.org/10.1108/EL-07-2013-0136>
- Sadiq, M. H. & Ahmed, N. S. (2019). Classifying and predicting students' performance using improved decision tree C4.5 in higher education institutes. *Journal of Computer Science*, 15(9), 1291-1306. <https://doi.org/10.3844/jcssp.2019.1291.1306>
- Safdar, M., Batool, S. H. & Mahmood, K. (2020). Relationship between self-efficacy and knowledge sharing: systematic review. *Global Knowledge, Memory and Communication*, 70(3), 254-271. <https://doi.org/10.1108/GKMC-11-2019-0139>
- Samson, S. (2014). Usage of e-resources: Virtual value of demographics. *The Journal of Academic Librarianship*, 40(6), 620-625. <https://doi.org/10.1016/j.acalib.2014.10.005>
- Scarnò, M. (2010). User's Behaviour inside a Digital Library. *International Journal of Decision Support System and Technology*, 2(2), 52-59. <http://doi.org/10.4018/jdsst.2010040106>
- Silwattananusarn, T. & Kulkanjanapiban, P. (2020). Mining and analyzing patron's book-loan data and university data to understand library use patterns. *International Journal of Information Science and Management (IJISM)*, 18(2), 151-172. Retrieved from https://ijism.ricest.ac.ir/article_698314_28f2b01436afc77b456a1e5d0fb16165.pdf
- Suresh, R., Anand, I., Vianesh, B. & Mohammad, H. R. (2018). Study of clustering algorithms for library management system. In *2018 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)*, (pp. 221-224). <http://doi.org/10.1109/ICCPEIC.2018.8525182>
- Tsai, C. S. & Chen, M. Y. (2008). Using adaptive resonance theory and data-mining techniques for materials recommendation based on the e-library environment. *The Electronic Library*, 26(3), 287-302. <https://doi.org/10.1108/02640470810879455>
- Tsuji, K., Kuroo, E., Sato, S., Ikeuchi, U., Ikeuchi, A., Yoshikane, F. & Itsumura, H. (2012). Use of library loan records for book recommendation. In *2012 IIAI International Conference on Advanced Applied Informatics*, (pp. 30-35). Fukuoka, Japan. <http://doi.org/10.1109/IIAI-AAI.2012.16>
- Uppal, V. & Chandwani, G. (2013). An Empirical Study of Application of Data Mining Techniques in Library System. *International Journal of Computer Applications*, 74(11), 42-46. <http://doi.org/10.5120/12933-0008>

- Vance, J. M. (2018). Mining reference chat transcripts to analyze noise complaints. *Reference Services Review*, 46(4), 578-586. <http://doi.org/10.1108/RSR-11-2017-0042>
- Wang, Y. & Shao, L. (2017). Understanding occupancy pattern and improving building energy efficiency through Wi-Fi based indoor positioning. *Building and Environment*, 114, 106-117. <https://doi.org/10.1016/j.buildenv.2016.12.015>
- Will, N. (2006). Data-mining: Improvement of university library services. *Technological Forecasting and Social Change*, 73(8), 1045-1050. <https://doi.org/10.1016/j.techfore.2006.05.006>
- Williams, L. A., Fox, L. M., Roeder, C. & Hunter, L. (2014). Negotiating a text mining license for faculty researchers. *Information Technology and Libraries*, 33(3), 5-21. <https://doi.org/10.6017/ital.v33i3.5485>
- Yaco, S., Brown, C. & Konrad, L. (2016). Linking special collections to classrooms: a curriculum-to-collection crosswalk. *The American Archivist*, 79(2), 417-437. <https://doi.org/10.17723/0360-9081-79.2.417>
- Yi, K., Chen, T. & Cong, G. (2018). Library personalized recommendation service method based on improved association rules. *Library Hi Tech*, 36(3), 443-457. <https://doi.org/10.1108/LHT-06-2017-0120>
- Zhang, C. (2014). Research on construction of universities sports management information system based on data mining. *Applied Mechanics and Materials*, 687-691, (2702-2705). <http://doi.org/10.4028/www.scientific.net/AMM.687-691.2702>