

Data Mining for Library Service Improvement by Using an R Tools Technique: A Review

Hanan M. Mourad¹, Sayyada Sara Banu², Bharti W. Gawali³ and Vaishali S. Khaparde⁴

^{1,3}Department of Computer Science & Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India.

²Department of Computer Science, Jazan University, Jizan, KSA.

⁴Department of Library & Information Science, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India.

Abstract

Due to the importance of data, different types and sources of data, different types of data and data rates rise daily, which affects changes in society. It is recognized that knowledge management is a vital process in academic libraries due to the importance of academic institutions and the different needs of students and researchers. The success of academic libraries in a competitive and challenging knowledge environment depends on their ability to use information and knowledge to meet the needs of the academic community. The main part of this paper I will analyze recent papers and identify the most used methods for research and data mining in academic libraries. Qualitative research method and case study design were adopted. For academic universities based on previous research related to research and journals, we will use data mining algorithms and deep research algorithms and artificial intelligence. Academic libraries are a vital and important source for students and researchers under any circumstance or changes that the student can through the academic library obtain academic subjects scientific research conferences the library The academy is a scientist for all who are interested in this field and don't forget the important aspect that data mining plays, as it helps in arranging and separating data.

Keywords: -data mining, deep search, Academic Libraries, Artificial intelligence.

Introduction

It is considered an important and indispensable institution. Libraries have a history of benefiting from technological innovation. In the past few years, these efforts have been used to push the boundaries of digital libraries and electronic publishing. Recently, professional librarians have been exploring new tools for managing and analyzing information.

With advances in technologies and data has increased dramatically, information, digitization, and intelligence The main direction of development has become the development of academic libraries, and its functions have gradually changed. Data extraction is the process of detecting relationships, trends and patterns of meaningful information by carefully analyzing the amount of data [1, 2]. Data mining is a multidisciplinary theory that combines theories and techniques of artificial intelligence,

database technology, pattern recognition, machine learning, statistics, and data visualization [3, 4, 5, 6]. Current trends used by library staff in responding to new challenges in the new age are the application of data mining technology to academic libraries, discovery of laws and potential associations of borrowing Information large number of readers, analysis and forecast demand for relevant books based on mining results, improve personalized book services for teachers and students. Internationally, the early search for data mining applied to the field of libraries in 1996 [7]. However, the scope of data used in these audit documents is still in 2014, does not include the latest data in the past three years. Moreover, the methods adopted were mostly bibliometric methods of analysis and classification methods. In addition to. Given that the overall level of the academic library readers are above the university level and each has a university focus on different disciplines, significant differences exist in the number of books in different disciplines. The library of colleges and universities has personal characteristics, which are different from the public library in the community. For this reason, the need for personalized service from university libraries is more urgent. Practical work of information from Chinese academic universities are still in the stage of learning and exploration. Thus, there is much room for improvement. Therefore, identify guidance for follow-up studies by summarizing timely data mining applications in academic libraries Analyze and predict the general trend and research hotspots in this particular area Important. The concept of the digital library is fascinating as it includes information technology, which could produce plenty of complex data for end-users. The emergence of digital libraries storing digitized data makes it possible to Search more easily and conveniently. Traditionally, the library used to play a passive role in that it merely provided books for borrowing. It is a crucial subject, however, for a library manager to think about how to guide readers to find what they want in an aggressive way and promote the borrowing rate at the same time.

Data mining concept Extensive discussions and development have been focused on the definition of data mining. The first and probably the most commonly used definition of Data mining was conceptualized in 2001 by Laneywho characterized Data mining using the 3Vs model: volume, velocity, and variety. Similarly, based on this model, Sagioglu and Sinanc (2013) presented an extensive review of Big Data research, especially the security issues. Furthermore, Lomotey and Deters (2014) extended the model defined by Laney (2001) to a 5V model (collect data, selection, process, transformed, and pattern models), as shown in Figure 1.

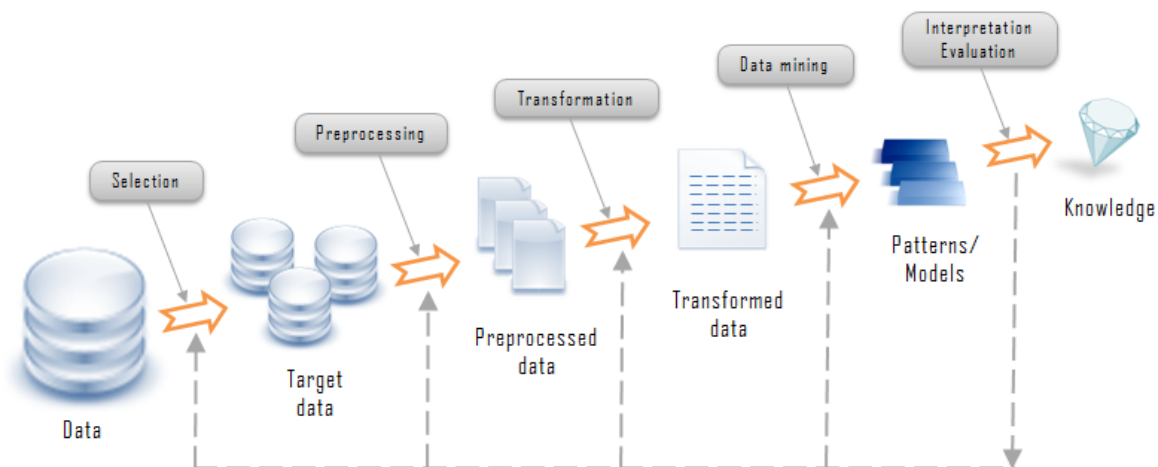


Figure 1:- The model that currently defines data mining

1- Educational Web Mining. There is growing interest in using data mining in the evaluation of web-based educational systems, making educational data mining (EDM) a rising and promising research field [Romero and Ventura 2007]. Data mining is the discovery and extraction of implicit knowledge from one or more large collection of data [Pahl and Donnellan 2002; Romero and Ventura 2007]. Educational data mining, as an emerging discipline, is concerned with applying data mining methods for exploring unique types of data that come from educational settings [Baker and Yacef 2009]. Increasingly educational learning environments, including educational digital library services, are accessed through the Web, thereby enabling a low-cost mechanism for collecting users' fine-grained behavior in real-time, and thus leaving behind a massive amount of data to analyze. Web mining, in response to this phenomenon, is a particular category of data mining problem that seeks to discover implicit patterns from usage of web documents and services [Chen and Chau 2004]. This study contributes to the field of educational web mining by investigating how to apply data mining to a particular online digital library service.

Knowledge Discovery and Data Mining Web mining typically follows the standard KDD process, entailing:

- 1) data cleaning and integration,
- 2) selection and transformation,
- 3) application of data mining algorithms,

4) evaluation and presentation [Han and Kamber 2005; Witten and Frank 2005]. Often the first two phases are combined and called data preprocessing [Cooley et al. 1997; Romero and Ventura 2007]. In general, web mining serves two purposes: description and prediction. Description aims at finding human interpretable patterns that describe the data. Prediction analyzes the existing data, and discovers relationships among the variables, in order to use such information to predict the unknown or future values of similar variables. Our study was intended to cluster Instructional Architect users in order to better understand typical teachers' online behaviors; as such, it falls into the description domain.

Classification method

Libraries classification method using data mining techniques. Classification framework: a comprehensive approach to assessing libraries in terms of the features and services they provide.

In the face of a rapidly changing landscape, characterized by shrinking budgets and dynamic services, libraries have recognized the need for evidence of their value.

Academic libraries are invited more than ever to demonstrate and justify their presence and contribution to corporate missions and goals (Association of Colleges and Research Libraries, 2010). Indeed, new trends and issues include academic libraries that influence a culture of increased

accountability for results, as libraries will be required to find better ways to document them in communications (ACRL, Research Planning and Review Committee, 2014). Nicholson (2004) and Siguenza-Guzman et al. (2015) Recognizing the need to evaluate libraries in a comprehensive and structured manner, suggested the use of a two-dimensional evaluation matrix. The four quadrants of the Comprehensive Assessment Matrix are next:

1. The internal perspective of the library system - process / service analysis: In this regard, the "library system" refers to everything that is part of the library offer, such as the organizational chart, electronic equipment, library staff, and utilities. The internal perspective of the library system includes an analysis of issues related to the operations and services that take place within the library.

2. External Perspective of the Library System - Quality Analysis: Quality

The collection and services are evaluated by the users. Thus, the second quarter, it assesses the robustness, appropriateness and usability of physical and digital resources by exploring user perceptions (Nicholson, 2004). Estimating methods for measuring service quality and collection Statistics include aggregation, suggestion boxes, web usability testing, user interface usability, and satisfaction surveys (Wright & White, 2007).

3. Library Insight - Collection Analysis: The third quadrant aims to assess the utility of the library's collection. Supporters This holistic approach suggests a combination of three evaluation methods; They are quote analysis, statistics provided by vendors, and citation databases. By doing this, libraries will gain extensive knowledge about the value of their collection and the importance of information.

4. Library Collection External Perspective - Usage Analysis: This last quarter evaluates users' behavior when tampering with the library system. Users are used to interact with the system to study users' preferences for customizing library services. Transaction log analysis, web usage analysis, in-depth log analysis, and usage statistics are the main techniques used for this.

3- Studies on improving reference service

Since challenges are confronted by libraries when providing reference service, studies are encouraged to help libraries survive conquering challenges. Weimer (2010) implements an innovative service in Alderman Library, University of Virginia. This service brings short message service (SMS) into the reference service system in Alderman Library. Scholars test how well SMS could function to provide reference service. In the end, continuous increment of usage in SMS is noticed and reference service are extended. This study highlights the significance to involve popular communication medium (in this case: mobile phone) to work for library reference service, which sets a good example for future researchers or librarians to pay attention on containing daily life resources to improve the service. Nunn and Ruane (2011) employ marketing theories to work for the improvement of reference service. They emphasize the importance to closely link users in order to manage issues caused by evolving user requirements, changing technology and increasing amount of students and long-distance citizens. Face-to-face communication is outlined in this study and librarians' social expertise is recognized as a key factor to

enhance the user awareness of library service. Human resource plays a role in improving reference service in this study. Contrary to Nunn and Ruane's study, Aguilar et al. (2011) consider that face-to-face communication is out of date and they think highly of the virtual environment. They launch initiatives to provide reference service in a virtual environment, which, according to the result of the initiatives, are more similar with current ways through which users approach information. And in the end, positive relationships with users are established. Todorinova et al. (2011) notice that data-driven changes are happening in academic libraries and involving data management for general pattern recognition could be a path to improving reference service. Saunders (2013) suggests learning a lesson from past bad examples. Two pieces of advice are put forward to improve the service at reference desk: arranging staff training, highlighting the role of evaluation. Saunders also suggests that training and evaluation should be integrated with each other and aimed more at customers. As such, the reference desk can effectively function. Aggarwal and Powers (2013) encourage a shared service model to increase reference service quality. This model was launched in the Career Education Cooperation and more access was reached with the help of this model. These studies clearly present one key approach to improve reference service: to maximize resources owned by libraries, such as enhancing skills attained by librarians, utilizing data generated at reference desk or accessing materials through shared service models. Under the situation that limited money or other sources can be invested in libraries, making full use of resource around library could be a good idea for service improvement. In addition, focusing on trends in daily life and users' needs could also be an effective method.

Literature review:-

This section presents an overview of the process-based perspective of knowledge management, extract data and processes in academic libraries, classification data from database, Association Rule Mining, clustering data in organization, research and implementation text in technique research.

1. With the continuous development of the era of big data, data mining technology is widely used in various systems. In fact, data mining technology is not a separate discipline and technology. It has integrated knowledge in many fields such as computer, mathematics, and information. Data mining, as the name implies, is to find the most useful information for this matter in a large number of databases[1].
2. The library's existing digital system, a large number of statistical data and forms are generated every moment. After these statistical data and forms are analyzed and processed, the internal information volume can be used for various services of the library. In particular, it will play a strong guiding role in procurement, collection, and consulting services[2].
3. Today library is a half and half of print and computerized assets. Many created and computerized libraries who have huge measure of advanced substance are presently attempting to give e-learning through their advanced libraries web entomb face, there by growing completely trustworthy information framework. In this article we will view different parts of e-figuring out how computerized libraries can add to e-learning[3].

4. we propose the application mechanism and strategy of user behavior data of university library, which mainly includes the improvement of the organization system of big data collection, the evaluation of the data scale and usability, positioning and locating the user's big data personalized service demand, and guaranteeing the regularity and effectiveness of the work of each layer in the organization system[4].
5. In the library's existing digital system, a large number of statistical data and forms are generated every moment. After these statistical data and forms are analyzed and processed, the internal information volume can be used for various services of the library. In particular, it will play a strong guiding role in procurement, collection, and consulting services[5].
6. The development of digital libraries is directly related to the development of information technologies, particularly Internet technologies. A digital library is an innovative library service that uses information technology, and nearly every major development stage of a digital library is accompanied by major technological changes[6].
7. In the social sciences, researchers search for information on the Web, but this is most often distributed on different websites, search portals, digital libraries, data archives, and databases. In this work, we present an integrated search system for social science information that allows finding information around research data in a single digital library[7].
8. Recently, Big Data studies have attracted considerable attention. However, Big Data analytics in academic libraries confront two fundamental challenges: the huge volume, velocity, and variety of data and the complexity of its techniques and algorithms. The primary aim of this study is to explore which techniques and tools can be applied in academic libraries to analyze Big Data, and then determine its profits in academic libraries[8].
9. The reflections on the future value of the library are presented by Wheel of Value higher order categorization. This approach proved useful in eliciting responses from participants in the face of recognized difficulty in getting beyond current views of the library and the approach is recommended to other universities looking to carry out a similar project[9].
10. The rapid development of network technology has brought great influence to library circles, How to reposition and retain readers of Libraries in the new period has aroused the thinking of industry. The consensus view is that libraries need to tap the needs of users to promote reader reflux[10].
11. Digital Libraries are an evolving area of research; Digital libraries defined as electronic information collections containing large and diverse repositories of digital objects, which can be accessed by a large number of geographically distributed users. Such repositories would exist in locations physically near or remote from the users [11].

12. The current work of library circulation service desk is being transferred from traditional pure manual operations to human–machine collaboration and artificial intelligence. It is urgent to study a set of general and reliable service desk distribution models to better optimize staffing and improve service efficiency and quality[12].
13. Recommender systems are important tools in library websites that assists the user to find the appropriate books. With the rapid library searching system. This research presents a book recommendation system for university libraries to support user interests development of internet technologies and the number of books has varied which waste of time and difficulty for finding from which are related in the same topic and faculty[13].
14. The recent Internet and Web technologies help higher educational institutions to design and offer online educational opportunities to meet the student and adult needs, such as, convenience and flexibility (Yukselturk, 2009). With the help of these technologies, the number of online degree programs and courses has significantly increased in the new century (Allen & Seaman, 2007). Despite the increasing in the number of online courses and programs, online learning suffers from several problems[14].
15. Today, one of the biggest challenges that E-management systems face is the explosive growth of operating data and to use this data to enhance services. Web usage mining has emerged as an important technique to provide useful management information from user's Web data[15].
16. This paper offers new understandings about the nature, form and value of personalisation in children's reading for pleasure, with a particular focus on personalisation in digital library systems, and how these systems currently position teachers. It argues for the reconceptualization of their professional positions within such systems which could afford more pedagogic value[16].
17. The use of Big Data as a resource can become very rife due to its application in educational analysis and data-driven decision-making, and it can even emerge as a vehicle for state transparency. Almost each sector has developed a fascination with the ostensibly new discovery of Big Data and its extraordinary capabilities to fuel analytical breakthroughs since 2012[17].
18. In the implementation and use of research information systems (RIS) in scientific institutions, text data mining and semantic technologies are a key technology for the meaningful use of large amounts of data. It is not the collection of data that is difficult, but the further processing and integration of the data in RIS. Data is usually not uniformly formatted and structured, such as texts and tables that can not be linked. These include various source systems with their different data formats such as project and publication databases, CERIF and RCD data model, etc.[18].
19. fulfill their information needs. The ever increasing nature of data in to information and to

- knowledge necessitated Big data analytics. For example if we take Polavaram Project being constructed in the state of Andhra Pradesh, a project engineer is having 2,523 records of physical and electronic books and journals, recordings, maps, and field trip notes, however, most of these records remain isolated from the Web, requiring a detailed study how this data could be effectively exposed for use with current big data and other technologies[19].
20. In recent years, a wide array of tools have emerged for the purposes of conducting educational data mining (EDM) and/or learning analytics (LA) research. In this article, we hope to highlight some of the most widely used, most accessible, and most powerful tools available for the researcher interested in conducting EDM/LA research[20].
 21. Digital competencies for developing and managing digital libraries fall into three main categories: digital competencies for developing digital libraries; digital competencies for managing digital libraries; and digital competencies to protect digital contents. The results revealed that training programs offered by Higher Education Commission (HEC), library associations, library schools, in-house trainings, use of online tutorials and trainings offered by skilled professionals are highly important and useful for university librarians to acquire digital competencies[21].
 22. The overall goal of the data mining process is to extract information from a data set and transforms to extract information from a data set and transform it into an understandable structure for further use.it into an understandable structure for further use. Data miningis the analysis step of the "knowledge discovery in databases" is the analysis step of the "knowledge discovery in databasesprocess, or KDD[22].
 23. Significant changes seen by society through the transmission of information around the globe and it is accountable to the evolution of information technology. Now it is possible to archiving and accessing knowledge in the digitized form besides preservation of traditional knowledge due to use of information technology. Demand for electronic information increasing day by day and at the same time traditional format of library becoming more and more expensive and complex to maintain[23].
 24. Data mining is a technology that blends traditional data analysis methods with sophisticated algorithms for processing large volumes of data. It has also opened exiting opportunities for exploring and analyzing new types of data and for analyzing old types of data in new ways. Data Mining is the process of automatically discovering useful information in large data repositories[24].
 25. Data mining technology is a process of discovering new knowledge using the means of network technology to find out valuable information that are hidden and cannot be found directly from the massive, random and no regular information in the data, to help people do effective analysis[25].
 26. Libraries and archives are the storehouse of knowledge assets and facilitate education,

research and overall development of the mankind. Since the mostly on paper, libraries are built in institutions or centers thus making them locality-bound and access-time restricted[26].

27. Highlights of this tutorial include the applications of digital libraries [2] and the underlying technologies [3], which include: Exploration, Evaluation, Integration, Complex Objects, Annotation/Subdocuments, Ontologies, Classification, Text Extraction, Security, Content-based Image Retrieval, Education, Social Networks, Bioinformatics/eScience/Simulation, and Geospatial Information[27].
28. Cloud computing facilitates end-users or small companies to use computational resources such as software, storage, and processing capacities belonging to other companies (cloud service providers). Cloud services include Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) [2]. Big corporate like Amazon, Google and Microsoft are providing cloud services in various forms. Amazon Web Services (AWS) provides cloud services that include Amazon Elastic Compute Cloud (EC2), Simple Queue Service (SQS) and Simple Storage Service (S3) [28].
29. Digital libraries allow online access to devices that contain digital knowledge. Libraries have traditionally worked on publishing data, however, to solve more complex queries it is necessary the connection to external data sources[29].
30. Both information retrieval (IR) and bibliometrics have long histories as distinct areas of investigation in information science. IR has focused on the storage, representation and retrieval of documents (text or other media) from the system and user perspectives[30].
31. Data mining, as a heatedly discussed term, has been studied in various fields. Its possibilities in refining the decision-making process, realizing potential patterns, and creating valuable knowledge have won attention of scholars and practitioners. However, there are less studies intending to combine data mining and libraries where data generation occurs all the time[31].
32. Research in digital libraries (DLs) has gained much interest across the globe. Most funding related to DL are available for building DLs, rather than producing digital librarians by developing the DL curricula and offering necessary funding to introduce state-of-the-art DL labs for future library professionals[32].
33. Based on the literature information resource, university library has various database resources, a large number of professionals in library information and science, good facilities and technical means. Thereby, university library has the ability to carry out patent information services. On December 13, 2016, State Intellectual Property Office of The P.R.C issued Rules of Intellectual Property Management for higher education institutions which explicitly defines university library as the supportive IP service organization and specifies the corresponding responsibilities[33].

34. Data Mining is the process of sifting through stores of data to extract previously unknown, valid patterns and relationships that provide useful information [1]. Once these patterns are found they can further be used to make certain decisions for development of businesses[34].
35. As more and more data accumulated, companies are eager to dig out the knowledge behind a large amount of data to support decision-making. The existing data mining tools such as IBM's IntelligenceMiner and SAS EnterpriseMiner, although they provide a comparatively rich mining function[35].
36. Data mining, also known as knowledge discovery in databases, can be defined as the process of analyzing large information repositories and of discovering implicit, but potentially useful information. Data mining has the capability to uncover hidden relationships and to reveal unknown patterns and trends by digging into large amounts of data[36].
37. Data Mining techniques may play pivotal role in effective and efficient management of E-Governance. In this paper an attempt has been made to review and compare the work done by various researchers in this field[37].
38. Data mining analysis is based normally on three techniques: classical statistics, artificial intelligence, and machine learning (Girija & Srivatsa, 2006). Classical statistics is mainly used for studying data, data relationships, as well as for dealing with numeric data in large databases (Hand, 1998)[38].
39. Educational data mining (EDM) is an exciting and rapidly growing new area that combines multiple disciplines toward understanding how students learn and toward creating better support for such learning[39].
40. This paper aims to provide the reader with a comprehensive background for understanding current knowledge on Learning Analytics (LA) and Educational Data Mining (EDM) and its impact on adaptive learning. It constitutes an overview of empirical evidence behind key objectives of the potential adoption of LA/EDM in generic educational strategic planning.[40]
41. Big Data concern large-volume, complex, growing data sets with multiple, autonomous sources. With the fast development of networking, data storage, and the data collection capacity, Big Data are now rapidly expanding in all science and engineering domains, including physical, biological and biomedical sciences[41].
42. Turning raw data into useful information and knowledge also enables educational institutions to improve teaching and learning practices, and to facilitate the decision-making process in educational settings. Thus, educational data mining is becoming an increasingly important research area with a specific focus to exploit the abundant data generated by various educational systems for enhancing teaching, learning and decision

making[42].

43. University library works as the information service center for students and teachers of a university and the services provided by a university library can be classified into many categories ranging from text books to advanced research papers in different formats such as paper-based resources, digital documents, compact disks and digital versatile disks[43].
44. Increasingly, education and training are delivered beyond the constraints of the classroom environment, and the increasingly widespread availability of online repositories, educational digital libraries, and their associated tools are major catalysts for these changes[44].
45. Traditionally, the majority of online instructors and institutional administrators rely on web-based course evaluation surveys to evaluate online courses (Hoffman, 2003). The data and information are then used to help inform online program effectiveness and generate information for program-level decision-making[45].
46. The advancement of medicine now relies upon the collection, management, storage, and analysis of large biological datasets. Data mining, statistical and machine learning techniques are the process by which new knowledge is extracted from a dataset[46].
47. Hand drawn pen and paper sketches are commonplace for capturing early phase designs and diagrams. Pen and paper offer an unconstrained space suitable for quick construction and allows for ambiguity[47].
48. -Compared with the traditional data mining system algorithms library, the introduction of Web Service technology into data mining system algorithm library realizes the separation of data, algorithm, and interface. The loose and interconnected mode greatly reduces the complexity of the development for the algorithms library in data mining system and is convenient for the dynamic management of data mining algorithm library[48].
49. Data mining is the latest tool available to explore the hidden information from the large amount of database, this can be done by applying the data mining in the library database. Data mining is usually employed on very large database[49].
50. Data mining is a computational process that extracts patterns that may be significant. It is used when one has lots of data and has the 'don't know where to start' issue. Data mining looks at the data in relationship to the other data within the collection which differentiates it from classic data retrieval where the system is responding to the external criteria of the search term[50].
51. Data mining techniques are used to recommend digital library services based on the user's profile and search history. First, similar users were clustered together, based on their profiles and search behavior. Then predictive classification for recommending appropriate services to them was used. It has been shown that users in the same cluster have a high probability of accepting similar services or their patterns[51].

52. Data mining is from a large number of heterogeneous data through dealingscreening process to obtain useful information. W. J. Frawley's definition of data mining is given by the public recognition - "data mining is from a large number of incomplete, noisy, fuzzy and random practical application data extracted knowledge of interest[52].
53. Group work is widespread in education. The growing use of online tools supporting group work generates huge amounts of data. We aim to exploit this data to support mirroring: presenting useful high-level views of information about the group, together with desired patterns characterizing the behavior of strong groups[53].
54. Data mining as a heatedly discussed term has been studied in various fields. Its possibilities in refining decision making process, realizing potential patterns and creating valuable knowledge have won attention from scholars and practitioners[54].
55. We discuss the use of web metrics at four digital libraries, the Instructional Architect, the Library of Congress, the National Science Digital Library, and WGBH Teachers' Domain. We describe practical issues involved in implementing and using web metrics to track web site performance[55].
56. Patent documents contain important research results that are valuable to the industrybusiness, law, and policy-making communities. If carefully analyzed, they can show technological details and relations, reveal business trends, inspire novel industrial solutions, or help make investment policy[56].
57. The humanities encompass what are popularly known as the liberal arts: literature, history, art history, philosophy, music, and language studies. Computational techniques in the humanities have a long history, dating back at least as far as Father Roberto Busa's use of punch- cards and an IBM computer to compile a concordance to the complete works of Thomas Aquinas in the late 1940s [57].
58. The intensive use of Information and Communication Technologies such as the Internet increases the possibilities for both content searching and delivery. This new paradigm has completely changed the vision in the distance education field[58].
59. Document clustering has been used for better document retrieval, document browsing, and text mining in digital library. In this paper, we perform a comprehensive comparison study of various document clustering approaches such as three hierarchical methods (single-link, complete-link, and complete link), Bisecting K-means, K-means, and Suffix Tree Clustering in terms of the efficiency, the effectiveness, and the scalability[59].
60. Data Mining refers to the extraction or "Mining" knowledge from large amount of data or Data Warehouse. To do this extraction data mining combines artificial intelligence, statistical analysis and database management systems to attempt to pull knowledge form stored data[60].

61. THE wide availability of ever-growing data sets from different domains has created a demand for automatic processes for extracting information from them. Data Mining (DM) is commonly defined as the extraction of patterns or models from observed data, usually as part of a more general process of extracting high-level, potentially useful knowledge, from low-level data, known as Knowledge Discovery in Databases[61].
62. Web usage mining is the application of data mining techniques to discover usage patterns from Web data, in order to understand and better serve the needs of Web-based applications. Web usage mining consists of three phases, namely preprocessing, pattern discovery, and pattern analysis[62].
63. In many applications, the data of interest comprises multiple sequences that each evolve over time. Examples include currency exchange rates, network traffic data from different network elements, demographic data from multiple jurisdictions, patient data varying over time[63].
64. In an Anglo-American-Scandinavian context, the public, research and university libraries have positioned themselves as the main vehicles for the organization, access, distribution and use of stored knowledge of quality[64].
65. The growing need of software infrastructure able to create, maintain and ease the evolution of scientific data, promotes the development of digital libraries in order to provide the user with fast and reliable access to data. In a world that is rapidly changing, the standard view of a digital library as a data repository specialized to a community of users and provided with some search tools is no longer tenable[65].
66. Electronic commerce (EC) and digital libraries (DL) are two increasingly important areas of computer and information sciences, with different user requirements but similar infrastructure requirements. In exploring strategic directions, we examine both requirements of the global information infrastructure that are a necessary prerequisite for EC and DL [Aho 1996] and specific requirements of EC and DL within the global infrastructure[66].
67. The process of locating and acquiring relevant information from libraries is getting more complicated due to the vast amount of information resources one has to plough through. To serve users purposefully, an academic library should be able to avail to users the tools and services that lessen the task of searching for information[67].
68. E-content revolution, technological advances, and ever-shrinking budgets oblige libraries to efficiently allocate their limited resources between collection and services. Unfortunately, re- source allocation is a complex process due to the diversity of data sources and formats required to be analyzed prior to decision making, as well as the lack of efficient methods of integration[68].

69. This case study in educational data mining (EDM) describes the motivation, methods, results, and best practices used while studying the usage data from users of a web-based tool, the Instructional Architect (IA.usu.edu). Our focus is twofold: (a) to highlight lessons learned for those new to data mining and (b) to introduce the analysis techniques of latent class modeling as an analysis tool[68].
70. Data is being collected and compiled in the global business environments. There is an urgent requirement for a new generation of computational theories and tools to assist humans in extracting useful information from the promptly growing volumes of digital data. DM is the process of applying these computational methods in showing unknown data formats in large data sets[69].
71. Increasingly, education and training are delivered beyond the constraints of the classroom environment, and educational digital libraries and their associated services are making major contributions to these changes [Choudhury et al. 2002]. With the rapid growth of e-learning environments and information networks, researchers as well as stakeholders need to ensure efforts and resources expended on the development of digital libraries are worthwhile in terms of their impact on targeted users[70].
72. Wind power generation is rapidly expanding and is becoming a noticeable contributor to the electric grid. The fact that most large- scale wind farms were developed in recent years has made studies of their performance overdue. Given the changing nature of the wind regime, wind farm power varies across all time scales[71].
73. The integration of multiple recommendation algorithms using various data and the real-time requirement are pressing problems in the development of e-commerce personalized service. This paper introduces recommendation methods and the timing and manners of recommendation result display, presents multiple recommendation algorithms that reflect the latest achievements in data mining research, designs a model of the e-commerce personalized recommendation system based on data mining[72].
74. The paper reviews applications of data mining in manufacturing engineering, in particular production processes, operations, fault detection, maintenance, decision support, and product quality improvement[73].
75. data mining is one kind of applications of the knowledge management. The aim of it is to discuss how to discover the useful information or rules from numerous data and use them to make decisions [7][9]. This technique is most likely applied on the field with huge data and commerce, such as business and medical science[74].

Conclusion: -

Because libraries contain large amounts of data called big data in academic libraries wireless

focused in this study, it is important to pay attention and conduct further studies to highlight this area of interest. In addition, many large databased technologies have not been taken. Big data can positively enable libraries to make more financially conscious or imaginative choices or suggestions that can fully meet user needs. Opinions on the data are expanding rapidly, and more analysts are willing Generally collect, mine, and sort data in new ways. Without seeing big data, searching for some data may become ineffective. The data collected by the library customers to use the department isIt is exceptionally useful in enhancing the overall customer experience and realizing the benefits of the user library. The ability to collect and analyze vast amounts of data will be prevalent in all institutions, including libraries. Therefore, big data may be suitable for linking big data or funding. A familiar database management system (database management system) or data inspection may be the dominant methodology. Future studies should examine the real stages or progress of a large data library. Possible extensions of this study include:

- Studying big data opportunities and challenges in academic libraries;
- A study of how classifying and extracting big data systematically works on economic value in academic libraries
- Propose guidance to librarians on how to develop a system and process in order to use big data in academic libraries
- Studying how academic libraries can benefit from data mining techniques in linking, classifying, and researching.

Key Findings:

It is beneficial that, relative to traditional and new digital libraries, users are a valuable resource if standing on the side of resource. The big data era allows the possibility to fully understand and connect users. Interaction between users and libraries is not only to meet the information needs of users but also continuously provide more user resources to libraries. By exploiting user resources, digital libraries can have a broader perspective of the construction of data resources. Obviously, this user-centric digital library transformation model can provide good opportunity for personalized service development. This change in the characteristics of library services can put forward higher requirements of utilization of big data and guide the direction of changes in library services.

Recommender systems are important tools in library websites that assists the user to find the appropriate books. With the rapid development of internet technologies and the number of books has varied which waste of time and difficulty for finding from library searching system. This research presents a book recommendation system for university libraries to support user interests which are related in the same topic and faculty. The main motive of this research is to develop the technique which recommends the most suitable books to users according to the faculty of the user profile with book category, and book loan or FUCL technique. This is based on the combined features of association rule mining. The results show that FUCL mining technique is suitable to apply for the recommender book tool in the library and has a higher accuracy value than other technique.

Bibliography

W. Wang, L. Meng, L. Wu, and J. Zhang, "Research and Application of Data Mining Technology in Library Office Information Construction," *J. Phys. Conf. Ser.*, vol. 1550, no. 3, 2020, doi: 10.1088/1742-6596/1550/3/032001.

J. Hu, J. Tang, and B. Zhang, "Research on Readers' Behavior of College Libraries Based on Data Mining Technology," *J. Phys. Conf. Ser.*, vol. 1345, no. 5, pp. 0–7, 2019, doi: 10.1088/1742-6596/1345/5/052010.

S. S. Hiremath, S. Lalasangi, and S. Acharya, "Role of digital libraries in E-Learning," *IP Indian J. Libr. Sci. Inf. Technol.*, vol. 4, no. 2, pp. 61–62, 2019, doi: 10.18231/j.ijlsit.2019.017.

R. Xu and Y. Yang, "Study of User Behavior Data Mining and Applications in University Libraries --Taking the Central China Normal University Library as an example," vol. 87, no. Icmeit, pp. 256–260, 2019, doi: 10.2991/icmeit-19.2019.44.

J. Hu, J. Tang, and B. Zhang, "Research on Readers' Behavior of College Libraries Based on Data Mining Technology," *J. Phys. Conf. Ser.*, vol. 1345, no. 5, 2019, doi: 10.1088/1742-6596/1345/5/052010.

S. Li, F. Jiao, Y. Zhang, and X. Xu, "Problems and Changes in Digital Libraries in the Age of Big Data From the Perspective of User Services," *J. Acad. Librariansh.*, vol. 45, no. 1, pp. 22–30, 2019, doi: 10.1016/j.acalib.2018.11.012.

D. Hienert, D. Kern, K. Boland, B. Zapilko, and P. Mutschke, "A digital library for research data and related information in the social sciences," *Proc. ACM/IEEE Jt. Conf. Digit. Libr.*, vol. 2019-June, pp. 148–157, 2019, doi: 10.1109/JCDL.2019.00030.

H. S. Al-barashdi and R. Al-karousi, "Big Data in academic libraries : literature review and future research directions," no. February, pp. 0–16, 2019, doi: 10.5339/jist.2018.13.

G. Boyce *et al.*, "The Journal of Academic Librarianship Visions of value : Leading the development of a view of the University Library in the 21st century," *J. Acad. Librariansh.*, vol. 45, no. 5, p. 102046, 2019, doi: 10.1016/j.acalib.2019.102046.

T. Xia and Y. Liu, "Application of Improved Association-rules Mining Algorithm in the Circulation of University Library," no. Icbdai, pp. 60–64, 2018, doi: 10.25236/icbdai.2018.010.

G. P. Khiste, R. K. Deshmukh, and A. P. Awate, "Literature audit of 'Digital Library': an overview," *Interdiscip. Multiling. Ref. J.*, no. February, pp. 403–411, 2018.

L. Sun and K. Yu, "Research on Library Queuing Model Based on Data Mining," *KnE Soc. Sci.*, vol. 3, no. 10, pp. 1452–1462, 2018, doi: 10.18502/kss.v3i10.3483.

- P. Jomsri, "FUCL mining technique for book recommender system in library service," *Procedia Manuf.*, vol. 22, pp. 550–557, 2018, doi: 10.1016/j.promfg.2018.03.081.
- E. Yukselturk, S. Ozekes, and Y. K. Türel, "Predicting Dropout Student: An Application of Data Mining Methods in an Online Education Program," *Eur. J. Open, Distance E-Learning*, vol. 17, no. 1, pp. 118–133, 2018, doi: 10.2478/eurodl-2014-0008.
- E. Optimizing *et al.*, "University of Huddersfield Repository WEB DATA MINING," 2018.
- N. Kucirkova and T. Cremin, "Personalised reading for pleasure with digital libraries: towards a pedagogy of practice and design," *Cambridge J. Educ.*, vol. 48, no. 5, pp. 571–589, 2018, doi: 10.1080/0305764X.2017.1375458.
- R. Article, "Big Data in academic libraries : literature review and future research directions," 2018.
- O. Azeroual *et al.*, "Text data mining and data quality management for research information systems in the context of open data and open science To cite this version : HAL Id : hal-01942077 Text data mining and data quality management for research information systems in the context of open data and open science," 2018.
- A. Kaladhar, "Application of Big Data Technology to Library data : A review," vol. 8, no. 64344, pp. 25–30, 2018.
- S. Slater, S. Joksimović, V. Kovanovic, R. S. Baker, and D. Gasevic, "Tools for Educational Data Mining: A Review," *J. Educ. Behav. Stat.*, vol. 42, no. 1, pp. 85–106, 2017, doi: 10.3102/1076998616666808.
- S. A. Khan and R. Bhatti, "Digital competencies for developing and managing digital libraries: An investigation from university librarians in Pakistan," *Electron. Libr.*, vol. 35, no. 3, pp. 573–597, 2017, doi: 10.1108/EL-06-2016-0133.
- S. Sengupta, "Applications of Data Mining in Library & Information Centres: An Overview," *Int. J. Curr. Res.*, vol. 9, no. 01, pp. 45246–45249, 2017.
- K. Thaker, "Digital Library and User's experience: A Literature Review," vol. 2, no. 1, pp. 1–8, 2017, [Online]. Available: www.ssarsc.org.
- P. Singh, "Data Mining: An Overview," *J. Knowl. Commun. Manag.*, vol. 7, no. 2, p. 214, 2017, doi: 10.5958/2277-7946.2017.00016.x.
- G. Li, "Application of Data Mining Technology in Information Analysis of University Library Literature," vol. 130, no. Mcei, pp. 310–313, 2016, doi: 10.2991/mcei-16.2016.65.
- P. Pr *et al.*, "National Digital Library : a Platform for Paradigm," 2016.

E. A. Fox, "Introduction to Digital Libraries," *Proc. ACM/IEEE Jt. Conf. Digit. Libr.*, vol. 2016-Septe, pp. 283–284, 2016, doi: 10.1145/2910896.2925429.

G. Gupta and D. R. Pathak, "Cloud Computing: 'Secured Service Provider for data mining,'" *Int. J. Eng. Comput. Sci.*, vol. 5, no. 5, 2016, doi: 10.18535/ijecs/v5i5.17.

M. Hallo, S. Luján-Mora, A. Maté, and J. Trujillo, "Current state of Linked Data in digital libraries," *J. Inf. Sci.*, vol. 42, no. 2, pp. 117–127, 2016, doi: 10.1177/0165551515594729.

D. Wolfram, "Bibliometrics, information retrieval and natural language processing: Natural synergies to support digital library research," *CEUR Workshop Proc.*, vol. 1610, pp. 6–13, 2016.

M. Zhan, "Exploring the Feasibility of Applying Data Mining for Library Reference Service Improvement," *Åbo Akad.*, 2016.

G. G. Chowdhury, "How to improve the sustainability of digital libraries and information Services?," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 10, pp. 2379–2391, 2016, doi: 10.1002/asi.23599.

Y. Wu, "Research on Patent Information Service in University Libraries Oriented to Whole Scientific Research Process : Practices from China," vol. 21, no. Icmesd, pp. 225–230, 2017.

N. Mlambo, "Data Mining : Techniques , Key Challenges and Approaches for Improvement," vol. 6, no. 3, pp. 59–65, 2016.

W. Yi, "Research on Key Technologies of Data Mining System Based on Web Service," no. Icmesd, pp. 225–230, 2016.

L. Siguenza-Guzman, V. Saquicela, E. Avila-Ordóñez, J. Vandewalle, and D. Cattrysse, "Literature Review of Data Mining Applications in Academic Libraries," *J. Acad. Librariansh.*, vol. 41, no. 4, pp. 499–510, 2015, doi: 10.1016/j.acalib.2015.06.007.

K. K. Verma and N. Shrivastava, "Exploring Role and Associated Challenges of Data Mining Technique and it ' s Implementation in," vol. 4, no. 8, pp. 3473–3478, 2015.

L. Siguenza-guzman, V. Saquicela, E. Avila-ordóñez, J. Vandewalle, and D. Cattrysse, "The Journal of Academic Librarianship Literature Review of Data Mining Applications in Academic Libraries," vol. 41, pp. 499–501, 2015.

K. R. Koedinger, S. D'Mello, E. A. Mclaughlin, Z. A. Pardos, and C. P. Rosé, "Data mining and education," *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 6, no. 4, pp. 333–353, 2015, doi: 10.1002/wcs.1350.

Z. Papamitsiou and A. A. Economides, "Learning analytics and educational data mining in practice: A systemic literature review of empirical evidence," *Educ. Technol. Soc.*, vol. 17, no. 4, pp. 49–64, 2014.

X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 1, pp. 97–107, 2014, doi: 10.1109/TKDE.2013.109.

W. He, "Examining students' online interaction in a live video streaming environment using data mining and text mining," *Comput. Human Behav.*, vol. 29, no. 1, pp. 90–102, 2013, doi: 10.1016/j.chb.2012.07.020.

"Investigation and Analysis on Library Service Quality," vol. 3, no. 10, pp. 15–23, 2013.

B. Xu, M. Recker, S. Journal, B. Xu, and M. Recker, "International Forum of Educational Technology & Society Teaching Analytics : A Clustering and Triangulation Study of Digital Library User Data Published by : International Forum of Educational Technology & Society Linked references are available on JSTOR ," vol. 15, no. 3, pp. 103–115, 2012.

A. J. Hung, Y. Hsu, and K. Rice, "International Forum of Educational Technology & Society Integrating Data Mining in Program Evaluation of K-12 Online Education Published by : International Forum of Educational Technology & Society Linked references are available on JSTOR for this article," *Int. Forum Educ. Technol. Soc.*, vol. 15, no. 3, pp. 27–41, 2012.

C. Edeki and S. Pandya, "Comparative Study of Data Mining and Statistical Learning Techniques for Prediction of Cancer Survivability," *Mediterr. J. Soc. Sci.*, vol. 3, no. November, pp. 49–56, 2012, doi: 10.5901/mjss.2012.v3n14p49.

R. Blagojevic, B. Plimmer, J. Grundy, and Y. Wang, "Using data mining for digital ink recognition: Dividing text and shapes in sketched diagrams," *Comput. Graph.*, vol. 35, no. 5, pp. 976–991, 2011, doi: 10.1016/j.cag.2011.07.002.

Y. B. Wu, "The application of web service technology in data mining system algorithm library," *Proc. 2011 Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2011*, vol. 4, pp. 2473–2477, 2011, doi: 10.1109/ICCSNT.2011.6182471.

V. A. Moses and C. S. Jeevan, "Library Management System Using Association Rule Mining," pp. 10–16, 2011.

K. Chiang, "Data mining , data fusion , and libraries," 2010.

A. Kovacevic, "Using data mining to improve digital library services," no. November, 2010, doi: 10.1108/02640471011093525.

G. G. P. Sekar*, "An Indian Journal Review," *Btairj*, vol. 3, no. 3, pp. 174–178, 2009.

D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. R. Zaane, "Clustering and sequential pattern mining of online collaborative learning data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 6, pp. 759–772, 2009, doi: 10.1109/TKDE.2008.138.

“Data Mining for Library Service Improvement : A Case Study of Turku City Library Ming Zhan Information Studies Faculty of Social Science , Business and Economics Åbo Akademi University Abstract,” pp. 1–11, 2008.

M. Khoo, M. Recker, J. Pagano, B. Palmer, A. L. Washington, and R. A. Donahue, “Using Web metrics to analyze digital libraries,” *Proc. ACM Int. Conf. Digit. Libr.*, pp. 375–384, 2008, doi: 10.1145/1378889.1378956.

Y. Tseng, “Text mining techniques for patent analysis,” no. September, 2007, doi: 10.1016/j.ipm.2006.11.011.

M. Kirschenbaum, “The remaking of reading: Data mining and the digital humanities,” ...*Next Gener. Data Min. Cyber-Enabled ...*, 2007, [Online]. Available: <http://www.csee.umbc.edu/~hillol/NGDM07/abstracts/talks/MKirschenbaum.pdf%5Cnhttp://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.959&rep=rep1&type=pdf>.

N. Ferran, J. Casadesús, M. Krakowska, and J. Minguillón, “Enriching e-learning metadata through digital library usage analysis,” *Electron. Libr.*, vol. 25, no. 2, pp. 148–165, 2007, doi: 10.1108/02640470710741296.

I. Yoo and X. Hu, “A comprehensive comparison study of document clustering for a biomedical digital library MEDLINE,” *Proc. ACM/IEEE Jt. Conf. Digit. Libr.*, vol. 2006, pp. 220–229, 2006, doi: 10.1145/1141753.1141802.

R. K. Dwivedi and R. P. Bajpai, “Use of Data Mining in the field of Library and Information Science : An Overview,” *2nd Int. CALIBER-2004*, pp. 512–519, 2004, [Online]. Available: <http://ir.inflibnet.ac.in:8080/jspui/handle/1944/375>.

M. C. Ferreira de Oliveira and H. Levkowitz, “From visual data exploration to visual data mining: A survey,” *IEEE Trans. Vis. Comput. Graph.*, vol. 9, no. 3, pp. 378–394, 2003, doi: 10.1109/TVCG.2003.1207445.

P. N. Srivastava, Jaideep; Cooley, R; Deshpande, M; Tan, “Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data,” *SIGKDD Explor.*, vol. 1, no. 2, pp. 12–23, 2000, doi: 10.1145/846183.846188.

B. K. Yi, N. D. Sidiropoulos, T. Johnson, H. V. Jagadish, C. Faloutsos, and A. Biliris, “Online data mining for co-evolving time sequences,” *Proc. - Int. Conf. Data Eng.*, pp. 13–22, 2000, doi: 10.1109/icde.2000.839383.

P. Ingwersen, “The role of libraries and librarians in organising digital information,” *Libri*, vol. 49, no. 1, pp. 11–15, 1999, doi: 10.1515/libr.1999.49.1.11.

G. Aloisio, M. Cafaro, and R. Williams, "The digital Puglia project: An active digital library of remote sensing data," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1593, pp. 563–572, 1999, doi: 10.1007/bfb0100617.

N. Adam *et al.*, "Strategic directions in electronic commerce and digital libraries: Towards a digital agora," *ACM Comput. Surv.*, vol. 28, no. 4, pp. 818–835, 1996, doi: 10.1145/242223.246871.

A. J. Alunga, "A Conceptual Data Mining Model (DMM) used in Selective Dissemination of Information (SDI): a case study of Strathmore University library," vol. 1, no. 2, pp. 12–21.

B. C. Palmer, "Applying Web Usage Mining to an Educational Digital Library Service," no. November, 2016.

K. Kasemsap, "The Role of Data Mining for Business Intelligence in Knowledge Management," pp. 12–14, 2015, doi: 10.4018/978-1-4666-6477-7.ch002.

B. Xu, "Understanding Teacher Users of a Digital Library Service : A Clustering Approach," vol. 3, no. 840745, pp. 1–28, 2011.

H. Zheng and A. Kusiak, "Prediction of Wind Farm Power Ramp Rates : A Data-Mining," vol. 131, no. August, pp. 2–9, 2009, doi: 10.1115/1.3142727.

Y. Fan, Y. Shen, and J. Mai, "Study of the Model of E-commerce Personalized Recommendation System Based on Data Mining," pp. 647–651, 2008, doi: 10.1109/ISECS.2008.106.

J. A. Harding, M. Shahbaz, and A. Kusiak, "Data Mining in Manufacturing :," vol. 128, no. November, pp. 969–976, 2006, doi: 10.1115/1.2194554.

M. Lin, "Enhancing Library Resources Usage Efficiency by Data Mining," pp. 382–387, 2004ss.

Iqbal, Peerzada Mohammad, and Abdul Majid Baba. "Currency of Research Articles for Select Major Search Engines in the Field of Library & Information Science." *IASET: International Journal of Library & Educational Science (IASET: IJLED) 2.2 (2016): 61-70.*

Patel, Preeti. "Security of Information with Biometric Applications." *IASET: International Journal of Library & Educational Science (IASET: IJLES) 2.2: 81-88.*

Bhatia, KETKI K., and M. A. Y. A. N. K. Trivedi. "MOOCs: Paradigm shift for libraries." *International Journal of Library & Educational Science 4.4 (2015): 67-80.*

CHAKRABARTI, ABHIJIT. "IMPLEMENTATION OF SIX–SIGMA IN ACADEMIC LIBRARIES FOR INCREASING SERVICE QUALITY." *International Journal of Library Science and Research (IJLSR) 4.6, Dec 2014, 1-12*

Sivakumar, N., and N. Tamilselvan. "Newspaper Reading Habits of College students: A case study of Kalaingar Karunanidhi Institute of Technology." *International Journal of Library science and Research* 5.2 (2015): 31-36.

Hariharan, S., and M. Tamizhchelvan. "Ethics of Library Professionals' Relation with Organisation and Professionalism in Self Financing Engineering College Libraries in Tamilnadu." *International Journal of Library Science and Research (IJLSR)*, 6 (3), 33 38 (2016).