*International Journal of Advanced Engineering, Management and Science (IJAEMS)*
*Peer-Reviewed Journal*
*ISSN: 2454-1311 | Vol-9, Issue-7; Jul, 2023*
*Journal Home Page: https://ijaems.com/*
*Article DOI: https://dx.doi.org/10.22161/ijaems.97.2*

# Needs of Data Mining in the Field of Library and Information Science: An Overall View

Anamika Shrivastava[1], Dr. Golam Ambia[2]

[1]Deputy Librarian, Prestige Institute of Management and Research, Gwalior, India
[2]Associate Professor, Sunrise University, Alwar, Rajasthan, India

Received: 28 Jun 2023; Received in revised form: 18 Jul 2023; Accepted: 25 Jul 2023; Available online: 31 Jul 2023

***Abstract***— *Increasing no of information sources, huge amount of data generate day by day, but unorganized data is not valuable to any use. Similarly, the users of the library have also become very demanding, and want the information and their needful data without any delay. Data mining has stepped promote organizational selection-making thru understanding statistics analyses. The data mining techniques that support these analyses may be separated into main functions; they are competent to either express the goal dataset or they can expect results via the usage of gadget mastering algorithms. These techniques are used to prepare and filter data. This paper short discuss about data mining and its need, application in the digital library to improve service quality.*

***Keywords***— ***Data Mining, Need of Data Mining, use of data mining Advantage and disadvantage of data mining future of data mining in field of library services.***

## I. INTRODUCTION

In this digital world availability of huge amount of Information gives a big challenge to LIS professionals. They have a big task to manage this information which is from many sources. The main responsibility for LIS professionals to manage all these information and select right information then provide library users. It is also a big task to collect information, manage and retrieval in right way. So now we need the technique that can reach the information and analyze information than review and provide user the particular information. Data Mining is a process of discovering interesting patterns and knowledge from large amounts of data. With the advancement of computer and network technologies,

a new Internet era by information and knowledge is ushering in. A new generation of tools and methodologies is urgently needed to help utilize the treasures of information and purify them into valuable knowledge. Knowledge discovery relies heavily on data mining. In the digital library, its techniques and processes offer a lot of application area and value. Data mining technology can aid in the development of huge volumes of data in detail, as well as the extraction of the inherent link between heterogeneous data in order to promote the digital library. This study defines data mining technologies, introduces the data mining process.

## II. LITERATURE REVIEW

Jiban K Pal (2011) found their study Data mining is a technique that could be to discover new that means in data, performs processing victimization refined data search capabilities and applied math algorithms, which can be utilized in any organization or system that must confirm the patterns or relationships implicit in an oversized

information warehouse for higher methods to best reach them. It are often fairly helpful to any company industries, monetary establishments, retailers, pharmaceutical companies, security agencies, government departments, on-line service suppliers, libraries, and individual researchers too.

Kovacevic, Ana and others (2009) found an answer for recommending digital library users a service from the library, primarily based not solely on applied math significance of service usage, but also considering the users' profiles. Our main analysis was targeted on serving to users to find relevant material additional simply. We have a tendency to achieved it by mistreatment data processing techniques on historical knowledge and by recommending the services that similar users would opt for. We first clustered the users supported their profiles at the side of their search behavior

Shantashree Sengupta (2017) concluded that there is requiring of records mining strategies that will remodel and simplify the working of library like classification, acquisition, circulation and referencing. Therefore, systematic efforts need to take vicinity to strengthen the application of records mining methods and algorithms for library databases. Also, it needs to be remembered that equipment for Data Mining are very effective and they require very skilled specialist who can put together the statistics and recognize the output.

Bedadyuti Sahoo and B. S. P. Mishra (2015) stat that the involvement of understanding people will make the statistics mining more relevant and fee added. It is additionally understood that, the essential factors of facts mining is to share the know-how and join thru planning to enhance the statistics enterprise for the organizational excellence and actuation.

Sherry Y. Chen and Xiaohui Liu (2004) focused in his paper on three primary useful techniques,

These are digital commerce, customized environments, and search engines. It must be referred to that data mining has additionally been utilized to different utility domains, such as bioinformatics, digital libraries, and web-based learning, etc. It is some other course for future lookup to look into what foremost features are required for each utility area and to enhance concrete standards for the assessment of their effectiveness.

Dwivedi, Roopesh K. and R.P. Bajpai (2004) consider that there's the necessity of information mining techniques which will plan and modify the working of library like classification, acquisition, circulation and referencing. the most use of information mining is in referencing however it is used for a few different work of library additionally. thus it's desperately required that systematic efforts are turn up to develop data processing techniques and algorithms for library database

## III. NEEDS OF DATA MINING IN DIGITAL LIBRARY ENVIRONMENT

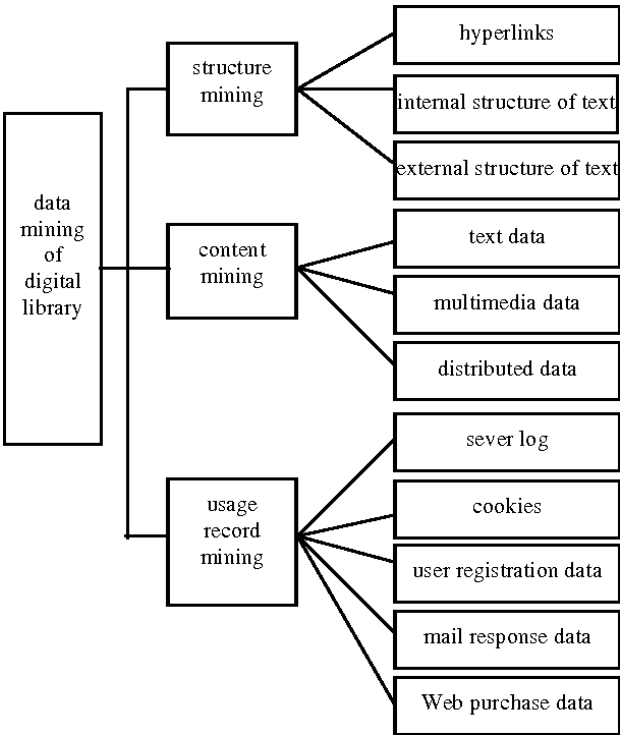
Now due to advance techniques data increased in a huge amount so the information collection and distribution atmosphere of the library are fully changed in to digital way. We can say about virtual library, a library wherein you may find out digital repositories, virtual collections, on-line databases, pix, audio, video, virtual documents, or special digital media codecs. Gadgets can incorporate digitized content material cloth collectively with prints or photographs, as well as regularly produced virtual content which includes word processor files or social media posts. Further to storing content material, digital libraries provide a way to prepare, are looking for, and retrieve the contents of a set. LIS specialist's ought to all digital competencies to maintain the virtual library similarly to technical. Library body of workers has information of hardware and software program to offer virtual, digital, and virtual services using virtual collections to the users. .as we all known According to the fifth law of library science, "the library is a growing organization" with invent of information science the number of library data is also rapidly increasing, for giving to user efficient service provides with short notice is the basic needs of successful library management and the extension of library services.

This is where library automation and e-Library come into play. However, automating the library or creating an e-Library is not enough. Unless and until we are able to uncover the hidden information from the enormous datasets, this is not the only option. the size of the database This can be accomplished by using data mining to analyses library data.

- Classification: - classification is the most important work in library management which is more time taking process but with the help of data mining we can design a computer software that will replace manual classification of library contents with automatic classification using data mining.
- Link analysis- Likewise the paper materials, where similar documents tend to have similar bibliographical references, and frequency of citation is often considered to reflect the quality or importance of document, link analysis assumes that higher-quality or otherwise more desirable
- documents will generally be linked to more frequently than other documents, and that links in

ac Document reveal something about the content of a document. Link analysis can place frequently

- linked-to-documents at the top of a list or identify documents that are associated with each other Sequence analysis- Sequence analysis uses statistical analysis to identify unlinked documents that users are likely to want to read together. It examines the paths that users follow when searching for information and can help identify which documents users are likely to want together
- Summarization- Though machine generated abstracts are inferior to human-generated ones in terms of readability and content, yet they can be very useful for helping users decide what items they need. Abstract-generating software typically works by identifying significant words or phrases based on position within documents association with critical phrases.
- Clustering- Clustering is similar to classification, except that the classes are determined by finding natural groupings in the data items based on probability analyses rather than by predetermined groupings. Clustering and classification are often used as a starting point for exploring further relationships in data. For example, many search ..




**Process of data mining**

The concept of Data mining has been with us on the grounds that long before the digital age Statistics mining is a finished interaction which mines difficult to understand, possible and affordable statistics from big statistics sets. Information mining technique contains 4 levels, recognizable evidence vicinity items, facts association, mining interaction and consequences articulation and research. The records mining cycle may be separated into three stages: data readiness, mining hobby, articulation and expertise of the results. The mining process is rehashed brief, which can't manipulate without the cooperation of the clients. Facts mining causes the mining of the huge data in full-size facts sets to grow to be less complicated, and the mining body of workers do not must undergo lengthy stretches of getting ready of the measurable or information examination. There exist precise contrasts and family members between records mining and statistics revelation. By using and big, information mining is a selected develop in the course the time spent records disclosure.

**Advantages of Data Mining**

- With the help of data mining organizations build up dependable statistics
- It's an emerald, value-effective solution compared to other facts applications
- It helps corporations make valuable manufacture and operational modifications
- data mining makes use of both new and legacy systems
- It facilitates organizations make knowledgeable decisions
- It facilitates locate credit score risks and fraud
- Data mining gives helps to records scientists smoothly examine big amounts of statistics rapidly.
- statistics scientists can use the facts to discover fraud, build danger models, and enhance product safety
- Data mining easily initiate automated predictions of behaviors and trends and discover hidden patterns

**Drawbacks to Data Mining**

Nothing's ideal, along with data mining. Some are the major problem in data mining.

- many data analytics device are complicated and hard to apply. Information scientists want the proper training to use the equipment successfully.
- Talking of the mechanism, one-of-a-kind ones work with varying styles of statistics mining, relying on the algorithms they employ. As a

result, information analysts need to be sure to select appropriate equipment.

- Records mining strategies aren't infallible, so there's usually the hazard that the records aren't absolutely accurate. This impediment is especially relevant if there's a loss of range within the dataset.
- Companies can probably promote the customer records they've gleaned to other businesses and agencies, raising privations issues.
- Records mining calls for huge databases, making the system hard to control.

## IV. DATA MINING FOR LIBRARY AND INFORMATION SCIENCE

In present digital world of information, we are all have encircled by huge amount of Substantial data, which percentage increased day by day. it may happen that one day we are confused in the right information, but we all are interested in to right information). The principal purpose at the back of this, all this information creates noise which makes it tough to mine. In short, we've generated heaps of amorphous records however experiencing failing big records initiatives because the useful records are deeply buried interior. Consequently, without effective equipment along with information mining, we cannot mine such records, and as a end result, we can no longer get any benefits from that facts. The major need of data mining in the Digital Library in recent years is due to the wide availability of huge amounts of the information and data and the imminent need for turning such data into useful information and knowledge. Data mining is the process of discovering interesting knowledge from large amount of data stored either in database there is a need of information mining methods that will overhaul and improve on the working of library like arrangement, securing, flow and referring to. Consequently, precise endeavors should happen to create the use of information digging procedures and calculations for library data sets. Additionally, it should be recollected that apparatuses for Information Mining are exceptionally strong and they require extremely talented expert who can set up the information and get the result. Information Mining draws out the examples and connections, yet at the same the importance and legitimacy of those examples should be made by the client.

We also understand data mining as other way by knowledge discovery in data (KDD), is a simple process to filter data in short to very large amount of data. Given the evolution of statistics warehousing era and the increase of massive statistics, adoption of facts mining strategies has unexpectedly extended over the past couple of many years, helping companies by transforming their uncooked records into beneficial understanding. But, regardless of the fact that that era continuously evolves to deal with information at a big-scale, leaders still face demanding situations with scalability and automation.

Without a doubt said, information mining refers to extracting or "mining" knowledge from big amount of records. Many humans treat information mining as a synonym for every other popularly used time period information Discovery in database, or KDD.

Need of Data Mining It tends to be sensibly valuable to any corporate enterprises, monetary foundations, Retailers, drug firms, security offices, government divisions, online specialist organizations, libraries, and individual scientists as well. It very well may be utilized for an assortment of utilizations in both public and private areas. Corporate enterprises and monetary establishments frequently use information mining to increment deals, lessen costs, further develop market execution, improve client base through creating models for credit scoring, risk evaluation, misrepresentation location, and so forth. In the past decade, data mining changes the discipline of information science, which investigates the properties of

Information and the methods and techniques used in the acquisition, analysis, organization, dissemination and use of information there is a wide range of data mining techniques, which has been successfully used in the field of information science.

**Future of data mining in the library working**

In future Data Mining can provide the new road map for the next generation of library by applying it for the following activities of library.

- Searching of Information (Reference Service)- Since the data of the library continuously growing with an exponential rate and the main problem is how one can reference the required information from the large amount of redundant information of the library. This can be possible by applying data mining techniques, so one can say that the data mining is the future of reference service.
- Classification- It will replace the manual classification of content of the library with the computer assisted classification, so that the classification task can be accomplished by less skilled person in a fast and efficient way. This will simplify the classification task of the library.
- Acquisition- As per third law of library science "Every book its reader". By applying the data

mining in the library data it can be easily find out the required contents that are necessary to acquire next. This will reduce the work of library staff related to the acquisition as well as the efficient use of budget allocated to the library.

## V. CONCLUSION

It concluded that the information mining may be viewed because of the herbal evolution of statistics era. To assist DL clients with getting treasured facts all the more efficaciously, we will make use of data mining techniques. Due to the fact records mining strategies are extraordinarily famous; several professionals have implemented them in one-of-a-kind spaces. In any case, few are focused across the place of DLs. Our principle goal is to make use of records mining techniques to prescribe explicit administrations to DL clients. As library is the storage facility of statistics moreover, the spot for dispersing records to the customers, it's far generally essential to make information mining in the ideal manner and use it to help the diverse sorts of clients in numerous angles. The association and synthesisation of data utilizing becoming programming is typically enormous as logical programming are effectively on hand looking for records dissecting and measurable obstruction

## REFERENCES

[1] Pal Jiban K.,Usefulness and applications of data mining in extracting information from different perspectives, Annals of Library and Information Studies Vol. 58, March 2011, pp. 7-16

[2] Ana Kovacevic, (2009) Using data mining to improve digital library services, The Electronic Library Vol. 28 No. 6, 2010 pp. 829-843 q Emerald Group Publishing Limited

[3] Dr. Shantashree Sengupta, Applications of data mining in library & information centres: an overview, International Journal of Current Research Vol. 9, Issue, 01, pp.45246-45249, January, 2017

[4] Leticia, Tweve. "Effects of Customer Satisfaction on Water Utility Business Performance: A Case Study of Mbinga Urban Water Supply and Sanitation Authority." International Journal of Engineering, Business and Management, 2022. https://doi.org/10.22161/ijebm.6.5.5.

[5] Bedadyuti Sahoo and B. S. P. Mishra, (2015)DATA MINING IS A PERPETUAL CONCEPT FOR LIBRARY AND INFORMATION SCIENCE: AN ESTIMATED VERVIEW, International Journal of Digital Library Services, Vol. 5, Issue 3, pp 15-21

[6] Sherry Y. Chen and Xiaohui Liu , (2004)The Contribution of Data Mining in

[7] Information Science Journal of Information Science ,

[8] Dwivedi, Roopesh K. and R.P. Bajpai, Use of Data Mining in the field of Library and Information Science :

[9] An Overview, 2nd International CALIBER-2004, New Delhi, 11-13 February, 2004 oo512-519

[10] Dhiman, Anil K., Data Mining and its use in Libraries, CALIBER-2003.

[11] http://www.ijodls.in/uploads/3/6/0/3/3603729/2535.pdf

[12] https://www.ibm.com/cloud/learn/data-mining

[13] https://ir.inflibnet.ac.in/bitstream/1944/375/1/04cali_67.pdf

[14] https://www.javatpoint.com/types-of-data-mining

[15] https://link.springer.com/chapter/10.1007/978-3-642-53703-5_30

[16] https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.426.2547&rep=rep1&type=pdf

[17] https://www.journalcra.com/sites/default/files/issue-pdf/20191.pdf

[18] https://www.lisedunetwork.com/data-mining/

[19] https://www.talend.com/resources/what-is-data-mining/

[20] https://www.simplilearn.com/what-is-data-mining-article

[21] https://www.semanticscholar.org/paper/Application-of-Data-Mining-Technology-in-Digital-Zhang/30e149d089975dae658c2fec9b62a77f4cf81316#citing-papers