Data Mining @ Information Age

M. Arunmozhi[1], R. Usha[2]
[1]Research Scholar, Bharathiar University, Librarian, KCS Kasi Nadar College of Arts and Science
[2]Librarian and Information Assistant, Archaeological Survey of India
E-Mail: marunmozhikutty@gmail.com, usha.ramur@gmail.com /ushangeorge@gmail.com

Abstract - In the information age, data turns to be the vital. Hence it is important to understand the data in order to face the future information challenges. This paper deals with the importance of data mining while explaining the concepts and life cycle involved. It extracts the basic gist of the topic presented in a user-friendly way. Further, in developing different stages of data mining followed by its extended application usage in practical business platform.

I. INTRODUCTION

Since, the information gateways are exploded with lots of information, its retention and regeneration is inevitable. In this information age, generation of new information happens rapidly and frequently. Hence, there is a dire need to analyse, summarize and interpret the developing information to improve business prospects. Hence, data mining plays a key role in analyzing the information in all possible dimensions and angle to understand the explore them. Data miners extract and process the data in a larger scale. This helps identify hidden relationships, different patterns and its associated inter-link and general norms that can foresee the correlations which can ease the business or research decisions and quick and authentic.

II. INFORMATION AGE

The patterns, associations, or relationships among all this info can offer information. maybe, analysis of retail purpose of sale dealings info can yield information on it product ar commerce and once. This includes:

- Operational or transactional info love, sales, cost, inventory, payroll, and accounting
- Reserve info, love trade sales, forecast info, and macro-economic information
- Meta info - info relating to the knowledge itself, love logical info vogue or information reference definitions

III. DEFINITION

Data may well be a method of developing the collected information into useful information. That information successively is used by the massive companies or info seekers to boost their business wings or content. However, processing does not restricted in extracting data however exploring {the information the knowledge the information} sets or patterns from a voluminous data for any analysis and work station implementation. The need for adapting processing has varied reasons.

IV. INFORMATION EXPLOSION

Information Explosion here suggests that not restricted to the pooling of information but its volume in terms of dimensions to boot matters. Limitations of Human Analysis Whenever somebody's brain processes the knowledge, these we've got an inclination torus the problems we could end up with a looking for advanced multifactor dependencies in info, and conjointly the dearth of judgment in such associate analysis.

Styles of Data Mining: Directed Data Mining and Undirected Data Mining
Directed data mining is a kind of queue approach (top-down) approach. This style is used to only when we know what are we looking for and the predicted results. Undirected data mining is a kind of stack approach (bottom – up) approach. This style involves patterns but decisive to the user point of view.

V. CYCLE OF DATA MINING PROCESS

1. Problem Definition
2. Data Preparation
3. Modeling
4. Evaluation
5. Deployment

Before we get exploring the stages of data mining, the data mining cycle needs to be understood. Basically, any data miners should be able to understand the problem and define the problem. This first steps almost common to data analytics and big data as well. After identifying the problem, the gathering of data takes place. Data can gathered from the readily available internal sources such as customer database and other different external sources. All those gathered data are stored in a sing place and common format to use. Next is to explore the data for evaluation and cleaning the data. This evaluation scrutinizes the data in all different angles and cleanses anything that is to necessary and inaccurate in terms of critical data fields.

Stages of the data-mining process
- Information gathering, e.g., info reposition, web crawl.
- Information cleansing: eliminate errors and/or phoned info.
- Feature extraction: obtaining only the attention-grabbing attributes of the knowledge, e.g., \date acquired" is probably not useful for agglomeration celestial objects, as in Sky cat.
- Pattern extraction and discovery. usually this will be} the stage that is usually thought of as \data mining," and is where we've got an inclination to shall concentrate our effort.

VI. APPLICATION

Data mining can be applied in almost all different disciplines including inter-disciplinary areas. This plays a vital role in

research and development area to arrive into a new set of data based on the already existing data from large databases.

Usage Processing

User information in any organization helps identify the patterns of user and user likely to turn non-user. Since data mining in this area extracts the usage statistics and other related information to draw a clear picture of what is happening with the business. It also compares the patterns with other related databases to obtain clarity.

Electronic Commerce

Electronic Commerce is a boon to the business environment. Because this had transformed the marketing fields and imposed a serious challenge to those who are lacking technological advancement skill set. This helps organizations to analyse the customers' behaviour developing marketing strategies to promote business in many avenues. Electronic Commerce embraces the following:
1. Consumer management
2. Retail business
3. Data point analysis.

Search Engines

Search engines are very much important in the role of data mining at this information era. Since the advent of internet and followed by technologies, many traditional search engines were playing an efficient role. However, it provides only limited assistance to the users who could be able to trace the relevant information. But, data mining search engines are developed with more advanced features to face the rising challenges in the current information age. Some advantages are:
(a) Ranking of pages
(b) Improvement of precision
(c) Citation analysis

VII.CONCLUSION

Within the past 10 years, there has been significant progress in the field of information science. Some of this progress represents improvements in existing techniques. One of these techniques is data mining, which can search for interesting relationships and global patterns from various types of resources. These relationships and patterns represent valuable knowledge about the objects and this is reflected by many applications in the field of information science. Data mining is widely used in many applications. The paper focuses on three main application domains, including electronic commerce, personalized environments, and search engines. It should be noted that data mining has also been applied to other application domains, such as bioinformatics, digital libraries, and web based learning, etc. It is another direction for future research to investigate what major functions are required for each application domain and to develop concrete criteria for the evaluation of their effectiveness.

References

[1] Agrawal, R., & Srikant, R. Fast algorithms for mining association rules. Proceedings of the 20th international conference on very large databases (1994), Santiago, Chile.

[2] Ankerst, M.. Visual Data Mining with Pixel-oriented Visualization Techniques. Proceedings of ACM SIGKDD Workshop on Visual Data Mining (2001).

[3] Bigus, J. Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support. New York: McGraw-Hill (1996).

[4] Borko, H. (1968) Information Science: What is it?. American Documentation. 19(1), 3-5.

[5] Bose, I. & Mahapatra, R. K.. Business data mining - a machine learning perspective. Information & Management. 39(3) (2001) 211-225.

[6] Brin, S. & Page, L. The Anatomy of a Lage-Scale Hypertextual Web Search Engine. Proceedings of the 7th International World-Wide Web Conference (1998), pp.107-117.

[7] Chakrabarti, S. Berg, M. van den, Dom, B. Focused Crawling: A New Approach for Topic-Specific Resource Discovery, Proceedings of the 8th International World Wide Web Conference (1999).

[8] Chakrabarti, S., Dom, B. E., Kumar, S.R., Raghavan, P., Rajagopalan, S., Tomkins, A., Gibson, D. & Kleinberg, J.M. Mining the web's link structure. Computer. 32 (1999) 60-67.

[9] Changchien, S., & Lu, T. Mining association rules procedure to support on-line recommendation by customers and products fragmentation. Expert Systems with Applications. 20(4) (2001) 325-335

[10] Chen, C. and Paul, R.J. Visualising a knowledge domain's intellectual structure. Computer, 34(3) (2001) 65-71.

[11] Chen, H. C. and Chen, A.L.P. A music recommendation system based on music data grouping and user interests. Proceedings of the CIKM'01, Atlanta, Georgia (2001), pp. 231–238.

[12] Chen, Y. L., Tang, K., Shen, R. J. & Hu, Y. H. Market basket analysis in a multiple store environment. Decision Support Systems, forthcoming.