INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

Analysis on Data Mining Applications and Trends

Apurva Sharma[1], Dr. I.S Hudiara[2]

[1]Apurva Sharma, Assistant Professor, MCA
IET Bhaddal, Ropar, Punjab

[2] Dr I.S Hudiara
(BE, MS Engineering, PhD, PDF, MBA-Finance), FIETE

Abstract: *In the present era information technology is used in every field which has led to the large quantity of data stored in various formats like records, documents, images, audio, video and many more other formats which are analyzed through various dimensions and the result is abridged into information used for decision support . For extracting knowledge or information from the large data repositories or from data warehouses, data mining is playing a vital role. The objective of this paper is to survey past, present and future trends in the field of data mining, with a focus on those which are thought to have the most promising and applicability to future data mining applications.*
Keywords: *Data Mining, Applications, Past, Present and Future data mining trends.*

1. Introduction

The development of Information Technology has generated large amount of databases and huge data in various areas. The research in databases and information technology has given rise to an approach to store and manipulate this precious data for further decision making. The data collected from various sources need a proper mechanism of extracting knowledge or information for making better decisions in the management of electronically managed stores and warehouses. There is a need to convert this data into information by analyzing it from various dimensions, identifying the relationships or grouping the data into categories [1].
Data mining is the process of finding the useful information from the large amount of data. The field of data mining has posed into new areas of human life with various integrations and advancements in the field of statistics, databases machine learning etc. The areas where we can apply the data mining techniques are customer relationship management, climate modeling, banking, retail etc. Hence this paper reviews the various trends and techniques of data mining.

2. Data Mining

The function of data mining is to discover hidden knowledge from large volumes of data stored in data warehouse. Each database contains a huge number of data and information. Extracting hidden knowledge form stored data and information in databases can be vital and important in business and marketing. This extracted knowledge can help and support companies and organizations to make better and more intelligent decisions [2].
Data Mining is an essential step in the knowledge discovery in databases process. The process of discovering useful knowledge from huge data is called knowledge discovery in database. The following steps are adopted for data knowledge discovery:

- Selection: Selecting data relevant to the analysis task from the database.

- Preprocessing: Removing noise and inconsistent data, combining multiple sources.
- Transformation: Transforming data into appropriate format to perform data mining
- Data mining: Choosing a data mining algorithm which is appropriate to pattern in the data; extracting data patterns.
- Evaluation: Evaluate the patters in knowledge by removing redundant or irrelevant patterns[3].

3. Applications of Data mining

Data mining is widely used in diverse areas. There are a number of commercial data mining systems available today and yet there are many challenges in this field. Some of the applications are:

- **Retail Industry:** Retail industry is having a huge amount of data of customer details, sales, stock etc. The use of data mining in retail is used for identifying the customer retention, customer satisfaction, improves the quality of customer service etc.

- **Telecommunication:** The telecommunication industry provide various services to their customers like voice, fax, pager, cellular phone, images, e-mail, computer and web data transmission and other data traffic. This creates a great demand from data mining in order to help understand business involved, make better use of resources, better quality of service, catch fraudulent activities and identify telecommunication patterns.

- **Finance:** Now a day's all the banks are giving the banking services to their customers like investment services, credit card services etc. The data which is collected in any bank is often complete , reliable and high quality which facilitates systematic data analysis and data mining.

- **Higher Education:** Higher education faces a lot of challenges like who will need additional assistance in order to graduate? Which student will enroll in particular course programs. Data mining has quickly emerged as a

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

highly desirable tool for using current reporting capabilities to uncover and understand hidden patterns in vast databases.

- **Text Mining and Web Mining:** The process of searching large volumes of documents from certain keywords or key phrases is known a Text mining. By the use of text mining we can easily derive certain patterns in the comments that may help identify a commonest of customer perceptions which is not captured by the other survey questions. Web mining is known as extension of text mining web mining integrates data and text mining within a website. It enhances the web site with intelligent behavior, such as suggesting related links or recommending new products to the consumer [4].

4. Past Trends in data mining

- **Machine Learning & Artificial Intelligence :** Artificial intelligence is built upon heuristics as opposed to statistics, and attempts to apply human-thought like processing to statistical problems. AI found a few applications at very high end scientific, government markets, but the required supercomputers of the era priced AI out of the reach of virtually everyone else. In the evolution of AI we can consider Machine learning, because it merge AI heuristics with advanced methods of statistics. It lets computer programs learn about the data they study and then apply learned knowledge to data.
- **Data Warehouse:** Enormous amount of data needs to be stored in a repository, and it also needs to be managed in an efficient manner. So, databases come into existence. Earlier, data was not huge so, data was managed in records and fields. In data mining, data is in bulk, so we need specialized servers for it. We call the term as Data Warehousing. Data warehousing also supports OLAP operations in order to support decision making [4].
- **Statistics:** There will be no data mining without statistics .Statistics are the substructure of most technologies on which data mining is built, for example, , standard distribution, standard deviation, discriminate analysis, regression analysis, cluster analysis, and confidence intervals, standard variance,. Certainly, classical statistical analysis plays a significant role within the heart of data mining tools and techniques.

5. Ongoing Trends

Data mining is formally defined as the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [5]. The field of data mining has been growing rapidly due to its broad applicability, achievements and scientific progress. A number of data mining applications have been successfully implemented in various domains like fraud detection, retail, health care, finance, telecommunication, and risk analysis...etc. are few to name. Advancements in data mining with various integrations and implications of methods and techniques have shaped the

present data mining applications to handle the various challenges.

Following are the current trends of data mining applications:

A. Bio-informatics: The most important application trend, deals with interpretation and mining of biological structures. Data mining tools are rapidly being used to cure of diseases like cancer and AIDS.

B. Brawl against Terrorism: After 9-11 attacks, new laws were imposed by many countries against fighting terrorism. These laws allow intelligence agencies to effectively fight against terrorist organizations.USA launched total Information awareness program with the goal of creating a huge database of that consolidate all the information on population. This program faced several problems, such as

a) The heterogeneity of database, the target database had to deal with text, audio, image and multimedia data.

b) The another problem was algorithms scalability.

C. Semantic Web: Web is totally unstructured so for this data mining is helping organized a web, which is called semantic web. The underlying technology is Resource Description Framework (RDF) which is used to describe resources.

D. Business Trends : Today's business environment is more energetic, so businesses must be able to react quicker, able to satisfied customer's need, must be more profitable, and offer high quality services to their customer's that ever before. Here, data mining serves as a fundamental technology in enabling customer's transactions more accurately, faster and meaningfully [4].

6. Future Trends

i) Hypertext and Hypermedia data mining: The hypertext and hypermedia data is a collection of data from online catalogues, digital libraries, and online information data bases which include hyperlinks, text markups and other forms of data. Some of the important data mining techniques used for hypertext and hypermedia data mining include classification (supervised learning), clustering (unsupervised learning), semi-structured learning, and social network analysis.

ii) Ubiquitous data mining: UBIQUITOUS DATA MINING (UDM) The origin of laptops, palmtops, cell phones, and wearable computers is making ubiquitous access to large quantity of data possible [4]. The Ubiquitous computing environments are subsequently giving rise to a new class of applications termed Ubiquitous Data Mining (UDM). UDM is the process of analysis of data for extracting useful knowledge from the data of ubiquitous computing [6]. Traditional data mining techniques that are drawn from the combination of ML and Statistics are presently employed in ubiquitous data mining [7].

iii) Multimedia data mining: Multimedia Data Mining is the mining and analysis of various types of data, including images, video, audio, and animation. As multimedia data mining incorporates the areas of text mining, as well as hypertext/hypermedia mining, these fields are closely related to Multimedia information, because its nature as a large collection of multimedia objects, must be represented

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

differently from conventional forms of data[4].The data mining techniques that are applied on multimedia data are rule based decision tree classification algorithms like Artificial Neural Networks, Instance-based learning algorithms, Support Vector Machines ,association rule mining, clustering methods etc.[8].

iv) Geographical or Spatial data mining: The spatial data includes astronomical data, satellite data and space craft data. Some of the data mining techniques and data structures which are used when analyzing spatial and related types of data include the use of spatial warehouses, spatial data cubes, spatial OLAP, and spatial clustering methods [8].

v) Temporal or Time Series data mining: Temporal Data Mining is a rapidly area of research that is at the intersection of several disciplines, including statistics (e.g., time series analysis), temporal pattern recognition, temporal databases, optimisation, visualisation, high-performance computing, and parallel computing. Typical examples include stock prices, currency exchange rates, the volume of product sales, biomedical measurements, weather data, etc, collected over monotonically increasing time. Rule induction algorithms such as Version Space [7], AQ15 [8], C4.5 rules are presently employed in Time series data mining applications [9].

7. Conclusion

In this paper various data mining trends and applications have been reviewed from its inception to the future. This review puts focus on various areas of data mining. Though only few areas are named here in this paper, yet they are those which are commonly forgotten. This paper provides a new outlook of a researcher regarding applications of data mining.

References

[1] Soni,S.(2015).A Literature Review on Data Mining and its Techniques. Indian Journal of Applied Research 5(6).

[2] Behrouzian Nejad, M.E. ,B Behrouzian Nejad ,H.S Nasab and N. Mousavian Far,2010.The role of customer relationship management in the organizations. Proceedings of 2010 1st Regional conference of Management, Young Researchers Club, Shoushtar Branch, Iran.

[3] Janakiraman, S., Uma Maheshwari, K.(2014). A Survey on Data Mining Techniques for Customer Relationship Management, International Journal of Engineering Business and Enterprise Applications7 (1): 55-61.

[4] Paidi A.N (2012). Data Mining: Future Trends and Applications. International Journal of Modern Engineering Research 2(6)

[5] Knowledge Discovery in Databases, AAAI Press / the MIT Press, Massachusetts Institute of Technology. ISBN 0– 26256097–6. MIT1996.

[6] Hsu, J. 2002. Data Mining Trends and Developments: The Key Data Mining Technologies and Applications for the 21st Century, The Proceedings of the 19th Annual Conference for Information Systems Educators (ISECON 2002), ISSN: 1542-7382. Available Online: http://colton.byuh.edu/isecon/2002/224b/Hsu.pdf.

[7] Shonali Krishnaswamy. 2005. Towards Situation awareness and Ubiquitous Data Mining for Road Safety: Rationale and Architecture for a Compelling Application (2005), Proceedings of Conference on Intelligent Vehicles and Road Infrastructure 2005, pages-16, 17.

[8] Kotsiantis, S., Kanellopoulos, D., Pintelas, P. 2004. Multimedia mining. WSEAS Transactions on Systems, No 3, s. 3263-3268.

[9] http://www.tutorialspoint.com/data_mining/dm_applications_trends.html.