Applications of Database Management in Reducing Workload of Processing Data

Mohd. Ahtesham Farooqui

Research Scholar (Computer Science), CMJ University, Shillong, Meghalaya

Abstract— Data Mining is one of the processes of analyzing the data or information from different perspectives and this is to discover the relationships among the separate data items. Data mining evolves from the discovery of regularities and frequent patterns in large data sets toward user-oriented, interactive and on-demand decision supporting. The data to be mined is generally located in the database and this is promising idea of integrating the data mining methods into DBMS (Database Management Systems). Today, data mining was very popular in the industry of Information Technology. This research discusses about the applications of data management system in order to reducing the workload of processing data in the organizations.

Index Terms— Database Management System (DBMS), Data Mining, Competitive Advantage

----------♦----------

1. INTRODUCTION TO DATABASE MANAGEMENT SYSTEM

A database is an integrated collection of fields, records, files and other objects. The Database Management System enables user to store, extract and modify the information or data from the database. A DBMS (Database Management System) is one of the software packages with computer programs which can have capacity to control the maintenance, use and creation of the database (Date, 2005). Database Management System is a set of programs which enables to extract, store and modify the information or data from the database and in addition to these, provides users with tools to delete, modify, access, analyze and add the data that stored in one location.

Database Management System provides the methods and technique to maintain the integrity of stored data, user access, security and recovering information or data if the system fails. Database Management Systems (DBMS) and applications provide the understanding of databases functionality and their role in the modern business environments. Industries introduced various concepts like relational conceptual data modeling, implementation issues, data management, database design and advanced database applications to perform the business functions. A fundamental and general knowledge of such concepts effectively deploy the commercial DBMS (Database Management System) in order to provide the needs of business organization.

2. INTRODUCTION TO DATA MINING

Data mining is a collection of process for analyzing and searching the meaningful information or data taken from large amount of data sets. The data mining is a technique of processing large volumes of data that is generally stored in a database, for searching new patterns and relationships contained in that data (Dartigue, Jang and Zeng, 2009). In other words, data mining is software that helps to analyze the data or information from large volumes of data stores and also to find out the new trends and techniques for the business activities. Data mining is the business processes that have capacity to interact with other business processes.

Data mining enhances several technical approaches such as clustering, learning classification rules, finding dependency networks, and summarization of data, analyzing alterations and predicting anomalies (Jiawei Han and Kamber, 2006). Data mining is used to discover the interesting and frequent patterns from huge amounts of data. Data mining is also known as KDD (Knowledge Discovery (mining) in Databases), information harvesting, data/pattern analysis, data dredging, knowledge extraction, business intelligence, data archeology, etc.

The business intelligence is mainly used in information technology industry to support the decision making process in the organization. The business intelligence has capacity to extract the information from the data and this process is called as the data mining. Data mining uses statistical

analysis to extract interesting events and trends, support decision making, create useful reports, etc (Berry, Michael and Linoff, 1997). Data mining exploits the huge amounts of data to achieve operational, scientific and business goals.

Data mining is the one of the processes to optimize the CRM (Customer Relationship Management) by enabling: customer acquisition; customer satisfaction; customer intimacy; customer up-sell cross; and customer retention (Gaber, Zaslavsky, and Krishnaswamy, (2005). Various applications are using data mining to achieve success in business operation. Some of the data mining applications are: stock market prediction; knowledge acquisition; credit assessment; fraud detection; fault diagnosis in the production systems; scientific discovery; medical discovery; target mailing; hazard forecasting; organizational restructuring; buying trends analysis and semantics based performance enhancement of the database management system (DBMS).

The following table illustrates the difference between database management system and data mining.

Area	DBMS	Data Mining
Task	Extraction of detailed and summary data	Knowledge discovery of hidden patterns and insights
Type of result	Information	Insight and Prediction
Method	Deduction (Ask the question, verify with data)	Induction (Build the model, apply it to new data, get the result)
Example question	Who purchased mutual funds in the last 3 years?	Who will buy a mutual fund in the next 6 months and why?

Table 1: Database management system Vs. Data Mining

Source: CCSU (2012): Introduction to Data Mining, retrieved on 3rd December 2012 from http://www.cs.ccsu.edu/~markov/ccsu_courses/datamining-1.html

3. ORGANIZATIONAL BENEFITS OF DATA MINING

Data Mining provides various benefits to the business organizations, society and individuals. The few benefits of the data mining for business organization are: time and money savings; finding new market opportunities; better CRM (customer relationship management); determine service and product features which are important to the customers; attracting and retaining the best customers; enhance productivity; reduce risk; develop insight to changing the customer requirements; distinguish preferred from other marginal customers; identify the customers interests about company's new services or products; target pricing; customize the marketing plans to particular markets; manage portfolios; and identify the customers who have high rate in purchasing particular service or products. In addition to these, data mining provides some benefits to individuals such as: understand the needs of customers; serve more customized products; better customer relationship; react to customers faster; rapid access to integrated information and systems. Data mining also provides benefits to society for identifying criminal activity and gaining intelligence information (Wang, 2003).

The following figure illustrates the data mining and business intelligence.

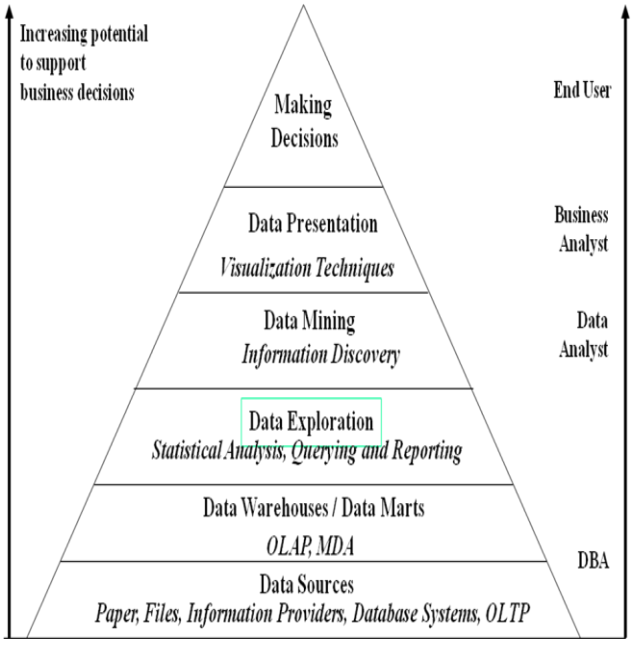

Figure 1: Data Mining and Business Intelligence

Source: Yang T H and Wang S H (2008): DBMS support of the Data Mining, retrieved on 3rd December 2012 from www.mis.nsysu.edu.tw/db-book/DBProject2007Fall/03/report.ppt

In addition to these, Data Mining provides some organizational benefits such as:

- Data mining involves the process of finding patterns that is useful for profitable business events.
- Data mining can be used to make analytical models for the decision making process.
- It is more flexible and data led discovery
- It helps organization to take strategic business related decisions.
- Business intelligence at many levels
- It provides timing feedback at each level of the organization
- It tends to satisfy various business needs (customer decision, effective decision making; competitive market, time saving)
- It helps organization in POS (point of sale) system, accounting and order processing system.

The following table illustrates the data mining challenges, solutions and benefits.

Business Challenges	Data Mining Solutions	Benefits
Understand the Characteristics of your customers	Segmentation and Profiling Analysis	Customer profitability and CRM optimization.
Attract new Customers	Target Response Model	Bring more customer for the same marketing costs
Up sell new Customers	CLTV Model	Identify long term profitability
Avoid High risk Customers	Risk and Approval Model	Avoid loss for the company
Make unprofitable customers become more profitable	Up-sell and cross-sell targeting Model	Increase profit from existing customers
Retain your profitable Customers	Retention or Churn models	Increase wallet share growth and overall profitability
Increase Customer Satisfaction	Market Research and customer profiling	Increase Retention and Dollar to dollar renewal
Increase Sales	Acquisition & Up-sell models	Increase Profits
Reduce Costs/expenses	Target Models	Attract customers
Win-back your lost customers	Win back Models	Increase Sales and profitability

Table: Data Mining Challenges, Solutions and Benefits

Source: Isson J P (2009), Data Mining: Implementation & Applications, Monster Worldwide.

4. MAJOR ISSUES IN DATA MINING

Data mining provides various advantages to the organization and other individuals. At the same time, it has some major issues.

Data mining has some major issues in mining methodology such as:

- Incorporation of background knowledge
- Distributed, incremental and parallel mining methods
- Pattern evaluation: interestingness problem
- Handling incomplete data and noise
- Performance: effectiveness, efficiency and scalability
- Knowledge fusion

According to Cook and Holder (2007), data mining has some major issues in user interaction:

- Ad-hoc mining and data mining query languages
- Visualization and expression of data mining results
- Interactive mining of knowledge at different and multiple levels of abstraction

According to Miller and Han (2001), data mining has some major issues in application and social impacts:

- Protection of data privacy, integrity and security
- Invisible data mining and domain-specific data mining

5. DATA MINING CHALLENGES

Data mining systems are facing with lot of pitfalls and problems. The data mining system will work good for the consistent data and also perform significant worse when there is a little noise that is added to the training set.

According to Yang and Wu (2006), there are some challenges in data mining and they are:

- Scaling up for high speed data streams and high dimensional data
- Developing the unifying theory
- Security, privacy and data integrity

- Mining time series data and sequence data
- Data mining in network setting
- Data mining for environmental and biological problems
- Data mining process related problems
- Distributed data mining
- Mining the complex knowledge from the complex data
- Mining multi-agent data
- Dealing with unbalanced, cost sensitive and non-static data

5. CASE STUDY

MobileMiner is a real mobile communication data set and this study discuss about the business solutions using the data mining technique. MobileMiner provides the platform for analytical tasks and here, the user profiles are extracted continuously from user's call and moving records. The user's profiles are very essential and valuable in the business. With profile mining platform, data mining tasks are effectively performed using the features of profiles.

The following figure illustrates the architecture of MobileMiner.

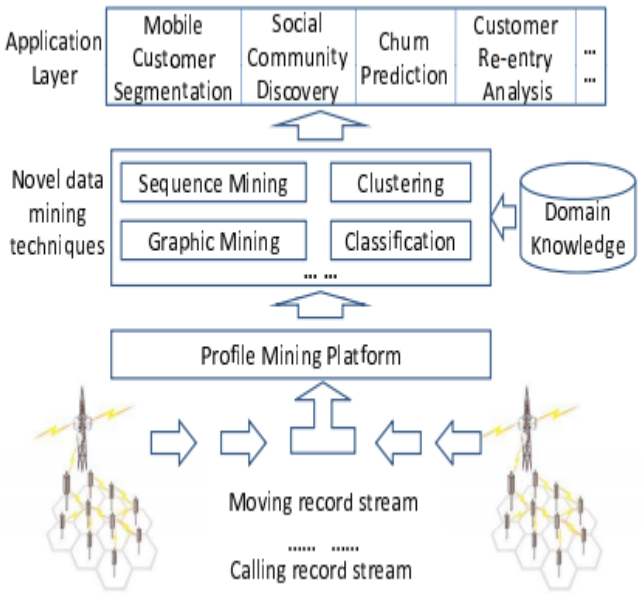

Figure: Architecture of MobileMiner.

Source: Wang et al (2009), MobileMiner: A Real World Case Study of Data Mining in Mobile Communication, SIGMOD'09, June 29–July 2, 2009, Providence, Rhode Island, USA.

The mobile user segmentation tasks will group the customers with the help of frequent moving patterns. The features that are used for grouping are obtained by the mining users' and identify the moving records continuously with profile platform. After knowing the moving patterns of various customer groups, the service provider will provide dynamically deploy resources and this to improve the service quality.

6. CONCLUSION

It can be understood that the main purpose of data mining is knowledge discovery and it mainly helps organization in decision making process. The data mining tools are used to find the frequent data patterns. This helps organization in decision making process. Today, most of the organizations are using the data mining techniques in their decision making process. The data mining techniques also help organizations in identifying their weaknesses and strength and also understanding their improvement areas. These processes help organization to make better decisions in order to focus the organization towards gaining strategic advantage.

It is concluded that, several organizations are being implementing the data mining projects successfully for long time. Mostly, organizations are obtaining several benefits and also they face many challenges in implementing and maintain the data mining process. However, organizations are using data mining for gaining strategic advantage.

REFERENCES

1. Dartigue, C., Jang, H.I. and Zeng, W. (2009): 'A New Data-Mining Based Approach for Network Intrusion Detection' , Seventh Annual Communication Networks and Services Research Conference, pp.372 – 377

2. Date, C J (2005): Database in Depth: Relational Theory for Practitioners. O'Reilly (2005), p. 152.

3. Yang T H and Wang S H (2008): DBMS support of the Data Mining, retrieved on 3rd December 2012 from www.mis.nsysu.edu.tw/db-book/DBProject2007Fall/03/report.ppt

4. Wang, J. (2003): Data Mining: Opportunities and Challenges, USA: Idea Group Publishing.

5. Yang Q and Wu X (2006): Challenging Problems in Data Mining Research, International Journal of Information Technology & Decision Making, Vol. 5, No. 4, 2006, 597-604.

6. Jiawei Han and Kamber M, (2006): Data Mining Concepts and Techniques, 2nd edition, Morgan Kauffman

7. Berry, Michael J. A. and Linoff G, (1997): Data Mining Techniques, New York: John Wiley & Sons, Inc.

8. Gaber M M, Zaslavsky A, and Krishnaswamy S,(2005): Mining data: streams: a review, SIGMOD Rec., 34(2):18–26, 2005

9. D. J. Cook and L. B. Holder (2007). Mining Graph Data. John Wiley & Sons, 2007

10. H. Miller and J. Han (2001): Geographic Data Mining and Knowledge Discovery. Taylor and Francis, 2001