Applications of Data Mining

Mohd. Ahtesham Farooqui

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

Abstract— Data mining is precisely described as the practice of analytical tools to determine information in a collection of data. Data mining is a method which discovers valuable patterns from group of data. Retail stores, banks, hospitals, and insurance companies are some of the organizations which have adapted technology of data mining to expand their businesses and achieving success. This paper aims to describe the data mining in various applications. Data mining technology is widespread in field of businesses because it lets them to acquire more knowledge about their clients and make smart marketing judgments.

Index Terms— Data Mining, Applications.

----------♦----------

1. INTRODUCTION

According to [1] defines that data mining as a process of mining of suitable data and patterns from vast data. It is also named as knowledge discovery process, knowledge gathering from data, knowledge withdrawal or information /configuration analysis.

[8], defines that data mining is a techniques for effective automated innovation of earlier unknown, fresh, effective, logical and beneficial patterns in massive databases. The patterns can be used in the process of decision making in an organization.

Data mining is a process in which the use of analytical tools to find out information in a pool of data. The analytical tools may comprise device learning, pattern acknowledgment, machine detection, statistics, artificial intelligence, computer user's interaction and data conception [2].

Data mining is a sound practice that is mainly used to explore large amount of data in order to retrieve useful data. This technique data mining concentrates to find patterns that were previously unidentified [3]. After finding the patterns they can be used for decision making of their businesses for further development.

There are three stages to deal with in finding the pattern they are exploration, pattern identification and deployment. In the first stage of exploration, data is cleaned and transmuted into another form and essential variables and then characteristics of data centered on the problem are determined. Then in the Pattern Identification as soon as the data is predicted and discovered, it is then developed and defined for the precise variables. Patterns are arrayed for final outcome in the final stage.

Data Mining is an evolution which aims to analyze data from diverse perspective and reviewing it into valuable data [4]. It is also known as Data or Knowledge Discovery. Data mining is also described as a technique of unearthing of the data create concealed information that otherwise would have escaped from the attention of the management. Gathered information is used to raise revenue and to cut expense. Data mining is technically a manner of finding correlations or patterns in the middle of mass of fields in huge relational databases.

Data mining supports organization to control relationships between internal factors such as cost, product setting, or staff services, and external factors such as financial pointers, competition, and client demographics [5]. And, it assists them to regulate the impact on transactions, customer satisfaction, and commercial profits. It also acts as a tool to view summary information of detail transactional data.

By implementing the process of data mining, a retailer can use point-of-sale archives of customer consumptions to refer directed promotions based on consumers purchase history. By mining demographic data from statement or warranty cards, the vendor could improve products and promotions to request to exact customer sections.

Data mining uses advance algorithm and it's faster and cheaper with huge storage capacity. The computer is in

charge for discovering the patterns by categorizing the fundamental rules and features in the data.

Data mining is the procedure of originating significant and actionable patterns, developments and summaries by altering through their website data using knowledge of pattern recognition such as instrument learning, neural systems and genetic algorithms [6]. Data mining is the usable knowledge from the warehoused server data.

Data mining systems assessed based on the multiple features such as data types, system issues, data sources, mining functions, coupling data mining with database, scalability, visualization tools, data mining query language and graphical user interface.

Data mining has various applications such as market division. By applying data mining the market division finds activities and patterns among customers that give the impression to buy the same products again and again. And to evaluate which customers are the most probable to discontinue purchasing the products or services of one and choose other seller.

Data mining applied in retail store aid to find which products are stolen the most. After determining applying measures to safeguard those products and identify those who are stealing them. Direct mail marketing is an older technique that has been followed for many years, concerns that combine it with data mining can experience great results [10]. Can determine which customer will reply positively to a direct mail marketing approach and can catch success of interactive marketing

Data mining to helps in increase profits of organization and can create new businesses and by automatically predicting customers behaviors and trends. Data mining predicts and analyze by looking at marketing strategies of past.

Data mining is a powerful tool for those who deal with finances. A financial institution like a bank can predict the number of nonpayment that will take place among their customers within a given span of time, and can also predict the amount of cheating that can occur.

Data mining potentially enables to find relationships that could allow making planned decisions by automatically recognizing patterns that were hidden in database.

In data mining the technique which is used for prediction is known as modeling. Modeling is a recent technique which is used for data loading and communication capabilities and to gather and stock huge amount of data and computation control to automate modeling techniques to work straightly on the existing data [11]. Neutral networks tool is used to detect the characteristics and knowledge of customers.

Data mining deals with real world bulk data. Process of data mining is interactive processes, which involve assembling of data from several sources, providing the data to data mining algorithm, analyzing the result and interpreting it and, applying the output to novel situations.

Data mining is grown into a pervasive technology in accomplishments as different as using old data to forecast the victory of a marketing campaign, observing for patterns in financial transactions to determine illegal activities or studying genome sequences. After this viewpoint, it was just a matter of while for the discipline to scope in that important area of computer security. Applications of data mining, in security of computer present a pool of study efforts on the usage of data mining in computer security.

Data mining is used to determine loyalty of customer, detection of fraud, analyzing trend, analyzing current market, detecting false [12]. Data mining is useful in predicting the events in the future offering clear information in a readable form and can be understood by all users. In current trend data mining is used in all organizations who has strong focus on end users. It assists organizations to determine the impact on sales, buyer satisfaction, and commercial profits. Companies implementing data mining can gain a lot and it is beneficial in all areas like sales, science, medical and engineering etc.

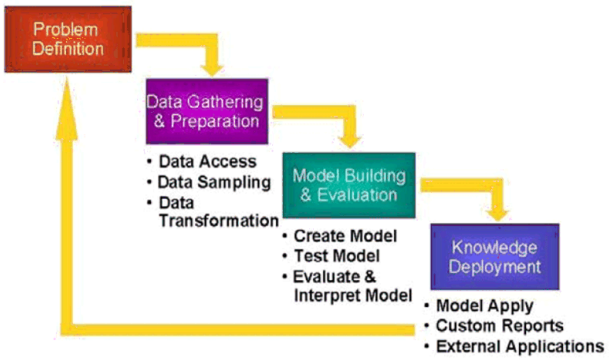

Figure 1: Process of Data Mining

Source:
http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm.

II. APPLICATIONS OF DATA MINING

Data mining techniques have been implemented successfully in several zones from business to science and sports. Data mining has been extensively used in database

marketing, retail data analysis, stock exchange, credit sanction, etc. Data mining procedures have been used in astronomy, medicine, molecular biology, geology, and many other fields. It has also been successfully used in health care management, tax cheating detection, money laundering observing, and also sports.

The applications of data mining are listed below.

a) Data Mining In Financial Market:

In the financial field the data mining plays a major role where planning, designing, building of data warehouses for multidimensional data investigation and data mining is taken place. Data mining helps in analyzing prediction of loan payment and customer credit policy. And assist in reaching the target by classifying and clustering of customers in the current market. And also supports in detecting money laundering and other financial crimes [9].

In the financial markets, neural networks is used in stock-price predicting, in option exchange, in bond rating, in portfolio administration, in product price forecast, in mergers and gaining, as well as in estimating economic disasters [15].

A software applications on the finance market that use data mining techniques for stock prediction is show in the below figure. The two lines in the graph denote the real and predicted stock values:

Figure 2: NETPROPHET by Neural Applications Corporation is a stock prediction using data mining technique.

Source:
http://www.pearsonhighered.com/samplechapter/0130862711.pdf

b) Data Mining In Retail Industry:

Through data mining multidimensional study of sales, consumers, products, time, and region is possible. It helps in analyzing efficiency of sales campaign and look at how to retain a customer. Data mining aids in recommending product and cross-referencing of things.

Retailers were put up in such a situation that they were very earlier to accept the data warehousing due to slim margin. Retailers then experienced better decision-support practices which lead to value-added efficiency in inventory management and financial prediction [16]. The timely implementation of data mining by retailers has offered them a better opening to take advantage of data mining.

The direct-mail business is an area where data mining, or data modeling, is extensively used. Nearly every category of retailer, as well as catalogers, consumer retail chains, grocers, issuers, business-to-business dealers, and packaged goods makers, uses direct marketing.

c) Data Mining In Telecommunication Industry:

Data mining implemented in the industry of telecommunication is to analyze telecommunication data in multiple aspects. It helps in analyzing fallacious pattern, identifying uncommon patterns and sequential pattern. It is used telecommunication data analysis and services with the help of visualization tools.

With the hyper-competitive character of this telecommunication industry, a necessity to appreciate customers, to keep them, and to prototype effective ways to advertise new products to these customers is motivating a demand for data mining in telecommunications where no request occurred in distant memory.

d) Data Mining In Banks:

In the banking industry, data mining is heavily used to model and predict credit fraud, to evaluate risk, to perform trend analysis, and to analyze profitability, as well as to help with direct marketing campaigns [14].

e) Data Mining In Health Care:

Data mining has been used broadly in the medical industry already [7]. Neuro-Medical Systems used neural networks to carry out a Pap smear diagnostic aid. Neural networks used to do protein analysis for drug improvement. The University of Rochester Cancer Center and the Oxford Transplant Center use Data SEEKER, a assessment tree knowledge, to help by means of their research. The Southern California Spinal Disorders Hospital uses Data Discovery to data excavation.

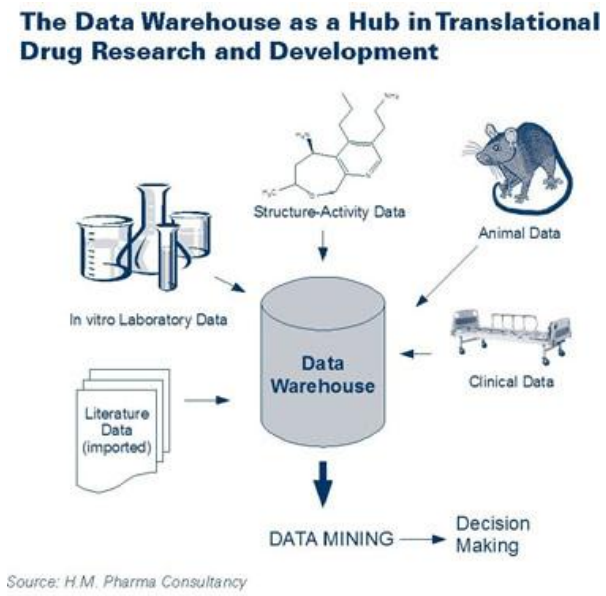

Figure 2: Data Mining in Health care

Source: [17]

f) Data Mining For Biological Data Analysis:

Through data mining semantic integration of heterogeneous, dispersed genomic and proteomic databases is possible [18]. It is used for aligning, indexing, searching similar data, comparing and analyzing multiple nucleotide/ protein sequences. Structural pattern and genetic network and protein trails can be identified and analyzed. Data mining facilitates in identifying co-occurring gene sequences and linking genes to various stages of virus development. Visualization tools are used in genetic data analysis.

III. DIFFERENT LEVELS OF ANALYSIS:

Few applications of data mining in advance planned benefit are: [19]

i. Artificial Neural Networks : Non-Linear Predictive Prototypes That Learn From Training And Its Similar To Biological Neural Networks In Structure.

ii. Genetic algorithms : Optimization methods that use procedures such as genetic combination, mutation, and natural collection in a design built on the ideas of natural evolution.

iii. Decision trees : Tree-shaped structures that denote sets of conclusions. These decisions make instructions for the grouping of a dataset. Explicit decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID) .These are decision tree methods used for sorting of a dataset. They run a set of rules that you can be applied to a new dataset to forecast which records will have a given result.

Nearest neighbor method : A technique that categorizes each record in a dataset centered on a grouping of the classes of the k record(s) most alike to it in the past dataset (where k 1). Occasionally called the k-nearest neighbor technique.

Rule induction : The retrieval of useful if-then rules from information based on statistical implication.

Data visualization : The visual clarification of difficult relationships in multidimensional data. Graphics tools are used to clarify data relationships.

IV. CONCLUSION

Data mining has developed which offers tools for data analysis and knowledge discovery. Extensive progress has been done in the increase of algorithms for mining knowledge from databases. This growth has not been coordinated by research on decision-making algorithms, which make use of the information obtained by the machine learning algorithms. This paper outlined a framework of applications in case of data mining. The tools of data mining used to discover data patterns that they make use of in understanding rules. Data mining helps industries to speed up process of analysis by concentrating on variables that are most important. The massive decrease in the performance and cost ratio of computer systems has made numerous organizations to start applying composite algorithms used in the techniques of data mining. In the current world industries are widely implementing data mining for their success.

REFERENCES

1) Ramageri M B (2010): DATA MINING TECHNIQUES AND APPLICATIONS , Indian Journal of Computer Science and Engineering Vol. 1 No. 4 301-305

2) Alluri R N (2005): Evaluation of data mining methods to support data warehouse administration and monitoring in SAP business warehouse, University of applied sciences - furtwangen, Germany, p 21.

3) Han J and Kamber M (2001): Data Mining: Concepets and Techniques. Morgan Kaufmann.

4) Hand D, Mannila H and Smyth P (2004): Princliples of data mining. MIT press, Cambridge.

5) Muller A J and Lernke F (2003): Selft organizing data mining: Extracting knowledge from data, Trafford publishing, British Columbia, Canada.

6) Pal K J (2011): Usefulness and applications of data mining in extracting information from different perspectives, Annals of Library and Information studies, Vol. 58, pp 7-16.

7) Ranjan J (2007): Application of data mining technique in pharmaceutical industry, Journal of theoretical applied information technology, Vol 13, No. 4.

8) Gupta G K (2006), Introduction to Data Mining with Case Studies, PHI Learning Private Limited, New Delhi, p 2.

9) Kusiak A (2006): Data mining: manufacturing and service applications, International Journal of production research, Vol. 44, No 18 -19.

10) Berson A and Smith, S.J (2005) : Data Warehousing, Data Mining & OLAP, Tata McGraw-Hill Edition, pp.5-6.

11) Ibrahim S R (1999): Data Mining of Machine Learning Performance Data, RMIT University.

12) Tuovinen L, Laurinenn P, Juutilainen I and Roning J: Data mining applications for diverse industrial application domain with smart archive, University of Oulu, Finland.

13) Data Mining (2004) Multi-Relational Data Mining <pdf>, Retrived on 30 November 2012, from http://www.kiminkii.com/thesis.pdf

14) Bhambri V (2011): Application of Data Mining in Banking Sector, IJCST, Vol 2, No. 2.

15) Dass R(n.d.,) DATA MINING IN BANKING AND FINANCE: A NOTE FOR BANKERS, Retrieved on 30 November 2012, from http://iimahd.ernet.in/publications/data/Note%20on%20Data%20Mining%20%26%20BI%20in%20Banking%20Sector.pdf

16) Bhasin L M (2006) Data Mining: A Competitive Tool in the Banking and Retail Industries <pdf>, Retrieved on 30 November 2012, from http://icai.org/resource_file/9935588-594.pdf

17) Aftarczuk K (2007): Evaluation of selected data mining algorithms implemented in Medical Decision Support Systems , Master Thesis Software Engineering ,Thesis no: MSE-2007-21.

18) Mucke (2009): Data Mining in Drug Development and Translational Medicine, Insight Pharma Reports.

19) Laudon K.C. & Laudon J.P. (2000). Management Information Systems: Organization and Technology in the Networked Enterprise. Sixth Edition; Prentice Hall.