

Applications of Data Mining

Simmi Bagga
Assistant Professor
Sant Hira Dass Kanya Maha Vidyalaya,
Kala Sanghian, Distt Kpt, India
(Email: simmibagga12@gmail.com)

Dr. G.N. Singh
Department of Physics and Computer
Science, Sudarshan Degree College,
Lalgaon, Rewa (MP), India

Abstract

Data Mining is a powerful tool capable of handling decision making and for forecasting future trends of market. Data Mining tools and techniques can be successfully applied in various fields in various forms. In this paper, we will explain various applications of Data Mining. We will also explain how Data Mining can integrate with other fields.

Keywords

DATA MINING, FINANCIAL DATA ANALYSIS, APPLICATIONS

1. Introduction

Data Mining is main concerned with the analysis of data and Data Mining tools and techniques are used for finding patterns from the data set. The main objective of Data Mining is to find patterns automatically with minimal user input and efforts. Data Mining is a powerful tool capable of handling decision making and for forecasting future trends of market. Data Mining tools and techniques can be successfully applied in various fields in various forms. Many Organizations now start using Data Mining as a tool, to deal with the competitive environment for data analysis. By using Mining tools and techniques, various fields of business get benefit by easily evaluate various trends and pattern of market and to produce quick and effective market trend analysis.

2. DATA MINING APPLICATIONS

Data Mining applications are widely used in Health industry, Auditing, Telecommunication industry, Retail industry etc.

2.1. Data Mining for Biomedical and DNA data analysis

In recent years, Data Mining has been widely used in area of Medical science such as Biomedical, DNA, Genetics and Medicine etc. In the area of Genetics, the important goal is to understand the mapping relationship between the variation in human DNA sequences and the disease susceptibility. Data Mining is very important tool to help improve the diagnosis, prevention and treatment of the diseases.

Due to growth in biomedical research, the large scale of genes patterns and functions has to be studied. Data Mining tools helps greatly to study DNA analysis and to find various patterns and functions. Data Mining helps the researches of Biomedical and DNA data analysis in the following ways:

- The DNA data is generally heterogeneous, highly distributed and uncontrolled in nature but this data is highly used in various areas of research in medical science. For the research purpose data must be systematic. Data Cleansing and integration approaches of Data Mining helps to systemize the data and store this data in the data base or data warehouse for further use in research.
- One of the main tasks related to DNA data analysis is to compare various sequences of the data and search for the similarities among this DNA data. Comparison mainly involves the gene sequence of healthy and diseased tissues to find the difference between these two types. This can be done by retrieving the gene sequence of both healthy and disease tissue classes and then finds the frequently occurring patterns of both classes. This analysis helps to find the similarities and dissimilarities in genetic sequence.
- In Biomedical research, it is studied that most of the diseases are triggered by combination of genes. Association analysis method is used to determine the co-occurrence of the group of genes and also we can study the interaction and relationships between the genes.
- There are different combinations of genes contributes to diseases but these different genes becomes activate at different level of stages. Path analysis is used to link the different genes to different stages of disease development. Path analysis performs an important role in genetic study.
- Visualization tools also play an important role in biomedical Data Mining. The visualization tools helps to present complex genes structure in graphs, trees and chains. The visual representation helps to better understanding of complex genes structures, for knowledge discovery and data exploration.

2.2. Data Mining as Financial Data Analysis

Financial data is mainly collected from banks and from other financial sectors. This financial data is usually reliable, complete and has high quality. Financial data need a systematic method for data analysis. Data Mining plays an important in analysis of financial data. Data Mining follows steps such as data collection and understanding, data refinement, model building and model evaluation and deployment. These steps help to deal with analysis of financial data. The proper analysis of financial data enables us to better decisions making capabilities according to the market analysis. Data Mining tools and techniques helps to analyze the financial data in the following ways:

- Data collected from the various financial institutes like banks are first collected in the data warehouse. Multidimensional data analysis techniques are used to analyze such data collected in data warehouse for its general properties.
- One of main task related to analyze about the prediction about loan payments and customer credit policies. For such analysis , Data Mining methods such as feature selection that helps to identifies the various features like customer

income level, payment to income ratio and credit history etc. By analyze of such features the bank can decide about the loan grant policies on the basis of relatively low risks.

- Clustering and Classification techniques of Data Mining help the financial institutes to cluster various customers that have the similar features. Effective clustering and filtering methods helps the bank to identify the customer groups, relate new customer with the present cluster and facilitate them some common benefits.
- Data Mining tools helps the financial institutes to detect the frauds or crimes by interrelate data from the various data bases and from the history transactions done by that customers. Data Visualization tools help to present data in the different format like graphs based on the certain attributes. By viewing data from different angles the bank can view the customers who have performed some illegal operations and then the detail investigation of these suspicious cases helps to find the frauds and crimes.

2.3. Data Mining for the Retail Industry

Data Mining plays an important role in the retail industry also. Retail industry involves large amount of data that includes transportation, sales and consumptions of goods and services. This data grows rapidly due to increase in purchase and sales in business. These days, E-commerce is growing fast with the growth of companies and also improving the online experience. Electronic commerce describes the buying and selling of products, services, and information via computer networks including the Internet. It can be viewed as a business concept that handles effectively and efficiently all previous business management and economic concepts. E-commerce streamlined the business processes, flatter organizational hierarchies and inter-firm collaboration. The databases have become treasure and essential e-commerce tools. To take advantage of the data base, Data Mining must be integrated into the e-commerce systems.

Retail Data mining helps in identifying customer behaviour, shopping patterns and distribution policies etc. As retail data in a very large in quantity, so we design data warehouse to store this large data and effective analysis of data. The main decision has to take while designing the data warehouse is dimension, level and preprocessing to perform the quality and efficient data mining.

The main requirement of the retail industry is timely information regarding the customer requirements, trends of market with the cost, profit and quality etc. To get these information effectively and efficiently, powerful multidimensional analysis and visualization tools are required. The retail industry also do some sort of advertisements, give discounts to customers to attract them. An effective analysis is required to get information to check the effect of advertisement on the sales. This analysis can be done by checking the amount of sale done during the sales period before the advertisement and the amount of sale after the advertisement.

The sequential pattern mining used to find the change in the customer consumption, adjustment of price and variety in various goods in order to attract customer. Association analysis also helps to get information in order to promote sales. In this way, various Data Mining tools help in the retail industry.

2.4. Data Mining for Telecommunication Industry

Telecommunication Industry has been growing very fast as the technology grows. These days Telecommunication services has grown from local and long distance voice communication services to fax, pager, cellular phones and e-mails. Now the telecommunication services have integrated with the computer, internet, and network and with other communication technologies. Due to the advancements in telecommunication technologies and to work these technologies effectively, Data Mining techniques integrated with these technologies to produce effective results. Data Mining helps to identify telecommunications patterns, fraud activities and also helps to better use of resources and improve the quality of services. Data Mining improves the telecommunication services in the following ways:

- Telecommunication data involves type of call, location of caller, location of called, time of call and duration of call etc. Multidimensional data analysis helps to identify and compare the system load, data traffic and profit etc. Analyst can view the charts and graphs of calling resources, destinations etc by using the visualization tools of Data Mining.
- The main problem faced by the telecommunication industry is due to the fraudulent activities. These fraudulent activities may involve fraudulent calls during busy hour, periodic calls etc. These activities may effect on the performance of the communication network. Data Mining methods such as cluster analysis, outlier analysis helps to detect fraudulent patterns and improves the efficiency of the communication services.
- The association and sequential pattern analysis helps to promote various telecommunication services.
- Visualization tools such as association visualization, clustering and association visualization shows very useful telecommunication data analysis.

The widespread changes in the adoption and utilization of new technologies in business, even small business has large number of financial transaction. It's the responsibility of the analyzer to analyze these transactions to detect frauds and errors in financial transactions. Due to change in business trends, it's very difficult and complicated to analyze financial transactions by manual methods. Due to limitations of in manual analysis of complex data, we use Data Mining tools and techniques in various areas to get effective results.

3. Conclusions

Data Mining can be used in various fields like retail industry, telecommunication industry etc. In Retail industry Data mining helps in identifying customer behaviour, shopping patterns and distribution policies etc. Data Mining also helps to identify telecommunications patterns, fraud activities and also helps to better use of resources and improves the quality of services. . Data mining tools helps greatly to study DNA analysis and to find various patterns and functions. In this we explained various applications of Data Mining.

References

- [1] Fayyad, U.M. "Data Mining and Knowledge Discovery: Making Sense Out of Data", *IEEE Expert* 11(5), 1996.
- [2] T. Imielinski and H. Mannila. A database perspective on knowledge discovery. *Communications of ACM*, 39, 1996.
- [3] J. Han and M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2000.
- [4] Elder, John F., IV and Daryl Pregibon, (1996), *Advances in Knowledge Discovery & Data Mining*, "A Statistical Perspective on KDD."
- [5] Chen, M. et al, 1996. *Data Mining: An Overview from a Database Perspective*. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8.
- [6] Han, J., & Kamber, M. (2006). *Data mining concepts and techniques*. Boston: Elsevier.
- [7] Dunham, M.H. (2003). *Data mining introductory and advanced topics*. Upper Saddle River, NJ: Pearson Education, Inc.
- [8] Yao, Y.Y. A step towards the foundations of data mining, in: *Data Mining and Knowledge Discovery: Theory, Tools, and Technology V*, Dasarathy, B.V. (ed.), The International Society for Optical Engineering.
- [9] Singh G.N, Bagga Simmi(2011) Three Phase Iterative Model of KDD, *International Journal of Information Technology and Knowledge Management*, Volume 4, No. 2, pp.695-697.
- [10] Singh G.N, Bagga Simmi (2011), Clustering Method for categorical and Numeric Data sets, *Global Journal in computer Science*.