

# Data Mining: Current Applications & Trends

Sidhant Sethi

*Vivekananda Institute of Professional Studies, Delhi, India*

Dheeraj Malhotra

*Vivekananda Institute of Professional Studies, Delhi, India*

Neha Verma

*Vivekananda Institute of Professional Studies, Delhi, India*

**Abstract-** Data mining is used in many areas. There are many data mining systems available and yet there are many challenges in this field. Data Mining helps us in future forecast of market trends and can itself do decision making. In order to analyze patterns and rules from database data mining is used. The techniques used includes classification, clustering, association, rule mining etc. Every individual technique has its own place and importance. Data Mining is used in fields like education.

**Keywords –** Data Mining, Applications of Data Mining, Data Mining Techniques

## I. INTRODUCTION

There is a huge amount of data available in today's world. This data is of no use until it is changed into some useful information. This information is required to be processed and scanned useful data has to be separated to be practically useful. Data Mining is defined as the procedure of extracting information from huge sets of data. In other words, we can say that data mining is mining knowledge from data. Data mining can also be defined as a process that analyses a large amount of data to find information that improves efficiency of business. Various industries are adopting data mining as a tool to their processes to have a competitive advantage and help the business to grow.

The main objective of Data Mining is to find patterns with minimum user input and efforts. Here is the list of steps which are involved in the Knowledge Discovery process. This process is shown in figure 1.

- Data Cleaning – Under this step, the noise and inconsistent data is removed.
- Data Integration – Under this step, multiple data sources are combined.
- Data Selection – Under this step, data relevant to the tasks are retrieved from the database.
- Data Transformation – Under this step, data is transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations.
- Data Mining – Under this step, intelligent methods are applied in order to extract data patterns.
- Pattern Evaluation – Under this step, data patterns are evaluated.
- Knowledge Presentation – Under this step, knowledge is represented.

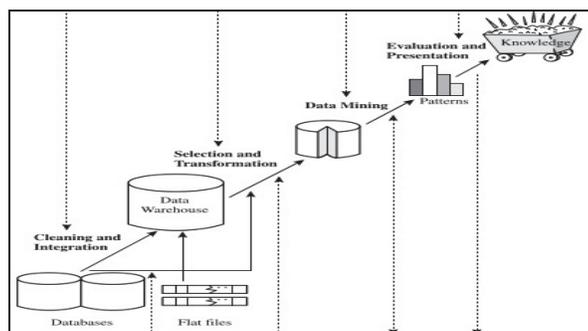


Figure 1: KDD Process

This paper is organized as follows: In section II we depict the techniques of data mining. In section III we discuss Literature related to the topic. In section IV we depict the objective. The applications of data mining, trends of data mining, conclusions and future work are presented in section V, VI, VII respectively.

## II. DATA MINING TECHNIQUES

Data mining also means collection of relevant information from unstructured data. So it enables to achieve specific objectives. The purpose of a data mining is normally either to create a descriptive model or a predictive model. A descriptive model presents the main characteristics of the data set. The purpose of a predictive model is to predict an unknown (often future) value of a specific variable; the target variable. The goal of predictive and descriptive model can be achieved using a variety of data mining techniques.

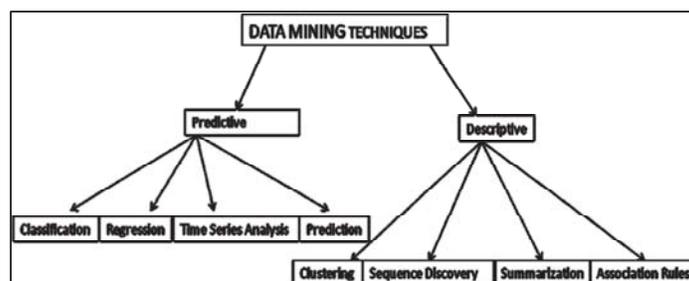


Figure 2: Techniques used in Data Mining

### A. Classification -

Classification is a form of data analysis that extracts models describing important data classes. Such models, called classifiers, predict categorical (discrete, unordered) class labels. This technique is based on supervised learning.

### B. Regression -

Regression is used to map data item to a real valued prediction variable i.e., regression can be used for prediction. In the regression techniques target value are familiar. For example, you can predict the child behaviour based on family history.

### C. Time Series Analysis -

Outcome of time series is the process where in we use statistical techniques to prototype and understand a time-dependent series of data points. Time series forecasting is a method of using a model to generate predictions for future events based on known past events. For example stock market.

### D. Prediction -

It is one of a data mining techniques that discover the relationship between independent variables and the relationship between dependent and independent variables. Prediction model based on continuous or ordered value.

### E. Clustering -

Clustering is a collection of similar data object. It is way finding similarities between data according to their characteristics. This technique based on the unsupervised learning. For example, image processing, pattern recognition, city planning.

### F. Clustering -

Clustering is a collection of similar data object. It is way finding similarities between data according to their characteristics. This technique based on the unsupervised learning. For example, image processing, pattern recognition, city planning.

### G. Summarization -

Summarization is abstraction of data. It is group of suitable task and gives a survey of data. For example, long distance race can be summarized total minutes, seconds and height.

#### H. Association –

Association is the most popular data mining techniques and finds most frequent item set. Association seeks to discover patterns in data which are based upon relationships between items in the same transaction. This method of data mining is utilized within the market based analysis in order to identify a set, or sets of products that consumers often purchase at the same time.

#### I. Association Rule –

Association is the most popular data mining techniques and finds most frequent item set. Association seeks to discover patterns in data which are based upon relationships between items in the same transaction. This method of data mining is utilized within the market based analysis in order to identify a set, or sets of products that consumers often purchase at the same time.

#### J. Sequence Discovery –

Uncovers relationships among data. It is set of object each associated with its own timeline of events. For example, scientific experiment, natural disaster and analysis of DNA sequence.

### III. LITREATURE REVIEW

Table 1: Literature Related to the Topic.

Author	Title	Summary
Smita,Priti Sharma	Use of Data Mining in Various Field	This paper provides an insight of data mining ,its techniques ,its applications in various fields & also conducts a formal review of the application of data mining such as the education sector, marketing, fraud detection, manufacturing and telecommunication.
Neelamadhab Padhy, Dr. Pragnyaban Mishra and Rasmita Panigrahi	The Survey of Data Mining Applications And Feature Scope	In this paper authors have focused a variety of techniques, approaches and different areas of the research which are helpful and marked as the important field of data mining technologies and briefly reviewed the various data mining applications.
Hardeep Kaur	A Review of Applications of Data Mining in the Field of Education	In this paper the author have focused only on the data mining application only in education field only.
Simmi Bagga Dr. G.N. Singh	Applications of Data Mining	This paper provides an insight of data mining ,it applications in various fields & also conducts a formal review of the application of data mining such as the education sector, marketing, fraud detection, manufacturing and telecommunication.
Nikita Jain, Vishal Srivastava	Data Mining Techniques	In this paper, the concept of data mining was summarized and its significance towards its methodologies was illustrated. The data mining based on Neural Network and Genetic Algorithm is researched in detail and the key technology and ways to achieve the data mining on Neural Network and Genetic Algorithm are also surveyed.
Gaurav Gupta, Geetika Hans, Tamanna Sehgal	Applications & Trends in Data Mining	In this paper authors have focused on different areas of data mining such as health care, finance, web mining, application exploration, and visual data mining.
Annan Naidu Paidi	Data Mining: Future Trends & Applications	In this paper authors have focused on different areas of data mining such as health care, finance, web mining, and higher education & also review on the future trends.
Dharminder Kumar, Deepak Bhardwaj	Rise of Data Mining: Current and Future Application Areas	In this paper authors have briefly reviewed the various data mining trends and applications from beginning to the future. Though very few areas are named here in this paper, yet they are those which are commonly forgotten. This paper provides a new perspective of a researcher regarding applications of data

		mining in social welfare.
Deepti Dalal	Current Data Mining Applications	In this paper authors have focused on different areas of data mining such as health care, finance, education, agriculture.

#### IV. OBJECTIVE

There is a large amount of data available in the Information Industry. It is necessary to analyze this huge amount of data and extract useful information from it. Data mining as a whole includes processes like data cleaning, data integration and transformation, presentation & evaluations of pattern. All these process in addition to extraction of information are used to make this information useful in many applications. The overall objective of this work is to review the various techniques, trends and applications of Data Mining.

#### V. APPLICATIONS OF DATA MINING

Data mining in various forms is used widely in many fields of today's world. Many institutions have now started using data mining in order to compete with the current environment of data analysis. In order to get quick and easy evaluations of trends and patterns of prevalent market and to produce a fast and useful market trend analysis various mining tools and techniques are used.

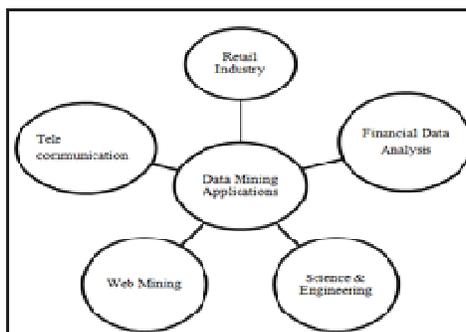


Figure 2: Major Data Mining Applications

##### A. 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 role in analysis of financial data. Data Mining follows steps such as data collection and understanding, data refinement, model building and model evaluation and deployment. The data available in financial institutes like bank is reliable and of premium quality that helps systematic data analysis and data mining. In the banking field, data mining is used to predict credit card fraud, to estimate risk, to predict the trend and profitability. In the financial markets, data mining technique such as neural networks used in stock forecasting, price prediction and so on.

Some of the specific cases are: Prediction of loan payments by creditors and analysis of specific customer credit policy, classification and clustering of customers for targeted marketing, Detection of money laundering and other financial crimes.

##### B. Data Mining for Retail Industry:

Data Mining has its great impact in retail industry in today's world because it gives large amount of data available on sales, customer purchasing history, goods transportation, consumption and services. It is natural that the quantity of data collected will continue to expand rapidly because of the increasing ease, availability and popularity of the web or e-commerce.

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 pre-processing to perform the quality and efficient data mining. Here is the list of examples of data mining in the retail industry is:

Design and Construction of data warehouses based on the benefits of data mining,

- Multidimensional analysis of sales, customers, products, time and region,
- Analysis of effectiveness of sales campaigns,
- Customer Retention

*C. Data Mining for Telecommunication Industry:*

Telecommunication Industry has been growing very rapidly as technology grows. These days, the telecommunication industry has evolved from local and global telephone services to provide many other comprehensive communication services such as cellular phones, email, internet access etc.

The telecommunication services have integrated with the computer, internet, and network and with other communication technologies. Data mining techniques can be integrated with these technologies to produce effective results. Data Mining helps to identify patterns, fraud activities and also helps to better use of resources and improve the quality of services. Here is the list of examples for which data mining improves telecommunication services are:

- Multidimensional Analysis of Telecommunication data.
- Fraudulent pattern analysis.
- Identification of unusual patterns.
- Multidimensional association and sequential patterns analysis.
- Mobile Telecommunication services.

*D. Data Mining for Science & Engineering:*

In the past, many scientific data analysis task handled small and homogeneous data sets which were analyzed using a “formulate hypothesis, build model and evaluate results” paradigm. In these cases, statistical techniques were typically employed for their analysis. Massive data collection and storage collections have recently changed the landscape of scientific data analysis. Today, scientific data can be analyzed at much quicker speeds and lesser costs.

This has resulted in accumulation of huge volumes of high dimensional data, stream data containing rich spatial information. Consequently, scientific applications are shifting from the “hypothesize and test” paradigm toward a “collect and store data, mine for new hypotheses, confirm with data or experimentation” process. Computer science generates unique kinds of data. For example, computer programs can be long and their execution often generates huge size traces. Data mining in computer science can be used to help monitor system status, improve system performance, isolate software bugs, detect software plagiarism, analyze computer system faults, uncover network intrusions, and recognize system malfunctions.

*E. Web Mining:*

Web Services and Web-based applications are growing at a rapid rate. This creates a huge amount of Web data having its own characteristics. This makes exploration in the area of Web Data Mining more demanding. Web Data Mining is an application of Data Mining dealing with extraction of knowledge from World Wide Web. Web mining can be categorized in 3 categories– Web usage mining, Web content mining and Web structure mining.

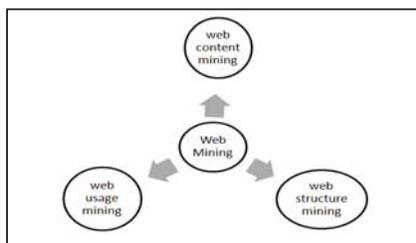


Figure 3: Major Types of Web Mining

Web Usage Mining is the application of data mining techniques to discover patterns from Web data in order to understand and better serve the needs of Web-based applications. Usage data captures the origin of users along with

their browsing behaviour at a Web site. It involves understanding of user behaviour when user interacts with the Web or the Web sites. Now, the user behaviour on the Web site can be reorganized according to the requirement.

Web structure mining is the process in which graph theory is used to analyze the node and connection structure of a web site on the basis of topology. Based on the variations of web structural data, web structure mining can be divided into two kinds: Extracting patterns from hyperlinks in the web and Mining the document structure. One of the techniques used for web structure mining is page rank in this technique the rank of a page is decided by the number of links pointing to the target node.

Web content mining is the extraction and integration of useful knowledge from content of Web page. The goal of Web content mining is to model and integrate the data on the Web so that more complex queries could be performed. Web content mining application includes classifying, clustering, comparing web site contents, modeling the document structure etc.

## VI. TRENDS IN DATA MINING

Data mining concepts are still evolving and here are the latest trends that we get to see in this field –

- Application Exploration.
- Scalable and interactive data mining methods.
- Integration of data mining with database systems, data warehouse systems and web database systems.
- Standardization of data mining query language.
- Visual data mining.
- New methods for mining complex types of data.
- Biological data mining.
- Data mining and software engineering.

## VII. CONCLUSION & FUTURE WORK

This research work is carried out as a partial fulfillment of Non University Examination System of MCA degree from GGS IP University, Delhi and provides an insight of data mining, its techniques, its applications in various fields & also conducts a formal review of the application of data mining such as the marketing, fraud detection, manufacturing and telecommunication. In future these techniques can be integrated with cloud computing frameworks such as Hadoop, Map Reduce and Spark etc for Big Data Analytics to assist in various applications like adaptive web page ranking systems, deriving useful association patterns for E Commerce and retail scenario etc.

## REFERENCES

- [1] N.Verma, "Improved Web Mining for E-commerce Website Restructuring" In IEEE International Conference Computational Intelligence & Communication Technology, PP 155-160, UP, India, Feb 2015.
- [2] N.Verma, D.Malhotra, M.Malhotra and J.Singh, "E-commerce website ranking using semantic web mining and neural computing" In Proceedings of International Conference on Advanced Computing Technologies and Applications, Mumbai, India, 2015. Procedia Computer Science, Science Direct. Elsevier, Vol. 45, 42-51. DOI = 10.1016/j.procs.2015.03.080
- [3] D.Malhotra, N.Verma, "An Ingenious Pattern Matching Approach to Ameliorate Web Page Rank" In International Journal of Computer Applications, New York, USA, 2013, Vol. 65, No 24, FCS, 33-39. DOI = 10.5120/11235-6543
- [4] D.Malhotra, "Intelligent Web Mining to Ameliorate Web Page Rank using Back Propagation Neural Network", In Proceedings of 5th International Conference, Confluence: The Next generation information Technology Summit, Noida, India, 2014, IEEE, 77-81. DOI = 10.1109/CONFLUENCE.2014.6949254
- [5] S.Sharma, P.Sharma, "Use of Data Mining in Various Fields", In IOSR Journal of Computer Engineering, IOSR-JCE Volume 16, Issue 3, Ver. V May-Jun. 2014.
- [6] N.Padhy, P.Mishra and R.Panigrahi, "The Survey of Data Mining Applications and Feature Scope", In International Journal of Computer Science, Engineering and Information Technology, IJCSET Vol.2, No.3, June 2012.
- [7] H.Kaur, "A Review of Applications of Data Mining in the Field of Education. In International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 4, April 2015.
- [8] S.Bagga, G.N.Singh, "Applications of Data Mining", In International Journal for Science and Emerging Technologies with Latest Trends 2012.
- [9] N.Jain, V. Srivastava, "DATA MINING TECHNIQUES", In International Journal of Research in Engineering and Technology, eISSN: 2319-1163 | pISSN: 2321-7308.

- [10] J.Han, and M. Kamber," Data mining concepts and techniques", Boston,2006.
- [11] A.N. Paidi, " Data Mining: Future Trends & Applications", In International Journal of Modern Engineering Research,Vol.2, Issue.6, 2012 pp-4657-4663 ISSN: 2249-6645.
- [12] G.Gupta,G.Hans, T. Sehgal," Applications & Trends in Data Mining".
- [13] D.Kumar,D. Bhardwaj," Rise of Data Mining: Current and Future Application Areas", In International Journal of Computer Science Issues, Vol. 8, Issue 5, No 1, September 2011 ISSN (Online): 1694-0814