

# DATA MINING AND BUSINESS INTELLIGENCE: CONCEPT AND COMPONENT

<sup>1</sup>Vivek Kumar, <sup>2</sup>Naveen Kumar, <sup>3</sup>Krishna Mohan Pandey

<sup>1</sup>IIMT Engineering College, Meerut, UP, (India)

<sup>2</sup>JPIET Meerut, UP, (India)

<sup>3</sup>IIMT Engineering College, Meerut, UP, (India)

## ABSTRACT

*Data mining is a technique for extracting valuable pattern on the data which come from multiple sources and make prediction for the good decision making. Data mining techniques are very useful to obtain valuable and profitable decision in business. Business intelligence is a process by which data is analyzed and valuable information is provided to corporate experts and business managers and end users make better decision for the business profit. This paper is for information about data mining and business intelligence.*

## I. DATA MINING

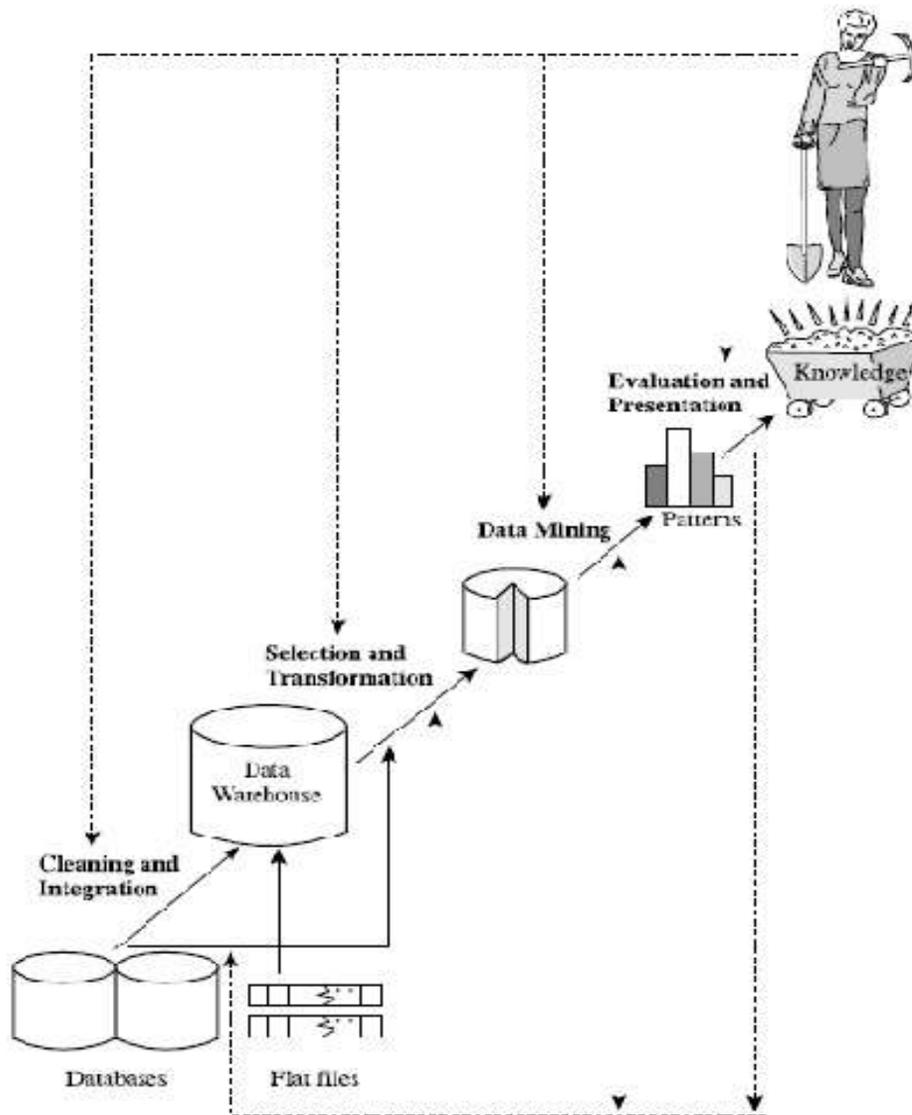
Data mining is the process of discovering valuable pattern and find out predictive models from data which is collected from multiple sources. A special term for Data mining is Knowledge Discovery in Databases", or KDD. Knowledge discovery has following steps

- Data cleaning is a preprocessing step and is used to remove noise data, identify and removing outliers, and correcting inconsistent data.
- Data integration also preprocessing step and the data may be collected from multiple sources and data integration means we combined that data sources because Resulting data are stored in a data warehouse.
- Data selection where data are retrieved from database and that data is ready to analysis.
- Data transformation means data is transformed and data is in appropriate form for mining and data transformation consist operation like Smoothing, Aggregation, Generalization of the data, Normalization, Attribute Construction.
- Data mining is an essential process where pattern are generated by the applying intelligent algorithm and that pattern are very helpful in decision making process in any business.
- Pattern evaluation means that all pattern which we have are not interesting some of the pattern are taking in consideration so question is that which kind of pattern are should be taken? To answer the this question, a pattern is interesting if
  - (1) It is easily understood by humans,
  - (2) Valid on new or test data with some degree of certainty,
  - (3) Potentially useful, and

(4) Novel.

A pattern is also interesting if it validates a hypothesis that the user sought to confirm. An interesting pattern represents knowledge.

- Knowledge presentation (where visualization and knowledge representation techniques are used to present the mined knowledge to the user). [1]



**Figure 1 Process of Knowledge Discovery.**

Business intelligence (BI) systems combine operational data with analytical tools to present complex and competitive information to planners and decision makers, in order to improve the timeliness and quality of the decision-making process [2].

Now a days data mining play vital role in business because business is going to more and more competitive and company need more accuracy in their decision making on the data which is coming from several sources .business intelligence is the process by which companies get more knowledge about their competitive company ,their

customer and as well as how they can make profit from the market knowledge. Data mining has a lot of technique by which the task of managers can go easy.

## II. CONTRIBUTION OF DATA MINING IN OTHER INDUSTRIES

- **Telecommunications** and **credit card** companies are two of the leaders in applying data mining to detect fraudulent use of their services.
- **Insurance companies** and **stock exchanges** are interested in applying data mining to reduce fraud.
- **Medical applications** use data mining to predict the effectiveness of surgical procedures, medical tests, or medications.
- **Financial firms** use data mining to determine market and industry characteristics as well as to predict individual company and stock performance.
- **Retailers** make use of data mining to decide which products to stock in particular stores (and even how to place them within a store), as well as to assess the effectiveness of promotions and coupons.
- **Pharmaceutical firms** mine large databases for chemical compounds and genetic material to discover substances that might be candidates for development as agents for the treatments of disease.[3]

## III. BUSINESS INTELLIGENCE COMPONENT

### On-Line Analytical Processing (OLAP)

It refers to the way in which business users can slice and dice their way through data using sophisticated tools that allow for the navigation of dimensions such as time or hierarchies. Online Analytical Processing or OLAP provides multidimensional, summarized views of business

data and is used for reporting, analysis, modeling and planning for optimizing the business. OLAP techniques and tools can be used to work with data warehouses or data marts designed for sophisticated enterprise intelligence systems. These systems process queries required to discover trends and analyze critical factors.

Reporting software generates aggregated views of data to keep the management informed about the state of their business. Other BI tools are used to store and analyze data, such as data mining and data warehouses; decision support systems and forecasting; document warehouses and document management; knowledge management ; mapping, information visualization, and dash boarding; management information systems, geographic information systems; Trend Analysis; Software as a Service (SaaS). [4]

### Advanced Analytics or Corporate Performance Management (CPM)

This set of tools allows business leaders to look at the statistics of certain products or services. For instance, a fast food chain may analyze the sale of certain items and make local, regional and national modifications on menu board offerings as a result. The data could also be used to predict in which markets a new product may have the best success.[5]

**Real Time BI:** whenever business operation occurs the real time business intelligence will deliver the information about the operation. Real time means near to zero latency and access to information whenever it is require.

**Data Warehouse:** According to William H. Inmon, a leading architect in the construction of data warehouse systems, “A **data warehouse** is a subject-oriented, integrated, time-variant, and nonvolatile collection of data in support of management’s decision making process”[6]

**Data Mart:** In data warehousing, there is a distinction between a data warehouse and a data mart. A data warehouse collects information about subjects that span the entire organization, such as customers, items, sales, assets, and personnel, and thus its scope is enterprise-wide. For data warehouses, the fact constellation schema is commonly used, since it can model multiple, interrelated subjects. A data mart, on the other hand, is a department subset of the data warehouse that focuses on selected subjects, and thus its scope is department wide. For data marts, the star or snowflake schemas are commonly used, since both are geared toward modeling single subjects, although the star schema is more popular and efficient.

A data mart contains a subset of corporate-wide data that is of value to a specific group of users. The scope is confined to specific selected subjects. For example, a marketing data mart may confine its subjects to customer, item, and sales. The data contained in data marts tend to be summarized.[1]

**Data Sources:** Data can be collected from many sources data may be historical, operational, external or existing data ware house

#### IV. CONCLUSION

Today’s consumer needs more services because they are more demanding. Data mining is applied in business to make business profitable. Companies have to improve herself to stay in the competitive zone. experts of company has to improve performance of services because in coming time it may be the consumers will know business information and can access and view all the market value of their segment. So the companies have to improve their performance and also improve their business intelligence capabilities to exist in competitive market.

#### REFERENCES

1. HanJiawei, KamberMicheline, “*Data Mining: Concepts and Techniques*”, 2nd edition, Morgan Kaufman Publishers, March 2006. ISBN 1-55860-901-6.
2. Negash, S.(2004). Business Intelligence. Communications of the Association for Information Systems, 13, 177-195.
3. Joyce Jackson. DATA MINING: A CONCEPTUAL OVERVIEW. Communications of the Association for Information Systems (Volume 8, 2002) 267-296
4. Jayanthi Ranjan. Business Intelligence: Concepts, Components, Techniques And Benefits . Journal of Theoretical and Applied Information Technology. Vol9. No 1. (pp 060-070)
5. <http://www.villanovau.com/resources/bi/overview-of-business-intelligence-bi-components/#.VV1vLFKunK8>
6. W. H. Inmon. *Building the Data Warehouse*. JohnWiley& Sons, 1996.
7. Inmon, W.H. (1999) ‘Building the Operational Data Store’, Wiley Publishers-New York, 2nd edition.