

DATA MINING

What motivated data mining? Why is it important?

The major reason that data mining has attracted a great deal of attention in information industry in recent years is due to the wide availability of huge amounts of data and the imminent need for turning such data into useful information and knowledge. The information and knowledge gained can be used for applications ranging from business management, production control, and market analysis, to engineering design and science exploration.

What is data mining?

Data mining refers to extracting or mining" knowledge from large amounts of data. There are many other terms related to data mining, such as knowledge mining, knowledge extraction, data/pattern analysis, data archaeology, and data dredging. Many people treat data mining as a synonym for another popularly used term, Knowledge Discovery in Databases".

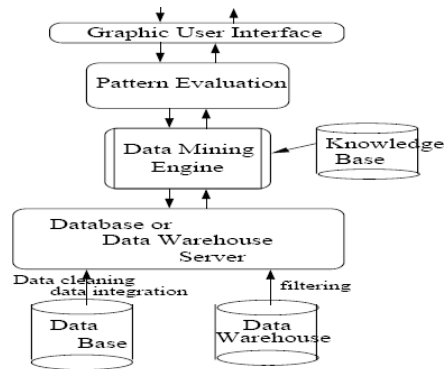
-Data mining is a process that is useful for the discovery of informative and analyzing the understanding of the aspects of different elements.

Architecture of a typical data mining system/Major Components

- A database, data warehouse, or other information repository, which consists of the set of databases, data warehouses, spreadsheets, or other kinds of information repositories containing the student and course information.
- A database or data warehouse server which fetches the relevant data based on users' data mining requests.
- A knowledge base that contains the domain knowledge used to guide the search or to evaluate the interestingness of resulting patterns. For example, the knowledge base may contain meta-data which describes data from multiple heterogeneous sources.
- A data mining engine, which consists of a set of functional modules

for tasks such as classification, association, classification, cluster analysis, and evolution and deviation analysis.

- A pattern evaluation module that works in tandem with the data mining modules by employing interestingness measures to help focus the search towards interestingness patterns.
- A graphical user interface that allows the user an interactive approach to the data mining system.



Architecture of a typical data mining system.

Classification of data mining systems

- **Classification according to the type of data source mined:** this classification categorizes data mining systems according to the type of data handled such as spatial data, multimedia data, time-series data, text data, World Wide Web, etc.
- **Classification according to the data model drawn on:** this classification categorizes data mining systems based on the data model involved such as relational database, object-oriented database, data warehouse, transactional, etc.
- **Classification according to mining techniques used:** This classification categorizes data mining systems according to the data analysis approach used such as machine learning, neural networks, genetic algorithms, statistics, visualization, database oriented or data warehouse-oriented, etc. Also with the degree of user interaction involved in the data mining process such as query-driven systems, interactive exploratory systems, or autonomous systems.

Few other processes which include in data mining are,

- Data Integration
- Data Cleaning
- Data Transformation
- Pattern Evaluation
- Data Presentation

The knowledge or information which is acquired through the data mining process can be made used in any of the following applications

-
1. Market Analysis
 2. Production Control
 3. Customer Retention
 4. Science Exploration
 5. Fraud Detection
 6. Sports
 7. Astrology
 8. Internet Web Surf-Aid

9.

Data Mining Applications: Some of the popular domains are,

1. Market Analysis and Management
2. Corporate Analysis & Risk Management
3. Fraud Detection

1. Market Analysis and Management:

- Customer Profiling
- Finding customer requirements
- Cross-market analysis
- Target marketing
- Determining customer purchasing pattern
- Provides summary information

2. Corporate Analysis and Risk Management:

- Finance Planning
- Asset Evaluation
- Resource Planning , Competition

3. Fraud Detection:

Frauds and malware is one of the most dangerous threats on the internet. It is almost a kind of crime that is increasing day after day. The fraud detection process can be mainly used through credit card services and telecommunication.

With the help of the services most of the important information like duration of the call, location, the time and day etc can be acquired which helps in big time.

Advantages of Data Mining

1.It helps to identify the shopping patterns:

Mining methods discover all the information about these shopping patterns. Moreover, this data mining process creates a space that determines all the unexpected shopping patterns. Therefore, this data mining can be beneficial while identifying shopping patterns.

2.Increases website optimization:

As most of the key factors of website optimization deal with information and analyzation, similarly, this mining provides such information which can utilize data mining techniques to increase website optimization.

3.It is beneficial for marketing campaigns:

It is also beneficial for marketing campaigns because it helps to identify the customer response over certain products available in the market and customer response through the marketing campaign increases profit for the growth of business.

4.Determining customer groups:

it also helps while determining customer groups. These new customer groups can be initiated through some sort of surveys and these surveys are one of the forms of mining where different types of information about unknown products and services are

gathered with the help of data mining.

5.It helps to measure profitability factors:

The data mining system provides all sorts of information about customer response and determining customer groups.

Therefore, it can be helpful while measuring all the factors of the profitability of the business.

6.Increases brand loyalty and customer satisfaction:

7.It increases customer loyalty:

This data mining process handles all customer satisfaction and customer loyalty regarding issues through in time service and response. Therefore, it creates an increased level of customer loyalty.

8.It identifies hidden profitability:

One of the most important elements of these data mining is considered as that it provides the determination of locked profitability therefore, this data mining provides clear identification of hidden profitability so that one can overcome the risk factor in their business.

9.Minimizes clients involvement:

Through **information technology allows** gathering information about key elements like marketing survey and research, assessing patterns of buying behavior is faster and economical compared to depends on their clients and customer for some additional information.

10.It is helpful to predict future trends:

It can be helpful while predicting future trends. And that is quite possible with the help of technology and behavioral changes adopted by the people.

11.It signifies customer habits:

Marketing industry need to understand all the matters of customer behaviour and their habits is possible with the help of data mining systems which helps in demand forecast reduces unsold stock and increases sales.

12. Quick fraud detection:

With the help of marketing analysis, one can also find out those fraudulent acts and products available in the market. Moreover, with the help of it one can understand the importance of accurate informations.