**Data Mining**

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

Data Mining is defined as extracting information from huge sets of data. In other words, we can say that data mining is the procedure of mining knowledge from data. The information or knowledge extracted so can be used for any of the following applications −

* Market Analysis
* Fraud Detection
* Customer Retention
* Production Control
* Science Exploration

**Data Mining Applications**

Data mining is highly useful in the following domains −

* Market Analysis and Management
* Corporate Analysis & Risk Management
* Fraud Detection

Apart from these, data mining can also be used in the areas of production control, customer retention, science exploration, sports, astrology, and Internet Web Surf-Aid.

**Market Analysis and Management**

Listed below are the various fields of market where data mining is used −

* **Customer Profiling** − Data mining helps determine what kind of people buy what kind of products.
* **Identifying Customer Requirements** − Data mining helps in identifying the best products for different customers. It uses prediction to find the factors that may attract new customers.
* **Cross Market Analysis** − Data mining performs association/correlations between product sales.
* **Target Marketing** − Data mining helps to find clusters of model customers who share the same characteristics such as interests, spending habits, income, etc.
* **Determining Customer purchasing pattern** − Data mining helps in determining customer purchasing pattern.
* **Providing Summary Information** − Data mining provides us various multidimensional summary reports.

**Corporate Analysis and Risk Management**

Data mining is used in the following fields of the Corporate Sector −

* **Finance Planning and Asset Evaluation** − It involves cash flow analysis and prediction, contingent claim analysis to evaluate assets.
* **Resource Planning** − It involves summarizing and comparing the resources and spending.
* **Competition** − It involves monitoring competitors and market directions.

**Fraud Detection**

Data mining is also used in the fields of credit card services and telecommunication to detect frauds. In fraud telephone calls, it helps to find the destination of the call, duration of the call, time of the day or week, etc. It also analyzes the patterns that deviate from expected norms.

**Knowledge discovery in databases (KDD)**

Knowledge discovery in databases (KDD) is the process of discovering useful knowledge from a collection of data. This widely used data mining technique is a process that includes data preparation and selection, data cleansing, incorporating prior knowledge on data sets and interpreting accurate solutions from the observed results.

Here is the list of steps involved in the knowledge discovery process −

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

The following diagram shows the process of knowledge discovery

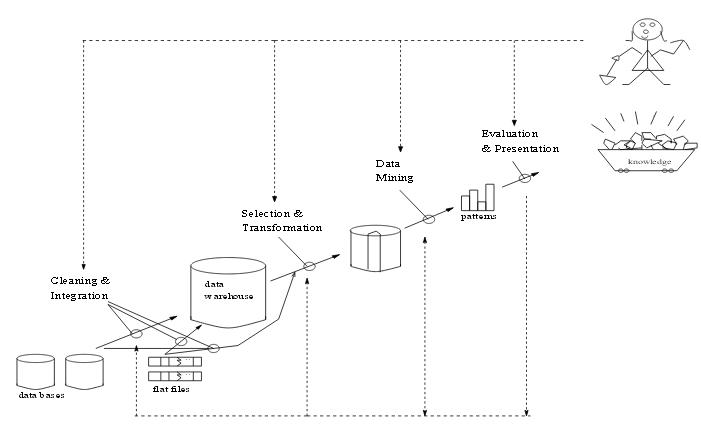


Fig 1: Data Mining as a process of knowledge discovery

**Architecture of data mining system**

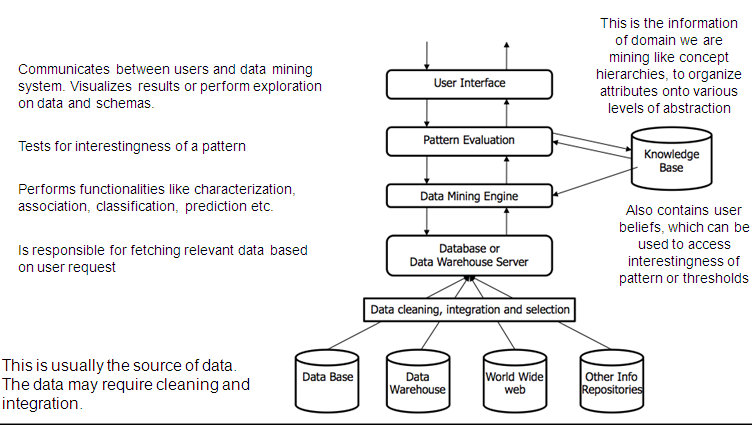
****

Fig 2: Architecture of data mining system