

Data Mining

"The process of extracting information to identify patterns, trends, and useful data that would allow the business to take the data-driven decision from huge sets of data is called Data Mining."

In other words, we can say that Data Mining is the process of investigating hidden patterns of information to various perspectives for categorization into useful data, which is collected and assembled in particular areas such as data warehouses, efficient analysis, data mining algorithm, helping decision making and other data requirement to eventually cost-cutting and generating revenue.

Types of Data Mining

Data mining can be performed on the following types of data:

1. Relational Database:

A relational database is a collection of multiple data sets formally organized by tables, records, and columns from which data can be accessed in various ways without having to recognize the database tables. Tables convey and share information, which facilitates data searchability, reporting, and organization.

2. Data warehouses:

A Data Warehouse is the technology that collects the data from various sources within the organization to provide meaningful business insights. The huge amount of data comes from multiple places such as Marketing and Finance. The extracted data is utilized for analytical purposes and helps in decision-making for a business organization. The data warehouse is designed for the analysis of data rather than transaction processing.

3. Data Repositories:

The Data Repository generally refers to a destination for data storage. However, many IT professionals utilize the term more clearly to refer to a specific kind of setup within an IT structure. For example, a group of databases, where an organization has kept various kinds of information.

4. Object-Relational Database:

A combination of an object-oriented database model and relational database model is called an object-relational model. It supports Classes, Objects, Inheritance, etc.

One of the primary objectives of the Object-relational data model is to close the gap between the Relational database and the object-oriented model practices frequently utilized in many programming languages, for example, C++, Java, C#, and so on.

5. Transactional Database:

A transactional database refers to a database management system (DBMS) that has the potential to undo a database transaction if it is not performed appropriately. Even though this was a unique capability a very long while back, today, most of the relational database systems support transactional database activities.

Advantages of Data Mining

- i. Data mining enables organizations to acquire knowledge-based data and make profitable modifications.
- ii. It is a cost-efficient statistical data application that aids in decision-making.
- iii. Data mining facilitates automated discovery of hidden patterns and prediction of trends.
- iv. It can be implemented in both new and existing platforms.
- v. The quick process allows new users to analyze large volumes of data efficiently.

Disadvantages of Data Mining

- i. Organizations may sell customer data for profit, as seen with American Express selling credit card purchases.
- ii. Some data mining analytics software is complex and requires advanced training to use effectively.
- iii. Choosing the right data mining tools can be a challenging task due to the variations in their algorithms and operation methods.
- iv. Data mining techniques are not always precise and can have serious consequences in certain situations.

Data Mining Applications

Data Mining is primarily used by organizations with intense consumer demands- Retail, Communication, Financial, marketing company, determine price, consumer preferences, product positioning, and impact on sales, customer satisfaction, and corporate profits. Data mining enables a retailer to use point-of-sale records of customer purchases to develop products and promotions that help the organization to attract the customer.

These are the following areas where data mining is widely used:

1. Healthcare:

Data mining in healthcare improves the system by using data and analytics to identify best practices, reduce costs, and enhance healthcare services. It employs machine learning, data visualization, and statistics to forecast patient categories and detect fraud.

2. Market Basket Analysis:

Market basket analysis uses a modeling method to understand purchase behavior. It identifies associations between groups of products, helping retailers understand buyer requirements and adjust store layouts accordingly. It enables comparisons between stores and customers in different demographic groups for analytical insights.

3. Education:

Education data mining explores knowledge from educational data to study student learning behavior, assess educational support impact, and promote learning science. It enables precise decision-making and prediction of student outcomes, allowing institutions to focus on effective teaching methods.

4. Manufacturing Engineering:

Data mining in manufacturing engineering helps find patterns in complex processes and optimize system-level design by analyzing relationships between product architecture, portfolio, and customer data needs. It aids in forecasting product development, cost, and expectations, among other tasks, leveraging knowledge as a valuable asset for manufacturing companies.

5. CRM:

Data mining in CRM helps businesses obtain and retain customers, enhance loyalty, and implement customer-oriented strategies. By collecting and analyzing customer data, organizations can establish strong relationships and leverage data mining technologies for effective analytics.

6. Fraud Detection:

Data mining in fraud detection saves money lost to fraud by identifying meaningful patterns in data. It provides an efficient alternative to traditional methods, safeguarding user data while classifying records as fraudulent or non-fraudulent. Using this information, a model is constructed to accurately identify fraudulent documents.

7. Lie Detection:

Data mining aids in lie detection, a challenging task in law enforcement. It involves analyzing unstructured text data, such as suspected terrorist communications, to identify meaningful patterns. By comparing information from previous investigations, data mining techniques help construct a model for detecting lies.

8. Financial Banking:

Data mining in financial banking utilizes vast amounts of transactional data from digitalization. It helps identify hidden trends, correlations, and market costs that may not be readily apparent to managers due to data volume and speed. By analyzing this data, managers can improve customer targeting, acquisition, retention, segmentation, and ensure profitability.