

DATA MINING - SYSTEMS

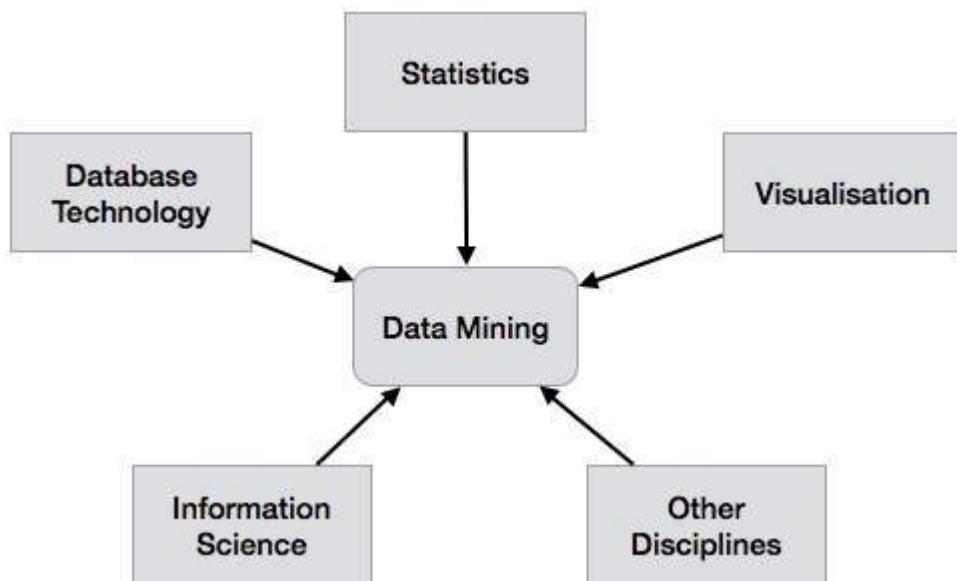
There is a large variety of data mining systems available. Data mining systems may integrate techniques from the following –

- Spatial Data Analysis
- Information Retrieval
- Pattern Recognition
- Image Analysis
- Signal Processing
- Computer Graphics
- Web Technology
- Business
- Bioinformatics

Data Mining System Classification

A data mining system can be classified according to the following criteria –

- Database Technology
- Statistics
- Machine Learning
- Information Science
- Visualization
- Other Disciplines



Apart from these, a data mining system can also be classified based on the kind of *a* databases mined, *b* knowledge mined, *c* techniques utilized, and *d* applications adapted.

Classification Based on the Databases Mined

We can classify a data mining system according to the kind of databases mined. Database system can be classified according to different criteria such as data models, types of data, etc. And the data mining system can be classified accordingly.

For example, if we classify a database according to the data model, then we may have a relational, transactional, object-relational, or data warehouse mining system.

Classification Based on the kind of Knowledge Mined

We can classify a data mining system according to the kind of knowledge mined. It means the data mining system is classified on the basis of functionalities such as –

- Characterization
- Discrimination
- Association and Correlation Analysis
- Classification
- Prediction
- Prediction
- Outlier Analysis
- Evolution Analysis

Classification Based on the Techniques Utilized

We can classify a data mining system according to the kind of techniques used. We can describe these techniques according to the degree of user interaction involved or the methods of analysis employed.

Classification Based on the Applications Adapted

We can classify a data mining system according to the applications adapted. These applications are as follows –

- Finance
- Telecommunications
- DNA
- Stock Markets
- E-mail

Integrating a Data Mining System with a DB/DW System

If a data mining system is not integrated with a database or a data warehouse system, then there will be no system to communicate with. This scheme is known as the non-coupling scheme. In this scheme, the main focus is on data mining design and on developing efficient and effective algorithms for mining the available data sets.

The list of Integration Schemes is as follows –

- **No Coupling** – In this scheme, the data mining system does not utilize any of the database or data warehouse functions. It fetches the data from a particular source and processes that data using some data mining algorithms. The data mining result is stored in another file.
- **Loose Coupling** – In this scheme, the data mining system may use some of the functions of database and data warehouse system. It fetches the data from the data repository managed by these systems and performs data mining on that data. It then stores the mining result either in a file or in a designated place in a database or in a data warehouse.
- **Semi-tight Coupling** - In this scheme, the data mining system is linked with a database or a data warehouse system and in addition to that, efficient implementations of a few data mining primitives can be provided in the database.
- **Tight coupling** – In this coupling scheme, the data mining system is smoothly integrated into the database or data warehouse system. The data mining subsystem is treated as one functional component of an information system.

>Loading [MathJax]/jax/output/HTML-CSS/jax.js