

## COMPONENTS, TYPES AND TECHNIQUES IN DATA MINING ARCHITECTURE

C. Aarthi\*<sup>1</sup>

\*<sup>1</sup>Affiliated, Madurai Kamaraj University, Computer Science, Fatima College (Autonomous),  
Madurai, Tamil Nadu, India.

DOI : <https://www.doi.org/10.56726/IRJMETS33261>

### ABSTRACT

Data mining is a field of intersection of computer science and statistics used to discover patterns in the information bank. The main aim of the data mining process is to extract the useful information from the dossier of data and mold it into an understandable structure for future use.

### I. INTRODUCTION

Data mining is a methodological approach for extracting information from enormous data sets. Data mining's primary goal is to find patterns, trends, or laws that contextually explain the behaviour of data. In contrast to previous data exploration techniques, the data mining method uses mathematical analysis to infer patterns and trends. For handling enormous amounts of data, data mining is a practical and very practical methodology.

### II. WHY DATA MINING IS IMPORTANT

Finding anomalies, trends, and correlations within huge data sets in order to forecast outcomes is known as data mining. You may use this information to lower risks, improve customer connections, raise profits, and more by employing a variety of strategies.

#### 1. Data Mining Architecture

A data source, data mining engine, data warehouse server, the pattern evaluation module, graphical user interface, and knowledge base are the key elements of data mining systems.

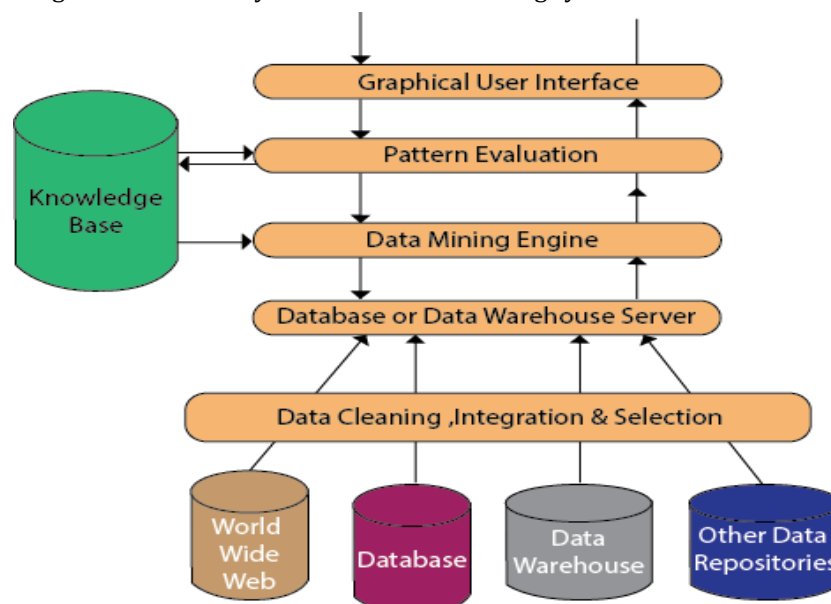


Fig 1: Data Mining Architecture

Basic working of Data Mining Architecture:

1. The user submits specific data mining requests, which are subsequently forwarded to data mining engines for pattern analysis.
2. These programmers use the existing database to try to locate the answer to the question.
3. The collected metadata is then given to the data mining engine for proper processing, which occasionally collaborates with modules for pattern assessment to get the desired outcome.

4. A appropriate interface is then used to send this result to the front end in a comprehensible format.

A detailed description of parts of data mining architecture

#### **Data Source**

You need data for a successful data mining process, and the data source might come from anywhere. It could be a database, data warehouse, the world wide web, etc.

#### **Database or Data warehouse Server**

The database or data warehouse server is a key component of the data mining architecture since it houses the cleaned and integrated data under a single schema, ready for processing.

#### **Data Mining Engine**

The Data Mining Engine consists of modules or tools for carrying out various operations on the data contained in the database or data warehouse server, such as data clustering, data categorization, prediction, and correlation analysis.

#### **Pattern Evaluation**

Using a threshold value, the pattern evaluation module investigates a pattern. It collaborates with a data mining engine and employs stake measures to discover interesting and useful patterns. Pattern evaluation may also collaborate with the mining module, which is the foundation of data mining techniques. It is suggested that the evaluation of pattern stakes be pushed as far into the mining procedure as possible in order to find desirable patterns and ensure an effective data mining process.

#### **GUI**

The data mining system and the user are connected by the GUI. The use of GUI makes data presentation simple and comprehensible while obscuring the intricate process of data analysis.

#### **Knowledge Base**

A knowledge base is characterised as a collection of general or domain-specific information obtained from data sources. It uses artificial intelligence and machine learning algorithms to store and retrieve vast volumes of structured data according to a set schema or "data model," which makes these tasks easier. The knowledge base provides data inputs and aids in pattern evaluation for the data mining engine.

### **2. Types of Data Mining Architecture**

There are four different types of architecture:

#### **No Coupling**

Data from specific data sources are retrieved using the no coupling data mining architecture. The database would otherwise be a very effective and precise means to get the data, but it does not do so. The no coupling data mining architecture is subpar and is only utilized for carrying out extremely basic data mining operations.

#### **Loose Coupling**

Data mining systems gather information from databases and store it in those systems using loosely coupled design. Memory-based data mining architecture is what this mining is for.

#### **Semi-Tight Coupling**

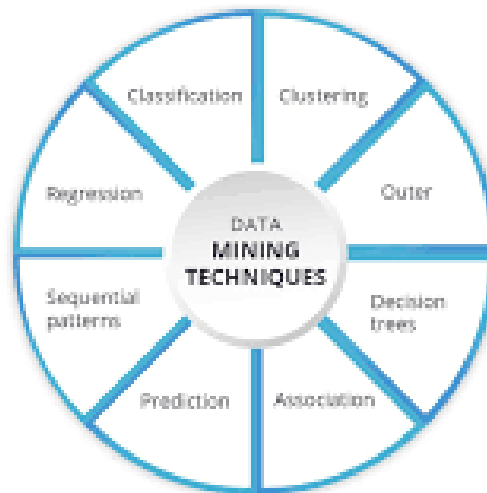
It frequently makes use of a number of beneficial aspects of data warehouse systems. Sorting, indexing, and aggregation are all included. For improved efficiency, this design allows for the database to hold an interim result.

#### **Tight Coupling**

A data warehouse is one of the key elements of this architecture, and its properties are used to carry out data mining operations. Scalability, speed, and integrated information are all provided by this design.

### **3. Data Mining Techniques**

There are several techniques in Data Mining architecture



**Fig 2: Data Mining Techniques**

- **Classification**

Based on machine learning, this. This is how we divide each item in a given collection into predetermined categories. Nevertheless, this approach uses mathematical methods. such as decision trees, linear programming, and neural networks, among others.

- **Clustering**

Clustering is a method that uses the shape of the item to automatically define several classes. The therefore created classes will subsequently be utilised to contain additional similar types of objects.

- **Outer detection (Outlier analysis)**

In this method, specific abnormalities (outliers) in the data set are found. You must be able to justify the existence of these anomalies within the general pattern. For instance, you could see a sudden spike in female purchasing activity among your male audience of purchasers.

- **Decision Tree**

We utilise it most frequently as a data mining approach. because of its straightforward structure. The decision tree's root serves as a condition. Each response points to particular information that enables us to build our ultimate conclusion on it.

- **Association**

Finding a link between two or more events or features is the goal of this data mining approach. It delves into a database system's fundamental model. Similar to when you buy a laptop, you are prompted to purchase a bag to go with it.

- **Prediction**

This technique is used to describe the connection between independent and dependent instances.

- **Sequence pattern**

This kind of data analysis looks for recurring patterns, regularities, or transaction trends across different informational threads during a predetermined time frame. Businesses can tell when a group of products is purchased together during a certain season of the year when it comes to sales. Due to this, businesses reward customers who have a real history of purchases with greater offers.

- **Regression**

Regression analysis is a technique for data mining that locates and investigates the connections between variables. In light of the existence of other variables, it may be used to calculate the probability of a certain variable.

**Below are some most useful data mining applications.**

- Healthcare. Data mining methods has the potential to transform the healthcare system completely. ...

- Education. ...
- Market basket analysis. ...
- Customer relationship management (CRM) ...
- Manufacturing engineering. ...
- Finance and banking. ...
- Fraud detection.

### **III. CONCLUSION**

If the types of data mining architecture systems employ a database for data retrieval, then these topics of the component, type, and techniques are the fundamental building blocks of data mining. Data mining system extracts information from a database in a loosely coupled data mining architecture. Additionally, it saves the outcome in such systems. The architecture of data mining is for memory-based data mining systems. Data mining professionals have dedicated their careers to better understanding how to process and draw conclusions from the enormous amount of data, and data mining techniques are dependent on a variety of techniques and technologies from the intersection of machine learning, database management, and statistics.

### **IV. REFERENCES**

- [1] <https://www.geeksforgeeks.org/types-and-part-of-data-mining-architecture/>
- [2] <https://data-flair.training/blogs/data-mining-architecture/>
- [3] <https://www.naukri.com/learning/articles/data-mining-architecture-in-data-mining-systems/>
- [4] Han, J, Kamber, M., & Pei, J. (2006). Data mining: concepts and techniques. Morgan kaufmann
- [5] <https://onix-systems.com/blog/8-data-mining-techniques-you-must-learn-to-succeed-in-business>