

## DATA MINING

### Introduction

#### DATA Mining Concepts

Data Mining - Data Mining is the process of extracting insights from large datasets using statistical and computational techniques. It can involve structured, semi-structured, or unstructured data stored in databases, data warehouses, or data lakes. Data mining involves many algorithms to accomplish different tasks.

\* TODAY, DATA is being generated at a rapid pace. Every time a click to purchase anything generates information. Data Mining is the process companies use to extract insights from this data to make smart decisions, understand customer behavior, and stay competitive.

DATA MINING - It is the process of extracting useful insights and knowledge from large datasets. It involves applying techniques like statistics, ML, and database systems to find hidden patterns, relationships, and trends. These insights help them to solve business problems, improve processes, and make future predictions.

#### COMMON APPLICATIONS OF DATA MINING-

- ① Customer Segmentation.
- ② Market Basket Analysis.
- ③ Predictive Modeling.

It is widely used across industries like finance, healthcare, retail, and telecommunication to make informed decisions. The core components of this field are ML, Statistics, database systems, data warehouse, information retrieval, high-performance computing, its applications, algorithms, visualization, and pattern recognition.

#### PROCESS OF DATA MINING-

- ① Data Cleaning And Integration - (Extracts relevant data)
- ② Data Pre-Processing - (Removes noisy data and missing values)
- ③ Pattern Recognition & ML. (Classification, Clustering, Regression).
- ④ Statistical Analysis (Mean, Mode, Median)
- ⑤ Evaluation and Interpretation.
- ⑥ Data Presentation & Visualization.