

Study Notes

■ Key Concepts

- **Data Mining (DM)**: Process of extracting **knowledge** from large amounts of data. ■
- **DM (Alt.)**: Searching raw data to find **patterns** & extract useful information. ■
- **DM Goal**: Discover hidden patterns & relationships for informed decisions or predictions. ■

■ Important Points

- DM uses various **statistical & computational techniques**. ■■
- Handles **structured, semi-structured, & unstructured data**. ■
- Data sources include databases & data warehouses. ■
- Techniques mentioned: classification & regression (for predictions). ■

■ Quick Facts

- Core idea: turn raw data into actionable knowledge. ■
- Focus is on finding what's *hidden* in the data. ■■
- Ultimate aim: make better choices or foresee outcomes. ■

■ Memory Tips

- **DM = KFLA**: **K**nowledge **F**rom **L**arge **A**mounts of Data.
- **DM Goal = HIP**: **H**idden **I**nformation for **P**redictions.

■ Reference Images

Image 1

Fundamental of Data Mining

(*) Association Rule Data Mining Concepts

Introduction:-

Data Mining Definition

Data Mining is the process of extracting knowledge from large amount of data using various statistical and computational techniques. The data can be structured, semi-structured or unstructured and can be stored in various form such as databases, data warehouses, and data lakes.

Definition:-

"Data Mining is the process of searching and analyzing a large batch of raw data in order to identify patterns and extract useful information".

Goal of data Mining

The primary goal of data Mining is to discover hidden patterns and relationships in the data that can be used to make informed decisions or predictions. This involves exploring the data using various techniques such as clustering, classification, regression

Scanned with CamScanner

Image 2

Date / /
Page No. 2.

Classification of Machine Learning

Machine learning implementations are classification into four major categories, depending on the nature of the learning "signal" or "response" available to a learning system which are as follows:

A. Supervised learning:

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs. The given data is labeled. Both classification and regression problems are supervised learning problems.

B. Unsupervised learning:

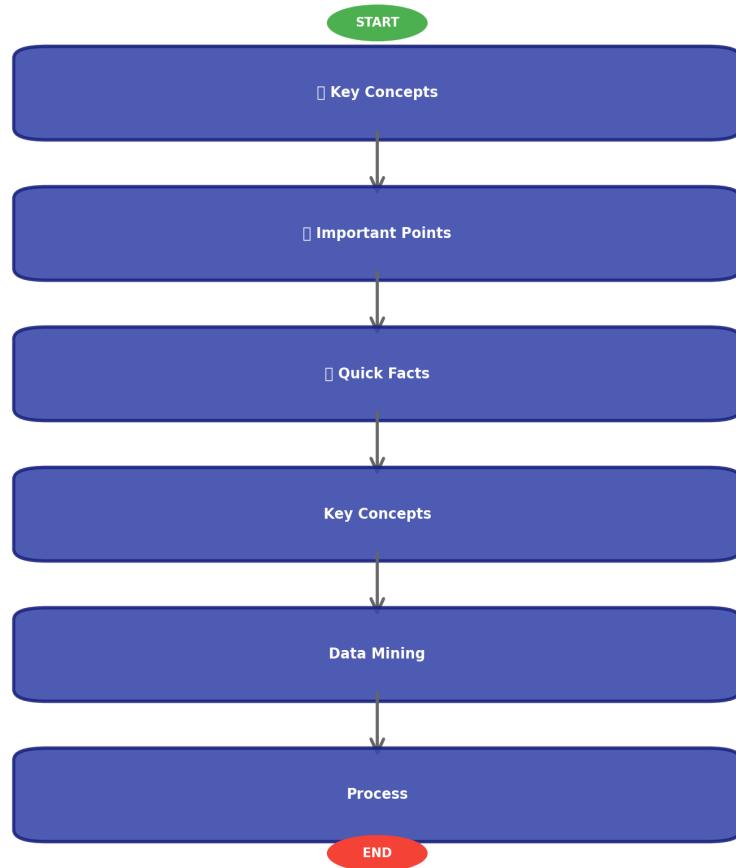
Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data

■ Visual Diagrams

Concept Mind Map



Process Flowchart



Concept Hierarchy

