Unit-1
Introduction

Arthur Samuel, an early American leader in the field of computer gaming and artificial intelligence, coined the term “Machine Learning ” in 1959 while at IBM.

Machine learning (ML) is **a branch of artificial intelligence (AI) that enables computers to “self-learn” from training data and improve over time, without being explicitly programmed**.

Machine learning algorithms are able to detect patterns in data and learn from them, in order to make their own predictions.

Definition of learning:

A computer program is said to *learn* from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks T, as measured by P , improves with experience E.

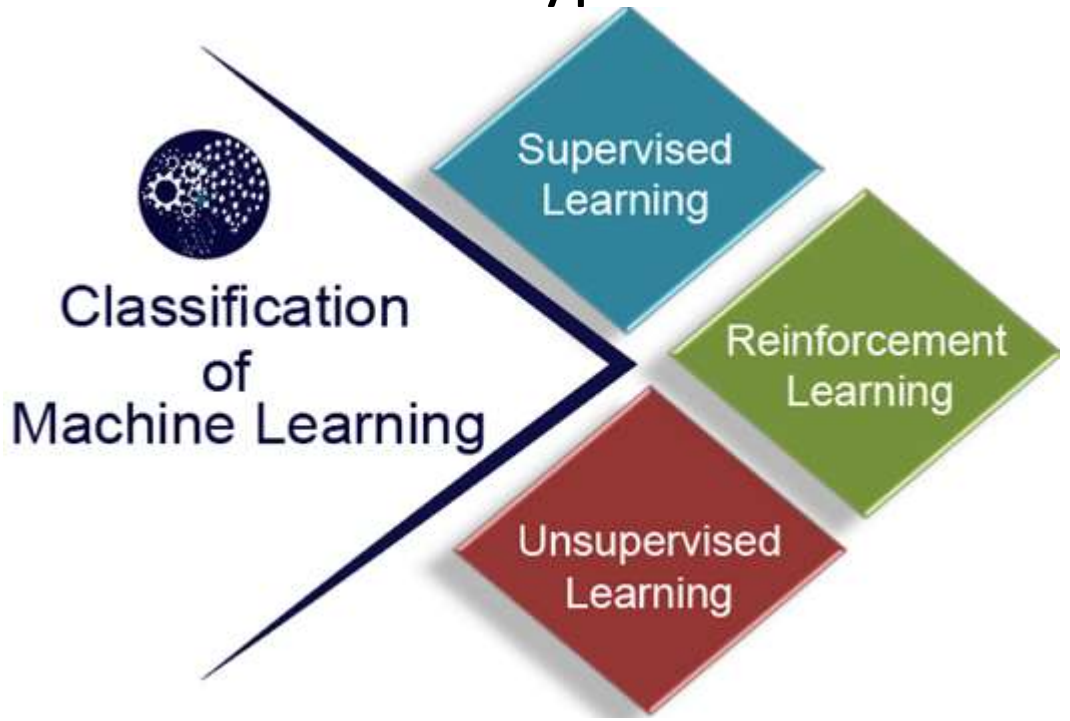
Examples

- Handwriting recognition learning problem
  - Task T : Recognizing and classifying handwritten words within images
  - Performance P : Percent of words correctly classified
  - Training experience E : A dataset of handwritten words with given classifications
- A robot driving learning problem
  - Task T : Driving on highways using vision sensors
  - Performance P : Average distance traveled before an error
  - Training experience E : A sequence of images and steering commands recorded while observing a human driver

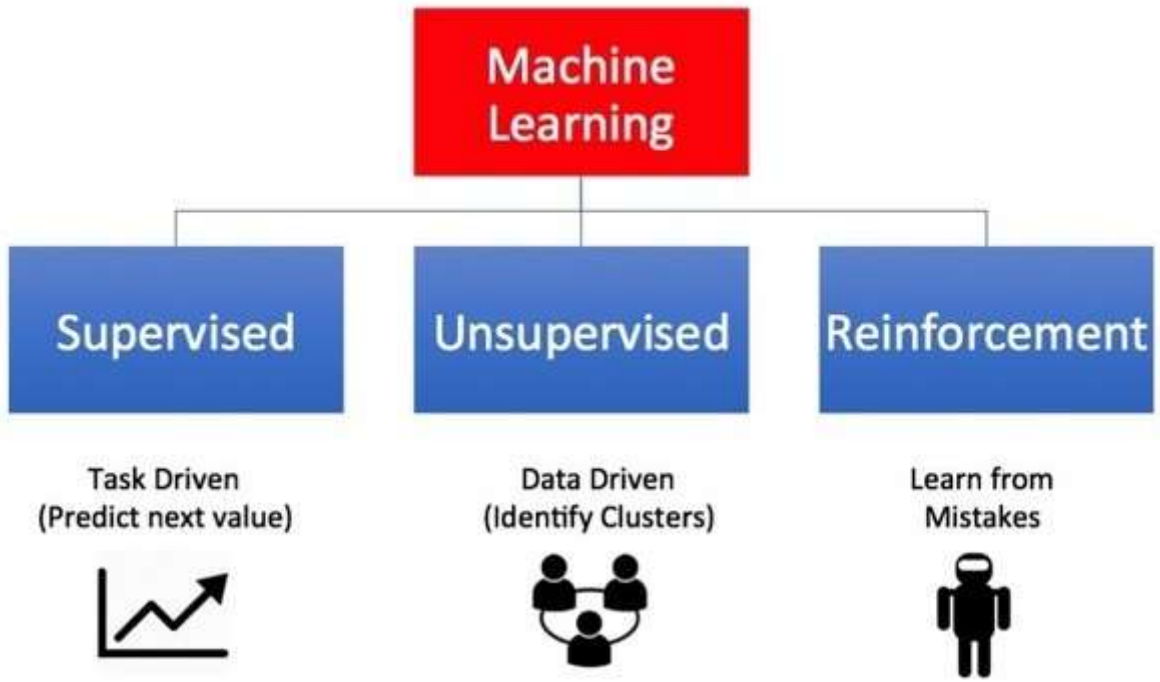
Types of Machine Learning

Machine learning can be classified into three types:

- **Supervised learning**
- **Unsupervised learning**
- **Reinforcement learning**



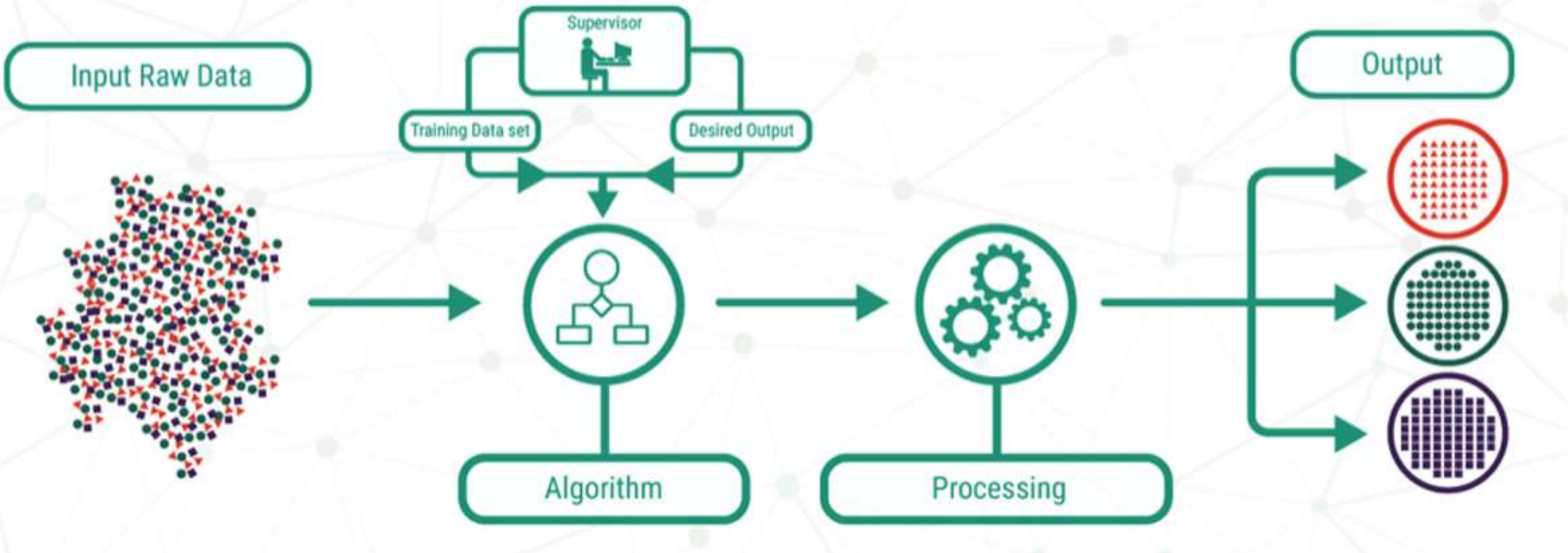
Types of Machine Learning



Supervised Learning

- Supervised learning is a type of machine learning method in which we provide sample labeled data to the machine learning system in order to train it, and on that basis, it predicts the output.
- The system creates a model using labeled data to understand the datasets and learn about each data, once the training and processing are done then we test the model by providing a sample data to check whether it is predicting the exact output or not.
- The goal of supervised learning is to map input data with the output data. The supervised learning is based on supervision, and it is the same as when a student learns things in the supervision of the teacher. The example of supervised learning is **spam filtering**.

Supervised Learning



Supervised learning can be grouped further in two categories of algorithms:

- **Classification**
- **Regression**

The main goal of the supervised learning technique is to map the input variable(x) with the output variable(y).

- **Classification**

Classification algorithms are used to solve the classification problems in which the output variable is categorical, such as "**Yes" or No, Male or Female, Red or Blue, etc**. The classification algorithms predict the categories present in the dataset. Some real-world examples of classification algorithms are **Spam Detection, Email filtering, etc.**

Some popular classification algorithms are given below:

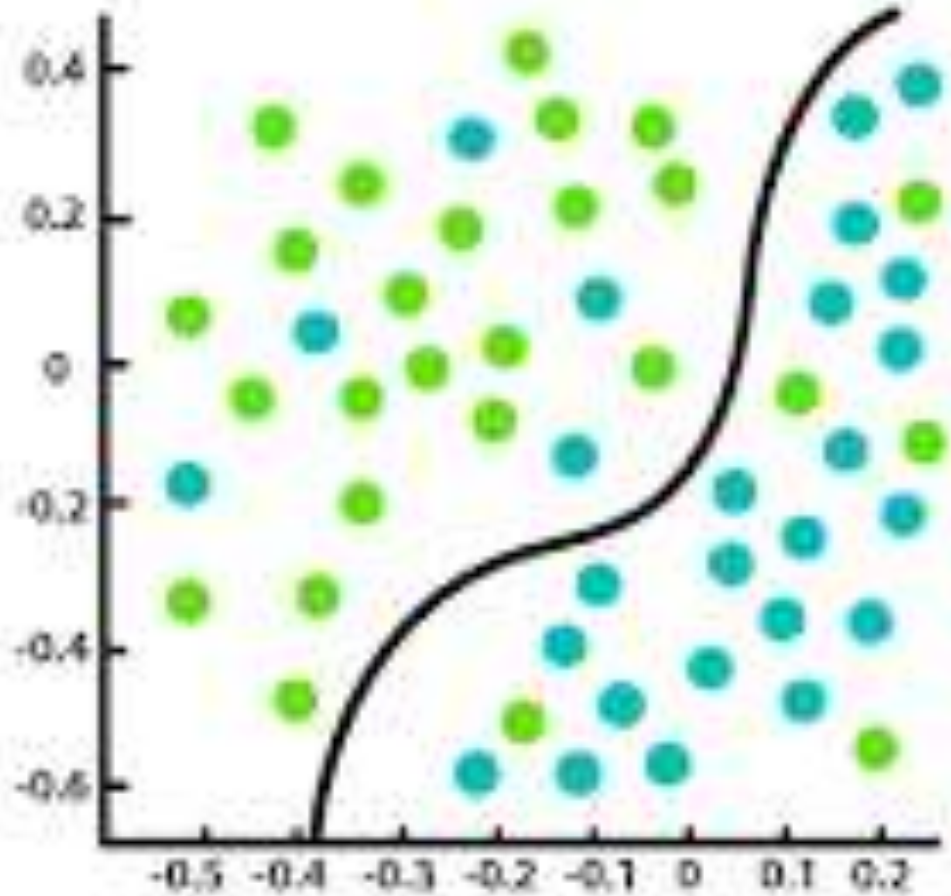
- Random Forest Algorithm
- Decision Tree Algorithm
- Logistic Regression Algorithm
- Support Vector Machine Algorithm

Regression

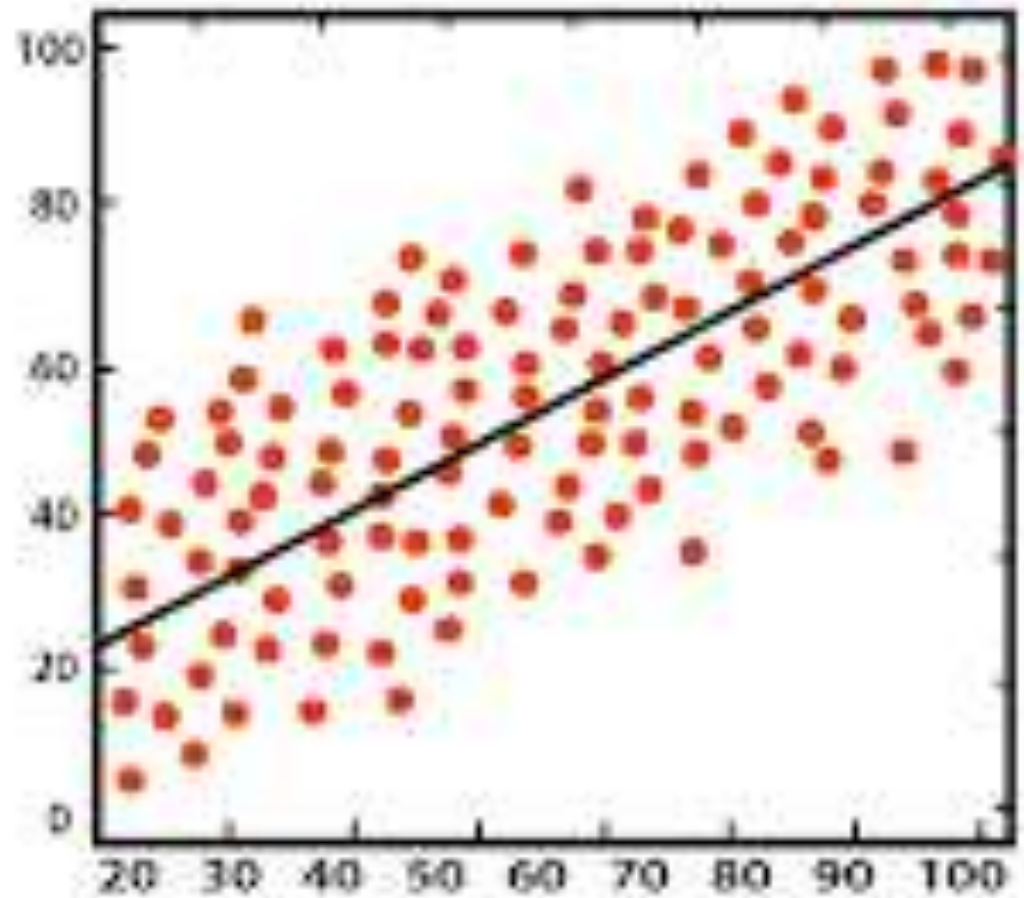
Regression algorithms are used to solve regression problems in which there is a linear relationship between input and output variables. These are used to predict continuous output variables, such as market trends, weather prediction, etc.

Some popular Regression algorithms are given below:

- Simple Linear Regression Algorithm
- Multivariate Regression Algorithm
- Decision Tree Algorithm
- Lasso Regression



Classification



Regression

Advantages:

- These algorithms are helpful in predicting the output on the basis of prior experience.

Disadvantages:

- These algorithms are not able to solve complex tasks.
- It may predict the wrong output if the test data is different from the training data.
- It requires lots of computational time to train the algorithm.

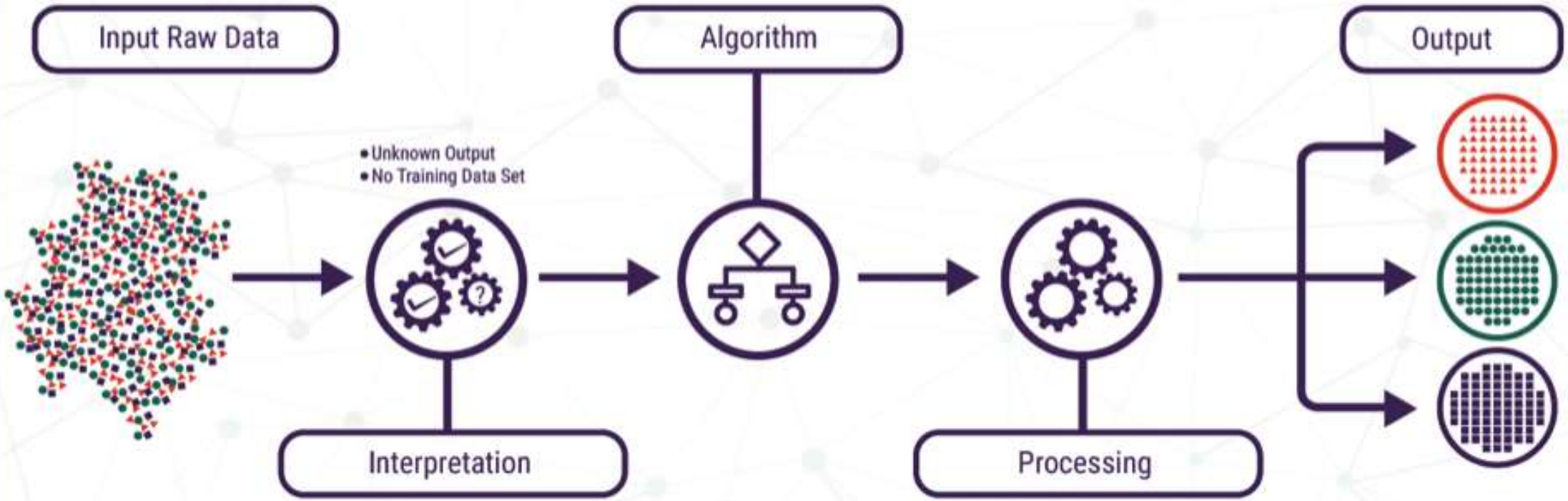
Applications of Supervised Learning:

Image Segmentation
Medical Diagnosis
Fraud Detection
Spam detection
Speech Recognition

Unsupervised Machine Learning

- The machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.
- In unsupervised learning, the models are trained with the data that is neither classified nor labelled, and the model acts on that data without any supervision.
- The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences. Machines are instructed to find the hidden patterns from the input dataset.
- So, now the machine will discover its patterns and differences, such as colour difference, shape difference, and predict the output when it is tested with the test dataset.

UNSUPERVISED LEARNING



Categories of Unsupervised Machine Learning

Unsupervised Learning can be further classified into two types

- Clustering
- Association

Clustering:

The clustering technique is used when we want to find the inherent groups from the data. It is a way to group the objects into a cluster such that the objects with the most similarities remain in one group and have fewer or no similarities with the objects of other groups. An example of the clustering algorithm is grouping the customers by their purchasing behavior.

Some of the popular clustering algorithms are given below:

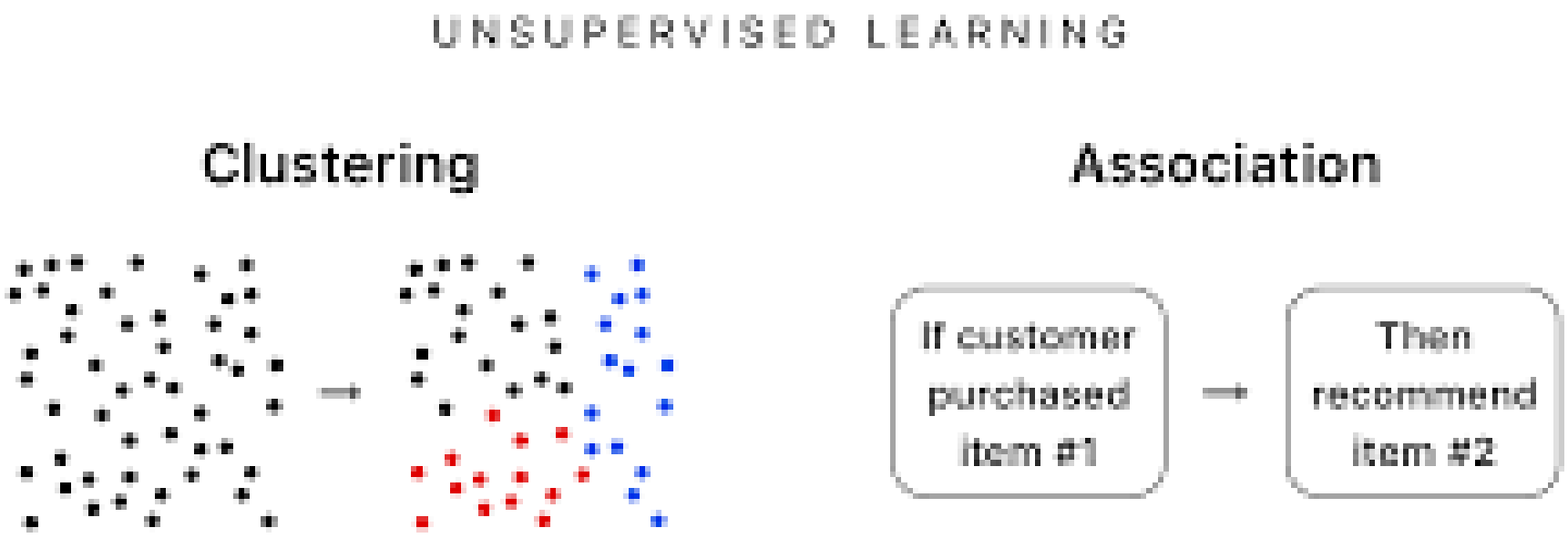
- K-Means Clustering algorithm
- Mean-shift algorithm
- DBSCAN Algorithm
- Principal Component Analysis
- Independent Component Analysis

Association:

Association rule learning is an unsupervised learning technique, which finds interesting relations among variables within a large dataset. The main aim of this learning algorithm is to find the dependency of one data item on another data item and map those variables accordingly so that it can generate maximum profit. This algorithm is mainly applied in Market Basket analysis, Web usage mining, continuous production, etc.

Some popular algorithms of Association rule learning are

- Apriori Algorithm,
- Eclat,
- FP-growth algorithm.



Advantages and Disadvantages of Unsupervised Learning Algorithm

Advantages:

- These algorithms can be used for complicated tasks compared to the supervised ones because these algorithms work on the unlabeled dataset.
- Unsupervised algorithms are preferable for various tasks as getting the unlabeled dataset is easier as compared to the labelled dataset.

Disadvantages:

- The output of an unsupervised algorithm can be less accurate as the dataset is not labelled, and algorithms are not trained with the exact output in prior.
- Working with Unsupervised learning is more difficult as it works with the unlabeled dataset that does not map with the output.

Applications of Unsupervised Learning:

- Network Analysis
- Recommendation Systems
- Anomaly Detection
- Singular Value Decomposition

Reinforcement Learning

Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance.

Agent gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning, agent is to maximize the rewards.

In this , there is no labelled data like supervised learning, and agents learn from their experiences only.

For example, a child learns various things by experiences in his day-to-day life.

Agent receives feedback in terms of punishment and rewards.


Reinforcement Learning in ML
Input Raw Data
Environment
Output
Reward
Best Action
Agent

Applications of Reinforcement Learning:

Due to its way of working, reinforcement learning is employed in different fields such as

- Game theory
- Operation Research
- Information theory
- Multi-agent systems

Categories of Reinforcement Learning:

Reinforcement learning is categorized mainly into two types of methods/algorithms:

- Positive Reinforcement Learning
- Negative Reinforcement Learning

Positive Reinforcement Learning:

Positive reinforcement learning specifies increasing the tendency that the required behavior would occur again by adding something. It enhances the strength of the behavior of the agent and positively impacts it.

Negative Reinforcement Learning:

Negative reinforcement learning works exactly opposite to the positive RL. It increases the tendency that the specific behaviour would occur again by avoiding the negative condition.

Real-world Use cases of Reinforcement Learning

- Video Games
- Resource Management
- Robotics
- Text Mining

How to practice reinforcement with kids?

POSITIVE

Making their favorite dish after they finish their homework.

Taking them to a park if they clean their room.

Clapping and cheering them every time they solve a math problem!

Complying with a request if they ask you politely.

NEGATIVE

Studying really hard to avoid getting failed in the exams.

Putting one's toys at the right place after playing to avoid getting them lost or misplaced.

Doing their homework on time to save their television privileges.

Eating healthy to avoid falling sick.

Advantages and Disadvantages of Reinforcement Learning

Advantages

- It helps in solving complex real-world problems which are difficult to be solved by general techniques.
- The learning model of RL is similar to the learning of human beings; hence most accurate results can be found.
- Helps in achieving long term results.

Disadvantage

- RL algorithms are not preferred for simple problems.
- RL algorithms require huge data and computations.
- Too much reinforcement learning can lead to an overload of states which can weaken the results.