Unit-4

MACHINE LEARNING

Overview of Machine Learning

- Machine Learning is continuously growing in the IT world and gaining strength in different business sectors.
- Although Machine Learning is in the developing phase, it is popular among all technologies.
- It is a field of study that makes computers capable of automatically learning and improving from experience.
- Hence, Machine Learning focuses on the strength of computer programs with the help of collecting data from various observations.

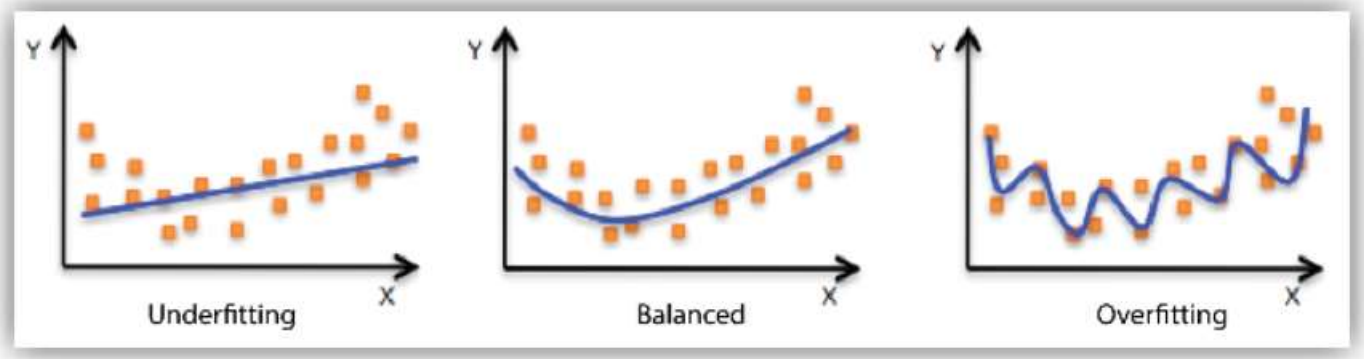
What is Machine Learning?

- Machine Learning is defined as a technology that is used to train machines to perform various actions such as predictions, recommendations, estimations, etc., based on historical data or past experience.

- Machine Learning enables computers to behave like human beings by training them with the help of past experience and predicted data.
- There are three key aspects of Machine Learning, which are as follows:
- **Task**: A task is defined as the main problem in which we are interested. This task/problem can be related to the predictions and recommendations and estimations, etc.
- **Experience**: It is defined as learning from historical or past data and used to estimate and resolve future tasks.
- **Performance**: It is defined as the capacity of any machine to resolve any machine learning task or problem and provide the best outcome for the same. However, performance is dependent on the type of machine learning problems.

– Over fitting

- Overfitting is an undesirable machine learning behavior that occurs when the machine learning model gives accurate predictions for training data but not for new data.
- When data scientists use machine learning models for making predictions, they first train the model on a known data set.
- Overfitting is a concept in data science, which occurs when a statistical model fits exactly against its training data.
- When this happens, the algorithm unfortunately cannot perform accurately against unseen data, defeating its purpose.



train/test splits

- Train test split technique is used to estimate the performance of machine learning algorithms which are used to make predictions on data not used to train the model.
- Splitting your dataset is essential for an unbiased evaluation of prediction performance.
- The train-test split procedure is appropriate when you have a very large dataset, a costly model to train, or require a good estimate of model performance quickly.
- How to use the scikit-learn machine learning library to perform the train-test split procedure.
- How to evaluate machine learning algorithms for classification and regression using the train-test split.
- It can be used for classification or regression problems and can be used for any supervised learning algorithm.

train/test splits

- The procedure involves taking a dataset and dividing it into two subsets.
- The first subset is used to fit the model and is referred to as the training dataset.
- The second subset is not used to train the model; instead, the input element of the dataset is provided to the model, then predictions are made and compared to the expected values.
- This second dataset is referred to as the test dataset.
- Train Dataset: Used to fit the machine learning model.
- Test Dataset: Used to evaluate the fit machine learning model.
- The objective is to estimate the performance of the machine learning model on new data: data not used to train the model.

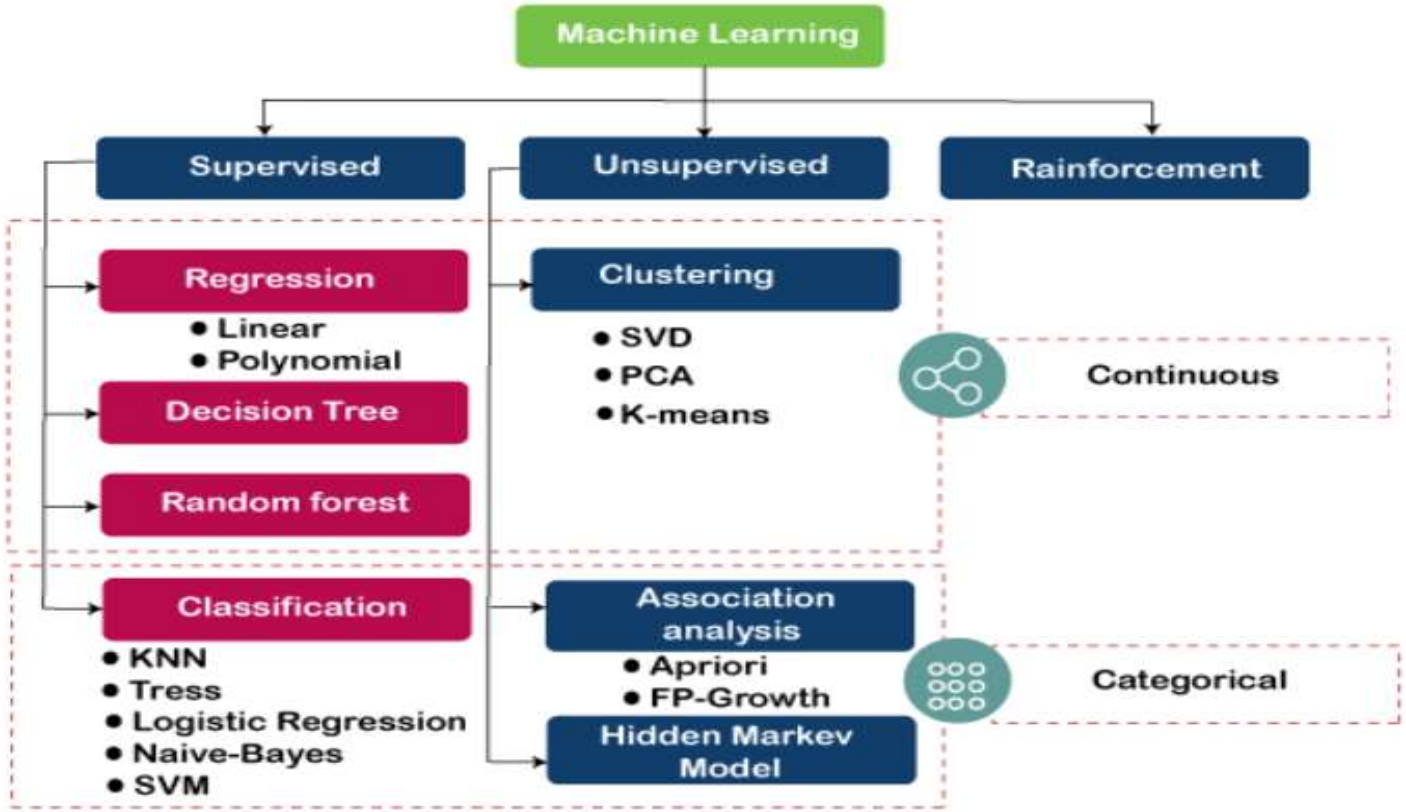
Types of Machine learning

- Machine Learning Algorithm can be broadly classified into three types:

1.Supervised Learning Algorithms

2.Unsupervised Learning Algorithms

3.Reinforcement Learning algorithm



Types of Machine learning

- Machine Learning Algorithm can be broadly classified into three types:

1.Supervised Learning Algorithms

- Supervised learning is a type of Machine learning in which the machine needs external supervision to learn.
- The supervised learning models are trained using the labeled dataset.
- Once the training and processing are done, the model is tested by providing a sample test data to check whether it predicts the correct output.
- The goal of supervised learning is to map input data with the output data.
- Supervised learning is based on supervision, and it is the same as when a student learns things in the teacher's supervision.
- The example of supervised learning is **spam filtering**.

Supervised learning can be divided further into two categories of problem:

- Classification
- Regression
- Examples of some popular supervised learning algorithms are Simple Linear regression, Decision Tree, Logistic Regression, KNN algorithm, etc

Types of Machine learning

- Machine Learning Algorithm can be broadly classified into three types:

2.Unsupervised Learning Algorithms

- It is a type of machine learning in which the machine does not need any external supervision to learn from the data, hence called unsupervised learning.
- The unsupervised models can be trained using the unlabelled dataset that is not classified, nor categorized, and the algorithm needs to act on that data without any supervision.
- In unsupervised learning, the model doesn't have a predefined output, and it tries to find useful insights from the huge amount of data.
- These are used to solve the Association and Clustering problems**. Hence further, it can be classified into two types:**
- Clustering
- Association
- Examples of some Unsupervised learning algorithms are **K-means Clustering, Apriori Algorithm, Eclat, etc.**

Types of Machine learning

- Machine Learning Algorithm can be broadly classified into three types:

3.Reinforcement Learning algorithm

- Reinforcement Learning is a feedback-based Machine learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions. For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or penalty.
- In Reinforcement Learning, the agent learns automatically using feedbacks without any labeled data, unlike supervised learning.
- Since there is no labeled data, so the agent is bound to learn by its experience only.
- RL solves a specific type of problem where decision making is sequential, and the goal is long-term, such as **game-playing, robotics**, etc.
- The agent interacts with the environment and explores it by itself. The primary goal of an agent in reinforcement learning is to improve the performance by getting the maximum positive rewards.
- The agent learns with the process of hit and trial, and based on the experience, it learns to perform the task in a better way. Hence, we can say that ***"Reinforcement learning is a type of machine learning method where an intelligent agent (computer program) interacts with the environment and learns to act within that."*** How a Robotic dog learns the movement of his arms is an example of Reinforcement learning.

•Reinforcement Learning

It is a core part of Artificial intelligence, and all AI agent works on the concept of reinforcement learning. Here we do not need to pre-program the agent, as it learns from its own experience without any human intervention.

Example: Suppose there is an AI agent present within a maze environment, and his goal is to find the diamond. The agent interacts with the environment by performing some actions, and based on those actions, the state of the agent gets changed, and it also receives a reward or penalty as feedback.