



# NATIONAL INSTITUTE OF TECHNOLOGY RAIPUR

## ASSIGNMENT 01

---

*Submitted To:*

Dr.Saurabh Gupta

Asst. Professor

Department of

Biomedical Engineering

*Submitted By :*

Deepanshu Patel

21111018

Fifth Semester

Biomedical Engineering

**Question 1: Explain the different types of Machine learning and also explain the five best algorithms of each type.**

**Answer :** Machine learning is a field of artificial intelligence that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. There are several types of machine learning, each with its unique characteristics:

- **Supervised Learning:** In supervised learning, the algorithm is trained on a labeled dataset, where each input is associated with a corresponding output. The goal is to learn a mapping between inputs and outputs so that the algorithm can make accurate predictions on new, unseen data.

**Algorithms:**

- 1) Linear Regression - It's like drawing a straight line that best fits the data to predict future values.
- 2) Decision Trees - Think of asking a series of yes-or-no questions to classify or predict something.
- 3) Random Forest - It's like combining the decisions of many trees to make more reliable predictions.
- 4) Support Vector Machines (SVM) - Imagine finding the best way to draw a line that separates different groups in the data.
- 5) Gradient Boosting Machines - It's like creating a team of decision-makers where each one corrects the mistakes of the previous one.

- **Unsupervised Learning:** Unsupervised learning deals with unlabeled data and aims to find patterns, relationships, or structure within the data.

**Algorithms:**

- 1) K-Means Clustering - It's like grouping similar things together based on their similarities.
- 2) Hierarchical Clustering - Think of organizing things in a tree-like structure where similar things are grouped together at different levels.
- 3) Principal Component Analysis (PCA) - It's like finding the most important features in the data to reduce complexity.
- 4) Independent Component Analysis (ICA) - It's like separating mixed signals into their individual sources.
- 5) t-Distributed Stochastic Neighbour Embedding (t-SNE) - Imagine visualizing high-dimensional data in a way that preserves its local structure.

- **Semi-Supervised Learning:** This type combines labelled and unlabeled data to improve learning when only a portion of the data is labelled.

**Algorithms:**

- 1) Self-Training - The model learns from labeled data and then uses its own predictions on unlabeled data to improve itself. It repeats this process iteratively.
- 2) Co-Training - Two models are trained on different parts of the data, and they share information to improve each other's predictions on unlabeled data.
- 3) Multi-View Learning - The model looks at the data from different

perspectives or views to get a more complete understanding and make better predictions.

4) Expectation-Maximization (EM) - It's like solving a puzzle with missing pieces. The model fills in the missing information step by step to improve its understanding.

5) Tri-Training - Three models collaborate and vote on the unlabeled data, and the most agreed-upon predictions become new labeled data. They help each other learn better.

- **Reinforcement Learning:** In reinforcement learning, the algorithm learns through trial and error. It interacts with an environment, gets rewards or penalties for its actions, and learns to make decisions to maximize the cumulative rewards over time.

### **Algorithms:**

1) Q-Learning - It's like learning from past experiences and estimating which actions are better in different situations.

2) Deep Q Networks (DQN) - Think of using a powerful brain (deep neural network) to make smarter decisions in complex situations.

3) Policy Gradient Methods - It's like finding a set of rules to maximize rewards by directly learning from experiences.

4) Proximal Policy Optimization (PPO) - Think of optimizing the rules of decision-making in a stable and efficient way.

5) Deep Deterministic Policy Gradients (DDPG) - It's like combining the best aspects of decision-making and using deep neural networks to handle continuous actions.

## Question 2: Explain Bagging and Boosting Ensemble Learning with an example.

**Answer:**

**Bagging (Bootstrap Aggregating):** Imagine you have a bunch of friends, and you want to make a decision. Instead of relying on just one friend's opinion, you decide to ask all of them separately and then take an average of their responses. This way, you get a more reliable and balanced decision.

In Bagging, we use the same idea with machine learning models. We take one learning algorithm (like a decision tree) and train it on different random subsets of the training data. Each subset is like a mini-group of data. Then, we let each model make its prediction on new data. The final prediction is the average of all these individual predictions (for problems like predicting prices) or the majority vote (for problems like classifying emails as spam or not).

**Example:** Random Forest is a popular bagging ensemble method. It creates an ensemble of decision trees, each trained on a bootstrapped subset of the data. The final prediction is obtained by averaging or voting on the predictions of individual trees.

**Boosting:** Imagine you are learning to play a game, and each time you make a mistake, a coach gives you special training on those specific mistakes you made. Over time, you become better and better at the game because the coach focuses on your weaknesses. In Boosting, we use a similar idea. We start with a weak learner (a simple model) and train it on the entire dataset. Then, we pay more attention to the examples that the model got wrong. We give these examples higher importance, so the next model we train will focus more on getting those tricky examples right. We keep repeating this process, with each new model trying to fix the errors made by the previous ones.

**Example:** AdaBoost (Adaptive Boosting) is a well-known boosting technique. In each iteration, it assigns higher weights to misclassified instances and trains a new model. The final prediction is obtained by weighing the predictions of all models based on their performance.

Both bagging and boosting techniques enhance the overall predictive power of machine learning models by reducing overfitting, improving generalisation, and capturing different aspects of the data.