SUPERVISED LEARNING

LINEAR REGRESSION METHOD:

Linear regression is a basic and commonly used type of predictive analysis used to explain the relationship between one dependent variable and one independent variables.

The linear regression equation with one dependent and one independent variable is defined by the formula:

$$y = a + b*x$$

where,

y = estimated dependent variable score
a = constant/intercept
b = regression coefficient/slope of x
x = independent variable.

MULTIPLE LINEAR REGRESSION METHOD:

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable.

LOGISTIC REGRESSION:

Logistic Regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.

K-NEAREST NEIGHBORS(KNN):

K-nearest neighbors (KNN) algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems. It's easy to implement and understand, but has a major drawback of becoming significantly slows as the size of that data in use grows.

SUPPORT VECTOR MACHINE(SVM):

Support Vector Machine(SVM) is a supervised learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space with the value of each feature being the value of a particular coordinate.

NAÏVE BAYES:

It is **a classification technique based on Bayes' Theorem with an assumption of independence among predictors**. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

DECISION TREE CLASSIFICATION:

Decision tree builds classification or regression models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with **decision nodes** and **leaf nodes**

RANDOM FOREST ALGORITHM:

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of *combining multiple classifiers to solve a complex problem and to improve the performance of the model.*

UNSUPERVISED LEARNING

K-MEANS CLUSTERING:

K-means clustering is **a type of unsupervised learning, which is used when you have unlabeled data** (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K

HIERARCHICAL CLUSTERING:

Hierarchical clustering is **an unsupervised clustering algorithm which involves creating clusters that have predominant ordering from top to bottom**.

PROBABILISTIC CLUSTERING:

A probabilistic model is **an unsupervised technique that helps us solve density estimation or "soft" clustering problems**. In probabilistic clustering, data points are clustered based on the likelihood that they belong to a particular distribution.

REINFORCED LEARNING

Q-LEARNING:

Q-learning is **a model-free, off-policy reinforcement learning that will find the best course of action, given the current state of the agent**. Depending on where the agent is in the environment, it will decide the next action to be taken.

POLICY OPTIMIZATION:

This methods view reinforcement learning as a numerical optimization problem where we optimize the expected reward with respect to the policy's parameters.

LEAST SQUARE METHOD:

The least squares method is a statistical procedure to find the best fit for a set of data points by minimizing the sum of the offsets or residuals of points from the plotted curve. Least squares regression is used to predict the behavior of dependent variables.