Machine Learning Algorithms

Ensemble methods

- Ensemble methods are powerful machine learning techniques that combine the predictions of multiple models to improve accuracy, robustness, and generalization compared to individual models.

- **Popular Ensemble Methods: 1. Bagging 2. Boosting**

- **Bagging :** Bagging is an ensemble machine learning technique that improves the stability and accuracy of machine learning algorithms. It involves creating multiple versions of a predictor by training each on a random subset of the training data (sampling with replacement). The final prediction is made by aggregating the predictions from all models, typically through averaging (for regression) or majority voting (for classification). Bagging reduces variance and helps prevent overfitting.

Random Forest is a supervised learning algorithm

- Imagine a forest full of decision trees, each a little different. That's kind of how Random Forest works! It's a machine learning technique that combines the predictions of many decision trees to make a better overall prediction. Here's a breakdown:
  What it is:
- Random Forest is a supervised learning algorithm, meaning it learns from labeled data to make predictions.
- It's an ensemble method, which combines multiple models (the trees) into one powerful model (the forest).

Random Forest is a supervised learning algorithm

- **How it works:**

1. **Grow the trees:** Random Forest trains many decision trees, each on a different subset of the data (with replacement, like Bagging).
2. **Randomness:** To make the trees more diverse, it also randomly selects a subset of features to consider at each split point in the tree.
3. **Vote or Average:** When making a prediction, for classification problems, the forest takes a majority vote from all the trees. For regression problems, it averages the predictions from all the trees.

- **Advantages:**

- **Accuracy:** By combining multiple trees, Random Forest is generally more accurate than a single decision tree.
- **Robustness:** The randomness helps prevent overfitting, a common problem in machine learning.
- **Flexibility:** It can handle both classification and regression tasks.
- **Easy to use:** It doesn't require a lot of fine-tuning compared to other algorithms.

- Overall, Random Forest is a powerful and versatile tool for various machine learning problems. It's like having a team of experts working together to make the best prediction possible.

- **Boosting**:

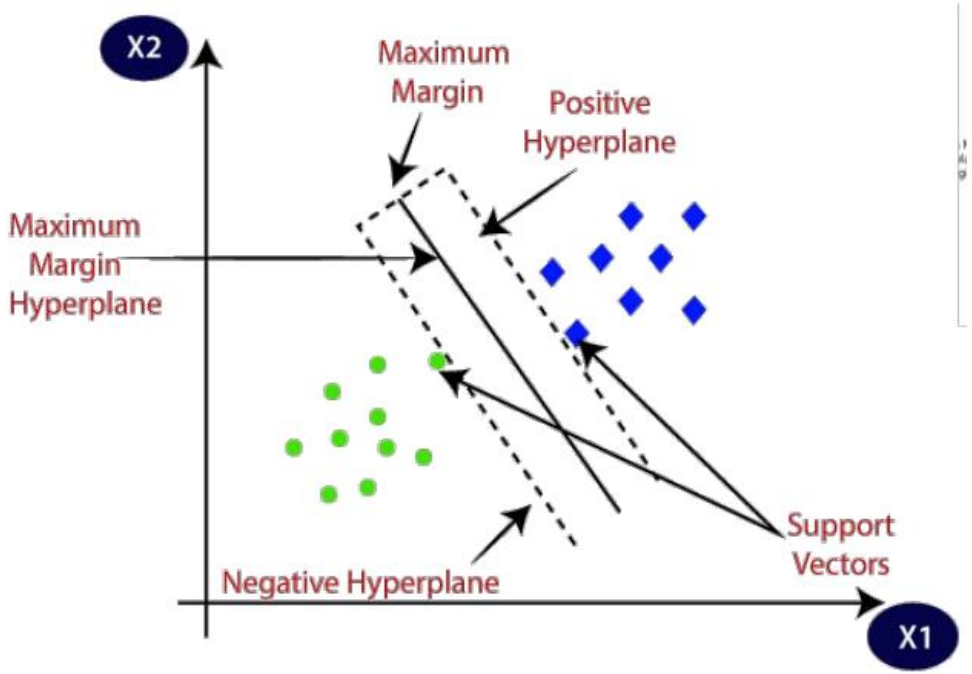
  Overview: Builds models sequentially, with each new model trying to correct the errors of the previous ones. Emphasizes harder-to-predict instances.

- **Examples**:
  - **AdaBoost**: Adjusts the weights of incorrectly classified instances so that subsequent models focus more on them.
  - **Gradient Boosting**: Optimizes a loss function by adding models that minimize the error of the current ensemble.
  - **XGBoost**: An efficient and scalable implementation of gradient boosting.

Support Vector Machine Algorithm

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine



Support Vector Machine Algorithm

Types of SVM

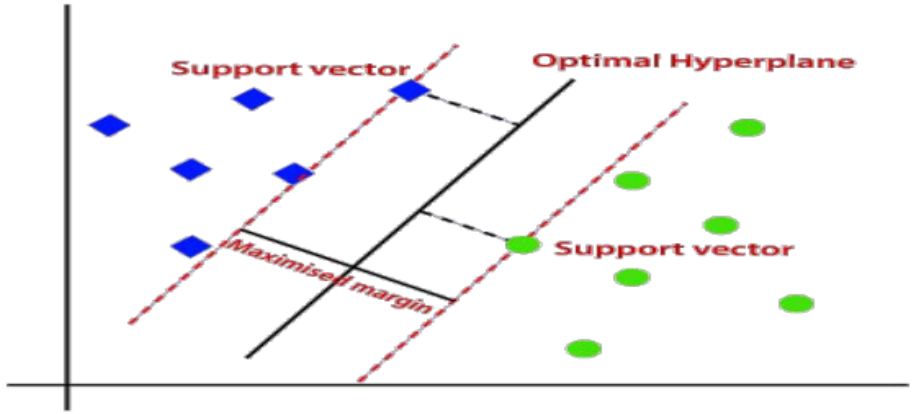
SVM can be of two types:

- **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.
- **Non-linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

Support Vector Machine Algorithm

How does SVM works?

- Linear SVM



Support Vector Machine Algorithm

How does SVM works?

- Non-Linear SVM