AdaBoost (Adaptive Boosting)

Introduction:

> AdaBoost is a machine learning ensemble technique that combines multiple weak classifiers to create a strong classifier. It was introduced by Yoav Freund and Robert Schapire in 1996.

Core Idea:

> The main idea behind AdaBoost is to sequentially train weak classifiers (often decision trees with a single split, called decision stumps) on different distributions of the training data and then combine their outputs with weighted votes.

Algorithm:

> Initialize Weights: Start with assigning equal weights to all training samples.

> Train Weak Learner: Train a weak classifier on the weighted training data.

> Compute Error: Calculate the error rate of the weak classifier on the weighted data.

> Update Weights: Increase the weights of the misclassified samples, making them more important for the next weak classifier.

> Combine Weak Classifiers: The final model is a weighted sum of the weak classifiers.

Advantages:

Improves Accuracy:

> By focusing on the hardest-to-classify examples, it significantly improves the accuracy of the combined model.

Versatile:

> Can be used with various types of weak learners.

Disadvantages:

Sensitive to Noisy Data:

> It can be susceptible to outliers and noisy data since it focuses heavily on misclassified points.

Overfitting:

> Although less prone to overfitting than many algorithms, it can still overfit if not carefully managed.

Applications:

> AdaBoost is used in various applications such as image recognition, text classification, and other areas where classification accuracy is crucial.

> AdaBoost represents a powerful and intuitive approach to boosting that has influenced many subsequent developments in ensemble learning techniques.