

BOOSTING ALGORITHM:

Boosting algorithms are a machine learning technique that combines multiple weak learners to create a strong learner. The goal is to improve the predictive power of data mining by reducing bias and variance in machine learning models.

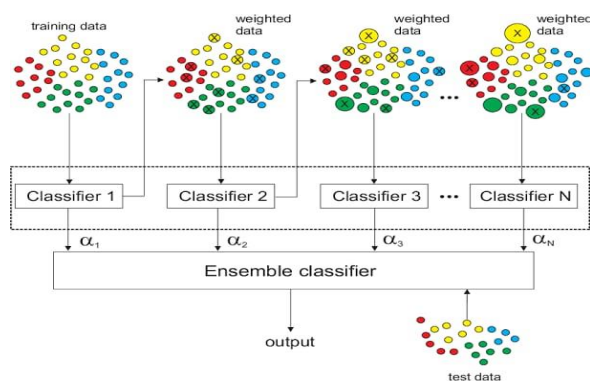
TYPES:

- ❖ Ada boosting
- ❖ XG boosting
- ❖ LG boosting

ADA BOOSTING:

AdaBoost is a boosting algorithm used in ensemble learning. It combines multiple weak classifiers to form a strong classifier by sequentially improving upon the misclassifications of prior classifiers. The algorithm assigns weights to each data point, emphasizing harder-to-classify points.

WORKING:



❑ **Step 1:** A weak classifier (usually a decision stump) is trained on the data.

❑ **Step 2:** The algorithm assigns more weight to the misclassified points from the weak classifier, forcing the next classifier to focus on these harder examples.

❑ **Step 3:** Subsequent weak classifiers are added sequentially, with each one trying to correct the mistakes made by the previous ones.

❑ **Step 4:** The final output is a weighted majority vote (or sum) of all the classifiers.

XG BOOSTING:

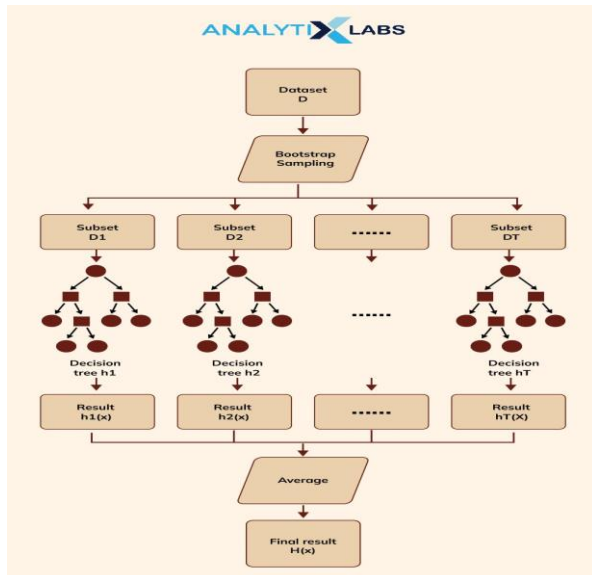
XGBoost is an optimized version of gradient boosting, designed for performance and speed. It is widely used in machine learning competitions and real-world applications due to its efficiency, scalability, and ability to handle a variety of data types. It is also known as extreme gradient boosting.

WORKING:

❑ **Step 1:** Like AdaBoost, XGBoost works by sequentially building models (trees). However, instead of assigning weights to misclassified examples, it builds new models to minimize the residual errors (the difference between predicted and actual values).

📌 **Step 2:** The algorithm uses gradient descent to optimize an objective function that measures the error, and each subsequent tree is built to correct the errors of the previous tree.

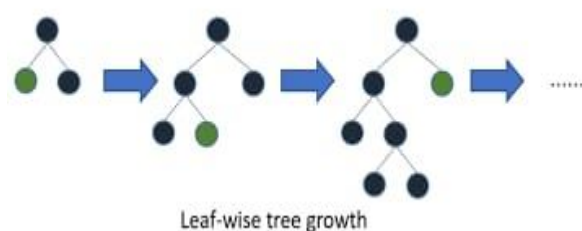
📌 **Step 3:** Regularization techniques (like L1 and L2 regularization) are used to prevent overfitting and ensure that the model generalizes well to new data.



LG BOOSTING:

LightGBM (LGBBoost) is a gradient boosting framework that is designed for high performance, particularly on large datasets. It is a faster and more memory-efficient alternative to traditional gradient boosting algorithms.

WORKING:



📌 **Step 1:** Similar to XGBoost, LightGBM uses gradient boosting to build a series of trees, with each tree correcting the errors of the previous one.

📌 **Step 2:** However, LightGBM introduces two key innovations:

- **Leaf-wise growth:** Instead of growing trees level-wise (as in traditional boosting), it grows them leaf-wise. It selects the leaf with the maximum delta loss to grow, which leads to deeper, more complex trees.
- **Histogram-based approach:** LightGBM discretizes continuous features into discrete bins (histograms), which reduces memory usage and speeds up computations.

📌 **Step 3:** These optimizations make LightGBM faster and more efficient, especially for large datasets with high-dimensional data.

ADVANTAGES:

- Improved Accuracy
- Versatility
- Reduction of Bias and Variance

DISADVANTAGES:

- Sensitivity to Outliers
- Computationally Expensive
- Overfitting

APPLICATIONS:

- Fraud Detection
- Credit Scoring and Risk Management
- Ranking Problems