
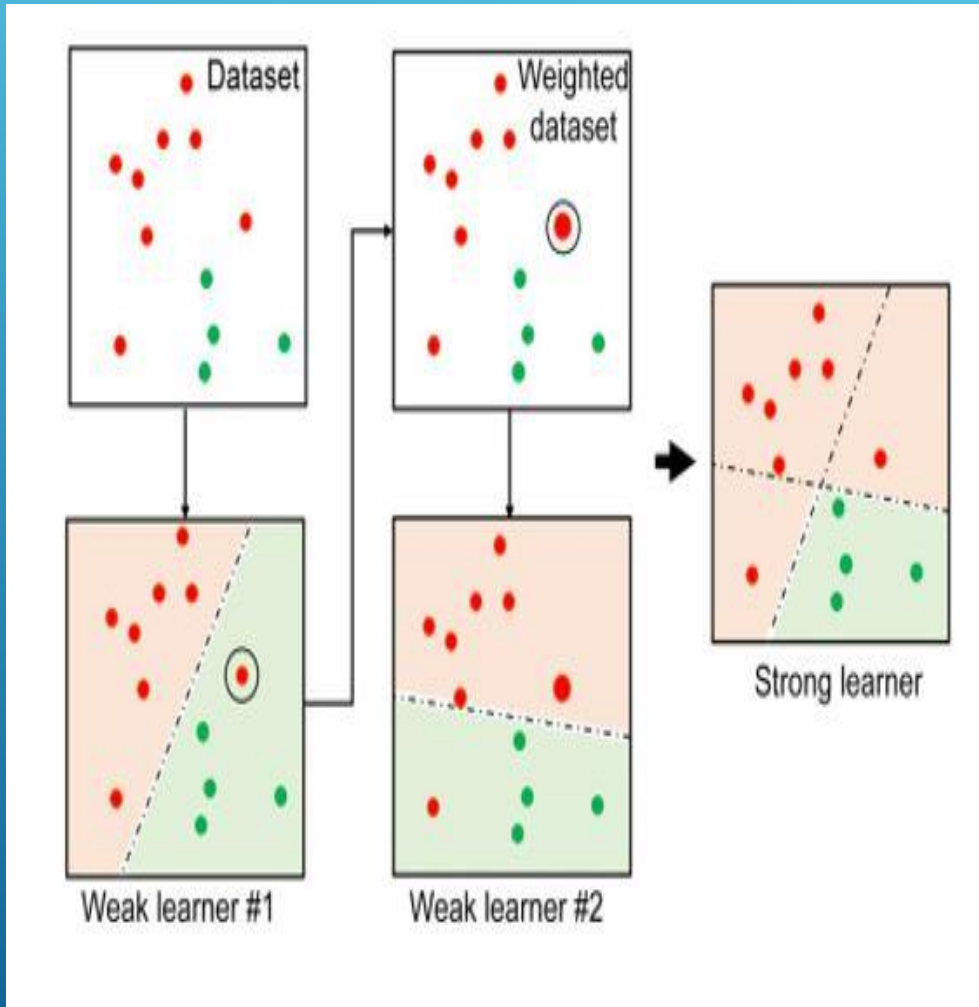


BOOSTING ALGORITHMS

- Ada Boost
 - XG Boosting
 - LG Boosting
- 
- A series of several thin, white, parallel diagonal lines extending from the bottom left towards the top right, located on the right side of the slide.

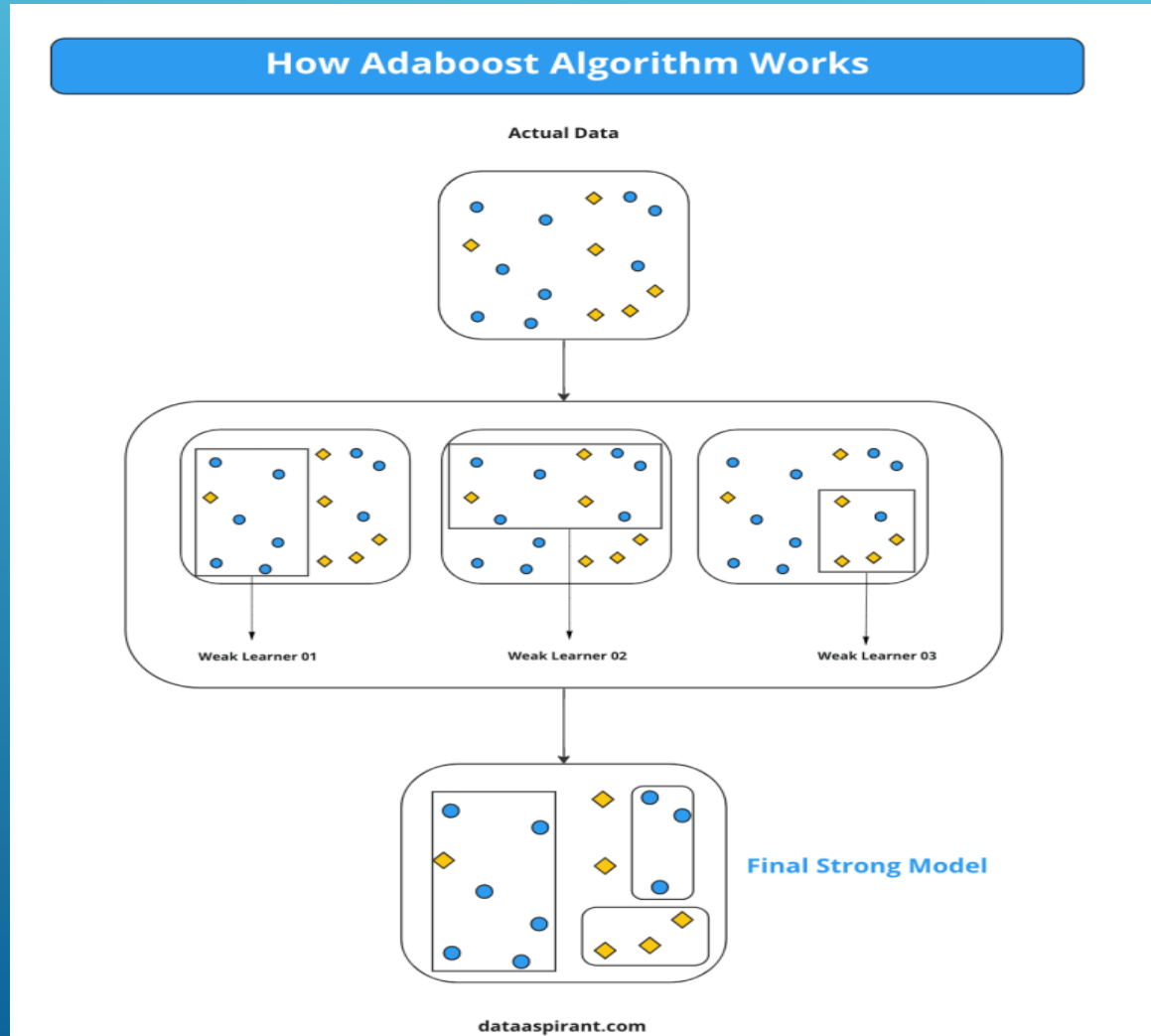
ADA BOOST



AdaBoost- Adaptive Boosting, is a powerful ensemble machine learning algorithm used for improving the accuracy of weak learners.

- It combines multiple weak classifiers to create a strong classifier.
- Weak learners are typically decision trees with limited depth.
- AdaBoost assigns weights to training examples, giving more weight to misclassified samples in each iteration.

HOW ADABOOST WORKS



1. Initialize equal weights for all training examples.
2. Train a weak classifier and compute its error.
3. Increase the weight of misclassified examples.
4. Repeat steps 2-3 for a predefined number of iterations.
5. Combine the weak classifiers into a strong classifier using weighted majority voting.

Examples of AdaBoost Applications

- Face Detection: AdaBoost is used in Viola-Jones face detection.
- Text Classification: It's applied to spam email classification.
- Medical Diagnosis: AdaBoost helps in disease diagnosis based on patient data.
- Stock Market Prediction: Used for predicting stock price movements.



Advantages & Disadvantages



Accuracy Improvement



Versatility



Robustness



Speed



Scalability



Sensitivity to Noisy Data



Complex Implementation



Computational Cost

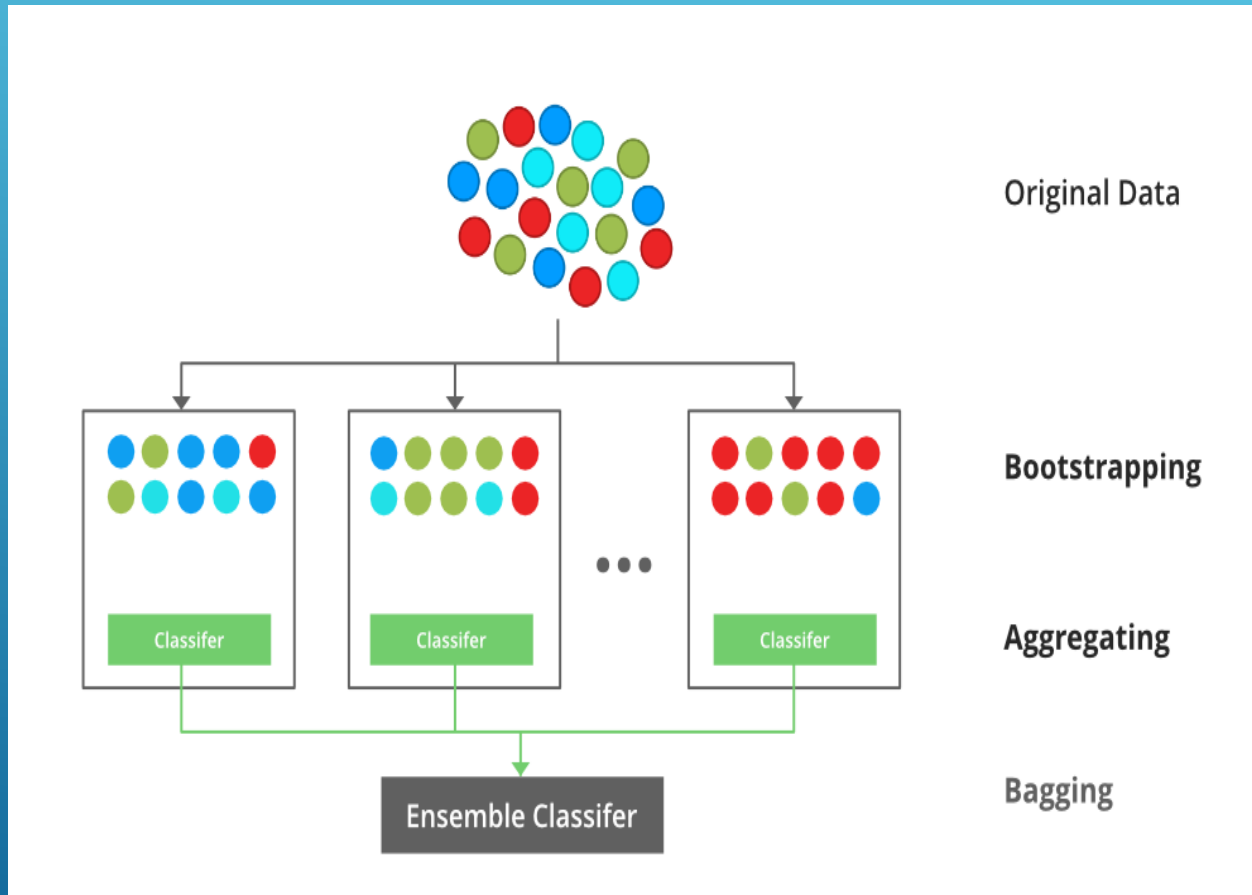


Bias towards Certain Features



Vulnerability to Outliers

XG BOOST- EXTREME GRADIENT BOOSTING




popular and powerful machine learning algorithm used for regression and classification tasks. It is based on the gradient boosting framework and is known for its efficiency, speed, and performance in a wide range of machine learning problems.

Key Features

- ▶ Gradient boosting framework.
 - ▶ Efficient and fast.
 - ▶ Handles missing data.
 - ▶ Regularization for preventing overfitting.
-
- ▶ Widely used for regression and classification.

How XGBoost Works


- ▶ Gradient Boosting:
 - ▶ Builds multiple decision trees sequentially.
 - ▶ Each tree corrects errors of the previous ones.
 - ▶ Regularization:
 - ▶ L1 (Lasso) and L2 (Ridge) regularization.
 - ▶ Reduces overfitting.
 - ▶ Parallel Processing:
 - ▶ Efficiently uses multiple CPU cores.
 - ▶ Distributed computing for scalability.
- 
- Several white lines of varying lengths and slopes are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

▶ **Use Cases and Benefits**

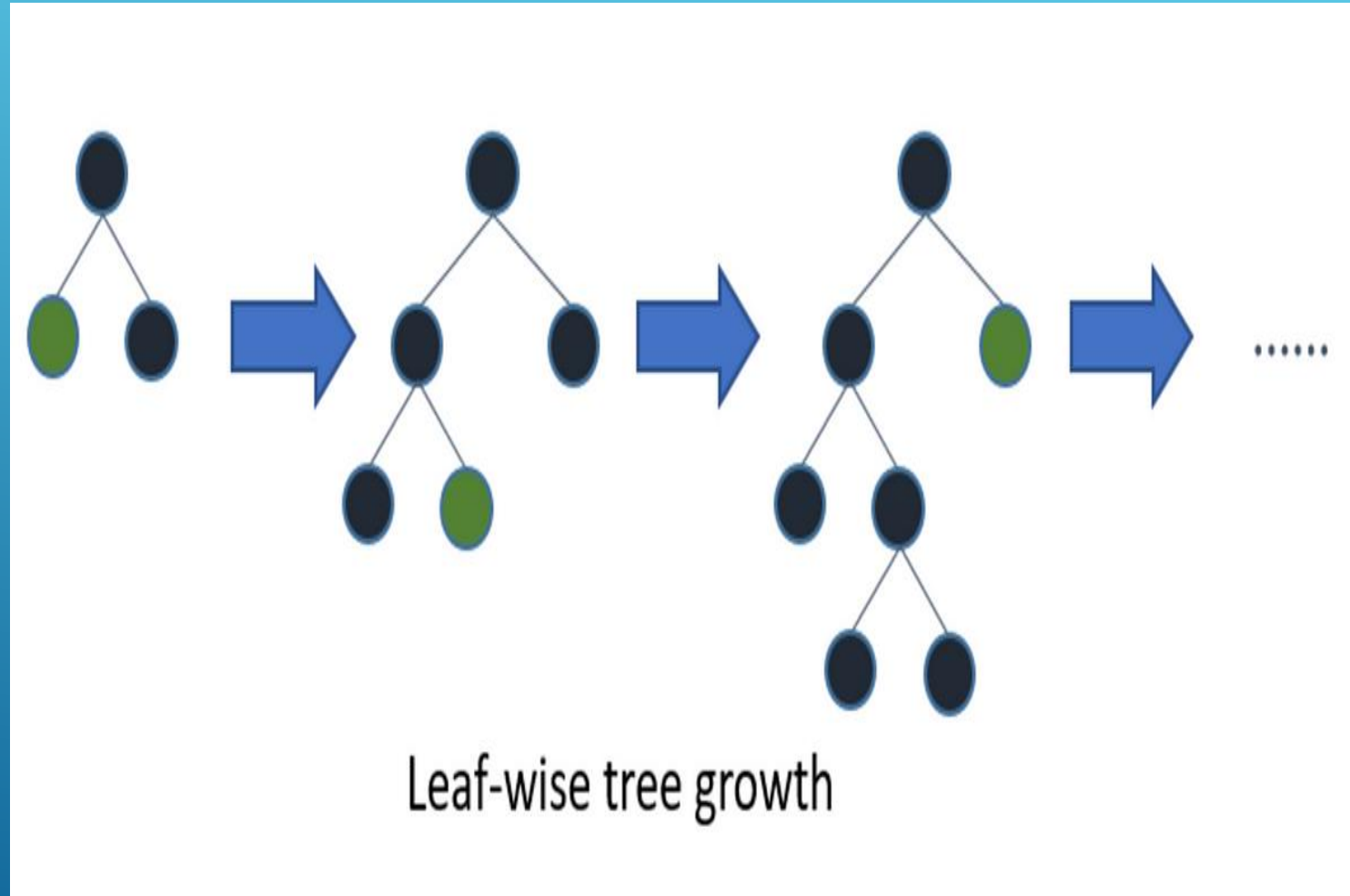
▶ Use Cases:

- ▶ Regression and classification tasks.
- ▶ Numeric and categorical data.
- ▶ Data with missing values.

▶ Benefits:


- ▶ Fast and efficient.
 - ▶ Excellent performance in competitions.
 - ▶ Feature importance analysis.
 - ▶ Early stopping for preventing overfitting.
- 
- Several white lines of varying lengths and slopes are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

LIGHT GRADIENT BOOSTING




LightGBM is a gradient boosting ensemble method that is used by the Train Using AutoML tool and is based on decision trees. As with other decision tree-based methods, LightGBM can be used for both classification and regression. LightGBM is optimized for high performance with distributed systems.


KEY FEATURES:

- ▶ Speed and efficiency.
 - ▶ Histogram-based algorithm for decision trees.
 - ▶ Categorical feature handling.
 - ▶ Parallel and GPU learning.
 - ▶ Cross-validation and regularization.
- 
- A series of white diagonal lines of varying lengths and thicknesses, located in the bottom right corner of the slide, creating a modern, abstract graphic element.


Why Choose LightGBM?

- ▶ **Speed and Efficiency:**
 - ▶ Faster training and prediction.
 - ▶ Memory-efficient histogram-based algorithm.
 - ▶ **Categorical Feature Handling:**
 - ▶ Built-in support for categorical data.
 - ▶ **Parallel and GPU Learning:**
 - ▶ Utilizes multiple CPU cores and GPUs for acceleration.
- 
- A series of white diagonal lines of varying lengths and thicknesses are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

HOW TO USE LIGHTGBM

- ▶ Data Preparation:
 - ▶ Load and split your dataset.
 - ▶ LightGBM Dataset:
 - ▶ Create a LightGBM dataset for efficient data handling.
 - ▶ Hyperparameter Definition:
 - ▶ Set parameters like objective, boosting type, and learning rate.
 - ▶ Model Training:
 - ▶ Train the model with your data.
 - ▶ Evaluation:
 - ▶ Evaluate model performance and make predictions.
- 
- Several white lines of varying lengths and orientations are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

Applications

- ▶ LightGBM can be used for a variety of machine learning tasks, including:
 - ▶ Classification: Predicting a category for a given input, such as predicting whether a customer will churn or not.
 - ▶ Regression: Predicting a continuous value for a given input, such as predicting the price of a house.
 - ▶ Ranking: Ranking items based on a given metric, such as ranking search results or product recommendations
- 
- Several white lines of varying lengths and orientations are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

Difference Between Boosting Algorithms

Algorithms	Gradient Boosting	AdaBoost	XGBoost	CatBoost	LightGBM
Year	–	1995	2014	2017	2017
Handling Categorical Variables	May require preprocessing like one-hot encoding	No	NO	Automatically handles categorical variables	No
Speed/Scalability	Moderate	Fast	Fast	Moderate	Fast
Memory Usage	Moderate	Low	Moderate	High	Low
Regularization	NO	No	Yes	Yes	Yes
Parallel Processing	No	No	Yes	Yes	Yes
GPU Support	No	No	Yes	Yes	Yes
Feature Importance	Available	Available	Available	Available	Available