Boosting Algorithms

LightGBM Boost Algorithm

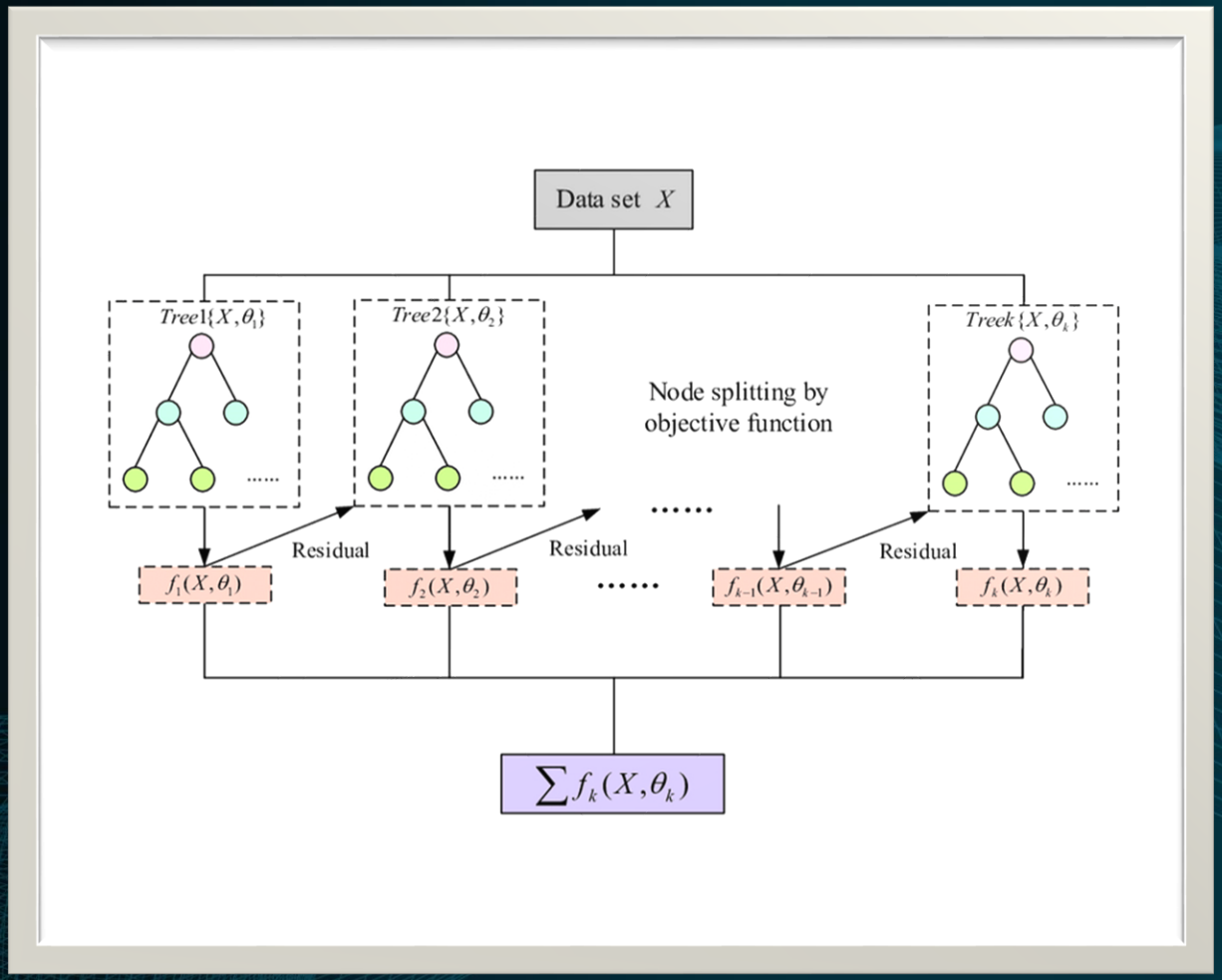
What is the LightGBM Boost Algorithm?

- **Light GBM** is a fast, distributed, high-performance **gradient boosting** framework that uses a **tree-based** learning algorithm. It also supports **GPU learning** and is thus widely used for data science application development.
- **Light GBM** splits the tree **leaf-wise** with the best fit whereas other boosting algorithms split the tree depth-wise or level-wise rather than leaf-wise. In other words, Light GBM grows trees vertically while other algorithms grow trees horizontally.

Working Steps of LightGBM Boost

- Step 1: Loading data into LightGBM.
- Step 2: Train a model using LightGBM.
- Step 3: Cross-validation and hyperparameter tuning. LightGBM with gridsearch.
- Step 4: LightGBM evaluation metrics. LightGBM with custom loss functions.

Flow chart of LightGBM Boost Algorithm



https://images.app.goo.gl/m5YwAuzkA9Q652pR9

Light GBM Boost Algorithm

Advantage	Disadvantage	Application
Faster training speed and higher efficiency compared to other boosting algorithms	Complexity in Tuning	Used in Classification and Regression Tasks
Can handle large-scale data with ease	Limited Interpretability	Used in ranking applications
Can deliver high accuracy in both classification and regression tasks	Relies on specific system dependencies which might not be readily available on all systems.	Applied in time series prediction tasks
Directly handles categorical features without needing one-hot encoding	Can be sensitive to noisy data	Effective in detecting anomalies in data due to its robustness and accuracy
Supports parallel learning, leading to faster computation times	Lead to overfitting, especially with small datasets.	Used in credit scoring, risk assessment, and portfolio optimization

Reference:

- https://www.kaggle.com/discussions/general/264327

THANK YOU!