Boosting Algorithm

1.ADA BOOSTING ALGORITHM

2.XG BOOSTING ALGORITHM\

3.LIGHT GBM ALGORITHM

1.Boosting Algorithms

WHY BOOSTING ALGORITHMS?

In machine learning, where datasets often exhibit high complexity, boosting algorithms play a major role in obtaining optimal solutions and achieving better model performance.

Ensemble Learning: Boosting is a type of ensemble learning where multiple weak learners are combined to form a strong learner.

Sequential Training: Boosting trains weak learners sequentially, with each new learner focusing on the mistakes of the previous ones.

Weighted Combination: The final prediction is a weighted sum of the predictions from all weak learners, with weights assigned based on the individual learner's performance.

Adaptability: Boosting is adaptable to different types of weak learners and can be used with various base models.

1.Analyse and draw stumps(Algorithm)

2.False prediction is highter weightage

3.repect step too until with right prediction called (sequentially)

BAGGING-work in parallel

BOOSTING-Sequentially

Boosting Algorithm

.Boosting Algorithm using Complex dataset

1.ENSEMBLE LEARNING(WEEK LEARNING)

.CAR AS (A)

.MORE THAN TWO WHEEL CONSIDER AS (A)

.SEAT BELT CONSIDER AS (A)

. 4 SITTING CONSIDER AS (A)

2.ENSEMBLE LEARNING(WEEK LEARNING)

.BIKE AS (B)

.TWO WHEEL IS PRESENT CONSIDER AS (B)

.2 SITTING CONSIDER AS (B)

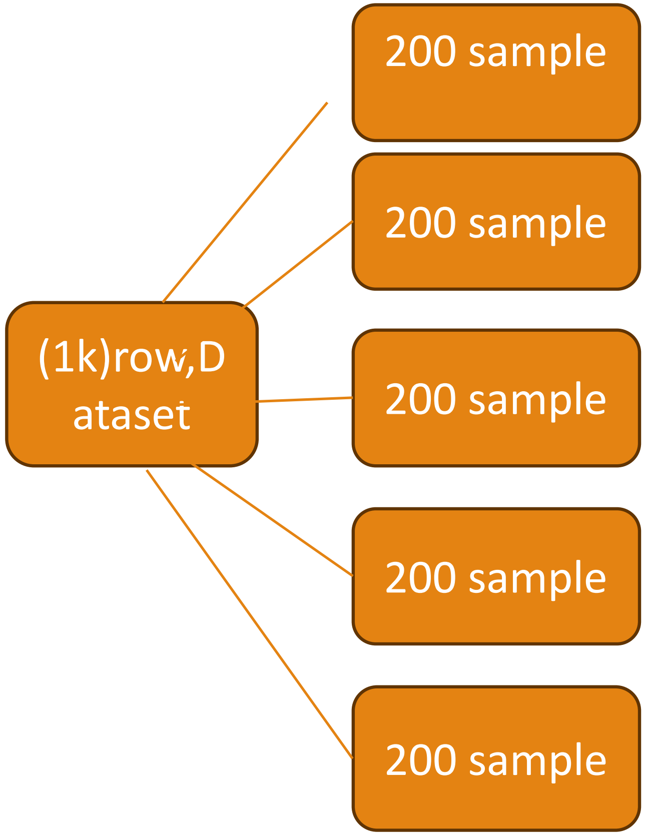
WeakLearning Preforming

In our rule-based system, we assign labels A, B, and C to represent CAR, BIKE, and unknow, respectively. The condition-checking process continues by evaluating the input and determining the majority to produce the correct output.

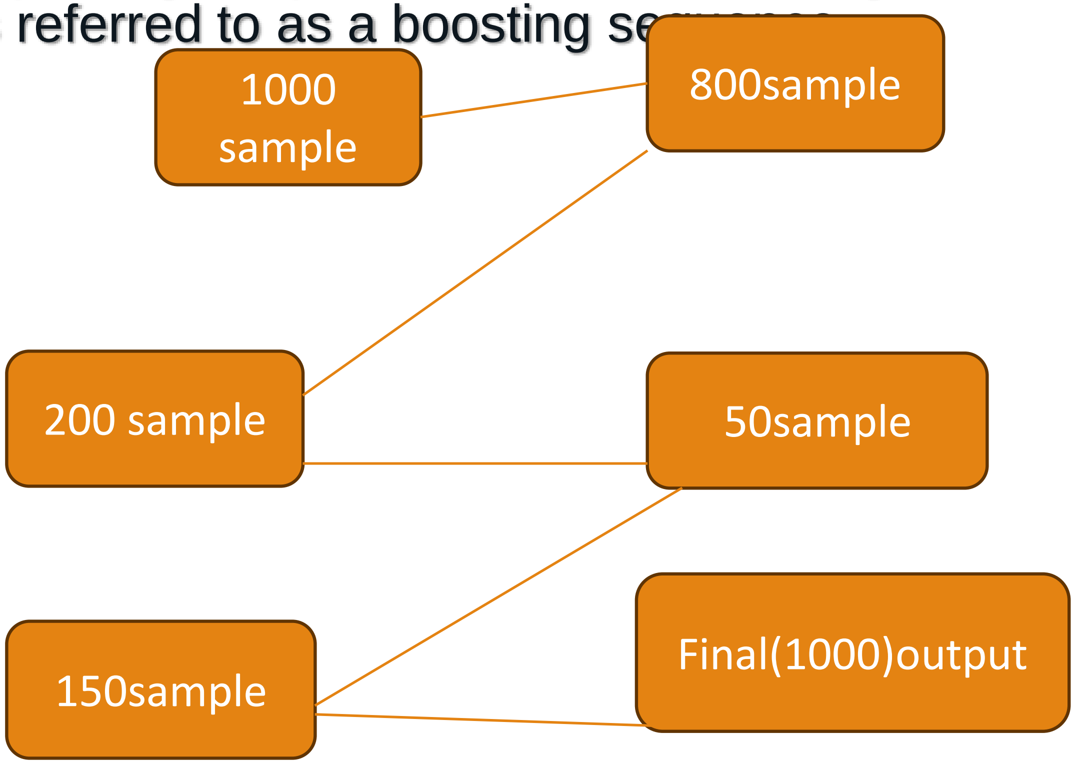
Bagging-Parallel

Boosting-sequence

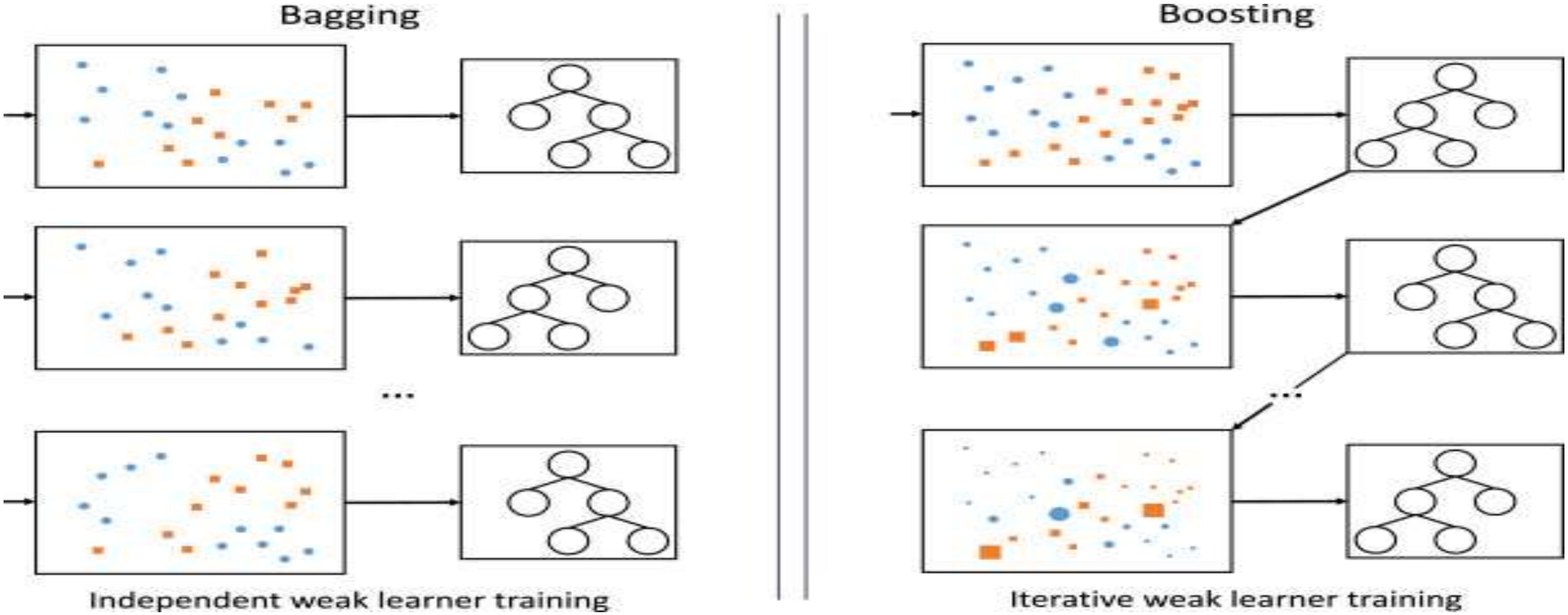
1.Dataset-output are Parallel(decision tree),if dataset have a 1000 row sample ,features sampling each 200 to div parallel(processing)



In the context of a dataset where the output is a sequence of decision trees, if there are 1,000 samples and 800 of them are correctly classified while 200 are mismatched, and the process involves giving more importance to the mismatched samples, iterating through them, and finding correct predictions for 50 samples initially and then repeating the process for the remaining 150 samples, this is referred to as a boosting se



Bagging-Parallel Boosting-sequence


Bagging
...
Independent weak learner training
Boosting
...
Iterative weak learner training

Boosting algorithm for classification + or -

Weighted Focus(d1) Assign higher weights to misclassified instances, making them more important for the next model

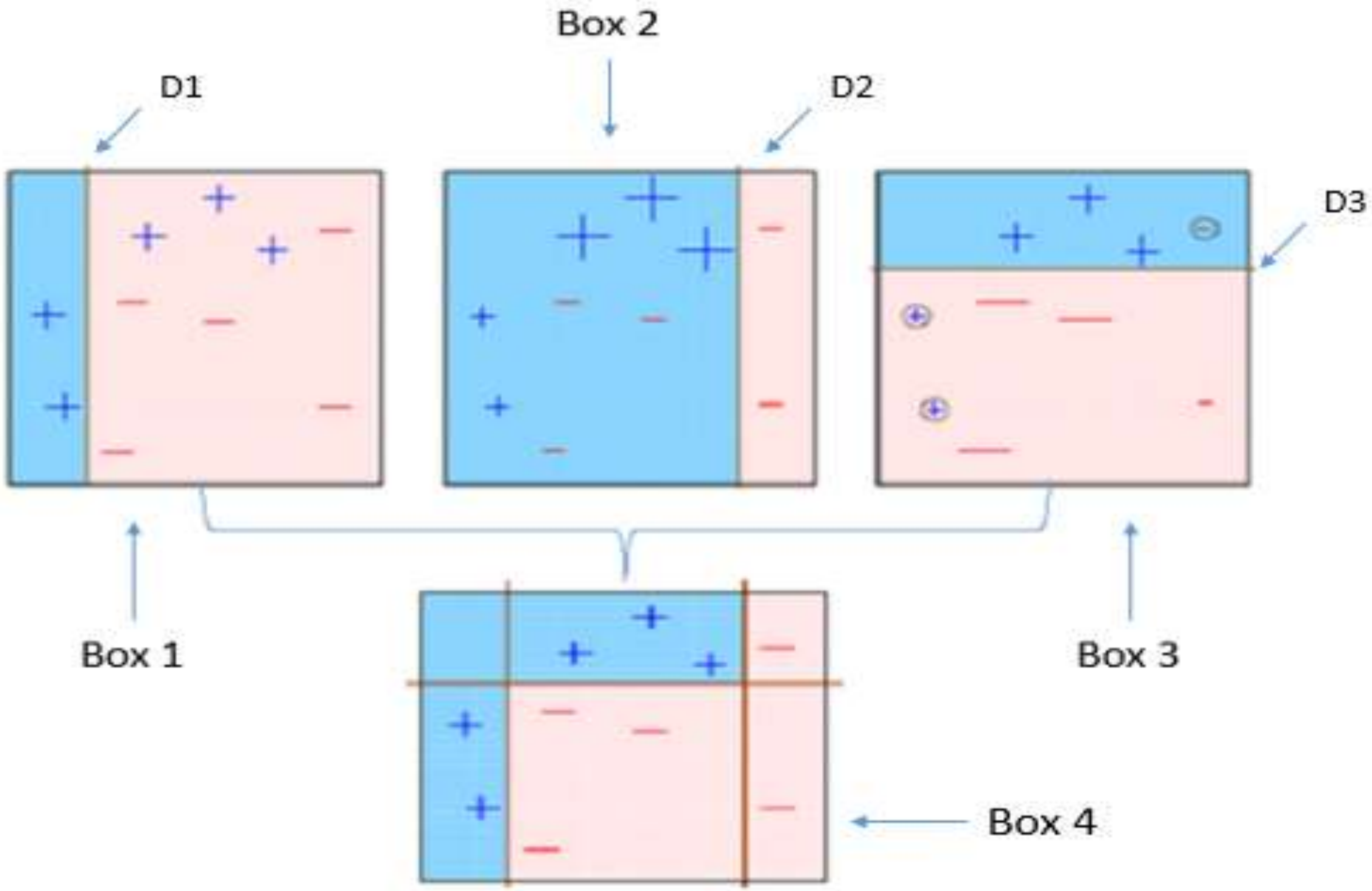
Build Next Model(d2) Train a new model, giving more emphasis to the previously misclassified I

Base Model(d3) Start with a simple model, often a shallow decision tree (a weak learner).nstances.

Combine Predictions(box4) Combine predictions from all models, giving more weight to better-performing models.

*The boosting process continues iteratively until a predefined number of weak learners have been trained or until a satisfactory level of accuracy is achieved. It does not necessarily aim for a (perfect) output,. final output is a combination of predictions from all the weak learners, with weights assigned based on their individual performance

* The goal is to create a strong, generalized model that performs well on unseen data.



Adaboosting algorithm

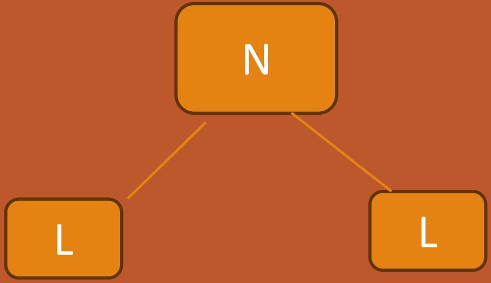
*AdaBoost (Adaptive Boosting) is an ensemble learning algorithm that combines multiple weak classifiers to create a strong classifier by giving more weight to misclassified instances in each iteration.

*meta-estimator-high model(fitting)-create_model,(fit) on a original dataset

*The weight of istances are adjusted according to the error of the current prediction

*In forest of tree made with AdaBOOSTING tree usually just node and two leaves

EX:

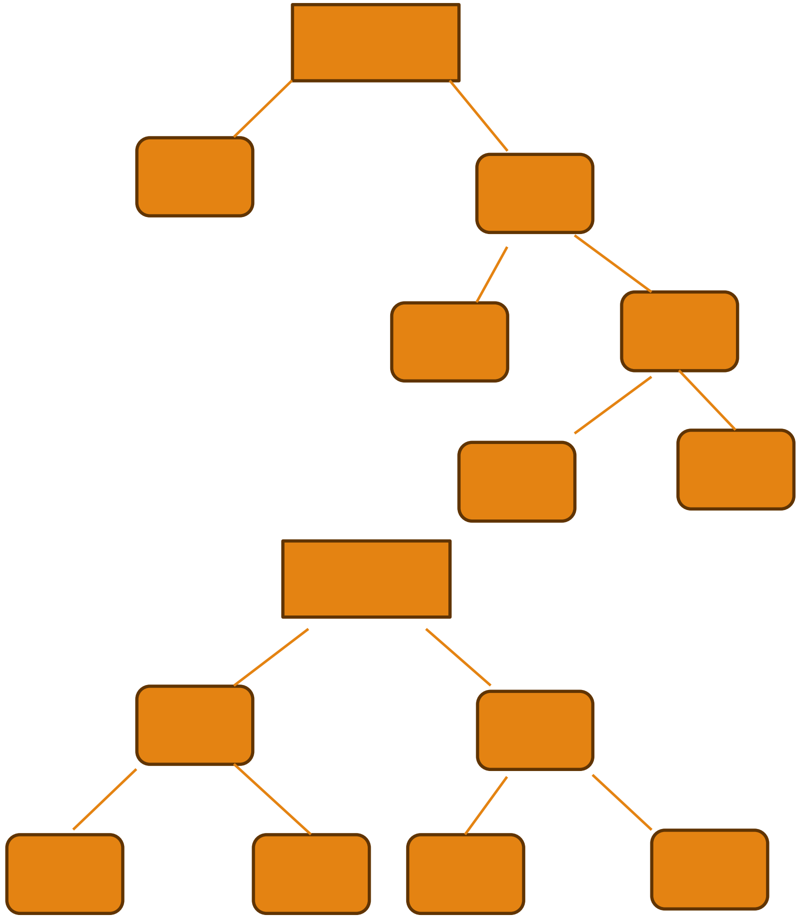


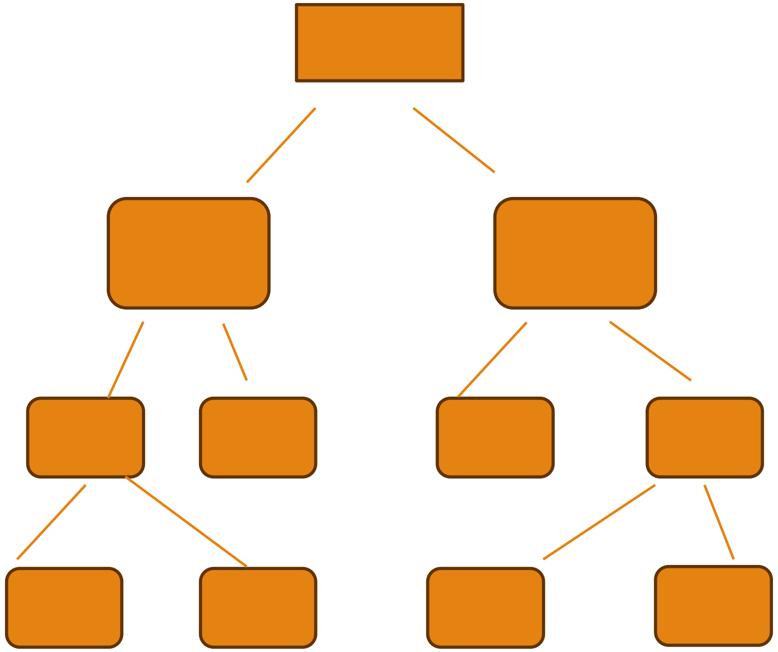
Randon Forest:

Because that is most common way to use Adaboosting

ex:

Random Forest





Decision Tree

STUMP

*Ada boosting combines a lot of [WEAK LEARNER”]

*Stump is made by taking previous stump mistake by own

*Stump can only use variable to make a decision

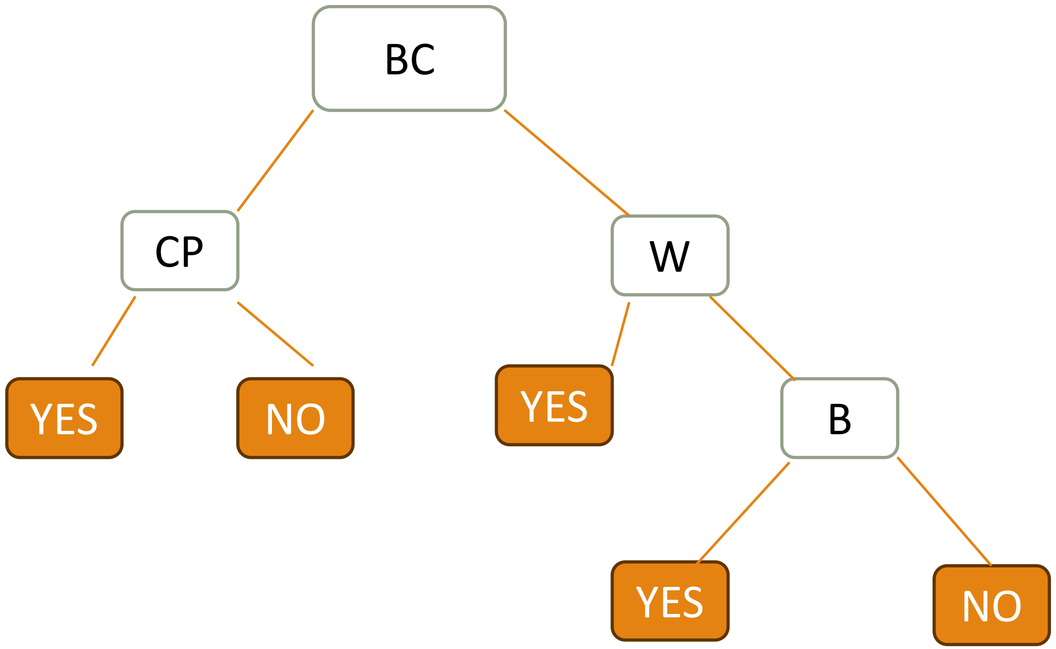
*sample weight=1/8

*start all sample get same weight

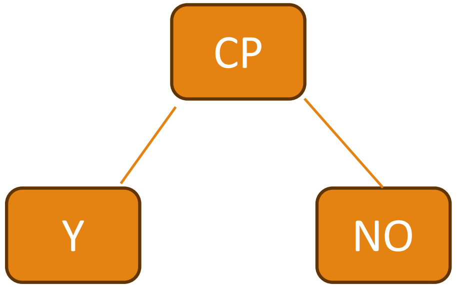
*amount of sa-1/2 log(1-total error/total error)

Chest pain	Blood cir	block	weight	heart	Sample weight
no	no	no	125	no	1/8
yes	yes	yes	180	yes	1/8
yes	yes	no	210	no	1/8
yes	no	yes	167	yes	1/8

Decision tree

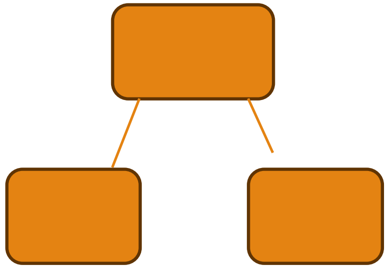


Stump can only use one variable to make a decision



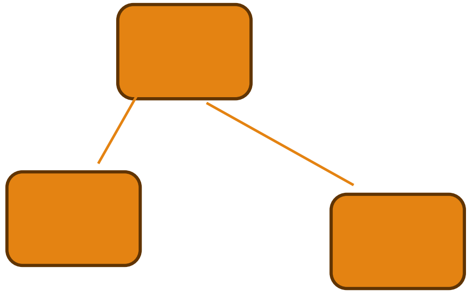
Stump are technically “weak learners”

1.



The error that the first stump makes

2.



Create other stump makes(error that the second stump makes ,then third stump make error then create other stump,then other until error of the current prediction

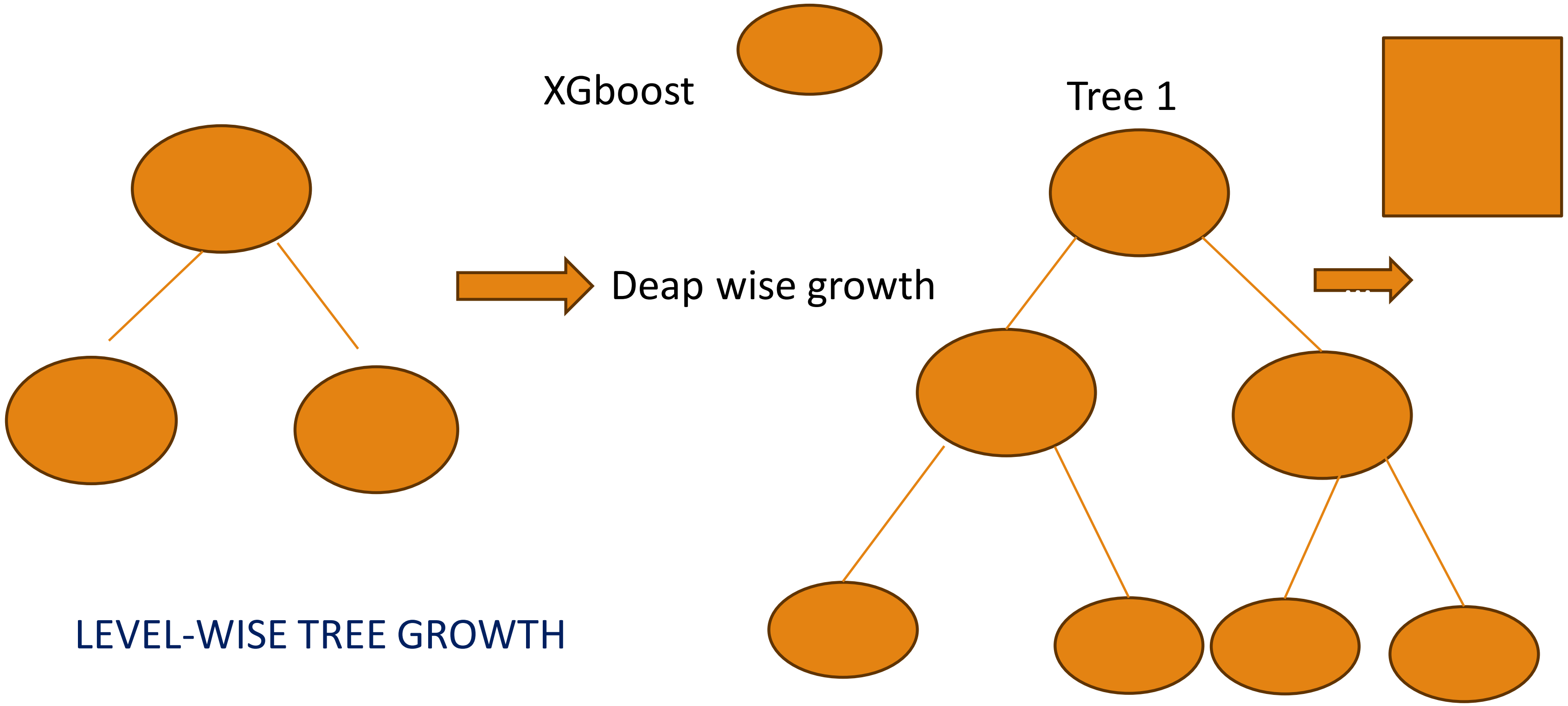
XG boosting algorithm

*XG Boost, short for eXtreme Gradient Boosting, is a powerful and popular machine learning algorithm used for both classification and regression tasks. It belongs to the ensemble learning category, where multiple weak learners (simple models) are combined to create a stronger, more accurate model.

Here's a simple explanation of XGBoost using a basic diagram:

1. Weak Learners (Base Models):
    1. XG Boost uses decision trees as its base or weak learners.
    2. A decision tree is a flowchart-like structure where each internal node represents a test on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label or a numerical value.
2. **Boosting Process**:
    1. XGBoost builds trees sequentially, and each new tree corrects errors made by the previous ones.
    2. The process starts with a simple tree, and subsequent trees are added to correct the errors or residuals of the existing ensemble.
3. **Gradient Boosting**:
    1. "Gradient" in XGBoost refers to the optimization method used to minimize the errors.
    2. The algorithm minimizes a loss function by adding new trees that correct the errors made by the existing ensemble.
4. **Regularization and Control Parameters**:
    1. XGBoost includes regularization terms in its objective function to control the complexity of the model and prevent overfitting.
    2. Control parameters allow you to customize the learning process, such as the learning rate

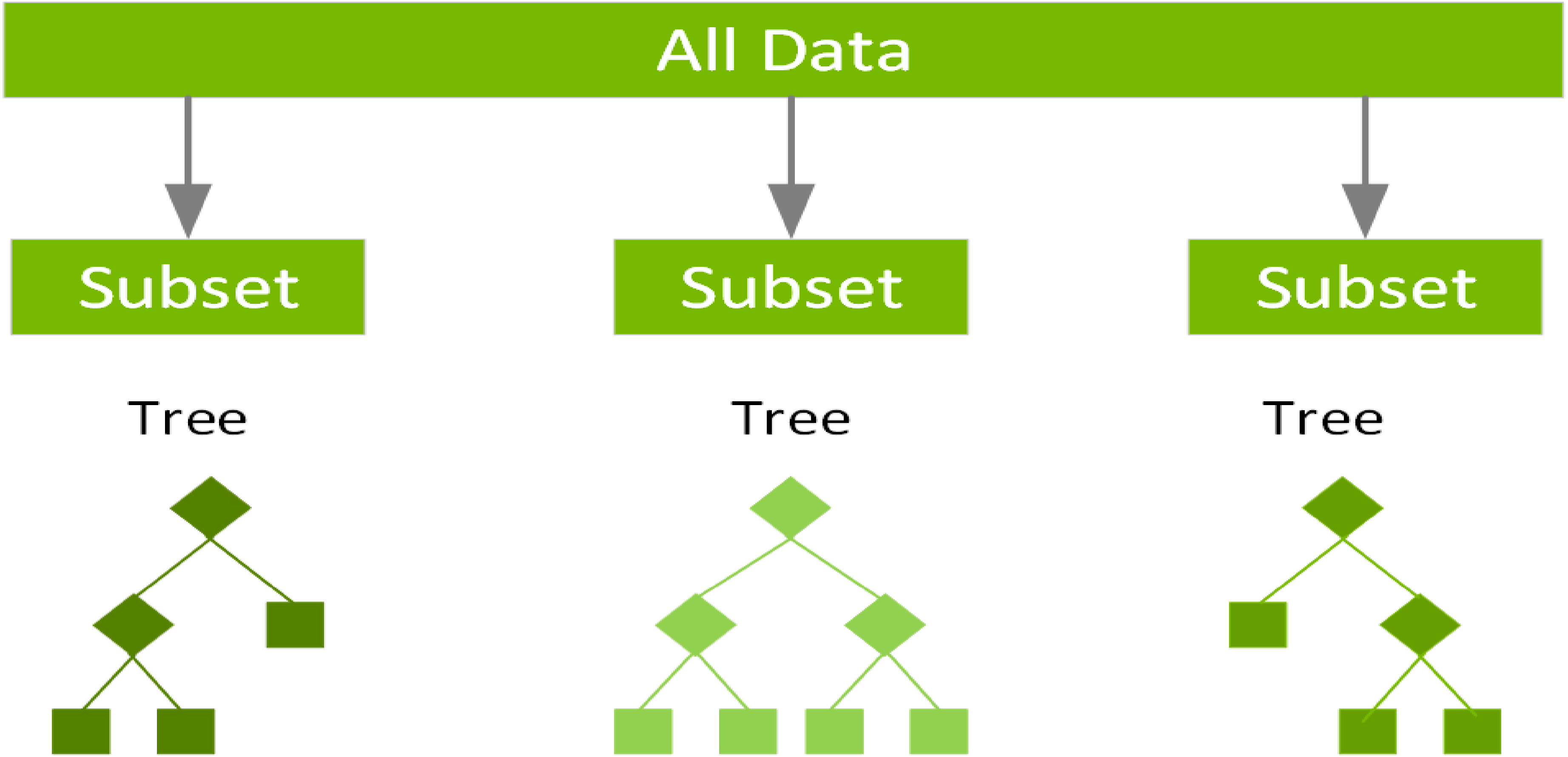
XGBoost is an ensemble learning algorithm that sequentially combines weak learners, typically decision trees, to construct a robust and highly accurate predictive model



Objective function=∑i=1nl(yi,y^i)+∑k=1k ω(fk)

•*N* is the number of training instances.
•*L*(*yi*,*yi*) is the loss function that measures the difference between the true label (*yi*) and the predicted value (*yi*).
•*K* is the number of trees in the ensemble.
•Ω(*fk*) is the regularization term for each tree (*fk*)

This flowchart illustrates the iterative nature of XG Boost, where each new model corrects the errors of the previous ones, leading to an ensemble that provides accurate predictions

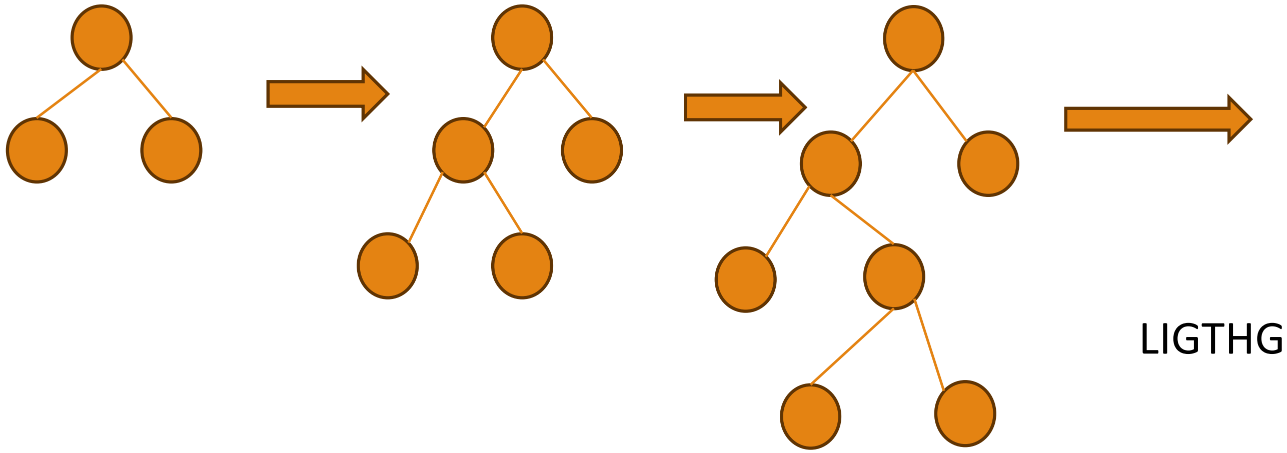


LIGHTGBM BOOSTING ALGORITHM

Light GBM, an efficient gradient boosting framework, excels in rapid model training and scalability, making it a preferred choice for large-scale machine learning tasks.

*Fast,distributed,high.performance gradient boosting framwork base on decision tree algorithm.

*It splits the tree leafwise.others do tree wise or level wise.



LIGTHGBM(Leaf-wise tree growth)

Can lead to over fiting which can be minimized by defining the depth for splitting
Many hyperparameter for splitting.