# Boosting → Decision Trees sequentially Connected

Boosting is a powerful machine learning technique used for improving the performance of weak learners (often referred to as base learners) to create a strong learner. The basic idea behind boosting is to combine multiple weak learners sequentially, with each new learner focusing on the mistakes made by the previous ones. This way, the ensemble gradually learns to correct its errors and improve overall accuracy.
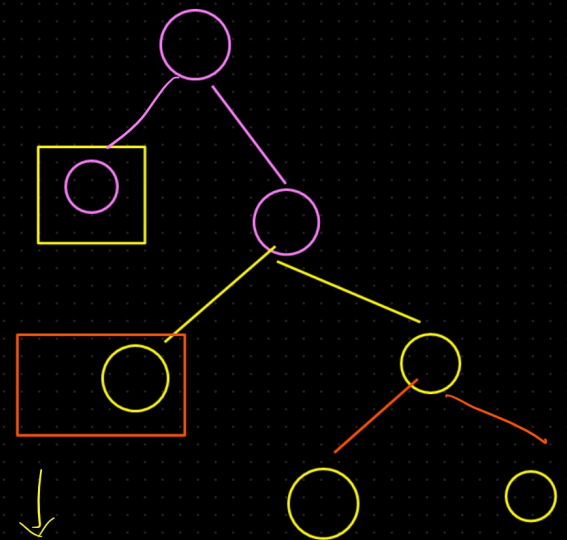
```
Dataset → DT 1 → DT 2 → DT 3 → - - - → DT n → ( Strong Learner )
```
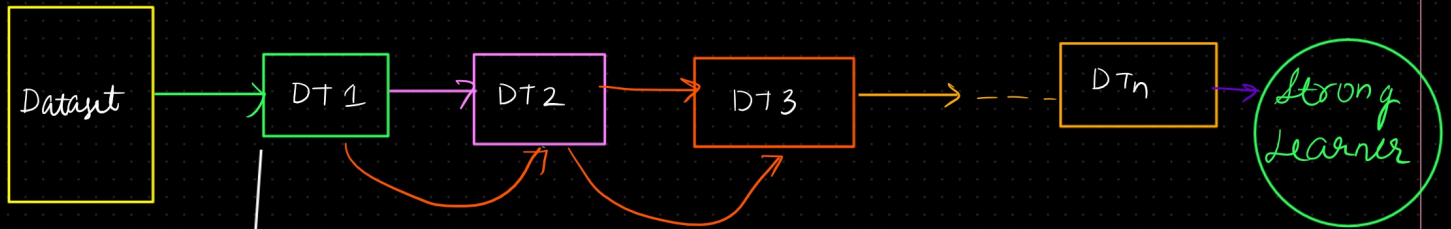
↳ Weak learner
↓

Haven't learnt much from the Training dataset

→ Assignment of the weight to the weak learner

$$f = \alpha_1(M_1) + \alpha_2(m_2) + \alpha_3(m_3) + - - - + \alpha_n(m_n)$$

$(\alpha_1, \alpha_2, \alpha_3 - - weight)$

↓

DT → Model Overfitting
↙

Training Acc ↑↑
Test Acc ↓↓

Low Bias
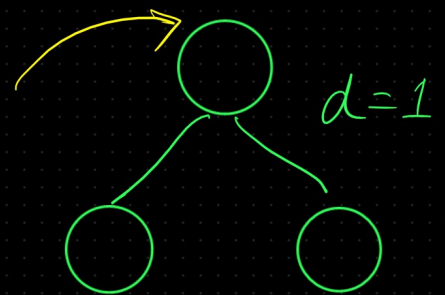High variance

⇓ Generalized Model
↓

Low Bias, Low Variance

Boosting Algorithm

① Adaboost (Adapative Boosting)

② Gradient Boosting

③ Xgboost

Adaboost

Step 1: Train a weak learner (D.T)

Decision stump

$d = 1$

Step 2: Calculate the error of the weak learner

Step 3: Update the weight of the data points based on the error

Step 4: Repeat steps 1-3 to train additional weak learner

Step 5: Combine the weak learners into a strong ensemble model.

This process leads to a final model that effectively combines the predictions of multiple weak learners, resulting in improved accuracy and generalization performance.