Ensemble Methods: Bagging and Boosting

After The session/s participants

- Understand the concept of Ensemble Methods
- Understand what Bagging and Boosting mean
- Understand how to implement Ensemble Methods in Python

Ensemble Methods

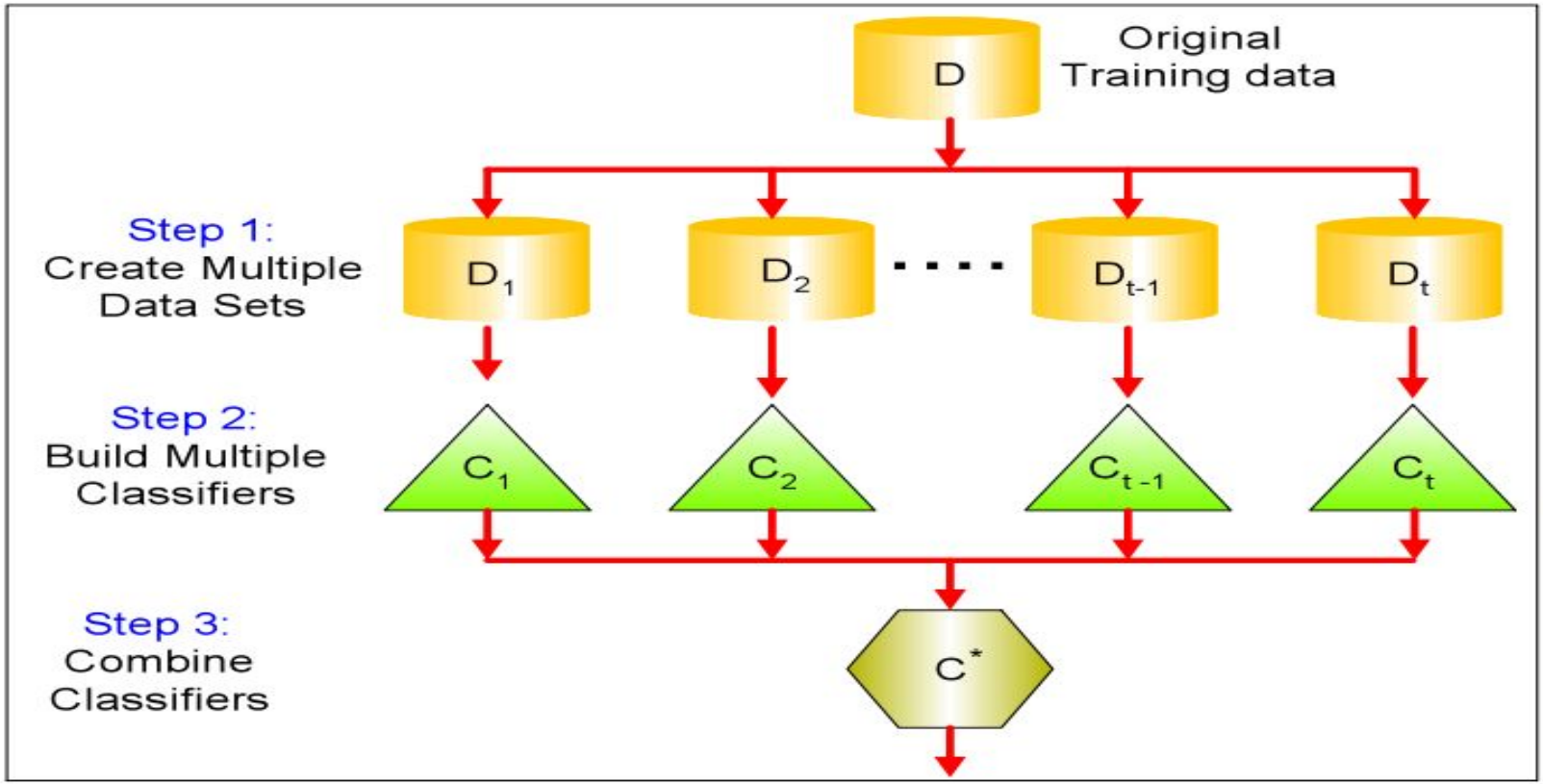
Main Idea : Make use of multiple models for making prediction instead of single model

Empirically tend to give better results than a single model.

You can have parallel ensemble methods and also sequential ensemble methods

Bagging : Bootstrap Aggregation

- Single Classifier is not the best of predictors
- Use Multiple Classifiers
- Committee of Experts : Very less likely that most experts are wrong
- Use the Wisdom of the Crowd!!!!



Slide Source: Lecture by David Sontag New York University

Bootstrapping is a sampling technique where samples are derived from the whole population (set) using the replacement procedure.

Aggregation in bagging is done to incorporate all possible outcomes of the prediction and randomize the outcome.

Bagging is advantageous since weak base learners are combined to form a single strong learner that is more stable than single learner

Reduces Variance

Computationally Expensive

Source: Ensemble Methods - Overview, Categories, Main Types (corporatefinanceinstitute.com)

Boosting

Learn from Previous Predictor’s mistake

Combine various base “ weak” learners to form a strong learner

Popular Boosting Techniques : AdaBoost and GradBoost (XGBoost)

AdaBoost

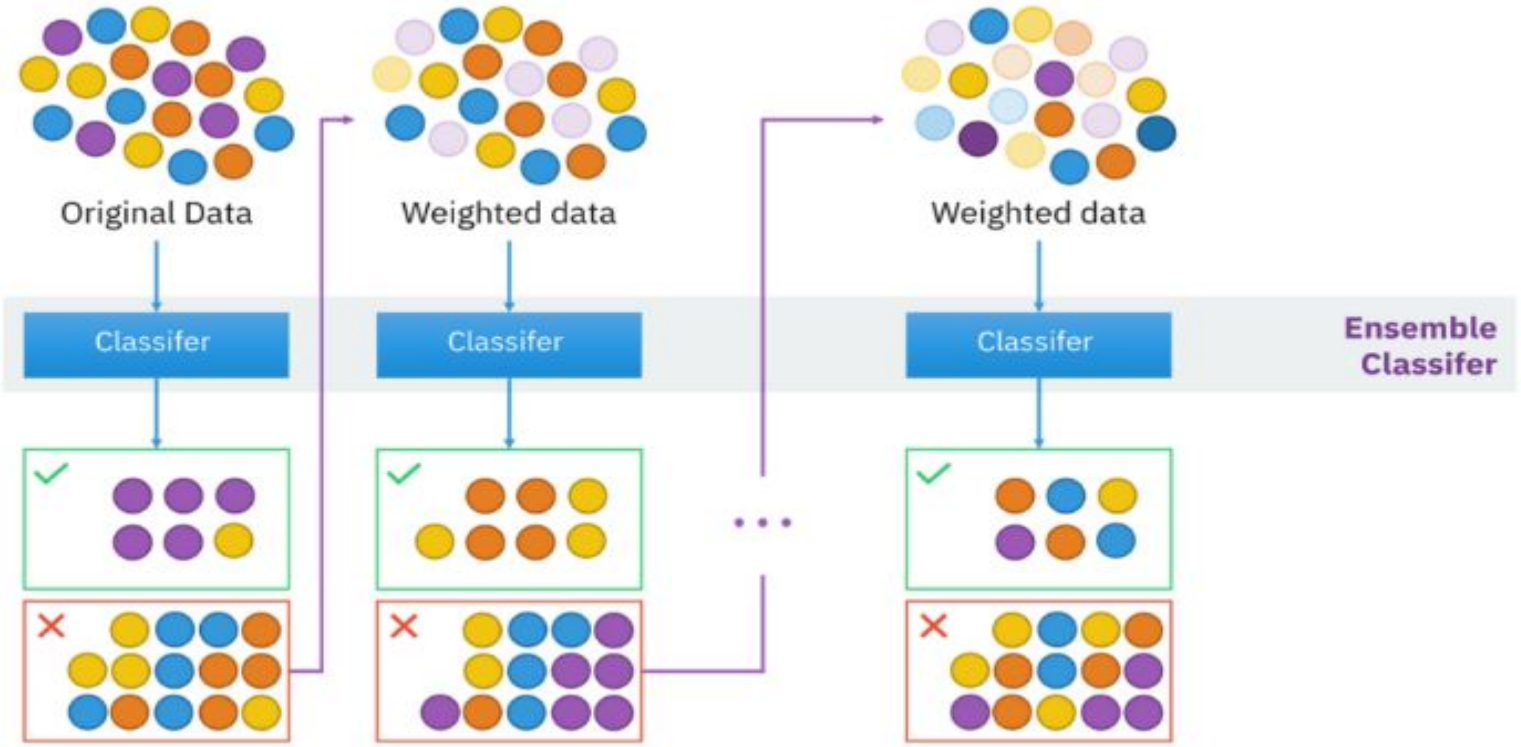


Image source: AdaBoost Algorithm in Machine Learning (almabetter.com)

Gradient Boosting

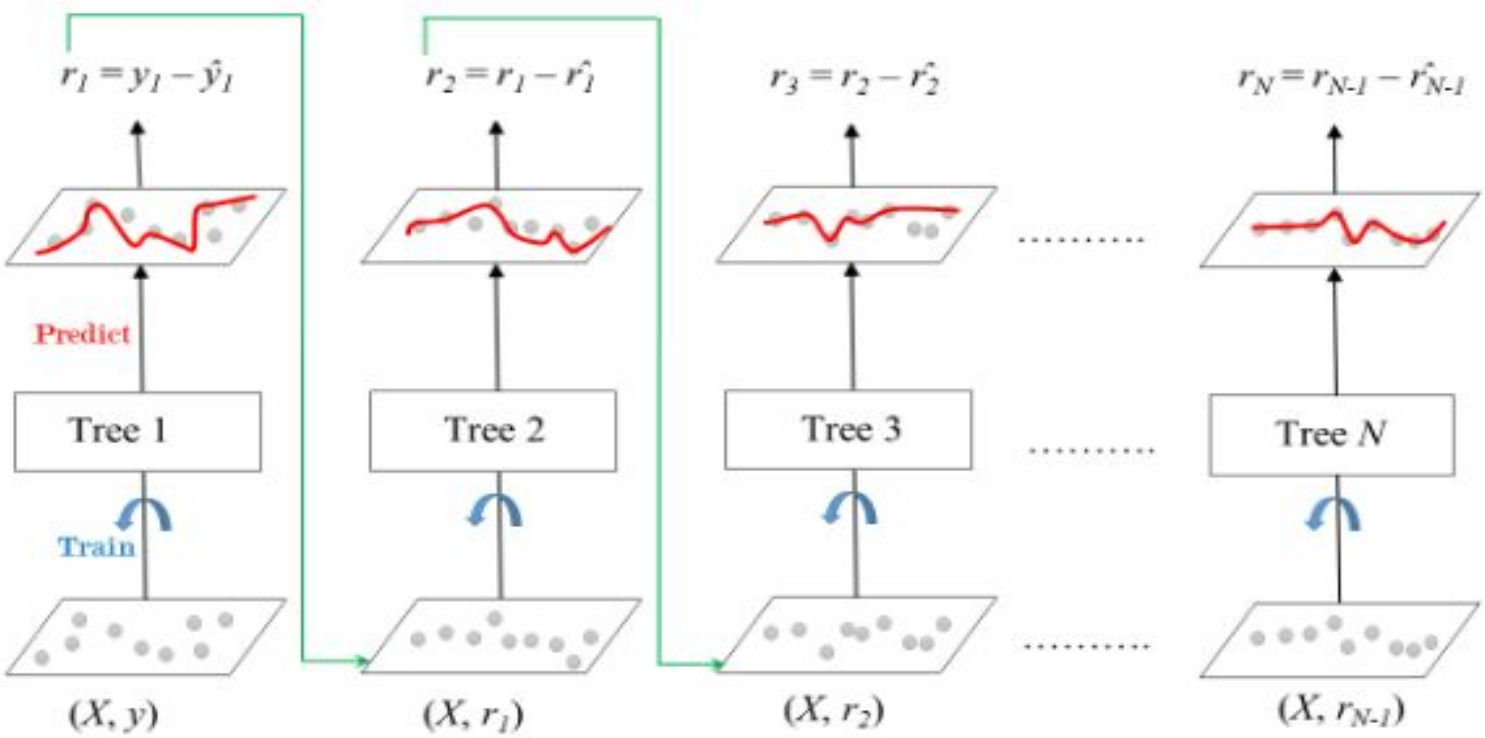


Image Source：[Gradient Boosting in ML - GeeksforGeeks]

Differences:

1. Adaboost changes weights of the points where as GradBoost uses residual errors of predecessors as labels

2. AdaBoost typically uses Decision stump as weak learner whereas GradBoost can use other variety of models (e.g. linear models)

3. AdaBoost is more susceptible to noise whereas GradBoost is more robust.