

# Gradient Boosting

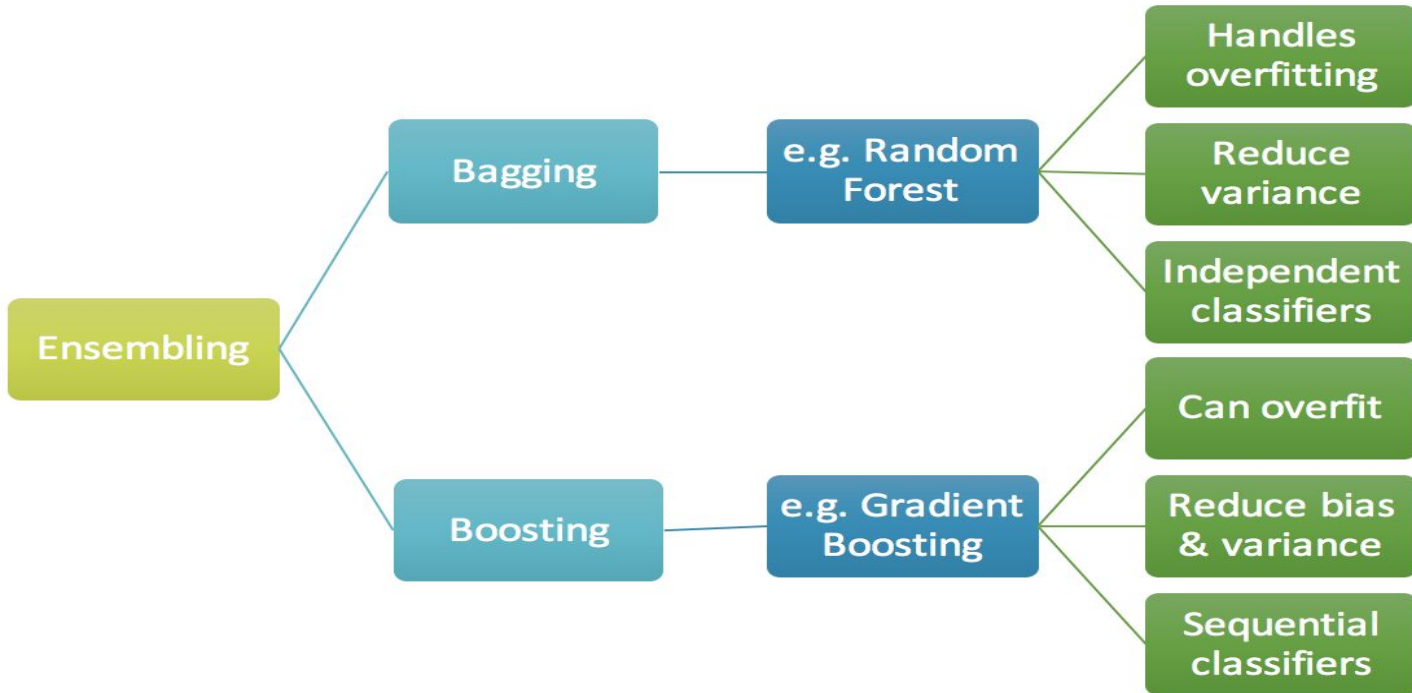
## Ensemble Technique:

- > It is a technique which uses multiple weak learners to produce a strong model for regression and classification.
- > In ensemble learning, many base models like classifiers and regressors are generated and combined together so that they give better results.

Bagging: Building multiple models (typically of the same type) from different subsamples of the training dataset.

Bagging works as follows:

- > Multiple subsets are created from the original dataset, selecting observations with replacement.
- > A base model (weak model) is created on each of these subsets.
- > The models run in parallel and are independent of each other.
- > The final predictions are determined by combining the predictions from all the models



# Gradient Boosting

-> Gradient Boosting is an supervised machine learning algorithm used for classification and regression problems.

-> Gradient Boosting is a popular boosting algorithm. In gradient boosting, each predictor corrects its predecessor's error.

-> It is an ensemble technique which uses multiple weak learners to produce a strong model for regression and classification. Here each predictor is trained using the residual errors of predecessor as labels.

-> Each tree predicts a label and final prediction is given by the formula.

## Steps:

1. Calculate the average/mean of the target variable.
2. calculate the residuals for each sample.

$\text{Residual} = \text{Actual Value} - \text{Predicted Value}$

3. use decision tree algorithm to train the model considering residual as label . We build a tree with the goal of predicting the Residuals.
4. Repeat steps 3 to 5 until the number of iterations matches the number specified by the hyper parameter(numbers of estimators)
5. Once trained, use all of the trees in the ensemble to make a final prediction as to value of the target variable.

The final prediction will be equal to the mean we computed in Step 1 plus all the residuals predicted by the trees that make up the forest multiplied by the learning rate.

final prediction = Average Price + LR\*Residual predicted by DT1 + LR\*Residual Predicted by DT2 + .....LR\*Residual Predicted by DT N

Here:

LR = Learning rate

DT = Decision tree