

Start-Tech Academy

Ensemble Methods

Types

- 1. Bagging
- 2. Random Forest
- 3. Boosting

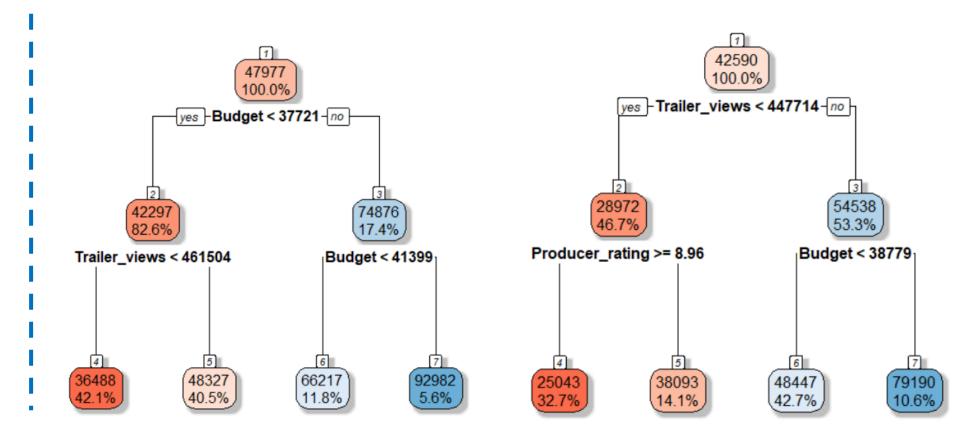
Problem with normal decision tree

• High Variance



Ensemble Methods

Example





Bagging

Methods

Concept:-

If N observations have variance sigma sq. (s^2), then variance of mean of these observations is (s^2)/N

Training Data 1

Training Data 2

Training Data 3

Training Data 4

Model 1

Model 2

Model 3

Model 4

Prediction 1

Prediction 2

Prediction 3

Prediction 4



Final Model



Bagging

Bootstrapping

7	9	5	4	3	
Sample 1 -	9	5	4	3	4
Sample 2 -	7	9	5	4	7
Sample 3 -	7	9	9	4	3



Bagging

Methods

- 1. While bagging pruning is not done, Full length trees are grown
- 2. Individual trees have high variance and low bias, averaging reduces the variance
- 3. In regression, we take the average of predicted values
- 4. In Classification, we take majority vote i.e. most predicted class will be taken as the final prediction

