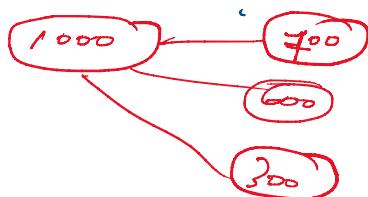


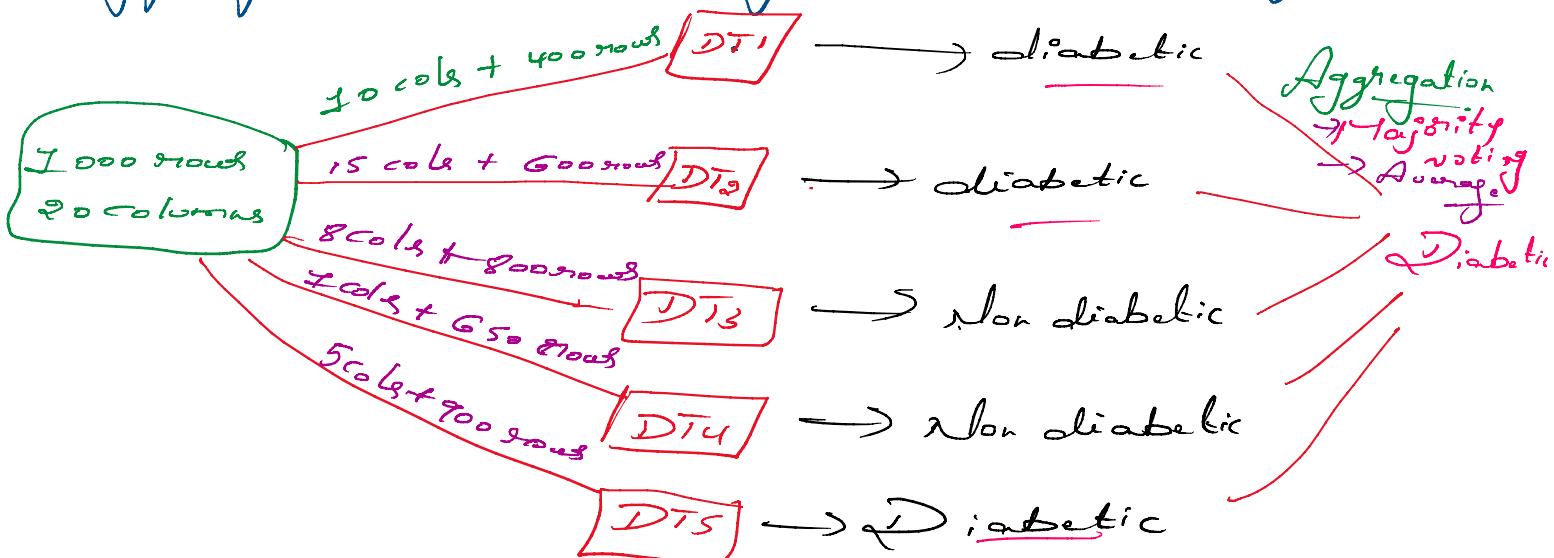
Random Forest

- * It comes under Supervised Learning.
- * It is used to solve classification and regression.
- * Random Forest constructs multiple trees using "Bootstrap aggregation"

Bootstrap → row sampling + column sampling with replacement



Aggregation → Combining all the outputs together.



5 Estimators

- * Random Forest build multiple trees parallelly using Bootstrap Aggregation.

- * Random Forest build multiple trees using Bootstrap Aggregation.
- * Aggregation
 - Majority voting [Classification]
 - Average [Regression]
- * If it is used to overcome the overfitting concept in Decision Tree.

Hyperparameters in Random Forest

No of estimators → No of trees should be odd.

Bootstrap → [True, False]

- Sampling without replacement.
- Sampling with replacement

Hyperparameters in Random Forest

- ① n_estimators → No of trees
- ② max_features → No of features / Columns selected to construct trees
 - [auto, sqrt]
 - auto → It will randomly select features to construct trees
 - sqrt → It will select square root of features
 $\sqrt{16} = 4$
 $\sqrt{25} = 5$
 $\sqrt{10} = 3.16$
- ③ Bootstrap → Sampling with or without replacement.

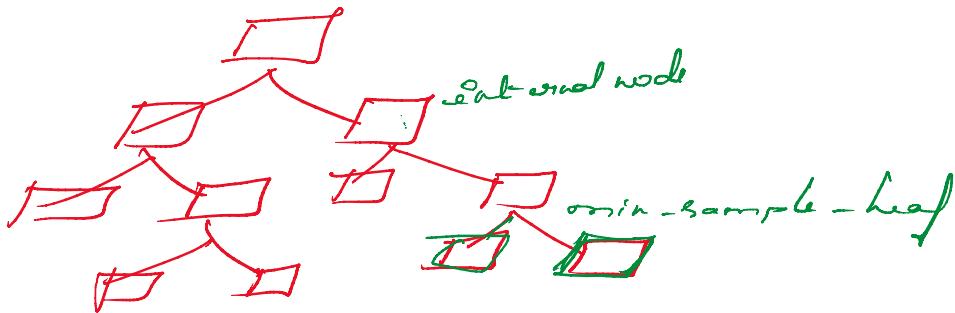
Bootstrap = [True, False]

Bootstrap = True → Sampling with replacement

Bootstrap = False → Sampling without replacement

- ④ max_depth → height / depth of tree
- ⑤ min_sample_split → minimum no of samples required at internal node split
- ⑥ min_sample_leaf,

- (5) min-sample-split → minimum no of samples required at internal node split
- (6) min-sample-leaf → minimum no of samples required at leaf node?



Hyperparameter tuning technique

→ Grid Search CV

Grid Search CV will train and evaluate the model for all the Combination

→ model, parameters, scoring, verbose, n_jobs, cv, X, Y

verbose=0 → no messages

verbose=1 & 2 → print the messages

n_jobs=1 → It will use only one processor

n_jobs=-1 → It will use all the processor

→ Randomized Search CV Break 9.45 to 9.55

Hyperparameter tuning technique which train and test model with specified number of Combinations

n_iter = no of Combinations to be chosen

Given n_iter=100 → It will train and test model with 100 random Combination