**Ensemble Learning –** It says do not depend on a single model for the result, instead of this try using multiple model. It might use different Algorithm as well as different Training data set. It has heterogeneous learning. Then combining the all sub result of each model resulting in the new model. This new model have high accuracy and predictive power with less error rate.

**Random Forest –** Similar to the trees in the forest, It has many decision trees (known as base learner) of various structures. It is a kind of ensemble classifier which uses decision tress algorithm in a randomized way. It requires the training data. It selects some rows and some features and creates some samples. There is a possibility it might miss some records while processing or training as records are picked randomly. In the end the class with majority of vote is predicted. Random Forest also correct for high variance (overfitting of data). It also tells the important feature of the current data. It can also detects non-linear relationships. Random Forest are comparatively stable and doesn’t change much. It can also deal with missing data

**Randomized Search CV –** It is used select the best decision tree among the provided decision tree to solve the particular problem. GridSearchCV takes more time than the RandomizedSearchCV.

**pandas.get\_dummies()** is used for data manipulation. It converts categorical data into dummy or indicator variables.