Week4 Homework

1. Hand tune every supervised machine learning algorithm implemented in "Uber_Rider_Churn_Supervised_Learning.ipynb", and
    a. Record the effect of every hyperparameter on bias and variance, following the format below:

Random Forest:

hyperparameter name	effect on bias (if increase)	effect on var (if increase)
n_estimators		
max_features		
max_depth		
min_samples_split		
min_samples_leaf		

    b. Hand tune every supervised machine learning algorithm to the best possible performance (AUC).
    c. Implement grid search code for every algorithm, and explore a few variations of hyperparameter sets. Compare grid search results with hand tuned results.
    d. Rank the algorithms based on best possible performance (best among hand tuned or grid search, results should be table or bar chart). Discuss why the ranking you have is reasonable based on the pros and cons of each algorithm.
2. What percentage of unique records in a bootstrap sample compared with original data as the number of records N -> infinity? Can you derive it?