## Assignment 2 - Question 2

2) Briefly describe how gradient boosting differs from bagging.

Definitions -

Bagging (stands for Bootstrap Aggregating) is a way to decrease the variance of your prediction by generating additional data for training from your original dataset using combinations with repetitions to produce multisets of the same cardinality/size as your original data. By increasing the size of the training data you can't improve the model predictive force, but just decrease the variance, narrowly tuning the prediction to expected outcome.

Boosting is a two-step approach, where one first uses subsets of the original data to produce a series of averagely performing models and then "boosts" their performance by combining them together using a particular cost function (=majority vote). Unlike bagging, in the classical boosting the subset creation is not random and depends upon the performance of the previous models: every new subset contains the elements that were (likely to be) misclassified by previous models.

Differences -

Boosting has higher variance than Bagging because of over fitting.

Boosting could generate a combined model with lower errors as it optimizes the advantages and reduces pitfalls of the single model.

Boosting determines weights for the training data to tip the scales in favor of the most difficult cases.

Source: Quora