CECS 451
Assignment 9
Total: 35 Points

General Instruction

- I recommend you can write your answer using LaTeX.
- Submit uncompressed file(s) in the Dropbox folder via BeachBoard (Not email).

---

1. Using `scikit learn`, evaluate the classification accuracy of the decision tree, bagging, AdaBoost, and Random forest.
   (a) Find the `Assn9.py` and use the 'Gini' index as the criterion.
   (b) (5 points) Complete the method `decision_tree` that generate a decision tree from `X_train, y_train` and predict `y` from `X_test`. This method should record its prediction accuracy at `tree_score`.
   (c) (10 points) Similarly, complete the method `bagging` that generate multiple decision trees using the bagging. This method should record its prediction accuracy at `bagging_score` by varying the parameter `n_estimators`. Draw a chart whose X-axis is `n_estimators` and Y-axis `bagging_score`, and the chart should have more than 20 data points of different X-axis values.
   (d) (5 points) Similarly, complete the method `boost` that generate multiple decision trees using the AdaBoost. Draw a chart whose X-axis is `n_estimators` and Y-axis `boost_score`, and the chart should have more than 20 data points of different X-axis values.
   (e) (15 points) Similarly, complete the method `forest` that generate multiple decision trees using the random forest. Draw a chart whose X-axis is `n_estimators`, Y-axis `max_features`, and Z-axis `forest_score`. The chart should have more than 100 data points of different pair of X-axis and Y-axis values.
   (f) Submit your `Assn9.py` and a report that includes the charts.