MAS DSE 220

Assignment #3

1. Train decision tree, forests of randomized trees and Boosting trained on the Titanic Data Set. Perform and plot the decision surface for each. Don't forget to perform data prep. (Hint: Kaggle has a nice description- https://www.kaggle.com/c/titanic).
   Perform feature importance analysis and plot histogram before the training. Several different configuration of each of the models/parameters should be explored, analyzed and plotted. Demonstrate how changes in parameters influences accuracy for different algorithms. Describe your process of parameter tuning and provide in detailed discussion of the results.

2. Perform Stacking Ensemble analysis on the Boston Housing Data set. You can use Python, Weka, KNIME or other tool to perform Stacking or write your own version in Python code. Include interesting plots and attribute importance analysis to support the choice of the final model configuration chosen. Any combination of any of the Machine learning algorithms we have covered in the class so far or you are already familiar with is acceptable to be used in the Stacking Ensemble.