Middle East Technical University
Electrical-Electronics Engineering Department

EE 583 Pattern Recognition
Homework 5

Using MATLAB, attempt the questions below:

1) Execute the **Grow a Classification Tree** example at https://www.mathworks.com/help/stats/classificationtree-class.html *step-by-step* via "Try This Example" option in your internet browser (or you may execute it in your local MATLAB software by using "View MATLAB Command"). Apply this algorithm to `fisheriris` data set, for its feature pairs (3,4). View and plot this tree. Set the maximum number of splits at 7 and grow/plot the tree once again. Compare the cross-validation classification errors of these two models by cross-validating the model by using 10-fold cross-validation. Repeat the test for Entropy/Deviance error for SplitCriterion.

2) Execute the **Train Boosted Classification Ensemble** example at https://www.mathworks.com/help/stats/classreg.learning.classif.classificationensemble-class.html *step-by-step* via "Try This Example" option in your internet browser (or you may execute it in your local MATLAB software by using "View MATLAB Command"). Apply this algorithm to `fisheriris` data set, for its feature pairs (3,4). Train a boosted ensemble of 25 classification trees using half of the measurements and AdaBoostM2 (i.e. multi-class Adaptive Boosting) method. Plot the first trained tree. Using the other half of the data, compare the classification performance of AdaBoost and the first tree. Comment on the results by varying LearnRate parameter.

3) Examine **TreeBagger** function at https://www.mathworks.com/help/stats/treebagger.html. Apply this algorithm to `fisheriris` data set, for its feature pairs (3,4) for combining the classification of 25 trees. Sample with replacement half of the available data during bootstrapping. Plot one of the trained trees. Using the other half of the data, compare the classification performance of AdaBoost and the first tree.

4) Execute **Jackknife Resampling** example at https://www.mathworks.com/help/stats/resampling-statistics.html#brjjmlm *step-by-step* via "Try This Example" option in your internet browser (or you may execute it in your local MATLAB software by using "View MATLAB Command"). Measure the uncertainty in estimating the correlation coefficient. First plot lawdata and calculate its sample correlation using corr function. Determine the jackknife estimate of correlation coefficient and find its average value. Determine the Jackknife estimate of the bias between sample correlation and true correlation. Repeat this step for median values of both lsat and gpa features rather than correlation coefficient.