

I 526/B659 Programming Assignment 2 - Due Friday November 4, 2016

Implement Adaboost and Bagging based on your prior decision tree code. Use the code developed in Programming Assignment 1 and tailor it to learn Adaboost and Bagging.

For Adaboost, try two different depths 1 and 2 for the trees and try 5 and 10 trees. For bagging, use a depth of 3 and 5 and try with 5 and 10 tree bags.

In particular, the input to your algorithm will include the training data set, the maximum depth of the tree and the number of trees. For example, if the depth is set to one, you will learn a decision tree with one test node, which is also called a decision stump. Test your implementation, with depth=1, and 2 respectively, on the following data set as described below (train on the training data and test with the testing data set).

Data set information: This data set is extracted from the UCI Mushroom data set (train and test) . Assume that there are 21 features (the first 21 columns) and the class labels are in the 4th column (bruises). There are 2 classes. Please refer to <https://archive.ics.uci.edu/ml/datasets/Mushroom> for details on the data set.

1. For bagging, try two depths 3 and 5 and two sets of bags 5 and 10
2. For Adaboost, try two depths 1 and 2. Try two bags 5 and 10 trees.
3. Report the confusion matrix for these four settings.
4. Now, use Weka's default Adaboost and bagging and present their results.

The total points on this homework is 20.