

Homework4. Problem 1 Random Forests

Jae Dong Hwang

Build Random Forest Model

Build a model with numTrees = 10, minSplit 2, use Bagging and FeatureRestriction=20. Create a table that has accuracies of the 10 individual trees, along with the accuracy of the full random forest (after the individual trees vote) on xTest.

Trees	Accuracies
Full	0.8629842180774749
Tree 0	0.8500717360114778
Tree 1	0.8436154949784792
Tree 2	0.8436154949784792
Tree 3	0.8443328550932568
Tree 4	0.8464849354375896
Tree 5	0.8493543758967002
Tree 6	0.8314203730272597
Tree 7	0.8393113342898135
Tree 8	0.8565279770444764
Tree 9	0.8651362984218077

Use Bagging(bootstrap): True Feature Restriction: 20 MinToSplit: 2 Seed for random: 10

Run parameter sweeps for numTrees in: [1, 20, 40, 60, 80], for each of these settings:

- Config 0 - minSplit = 2, Bagging, and FeatureRestriction = 20
- Config 1 - minSplit of 50, Bagging, and FeatureRestriction = 20
- Config 2 - minSplit = 2, NO Bagging, and FeatureRestriction = 20
- Config 3 - minSplit = 2, Bagging, and NO FeatureRestriction = 0

Produce a plot with numTrees on x-axis and the hold-out set accuracy for each of these variations on the y-axis (that means 4 lines).

