

MACHINE LEARNING CBC

ASSIGNMENT TWO

Xin Yan Kwek, Darcy Qiu, Charles Koh, and C.S. Rudolf Lai

Pruning

The `pruning_example` function executes three MATLAB tree functions in this sequence: `treefit`, `treetest`, and `treeprune`. First it uses `treefit` to create a decision tree for predicting responses (targets) as a function of predictors (examples). The tree is created such that impure nodes will be split as long as there are observations. Using `treetest`, the expectation of misclassification costs over all terminal nodes is then computed using both the resubstitution method and the 10-fold cross-validation method. Along with the cost, the standard error of each cost value, the terminal node count for each subtree, and an *estimated best level of pruning* is also computed. Using the two estimated best level, we then use `treeprune` to prune the tree, coming up with two new trees.