

Assignment #4: Tree models

Textbook Chapter 8- Exercise Question #8

In the lab, a classification tree was applied to the **Carseats** data set after converting **Sales** into a qualitative response variable. Now we will seek to predict **Sales** using regression trees and related approaches, treating the response as a quantitative variable.

- (a) Split the data set into a training set and a test set.
- (b) Fit a regression tree to the training set. Plot the tree, and interpret the results. What test error rate do you obtain?
- (c) Use cross-validation in order to determine the optimal level of tree complexity. Does pruning the tree improve the test error rate?
- (d) Use the bagging approach in order to analyze this data. What test error rate do you obtain? Use the **importance ()** function to determine which variables are most important.
- (e) Use random forests to analyze this data. What test error rate do you obtain? Use the **importance ()** function to determine which variables are most important. Describe the effect of m , the number of variables considered at each split, on the error rate obtained.

Submit through link: eCampus -> Assignments

Deadline: April 7, Thursday @11:59pm