

ch08_notes

August 10, 2019

- Table of Contents
- 8 Tree-Based Methods
 - 8.1 The Basics of Decision Trees
 - 8.1.1 Regression Trees
 - 8.1.2 Classification Trees
 - 8.1.3 Trees Versus Linear Models
 - 8.1.4 Advantages and Disadvantages of Trees
 - 8.2 Bagging, Random Forests, Boosting
 - 8.2.1 Bagging
 - 8.2.2 Random Forests
 - 8.2.3 Boosting
 - 8.3 Footnotes
 - 8.3.1 blah

1 Tree-Based Methods

1.1 The Basics of Decision Trees

1.1.1 Regression Trees

Overview

- There are two main steps:
 - Partition predictor space \mathbb{R}^p into regions R_1, \dots, R_M .
 - For all $X = (X_1, \dots, X_p) \in R_m$, predict the average over the responses in R_m

$$\hat{f}(X) = \hat{y}_{R_m} := \frac{1}{N_m} \sum_{i: y_i \in R_m} y_i$$

where $N_m = |\{y_i \mid y_i \in R_m\}|$

- In practice, we take the regions of the partition to be rectangular for simplicity and ease of interpretation.