

Saarland University

The Elements of Statistical Learning

Assignment 6

Due Date: 17.01.2018

Thibault SCHOWING Mat. 2571837

Sarah MCLEOD Mat. 2566398

January 12, 2018

Problem 1

Question How do you have to change the procedure of generating a MARS model to make a decision tree?

Answer Where MARS will create different linear estimation between its knots, Decision tree will take the average of the output values. In order to make a decision tree out of the MARS model, instead of using a function between two knots, the average of the output (between the knots) can be returned.

Question Can you argue on the basis of the relationship between MARS and decision trees revealed in (a) what is an advantage of MARS over decision trees and what is an advantage of decision trees over MARS?

Answer For numeric data, MARS would tend to be better than Decision Trees because the reflected pairs (hinges) would adapt more accurately simply to the underlying structure of the data (for instance local linearity) than the constant segmentation realised by decision trees. Think about a diagonal set of points that would have to be separated in many sub-trees, with MARS a single line and two knots can estimate it accurately.

The decision trees are more faster to create and, especially for categorical data, the fits of partitioning would be better. Even if MARS adapts to non linear data, depending on the exact shape of the data, a decision tree could separate more accurately the different regions.

Problem 2

Problem 3

Problem 4