STATS 202: Data Mining and Analysis
Instructor: Linh Tran

HOMEWORK # 3
Due date: August 6, 2021

Stanford University

Introduction

Homework problems are selected from the course textbook: *An Introduction to Statistical Learning*.

Problem 1 (7 points)

Chapter 6, Exercise 3 (p. 260).

Problem 2 (7 points)

Chapter 6, Exercise 4 (p. 260).

Problem 3 (7 points)

Chapter 6, Exercise 9 (p. 263). Don't do parts (e), (f), and (g).

Problem 4 (7 points)

Chapter 7, Exercise 1 (p. 297).

Problem 5 (7 points)

Chapter 7, Exercise 8 (p. 299). Find at least one non-linear estimate which does better than linear regression, and justify this using a t-test or by showing an improvement in the cross-validation error with respect to a linear model. You must also produce a plot of the predictor X vs. the non-linear estimate $\hat{f}(X)$.

Problem 6 (7 points)

Chapter 9, Exercise 1 (p. 368).

Problem 7 (8 points)

Chapter 9, Exercise 8 (p. 371).