

Homework 2
Statistical Learning, Spring Term 2019
STAT760

The website for the Statistical Learning book has a repository of data:

<https://web.stanford.edu/~hastie/ElemStatLearn/>

Under the “Data” button we can find several datasets. The prostate cancer data set has eight predictors (columns 1--8)

lcavol
lweight
age
lbph
svi
lcp
gleason
pgg45

and one outcome (column 9), the lpsa measurement.

Problem 1

- a) Train a multivariate regression predictor for lpsa using from one to eight predictors. Graph the training error for the eight possible regressions. Use all the data, without dividing into train & test sets.
(15 points)

Program the classifier using Matlab, R, Python, or any other language of your choice. Do not use a library function for regression.

Problem 2 (5 points)

Solve the problems 3 and 4 from Chapter 3 of the ISL book (Section 3.7).

<http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Seventh%20Printing.pdf>