

Homework 1
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 “Zip code” dataset has a database of handwritten digits. The button “Info” describes the format. Each digit is represented by 256 gray values. There is a training set and a test set that you can use for Problem 1.

Problem 1

- a) Train a nearest neighbor classifier for handwritten digits with the training set provided, and test it using the test set.
(15 points)

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

Problem 2 (5 points)

Solve the problems 1,2,3 from Chapter 2 of the ISL book.

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