

Homework problem 2

Problem ESL 2.7 Compare classification performance of k-NN and linear regression on the `zipcode`¹ data, on the task of separating the digits 2 and 3. Use $k \in \{1, 3, 5, 7, 15\}$. Plot training and test error for k-NN choices and linear regression. Comment on the shape of the training and test error graphs of k-NN, and the training and test errors of linear regression.

¹Training: <http://statweb.stanford.edu/~hastie/ElemStatLearn/datasets/zip.train.gz>
Testing: <http://statweb.stanford.edu/~hastie/ElemStatLearn/datasets/zip.test.gz>
Info: <http://statweb.stanford.edu/~hastie/ElemStatLearn/datasets/zip.info>