1. Compare the classification performance of linear regression and k–nearest neighbor classification on the zipcode data. In particular, consider only the 2’s and 3’s, and k = 1, 3, 5, 7 and 15. Show both the training and test error for each choice. The zipcode data are available from the book website www-stat.stanford.edu/ElemStatLearn.

Linear Regression:

For linear regression R2 is 90.05% which means model can explain 90.05% variation in the training data. The RMSE error for training data is 0.1575174 and for test data is 0.3894424.

K nearest neighbor:

Following are the Error rates for different values of K;

