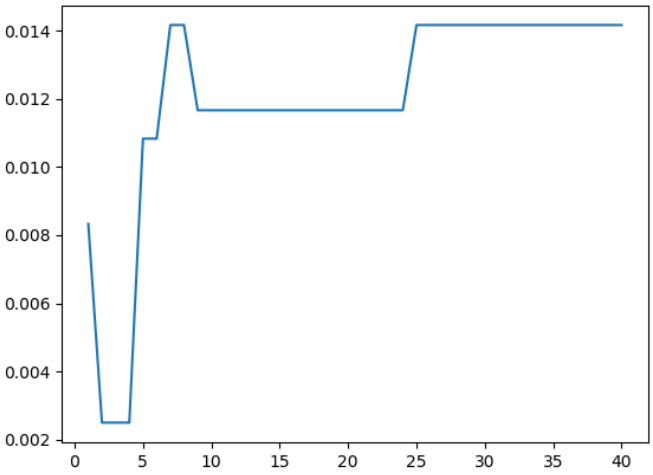
Homework 11

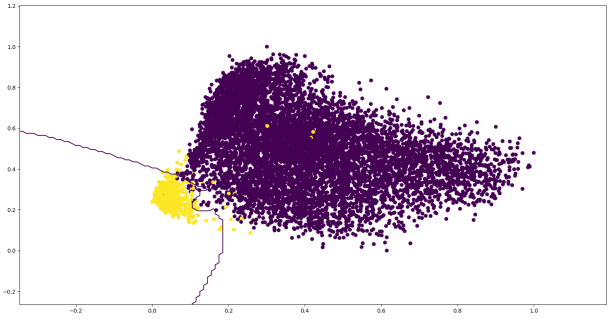
Karan Sarkar
sarkak2@rpi.edu

November 26, 2019

1. (a) We chose $k = 3$.

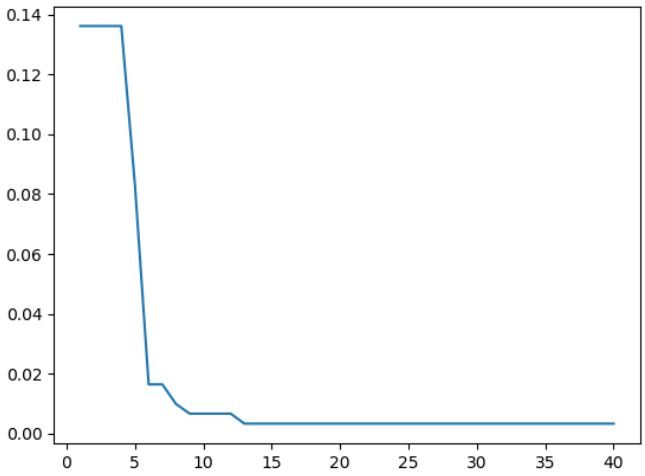


(b) For $k = 3$, we got in-sample error of 0.003323 and cross validation error of 0.00344.

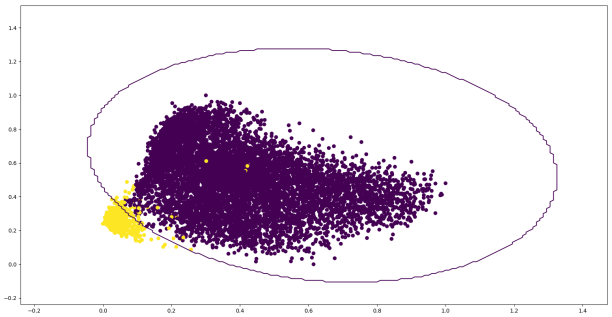


(c) We got a test error of 0.00644.

2. (a) We chose $k = 10$.



(b) For $k = 10$, we got in-sample error of 0.00332 and cross validation error of 0.00333.



(c) We got a test error of 0.00722.

3. The linear model with polynomial features got a test error of 0.03. It performed poorly compared to the KNN and RBF methods. I believe the RBF method is superior because it provides optimal results while retaining scalability. The KNN algorithm is likely not scalable to large amounts of data. An increase in complexity for the RBF is just linear while increasing the complexity of the polynomial transform grows rapidly.