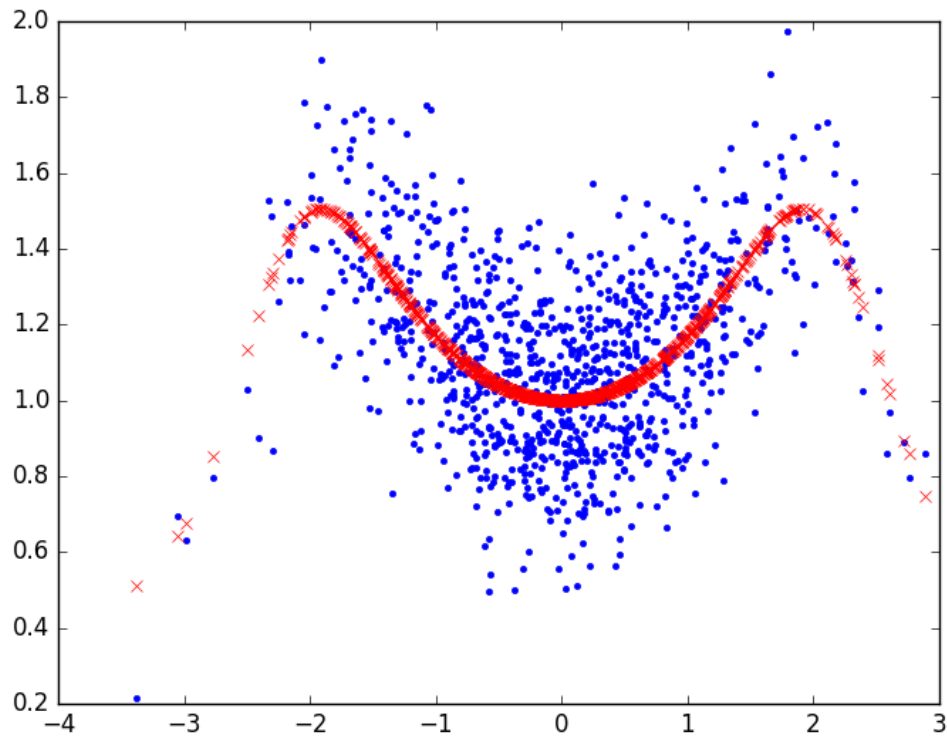Homework 4

Philip Hasse

11/1/16

Solution to Problem (1a):

I implemented the new generated dataset and changed the amount of generated points to 1000.

Solution to Problem (1b):

After running a 10-fold CV, the top 3 k-neighbors values changes every time depending on the dataset generated; however for the most recent values were 1, 277, and 283.

Solution to Problem (1c):

The best value for CV E_{out} in my sample was 0.2985. However, this value is subject to changed based on the dataset generated.