

Kenneth Kitahata & Spencer Tu

February 19th, 2020

STAT 339

HW1a Write Up

All of our functions from 1a. – 1b. work. However, when we try to run cross-validation function using 10 folds on the S1 and S2 datasets for odd values of K, our function crash within the for loop that we are writing and unfortunately, we were not able to fix the bug so there is no graph attached. Please take a look at our source code attached.

1 (f): Comment on how the complexity of the dataset affects the optimal choice of K.

We think the more complex the data is, the higher the optimal choice. When there are more labels to choose from, it is important to look at more neighbors before giving the new data a label.