

Recitation Week 14 Problems

Pick any dataset you are familiar with (except for the Iris and Diamonds data) and build a simple KNN model. This should include the following steps (code for these can be found in the notes from class):

- 1) Pick an outcome variable (what do you want to predict?) and a set of characteristics you want to predict based on.
- 2) Split your dataset into test and train. Standardize your variables. Bonus: why is it important to standardize variables?
- 3) Run the KNN and examine your predictions. Make a plot of your predictions - give a geometric interpretation of KNN algorithm.
- 4) Examine / interpret the confusion matrix, and calculate the accuracy.