MIT 801 – Statistical learning

Assignment 2

Due date: as indicated on ClickUp

Question 1:

Give a detailed discussion of hierarchical clustering. Consider both single and complete linkage.

Marks: [10]

Question 2:

Consider the paper "Robust clustering using a k-NN mode seeking ensemble". This paper consider clustering using mean shift mode seeking and k-NN mode seeking. For the purposes of this assignment, ***we only focus on k-NN mode seeking excluding the ensemble approach***.

1. Give a detailed discussion of
   a. The concept of mode seeking in a cluster setup
   b. K-NN based density estimation
   c. The k-NN mode seeking algorithm
2. Consider the datasets mode1.xlsx and mode2.xlsx. Mode1 considers a univariate clustering case whereas mode2 considers a bivariate case. For each of the datasets, use k-NN mode seeking based clustering to do an appropriate clustering of the observations. That is
   a. Give the estimated k-NN density estimate
   b. Give the k-NN based clustering solution
   c. Interpret the results

Marks: [30]

Total: [40]