

CS251 HW 8 | Mon Apr 15, 2019 | Week 11

Name:

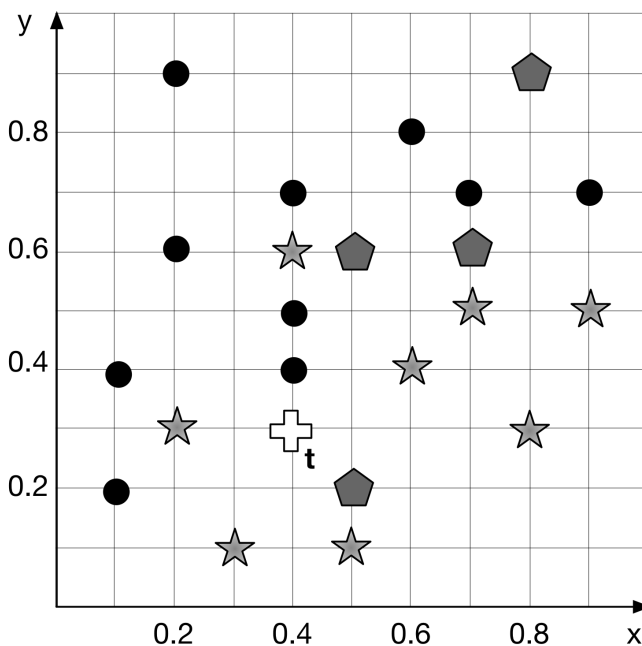
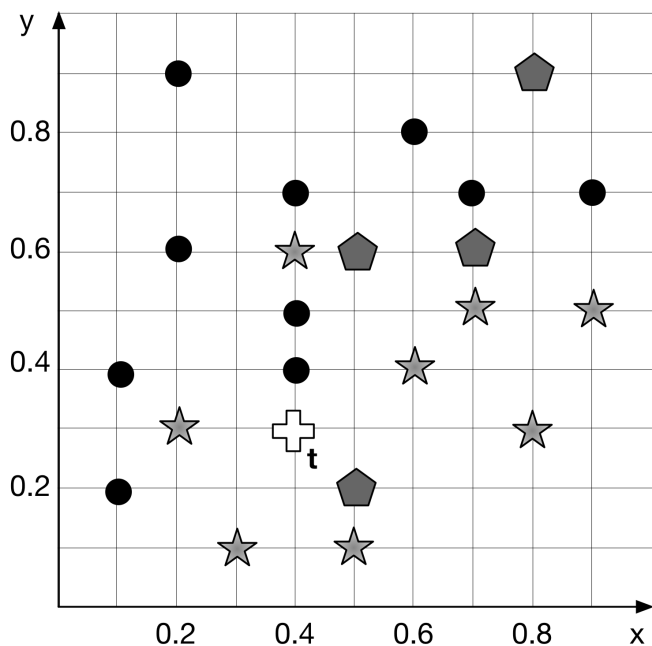
Use city-block distance (L^1 norm) throughout.

Question 1: K Nearest Neighbors (KNN)

- Use the majority vote version of KNN to assign the class of a test point on this homework.
- In this question there are three output classes: *Circle*, *Star*, and *Pentagon*.

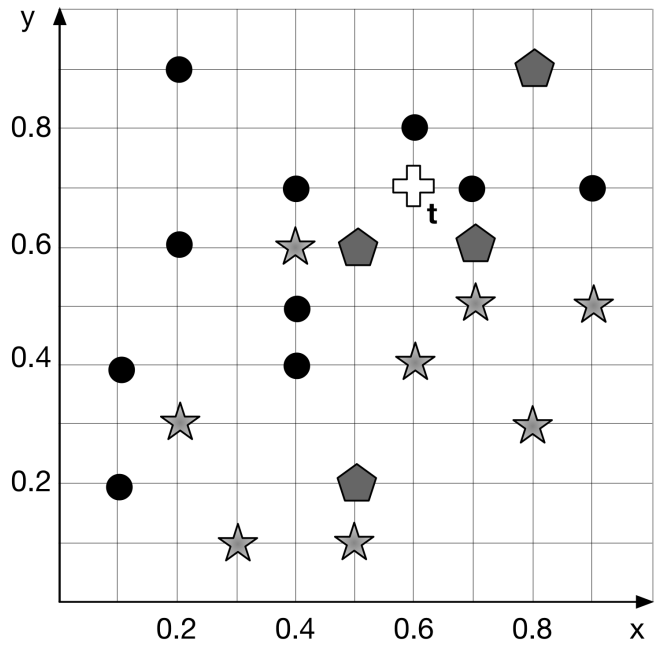
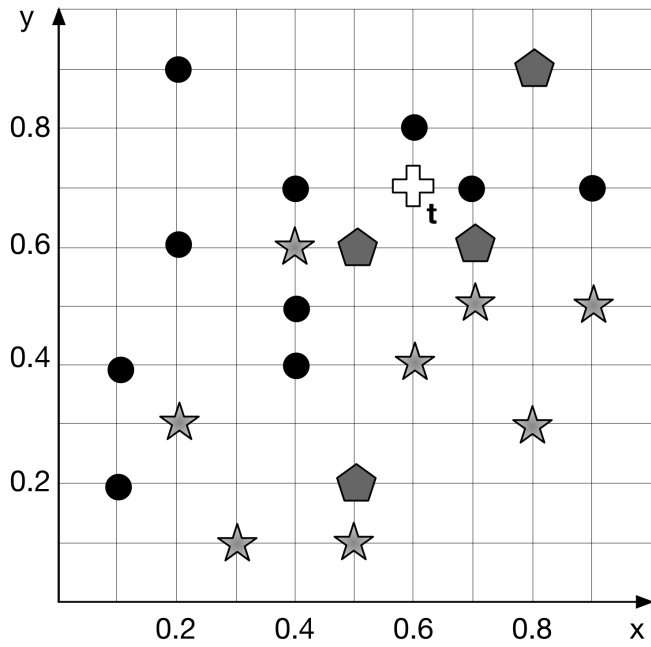
a) Test point $t = (0.4, 0.3)$. What is the predicted class when $K = 4$? Draw your neighborhood on the **left** plot.

b) Test point $t = (0.4, 0.3)$. What is the predicted class when $K = 8$? Draw your neighborhood on the **right** plot.



c) Test point $t = (0.6, 0.7)$. What is the predicted class when $K = 5$? Draw your neighborhood on the **left** plot.

d) Test point $t = (0.6, 0.7)$. What is the predicted class when $K = 9$? Draw your neighborhood on the **right** plot.



Question 2: Hierarchical clustering

a) Perform a hierarchical clustering using the single-linkage metric on the following matrix representing the distance between the following places in Maine. For full credit, be sure to include:

- List of cluster merge distances.
- Updated distance matrix after each cluster merge.
- The cities in each cluster after each merge.

$$dist = \begin{bmatrix} & A & B & C & D \\ A & 0 & 2 & 5 & 4 \\ B & 2 & 0 & 3 & 4 \\ C & 5 & 3 & 0 & 1 \\ D & 4 & 4 & 1 & 0 \end{bmatrix}$$

Key: *A*: Augusta, *B*: Bagnor, *C*: Corinth, *D*: Dexter.

b) Sketch a dendrogram of your results from (a). Your plot doesn't need to be drawn perfectly to scale. Label the x and y axes.

c) How many clusters would we have if we used a merge height threshold of $T = 1.5$?