
CSE497 Fall 2018 Homework 5

Assigned: Monday, November 26, 2018

Due: Monday, December 3, 2018 (by midnight, submit online via Canvas)

Maximum: 100 point

Note: This assignment is to be done by an individual student, no team work allowed.

1. (30%) Perform K-means on the data shown in Table 1. Use $K=2$ and suppose A and C are randomly selected as initial means. Please show the clustering results of nodes of each round until converged.

Table 1

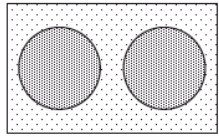
i	X_1	X_2
A	1	1
B	1	0
C	0	2
D	2	4
E	3	5

2. (50%) Use the **Similarity** Matrix in Table 2 to perform *MIN (Single Link)* and *MAX (Complete Link)* hierarchical clustering. Note that MIN/MAX Hierarchical Clustering schemes are named based on **dissimilarity** but here the matrix below contains the **similarity** measure. Show your results by drawing a dendrogram. The dendrogram should clearly show the order in which the points are merged.

Table 2 Similarity Measure

	p1	p2	p3	p4	p5
p1	1.00	0.10	0.41	0.55	0.35
p2	0.10	1.00	0.64	0.47	0.98
p3	0.41	0.64	1.00	0.44	0.85
p4	0.55	0.47	0.44	1.00	0.76
p5	0.35	0.98	0.85	0.76	1.00

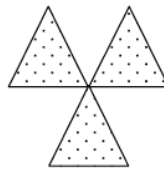
3. (20%) Identify possible clusters in the figure below using DBSCAN. Please indicate *the number of clusters* for each case and *briefly explain* your reasoning. Note that darkness or the number of dots in an area visually indicates its density.



(a)



(b)



(c)



(d)