Imaduddin Sheikh

CS 422 - Homework 9

1. Exercises

1.1. Tan, Chapter 7 (Cluster Analysis: Basic Concepts and Algorithms)

Q.7)

In order to minimize the squared error, more clusters need to be allocated to the less dense regions so the answer should be **(b)**.

Q.12)

(a)

The algorithm is computationally efficient as the data is scanned only once, and each object is, at most once, compared to the final set of centroids. It will always produce the same set of clusters for a fixed ordering of objects, even though the algorithm is dependent on order. However, the disadvantages of the leader algorithm involve not being able to set up the parameter for the resulting number of clusters, it can be done indirectly though. Relatively, the K-means algorithm performs better clustering when measured by SSE.

(b)

Measure the distances between the points and utilize the knowledge of their distribution to determine an optimal value of the threshold to set to. The algorithm can be tweaked such that it performs clustering for numerous thresholds.

Q.16)

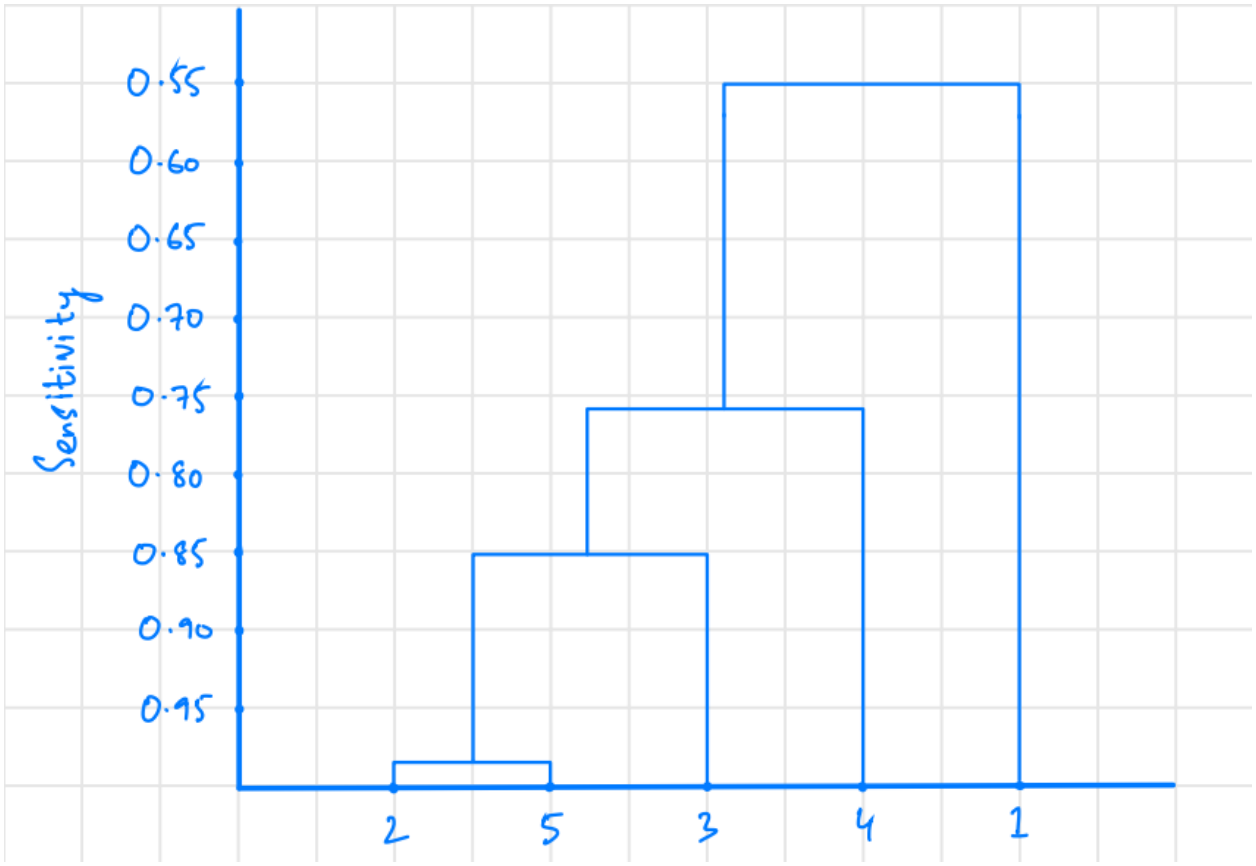
SINGLE LINK

	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

	p1	{p2, p5}	p3	p4
p1	1.00	0.35	0.41	0.55
{p2, p5}	0.35	1.00	0.85	0.76
p3	0.41	0.85	1.00	0.44
p4	0.55	0.76	0.44	1.00

	p1	{{p2, p5}, p3}	p4
p1	1.00	0.41	0.55
{{p2, p5}, p3}	0.41	1.00	0.76
p4	0.55	0.76	1.00

	{{{p2, p5}, p3}, p4}	p1
{{{p2, p5}, p3}, p4}	1.00	0.55
p1	0.55	1.00



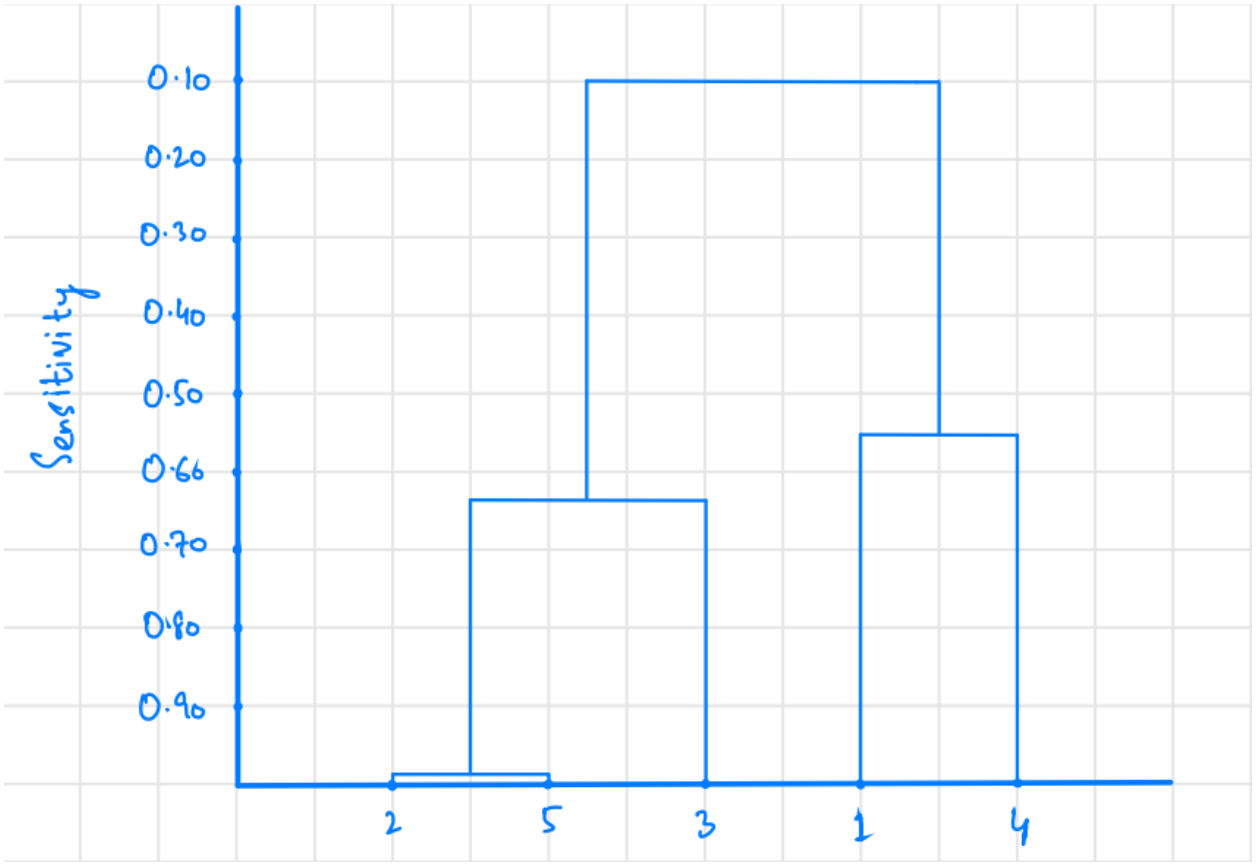
COMPLETE LINK

	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

	p1	{p2, p5}	p3	p4
p1	1.00	0.10	0.41	0.55
{p2, p5}	0.10	1.00	0.64	0.47
p3	0.41	0.64	1.00	0.44
p4	0.55	0.47	0.44	1.00

	p1	{{p2, p5}, p3}	p4
p1	1.00	0.10	0.55
{{p2, p5}, p3}	0.10	1.00	0.44
p4	0.55	0.44	1.00

	{p1, p4}	{{p2, p5}, p3}
{p1, p4}	1.00	0.10
{{p2, p5}, p3}	0.10	1.00


Sensitivity
0·10
0·20
0·30
0·40
0·50
0·66
0·70
0·80
0·90
2
5
3
1
4