Unit 10 - Week 8

Assignment 8

The due date for submitting this assignment has passed. Due on 2019-09-25, 23:59 IST.
As per our records you have not submitted this assignment.

1) What is true about K-Mean Clustering? 2 points

1. K-means is extremely sensitive to cluster center initializations
2. Bad initialization can lead to Poor convergence speed
3. Bad initialization can lead to bad overall clustering

a. 1 and 2
b. 1 and 3
c. All of the above
d. 2 and 3

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
c

2) In which of the following cases will K-Means clustering fail to give good results? (Choose all the correct answers) 2 points

a. Data points with outliers
b. Data points with round shapes
c. Data points with non-convex shapes
d. Data points with different densities

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
a
c
d

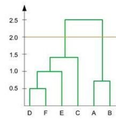
3) Which of the following clustering algorithms suffers from the problem of convergence at local optima? (Choose all the correct answers) 2 points

1. K-Means clustering algorithm.
2. Agglomerative clustering algorithm.
3. Expectation-Maximization clustering algorithm
4. Diverse clustering algorithm

1
2
3
4

No, the answer is incorrect.
Score: 0
Accepted Answers:
1
3

4) **In the figure below, if you draw a horizontal line on y-axis for y=2. What will be the number of clusters formed?** 2 points



a. 1
b. 2
c. 3
d. 4

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
c

5) **Assume, you want to cluster 7 observations into 3 clusters using K-Means clustering algorithm. After first iteration the clusters: C1, C2, C3 has the following observations:** 2 points

C1: {(1,1), (4,4), (7,7)}

C2: {(0,4), (4,0)}

C3: {(5,5), (9,9)}

What will be the cluster centroids if you want to proceed for second iteration?

a. C1: (4,4), C2: (2,2), C3: (7,7)
b. C1: (2,2), C2: (0,0), C3: (5,5)
c. C1: (6,6), C2: (4,4), C3: (9,9)
d. None of these

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
a

6) Following Question 5, what will be the Manhattan distance for observation (9, 9) from cluster centroid C1 in the second iteration? 2 points

a. 10
b. 5
c. 6
d. 7

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
a

7) **If two variables V1 and V2 are used for clustering. Which of the following are true for K-means clustering with k =3?** 2 points

1. **If V1 and V2 has a correlation of 1, the cluster centroids will be in a straight line**
2. **If V1 and V2 has a correlation of 0, the cluster centroids will be in straight line**

Choose the correct answer?

a. 1 Only
b. 2 Only
c. Both 1 and 2
d. None of the above

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
a

8) Which of the following is not a clustering approach? 2 points

a. Hierarchical
b. Partitioning
c. Bagging
d. Density-Based

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
c

9) 2 points

Given six points with the following attributes:

point	x coordinate	y coordinate
p1	0.4005	0.5306
p2	0.2148	0.3854
p3	0.3457	0.3156
p4	0.2652	0.1875
p5	0.0789	0.4139
p6	0.4548	0.3022

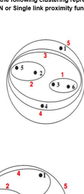
Table : X-Y coordinates of six points.

	p1	p2	p3	p4	p5	p6
p1	0.0000	0.2357	0.2218	0.3688	0.3421	0.2347
p2	0.2357	0.0000	0.1483	0.2042	0.1388	0.2540
p3	0.2218	0.1483	0.0000	0.1513	0.2843	0.1100
p4	0.3688	0.2042	0.1513	0.0000	0.2932	0.2216
p5	0.3421	0.1388	0.2843	0.2932	0.0000	0.3921
p6	0.2347	0.3540	0.1100	0.2216	0.3921	0.0000

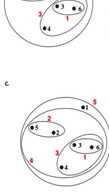
Table : Distance Matrix for Six Points

Which of the following clustering representations and dendrogram depicts the use of MIN or Single link proximity function in hierarchical clustering:

A.



B.



C.

D.

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
a

10) **Which of the following clustering algorithms is the most sensitive to outliers?** 2 points

a. K-means clustering algorithm
b. K-medians clustering algorithm
c. K-modes clustering algorithm
d. None of the above

a
b
c
d

No, the answer is incorrect.
Score: 0
Accepted Answers:
a