Vidyavardhini's College of Engineering and Technology

Department of Artificial Intelligence & Data Science

AY: 2024-25

Class:	TE	Semester:	V
Course Code:		Course Name:	DWM

Name of Student:	Sainath Khot
Roll No. :	20
Assignment No.:	4
Title of Assignment:	
Date of Submission:	
Date of Correction:	

Evaluation

Performance Indicator	Max. Marks	Marks Obtained
Completeness	5	
Demonstrated Knowledge	3	
Legibility	2	
Total	10	

Performance Indicator	Exceed Expectations (EE)	Meet Expectations (ME)	Below Expectations (BE)
Completeness	5	3-4	1-2
Demonstrated Knowledge Legibility	3	2	1
Legibility	2	1	0

Checked by

Name of Faculty :

Signature :

Date :

DWM Assignment 4

Q.

Points	A	B
P_1	2	4
P_2	8	2
P_3	9	3
P_4	1	5
P_5	8.5	1

Sol

P_1, P_3, P_5 are centre of each cluster Number of cluster $K = 3$

Initial cluster $K = 3$

$C_2 = P_3 (9, 3)$ $C_3 = P_5 (8.5, 1)$

Euclidean distance formula

$$\text{Dist}[(x, y), (a, b)] = \sqrt{(x-a)^2 + (y-b)^2}$$

iteration 1

$P_1 (2, 4)$

$\text{Distance}[(2,4), (2,4)] = \sqrt{(2-2)^2 + (4-4)^2} = 0$

$\text{Dist}[(2,4), (9,3)] = \sqrt{(2-9)^2 + (4-3)^2} = 7.07$

$\text{Dist}[(2,4), (8.5,1)] = \sqrt{(2-8.5)^2 + (4-1)^2} = 7.15$

$P_1 (2, 4)$ belongs to cluster C_1

$P_2 (8, 2)$

$\text{Distance}[(8,2), (2,4)] = \sqrt{(8-2)^2 + (2-4)^2} = 6.32$

$\text{Dist}[(8,2), (9,3)] = \sqrt{(8-9)^2 + (2-3)^2} = 1.41$

$\text{Dist}[(8,2), (8.5,1)] = \sqrt{(8-8.5)^2 + (2-1)^2} = 1.11$

$P_2 (8, 2)$ belongs to cluster C_3

P_3 (9, 3)

Dist [(9,3) (2,4)] : $\sqrt{(9-2)^2 + (3-4)^2}$ = 7.07

Dist [(9,3) (9,3)] : $\sqrt{(9-9)^2 + (3-3)^2}$ = 0

Dist [(9,3) (8.5,1)] : $\sqrt{(9-8.5)^2 + (3-1)^2}$ = 2.06

P_3 (9,3) belongs to cluster C_2

P_4 (1,5)

Distance [(1,5) (2,4)] : $\sqrt{(1-2)^2 + (5-4)^2}$ = 1.414

Dist [(1,5), (9,3)] : $\sqrt{(1-9)^2 + (5-3)^2}$ = 8.24

Dist [(9,5), (8.5,1)] : $\sqrt{(1-8.5)^2 + (5-1)^2}$ = 8.5

P_4 (1,5) belongs to cluster C_1

P_5 (8.5, 1)

Dist [(8.5,1), (2,4)] : $\sqrt{(8.5-2)^2 + (1-4)^2}$ = 7.16

Dist [(1,5), (9,3)] : $\sqrt{(1-9)^2 + (5-3)^2}$ = 8.24

Dist [(1,5), (8.5,1)] : $\sqrt{(1-8.5)^2 + (5-1)^2}$ = 8.5

P_5 (8.5, 1) belongs to cluster C_3

After iteration 1

cluster C_1 : [P_1 (2,4), P_4 (1,5)]

cluster C_2 : [P_3 (9,3)]

cluster C_3 : [P_2 (8,2), P_5 (8.5, 1)]

Iteration 2:

Centre of new clusters

Cluster C_1 : $\left[\frac{2+1}{2}, \frac{4+5}{2}\right]$ = (1.5, 4.5)

Cluster C1 : (9, 3)

Cluster C2 : $\left(\frac{8 + 8.5}{2}, \frac{2+1}{2} \right) = (8.25, 1.5)$

P_1 (2, 4)

Dist [(2, 4) (1.5, 4.5)] = $\sqrt{(2-1.5)^2 + (4-4.5)^2} = 0.70$

Dist [(2, 4) (9, 3)] = $\sqrt{(2-9)^2 + (4-3)^2} = 7.07$

Dist [(2, 4), (8.25, 1.5)] = $\sqrt{(2-8.2)^2 + (4-1.5)^2} = 6.73$

P_2 (2, 4) belongs to cluster C1

P_3 (8, 2)

Dist [(8, 2), (1.5, 4.5)] = $\sqrt{(8-1.5)^2 + (2-4.5)^2} = 6.96$

Dist [(8, 2), (9, 3)] = $\sqrt{(8-9)^2 + (2-3)^2} = 1.41$

Dist [(8, 2), (8.25, 1.5)] = $\sqrt{(8-8.25)^2 + (2-1.5)^2} = 0.55$

P_2 (8, 2) belongs to cluster C3

P_3 (9, 3)

Dist [(9, 3), (1.5, 4.5)] = $\sqrt{(9-1.5)^2 + (3-4.5)^2} = 6.96$

Dist [(9, 3), (9, 3)] = $\sqrt{(9-9)^2 + (3-3)^2} = 0$

Dist [(9, 3) (8.25, 1.5)] = $\sqrt{(9-8.25)^2 + (3-1.5)^2} = 1.67$

P_3 (9, 3) belongs to cluster C2

P_4 (1, 5)

Dist [(1, 5) (1.5, 4.5)] = $\sqrt{(1-1.5)^2 + (5-4.5)^2} = 0$

Dist [(1, 5) (8.25, 1.5)] = $\sqrt{(1-8.7)^2 + (5-1.5)^2} = 8.05$

P_4 (1, 5) belongs to cluster C1

P_5 (8.5 , 1)

Dist [(8.5, 1) , (1.5 , 4.5)] : $\sqrt{(8.5 - 1.5)^2 + (1 - 4.5)^2}$

Dist [(8.5 , 1) , (9, 3)] : $\sqrt{(8.5 - 9)^2 + (1 - 3)^2}$ = 2.06

Dist [(8.5, 1) , (8.25 , 1.5)] : $\sqrt{(8.5 - 8.25)^2 + (1 - 1.5)^2}$ 0.

P_5 (8.5 , 1) belongs to cluster C3

After Iteration 2

Cluster C_1 : { P_1 (2,4) , P_4 (1,5) }

Cluster C_2 : { P_3 (9,3) }

Cluster C_3 : { P_2 (8,2) , P_5 (8.5 , 1) }

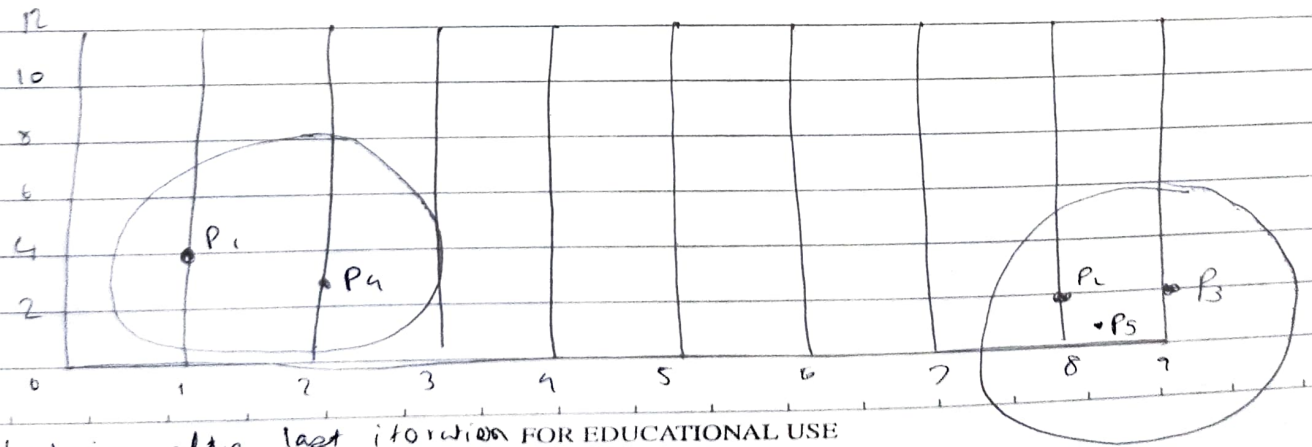
~~cluster~~ {

Comparing the clustering of iteration 1 and iteration 4, we find that object does not name cluster anymore thus there no computation of the K-mean clustering has reached its stability and no more iteration is needed So we final clusters are :

Cluster C_1 = { P_1 (2,4) , P_4 (1,5) }

Cluster C_2 = { P_3 (9,3) }

cluster C_3 = { P_2 (8,2) , P_5 { 8.5 , 1) }



Clustering after last iteration

2. Given

Object	X	Y
0	8	7
1	3	7
2	4	9
3	9	6
4	8	5
5	5	8
6	7	3
7	8	4
8	7	5
9	4	5

Step 1 : We select random representative objects
C_1 (4, 9) C_2 (8, 4)

(Object) i	x	y	C_1		C_2 (8, 4)	C
0	8	7	4	9	$\|8-4\| + \|7-9\| =$	6
1	3	7	4	9	$\|3-4\| + \|7-9\| =$	3
~~2~~	4	9	4	9	$\|4-4\| + \|9-9\| =$	
3	9	6	4	9	$\|9-4\| + \|6-9\| =$	8
4	8	5	4	9	$\|5-5\| + \|5-9\| =$	8
5	5	8	4	9	$\|5-4\| + \|8-9\| =$	2
6	7	3	4	9	$\|7-4\| + \|3-9\| =$	9
~~7~~	8	4	4	9	$\|8-4\| + \|4-9\| =$	
8	7	5	4	9	$\|7-4\| + \|5-9\|$	7
9	4	5	4	9	$\|4-4\| + \|5-9\|$	4

Object(i)	x	y	C_1	C_2	Dist / cost	C
0	8	7	8	4	$\|8-8\| + \|7-4\|$	3
1	3	7	8	4	$\|3-8\| + \|7-4\|$	8
3	9	6	8	4	$\|9-8\| + \|6-4\|$	3
4	8	5	8	4	$\|8-8\| + \|5-4\|$	1
5	5	8	8	4	$\|5-8\| + \|8-4\|$	7
6	7	3	8	4	$\|7-8\| + \|3-4\|$	2
8	7	5	8	4	$\|7-8\| + \|5-4\|$	2
9	4	5	8	4	$\|4-1\| + \|5-4\|$	5

Compare cost of cost (1) and cost (2) for every i and select the minimum one

Step 2 The cluster are

Cluster 1: {(3,7), (4,9), (5,8), (4,5)}

cluster 2 {(8,7), (9,6), (8,5), (7,3), (8,4), (7,5)}

Calculate total cost

$$\text{cost}(x,c) = \sum_{i=1}^{d} |x_i - c_i|$$

Total cost = {cost((8,4),(8,7)), cost((8,4)(9,6)), cost((8,4),(8,5)), cost((8,4),(7,3)), cost((8,4),(7,5)), cost((9,9),(3,7)), cost((9,9)(5,8)), cost((9,9),(9,5))}

$= (3+3+1+2+2) + (3+2+4)$

$11 + 9 = 20$

Step 3 Select one of non-medoids O

Let O = (8,5) i.e object (4)

So new medoid are C1 (4,9) and O' (8,5)

object (i)	x	y	O'		Dist / cost	C
	8	7	8	5	|8-8| + |7-5|	2
	3	7	8	5	|3-8| + |7-5|	7
	9	6	8	5	|9-8| + |6-5|	2
	5	8	8	5	|5-8| + |8-5|	6
	7	3	8	5	|7-8| + |3-5|	3
	8	4	8	5	|8-8| + |4-5|	1
	7	5	8	5	|7-8| + |5-5|	1
	4	5	8	5	|4-8| + |5-5|	4

Object (i)	x	y	C1		Dist / cost	C
0	8	7	4	9	|8-4| + |7-9|	6
1	3	7	4	9	|3-4| + |7-9|	3
3	9	6	4	9	|9-4| + |6-9|	8
5	5	8	4	9	|5-4| + |8-9|	2
6	7	3	4	9	|7-4| + |3-9|	9
7	8	4	4	9	|8-4| + |4-9|	9
8	7	5	4	9	|7-4| + |5-9|	7
9	4	5	4	9	|4-4| + |5-9|	4

Compare the cost of cost (C1) & cost (O') every (i) and select the min one

Again create the cluster

cluster 1 : {(3,4), (3,8), (4,5), (4,9)}

cluster 2 : {(8,7), (9,6), (7,7), (8,4), (7,5), (8,5)}

current total cost : (2+2+3+1+1) + (3+2+4)

= 9+9 = 18

Step 4 : So cost of swapping medoid from C2 to O' is

S = current total cost - past total cost

= 18 - 20 = -2 < 0

So moving O' would be a good idea

Now, move cluster again

Select one of Non-medoid O''

Let O'' = (7,3) i.e object 6

So new medoid are C1(4,9) & O''(7,3)

Object (i)	x	y	C_i	Dist / cost	C
0	8	7	4 9	|8-4| + |7-9|	5
1	3	7	4 9	|3-4| + |7-9|	3
3	9	6	4 9	|9-4| + |6-9|	8
4	8	5	4 9	|8-4| + |5-9|	8
5	5	8	4 9	|5-4| + |8-9|	2
7	8	4	4 9	|8-4| + |4-9|	9
8	7	5	9 9	|7-9| + |5-9|	7
9	4	5	4 9	|4-4| + |5-9|	4

ect(i)	x	y	O''	Dist / cost	C
0	8	7	7 3	|8-7| + |7-3|	5
1	3	7	7 3	|3-7| + |7-3|	8
3	9	6	7 3	|9-7| + |6-3|	5
4	8	5	7 3	|8-7| + |5-3|	3
5	5	8	7 3	|5-7| + |8-3|	7
7	4	4	7 3	|4-7| + |4-3|	2
8	7	5	7 3	|7-7| + |5-3|	2
9	4	5	7 3	|4-7| + |5-3|	5

Again create the cluster

Cluster 1 : { (3,7) , (5,8) , (9,5) , (4,9) }

Cluster 2 : { (8,7) , (9,6) , (8,5) , (8,4) , (7,5) , (7,3) }

Current total cost = (3+2+4) + (5+5+3+2+2)

= 9 + 17 = 26

So cost of swapping medoid from O' to O'' is

S = current total cost - past total cost

S = 26 - 18 = 8 > 0