20181CSE0621
Sai Ram

20181CSE0621
Sai Ram.K
7-CSE-10

Part-C

Q.2]

X	1	2	2	3	4	5
Y	1	1	3	2	3	5

Step1: Choosing cluster centers,
$v_1 = (2,1)$ and $v_2 (2,3)$

Data Point	Distance from v_1	Distance from v_2	Allocated
$a_1(1,1)$	1	2.24	v_1
$a_2(2,1)$	0	2	v_1
$a_3(2,3)$	2	0	v_2
$a_4(3,2)$	1.41	1.41	v_1
$a_5(4,3)$	2.83	2	v_2
$a_6(5,5)$	5	3.61	v_2

* Distance is calculated by euclidean measure
(x_1,x_2) & (y_1,y_2), $d = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2}$

$\therefore a_1(1,1) :- d = \sqrt{(1-2)^2+(1-1)^2} = 0$

Step3: Cluster 1 of $v_1 = \{a_1, a_2, a_4\}$
Cluster 2 of $v_2 = \{a_3, a_5, a_6\}$

Step4: Recalculate cluster centers:

$v_1 = \frac{1}{3}[(1,1)+(2,1)+(3,2)] = (2, 1.33)$

$v_2 = \frac{1}{3}[(2,3)+(4,3)+(5,5)] = \frac{1}{3}[(11,11)]$

$= (3.67, 3.67)$

2018ICSE0621
Sai Ram

classmate
Date
Page

Repeating step2 until we get same elements

Data Point	Distance from V_1	Distance from V_2	Allocated
a_1	1.05	3.78	V_1
a_2	0.33	3.15	V_1
a_3	1.67	1.8	V_1
a_4	1.204	1.8	V_1
a_5	2.605	0.75	V_2
a_6	4.711	1.88	V_2

→ Cluster 1 of $V_1 = \{a_1, a_2, a_3, a_4\}$
Cluster 2 of $V_2 = \{a_5, a_6\}$

Recalculating centres,

$V_1 = \frac{1}{4}[a_1 + a_2 + a_3 + a_4] = \frac{1}{4}[(1,1) + (2,1) + (2,3) + (3,2)]$

$= (2, 1.75)$

$V_2 = \frac{1}{2}[a_5 + a_6] = \frac{1}{2}[(4,3) + (5,5)]$

$= (4.5, 4)$

Repeating step2 again as centres not same.

Data Point	Distance from V_1	Distance from V_2	Allocated
a_1	1.25	4.61	V_1
a_2	0.75	3.9	V_1
a_3	1.25	2.69	V_1
a_4	1.03	2.5	V_1
a_5	2.36	1.12	V_2
a_6	4.42	1.12	V_2

20181CSE0621
Sai Ram

→ Cluster 1 of $V_1 = \{a_1, a_2, a_3, a_4\}$
Cluster 2 of $V_2 = \{a_5, a_6\}$

* Clusters are the same as previous iteration hence we can stop.

⟹ The final clusters are

Cluster 1 : $\{(1,1), (2,1), (2,3), (3,2)\}$

Cluster 2 : $\{(4,3), (5,5)\}$