Nama : RIZQILLAH
NIM : 1957301020
Kelas : TI 4B
MK : Data Mining

K-MEANS CLUSTERING

Dataset :

Titik	X	Y
1	2	1
2	3	2
3	2	2
4	2	3
5	1	3
6	3	1
7	4	2
8	6	3
9	5	4
10	4	4

Untuk mengaplikasikan klastering, dibutuhkan klaster, disini akan menggunakan 4 cluster beserta nilai random yang berbeda, yaitu sebagai berikut:

Cluster	X	Y
K1	2	2
K2	3	4
K3	7	4
K4	1.5	3

Rumus persamaan Euclidean Distance

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

Perhitungan Cluster 1 (K1) :
K1(x,y) = K1(2,2)

Cluster 1 Titik 1 (2, 1) = $\sqrt{(2-2)^2 + (1-2)^2} = 1$
Cluster 1 Titik 2 (3, 2) = $\sqrt{(3-2)^2 + (2-2)^2} = 1$
Cluster 1 Titik 3 (2, 2) = $\sqrt{(2-2)^2 + (2-2)^2} = 0$
Cluster 1 Titik 4 (2, 3) = $\sqrt{(2-2)^2 + (3-2)^2} = 1$
Cluster 1 Titik 5 (1, 3) = $\sqrt{(1-2)^2 + (3-2)^2} = 1{,}4$
Cluster 1 Titik 6 (3, 1) = $\sqrt{(3-2)^2 + (1-2)^2} = 1{,}4$
Cluster 1 Titik 7 (4, 2) = $\sqrt{(4-2)^2 + (2-2)^2} = 2$
Cluster 1 Titik 8 (6, 3) = $\sqrt{(6-2)^2 + (3-2)^2} = 4{,}12$
Cluster 1 Titik 9 (5, 4) = $\sqrt{(5-2)^2 + (4-2)^2} = 3{,}6$
Cluster 1 Titik 10 (4, 4) = $\sqrt{(4-2)^2 + (4-2)^2} = 2{,}83$

Perhitungan Cluster 2 (K2) :
K2(x,y) = K2(3,4)

Cluster 2 Titik 1 (2, 1) $= \sqrt{(2-3)^2 + (1-4)^2} = 3{,}16$
Cluster 2 Titik 2 (3, 2) = $\sqrt{(3-3)^2 + (2-4)^2} = 2$
Cluster 2 Titik 3 (2, 2) = $\sqrt{(2-3)^2 + (2-4)^2} = 2{,}24$
Cluster 2 Titik 4 (2, 3) = $\sqrt{(2-3)^2 + (3-4)^2} = 1{,}41$
Cluster 2 Titik 5 (1, 3) = $\sqrt{(1-3)^2 + (3-4)^2} = 2{,}24$
Cluster 2 Titik 6 (3, 1) = $\sqrt{(3-3)^2 + (1-4)^2} = 3$
Cluster 2 Titik 7 (4, 2) = $\sqrt{(4-3)^2 + (2-4)^2} = 2{,}24$
Cluster 2 Titik 8 (6, 3) = $\sqrt{(6-3)^2 + (3-4)^2} = 3{,}16$
Cluster 2 Titik 9 (5, 4) = $\sqrt{(5-3)^2 + (4-4)^2} = 2$
Cluster 2 Titik 10 (4, 4) = $\sqrt{(4-3)^2 + (4-4)^2} = 1$

Perhitungan Cluster 3 (K3) :
K3(x,y) = K3(7,4)

Cluster 3 Titik 1 (2, 1) $= \sqrt{(2-7)^2 + (1-4)^2} = 5{,}83$
Cluster 3 Titik 2 (3, 2) = $\sqrt{(3-7)^2 + (2-4)^2} = 4{,}47$
Cluster 3 Titik 3 (2, 2) = $\sqrt{(2-7)^2 + (2-4)^2} = 5{,}39$
Cluster 3 Titik 4 (2, 3) = $\sqrt{(2-7)^2 + (3-4)^2} = 5{,}1$
Cluster 3 Titik 5 (1, 3) = $\sqrt{(1-7)^2 + (3-4)^2} = 6{,}08$
Cluster 3 Titik 6 (3, 1) = $\sqrt{(3-7)^2 + (1-4)^2} = 5$
Cluster 3 Titik 7 (4, 2) = $\sqrt{(4-7)^2 + (2-4)^2} = 3{,}61$
Cluster 3 Titik 8 (6, 3) = $\sqrt{(6-7)^2 + (3-4)^2} = 1{,}41$
Cluster 3 Titik 9 (5, 4) = $\sqrt{(5-7)^2 + (4-4)^2} = 2$
Cluster 3 Titik 10 (4, 4) = $\sqrt{(4-7)^2 + (4-4)^2} = 3$

Perhitungan Cluster 4 (K4) :
K4(x,y) = K4(1.5,3)

Cluster 4 Titik 1 (2, 1) $= \sqrt{(2-1.5)^2 + (1-3)^2} = 2{,}06$
Cluster 4 Titik 2 (3, 2) = $\sqrt{(3-1.5)^2 + (2-3)^2} = 1{,}80$
Cluster 4 Titik 3 (2, 2) = $\sqrt{(2-1.5)^2 + (2-3)^2} = 1{,}12$
Cluster 4 Titik 4 (2, 3) = $\sqrt{(2-1.5)^2 + (3-3)^2} = 0{,}5$
Cluster 4 Titik 5 (1, 3) = $\sqrt{(1-1.5)^2 + (3-3)^2} = 0{,}5$
Cluster 4 Titik 6 (3, 1) = $\sqrt{(3-1.5)^2 + (1-3)^2} = 2{,}5$

Cluster 4 Titik 7 (4, 2) = $\sqrt{(4-1.5)^2+(2-3)^2} = 2{,}69$
Cluster 4 Titik 8 (6, 3) = $\sqrt{(6-1.5)^2+(3-3)^2} = 4{,}5$
Cluster 4 Titik 9 (5, 4) = $\sqrt{(5-1.5)^2+(4-3)^2} = 3{,}64$
Cluster 4 Titik 10 (4, 4) = $\sqrt{(4-1.5)^2+(4-3)^2} = 2{,}69$

No	X	Y	Cluster 1 (2,2)	Cluster 2 (3,4)	Cluster 3 (7,4)	Cluster 4 (1.5,3)	Jarak Terdekat	Kelompok Cluster
1	2	1	1	3,16	5,83	2,06	1	1
2	3	2	1	2	4,47	1,80	1	1
3	2	2	0	2,4	5,39	1,12	0	1
4	2	3	1	1,41	5,1	0,5	0,5	4
5	1	3	1,4	2,24	6,08	0,5	0,5	4
6	3	1	1,4	3	5	2,5	1,4	1
7	4	2	2	2,24	3,61	2,69	2	1
8	6	3	4,12	3,16	1,41	4,5	1,41	3
9	5	4	3,6	2	2	3,64	2	2
10	4	4	2,83	1	3	2,69	1	2

Adapun rumus untuk mencari jarak terdekat dari nilai Cluster yaitu dengan cara mencari minimum dari ke-4 cluster yang ada, maka hasilnya seperti di atas.

Kesimpulan :
K1 = 1,2,3,6,7
K2 = 9,10
K3 = 8
K4 = 4,5

Iterasi selanjutnya adalah menentukan nilai tiap klaster dari hasil pertambahan dari setiap nilai x dan y dari kelompok klaster.

Rumusnya menentukan nilai klaster selanjutnya: $K_j = (i_1 + i_2 + i_3 \ldots + i_n)/n$

Keterangan:
- K adalah kelompok klaster
- j adalah urutan klaster
- i adalah urutan dari setiap himpunan data klaster
- n adalah banyaknya himpunan data klaster

Maka nilai x untuk klaster terbaru sebagai berikut:
K1 = (2+3+2+3+4)/5 = 2.8
K2 = (5+4)/2 = 4.5
K3 = (6)/1 = 6
K4 = (2+1)/2 = 1.5

Nilai y untuk klaster terbaru sebagai berikut:
K1 = (1+2+2+1+2)/5 = 1.6
K2 = (4+4)/2 = 4
K3 = (3)/1 = 3
K4 = (3+3)/2 = 3

Cluster	X	Y
K1	2.8	1.6
K2	4.5	4
K3	6	3
K4	1.5	3

Lakukan perhitungan menggunakan persamaan Euclidean Distance.

Rumus persamaan Euclidean Distance

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

Perhitungan Cluster 1 (K1) :
K1(x,y) = K1(2,2)

Cluster 1 Titik 1 (2, 1) = $\sqrt{(2-2.8)^2 + (1-1.6)^2} = 1$
Cluster 1 Titik 2 (3, 2) = $\sqrt{(3-2.8)^2 + (2-1.6)^2} = 0{,}45$
Cluster 1 Titik 3 (2, 2) = $\sqrt{(2-2.8)^2 + (2-1.6)^2} = 0.9$
Cluster 1 Titik 4 (2, 3) = $\sqrt{(2-2.8)^2 + (3-1.6)^2} = 1{,}61$
Cluster 1 Titik 5 (1, 3) = $\sqrt{(1-2.8)^2 + (3-1.6)^2} = 2{,}28$
Cluster 1 Titik 6 (3, 1) = $\sqrt{(3-2.8)^2 + (1-1.6)^2} = 0{,}63$
Cluster 1 Titik 7 (4, 2) = $\sqrt{(4-2.8)^2 + (2-1.6)^2} = 1{,}26$
Cluster 1 Titik 8 (6, 3) = $\sqrt{(6-2.8)^2 + (3-1.6)^2} = 3{,}49$
Cluster 1 Titik 9 (5, 4) = $\sqrt{(5-2.8)^2 + (4-1.6)^2} = 3{,}26$
Cluster 1 Titik 10 (4, 4) = $\sqrt{(4-2.8)^2 + (4-1.6)^2} = 2{,}68$

Perhitungan Cluster 2 (K2) :
K2(x,y) = K2(3,4)

Cluster 2 Titik 1 (2, 1) $= \sqrt{(2-4.5)^2+(1-4)^2} = 3{,}90$
Cluster 2 Titik 2 (3, 2) = $\sqrt{(3-4.5)^2+(2-4)^2} = 2.5$
Cluster 2 Titik 3 (2, 2) = $\sqrt{(2-4.5)^2+(2-4)^2} = 3{,}2$
Cluster 2 Titik 4 (2, 3) = $\sqrt{(2-4.5)^2+(3-4)^2} = 2{,}69$
Cluster 2 Titik 5 (1, 3) = $\sqrt{(1-4.5)^2+(3-4)^2} = 3{,}64$
Cluster 2 Titik 6 (3, 1) = $\sqrt{(3-4.5)^2+(1-4)^2} = 3{,}35$
Cluster 2 Titik 7 (4, 2) = $\sqrt{(4-4.5)^2+(2-4)^2} = 2{,}06$
Cluster 2 Titik 8 (6, 3) = $\sqrt{(6-4.5)^2+(3-4)^2} = 1{,}8$
Cluster 2 Titik 9 (5, 4) = $\sqrt{(5-4.5)^2+(4-4)^2} = 0{,}5$
Cluster 2 Titik 10 (4, 4) = $\sqrt{(4-4.5)^2+(4-4)^2} = 0{,}5$

Perhitungan Cluster 3 (K3) :
K3(x,y) = K3(7,4)

Cluster 3 Titik 1 (2, 1) $= \sqrt{(2-6)^2+(1-3)^2} = 4{,}47$
Cluster 3 Titik 2 (3, 2) = $\sqrt{(3-6)^2+(2-3)^2} = 3{,}16$
Cluster 3 Titik 3 (2, 2) = $\sqrt{(2-6)^2+(2-3)^2} = 4{,}12$
Cluster 3 Titik 4 (2, 3) = $\sqrt{(2-6)^2+(3-3)^2} = 4$
Cluster 3 Titik 5 (1, 3) = $\sqrt{(1-6)^2+(3-3)^2} = 5$
Cluster 3 Titik 6 (3, 1) = $\sqrt{(3-6)^2+(1-3)^2} = 3{,}6$
Cluster 3 Titik 7 (4, 2) = $\sqrt{(4-6)^2+(2-3)^2} = 2{,}24$
Cluster 3 Titik 8 (6, 3) = $\sqrt{(6-6)^2+(3-3)^2} = 0$
Cluster 3 Titik 9 (5, 4) = $\sqrt{(5-6)^2+(4-3)^2} = 1{,}41$
Cluster 3 Titik 10 (4, 4) = $\sqrt{(4-6)^2+(4-3)^2} = 2{,}24$

Perhitungan Cluster 4 (K4) :
K4(x,y) = K4(1.5,3)

Cluster 4 Titik 1 (2, 1) $= \sqrt{(2-1.5)^2+(1-3)^2} = 2{,}06$
Cluster 4 Titik 2 (3, 2) = $\sqrt{(3-1.5)^2+(2-3)^2} = 1{,}80$
Cluster 4 Titik 3 (2, 2) = $\sqrt{(2-1.5)^2+(2-3)^2} = 1{,}12$
Cluster 4 Titik 4 (2, 3) = $\sqrt{(2-1.5)^2+(3-3)^2} = 0{,}5$
Cluster 4 Titik 5 (1, 3) = $\sqrt{(1-1.5)^2+(3-3)^2} = 0{,}5$
Cluster 4 Titik 6 (3, 1) = $\sqrt{(3-1.5)^2+(1-3)^2} = 2{,}5$

Cluster 4 Titik 7 (4, 2) = $\sqrt{(4-1.5)^2+(2-3)^2} = 2{,}69$
Cluster 4 Titik 8 (6, 3) = $\sqrt{(6-1.5)^2+(3-3)^2} = 4{,}5$
Cluster 4 Titik 9 (5, 4) = $\sqrt{(5-1.5)^2+(4-3)^2} = 3{,}64$
Cluster 4 Titik 10 (4, 4) = $\sqrt{(4-1.5)^2+(4-3)^2} = 2{,}69$

No	X	Y	Cluster 1 (2.8,1.6)	Cluster 2 (4.5,4)	Cluster 3 (6,3)	Cluster 4 (1.5,3)	Jarak Terdekat	Kelompok Cluster
1	2	1	1	3,90	4,47	2,06	1	1
2	3	2	0,45	2,5	3,16	1,80	0,45	1
3	2	2	0,9	3,2	4,12	1,12	0,9	1
4	2	3	1,61	2,69	4	0,5	0,5	4
5	1	3	2,28	3,64	5	0,5	0,5	4
6	3	1	0,63	3,35	3,6	2,5	0,63	1
7	4	2	1,26	2,06	2,24	2,69	1,26	1
8	6	3	3,49	1,8	0	4,5	0	3
9	5	4	3,26	0,5	1,41	3,64	0,5	2
10	4	4	2,68	0,5	2,24	2,69	0,5	2

Kesimpulan :
K1 = 1,2,3,6,7
K2 = 9,10
K3 = 8
K4 = 4,5

Dikarenakan klaster tidak ada yang berubah, maka iterasi pemisahan klaster telah selesai dilakukan.