

Machine Learning 2017/2018

1. a) Dari grafik yang diberikan dapat disimpulkan bahwa nilai  $k$  terbaik adalah 4, mengapa demikian? Karena pada grafik dpt dilihat nilai SSE yang paling menurun drastis terdapat pada  $k = 4$ , hal ini menunjukkan bahwa perubahan  $k$  menjadi 4 secara maksimal mengurangi nilai error pada clustering.

Nb: jika pada average silhouette yang terbaik pada titik tertinggi,

b). Tidak, strategi yang dapat dilakukan dengan algoritma k-means untuk mendapat hasil clustering yang memadai adalah dengan menambahkan dimensi fitur.

2).

	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$
$P_1$	1				
$P_2$	0,1	1			
$P_3$	0,41	0,64	1		
$P_4$	0,55	0,47	0,44	1	
$P_5$	0,55	0,98	0,95	0,76	1

① Use single - 0 case nilai similarity

•  $\max(D) = 1 = \max(d_{1,5}) = 0,98$

↳ gabungkan 2 dan 5

• menghitung jarak (2 dan 5) dengan kelompok lain :

$d_{(25)1} = \min(d_{21}, d_{51}) = \min(0,1, 0,55) = 0,1$

$d_{(25)3} = \min(d_{23}, d_{53}) = 0,64$

$d_{(25)4} = \min(d_{24}, d_{54}) = 0,47$

D	1	(25)	3	4
1	1	0,1	0,41	0,55
(25)	0,1	1	0,64	0,47
3	0,41	0,64	1	0,44
4	0,55	0,47	0,44	1

•  $\max(D) = \max(d_{3,4}) = 0,44$

• menghitung jarak (2,3 dan 5) dengan kelompok lain :

$d_{(235)1} = \min(d_{21}, d_{31}, d_{51}) = 0,1$

$d_{(235)4} = \min(d_{24}, d_{34}, d_{54}) = 0,44$

D	1	(235)	4
1	1	0,1	0,55
(235)	0,1	1	0,44
4	0,55	0,44	1

•  $\max(D) = \max(d_{1,4}) = 0,55$

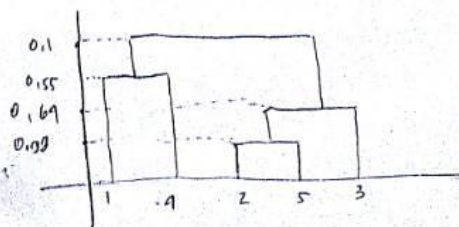
• menghitung jarak (4) dengan kelompok lain :

$d_{(14)(235)} = \min(d_{12}, d_{13}, d_{15}, d_{42}, d_{43}, d_{45}) = 0,1$

D	(14)	(235)
(14)	1	0,1
(235)	0,1	1

• kemudian didapat cluster 1 = (1,4) ; cluster 2 = (2,3,5)  
dengan similarity terbesar = 0,1

histogram :



- 3 a. cohesion : seberapa dekat jarak antar objek dalam sebuah cluster  
 separation : seberapa jauh jarak (beda) objek-objek antar cluster

$$\begin{aligned}
 b). \quad TSS &= SSE + SSB \\
 \sum_{i=1}^K \sum_{x \in C_i} (x - c)^2 &= \sum_{i=1}^K \sum_{x \in C_i} (x - c_i)^2 + \sum_{i=1}^K |C_i| (c - c_i)^2 \\
 &= \sum_{i=1}^K \sum_{x \in C_i} (x - c_i)^2 + \sum_{i=1}^K \sum_{x \in C_i} (c - c_i)^2 \\
 &= \sum_{i=1}^K \sum_{x \in C_i} (x - c_i)^2 - 2 \sum_{i=1}^K \sum_{x \in C_i} (x - c_i)(c - c_i) + \sum_{i=1}^K \sum_{x \in C_i} (c - c_i)^2 \\
 &= \sum_{i=1}^K \sum_{x \in C_i} ((x - c_i) - (c - c_i))^2 \\
 &= \sum_{i=1}^K \sum_{x \in C_i} (x - c)^2
 \end{aligned}$$

- 4). kelompok (2) =  $(-1, 1, 1, -1), (-1, 1, 1, 1), (1, -1, 1, 1), (1, 1, -1, -1)$  dan  $(-1, 1, -1, 1)$

learning rate ( $\alpha$ ) = 0,5, radius neighborhood ( $\rho$ ) = 0, nilai inisialisasi bobot :  $w = \begin{bmatrix} -0,5 & 0,2 \\ 0,3 & -0,7 \\ -0,7 & 0,4 \\ 0,3 & -0,5 \end{bmatrix}$

$$D(j) = \sum_i (w_{ij} - x_i)^2$$

Step 0

$$D(1) = (-0,5 - (-1))^2 + (0,3 - 1)^2 + (-0,7 - (-1))^2 + (0,3 + 1)^2 = 5,32$$

$$D(2) = (0,2 - 1)^2 + (-0,7 - (-1))^2 + (0,4 - 1)^2 + (-0,5 + 1)^2 = 4,14 < 0$$

•  $D(1) > D(2)$  maka pilih  $j = 2$

$$w_{ij}^{(new)} = w_{ij}^{(old)} + \alpha [x_i - w_{ij}^{(old)}]$$

$$\begin{aligned}
 w_{12}^{(new)} &= w_{12}^{(old)} + 0,5 [x_1 - w_{12}^{(old)}] \\
 &= 0,2 + 0,5 (-1 - 0,2) = -0,1
 \end{aligned}$$

$$w_{22}^{(new)} = -0,7 + 0,5 (1 + 0,7) = 0,15$$

$$w_{32}^{(new)} = 0,4 + 0,5 (1 - 0,4) = 0,7$$

$$w_{42}^{(new)} = -0,5 + 0,5 (-1 + 0,5) = -0,75$$

$$w_{new} = \begin{bmatrix} -0,5 & -0,1 \\ 0,3 & 0,15 \\ -0,7 & 0,7 \\ 0,3 & -0,75 \end{bmatrix}$$



① Step = 1

vektor  $z = [-1, -1, 1, 1]$

$$D(1) = (-0,15+1)^2 + (0,3+1)^2 + (-0,7-1)^2 + (0,3-1)^2 = 5,32$$

$$D(2) = (-0,4+1)^2 + (0,15+1)^2 + (0,7-1)^2 + (-0,75-1)^2 = 9,84$$

•  $D(1) > D(2)$  maka  $j = 2$

$$\left. \begin{aligned} W_{12}(\text{new}) &= -0,4 + 0,15(-1+0,9) = -0,7 \\ W_{22}(\text{new}) &= 0,15 + 0,15(-1-0,15) = -0,43 \\ W_{32}(\text{new}) &= 0,7 + 0,15(1-0,7) = 0,85 \\ W_{42}(\text{new}) &= -0,75 + 0,15(1+0,75) = 0,13 \end{aligned} \right\} W_{\text{new}} = \begin{bmatrix} -0,15 & -0,7 \\ 0,3 & -0,43 \\ -0,7 & 0,85 \\ 0,3 & 0,13 \end{bmatrix}$$

② Step = 2

vektor  $z = (1, -1, 1, 1)$

$$D(1) = (-0,15-1)^2 + (0,3+1)^2 + (-0,7-1)^2 + (0,3-1)^2 = 7,32$$

$$D(2) = (-0,7-1)^2 + (0,33+1)^2 + (0,85-1)^2 + (0,13-1)^2 = 3,99$$

•  $D(1) > D(2)$  maka  $j = 2$

$$\left. \begin{aligned} W_{12}(\text{new}) &= (-0,7) + 0,15(1+0,7) = 0,15 \\ W_{22}(\text{new}) &= -0,43 + 0,15(-1+0,43) = -0,15 \\ W_{32}(\text{new}) &= 0,85 + 0,15(1-0,85) = 0,93 \\ W_{42}(\text{new}) &= 0,13 + 0,15(1-0,13) = 0,157 \end{aligned} \right\} W_{\text{new}} = \begin{bmatrix} -0,15 & 0,15 \\ 0,3 & -0,15 \\ -0,7 & 0,93 \\ 0,3 & 0,157 \end{bmatrix}$$

③ Step = 3

vektor  $z = (1, 1, -1, 1)$

$$D(1) = (-0,15-1)^2 + (0,3-1)^2 + (-0,7+1)^2 + (0,3+1)^2 = 4,12$$

$$D(2) = (0,15-1)^2 + (-0,15-1)^2 + (0,93+1)^2 + (0,157+1)^2 = 8,23$$

•  $D(1) < D(2)$  maka  $j = 1$

$$\left. \begin{aligned} W_{11}(\text{new}) &= -0,15 + 0,15(1+0,15) = 0,25 \\ W_{21}(\text{new}) &= 0,3 + 0,15(1-0,3) = 0,65 \\ W_{31}(\text{new}) &= -0,85 \\ W_{41}(\text{new}) &= -0,35 \end{aligned} \right\} W_{\text{new}} = \begin{bmatrix} 0,25 & 0,15 \\ 0,65 & -0,15 \\ -0,85 & 0,93 \\ -0,35 & 0,157 \end{bmatrix}$$

④ Step 4

vektor  $z = (-1, 1, -1, 1)$

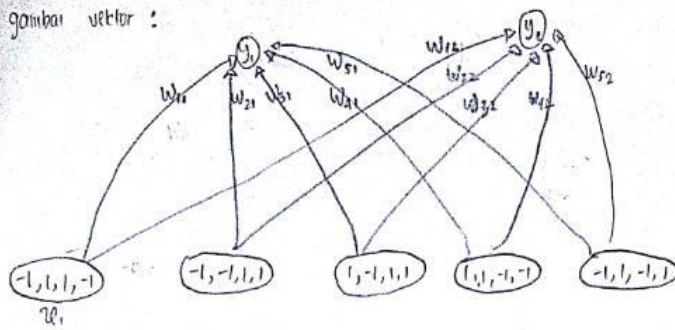
$$D(1) = (0,25+1)^2 + (0,65-1)^2 + (-0,85+1)^2 + (-0,35-1)^2 = 3,53$$

$$D(2) = (0,15+1)^2 + (-0,15-1)^2 + (0,93+1)^2 + (0,157-1)^2 = 6,55$$

•  $D(1) < D(2)$  maka  $j = 1$

$$\left. \begin{aligned} W_{11}(\text{new}) &= 0,25 + 0,15(-1-0,15) = -0,38 \\ W_{21}(\text{new}) &= 0,65 + 0,15(1-0,65) = 0,83 \\ W_{31}(\text{new}) &= -0,85 + 0,15(1+0,85) = -0,93 \\ W_{41}(\text{new}) &= -0,35 + 0,15(1+0,35) = 0,13 \end{aligned} \right\} W_{\text{new}} = \begin{bmatrix} -0,38 & 0,15 \\ 0,83 & -0,15 \\ -0,93 & 0,03 \\ 0,13 & 0,157 \end{bmatrix}$$

gambar vektor :



5) a. pada principal component, memilih eigenvector dengan eigenvalue tertinggi, maka akan dipilih yang ke (1) ->

$$\text{eigenvalue} = 9,279$$

$$\text{eigenvector} = \begin{bmatrix} 0,1222 \\ 0,9925 \end{bmatrix}$$

b). rata-rata data -1 = 7      Data Adjust = tiap data - rata-rata  
rata-rata data -2 = 6      =  $x - \bar{x}$

$$\text{Data Adjust} = \begin{bmatrix} 1 & 0 \\ 2 & -3 \\ -3 & 1 \\ -2 & -4 \\ 0 & 4 \\ -2 & 2 \end{bmatrix}$$

$$\text{Row Data Adjust} = (\text{Data Adjust})^T$$

$$= \begin{bmatrix} 1 & 2 & -3 & -2 & 0 & -2 \\ 0 & -3 & 1 & -4 & 4 & 2 \end{bmatrix}_{6 \times 2}$$

$$\text{Row Feature Vector} = (\text{eigenvectors})^T$$

$$= \begin{bmatrix} 0,1222 & 0,9925 \end{bmatrix}$$

1x2      2x6

$$\text{Final Data} = \text{Row Feature Vector} \times \text{Row Data Adjust}$$

$$= \begin{bmatrix} 0,1222 & 0,9925 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 & -2 & 0 & -2 \\ 0 & -3 & 1 & -4 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 0,1222 & -2,7325 & 0,6257 & -4,2144 \\ 3,97 & 1,2294 \end{bmatrix}$$

$$\text{hasil akhir} = (\text{Final Data})^T = \begin{bmatrix} 0,1222 \\ -2,7325 \\ 0,6257 \\ 3,97 \\ 2,1229 \end{bmatrix}$$

-2