

1.) a) Perbedaan clustering for understanding dan clustering for utility:

- Pada konteks clustering for understanding, cluster adalah kelas potensial dan analisis cluster merupakan studi tentang teknik untuk menemukan kelas secara otomatis. Contohnya misal pada dunia bisnis, seorang pebisnis akan mengumpulkan informasi tentang pelanggannya saat ini yang potensial. Dari informasi yang didapatkan tersebut, secara langsung pebisnis dapat mengelompokkan pelanggannya menjadi beberapa kelompok yang dapat digunakan sebagai analisis tambahan dalam kegiatan pemasaran.
- Sedangkan pada konteks clustering for utility, analisis cluster merupakan studi tentang teknik untuk menemukan atau menentukan objek data yang mewakili objek lain di dalam suatu cluster secara representatif. Contohnya yaitu saat proses menentukan tetangga yang paling dekat.

b) Ya, karena semua titik yang memiliki bentuk sama akan menjadi satu cluster.

2.)

	P ₁	P ₂	P ₃	P ₄
P ₁	1,00	0,10	0,41	0,55
P ₂	0,10	1,00	0,64	0,47
P ₃	0,41	0,64	1,00	0,44
P ₄	0,55	0,47	0,44	1,00

 menggunakan complete linkage AHC

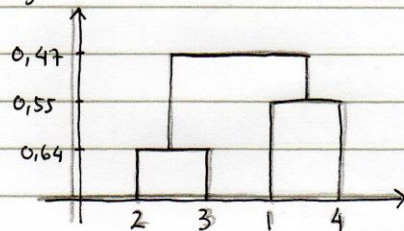
- gabungkan P₂ dan P₃, karena punya similaritas paling besar

	P ₁	(P ₂ , P ₃)	P ₄	
P ₁	1,00	0,41	0,55	$D(P_1, (P_2, P_3)) = \max(D(P_1, P_2), D(P_1, P_3)) = \max(0,10; 0,41) = 0,41$
(P ₂ , P ₃)	0,41	1,00	0,47	$D((P_2, P_3), P_4) = \max(D(P_2, P_4), D(P_3, P_4)) = \max(0,47; 0,44) = 0,47$
P ₄	0,55	0,47	1,00	

- gabungkan P₁ dan P₄

	(P ₁ , P ₄)	(P ₂ , P ₃)	
(P ₁ , P ₄)	1,00	0,47	$D((P_1, P_4), (P_2, P_3)) = \max(D(P_1, P_2), D(P_1, P_3), D(P_4, P_2), D(P_4, P_3)) = \max(0,10; 0,41; 0,47; 0,44) = 0,47$
(P ₂ , P ₃)	0,47	1,00	

- Dendrogram



- 4) a) True
b) False
c) False
d) 10

5) a) Ketika dimensi meningkat, kinerja classifier juga meningkat hingga jumlah fitur optimal tercapai. Apabila ~~ter~~ telah optimal dan kemudian ditambahkan fitur lagi maka performanya akan semakin menurun. Solusinya jika mau menambahkan fitur maka harus menambahkan data juga sehingga tidak terjadi overfitting.

b.) \rightarrow eigenvalue = $\begin{pmatrix} 9,274 \\ 4,326 \end{pmatrix}$

\rightarrow eigenvector = $\begin{pmatrix} 0,1222 & -0,9925 \\ 0,9925 & 0,1222 \end{pmatrix} \Rightarrow$ diurutkan secara descending berdasarkan eigenvalue

\rightarrow karena akan direduksi menjadi 1 dimensi, maka ambil eigenvalue yang paling besar yaitu 9,274

\rightarrow row feature vector : eigenvector^T
= $(0,1222 \quad 0,9925)$

\rightarrow Final data = row feature vector \times row data adjust

$$= (0,1222 \quad 0,9925) \begin{pmatrix} 1 & 2 & -3 & -2 & 0 & 2 \\ 0 & -3 & 1 & -4 & 4 & 2 \end{pmatrix}$$

$$= (0,1222 \quad -2,7331 \quad 0,6259 \quad -4,2144 \quad 3,97 \quad 2,2294)$$

Data1 2 3 4 5 6

\rightarrow Hasil reduksi 1 dimensi

\times

$$\begin{pmatrix} 0,1222 \\ -2,7331 \\ 0,6259 \\ -4,2144 \\ 3,97 \\ 2,2294 \end{pmatrix}$$