

PEMBELAJARAN MESIN - A
TUGAS HIERARCHICAL CLUSTERING



Dosen Pengampu:

Dini Adni Navastara, S.Kom, M.Sc.
NIP. 198510172015042001

Disusun oleh:

Faa'iz Haikal Hilmi (5025221219)

**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
2025**

Gunakan *Agglomerative Clustering* dengan pendekatan *single link*, *complete link* dan *group average link* untuk data yang ditunjukkan pada *distance matrix* (matriks jarak) berikut ini. Gambarkan dendrogramnya!

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
P1	0									
P2	1,41	0								
P3	6,40	6,71	0							
P4	7,28	6,71	4,24	0						
P5	6,32	7,07	2,24	6,40	0					
P6	9,85	9,43	5,10	2,83	7,28	0				
P7	2,00	1,41	5,39	5,39	6,00	8,06	0			
P8	8,06	7,28	5,66	1,41	7,81	3,16	6,08	0		
P9	4,47	5,10	2,24	5,39	2,00	7,00	4,00	6,71	0	
P10	4,47	4,24	3,00	3,00	4,47	5,39	2,83	4,12	2,83	0

1. Pendekatan *single link*

Untuk pendekatan *single link*, jarak antara *cluster* A dan B adalah

$$d(A,B) = \min_{\{x \in A, y \in B\}} d(x,y)$$

• *Cluster* P1 dan P2

	P1,P2	P3	P4	P5	P6	P7	P8	P9	P10
P1,P2	0								
P3	6,40	0							
P4	6,71	4,24	0						
P5	6,32	2,24	6,40	0					
P6	9,43	5,10	2,83	7,28	0				
P7	1,41	5,39	5,39	6,00	8,06	0			
P8	7,28	5,66	1,41	7,81	3,16	6,08	0		
P9	4,47	2,24	5,39	2,00	7,00	4,00	6,71	0	
P10	4,24	3,00	3,00	4,47	5,39	2,83	4,12	2,83	0

• *Cluster* P7 dan P1,P2

	P1,P2,P7	P3	P4	P5	P6	P8	P9	P10
P1,P2,P7	0							
P3	5,39	0						
P4	5,39	4,24	0					
P5	6,00	2,24	6,40	0				
P6	8,06	5,10	2,83	7,28	0			
P8	6,08	5,66	1,41	7,81	3,16	0		
P9	4,00	2,24	5,39	2,00	7,00	6,71	0	
P10	2,83	3,00	3,00	4,47	5,39	4,12	2,83	0

● *Cluster P4 dan P8*

	P1,P2,P7	P3	P4,P8	P5	P6	P9	P10
P1,P2,P7	0						
P3	5,39	0					
P4,P8	5,39	4,24	0				
P5	6,00	2,24	6,40	0			
P6	8,06	5,10	2,83	7,28	0		
P9	4,00	2,24	5,39	2,00	7,00	0	
P10	2,83	3,00	3,00	4,47	5,39	2,83	0

● *Cluster P5 dan P9*

	P1,P2,P7	P3	P4,P8	P5,P9	P6	P10
P1,P2,P7	0					
P3	5,39	0				
P4,P8	5,39	4,24	0			
P5,P9	4,00	2,24	5,39	0		
P6	8,06	5,10	2,83	7,00	0	
P10	2,83	3,00	3,00	2,83	5,39	0

● *Cluster P3 dan P5,P9*

	P1,P2,P7	P3,P5,P9	P4,P8	P6	P10
P1,P2,P7	0				
P3,P5,P9	4,00	0			
P4,P8	5,39	4,24	0		
P6	8,06	5,10	2,83	0	
P10	2,83	2,83	3,00	5,39	0

- Cluster P6 dan P4,P8

	P1,P2,P7	P3,P5,P9	P4,P6,P8	P10
P1,P2,P7	0			
P3,P5,P9	4,00	0		
P4,P6,P8	5,39	4,24	0	
P10	2,83	2,83	3,00	0

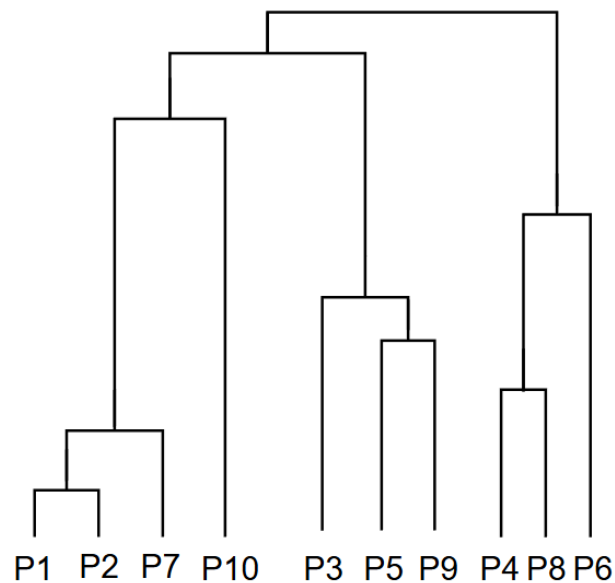
- Cluster P10 dan P1,P2,P7

	P1,P2,P7,P10	P3,P5,P9	P4,P6,P8
P1,P2,P7,P10	0		
P3,P5,P9	2,83	0	
P4,P6,P8	3,00	4,24	0

- Cluster P1,P2,P7,P10 dan P3,P5,P9

	P1,P2,P7,P10,P3,P5,P9	P4,P6,P8
P1,P2,P7,P10,P3,P5,P9	0	
P4,P6,P8	3,00	0

Jadi, dapat digambarkan hierarki *clustering* di atas menggunakan dendrogram sebagai berikut.



2. Pendekatan *complete link*

Untuk pendekatan *complete link*, jarak antara *cluster* A dan B adalah

$$d(A,B) = \max_{\{x \in A, y \in B\}} d(x,y)$$

• *Cluster* P1 dan P2

	P1,P2	P3	P4	P5	P6	P7	P8	P9	P10
P1,P2	0								
P3	6,71	0							
P4	7,28	4,24	0						
P5	7,07	2,24	6,40	0					
P6	9,85	5,10	2,83	7,28	0				
P7	2,00	5,39	5,39	6,00	8,06	0			
P8	7,67	5,66	1,41	7,81	3,16	6,08	0		
P9	8,06	2,24	5,39	2,00	7,00	4,00	6,71	0	
P10	4,47	3,00	3,00	4,47	5,39	2,83	4,12	2,83	0

• *Cluster* P4 dan P8

	P1,P2	P3	P4,P8	P5	P6	P7	P9	P10
P1,P2	0							
P3	6,71	0						
P4,P8	8,06	5,66	0					
P5	7,07	2,24	7,81	0				
P6	9,85	5,10	3,16	7,28	0			
P7	2,00	5,39	6,08	6,00	8,06	0		
P9	5,10	2,24	6,71	2,00	7,00	4,00	0	
P10	4,47	3,00	4,12	4,47	5,39	2,83	2,83	0

● *Cluster P7 dan P1,P2*

	P1,P2,P7	P3	P4,P8	P5	P6	P9	P10
P1,P2,P7	0						
P3	6,71	0					
P4,P8	8,06	5,66	0				
P5	7,07	2,24	7,81	0			
P6	9,85	5,10	3,16	7,28	0		
P9	5,10	2,24	6,71	2,00	7,00	0	
P10	4,47	3,00	4,12	4,47	5,39	2,83	0

● *Cluster P5 dan P9*

	P1,P2,P7	P3	P4,P8	P5,P9	P6	P10
P1,P2,P7	0					
P3	6,71	0				
P4,P8	8,06	5,66	0			
P5,P9	7,07	2,24	7,81	0		
P6	9,85	5,10	3,16	7,28	0	
P10	4,47	3,00	4,12	4,47	5,39	0

● *Cluster P3 dan P5,P9*

	P1,P2,P7	P3,P5,P9	P4,P8	P6	P10
P1,P2,P7	0				
P3,P5,P9	7,07	0			
P4,P8	8,06	7,81	0		
P6	9,85	7,28	3,16	0	
P10	4,47	4,47	4,12	5,39	0

- *Cluster* P10 dan P4,P8

	P1,P2,P7	P3,P5,P9	P4,P8,P10	P6
P1,P2,P7	0			
P3,P5,P9	7,07	0		
P4,P8,P10	8,06	7,07	0	
P6	9,85	7,28	5,39	0

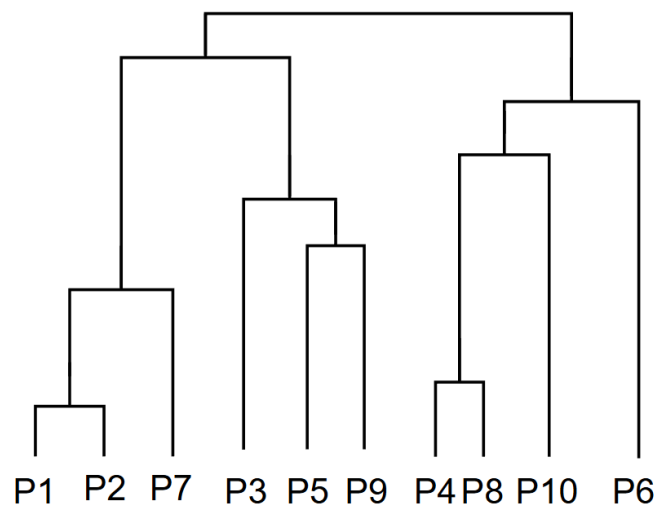
- *Cluster* P6 dan P4,P8,P10

	P1,P2,P7	P3,P5,P9	P4,P8,P6,P10
P1,P2,P7	0		
P3,P5,P9	7,07	0	
P4,P8,P6,P10	9,85	7,28	0

- *Cluster* P1,P2,P7 dan P3,P5,P9

	P1,P2,P3,P5,P7,P9	P4,P8,P6,P10
P1,P2,P3,P5,P7,P9	0	
P4,P8,P6,P10	9,85	0

Jadi, dapat digambarkan hierarki *clustering* di atas menggunakan dendrogram sebagai berikut.



3. Pendekatan *group average link*

Untuk pendekatan *group average link*, jarak antara *cluster* A dan B adalah

$$d(A, B) = \frac{1}{|A||B|} \sum_{\{x \in A\}} \sum_{\{y \in B\}} d(x, y)$$

- *Cluster* P1 dan P2 (**C₁₂**)

- $d(P3, C_{12}) = \frac{1}{2}[d(P3, P1) + d(P3, P2)] = 6,555$
- $d(P4, C_{12}) = \frac{1}{2}[d(P4, P1) + d(P4, P2)] = 6,995$
- $d(P5, C_{12}) = \frac{1}{2}[d(P5, P1) + d(P5, P2)] = 6,695$
- $d(P6, C_{12}) = \frac{1}{2}[d(P6, P1) + d(P6, P2)] = 9,64$
- $d(P7, C_{12}) = \frac{1}{2}[d(P7, P1) + d(P7, P2)] = 1,705$
- $d(P8, C_{12}) = \frac{1}{2}[d(P8, P1) + d(P8, P2)] = 7,67$
- $d(P9, C_{12}) = \frac{1}{2}[d(P9, P1) + d(P9, P2)] = 4,785$
- $d(P10, C_{12}) = \frac{1}{2}[d(P10, P1) + d(P10, P2)] = 4,355$

	C ₁₂	P3	P4	P5	P6	P7	P8	P9	P10
C ₁₂	0								
P3	6,555	0							
P4	6,995	4,24	0						
P5	6,695	2,24	6,40	0					
P6	9,64	5,10	2,83	7,28	0				
P7	1,705	5,39	5,39	6,00	8,06	0			
P8	7,67	5,66	1,41	7,81	3,16	6,08	0		
P9	4,785	2,24	5,39	2,00	7,00	4,00	6,71	0	
P10	4,355	3,00	3,00	4,47	5,39	2,83	4,12	2,83	0

- *Cluster* P4 dan P8 (**C₄₈**)

- $d(C_{12}, C_{48}) = \frac{1}{2}[d(C_{12}, P4) + d(C_{12}, P8)] = 7,3325$
- $d(P3, C_{48}) = \frac{1}{2}[d(P3, P4) + d(P3, P8)] = 4,95$
- $d(P5, C_{48}) = \frac{1}{2}[d(P5, P4) + d(P5, P8)] = 7,105$
- $d(P6, C_{48}) = \frac{1}{2}[d(P6, P4) + d(P6, P8)] = 2,995$
- $d(P7, C_{48}) = \frac{1}{2}[d(P7, P4) + d(P7, P8)] = 5,735$
- $d(P9, C_{48}) = \frac{1}{2}[d(P9, P4) + d(P9, P8)] = 6,05$

$$- d(P10, C_{48}) = \frac{1}{2}[d(P10, P4) + d(P10, P8)] = 3,56$$

	C ₁₂	P3	C ₄₈	P5	P6	P7	P9	P10
C ₁₂	0							
P3	6,555	0						
C ₄₈	7,3325	4,95	0					
P5	6,695	2,24	7,105	0				
P6	9,64	5,10	2,995	7,28	0			
P7	1,705	5,39	5,735	6,00	8,06	0		
P9	4,785	2,24	6,05	2,00	7,00	4,00	0	
P10	4,355	3,00	3,56	4,47	5,39	2,83	2,83	0

• Cluster P7 dan C₁₂ (C₁₂₇)

$$\begin{aligned}
 - d(P3, C_{127}) &= \frac{1}{2}[d(P3, C_{12}) + d(P3, P7)] = 5,9725 \\
 - d(C_{48}, C_{127}) &= \frac{1}{2}[d(C_{48}, C_{12}) + d(C_{48}, P7)] = 6,53375 \\
 - d(P5, C_{127}) &= \frac{1}{2}[d(P5, C_{12}) + d(P5, P7)] = 6,3475 \\
 - d(P6, C_{127}) &= \frac{1}{2}[d(P6, C_{12}) + d(P6, P7)] = 8,85 \\
 - d(P9, C_{127}) &= \frac{1}{2}[d(P9, C_{12}) + d(P9, P7)] = 4,3925 \\
 - d(P10, C_{127}) &= \frac{1}{2}[d(P10, C_{12}) + d(P10, P7)] = 3,5925
 \end{aligned}$$

	C ₁₂₇	P3	C ₄₈	P5	P6	P9	P10
C ₁₂₇	0						
P3	5,9725	0					
C ₄₈	6,53375	4,95	0				
P5	6,3475	2,24	7,105	0			
P6	8,85	5,10	2,995	7,28	0		
P9	4,3925	2,24	6,05	2,00	7,00	0	
P10	3,5925	3,00	3,56	4,47	5,39	2,83	0

● *Cluster P5 dan P9 (C₅₉)*

- $d(C_{127}, C_{59}) = \frac{1}{2}[d(C_{127}, P5) + d(C_{127}, P9)] = 5,37$
- $d(P3, C_{59}) = \frac{1}{2}[d(P3, P5) + d(P3, P9)] = 2,24$
- $d(C_{48}, C_{59}) = \frac{1}{2}[d(C_{48}, P5) + d(C_{48}, P9)] = 6,5775$
- $d(P6, C_{59}) = \frac{1}{2}[d(P6, P5) + d(P6, P9)] = 8,85$
- $d(P10, C_{59}) = \frac{1}{2}[d(P10, P5) + d(P10, P9)] = 3,5925$

	C ₁₂₇	P3	C ₄₈	C ₅₉	P6	P10
C ₁₂₇	0					
P3	5,9725	0				
C ₄₈	6,53375	4,95	0			
C ₅₉	5,37	2,24	6,5775	0		
P6	8,85	5,10	2,995	7,14	0	
P10	3,5925	3,00	3,56	3,65	5,39	0

● *Cluster P3 dan C₅₉ (C₃₅₉)*

- $d(C_{127}, C_{359}) = \frac{1}{2}[d(C_{127}, P3) + d(C_{127}, C_{59})] = 5,67125$
- $d(C_{48}, C_{359}) = \frac{1}{2}[d(C_{48}, P3) + d(C_{48}, C_{59})] = 5,76375$
- $d(P6, C_{359}) = \frac{1}{2}[d(P6, P3) + d(P6, C_{59})] = 6,12$
- $d(P10, C_{359}) = \frac{1}{2}[d(P10, P3) + d(P10, C_{59})] = 3,325$

	C ₁₂₇	C ₃₅₉	C ₄₈	P6	P10
C ₁₂₇	0				
C ₃₅₉	5,67125	0			
C ₄₈	6,53375	5,76375	0		
P6	8,85	6,12	2,995	0	
P10	3,5925	3,325	3,56	5,39	0

- *Cluster* P6 dan C₄₈ (**C₄₆₈**)

- $d(C_{127}, C_{468}) = \frac{1}{2}[d(C_{127}, C_{48}) + d(C_{127}, P6)] = 7,691875$
- $d(C_{359}, C_{468}) = \frac{1}{2}[d(C_{359}, C_{48}) + d(C_{359}, P6)] = 5,941875$
- $d(P10, C_{468}) = \frac{1}{2}[d(P10, C_{48}) + d(P10, P6)] = 4,475$

	C ₁₂₇	C ₃₅₉	C ₄₆₈	P10
C ₁₂₇	0			
C ₃₅₉	5,67125	0		
C ₄₆₈	7,691875	5,941875	0	
P10	3,5925	3,325	4,475	0

- *Cluster* P10 dan C₃₅₉ (**C₃₅₉₁₀**)

- $d(C_{127}, C_{35910}) = \frac{1}{2}[d(C_{127}, C_{359}) + d(C_{127}, P10)] = 4,631875$
- $d(C_{468}, C_{35910}) = \frac{1}{2}[d(C_{468}, C_{359}) + d(C_{468}, P10)] = 5,2084375$

	C ₁₂₇	C ₃₅₉₁₀	C ₄₆₈
C ₁₂₇	0		
C ₃₅₉₁₀	4,631875	0	
C ₄₆₈	7,691875	5,2084375	0

- *Cluster* C₁₂₇ dan C₃₅₉₁₀ (**C₁₂₃₅₇₉₁₀**)

- $d(C_{468}, C_{12357910}) = \frac{1}{2}[d(C_{468}, C_{127}) + d(C_{468}, C_{35910})] = 6,45015625$

	C ₁₂₃₅₇₉₁₀	C ₄₆₈
C ₁₂₃₅₇₉₁₀	0	
C ₄₆₈	6,45015625	0

Jadi, dapat digambarkan hierarki *clustering* di atas menggunakan dendogram sebagai berikut.

