

Machine Learning com Python

Prof. Luciano Galdino

Agrupamento Hierárquico (Agglomerative Hierarchical Clustering)

Aplicado em problemas de aprendizagem não supervisionada.

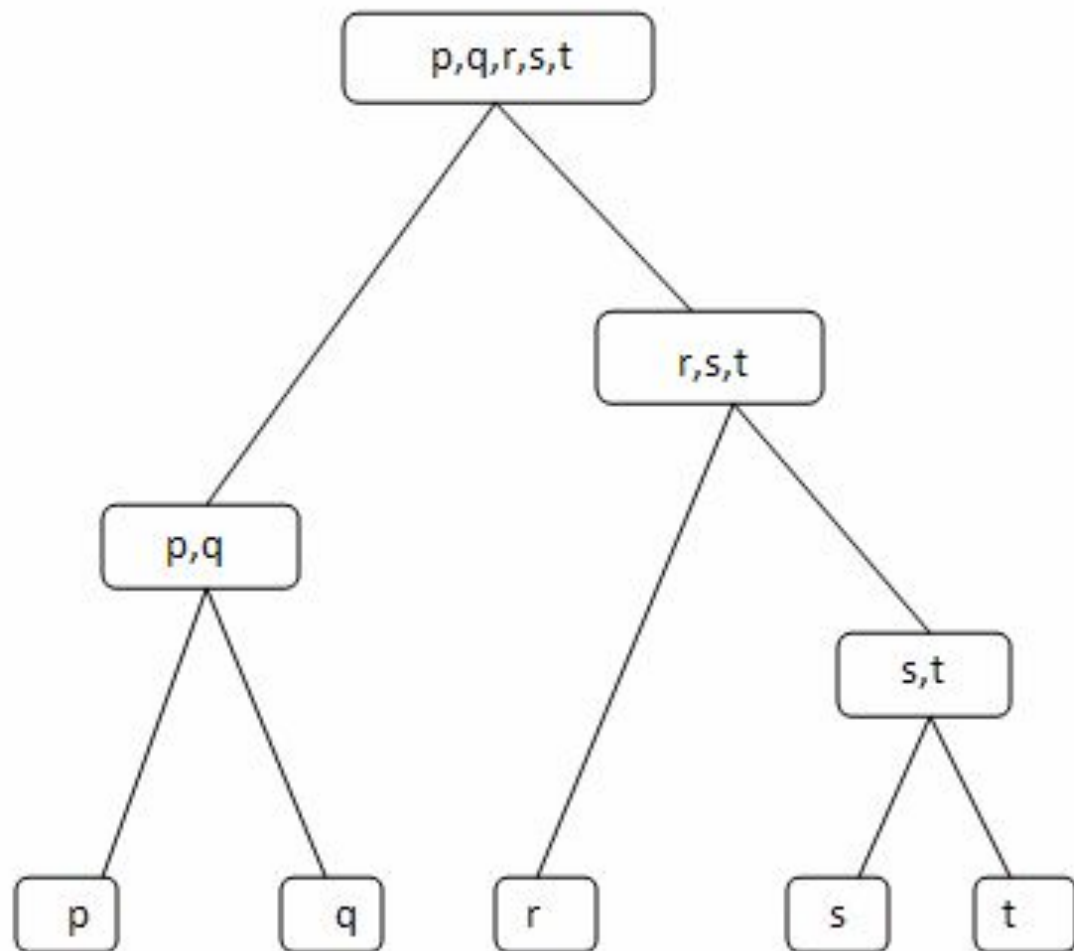
Busca construir uma hierarquia de grupos, de forma aninhada, mesclando-os ou dividindo-os sucessivamente.

Esquematizada como forma de uma árvore e representada graficamente através de um dendrograma.

A raiz da árvore reúne todas as amostras, sendo as folhas com apenas uma amostra.



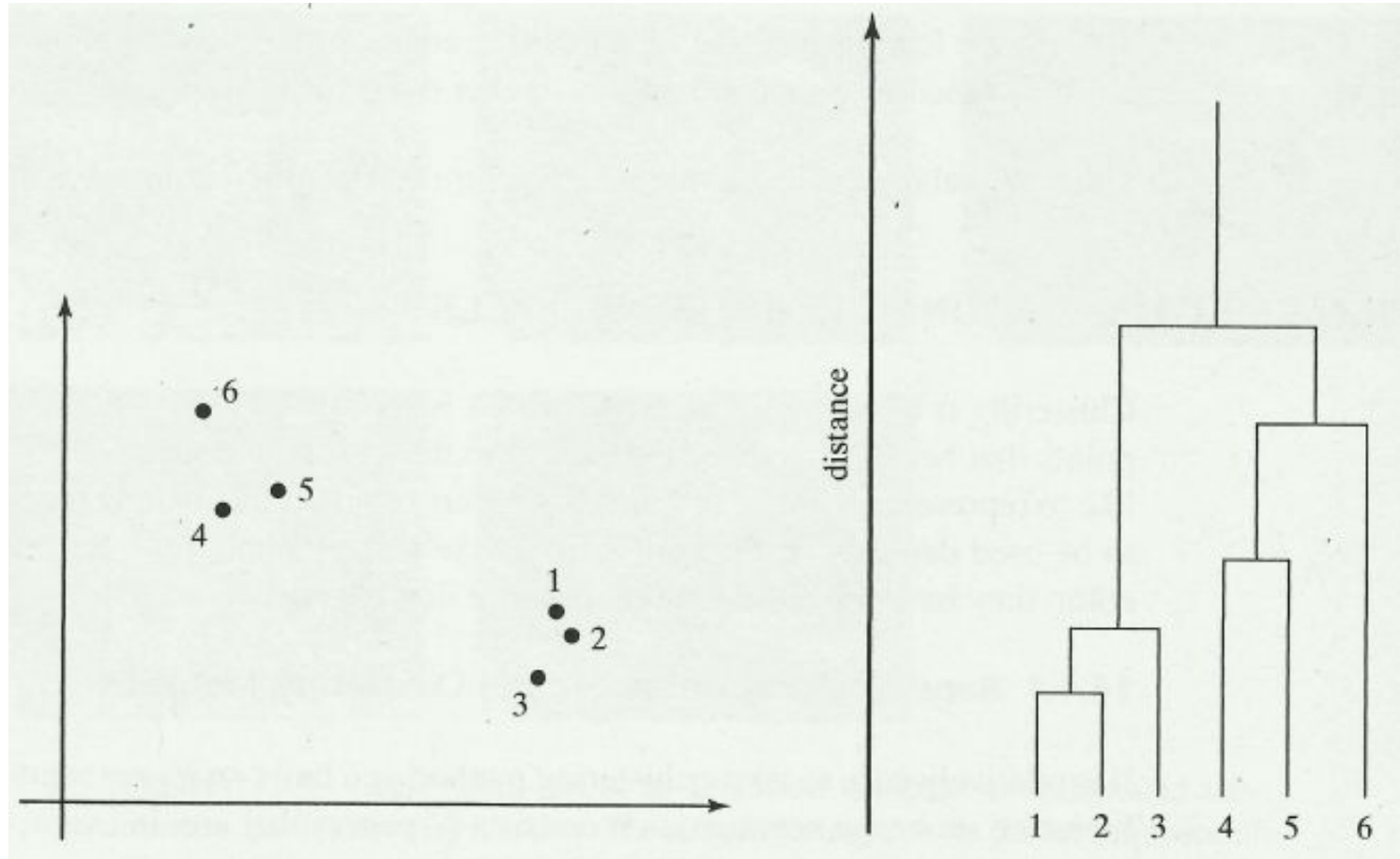
Agglomerative
methods



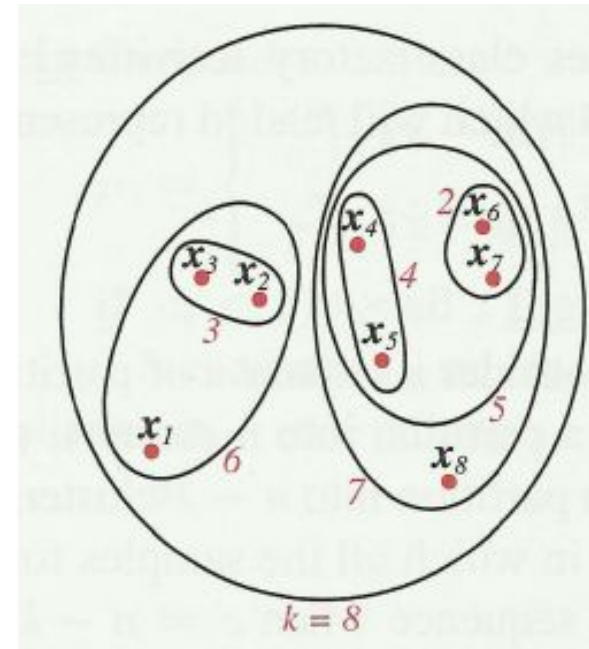
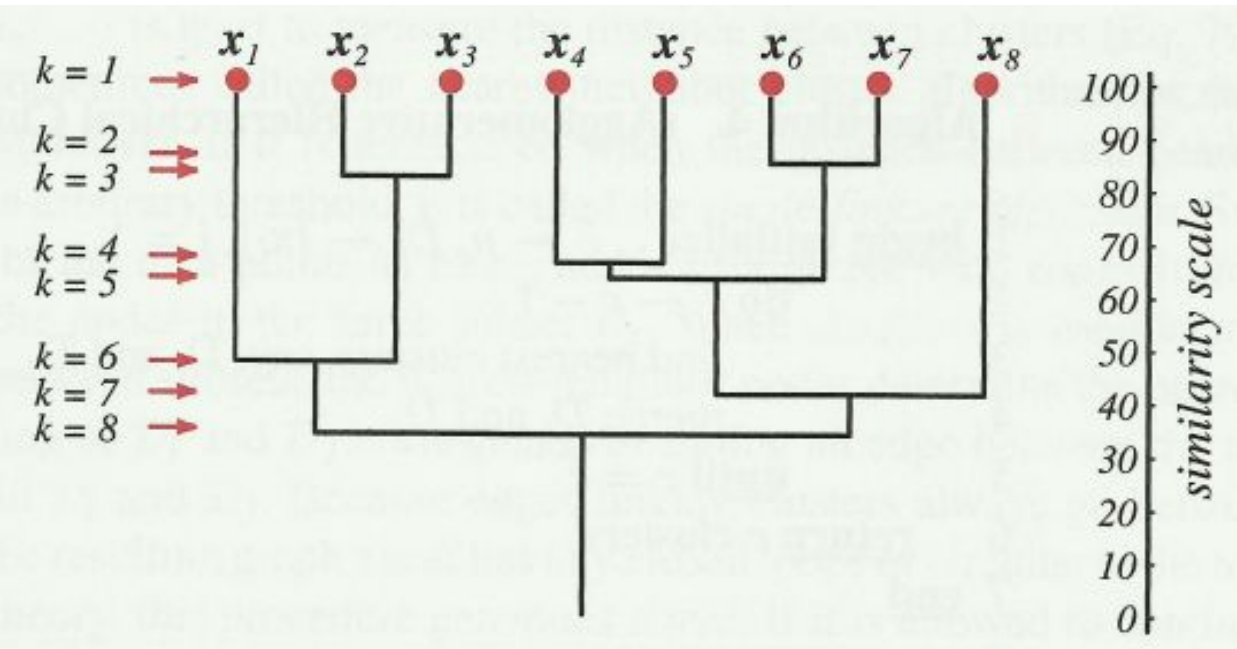
Divisive
methods



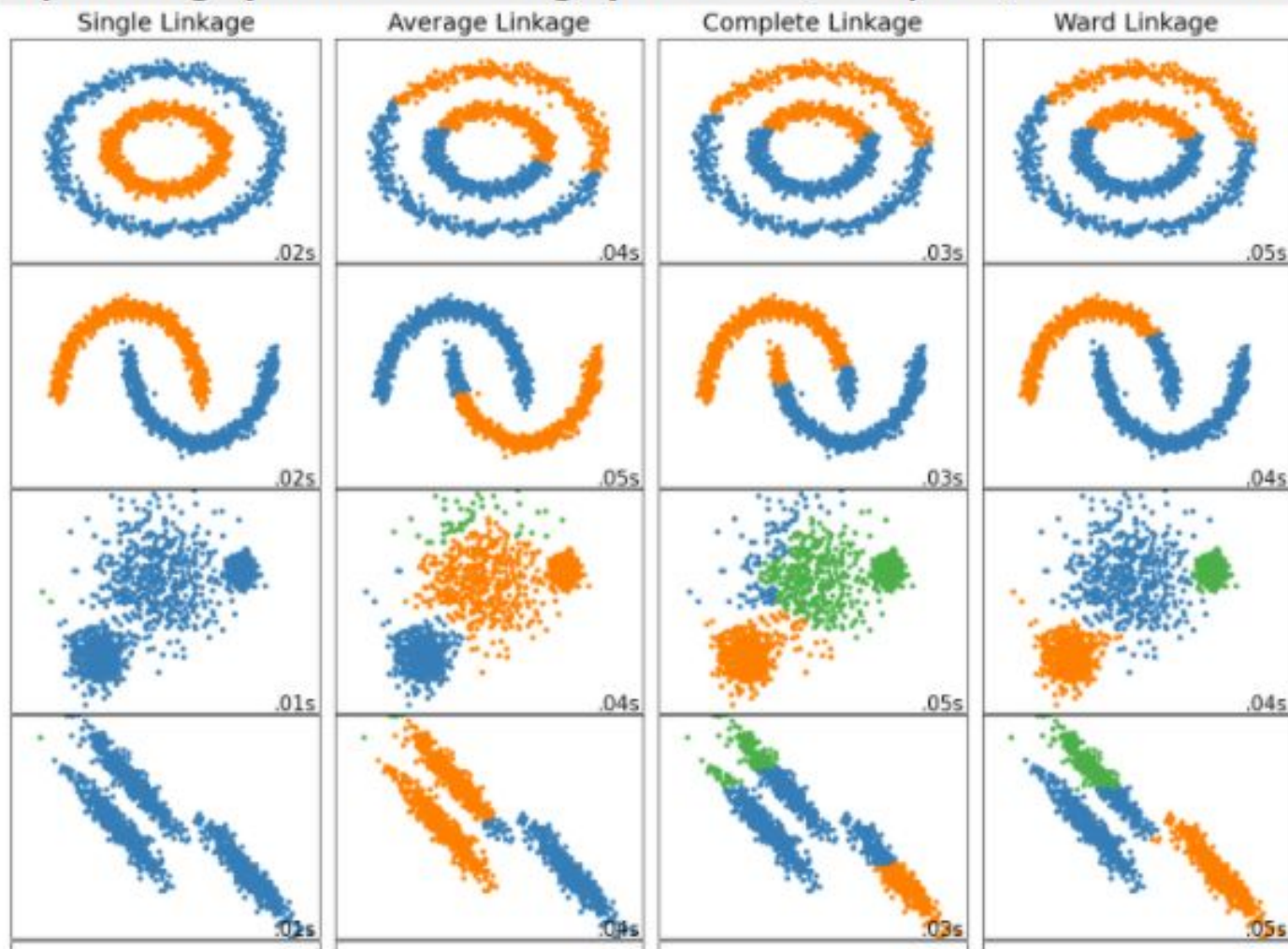
Dendrograma por distância



Dendograma por similaridade



Tipo de ligação diferente: Ligação de ala, completa, média e única



TIPOS DIFERENTES DE LIGAÇÕES

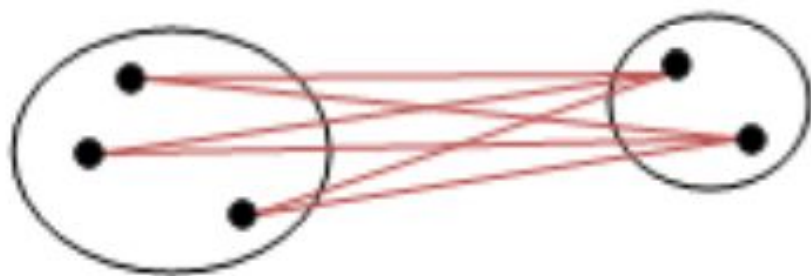
Ligação única (single): a distância entre dois clusters é a mínima distância entre observações de pares de clusters.

Ligação completa (complete): a distância entre dois clusters é a máxima distância entre observações de pares de clusters.

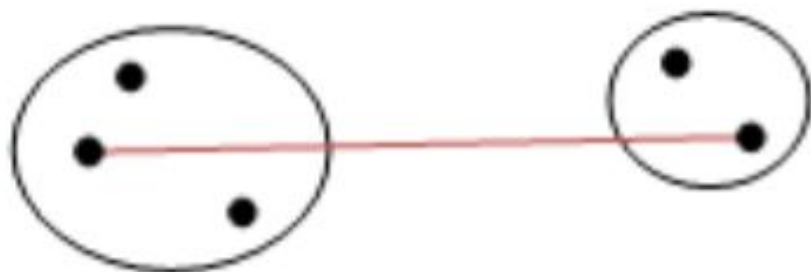
Ligação média (average): a distância clusters é a distância média entre cada ponto em um cluster para cada particular em outro.

Ward: A distância entre clusters é a soma das diferenças quadradas em todos os clusters.

Average linkage



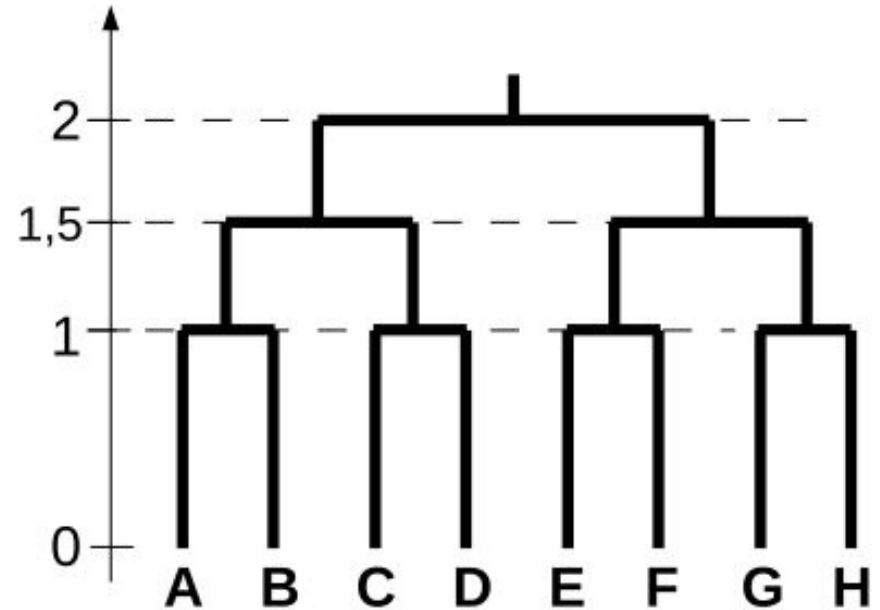
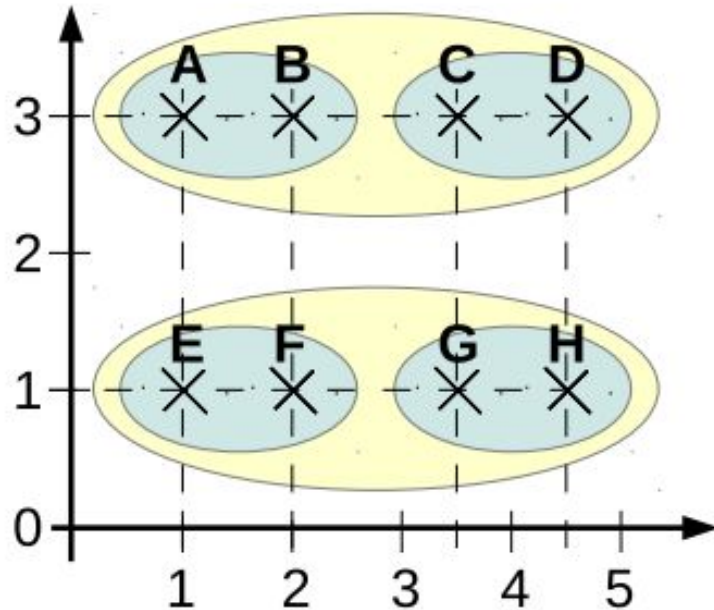
Complete linkage



Single linkage



Single Linkage (vizinhos mais próximos)



Complete Linkage (vizinhos mais distantes)

