



Fundamentos de IA y Machine Learning

Lección 5: Aprendizaje no supervisado -
Clustering

Aprendizaje no supervisado - Clustering

CASO PRÁCTICO

Vamos a dividir el caso práctico en tres problemas similares a los vistos en clase.

Problema I

Se pide agrupar un total de 8 patrones bidimensionales en tres *clusters* ($k = 3$). Los patrones son los siguientes: A1 (2,10), A2(2,5), A3(8,4), A4(5,8), A5(7,5), A6(6,4), A7(1,2) y A8(4,9). Los centroides iniciales son los puntos A1, A3 y A4. La métrica de distancia utilizada será la distancia euclídea.

Se pide:

1. Representar los *clusters* creados y la posición de los centroides después de cada iteración.
2. El valor de la métrica *SSE*.
3. Comparar los resultados con los obtenidos en el ejemplo de la sección 2.3. del manual de la lección 5.
4. ¿Á que se debe la diferencia de resultados?

Problema II

Dada la siguiente matriz de distancias entre cinco patrones:

	A	B	C	D	E
A	0	1	2	9	10
B		0	3	7	5
C			0	4	6
D				0	8
E					0

Se pide:

1. Aplicar un clustering jerárquico con el método de enlace simple.
2. Aplicar un clustering jerárquico con el método de enlace completo.
3. ¿Se podría usar el método de enlace medio con estos datos? Justifica la respuesta.

Problema III

Dados los mismos puntos del Problema I.

Se pide:

1. Aplicar DBSCAN considerando $M = 3$ y $\epsilon = \sqrt{2}$.
2. Aplicar DBSCAN considerando $M = 3$ y $\epsilon = \sqrt{10}$.
3. ¿En qué afecta el cambio del parámetro ϵ ?