

## EJERCICIOS DE CLUSTERING DE PATRONES

1. La tabla adjunta contiene 8 casos bidimensionales a partir de los cuales se trata de agrupar los diferentes ejemplos en tres grupos. ¿Cómo quedarían clasificados estos ejemplos de acuerdo al algoritmo K-means? Detalla cada uno de los pasos que haría aplicados a este ejemplo. ¿Se te ocurre otro algoritmo de aprendizaje automático que podrías usar para este caso? ¿Cuál y por qué?

Caso	X1	X2
1	1	1
2	2	4
3	3	2
4	3	5
5	4	4
6	4	7
7	6	4
8	6	6

2. Supón un conjunto S consistente en 6 puntos. Sean estos puntos los puntos  $a=(0,0)$ ,  $b=(8,0)$ ,  $c=(16,0)$ ,  $d=(0,6)$ ,  $e=(8,6)$ ,  $f=(16,6)$ . Supón que aplicamos el algoritmo k-means a este conjunto, usando como criterio de similitud la distancia métrica euclídea (es decir, la distancia en línea recta entre dos puntos en un espacio de dos dimensiones) Supón a su vez que en caso de igualdad de distancias entre dos puntos se resolverá a favor del centroide del clúster que esté situado más a la izquierda/abajo.

Sean estas dos definiciones:

- Una configuración de k-semillas es un subconjunto de k puntos pertenecientes a S que forman los primeros centros de los clústers.
- Una k-partición es una partición de S en k subconjuntos no-vacíos.

Claramente cualquier k-partición induce de la misma manera un conjunto de k centroides. Se dice que una k-partición es estable si una repetición de iteraciones del algoritmo k-means con los centroides inducidos la deja invariable.

Rellena la siguiente tabla, diciendo si cada una de las siguientes particiones son o no estables.

3-partición	¿Estable?
{a,b,e}, {c,e}, {f}	
{a,b}, {d,e}, {c,f}	
{a,d}, {b,e}, {c,f}	
{a}, {d}, {b,c,e,f}	
{a,b}, {d}, {c,e,f}	
{a,b,d}, {c}, {e,f}	