Sistemas Inteligentes
Cuestiones y ejercicios del bloque 2, tema 4
Aprendizaje no supervisado: algoritmo k-medias

Escola Tècnica Superior d'Informàtica
Dep. de Sistemes Informàtics i Computació
Universitat Politècnica de València

18 de noviembre de 2024

1. Cuestiones

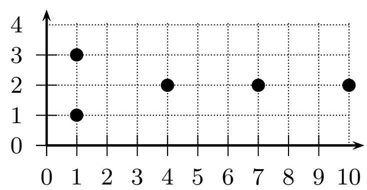
1 $\boxed{\text{C}}$ Durante la ejecución del algoritmo *c-means* se obtiene la siguiente partición en dos grupos $X_1 = \{(0,0),(1,0),(2,1)\}$ y $X_2 = \{(0,1),(1,2),(2,2)\}$. Calcula la Suma de Errores Cuadráticos (SEC) de dicha partición.

A) 8/3
B) 4/3
C) 16/3
D) 5/3

2 $\boxed{\text{B}}$ Indica cuál de las siguientes afirmaciones con respecto a la Suma de Errores Cuadráticos (SEC) es la correcta:

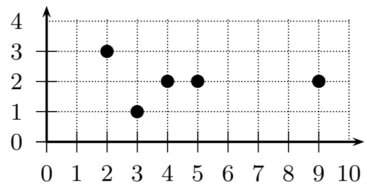
A) La versión de Duda-Hart del *c-means* garantiza un mínimo global de la SEC.
B) No existe ningún algoritmo de coste polinómico que garantice un mínimo global de la SEC.
C) La versión de Duda-Hart del *c-means* garantiza una SEC nula.
D) La versión "popular" del *c-means* garantiza un mínimo local de la SEC.

3 $\boxed{\text{B}}$ La menor suma de errores cuadráticos con la que los puntos de la figura a la derecha pueden agruparse en dos clústers es:



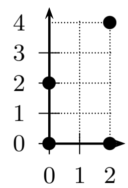
A) Menor que 10
B) Entre 10 y 15
C) Entre 15 y 20
D) Mayor que 20

4 $\boxed{\text{B}}$ La menor suma de errores cuadráticos con la que los puntos de la figura a la derecha pueden agruparse en dos clústers es:



A) Menor que 5.
B) Mayor que 5 y menor que 10.
C) Mayor que 10 y menor que 15.
D) Mayor que 15.

5 $\boxed{\text{A}}$ Los puntos de la figura a la derecha están siendo agrupados mediante el algoritmo C-Medias y, tras cierta iteración del algoritmo, se tiene la partición $\Pi = \{X_1 = \{(0,0),(0,2)\}, X_2 = \{(2,0),(2,4)\}\}$, medias $\mathbf{m}_1 = (0,1)$ y $\mathbf{m}_2 = (2,2)$, y SEC (suma de errores cuadráticos) $J = 10$. Si el punto $(2,0)$ se cambia de grupo, entonces:



A) La nueva SEC será menor que 6.
B) La nueva SEC estará entre 6 y 10.
C) La nueva SEC será mayor que 10.
D) No conviene cambiar ese punto de grupo porque los grupos se quedarían con tallas desequilibradas.

6 $\boxed{\text{A}}$ Sean dos clases, A y B, de las que se dispone de los siguientes prototipos: $A = \{(0,2),(1,1),(1,3),(2,2)\}$; y $B = \{(3,2),(3,3),(4,2),(4,3)\}$. Supóngase estos dos conjuntos de prototipos constituyen dos clústers que resultan de un proceso de agrupamiento no supervisado. La SEC, J, correspondiente a dicho agrupamiento es:

A) $J \leq 6$
B) $6 < J \leq 8$
C) $8 < J \leq 10$
D) $J > 10$

7 $\boxed{\text{D}}$ La diferencia principal entre el aprendizaje supervisado (AS) y no-supervisado (ANS) es que:

A) en el AS se conocen las clases correctas de los datos de test, mientras que en el ANS solo se conocen las de entrenamiento.
B) en el AS siempre hay un operador humano que supervisa los resultados de forma que el sistema solo sirve de ayuda o asistencia, mientras que en el ANS todo se realiza de forma totalmente automática.
C) el ANS es un proceso iterativo mientras que el AS se realiza en un paso.
D) en el AS se conocen las etiquetas de clase correctas de todos los datos de aprendizaje, mientras que en el ANS no se conocen.

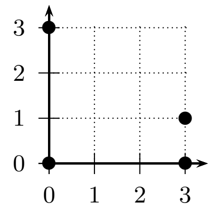
8 $\boxed{\text{B}}$ El algoritmo C-medias es una técnica de *clustering* particional que aplicamos en reconocimiento del habla para...

A) Transformar la señal al dominio temporal-frecuencial.
B) Diseñar *codebooks*.
C) Entrenar modelos de Markov.
D) Ninguna de las anteriores.

9 $\boxed{\text{B}}$ Sean dos clases, A y B, de las que se dispone de los siguientes prototipos: $A = \{(2,1),(1,2),(2,3),(3,2)\}$ y $B = \{(4,3),(5,3),(3,5),(6,5)\}$. Supóngase estos dos conjuntos de prototipos constituyen dos clústers que resultan de un proceso de agrupamiento particional. La SEC sería:

A) $SEC < 4$
B) $SEC > 12$
C) $SEC = 11$
D) $4 < SEC < 10$

10 $\boxed{\text{C}}$ Los puntos de la figura a la derecha están siendo agrupados mediante el algoritmo C-Medias y, tras cierta iteración del algoritmo, se tiene la partición $\Pi = \{X_1 = \{(0,0),(0,3),(3,0)\}, X_2 = \{(3,1)\}\}$. Sea J' la suma de errores cuadráticos de esta partición y sea J la suma de errores cuadráticos de la partición que se obtiene al cambiar de grupo el punto $(3,0)$. Entonces:

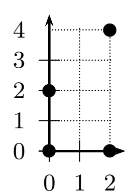


A) $J \geq J'$
B) $\frac{1}{2}J' \leq J < J'$
C) $\frac{1}{4}J' \leq J < \frac{1}{2}J'$
D) $J < \frac{1}{4}J'$

11 $\boxed{\text{C}}$ Cuál de la siguientes afirmaciones en relación al aprendizaje no supervisado es FALSA:

A) El objetivo del aprendizaje no supervisado es agrupar en grupos “naturales” los datos disponibles
B) Una medida muy empleada para medir la calidad de un agrupamiento particional es la Suma de Errores Cuadráticos (SEC)
C) El algoritmo c-medias garantiza un mínimo global del SEC
D) Se emplea, por ejemplo, en Reconocimiento Automático del habla para representar una señal acústica como una secuencia de símbolos asociados a los “codewords”

12 $\boxed{\text{B}}$ Los puntos de la figura a la derecha están siendo agrupados mediante el algoritmo C-Medias y, tras cierta iteración del algoritmo, se tiene la partición $\Pi = \{X_1 = \{(0,0),(0,2)\}, X_2 = \{(2,0),(2,4)\}\}$, medias $\mathbf{m}_1 = (0,1)$ y $\mathbf{m}_2 = (2,2)$, y SEC (suma de errores cuadráticos) $J = 10$. Si el punto $(2,0)$ se cambia de grupo, entonces:



A) La nueva SEC será menor que 5.
B) La nueva SEC estará entre 5 y 7.
C) La nueva SEC será mayor que 7 pero menor que 10
D) Ese punto no se puede cambiar porque deja uno de los grupos con sólo un punto.

13 $\boxed{\text{C}}$ Sea $X = \{1, 3, 4.5\}$ un conjunto de 3 datos unidimensionales a agrupar en dos clústers mediante alguna técnica de clustering particional. Más concretamente, se desea optimizar el criterio SEC (suma de errores cuadráticos) y emplear el algoritmo C-medias, si bien no se ha decidido si emplearemos la versión *popular* o la versión de *Duda y Hart (DH)*. Sea $\Pi^0 = \{X_1 = \{1, 3\}, X_2 = 4.5\}$ una partición inicial en dos clústers y SEC $J^0 = 2$. Indica cuál de las siguientes afirmaciones es cierta:

A) Tanto la versión popular como la DH terminarán sin modificar la partición inicial.
B) La versión popular terminará con una partición mejorada, pero no la versión DH, que terminará sin modificar la partición inicial.
C) La versión DH terminará con una partición mejorada, pero no la versión popular, que terminará sin modificar la partición inicial.
D) Ambas versiones terminarán con particiones mejoradas respecto a la partición inicial.

14 $\boxed{\text{A}}$ (Examen de SIN del 18 de Enero de 2013)
El criterio de clustering particional “Suma de Errores Cuadráticos” es apropiado cuando los datos forman clústers:

A) Hiperesféricos y de tamaño similar.
B) Hiperesféricos y de cualquier tamaño.
C) Alargados y de tamaño similar.
D) Alargados y de cualquier tamaño.

15 $\boxed{\text{C}}$ (Examen de SIN del 30 de Enero de 2013)
Se tienen 3 datos unidimensionales: $x_1 = 0$, $x_2 = 20$ y $x_3 = 35$. Se ha establecido una partición de estos datos en dos clústers, $\Pi = \{X_1 = \{x_1, x_2\}, X_2 = \{x_3\}\}$. La Suma de Errores Cuadráticos (SEC) de esta partición es:

A) $J(\Pi) = 0$
B) $0 < J(\Pi) \leq 150$
C) $150 < J(\Pi) \leq 300$ $\quad J(\Pi) = (x_1 - m_1)^2 + (x_2 - m_1)^2 + (x_3 - m_2)^2 = (0 - 10)^2 + (20 - 10)^2 + (35 - 35)^2 = 200$
D) $J(\Pi) > 300$

16 $\boxed{\text{B}}$ (Examen de SIN del 30 de Enero de 2013)
Tras aplicar la versión “correcta” (“Duda y Hart”) del algoritmo C-medias a la partición de la cuestión anterior (Π), la partición resultante (Π^*) es: $\quad \Delta J = \frac{n_2}{n_2+1}|x_2 - m_2|^2 - \frac{n_1}{n_1-1}|x_2 - m_1|^2$

A) $\Pi^* = \Pi$. $\quad \Delta J = 0$
B) $\Pi^* = \{X_1 = \{x_1\}, X_2 = \{x_2, x_3\}\}$. $\quad \Delta J = \frac{1}{2}|20 - 35|^2 - \frac{2}{1}|20 - 10|^2 = 112.5 - 200 = -87.5$
C) $\Pi^* = \{X_1 = \{x_2\}, X_2 = \{x_1, x_3\}\}$. $\quad \Delta J = \frac{1}{2}|0 - 35|^2 - \frac{2}{1}|0 - 10|^2 = 612.5 - 200 = 412.5$
D) Ninguna de las anteriores.

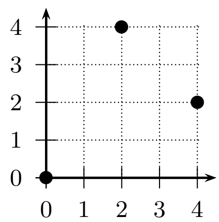
17 $\boxed{\text{D}}$ Indica cuál de la siguientes afirmaciones sobre *Clustering* es correcta:

A) Se suele emplear el algoritmo *Perceptrón* a partir de datos de entrenamiento *con* etiquetas de clase.
B) Se suele emplear el algoritmo *Perceptrón* a partir de datos de entrenamiento *sin* etiquetas de clase.
C) Se suele emplear el algoritmo *C-Medias* a partir de datos de entrenamiento *con* etiquetas de clase.
D) Se suele emplear el algoritmo *C-Medias* a partir de datos de entrenamiento *sin* etiquetas de clase.

18 $\boxed{\text{D}}$ El criterio de clustering particional “Suma de Errores Cuadráticos” es apropiado cuando los datos forman clústers:

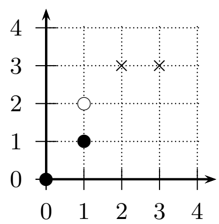
A) No alargados.
B) Alargados y de cualquier tamaño.
C) Alargados y de tamaño similar.
D) Ninguna de las anteriores.

19 $\boxed{\text{B}}$ La menor suma de errores cuadráticos con la que pueden agruparse en dos clústers los puntos a la derecha es un valor:



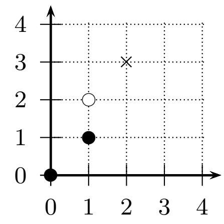
A) Entre 0 y 3.
B) Entre 3 y 6. $\quad J = 4$
C) Entre 6 y 9.
D) Mayor que 9.

20 $\boxed{\text{B}}$ La figura a la derecha muestra una partición de 5 puntos bidimensionales en 3 clústers (representados mediante los símbolos $\bullet$, $\circ$ y $\times$). Considera todas las posibles transferencias de clúster de cada punto (en un clúster no unitario). En términos de suma de errores cuadráticos (J):



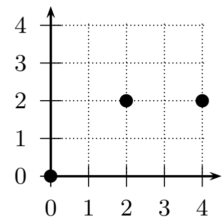
A) Ninguna transferencia permite mejorar J.
B) Sólo se puede mejorar J transfiriendo $(1, 1)^t$ del clúster $\bullet$ al $\circ$.
C) Sólo se puede mejorar J transfiriendo $(2, 3)^t$ del clúster $\times$ al $\circ$.
D) Las dos transferencias anteriores permiten mejorar J.

21 $\boxed{\text{C}}$ La figura a la derecha muestra una partición de 4 puntos bidimensionales en 3 clústers (representados mediante los símbolos $\bullet$, $\circ$ y $\times$). La suma de errores cuadráticos de esta partición es $J = 1$. Si se ejecuta el algoritmo C-medias (de Duda y Hart) a partir de la misma:



A) No se realizará niguna transferencia de clúster.
B) Se transferirá un único punto, obteniéndose una partición de J entre $\frac{2}{3}$ y 1.
C) Se transferirá un único punto, obteniéndose una partición de J entre 0 y $\frac{2}{3}$. J=0.5
D) Se realizarán dos transferencias de clúster, obteniéndose una partición de J nula.

22 $\boxed{\text{A}}$ La menor suma de errores cuadráticos con la que pueden agruparse en dos clústers los puntos a la derecha es un valor:



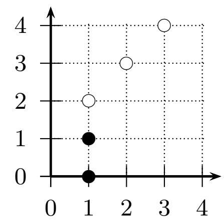
A) Entre 0 y 3. $J = 2$

B) Entre 3 y 6.

C) Entre 6 y 9.

D) Mayor que 9.

23 $\boxed{\text{C}}$ La figura a la derecha muestra una partición de 5 puntos bidimensionales en 2 clústers (representados mediante los símbolos $\bullet$ y $\circ$). Considera todas las posibles transferencias de clúster de cada punto. La transferencia más provechosa en términos de suma de errores cuadráticos (SEC) conduce a un incremento de SEC (ΔJ):



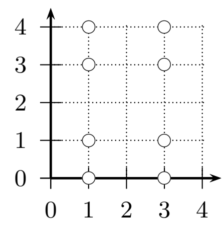
A) $\Delta J > 0$

B) $0 \geq \Delta J > -1$

C) $-1 \geq \Delta J > -2$ $\Delta J = -1.5$ $(J = 4.5 \rightarrow J = 3)$

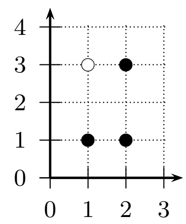
D) $-2 \geq \Delta J$

24 $\boxed{\text{B}}$ La figura a la derecha muestra 8 puntos bidimensionales. La menor "Suma de Errores Cuadráticos", J, con la que pueden agruparse estos puntos en dos clústers es:



A) $0 \leq J \leq 7$
B) $7 < J \leq 14$ $J = 10$
C) $14 < J \leq 21$
D) $21 < J$

25 $\boxed{\text{D}}$ La figura a la derecha muestra una partición de 4 puntos bidimensionales en 2 clústers (representados mediante los símbolos $\bullet$ y $\circ$). La suma de errores cuadráticos (SEC) de esta partición es $J = \frac{30}{9}$. La transferencia del punto $(2, 3)^t$ del clúster $\bullet$ al $\circ$ conduce a un incremento de la SEC, ΔJ, tal que:

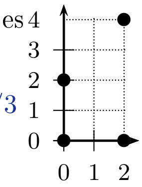


A) $\Delta J > 0$
B) $0 \geq \Delta J > -1$
C) $-1 \geq \Delta J > -2$
D) $-2 \geq \Delta J$ $\Delta J = -\frac{21}{9} = -2.33$ $(J = \frac{30}{9} \rightarrow J = 1)$

26 $\boxed{\text{B}}$ Dos versiones bien conocidas del algoritmo C-medias son la de *Duda y Hart* (DH) y la "popular". Suponiendo que ambas versiones se aplican a partir de un misma partición inicial, indica cuál de las siguientes afirmaciones sobre sus resultados es cierta:

A) Ambas versiones obtendrán la misma partición optimizada.
B) La versión DH obtendrá una partición final que no podrá mejorarse mediante la versión popular.
C) La versión popular obtendrá una partición final que no podrá mejorarse mediante la versión DH.
D) La partición final obtenida mediante la versión DH podrá mejorarse mediante la versión popular, y viceversa.

27 $\boxed{\text{A}}$ Considérese la partición $\Pi = \{X_1 = \{(0,0)^t, (0,2)^t\}, X_2 = \{(2,0)^t, (2,4)^t\}\}$ de los puntos de la figura a la derecha. Las medias de esta partición son $\mathbf{m}_1 = (0,1)^t$ y $\mathbf{m}_2 = (2,2)^t$. Su suma de errores cuadráticos, SEC, es 10. Si el punto $(0,2)^t$ se cambia de grupo, entonces:



A) La nueva SEC será mayor que 10. $\|(0,2)^t - (4/3,2)^t\|^2 + \|(2,0)^t - (4/3,2)^t\|^2 + \|(2,4)^t - (4/3,2)^t\|^2 = 32/3$
B) La nueva SEC será mayor que 8 y no mayor que 10.
C) La nueva SEC será mayor que 6 y no mayor que 8.
D) La nueva SEC no será mayor que 6.

28 $\boxed{\text{D}}$ Se tiene un problema de clasificación para el cual se ha aprendido un clasificador. El intervalo de confianza al 95 % para la probabilidad de error de dicho clasificador se ha estimado empíricamente, a partir de un cierto conjunto de muestras de test. Indica cuál de las siguientes opciones permitiría reducir el tamaño del intervalo estimado:

A) Reducir significativamente el conjunto de test.
B) Mantener el conjunto de test y re-entrenar el clasificador con el algoritmo C-medias de Duda y Hart.
C) Mantener el conjunto de test y re-entrenar el clasificador con el algoritmo C-medias convencional ("popular").
D) Aumentar significativamente el conjunto de test.

29 $\boxed{\text{B}}$ Indica cuál de las siguientes afirmaciones sobre aprendizaje supervisado (AS) y no-supervisado (ANS) es correcta:

A) Tanto en ANS como en AS se requieren datos de entrenamiento sin etiqueta de clase.
B) En ANS se requieren datos de entrenamiento sin etiqueta de clase; en AS, con etiqueta.
C) En ANS se requieren datos de entrenamiento con etiqueta de clase; en AS, sin etiqueta.
D) Tanto en ANS como en AS se requieren datos de entrenamiento con etiqueta de clase.

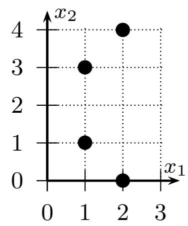
30 $\boxed{\text{D}}$ Considérese el algoritmo C-medias en su versión convencional o "popular" (CM), así como en su versión de Duda y Hart (DH). Aunque ambas optimizan la *suma de errores cuadráticos* (SEC), sus resultados pueden diferir pues:

A) DH mimimiza la SEC y CM la maximiza.
B) DH maximiza la SEC y CM la minimiza.
C) Ambas maximizan la SEC, si bien DH puede alcanzar mejores soluciones que CM.
D) Ninguna de las anteriores.

31 $\boxed{\text{D}}$ Supongamos que hemos aplicado el algoritmo C-medias a un conjunto de puntos bi-dimensionales para obtener un agrupamiento (partición) en dos *clústers*. Tras una serie de iteraciones del algoritmo C-medias tenemos el agrupamiento: $\{\ \{(0,1)^\mathbf{t}, (0,2)^\mathbf{t}\},\ \{(0,3)^\mathbf{t}, (0,5)^\mathbf{t}, (0,6)^\mathbf{t}, (0,7)^\mathbf{t}, (1,6)^\mathbf{t}, (-1,6)^\mathbf{t}\}\ \}$. Indica la respuesta *correcta*.

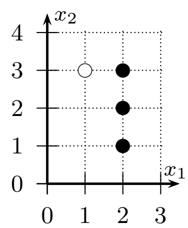
A) La suma de errores cuadráticos (SEC) es 15 y puede llegar a ser 8.
B) La SEC es 15 y cuando el algoritmo converja será 12.
C) La SEC es 12 y cuando el algoritmo converja será 10.
D) La SEC es 12 y cuando el algoritmo converja será 6.

32 $\boxed{\text{D}}$ En la figura de la derecha se representan 4 muestras de bidimensionales. ¿Cuál es el número de clústers que minimiza la suma de errores cuadráticos para dicho conjunto de muestras?



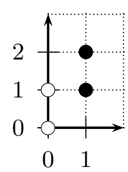
A) 1
B) 2
C) 3
D) 4 $\quad J = 0$

33 $\boxed{\text{D}}$ La figura a la derecha muestra una partición de 4 puntos bidimensionales en 2 clústers (representados mediante los símbolos $\bullet$ y $\circ$). La transferencia del punto $(2,3)^t$ del clúster $\bullet$ al $\circ$ conduce a una variación de la SEC, ΔJ, tal que:



A) $\Delta J > 0$.
B) $0 \geq \Delta J > -\frac{1}{2}$.
C) $-\frac{1}{2} \geq \Delta J > -1$.
D) $-1 \geq \Delta J$. $\qquad \Delta J = \frac{1}{2} - \frac{3}{2} = -1$

34 $\boxed{\text{A}}$ En la figura de la derecha se muestra una partición de 4 puntos bidimensionales de 2 clústers. La transferencia del punto $(1,1)^t$ del clúster $\bullet$ al clúster $\circ$



A) produce un incremento en la SEC.
B) produce un decremento en la SEC.
C) no altera la SEC.
D) produce una SEC negativa.

35 $\boxed{\text{A}}$ Considérese el algoritmo C-medias de Duda y Hart. Indicar cuál de las siguientes afirmaciones es *correcta*:

A) Su buena eficacia computational se consigue gracias al cálculo incremental de la variación de distorsión y de los vectores media de clúster.
B) Determina el número de clústers que minimiza la suma de errores cuadráticos (SEC).
C) Cuando un clúster se queda vacío, dicho clúster se elimina.
D) Ninguna de las anteriores.

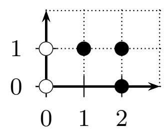
36 $\boxed{\text{D}}$ Considérese el conjunto de aprendizaje formado por los 6 datos tridimensionales de la tabla a la derecha. Se cree que una partición natural de dicho conjunto en 2 clústers consiste en agrupar los primeros 4 datos en un clúster y los 2 últimos en el otro. La suma de errores cuadráticos de dicha partición, J, es:

A) $J < 3$
B) $3 \leq J < 6$
C) $6 \leq J < 12$
D) $12 \leq J$ $\quad(1+2+1+2)+(3+3)=6+6=12$

$\mathbf{x}_n = (x_{n1}, x_{n2}, x_{n3})^t$

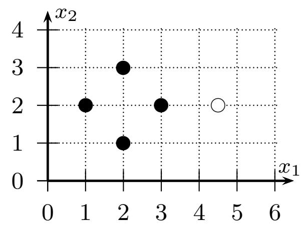
n	x_{n1}	x_{n2}	x_{n3}
1	0	1	1
2	2	1	0
3	1	2	1
4	1	0	2
5	4	6	4
6	6	4	6

37 $\boxed{\text{C}}$ En la figura de la derecha se muestra una partición de 5 puntos bidimensionales de 2 clusters. La transferencia del punto $(1,1)^t$ del cluster $\bullet$ al cluster $\circ$



A) produce un incremento en la SEC.
B) produce un decremento en la SEC.
C) no altera la SEC.
D) produce una SEC negativa.

38 $\boxed{\text{C}}$ En la figura de la derecha se representan 5 muestras bidimensionales particionadas inicialmente en dos clústers ($\bullet$ y $\circ$). ¿Cuál sería el resultado de la aplicación de una iteración del algoritmo C-medias en su versión convencional?, ¿y en su versión Duda y Hart (D&H)?

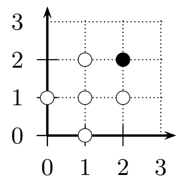


A) Ambas versiones transfieren la muestra $(3,2)$.
B) Sólo la versión convencional transfiere la muestra $(3,2)$.
C) Sólo la versión D&H transfiere la muestra $(3,2)$.
D) Ninguna de las dos versiones transfiere la muestra $(3,2)$.

39 $\boxed{\text{A}}$ Se aplica el algoritmo C-medias de Duda y Hart a un conjunto de N vectores no etiquetados y se obtiene una partición de dicho conjunto en C subconjuntos disjuntos cuya suma de errores cuadráticos, SEC, es R. ¿Cuál de las siguientes afirmaciones es *falsa*?:

A) Si $C \geq N/2$, $R = 0$.
B) Si $C = N$, $R = 0$.
C) Si $C \leq N$, C-medias termina en un número finito de iteraciónes y R es un mínimo local de la SEC.
D) Ninguna de las anteriores.

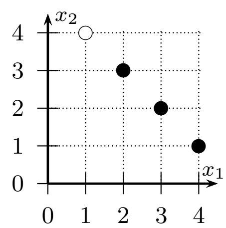
40 $\boxed{\text{C}}$ En la figura de la derecha se muestra una partición de 6 puntos bidimensionales en 2 clústers, $\circ$ y $\bullet$, obtenida mediante el algoritmo C-medias (convencional o "popular"). Si transferimos los puntos $(1,2)^t$ y $(2,1)^t$ del clúster $\circ$ al clúster $\bullet$, entonces:



A) se produce un incremento de la SEC.
B) no se altera la SEC.
C) se produce un decremento de la SEC.
D) se produce una SEC igual a 0.

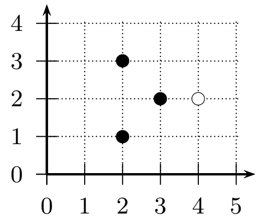
$$J' = J'_{\circ} + J'_{\bullet} = 4 + 0 = 4$$
$$J = J_{\circ} + J_{\bullet} = \tfrac{4}{3} + \tfrac{4}{3} = \tfrac{8}{3}$$
$$\Delta J = J - J' = \tfrac{8}{3} - 4 = -\tfrac{4}{3} < 0$$

41 $\boxed{\text{D}}$ La figura a la derecha muestra una partición de 4 puntos bidimensionales en 2 clústers (representados mediante los símbolos $\bullet$ y $\circ$). La transferencia del punto $(2,3)^t$ del clúster $\bullet$ al $\circ$ conduce a una variación de la SEC, ΔJ, tal que:



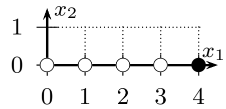
A) $\Delta J > 0$.
B) $0 \geq \Delta J > -1$.
C) $-1 \geq \Delta J > -2$.
D) $-2 \geq \Delta J$. $\quad \Delta J = \frac{2}{2} - \frac{6}{2} = -2$

42 $\boxed{\text{B}}$ La figura a la derecha muestra una partición de 4 puntos bidimensionales en 2 clústers (representados mediante los símbolos $\bullet$ y $\circ$). La transferencia del punto $(3,2)^t$ del cluster $\bullet$ al cluster $\circ$:



A) produce un incremento en la SEC.
B) produce un decremento en la SEC. $\Delta J = 0.5 - 0.67335 = -0.17335$
C) no altera la SEC.
D) produce una SEC negativa.

43 $\boxed{\text{C}}$ La figura a la derecha muestra una partición de 5 puntos bidimensionales en 2 clústers (representados mediante los símbolos $\bullet$ y $\circ$). Si, durante la ejecución del algoritmo C-medias, el punto $(3,0)$ se cambia del clúster $\circ$ al $\bullet$, entonces (indica cuál de la siguientes afirmaciones es cierta):



A) Las medias de clúster no cambian.
B) La suma de errores cuadráticos crece.
C) La suma de errores cuadráticos decrece.
D) Solo cambia la suma de errores cuadráticos de uno de los clústers.

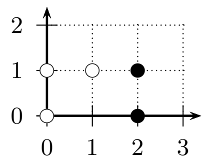
44 $\boxed{\text{B}}$ Se tiene una partición de un conjunto de datos 3-dimensionales en un número de clústers dado, $C \geq 2$. Considérese la transferencia del dato $\mathbf{x} = (3,2,9)^t$ de un clúster i a otro j, $j \neq i$. Se sabe que el clúster i contiene 3 datos (contando $\mathbf{x}$) y el j 4. Asimismo, se sabe que la media del clúster i es $\mathbf{m}_i = (7,3,3)^t$ y la del j $\mathbf{m}_j = (7,6,7)^t$. Si se realiza dicha transferencia, se producirá un incremento de la suma de errores cuadráticos, ΔJ, tal que: $\Delta J = -50.7$

A) $\Delta J < -70$
B) $-70 \leq \Delta J < -30$
C) $-30 \leq \Delta J < 0$
D) $\Delta J \geq 0$

45 $\boxed{\text{B}}$ Se tiene una partición de un conjunto de datos 3-dimensionales en un número de clústers dado, $C \geq 2$. Considérese la transferencia del dato $\mathbf{x} = (4,5,2)^t$ de un clúster i a otro j, $j \neq i$. Se sabe que el clúster i contiene 3 datos (contando $\mathbf{x}$) y el j 2. Asimismo, se sabe que la media del clúster i es $\mathbf{m}_i = (10,8,4)^t$ y la del j $\mathbf{m}_j = (7,7,1)^t$. Si se realiza dicha transferencia, se producirá un incremento de la suma de errores cuadráticos, ΔJ, tal que: $\Delta J = -64.2$

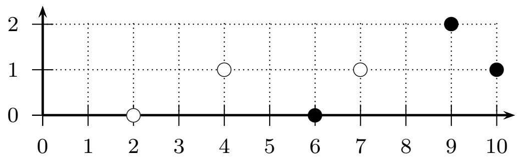
A) $\Delta J < -70$
B) $-70 \leq \Delta J < -30$
C) $-30 \leq \Delta J < 0$
D) $\Delta J \geq 0$

46 $\boxed{\text{D}}$ En la figura de la derecha se muestra una partición de 5 puntos bidimensionales de 2 clusters. La transferencia del punto $(1,1)^t$ del cluster $\circ$ al cluster $\bullet$:



A) Produce un incremento de la Suma de Errores Cuadráticos (SEC).
B) Produce un decremento de la SEC.
C) Produce una SEC negativa.
D) No altera la SEC.

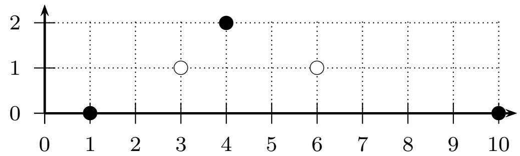
47 $\boxed{\text{D}}$ La figura siguiente muestra una partición de 6 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



Si intercambiamos de clúster los puntos $(10,1)^t$ y $(7,1)^t$, se produce una variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio), tal que:

A) $\Delta J < -7$. $\Delta J = 42.0 - 24.0 = 18.0$
B) $-7 \leq \Delta J < 0$.
C) $0 \leq \Delta J < 7$.
D) $\Delta J \geq 7$.

48 $\boxed{\text{B}}$ La figura siguiente muestra una partición de 5 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



Si intercambiamos de clúster los puntos $(1,0)^t$ y $(3,1)^t$, se produce una variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio), tal que:

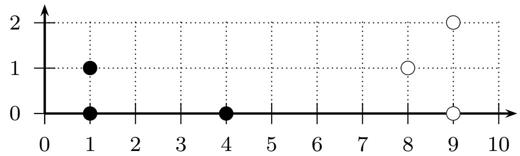
A) $\Delta J < -7$. $\quad\quad$ $\Delta J = 43.7 - 49.2 = -5.5$

B) $-7 \leq \Delta J < 0$.

C) $0 \leq \Delta J < 7$.

D) $\Delta J \geq 7$.

49 $\boxed{\text{D}}$ La figura siguiente muestra una partición de 6 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



Si transferimos de clúster el punto $(1,0)^t$, se produce una variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio), tal que:

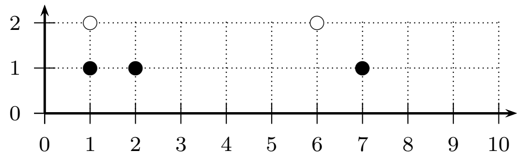
A) $\Delta J < -7$. $\quad\quad$ $\Delta J = 52.5 - 9.3 = 43.2$

B) $-7 \leq \Delta J < 0$.

C) $0 \leq \Delta J < 7$.

D) $\Delta J \geq 7$.

50 $\boxed{\text{A}}$ La figura siguiente muestra una partición de 5 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



Si transferimos de clúster el punto $(7,1)^t$, se produce una variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio), tal que:

A) $\Delta J < -7$. $\quad\quad$ $\Delta J = 21.8 - 33.2 = -11.3$

B) $-7 \leq \Delta J < 0$.

C) $0 \leq \Delta J < 7$.

D) $\Delta J \geq 7$.

51 $\boxed{\text{B}}$ Se tiene una partición de un conjunto de datos 3-dimensionales en un número de clústers dado, $C \geq 2$. Considérese la transferencia del dato $\mathbf{x} = (1,6,9)^t$ de un clúster j a otro i, $j \neq i$. Se sabe que el clúster j contiene 2 datos (contando $\mathbf{x}$) y el i 3. Asimismo, se sabe que la media del clúster j es $\mathbf{m}_j = (8,2,8)^t$ y la del i $\mathbf{m}_i = (10,8,9)^t$. Si se realiza dicha transferencia, se producirá un incremento de la suma de errores cuadráticos, ΔJ, tal que: $\quad$ $\Delta J = -68.2$

A) $\Delta J < -70$

B) $-70 \leq \Delta J < -30$

C) $-30 \leq \Delta J < 0$

D) $\Delta J \geq 0$

52 $\boxed{\text{D}}$ Se tiene una partición de un conjunto de datos 3-dimensionales en un número de clústers dado, $C \geq 2$. Considérese la transferencia del dato $\mathbf{x} = (3,6,4)^t$ de un clúster j a otro i, $j \neq i$. Se sabe que el clúster j contiene 3 datos (contando $\mathbf{x}$) y el i 3. Asimismo, se sabe que la media del clúster j es $\mathbf{m}_j = (3,3,2)^t$ y la del i $\mathbf{m}_i = (7,6,9)^t$. Si se realiza dicha transferencia, se producirá un incremento de la suma de errores cuadráticos, ΔJ, tal que: $\quad$ $\Delta J = 11.2$

A) $\Delta J < -70$

B) $-70 \leq \Delta J < -30$

C) $-30 \leq \Delta J < 0$

D) $\Delta J \geq 0$

53 $\boxed{\text{D}}$ Se tiene una partición de un conjunto de datos 3-dimensionales en un número de clústers dado, $C \geq 2$. Considérese la transferencia del dato $\mathbf{x} = (4, 3, 5)^t$ de un clúster i a otro j, $j \neq i$. Se sabe que el clúster i contiene 4 datos (contando $\mathbf{x}$) y el j 3. Asimismo, se sabe que la media del clúster i es $\mathbf{m}_i = (3, 8, 8)^t$ y la del j $\mathbf{m}_j = (10, 9, 10)^t$. Si se realiza dicha transferencia, se producirá un incremento de la suma de errores cuadráticos, ΔJ, tal que: $\Delta J = 26.1$

A) $\Delta J < -70$

B) $-70 \leq \Delta J < -30$

C) $-30 \leq \Delta J < 0$

D) $\Delta J \geq 0$

54 $\boxed{\text{D}}$ Se tiene una partición de un conjunto de datos 3-dimensionales en un número de clústers dado, $C \geq 2$. Considérese la transferencia del dato $\mathbf{x} = (4, 10, 4)^t$ de un clúster i a otro j, $j \neq i$. Se sabe que el clúster i contiene 4 datos (contando $\mathbf{x}$) y el j 2. Asimismo, se sabe que la media del clúster i es $\mathbf{m}_i = (1, 8, 2)^t$ y la del j $\mathbf{m}_j = (10, 2, 10)^t$. Si se realiza dicha transferencia, se producirá un incremento de la suma de errores cuadráticos, ΔJ, tal que: $\Delta J = 68.0$

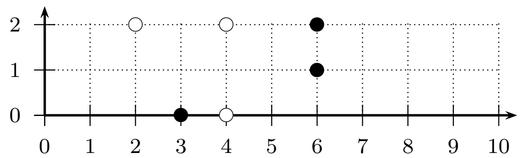
A) $\Delta J < -70$

B) $-70 \leq \Delta J < -30$

C) $-30 \leq \Delta J < 0$

D) $\Delta J \geq 0$

55 $\boxed{\text{A}}$ La figura siguiente muestra una partición de 6 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



¿Qué punto al ser transferido de clúster minimiza la variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio)? $\Delta J = 7.2 - 13.3 = -6.1$

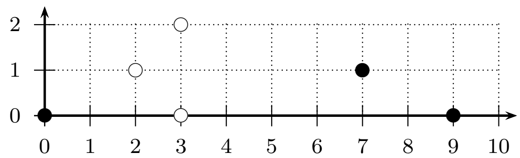
A) $(3, 0)^t$

B) $(6, 2)^t$

C) $(4, 0)^t$

D) $(2, 2)^t$

56 $\boxed{\text{A}}$ La figura siguiente muestra una partición de 6 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



¿Qué punto al ser transferido de clúster minimiza la variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio)? $\Delta J = 11.2 - 48.0 = -36.8$

A) $(0, 0)^t$

B) $(9, 0)^t$

C) $(2, 1)^t$

D) $(3, 0)^t$

2. Problemas

1. Se tienen los siguientes 5 vectores bidimensionales:

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 7 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} 8 \\ 2 \end{pmatrix} \quad \text{y} \quad \mathbf{x}_5 = \begin{pmatrix} 8 \\ 6 \end{pmatrix}$$

Se desea agrupar estos vectores de manera no-supervisada en 2 clases. Partiendo de la partición

$$\Pi = \{X_1 = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}, X_2 = \{\mathbf{x}_4, \mathbf{x}_5\}\}$$

realiza una traza de ejecución de una iteración del bucle principal del algoritmo *c-medias*.

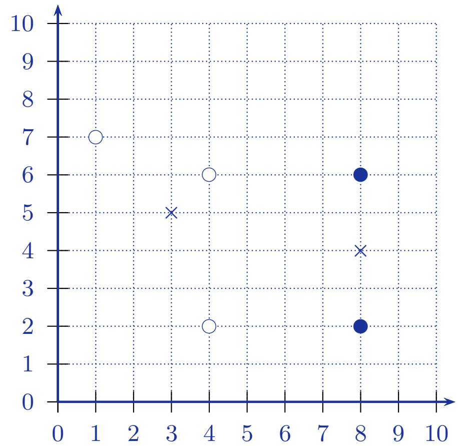
$$\mathbf{m}_1 = \frac{1}{3}(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$\mathbf{m}_2 = \frac{1}{2}(\mathbf{x}_4 + \mathbf{x}_5) = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$$

$$J_1 = \|\mathbf{x}_1 - \mathbf{m}_1\|^2 + \|\mathbf{x}_2 - \mathbf{m}_1\|^2 + \|\mathbf{x}_3 - \mathbf{m}_1\|^2 = 8 + 10 + 2 = 20$$

$$J_2 = \|\mathbf{x}_4 - \mathbf{m}_2\|^2 + \|\mathbf{x}_5 - \mathbf{m}_2\|^2 = 4 + 4 = 8$$

$$J = J_1 + J_2 = 28$$



Si transferimos $\mathbf{x}_n \in X_i$ a X_j, entonces $\quad \Delta J = \frac{|X_j|}{|X_j|+1}\|\mathbf{x}_n - \mathbf{m}_j\|^2 - \frac{|X_i|}{|X_i|-1}\|\mathbf{x}_n - \mathbf{m}_i\|^2$

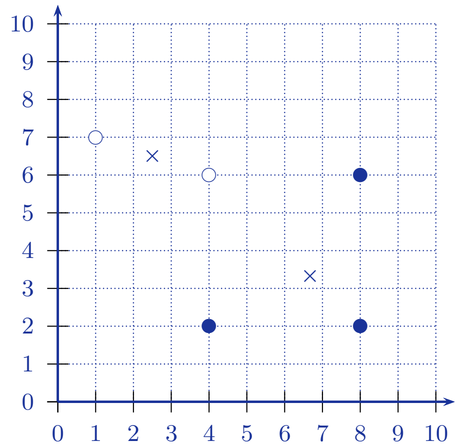
¿Transferimos $\mathbf{x}_1$ de X_1 a X_2? : $\Delta J = \frac{2}{3} \cdot 58 - \frac{3}{2} \cdot 8 = \frac{80}{3} > 0 \Rightarrow$ NO

¿Transferimos $\mathbf{x}_2$ de X_1 a X_2? : $\Delta J = \frac{2}{3} \cdot 20 - \frac{3}{2} \cdot 10 = -\frac{5}{3} < 0 \Rightarrow$ SÍ

$$\mathbf{m}_1 = \mathbf{m}_1 - \frac{\mathbf{x}_2 - \mathbf{m}_1}{|X_1| - 1} = \begin{pmatrix} 5/2 \\ 13/2 \end{pmatrix} = \begin{pmatrix} 2.5 \\ 6.5 \end{pmatrix}$$

$$\mathbf{m}_2 = \mathbf{m}_2 + \frac{\mathbf{x}_2 - \mathbf{m}_2}{|X_2| + 1} = \begin{pmatrix} 20/3 \\ 10/3 \end{pmatrix} = \begin{pmatrix} 6.67 \\ 3.33 \end{pmatrix}$$

$$J = J + \Delta J = \frac{79}{3} = 26.33$$



¿Transferimos $\mathbf{x}_3$ de X_1 a X_2? : $\Delta J = \frac{3}{4} \cdot \frac{128}{9} - \frac{2}{1} \cdot \frac{10}{4} = \frac{17}{3} = 5.67 > 0 \Rightarrow$ NO

¿Transferimos $\mathbf{x}_4$ de X_2 a X_1? : $\Delta J = \frac{2}{3} \cdot \frac{151}{2} - \frac{3}{2} \cdot \frac{32}{9} = \frac{805}{16} = 50.31 > 0 \Rightarrow$ NO

¿Transferimos $\mathbf{x}_5$ de X_2 a X_1? : $\Delta J = \frac{2}{3} \cdot \frac{61}{2} - \frac{3}{2} \cdot \frac{80}{9} = 7 > 0 \Rightarrow$ NO

(AQUÍ TERMINA LA RESPUESTA AL ENUNCIADO). El algoritmo continúa como sigue:

¿Transferimos $\mathbf{x}_1$ de X_1 a X_2? : $\Delta J = \frac{3}{4} \cdot \frac{410}{9} - \frac{2}{1} \cdot \frac{5}{2} = \frac{175}{6} = 29.17 > 0 \Rightarrow$ NO

¿Transferimos $\mathbf{x}_2$ de X_2 a X_1? : $\Delta J = \frac{2}{3} \cdot \frac{45}{2} - \frac{3}{2} \cdot \frac{80}{9} = \frac{5}{3} = 1.67 > 0 \Rightarrow$ NO

No es necesario seguir ya que no se realizará ninguna transferencia. La partición optimizada es:

$$\Pi = \{X_1 = \{\mathbf{x}_1, \mathbf{x}_3\}, X_2 = \{\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5\}\}$$