Quiz: Unsupervised Learning

Question 1:

- 1. For which of the following tasks might K-means clustering be a suitable algorithm? Select all that apply.
 - Given a database of information about your users, automatically group them into different market segments.

✓ Correcto

You can use K-means to cluster the database entries, and each cluster will correspond to a different market segment.

Given sales data from a large number of products in a supermarket, figure out which products tend to form coherent groups (say are frequently purchased together) and thus should be put on the same shelf.

✓ Correcto

If you cluster the sales data with K-means, each cluster should correspond to coherent groups of items.

Results: Correct

Question 2:

- 2. Suppose we have three cluster centroids $\mu_1=\begin{bmatrix}1\\2\end{bmatrix}$, $\mu_2=\begin{bmatrix}-3\\0\end{bmatrix}$ and $\mu_3=\begin{bmatrix}4\\2\end{bmatrix}$. Furthermore, we have a training example $x^{(i)}=\begin{bmatrix}3\\1\end{bmatrix}$. After a cluster assignment step, what will $c^{(i)}$ be?
 - $c^{(i)} = 3$
 - $\bigcirc \ c^{(i)}$ is not assigned
 - $\bigcirc c^{(i)} = 2$
 - $\bigcirc \ c^{(i)} = 1$

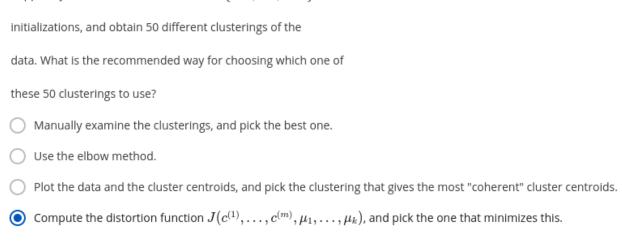
✓ Correcto

 $x^{(i)}$ is closest to μ_3 , so $c^{(i)}=3$

Results: Correct

Question 3:

3.	K-means is an iterative algorithm, and two of the following steps are repeatedly carried out in its inner-loop. Which two?
	Feature scaling, to ensure each feature is on a comparable scale to the others.
	Using the elbow method to choose K.
	$lacksquare$ Move the cluster centroids, where the centroids μ_k are updated.
	✓ Correcto
	The cluster update is the second step of the K-means loop.
	$lacksquare$ The cluster assignment step, where the parameters $c^{(i)}$ are updated.
	✓ Correcto
	This is the correst first step of the K-means loop.
Results: Correct	
Question 4	
4.	Suppose you have an unlabeled dataset $\{x^{(1)},\dots,x^{(m)}\}$. You run K-means with 50 different random
	initializations, and obtain 50 different clusterings of the



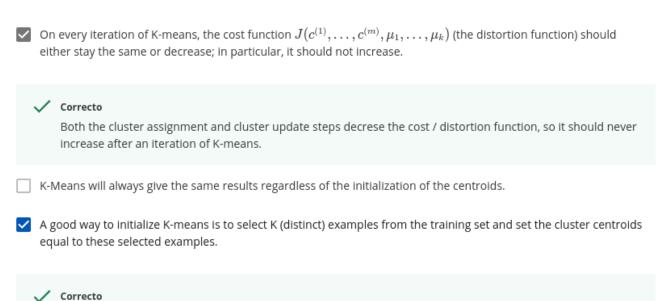


 $\hbox{A lower value for the distortion function implies a better clustering, so you should choose the clustering with}\\$

Results: Correct

Question 5

5. Which of the following statements are true? Select all that apply.



Results: Correct