Unsupervised Learning

LATEST SUBMISSION GRADE

100%

1. For which of the following tasks might K-means clustering be a suitable algorithm? Select all that apply. **1 / 1 point**

- [x] Given a database of information about your users, automatically group them into different market segments.

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

- [x] 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.

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

- [ ] Given historical weather records, predict the amount of rainfall tomorrow (this would be a real-valued output)

- [ ] Given sales data from a large number of products in a supermarket, estimate future sales for each of these products.

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 $x^{(i)} = \begin{bmatrix}-2\\1\end{bmatrix}$. After a cluster assignment step, what will $c^{(i)}$ be? **1 / 1 point**

- (•) $c^{(i)} = 2$
- ( ) $c^{(i)} = 1$
- ( ) $c^{(i)} = 3$
- ( ) $c^{(i)}$ is not assigned

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

- [ ] Given historical weather records, predict the amount of rainfall tomorrow (this would be a real-valued output)

- [ ] Given sales data from a large number of products in a supermarket, estimate future sales for each of these products.

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 $x^{(i)} = \begin{bmatrix}-2\\1\end{bmatrix}$. After a cluster assignment step, what will $c^{(i)}$ be? **1 / 1 point**

- (●) $c^{(i)} = 2$
- ( ) $c^{(i)} = 1$
- ( ) $c^{(i)} = 3$
- ( ) $c^{(i)}$ is not assigned

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

- [ ] Given historical weather records, predict the amount of rainfall tomorrow (this would be a real-valued output)
- [ ] Given sales data from a large number of products in a supermarket, estimate future sales for each of these products.

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 $x^{(i)} = \begin{bmatrix}-2\\1\end{bmatrix}$. After a cluster assignment step, what will $c^{(i)}$ be? **1 / 1 point**

- (●) $c^{(i)} = 2$
- ( ) $c^{(i)} = 1$
- ( ) $c^{(i)} = 3$
- ( ) $c^{(i)}$ is not assigned

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

5. Which of the following statements are true? Select all that apply. **1 / 1 point**

- [ ] Once an example has been assigned to a particular centroid, it will never be reassigned to another different centroid
- [x] A good way to initialize K-means is to select K (distinct) examples from the training set and set the cluster centroids equal to these selected examples.

✓ **Correct**
This is the recommended method of initialization.

- [x] On every iteration of K-means, the cost function $J(c^{(1)}, \ldots, c^{(m)}, \mu_1, \ldots, \mu_k)$ (the distortion function) should either stay the same or decrease; in particular, it should not increase.

✓ **Correct**
Both the cluster assignment and cluster update steps decrese the cost / distortion function, so it should never increase after an iteration of K-means.

- [ ] K-Means will always give the same results regardless of the initialization of the centroids.