

Solution

1. Clustering is considered what type of learning? Supervised, semi-supervised, unsupervised, or reinforcement? Circle one and explain. (2 points)

no labels.

2. K-Means will always converge. True or false? Circle one. (1 point)

3. The K in K-Means is a parameter that is determined by the training data. True or false. Circle one and explain. (2 points)

set by programmer

4. Answer the following questions given the data below. The features are in 1D.

$s_1 = -4$
 $s_2 = -3$
 $s_3 = -1$
 $s_4 = 0$
 $s_5 = 2$
 $s_6 = 3$
 $s_7 = 4$

- (a) Use K-Means with $K=3$ to cluster the data. Start with $\mu_1 = 2$ and $\mu_2 = 3$, and $\mu_3 = 4$. What are your final clusters, C_1 and C_2 ? (4 points)

iter 1 $\mu_1 = 2$ $\mu_2 = 3$ $\mu_3 = 4$ $C_1 = \{-4, -3, -1, 0, 2\}$ $C_2 = \{3\}$ $C_3 = \{4\}$
 iter 2 $\mu_1 = -1\frac{1}{5}$ $\mu_2 = 3$ $\mu_3 = 4$ $C_1 = \{-4, -3, -1, 0\}$ $C_2 = \{2, 3\}$ $C_3 = \{4\}$
 iter 3 $\mu_1 = -2$ $\mu_2 = 2\frac{1}{2}$ $\mu_3 = 4$ $C_1 = \{-4, -3, -1, 0\}$ $C_2 = \{2, 3\}$ $C_3 = \{4\}$
 iter 4 C_1, C_2, C_3 no change from 2-3 so μ s no change done ✓

- (b) Is this the best clustering you think could have been achieved with K-Means? Explain. (1 point)

No. $\{-4, -3\}$ $\{-1, 0\}$ $\{2, 3, 4\}$
 makes most sense.

1. Clustering is considered what type of learning? Supervised, semi-supervised, unsupervised, or reinforcement? Circle one and explain. (2 points)

No labels provided with training data.

2. K-Means will always converge to the same cluster centers. True or false. Circle one and explain. (2 points)

guaranteed to converge, but not to same result.

3. How are K-Means and K-NN similar in the way they use features? (2 points)

They treat all features equally.
They love garbage!

4. Answer the following questions given the data below. The features are in 1D.

$s_1 = -4$
 $s_2 = -3$
 $s_3 = -1$
 $s_4 = 0$
 $s_5 = 2$
 $s_6 = 3$
 $s_7 = 4$

- (a) Use K-Means with $K=2$ to cluster the data. Start with $\mu_1 = -2$ and $\mu_2 = 1$. What are your final clusters, C_1 and C_2 ? (3 points)

iter 1: $\mu_1 = -2$ $\mu_2 = 1$ $C_1 = \{-4, -3, -1\}$ $C_2 = \{0, 2, 3, 4\}$

iter 2: $\mu_1 = -2\frac{2}{3}$ $\mu_2 = 2\frac{1}{4}$ $C_1 = \{-4, -3, -1\}$ $C_2 = \{0, 2, 3, 4\}$

iter 3: $C_1 \hat{=} C_2$ no change from iter 1 to 2
 so $\mu_1 \hat{=} \mu_2$ are same as iter 2. done ✓

- (b) How would you classify the following test point with your clusters above? $s_t = 1$. (1 point)

s_t would go into cluster 2.