CS7641 ML Practice Quiz
Module UL 2: Clustering
Created by Kyle Nakamura with the help of ChatGPT

Question 1
In unsupervised learning, what is the main focus when dealing with unlabeled data?
A. To classify the data based on predefined labels.
B. To make predictions about future data instances.
C. To find a more compact way to describe the data.
D. To use labeled training data to generalize labels.
E. To strictly follow the rules of metric spaces in data organization.

Question 2
What is the primary goal of clustering in unsupervised learning?
A. To assign each object to its own unique partition.
B. To classify data based on supervised learning techniques.
C. To group a set of objects based on their relationships using inter-object distances.
D. To find the smallest possible partition for all objects.
E. To align all data objects along a predefined metric.

Question 3
How does single-linkage clustering determine which clusters to merge?
A. By merging clusters based on the largest distance between points in the clusters.
B. By considering the average distance between all points in the clusters.
C. By merging clusters that have the smallest intercluster distance, defined by the distance between the closest points in the clusters.
D. By randomly selecting clusters to merge until the desired number of clusters is reached.
E. By merging clusters based on a predetermined number of desired clusters.

Question 4

What is the primary characteristic of K-means clustering?

A. It groups points into clusters based on the furthest distance from cluster centers.

B. It begins by randomly assigning points to clusters without recalculating centers.

C. It clusters points based on the closest distance to randomly selected cluster centers and iteratively recalculates centers.

D. It relies on pre-defined cluster centers without iterations.

E. It only considers the mean of the points for clustering without updating cluster assignments.

Question 5

Which of the following algorithms is most similar to K-means in terms of optimization?

A. Genetic Algorithms

B. Simulated Annealing

C. Randomized Hill Climbing

D. Random Restart Hill Climbing

E. Taboo Search

Answer Key

1. C
2. C
3. C
4. C
5. C