## CIS662 Quiz 3 (10/18/2023)

## [Closed-book, open-notes, no communication. Answer each question in the space provided.]

1. (2 points) **Give a 1-dimensional example with k=2, to illustrate that the results of k-means clustering may differ, depending on initialization. Each data point (and each prototype’s initial position) would be just a number.**

Data: 1, 2, 3

First set of initial positions of the two prototypes (cluster centroids): 1, 2

Prototype positions after one iteration of k-means clustering: 1, 2.5

Final positions of prototypes (after convergence of k-means clustering): 1, 2.5

Second set of initial positions of the two prototypes (cluster centroids): 2, 3

Prototype positions after one iteration of k-means clustering: 1.5, 3

Final positions of prototypes (after convergence of k-means clustering): 1.5, 3

1. (1 point) **Mention an advantage of hierarchical clustering over k-means clustering.**

An entire tree (dendrogram) is obtained, interpretable in various ways in addition to obtaining clusters.

Different solutions with different numbers of clusters (k) can be obtained from the tree.

1. (1 point) **What is the expected behavior of a simpler version of the EM algorithm (with Gaussian mixtures), in which the parameters include variances for all dimensions, but not covariances?**

This simplified algorithm will not be able to find clusters whose axes are oriented at an angle to the original dimensions, e.g., if a 2D data set contains two clusters {(2,2), (3,3), (4, 4), (5,5), (6,6)} and {(6, 5), (7, 4), (8, 3)}.

## CIS662 Quiz 3 (10/18/2023)

## [Closed-book, open-notes, no communication. Answer each question in the space provided.]

1. **(1 point) In the mixture model used with the EM algorithm, each Gaussian’s parameters include the centroid and the terms in a covariance matrix. How would the algorithm behave, if we ignored covariances and only used (and updated) centroids and variances for all dimensions?**

This simplified algorithm will not be able to find clusters whose axes are oriented at an angle to the original dimensions, e.g., if a 2D data set contains two clusters {(2,2), (3,3), (4, 4), (5,5), (6,6)} and {(6, 5), (7, 4), (8, 3)}.

1. (1 point) **Mention an advantage of k-means clustering over hierarchical clustering.**

Requires less computational effort and far less space (especially if the data set is large).

1. (2 points) **Give a 1-dimensional example with k=2, to illustrate that the results of k-means clustering may differ, depending on initialization. Each data point (and each prototype’s initial position) would be just a number.**

Data: 1, 2, 3

First set of initial positions of the two prototypes (cluster centroids): 1, 2

Prototype positions after one iteration of k-means clustering: 1, 2.5

Final positions of prototypes (after convergence of k-means clustering): 1, 2.5

Second set of initial positions of the two prototypes (cluster centroids): 2, 3

Prototype positions after one iteration of k-means clustering: 1.5, 3

Final positions of prototypes (after convergence of k-means clustering): 1.5, 3