HW7

Seth Jaksik
1001541359

November 23, 2020

Task 2

This clustering configuration can not be the final result of the k-means algorithm. The mean of the Blue cluster would be approximately the center of the square, and the mean of the Red cluster would be the red dot since there is only one. Then, the L_2 distance between the nearest blue dots to the red one would be smaller than their L_2 distance to the mean of the Blue cluster at the center of the circle, meaning they would be part of the red cluster. This makes the current clustering configuration not the final result of the k-means algorithm.

Task 3

a)

The EM algorithms will not always give the same result. Since EM algorithms use a mixture of gaussians, using a different set of initialization values will produce differing results.

b)

The agglomerative clustering with the d_{min} distance will always produce the same result on the same set of data. This is because there are no ties and the minimum distance between any intermediate clusters will always be the same since we are always using the same data sets. The min distance at each iteration will always be the same.

Task 4

a)

Iteration 1:

$(2, 4)(7)(11)(16)(22)(29)(37)$

Iteration 2:

$(2, 4, 7)(11)(16)(22)(29)(37)$

Iteration 3:

$(2, 4, 7, 11)(16)(22)(29)(37)$

Iteration 4:

$(2, 4, 7, 11, 16)(22)(29)(37)$

Iteration 5:

$(2, 4, 7, 11, 16, 22)(29)(37)$

Iteration 6:

$(2, 4, 7, 11, 16, 22, 29)(37)$

Iteration 7:

$(2, 4, 7, 11, 16, 22, 29, 37)$

b)

Iteration 1:

$(2, 37)(4)(7)(11)(16)(22)(29)$

Iteration 2:

$(2, 29, 37)(4)(7)(11)(16)(22)$

Iteration 3:

$(2, 22, 29, 37)(4)(7)(11)(16)$

Iteration 4:

$(2, 16, 22, 29, 37)(4)(7)(11)$

Iteration 5:

$(2, 11, 16, 22, 29, 37)(4)(7)$

Iteration 6:

$(2, 7, 11, 16, 22, 29, 37)(4)$

Iteration 7:

$(2, 4, 7, 11, 16, 22, 29, 37)$