Name: Sujan Shrestha
UTA ID:1001752468

Assignment 6 Written Task

TASK2

The clustering cannot be final result of K-means algorithm. This is because when we calculate the mean for the shown cluster mean for the cluster of red dots is the red dot itself whereas for the blue dots is around the center. So, we have to reassign the value for the dots by calculating the Euclidean distance between the mean and the dots. Therefore, some blue dots have the Euclidean distance less with the mean red dots rather than the one in the center. Thus, the blue dots that are close to red dots will be assigned as the red cluster changing the given clustering.

TASK3

Part a:

No, the EM algorithm will now always give the same results when applied to the same dataset with same K. This is because to initialize EM algorithm, we initialize μi , σi , and wi values randomly in different ways. So, this will result in giving different results when run multiple times.

Part b:

Yes, agglomerative algorithm will always give the same results when applied to same dataset. This is because agglomerative algorithm clustering merges at each step based on the minimum distance between two closest cluster. For the same dataset the distance between two closest cluster will remain same. Thus, agglomerative algorithm outputs same result given the same dataset.

TASK5

Part a:

Given data set:
2, 4, 7, 11, 16, 22, 29, 37
Applying agglomerative clustering step by step using dmin;
[2, 4] 7, 11, 16, 22, 29, 37
[2, 4, 7], 11, 16, 22, 29, 37
[2, 4, 7, 11], 16, 22, 29, 37
[2, 4, 7, 11, 16], 22, 29, 37
[2, 4, 7, 11, 16, 22], 29, 37
[2, 4, 7, 11, 16, 22, 29], 37
[2, 4, 7, 11, 16, 22, 29, 37]

Part b:

Given data set:

2, 4, 7, 11, 16, 22, 29, 37

Applying agglomerative clustering step by step using dmax;

[2, 4] 7, 11, 16, 22, 29, 37

[2, 4], [7, 11], 16, 22, 29, 37

[2, 4], [7, 11], [16, 22], 29, 37

[2, 4], [7, 11], [16, 22], [29, 37]

[2, 4, 7, 11], [16, 22], [29, 37]

[2, 4, 7, 11, 16, 22], [29, 37]

[2, 4, 7, 11, 16, 22, 29, 37]