Name: Naveenraj Palanisamy

NetId: NXP154130

Machine Learning Assignment (CS 6375.001)-(EM)

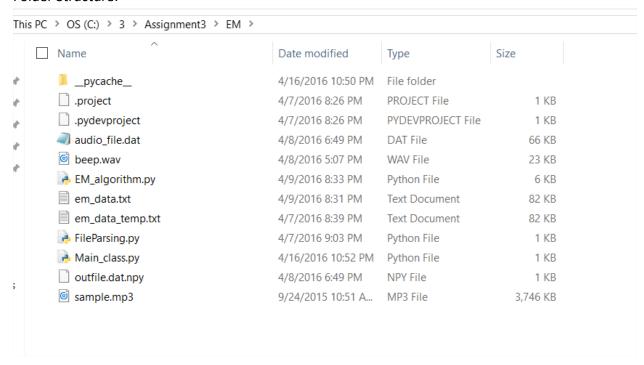
(Note: Runs in python 3.5)

(Note: File Path should be full path)

(Note: This is done as optional home work)

EM:

Folder Structure:



Main_Class.py is the file need to run:

Running the program steps: ->

- 1) Go to 'command prompt'
- 2) Go to folder where python is installed.
- 3) Given command as (python.exe 'full path to-> Main_class.py' Number_of_Iterations varience_value K_value em_Datalcation)

Sample Run:

```
Go to
<sup>Go to</sup>[25.705, 5.8859, 15.6422, 6.9153]
    [0.6749, 0.9238, 0.6637, 7.8733]
mber
   [0.3264, 0.3188, 0.3175, 0.0372]
   [25.5161, 5.5498, 15.5004, 12.3982]
   [0.8838, 0.9049, 0.9087, 7.753]
   [0.3285, 0.3244, 0.3217, 0.0254]
   [25.4966, 5.5088, 15.4638, 13.5277]
   [0.9381, 0.9502, 0.9581, 12.1225]
   [0.3291, 0.3288, 0.3278, 0.0143]
   [25.4894, 5.5093, 15.4554, 15.0722]
   [0.9508, 0.9516, 0.9546, 21.8478]
nple
3\Ass[0.3313, 0.3316, 0.3313, 0.0058]
   [25.4865, 5.51, 15.4527, 15.8182]
   [0.9731, 0.9722, 0.9731, 55.6572]
   C:\Users\NAVE>
```

Algorithm is tested for different initialization values and algorithm works mainly based on the initialization parameter.

Different initialization values are passed as comment line argument.

Initially took values as below:

K: 2

Iterations: 4

Variance: 0.5

Take variance as 1 and run for 3 clusters. Means values are there in image.

According to me setting variance to 1 seems to be best, because it will reduce the change of cluster without any assignment in the initial iterations itself.