**Sai Anirudh Yellanki**

**18XJ1A0548**

**ML Assignment 4 - Fuzzy C Means Clustering**

**NOTE-** **A Brief about the inputs (No need to modify for the code to work, but, can be tweaked.)**

The training size and testing size can be modified within the code.

The code can work on any dataset but is defaulted to Data Set 2, this can be modified by changing the sheet name in the FuzzyCMeansClustering function. Also, if there are rows without data values but with labels like X,Y, at the top, a parameter named rows\_to\_skip can be changed in the same FuzzyCMeansClustering function. For example, Dataset 2 has no rows to skip => 0, Dataset 5 has 2 non data rows => 2 rows\_to\_skip

The minimum and maximum no of clusters can be changed in the same function FuzzyCMeansClustering. It is defaulted to 2 (included), 11 (excluded) => 1 < c < 11.

The code takes in the Number of Training samples and the rest are for testing data. It doesn’t take the ratio. The Number of training samples can be modified in the same function FuzzyCMeansClustering.

**NOTE - Files generated**

2 files are generated by the clustering program(FuzzyCMeansClustering()), these files are read by the classification program(FuzzyCMeansClassification()) and are classified.

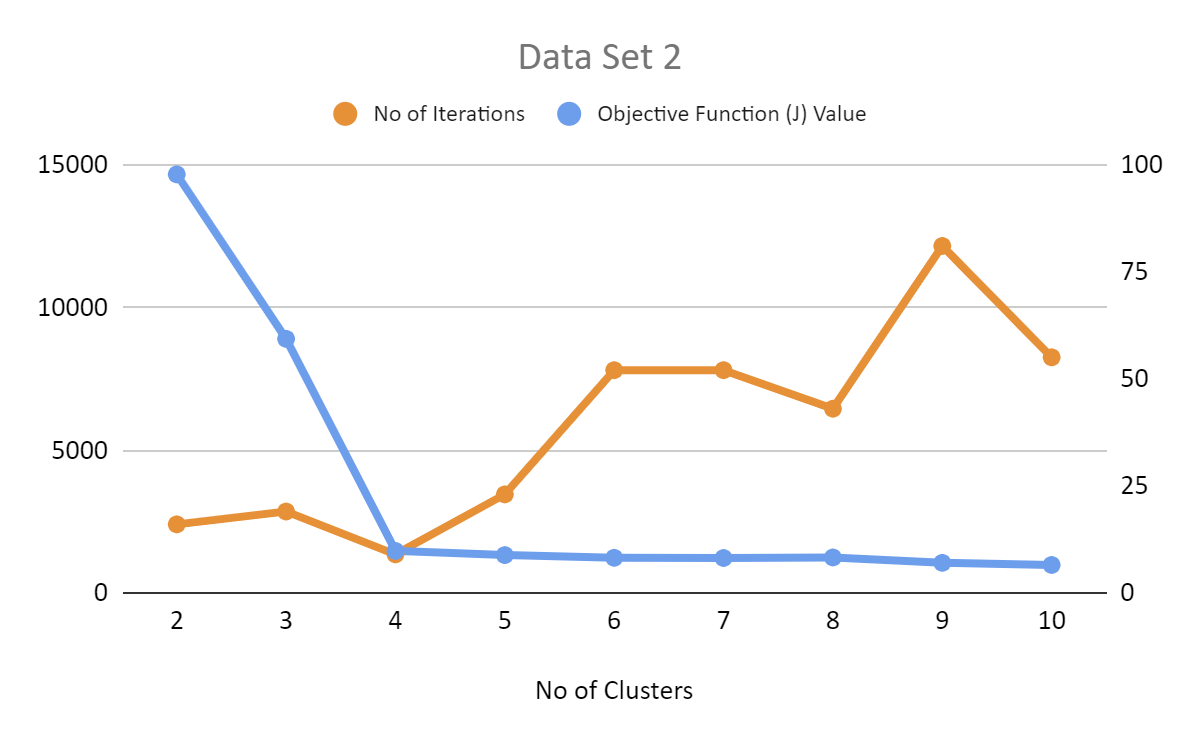
**NOTE - Also included Datasets.xlsx excel file in the Submission folder as it has been said to submit code with “All relevant data to execute the code”**

**Plot of No of clusters vs (No of Iterations and Objective Function (J) Value)**

This is the plot of the number of clusters, no of iterations and the objective function value

Note:- This plot is from excel. The code also displays a plot on running it.

The code to write to excel has been commented.



**The Classification Clustering Plot**

This is the Classification result for the 140 testing data points. The Data Set 2 can be classified into 4 clusters as it can be seen in the Image.

