Lab Assignment 8 **Fuzzy C-means** **Sept 15,2020** 20 marks

CO 3	Develop various classification and clustering techniques.	L4

Fuzzy C-means

In **Fuzzy clustering** each element has a set of membership coefficients corresponding to the degree of being in a given cluster. Points close to the center of a cluster may be in the cluster to a higher degree than points in the edge of a cluster. The degree, to which an element belongs to a given cluster, is a numerical value varying from 0 to 1.I.Fuzzy clustering algorithms seeks to **minimize cluster memberships** and distances, but we will focus on Fuzzy C-Means Clustering algorithm. Fuzzy c-means developed in 1973 and improved in 1981. It's very similar to k-means algorithm in a structure way:

1. Choose number of clusters.

2. Assign coefficients randomly to each data point for being in the clusters.

3. Repeat until algorithm converged (Objective Function C minimizes cluster memberships and distances):

4.Compute the centroid for each cluster

5.Compute each data points' coefficients of being in the clusters.

The main difference with k-means cluster is that **objective function for fuzzy c-means algorithm** allows different cluster membership with probability values, where k-means cluster has strict objective function allows only one cluster membership.

Each question 4 marks.

SPECTF is a dataset to be used.
Analyze the Clusters in data using **Fuzzy C-means**
Do the **Exploratory data analysis** on the data

I. Reading and Understanding the Data
II. Data Cleaning
III. Data Preparation
IV. Model Building using **Fuzzy C-means**
V. Final Analysis-
    a. What is the inference from the clusters formed?
    b. Find the clustering parameters
        - ❖ Estimated number of clusters
        - ❖ Homogeneity
        - ❖ Completeness
        - ❖ Adjusted Rand Index
        - ❖ Adjusted Mutual Information
        - ❖ Silhouette Coefficient