DS481 Data Science
Assignment 04

Due Date: 11th May 2020 **Total Marks: 100**

---

1. Using kmeans algorithm and Euclidean distance to cluster the following 8 points into 3 clusters. Using A1 = (2,10), A2 = (2,5), A3 = (8,4), A4 = (5,8), A5 = (7,5), A6 = (6,4), A7 = (1,2) , A8 = (4,9). Consider initial seeds as A1, A4, and A7. Run algorithm for 1 iteration only. At the end of iteration 1, show [**40 Points**]

    - The new clusters (i.e. the examples belong to each cluster)
    - The center of the new clusters
    - Draw 10×10 space and all 8 points and show the clusters after 1st iteration and the new centroids
    - Without running algorithm again, guess how many more iterations are required to converge. Draw the result of each iteration

2. Using hierarchical clustering algorithms (Single, Complete, Group Average and Distance b/w centroids) and Euclidean distance to cluster the following 8 points into 3 clusters. Using A1 = (2,10), A2 = (2,5), A3 = (8,4), A4 = (5,8), A5 = (7,5), A6 = (6,4), A7 = (1,2) , A8 = (4,9). [**40 Points**]

3. Review a paper of your choice on Soft Clustering e.g. Fuzzy Kmeans (One Page Summary) [**20 Points**]