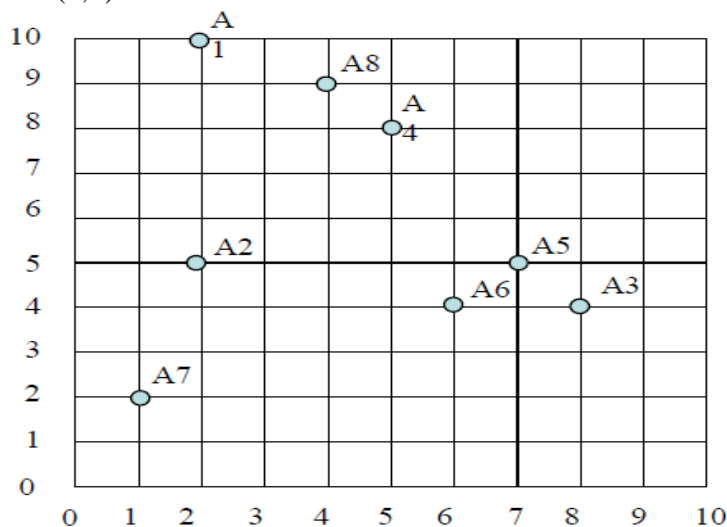# Practical Data Science/Analytics (Clustering)

Use the clustering algorithms discussed in class to solve the following problems.

Problem 1: Iterative K-Means Clustering

Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into 3 clusters: A1=(2,10), A2=(2,5), A3=(8,4), A4=(5,8), A5=(7,5), A6=(6,4), A7=(1,2), A8=(4,9).

Suppose that the initial seeds (centers of each cluster) are A1, A4 and A7. Run the k-means algorithm for 1 epoch(or pass) only. At the end of this epoch show:

a) The new clusters (i.e. the examples belonging to each cluster)
b) The centers of the new clusters
c) Draw a 10 by 10 space with all the 8 points and show the clusters after the first epoch and the new centroids.
d) How many more iterations are needed to converge? Draw the result for each epoch.

Problem 2: Iterative K-Medoid Clustering

Show the result of each epoch on the data given in problem1 by applying K-Medoid algorithm.

Practical Data Science/Analytics (Clustering)

Problem 3: Hierarchical Clustering

Show the result of applying single-link, complete-link, average-link & centroid agglomerative clustering algorithms on the data given in problem1 and also show the dendograms resulting from each algorithm.

Problem 4: Density based DBSCAN Clustering

If Epsilon is 2 and minpoint is 2, what are the clusters that DBScan would discover with the data given in problem1. Draw the 10 by 10 space and illustrate the discovered clusters. What if Epsilon is increased to sqrt(10)?