Systems and Biomedical Engineering Department **Fourth Year / AI in Medicine**
Faculty of Engineering
Cairo University **Due Date 20/01/2021**

Assignment #7 K-means Clustering and Dimensionality reduction

In this assignment, you are to implement K-means Clustering and Dimensionality reduction. The different tasks to reach your goal can be listed as follows:

1. Write your own PCA function that will return:
   - The sorted eigenvectors of principal components
   - The corresponding eigenvalues
   - The transformed data.
   Note, please make sure that the data is centered before transformation(i.e., the sample mean subtracted out) but not normalized.
2. Write your own K-means clustering, where the Euclidian distance is used to evaluate the distance/ similarity between the dataset. The input to this function should be:
   - The number of clusters
   - The dataset.
   Whereas , the output should be:
   - The list of cluster centers
3. Apply the k-means clustering with 4 clusters and Scatter Plot the points clusters along with the cluster center(use symbols to distinguish the cluster center from cluster point and the color codes to distinguish between the 4 clusters,
4. Apply PCA to the Old Faithful dataset as a preprocessing then Use only the first component once and the first two components one other time. Draw the aforementioned scatter plot in both cases and comment on the results.

General Instructions

1- This is an individual task.
2- The source code as well as the report describing your functions and output should be submitted through Google Classroom.
3- Make sure to clearly add comments to your code describing each function.
4- The due date for the submission of this phase is Sunday, January, 20, 2021 at 12:00 am.
5- Please Review the definition of cheating in the first presentation.