Team assignment #4 - clustering (each person is required to upload the same team report)

Use the following two methods to cluster the given dataset (D1_2.amt).

1) K-means
2) EM and Gaussian mixture models (GMM)

Due date: 2020/4/16 11:59pm

Report: No more than 4 pages in IEEE journal publication format. Your report pdf file should be zipped together with your code prior to your upload (aka you only upload one item, not two).

- Your report should be self-contained, meaning that a reader with reasonable engineering background but without specific knowledge on K-means or EM and Gaussian mixture models can understand the basic steps you took to perform clustering on the given dataset.
- Respectively build your clustering models with 3, 5 and 7 clusters (3 clustering models for K-means and 3 for GMM, respectively). For each model, perform a PCA and scatter plot the first two principal components with different classes labeled with different colors. Also, plot the K-means cost function over iterations, and plot the log-likelihood values over EM iterations.

Code: clearly annotate your code so your TA can verify your results.

File name convention: Team#-Assignment4-Clustering