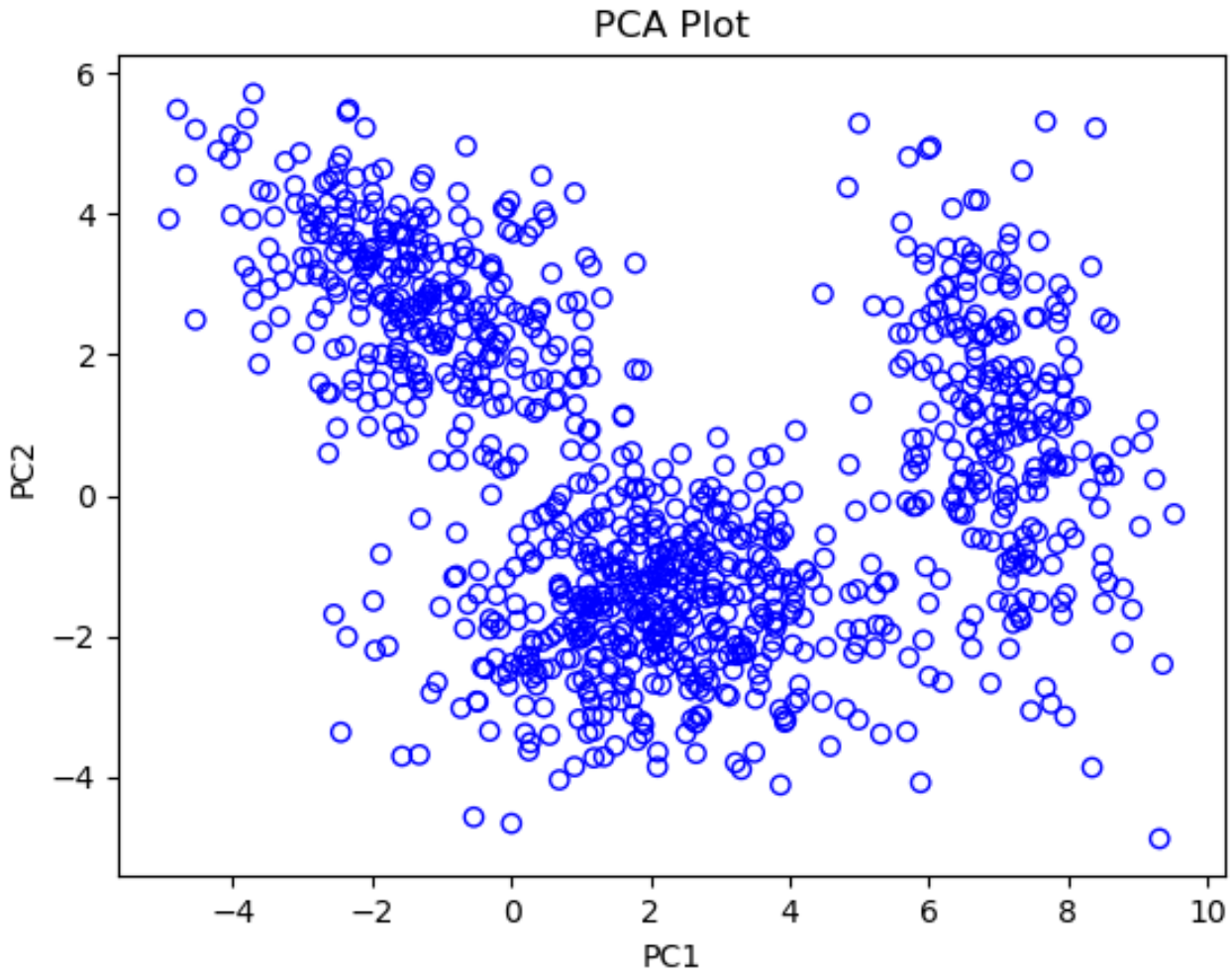
CSE 472 (Machine Learning Sessional)

Assignment 3: Dimensionality Reduction using Principal Component Analysis and Clustering using Expectation-maximization Algorithm

Moaz Mahmud (1505064)

PCA implementation



EM implementation

From PCA, number of clusters in the data in two principal components,

$$\boxed{k = 3}$$

$$\boldsymbol{\mu_1} = \begin{pmatrix} 7.07860183 \\ 0.99051031 \end{pmatrix} \qquad \boldsymbol{\Sigma_1} = \begin{pmatrix} 0.82725394 & -0.57770596 \\ -0.57770596 & 3.36105752 \end{pmatrix}$$

$$\boldsymbol{\mu_2} = \begin{pmatrix} 2.1049187 \\ -1.63644085 \end{pmatrix} \qquad \boldsymbol{\Sigma_2} = \begin{pmatrix} 2.5259442 & 0.16589947 \\ 0.16589947 & 1.09054177 \end{pmatrix}$$

$$\boldsymbol{\mu_3} = \begin{pmatrix} -1.30281985 \\ 2.79003276 \end{pmatrix} \qquad \boldsymbol{\Sigma_3} = \begin{pmatrix} 1.99641987 & -0.87684847 \\ -0.87684847 & 1.47243505 \end{pmatrix}$$

$$\boldsymbol{w_1} = 0.24037177 \qquad \boldsymbol{w_2} = 0.45676345 \qquad \boldsymbol{w_3} = 0.30286478$$

Python Format

```
np.array([[ 7.07860183,  0.99051031],
          [ 2.1049187 , -1.63644085],
          [-1.30281985,  2.79003276]])

np.array([[[ 0.82725394, -0.57770596],
           [-0.57770596,  3.36105752]],

           [[ 2.5259442 ,  0.16589947],
            [ 0.16589947,  1.09054177]],

           [[ 1.99641987, -0.87684847],
           [-0.87684847,  1.47243505]]])

np.array([0.24037177, 0.45676345, 0.30286478])
```