Assignment-4(Report)

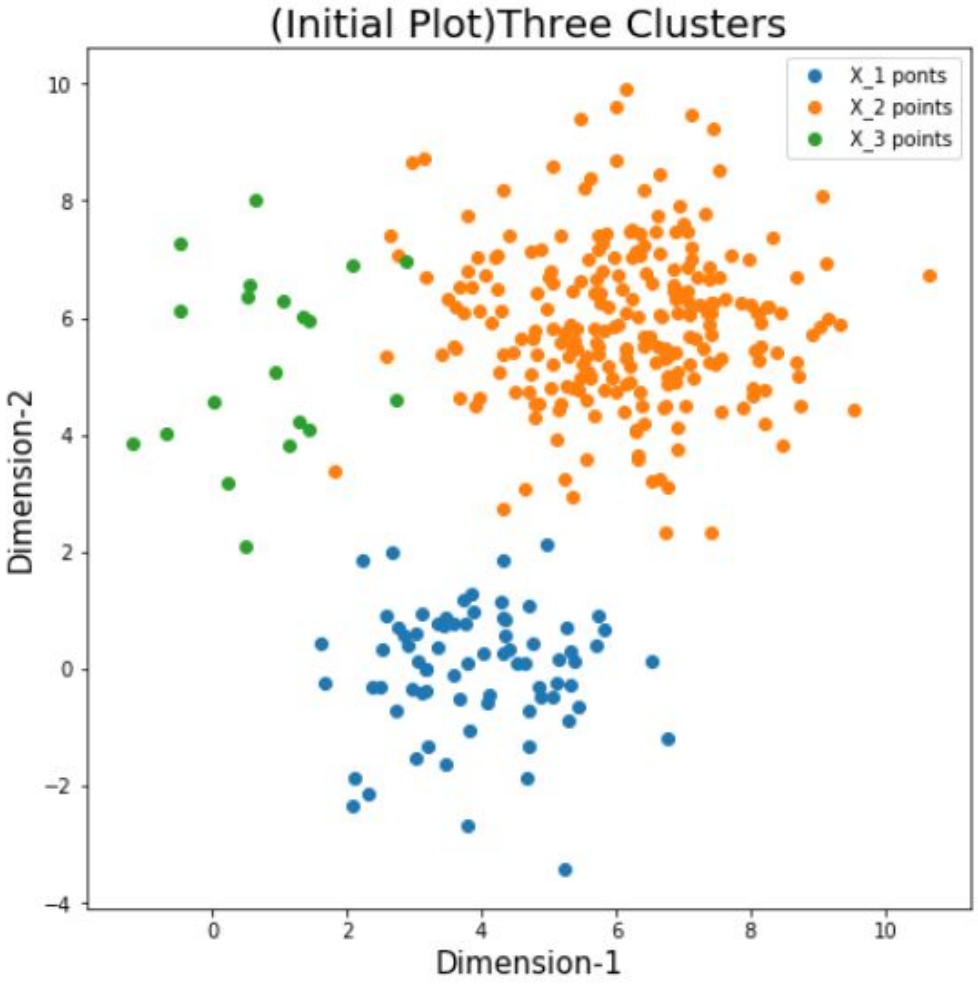
Task-1 :

Generated three clusters as mentioned in the question. The plot of the clusters is shown below.

X_1 = np.random.multivariate_normal(mean=[4, 0], cov=[[1, 0], [0, 1]], size=75)

X_2 = np.random.multivariate_normal(mean=[6, 6], cov=[[2, 0], [0, 2]], size=250)

X_3 = np.random.multivariate_normal(mean=[1, 5], cov=[[1, 0], [0, 2]], size=20)

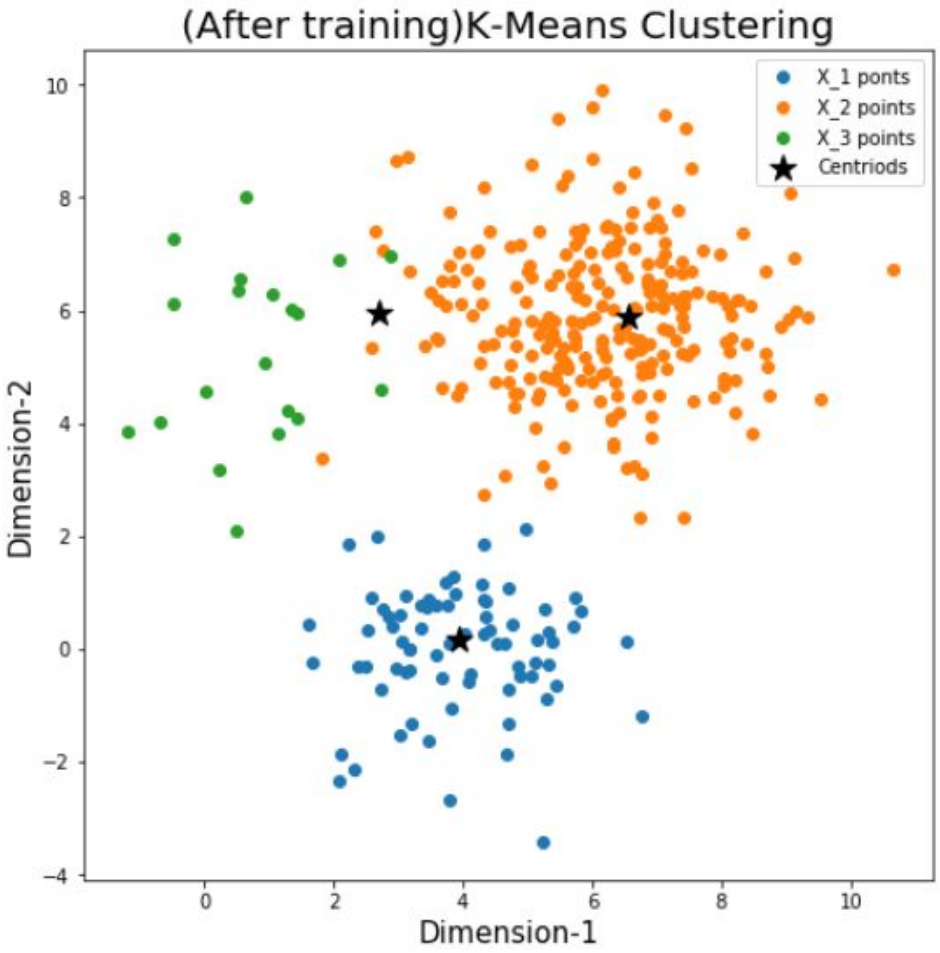


(three different colors indicate the kind of cluster they belong to)

Task-2:

Given the generated samples as input to the k-means clustering algorithm with the number of clusters as 3, for training. The k-means did well in finding the three centroids as there are almost close to their initial cluster they belong to. The plot of the k-means clustering is shown below. The initialization of the centroid was done at random.

```
The centriods found for three different clusters are
[[6.56298766 5.89701945]
 [3.95071338 0.16026379]
 [2.71242971 5.95924595]]
```

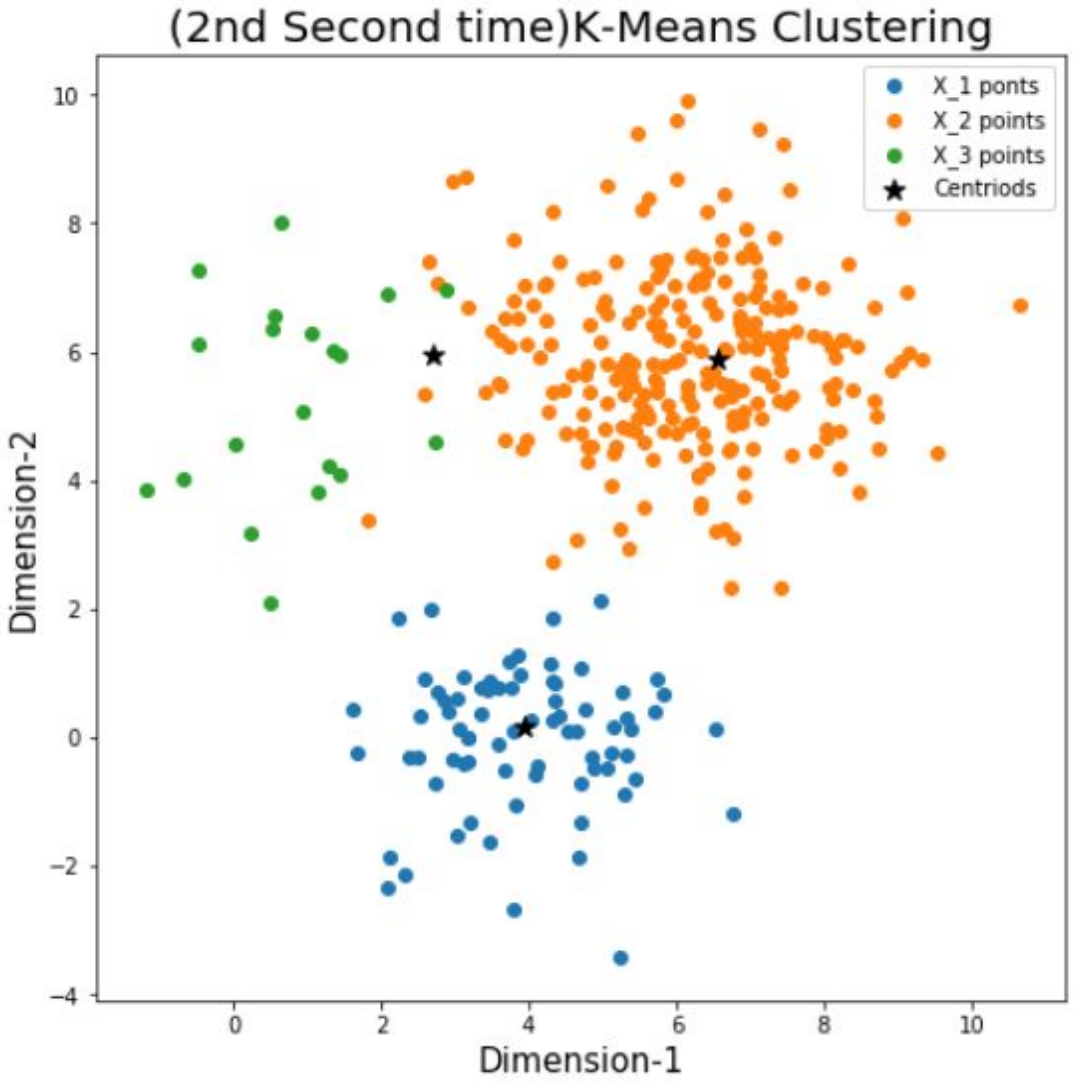


(We can see that black stars(centroids) found were close enough to the initial clusters),

For another random different position of the center, The plot of that is shown below

```
The centriods found for three different clusters are
[[2.71242971 5.95924595]
 [6.56298766 5.89701945]
 [3.95071338 0.16026379]]
```

(We can see that the found centroids are similar but not the same. It also indicates the centroids converged well),

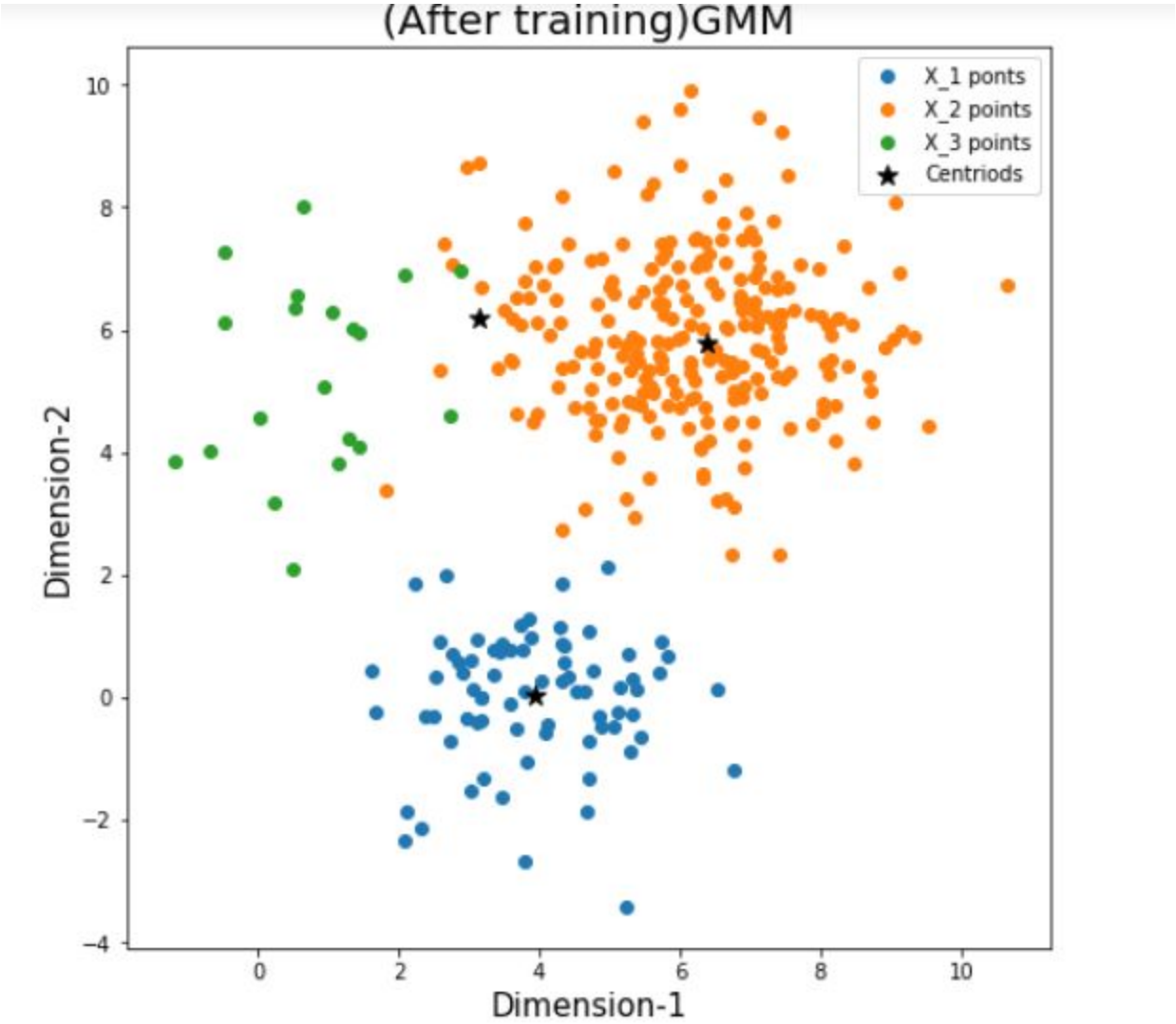


(The second one also done well like the first one. It was able to find good centroids although some points are misclassified),

Task-3:

Given the generated samples as input to the GMM clustering algorithm with the number of clusters as 3, for training. The GMM also did well in finding the three centroids as there are almost close to their initial cluster they belong to. The plot of the GMM clustering is shown below. The initialization of the centroid was done at random.

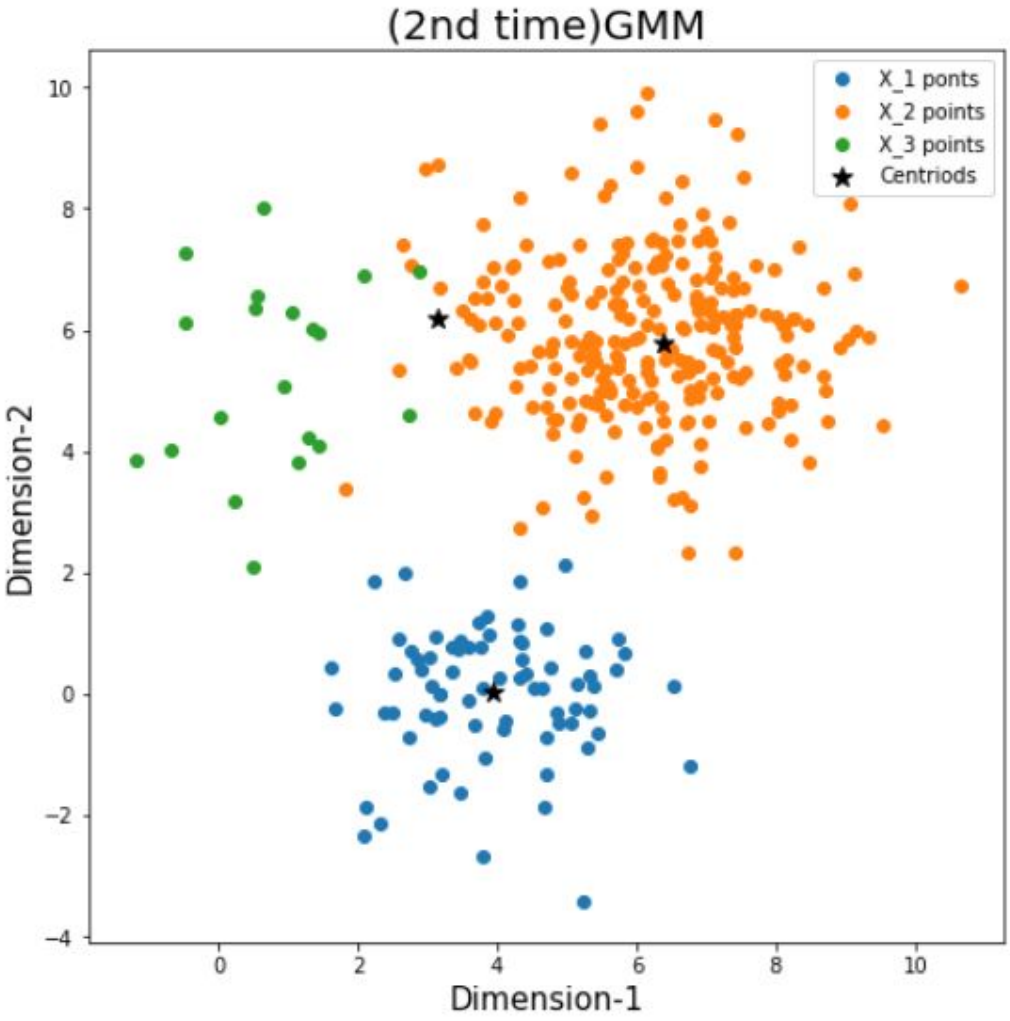
```
The centriods of the GMM-model are
[[3.13874139 6.18202053]
 [3.93666407 0.02417553]
 [6.37208084 5.77807491]]
```



For another random initialization,
The centroids found this time is,

```
[[3.93666407 0.02417553]
 [3.13874139 6.18202053]
 [6.37208084 5.77807491]]
```

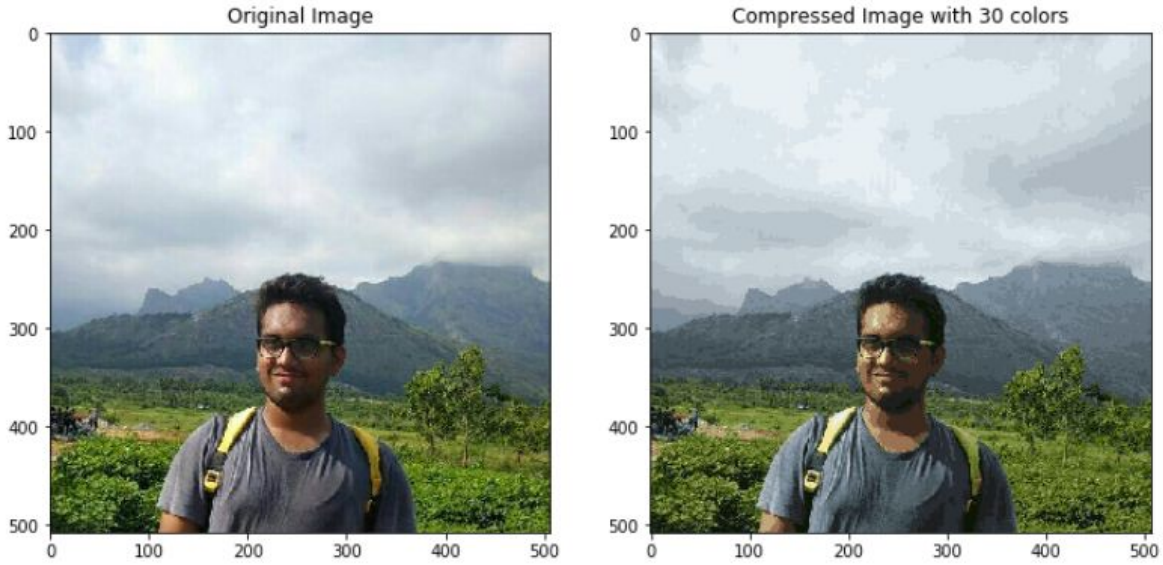
The plot of the GMM-model,



(GMM model did considerably well.)

Now, on comparing K-means and GMM, K-means has performed well on the sample generated. K-means was having a higher score than the GMM, the number of points correctly classified was higher in K-means model than the GMM model.

Task-4(Image Compression):



We can see that the compressed image is almost the same as the original one.

Yes, the compressed image size gets decreased.

The initial image size was 508*507*3
So the total number of bits required was **508*507*3*8** (each pixel is 8 bit) **= 6181344** bits.
Therefore the initial image needed **6181344** bits.

Now, the compressed image needs **30*3*8**(for the 30 different pixels) **+ 508*507*4**(4 bits are enough to map the 30 different pixels) **= 1030944**
The compressed image needs **1030944** bits.

We can see that there was a significant drop of **5150400** bits and the compressed image size is significantly smaller than the initial image