**Assignment 4**

Machine Learning, SS2021

|  |  |  |
| --- | --- | --- |
| **Team members** | | |
| Last name | First name | Matriculation Number |
| Rodiga | Luca | 1620972 |
| Ertin | Alpertunga | 12045064 |

1.1)

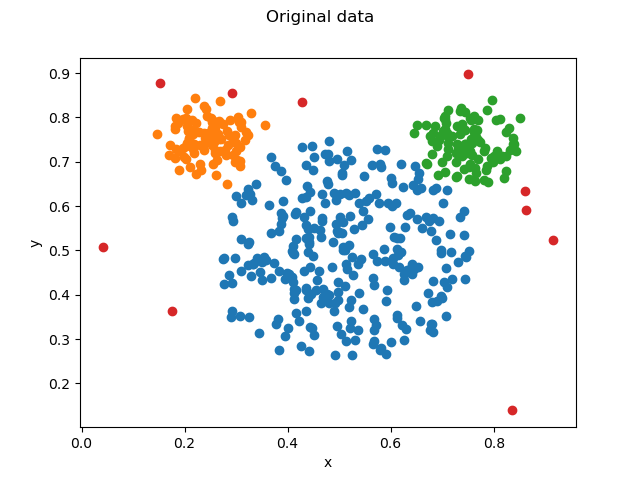
￼
￼￼￼
￼￼￼
￼￼￼￼￼￼￼￼￼
￼
Ink Drawings


￼
￼￼￼
￼￼￼
￼￼￼￼￼￼￼￼￼
￼
Ink Drawings

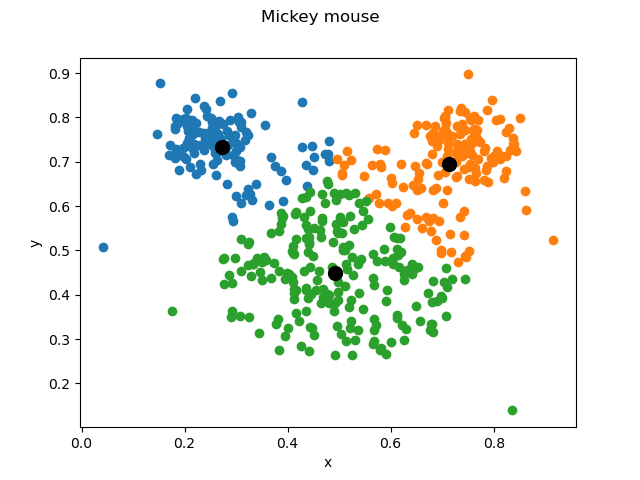

￼
￼￼￼￼￼
￼￼￼
￼￼￼￼
￼￼￼￼￼
￼￼
￼￼￼
￼
￼

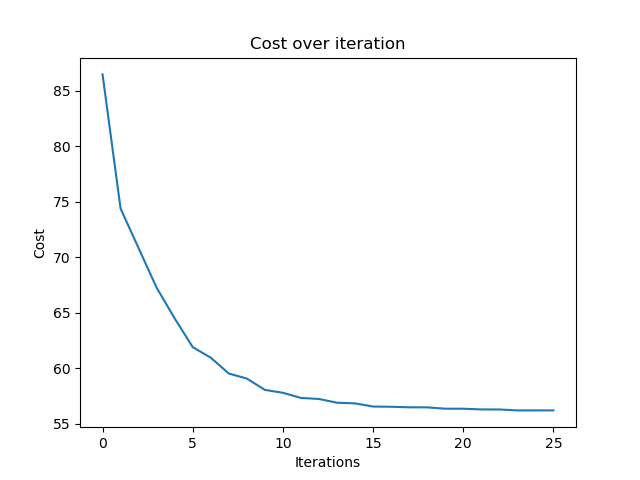

￼￼￼￼￼￼￼
￼


2.1)



K = 3



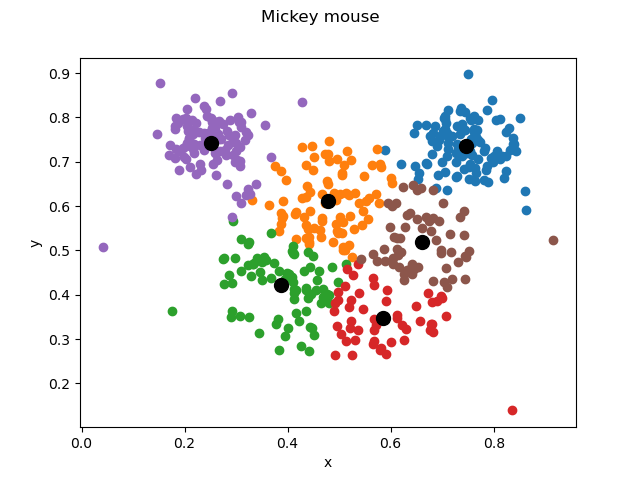


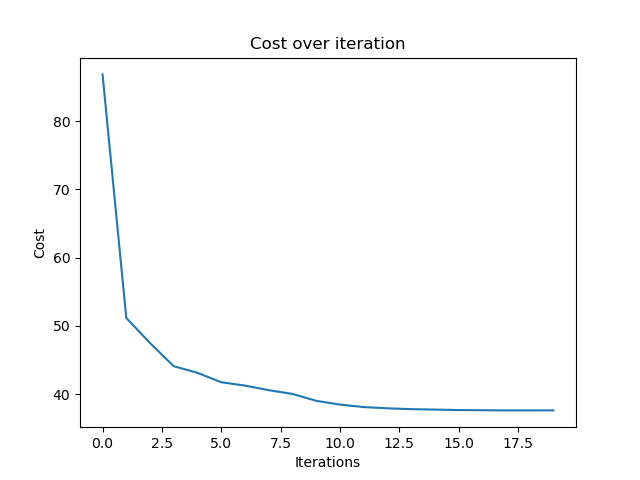
K-Means with K= 3 offered the most accurate clustering, compared to the original data plot, as we still have one cluster for each ear and one cluster for the head.

The algorithm converged in iteration 12.

Compared to the original data, it can be seen, that it was not possible for K-Means and K = 3 to achieve the sharp edges between the head and the ears, because the outliers distorted the centroids.

K= 6





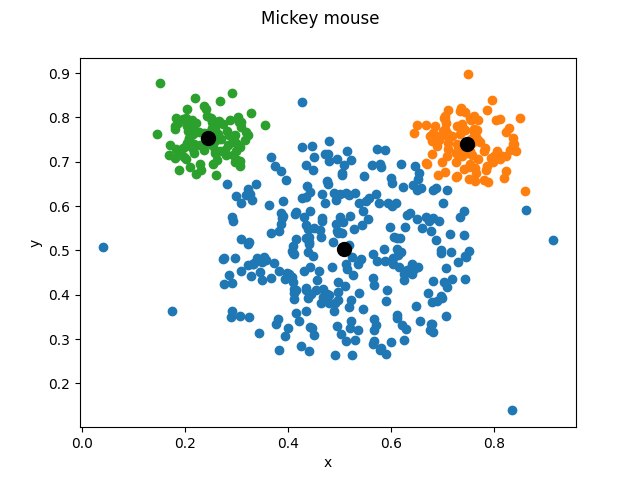
K-Means with K = 6 allows for a better separation of the head and the ears compared to K = 3, but we do have 4 clusters for actually only one head in the original data.

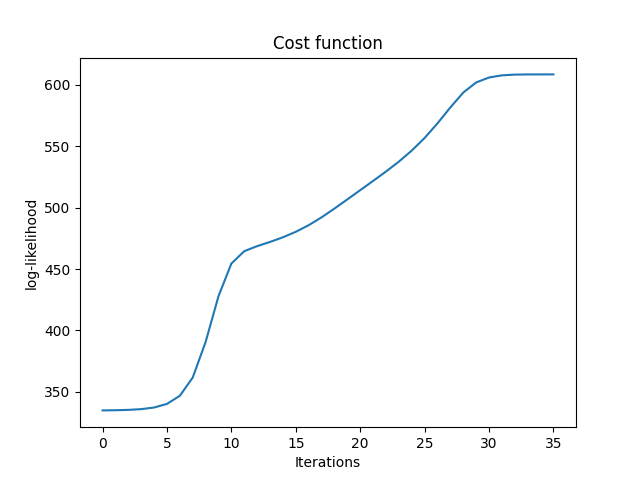
2.2)

When we compare the below plots to the original image, we see that the ears are assigned nicely to two components. We also see some of the outliers will be assigned to the components with respect to covariance, in the K=3 case this component has the largest pi value.

For K=3 and initial pi=[0.33333333 0.33333333 0.33333333]

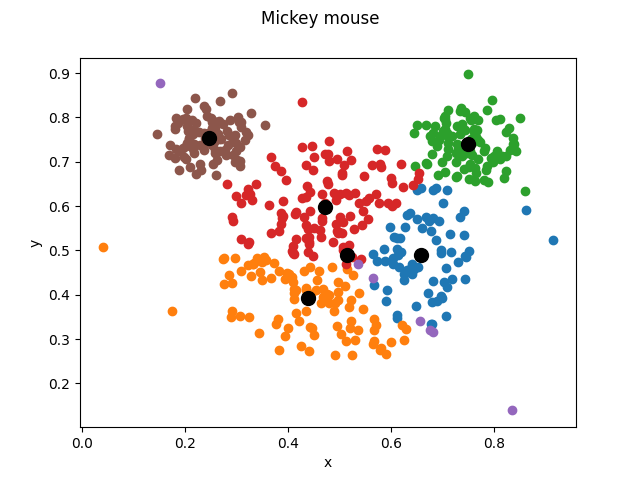
Final pi=[0.60067642 0.20135682 0.19796676]

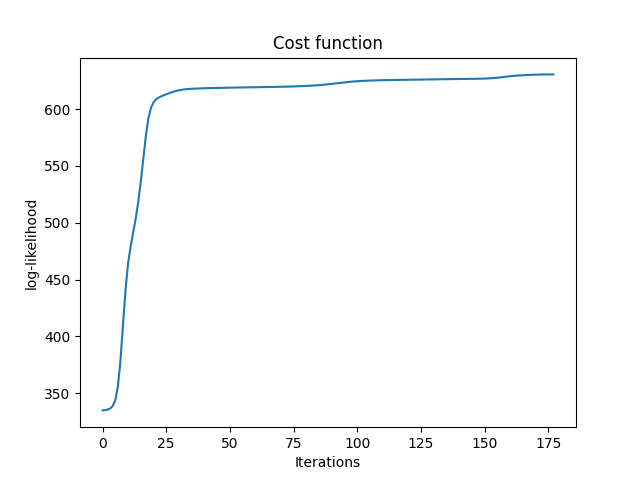




For K=6 and initial pi=[0.16666667 0.16666667 0.16666667 0.16666667 0.16666667 0.16666667]

Final pi=[0.15113238 0.18867002 0.20041319 0.24602091 0.01658345 0.19718006]

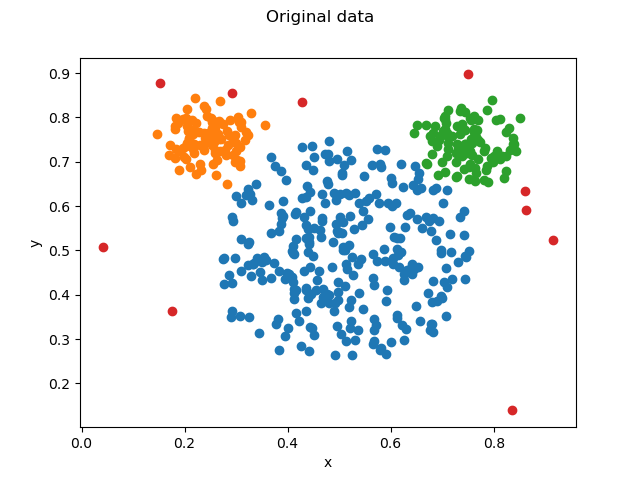




2.3)

2.3.1) Although EM takes more iterations, it deals with the ear shapes better and can handle the outliers better.

2.3.2) Yes, as you can see the red dots in the original data are outliers.



K-Means is **not** robust against outliers, because centroids are always calculated with the mean of all assigned datapoints of the cluster and therefore, also always take the outliers into account, which can hinder performance, if there are too many outliers.

EM can deal with the outliers a little better than K-means because they will not be assigned to the closest means, however it is not hundred percent robust.

2.3.3) K-Means uses hard assignment to clusters while EM uses soft assignments.

2.3.4) For the initial values of pi the discrete uniform distribution is used. At the end of our algorithm the pi values correlate to the sizes of the clusters. For example the largest cluster (face of the mickey mouse) relates to the largest pi.

2.3.5) Since we use the L2 norm as our cost function K-means tends to produce circular shapes here. Since the variance is not important we can assume them to be a zero matrix.