Ex1 - Lior Ziv

November 16, 2017

MNIST Digits

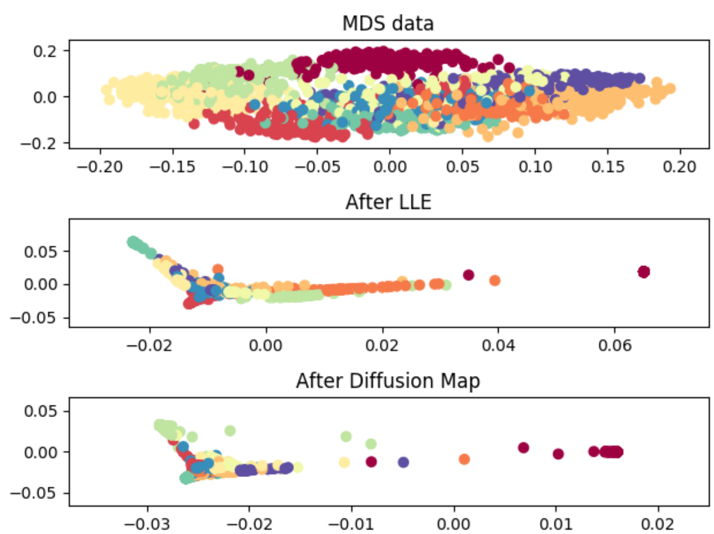

Figure 1:

- Here we are trying to take images of digits (64*64) reduce their dimension to 2d and display the points over a scatter plot,

  we expect points with the same color(meaning they came from the same label - digit) to appear close in the plot.

- We can see that MDS compared to LLE and Diffusion map takes most of the points and put them close together with less overlap between digits but in a relatively big cloud.

Swiss Roll

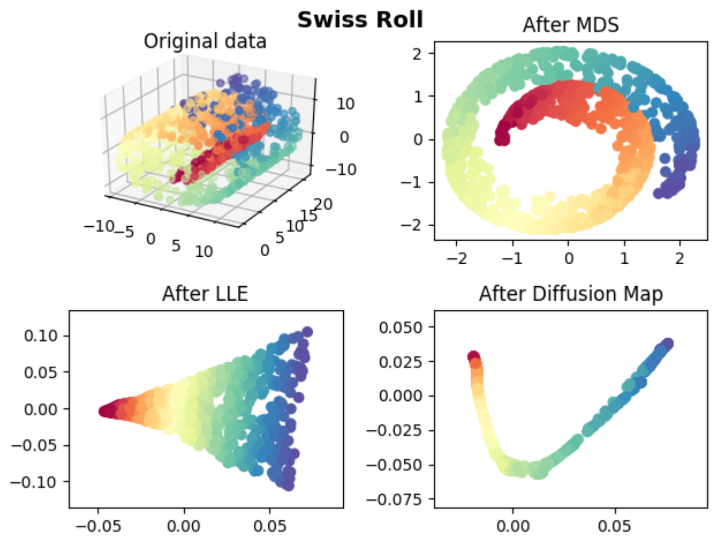

Figure 2:

- In manifold learning we are recovering a global shape from local difference affinity while using local differences.

  In the case of the swiss roll euclidean distance is not the best therefor a PCA won't do the job.

  - LLE works good in this case since it actually expresses each point in a linear sub-space created by it's nearest neighbors
  - Diffusion Map takes into account all the relation between points x and y while calculating the distance and serves as a better notion of proximity than just Euclidean distance(in LLE we lose that notion)
  - MDS works best when the data is approximately on a subspace, but in our case we have many sub spaces therefore MDS result is not satisfying

Faces

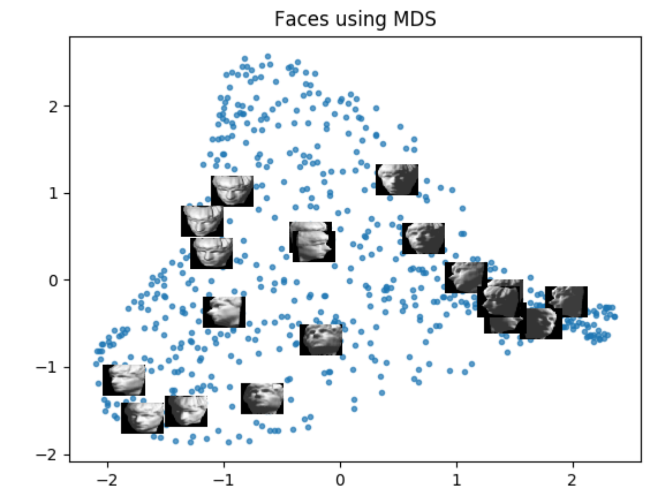

Figure 3:

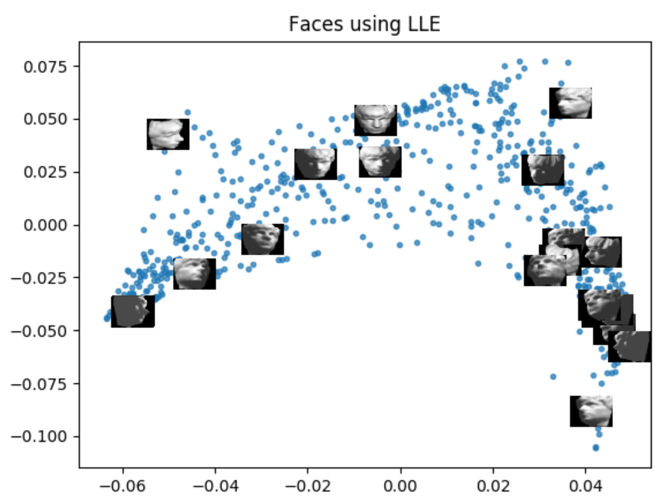

Figure 4:

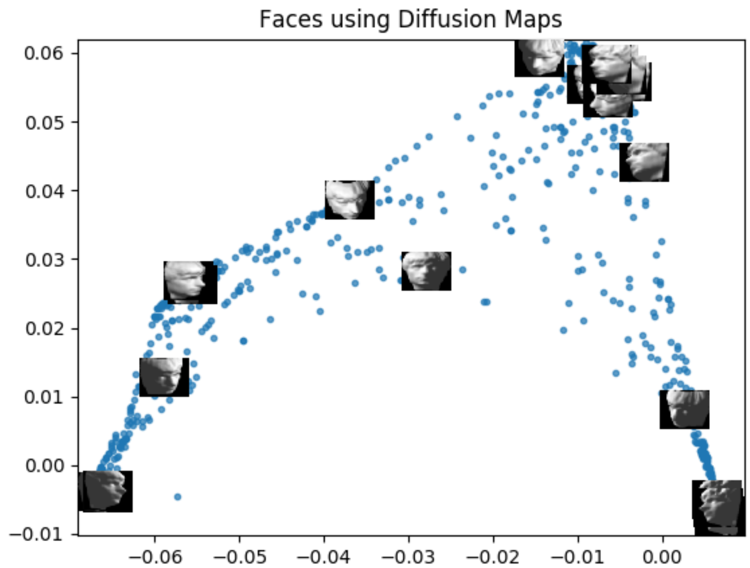

Figure 5:

- Using dimension reduction on this faces set is suppose to visualize in 2D the biggest motifs in those pictures, we can actually see that the axis express the faces rotation

Scree Plot

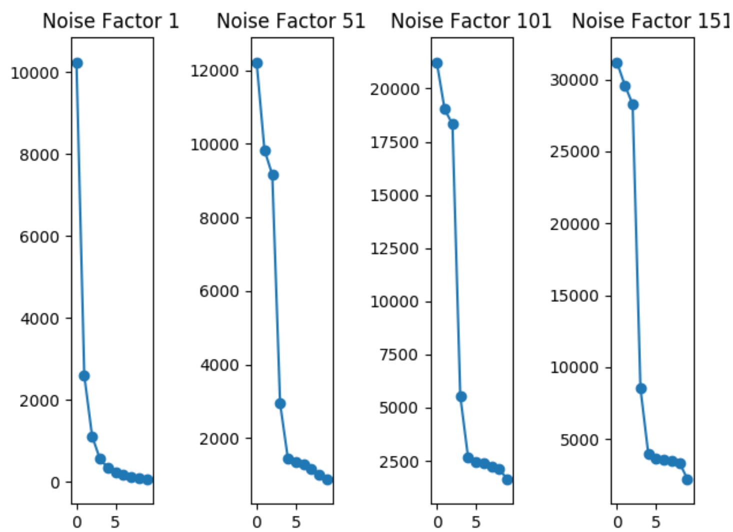

Figure 6:

- We can see the “knee” we are getting, first when we have less noise at the 2D and as the noise grows it gets much more tough to

  MDS algorithm to get the actual dimension of the embedded data.

Lossy Compression Distance

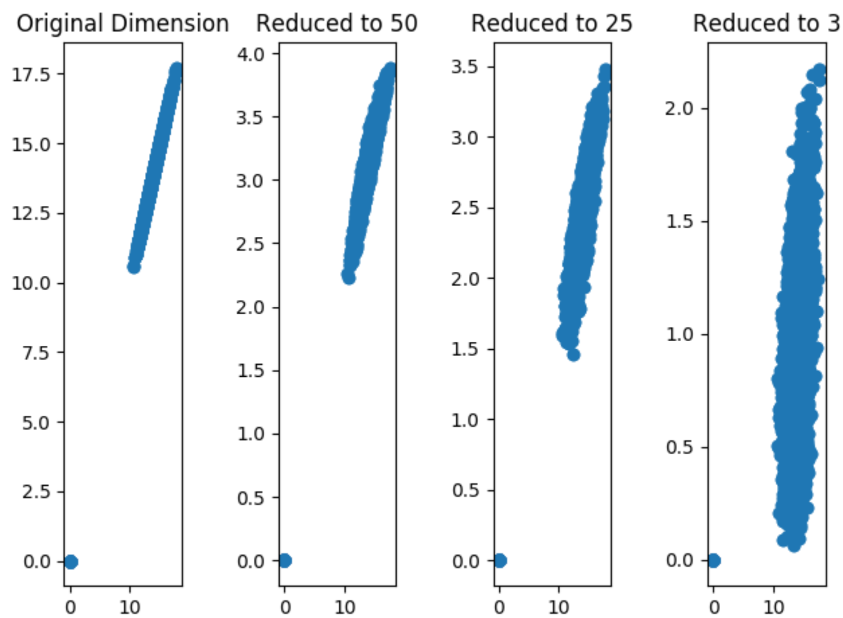

Figure 7:

- As we reduce the dimension of a given data the distance matrix deviate more from the original one, meaning the distance between points change and grow as we reduce the dimension to a lower one, since the original data losses more and more information.

Parameter Tweaking

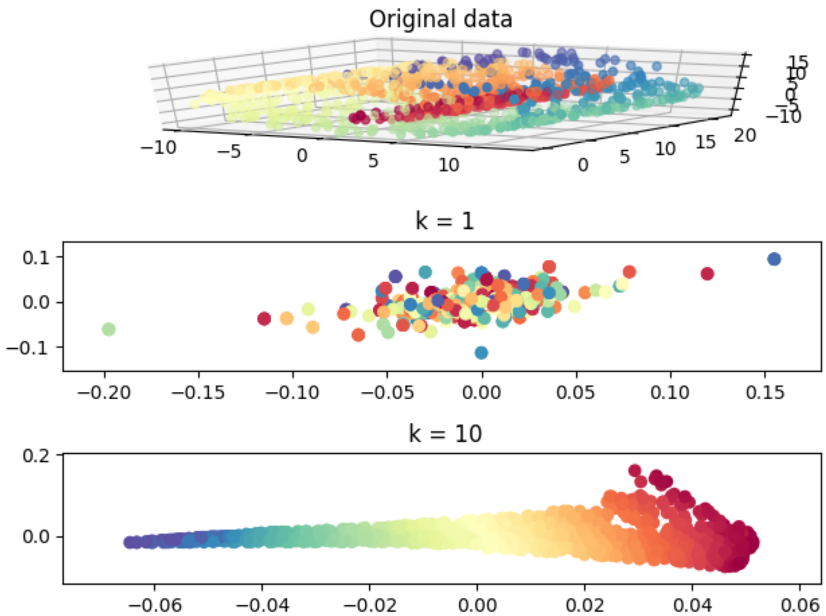

Figure 8:

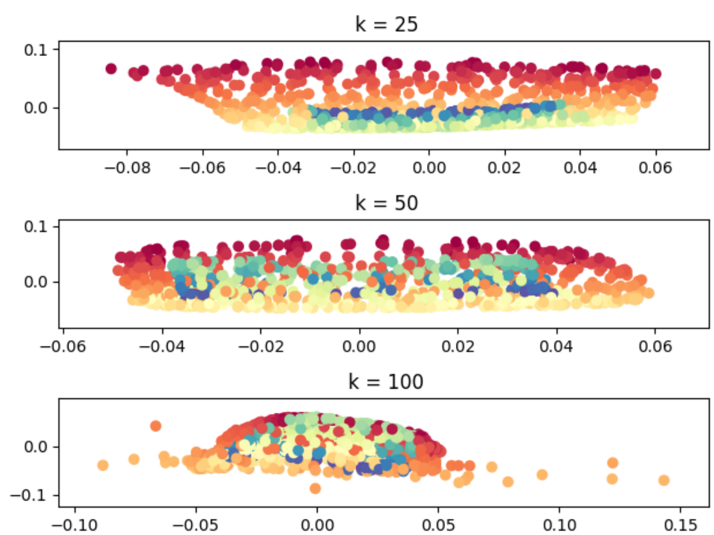

Figure 9:

- Here I performed LLE with different neighbors number using the swiss roll data set
  - For 1 neighbor we see that each point got her own color which is not interesting since we can't really separate the data - overfitting
  - Above 1 we start to see a behavior we would like to have since we can actually separate the data to different groups and describe it with less information
  - As the number of neihbors grows we will get underfitting since we will separate to less and less groups meaning we will think to really different pionts are close to each other, as we see in the 100 example the points start to shrink and get closer and closer to each other

feedback

1. It took me about 3 days
2. Just with choosing the parameter but actually I think I learned from it :)
3. No