

CMPE 462 PROJECT 3

Eray Kurtulus, Şadi Uysal
2015400159, 2015400162

June 2020

Task 1

In this task, we are asked to implement k-means clustering algorithm with the number of iterations as the stopping condition and use it on the provided data.

Below is the provided data and their labels represented by their color:

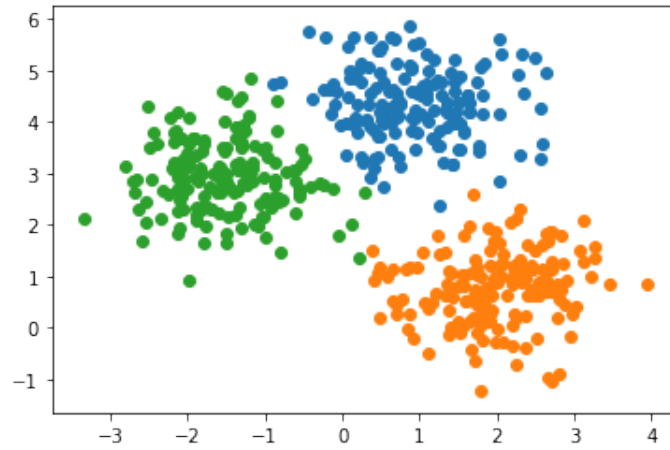


Figure 1: The provided data, colored according to their label

Below, we have shown how our implementation has clustered the provided data in 8 iterations:

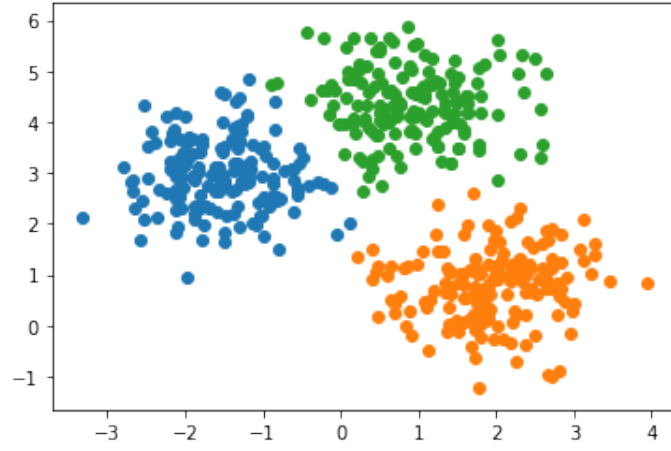


Figure 2: Our implementation, 8 iterations

And here is how our implementation works in 1-to-9 iterations:

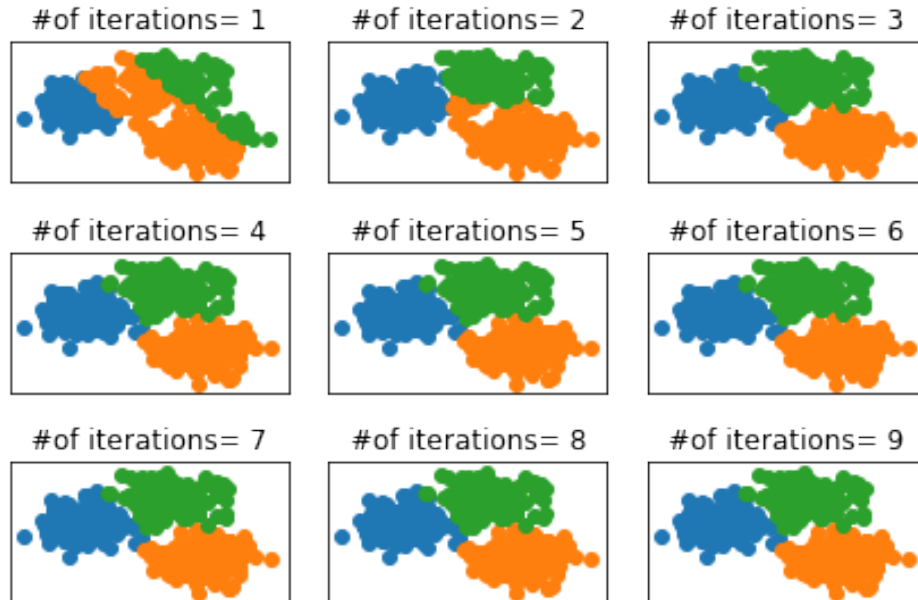


Figure 3: Our implementation

Task 2

In this task, we are asked to implement principal component analysis (PCA) and apply it to the provided dataset which contains 3000 16x16 images of handwritten digits.

Our PCA implementation consists of the following steps:

- Standardizing the dataset
- Forming the covariance matrix
- Computing its eigenvectors
- Using the first d eigenvectors to form the d Principal Components
- Forming the transformation matrix G

In the figures below, we have visualized the PCA-applied and reconstructed images at indices $i = 0, 500, 1000, 2000$ for $d = 50, 100, 200, 256$

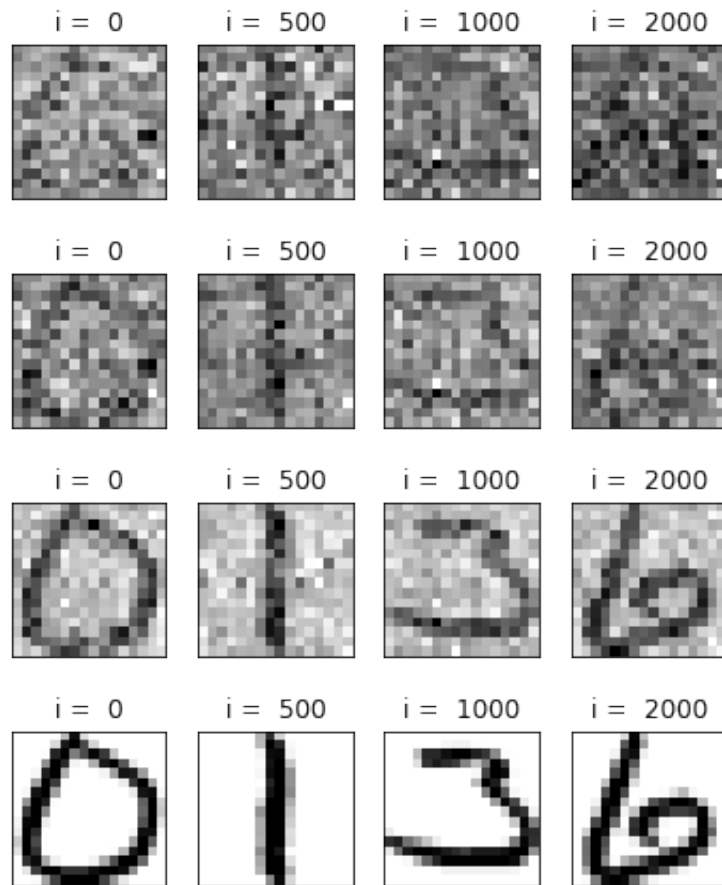


Figure 4: Reconstructed images, $d=50,100,200,256$

These visualized results are exciting to see for us, as this was our first time compressing images. It is also quite an easy process. However, the image quality quickly deteriorates with compression.