

CSE 472: Machine Learning Sessional

Report on Assignment 4:

Principal Component Analysis (**PCA**) and
Expectation Maximization(**EM**) algorithm

Sabah Ahmed
1905118

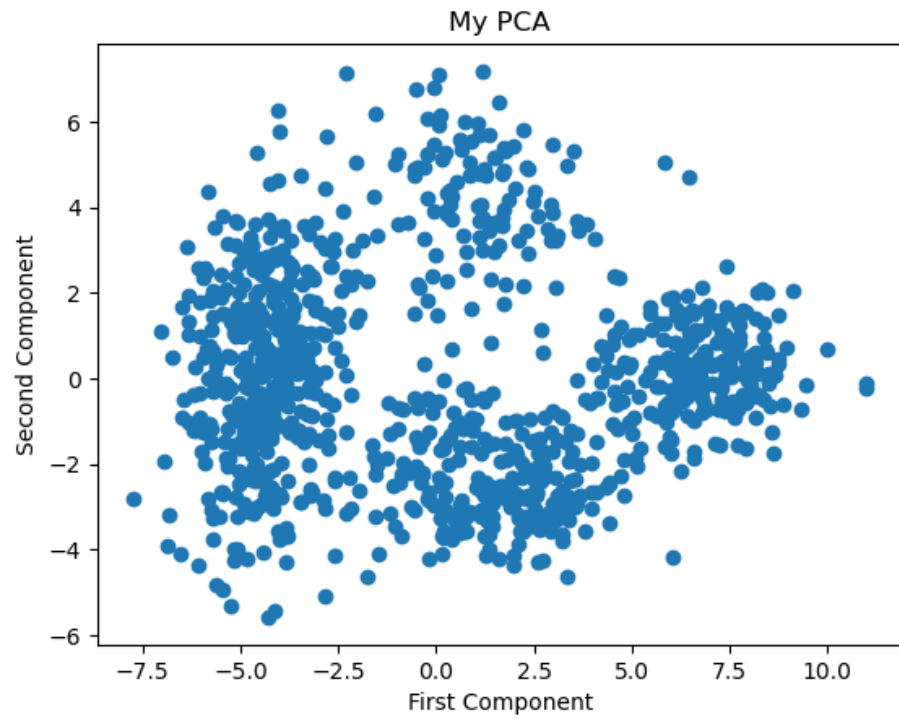
November 21, 2024

1 Instructions For Running Code

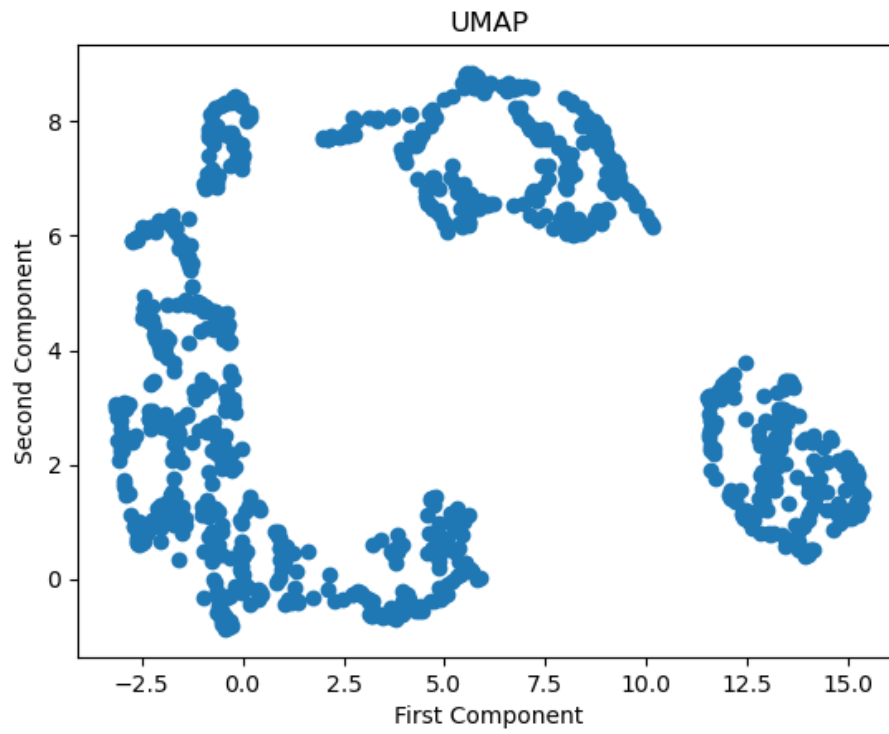
- Keep the **pca_data.txt** file in the same folder as **1905118_pca.ipynb** file.
- Keep the **em_data.txt** file in the same folder as **1905118_em.ipynb** file.
- Hit Run All in **1905118_pca.ipynb** file to generate the results for **PCA** task.
- Hit Run All in **1905118_em.ipynb** file to generate the results for **EM** task.

2 Principal Component Analysis (PCA)

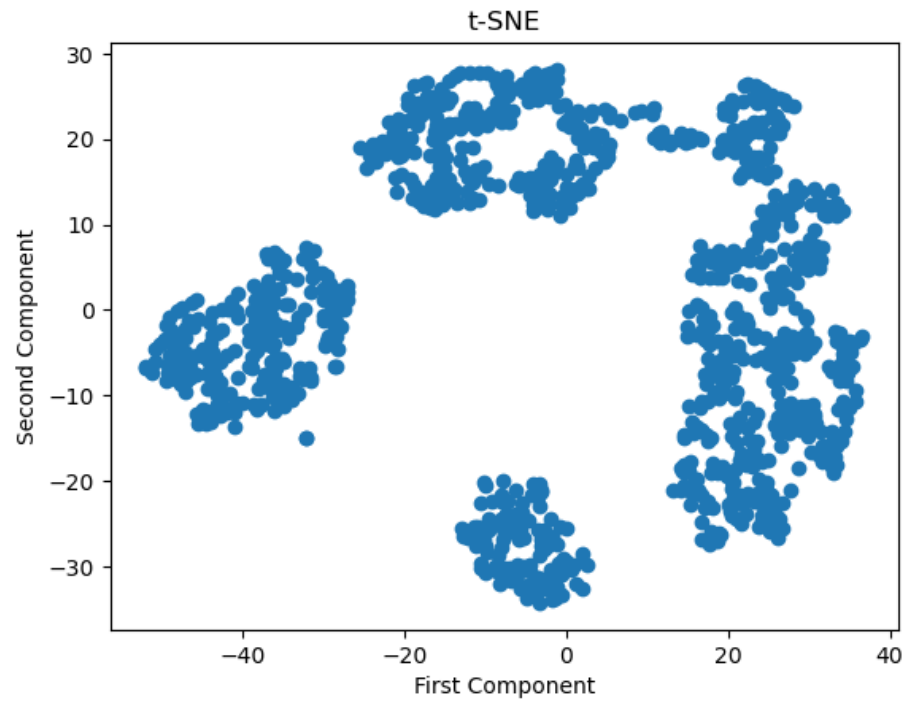
2.1 Plot of PCA (own implementation)



2.2 Plot of UMAP (library)



2.3 Plot of tSNE (library)

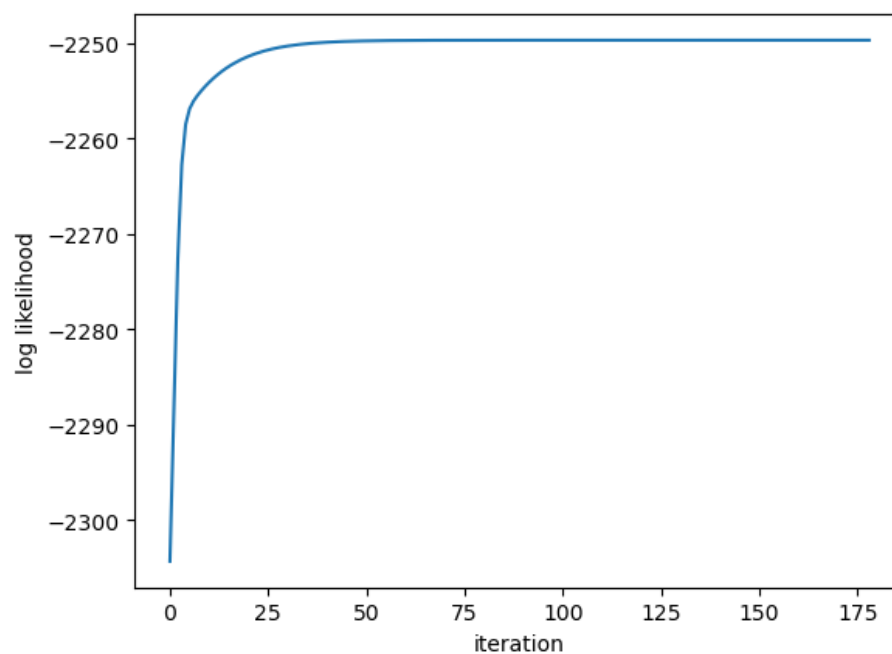


3 Expectation Maximization (EM)

3.1 Estimation

- Estimated Mean Number of Children for Family with Family Planning Advice: **1.7824294239512446**
- Estimated Mean Number of Children for Family without Family Planning Advice: **4.910737191986101**
- Estimated Proportions of Families with Family Planning Advice: **35.60%**
- Estimated Proportions of Families without Family Planning Advice: **64.40%**

3.2 Log Likelihood Convergence Plot



3.3 Plot of Frequence Histogram and Estimated Distribution

The estimated probability distribution is given by:

$$P(x) = 0.64398094 \cdot \frac{(4.91073719)^x e^{-4.91073719}}{x!} + 0.35601906 \cdot \frac{(1.78242942)^x e^{-1.78242942}}{x!}$$

