INDEX

Machine Learning Offline 4 Report

1 Introduction

Principal component analysis (PCA) and the expectation-maximization (EM) algorithm are two of the most widely used unsupervised methods in machine learning. In this assignment, we implement PCA for dimensionality reduction and implement the EM algorithm for the Poisson mixture model.

2 Code Run Instruction

1. Install the necessary packages
    - `numpy`
    - `pandas`
    - `matplotlib`
    - `scipy`
    - `scikit-learn`
    - `umap-learn`
2. Just run all the cells from the begining to the end.

3 Plots for PCA

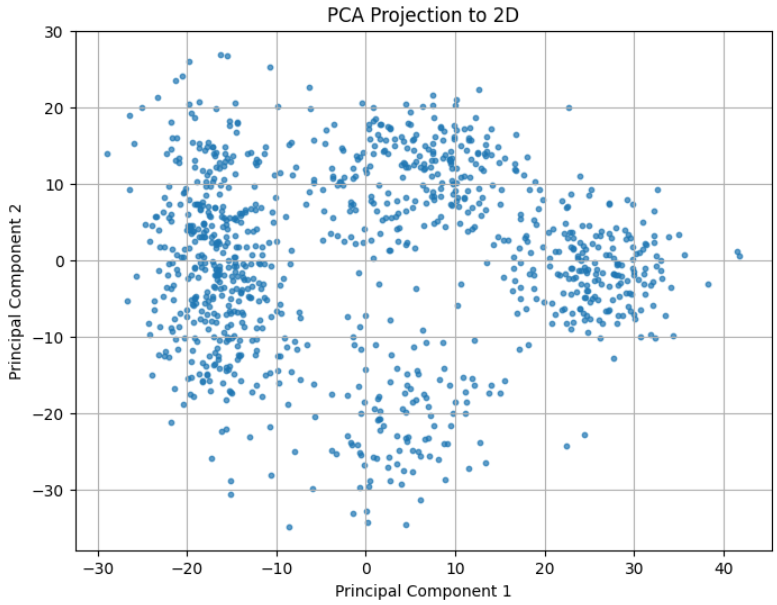

Figure 1: PCA Plot

4 Plot for UMAP

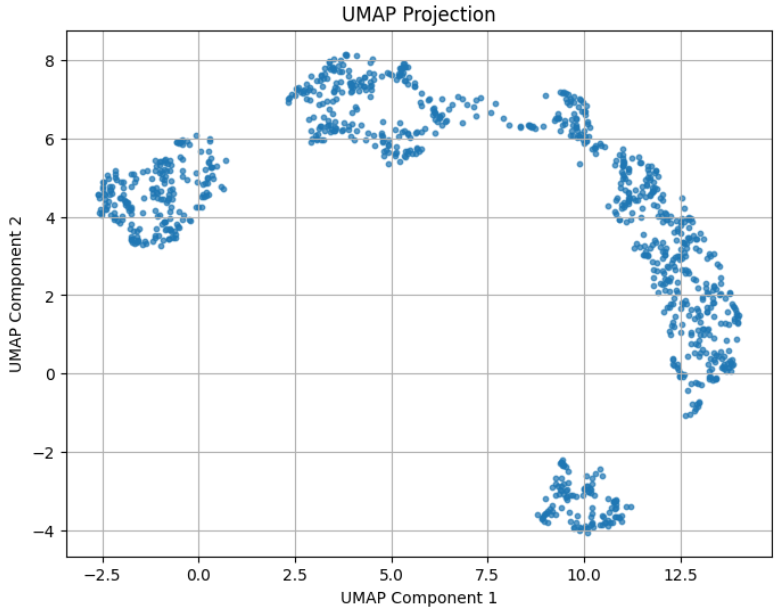

Figure 2: UMAP Plot

5 Plot for tSNE

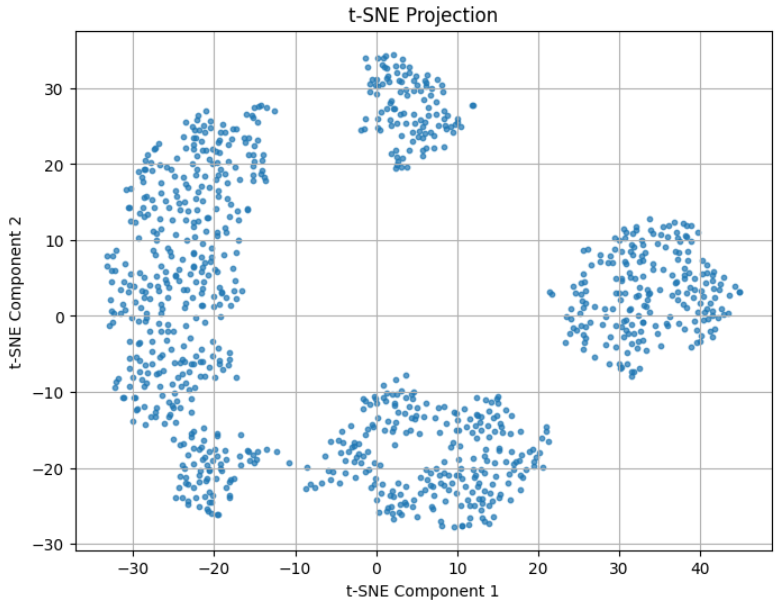

Figure 3: tSNE Plot

6 Results from EM Algorithm

- **Mean number of children with family planning** (λ_A): *1.7824*
- **Mean number of children without family planning** (λ_B): *4.9107*
- **Proportion of families with family planning** (π_A): *0.3560*
- **Proportion of families without family planning** ($1 - \pi_A$): *0.6440*