Assignment 3, Part II: Representation Learning - Variational Autoencoder

Abhinav Rajesh Shripad(2022CS11596)
Jahnabi Roy(2022CS11094)

November 11, 2024

1 Introduction

This report presents an implementation and analysis of a Variational Autoencoder (VAE) applied to MNIST digit data, incorporating Gaussian Mixture Model (GMM) clustering in the latent space.

2 Reconstructed Image

In this section, we present the reconstructed images generated by the VAE.

Figure 1: Reconstructed images generated by the VAE

3 Generated Image

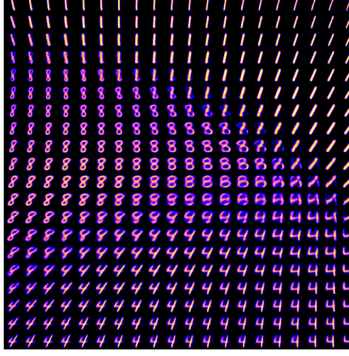In this section, we present the Generated images generated by our model

Figure 2: Generated images generated by the VAE

4   Latent Space Visualization and Explanationn

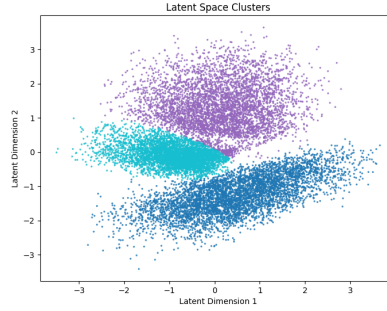In this section, we present the Latent Space Visualization.

Figure 3: Latent Space

The scatter plot shown in Figure 3 illustrates the clustering in the latent space of a Variational Autoencoder (VAE) for the MNIST digits 1, 4, and 8. The latent space is represented by two dimensions, labeled as *Latent Dimension 1* and *Latent Dimension 2*.

The following observations can be made from the clustering patterns:

- **Distinct Clusters:** Three distinct clusters are visible, each represented by a different color (blue, cyan, and purple). Each cluster corresponds to a different digit from the dataset (1, 4, or 8).

- **Cluster Separation:** The clusters are well-separated in the latent space, which indicates that the Variational Autoencoder (VAE) has effectively learned to encode distinctive features of each digit class. The separation of clusters implies that the latent representations are informative for distinguishing these specific digits.

- **Cluster Shape and Spread:**

    - The cluster at the bottom (blue) has a broader spread along *Latent Dimension 1*, suggesting that there is more variation in this dimension for the corresponding digit. This could be due to variability in how this digit (likely digit 8) appears in the dataset.

    - The central cluster (cyan) is more compact and centered, indicating less variance in the latent space dimensions. This cluster likely represents a digit with more consistent features, possibly the digit 4.

    - The top cluster (purple) has a slightly elongated shape along the vertical axis (*Latent Dimension 2*), indicating variability in that dimension. This cluster might correspond to the digit 1, which has a simpler and more linear structure, leading to more spread in one dimension.

- **Gaussian Structure:** Each cluster resembles a Gaussian distribution, which aligns with the assumptions of the GMM. The clusters' shapes and spreads suggest that the latent space has a normal distribution-like structure for each digit.