COL333 – Principles of Artificial Intelligence

Assignment 3 – Part 2 – Report

Aneeket Yadav (2022CS11116) & Sanyam Garg(2022CS11078)

Introduction

This report outlines a summary of our approaches aimed at designing a VAE to reconstruct images of digits - {1,4,8} from MNIST dataset and using it for the downstream task of prediction using a Gaussian Mixture Model. Due to the absence of a sufficiently large test set, we conducted our experiments by splitting the original test set into a train and test set by splitting it in an 80-20 ratio.

Due to the simplicity of the dataset, the encoder and decoder of our VAE, each have a single hidden layer of 324 neurons. We tried sizes 512,324,256,144 for this layer and 256 worked best(in terms of maximising both SSIM and accuracy).

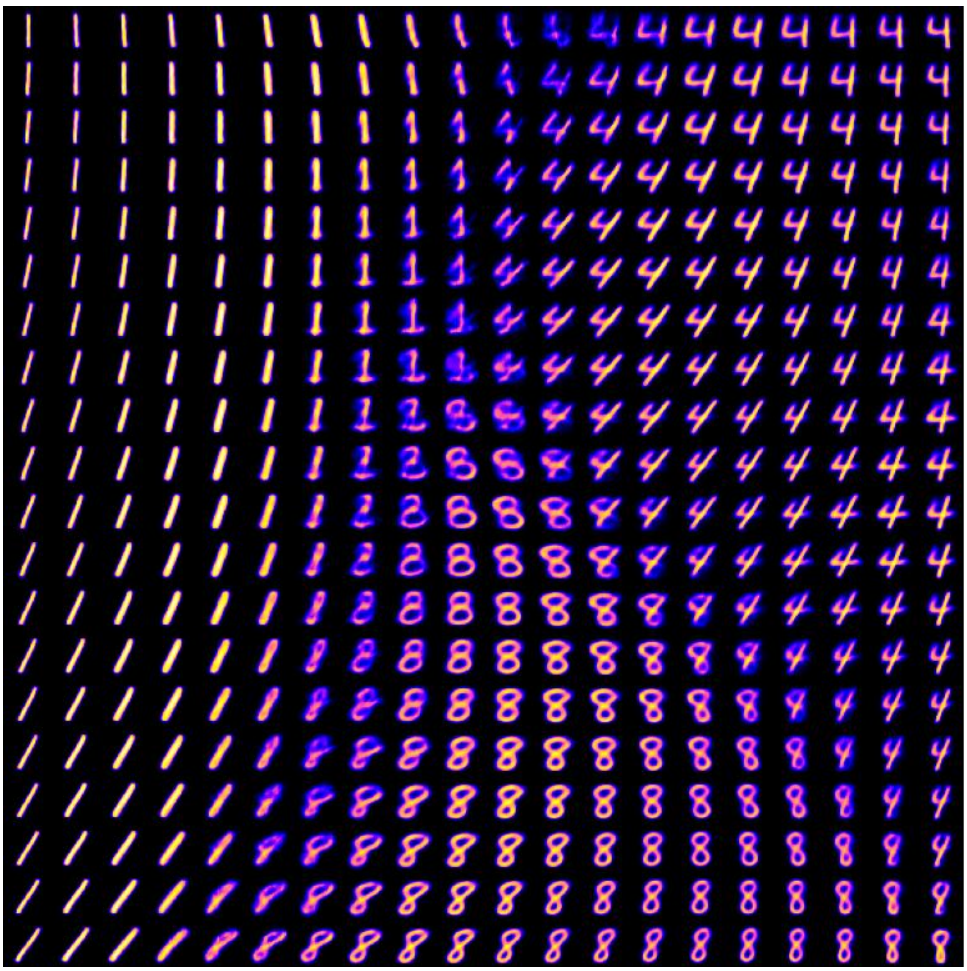
Key considerations-

- There was a tradeoff between SSIM and test accuracy. Models with two or more hidden layers were great at maximising SSIM(between 0.72 and 0.76) but yielded poor train accuracy which fluctuated greatly between 45% and 85%. As number of epochs increased, SSIM increased but test as well as train accuracy decreased. This was also clear by observing the highly intersecting clusters in the GMM graph.
- AdamW yielded slighly better results than Adam
- Results obtained by training on the 80% train set do not generalise well when trained on the full dataset. In the prior case, hidden layer size of 256 yielded better results than 324. However, the opposite was true when trained on the full train set.
- 100 iterations were appropriate for the GMM. Increasing as well as decreasing the iterations yielded worse performance.
- The model achieved 97.28 % accuracy on the train set.
- On the provided validation set, current model achieves-
    - SSIM = 0.6984
    - 1 – MSE = 0.9736
- We have used seed value of 0 for our experiments

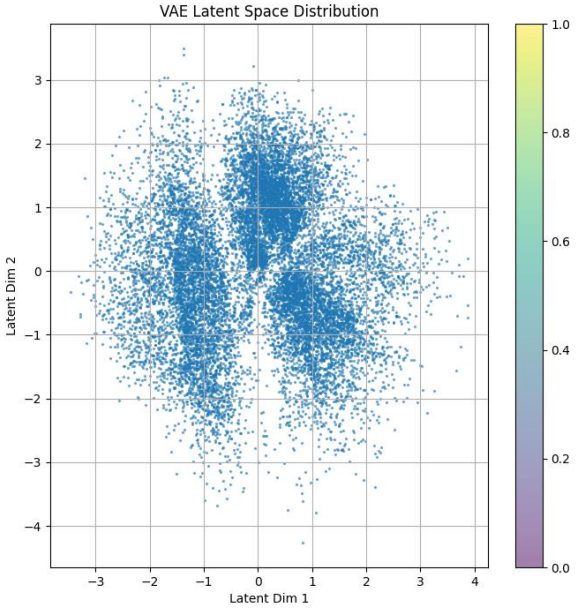
Visualisations

1. Reconstructed images -

2. 2-D manifold



3. Points in latent space



4. Clusters formed by GMM-