Practice Programming Assignment 2

Prathosh A. P.

August 17, 2025

1. Implement a vanilla VAE with MSE for conditional likelihood. Report results with different numbers of samples of z during training for the input to the decoder. Plot 10 by 10 grids of both reconstructions and generations. Plot the loss curves for likelihood, KL, and the combined terms. Compute FID with 1000 generated images.
2. Build a CNN-based classifier using the images in the training part of the dataset and report the accuracy on the test set
3. Now, using the VAE trained, perform a posterior inference on all the images and use the thus obtained latent vectors with an MLP for classification. Quantify and document your observations on the advantages/disadvantages of using latent space, in terms of classification accuracy and network size.
4. Implement a VQ-VAE with discrete latent space. Perform posterior inference on all the images, post-training. Build a classifier with these latent vectors and calculate the accuracy.
5. Fit a GMM on the vectors from the latent space obtained via posterior inference. Subsequently, sample new latents from the GMM, pass it through the decoder, and plot a grid of 10 by 10 images.

General Instructions:

1. For the first question, both the animal and butterfly datasets are to be used. Second question onwards, use only the Butterfly dataset.
2. You need to resize all images to 128x128 pixels before implementing.
3. Use Google collab with Jupiter notebook for all the computing.
4. You are supposed to submit a single Jupiter notebook with all the solutions made into separate blocks.
5. Use Pytorch for building neural networks. You are supposed to directly use the off-the-shelf functions for the models asked.