Assignment 1 ADRL - 2022

Instructor: Prathosh A.P., IISc

September 4, 2022

Variational Auto-Encoders and Family

1. Implement a standard VAE with MSE loss for the likelihood on the following datasets (a) CelebA (https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html) (b) https://github.com/deepmind/dsprites-dataset. Calculate the marginal likelihood scores for both the case using the procedure outlines in Appendix D in this paper (https://arxiv.org/pdf/1312.6114.pdf). Plot 100 random samples of the generated images in a 10x10 grid for both the datasets post training.

2. For the same datasets in question 1, implement a β-VAE. Plot the variations in the generated images with latent traversals as in Figure 4 and 7 in this paper (https://openreview.net/pdf?id=Sy2fzU9gl) for at least 3 values of β (including $\beta = 1$). Choose your βs such that reconstruction vs disentanglement trade-off is apparent. Compute and report the the disentangle metric as in Figure 6 in - https://openreview.net/pdf?id=Sy2fzU9gl for all three values of β**.**

3. Implement a VQ-VAE on the tiny imagenet dataset (https://www.kaggle.com/datasets/akash2sharma/tiny-imagenet), plot the reconstructions of randomly sampled 10x10 images in a grid. Repeat the experiment with three different choices for the size of the dictionary with same latent dimensions. Post training, (a) Fit a GMM on the latent space of the VQ-VAE using the EM algorithm, (b) Fit a simple VAE on the latent space of the VQ-VAE. After both step (a) and (b), sample latent points from the trained GMM/VAE, pass it through the Decoder of the VQ-VAE and plot the generated images in a 10x10 grid in both the cases.

Adversarial Learning and Family

1. Implement a naive Deep-Convolutional Generative Adversarial Network (DC-GAN) on the Bitmoji dataset (https://www.kaggle.com/datasets/romaingraux/bitmojis). Plot a 10x10 generated image grid post training. Plot the training loss curves for both Generator and Discriminator functions and comment on their behavior.

Sample 1000 images from the generator and estimate the Fréchet Inception Distance (https://github.com/mseitzer/pytorch-fid) between the generated 1000 samples and 1000 samples from the real data.

2. Repeat the above experiment with a Bi-GAN and LS-GAN. In addition, plot two instances of 10x10 grids of image generations with traversals in the latent space for all three cases of DCGAN, BiGAN and LSGAN. Document your observations.

3. Implement a conditional Wasserstein GAN and conditional DC GAN on the SVHN dataset (https://www.kaggle.com/datasets/hugovallejo/street-view-house-numbers-svhn-dataset-numpy) with conditioning on the digit class. Compute FID in both the cases, plot the generated images in grids, and plot the training curves for generator and discriminator for both the cases and observe the presence/absence of vanishing discriminator gradients for the case of DCGAN/WGAN.

4. Implement a Cycle-GAN for the following pairs of datasets CELEBA - Bitmoji and SVHN - MNIST. Plot the image translation results for different pairs for both the datasets. Use a Wasserstein GAN as your base model.