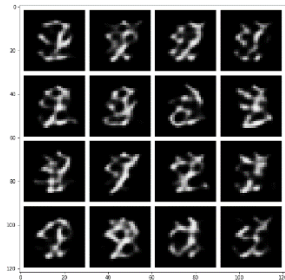# PROBLEM 1 SOLUTION

MEHAK PIPLANI

1. Auto-encoder samples and AE sampling inline question answer.

Inline Question: Describe your observations, why do you think they occur?
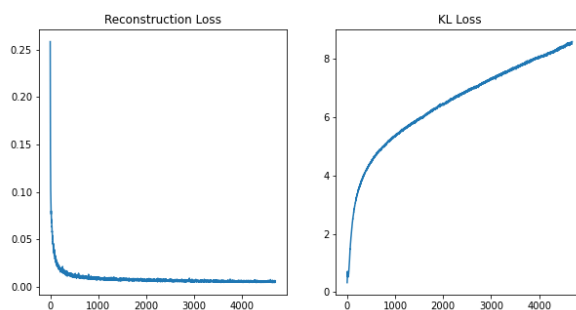
Answer:

My observations are that Auto encoder basically outputs noise when we are trying to visualize samples based on random normal input. This is probably happening because the distribution of the dataset learned by the model is different from the normal distribution, we are using to sample the images.
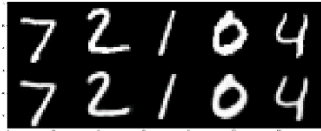
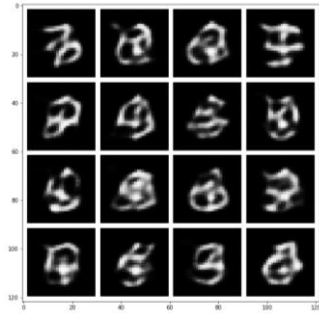2. VAE training curves, reconstructions and samples for:
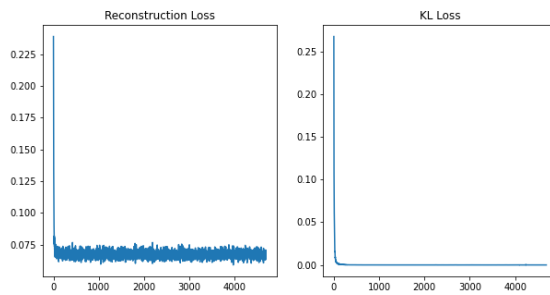
- β=0

   TRAINING CURVES:

   RECONSTRUCTION SAMPLES:

VISUALIZING SAMPLES:

- β=10

TRAINING CURVES:

RECONSTRUCTION SAMPLES:

VISUALIZING SAMPLES:

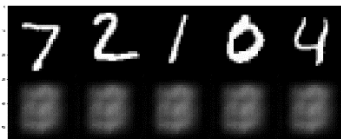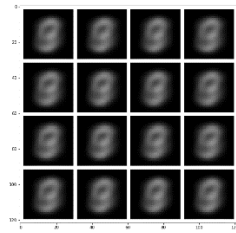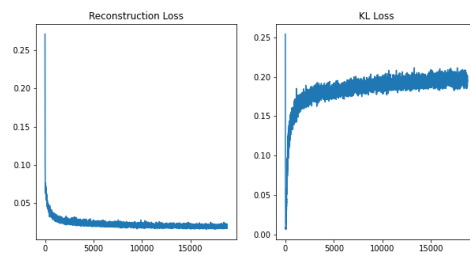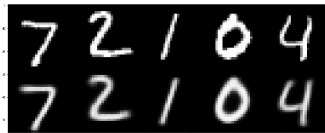- your tuned β (also listing the tuned value for β)

  My tuned β = 0.075
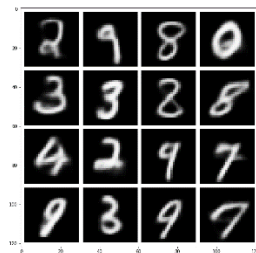
  TRAINING CURVES:

  RECONSTRUCTION SAMPLES:

  VISUALIZING SAMPLES:

3. Answers to all inline questions in VAE section

Inline Question: What can you observe when setting $\beta=0$? Explain your observations!

Answer: At $\beta=0$, there is no contribution of the KL Divergence loss in the model and hence, the VAE is almost like AE. The reconstruction samples are good as the model is learning the weights according to the reconstruction loss only but when we visualize samples from gaussian distribution, we get noise due to the difference in the distribution learned by the model and the distribution used for sampling as model is not learning the training sample's distribution using the KL divergence loss.

**Inline Question: What can you observe when setting $\beta=10$? Explain your observations!**

Answer: At $\beta=10$, high amount of weightage is given to the loss minimizing the gap between the latent distribution and the distribution learned by the model. Thereby, the reconstruction loss is not decreasing and hence the model is not learning the features to reconstruct the samples from the dataset and hence, outputs a sort of gaussian blur on the input image.

Inline Question: Characterize what properties you would expect for reconstructions (1pt) and samples (2pt) of a well-tuned VAE!

Answer: The characteristics that I expect for reconstructions is that the output images are close to the image given to the model (reconstruction loss should be low). The expectations for samples would be to be able to construct meaningful and interpretable images from the random sample given to the model such that they are like the images in the dataset, the model was trained on. The "good result" would have the following properties for encodings, all of which are as close as possible to each other while still being distinct, allowing smooth interpolation, and enabling the construction of new samples.

4. Three representative interpolation comparisons that show AE and VAE embedding interpolation between the same images

Variational Auto-Encoder
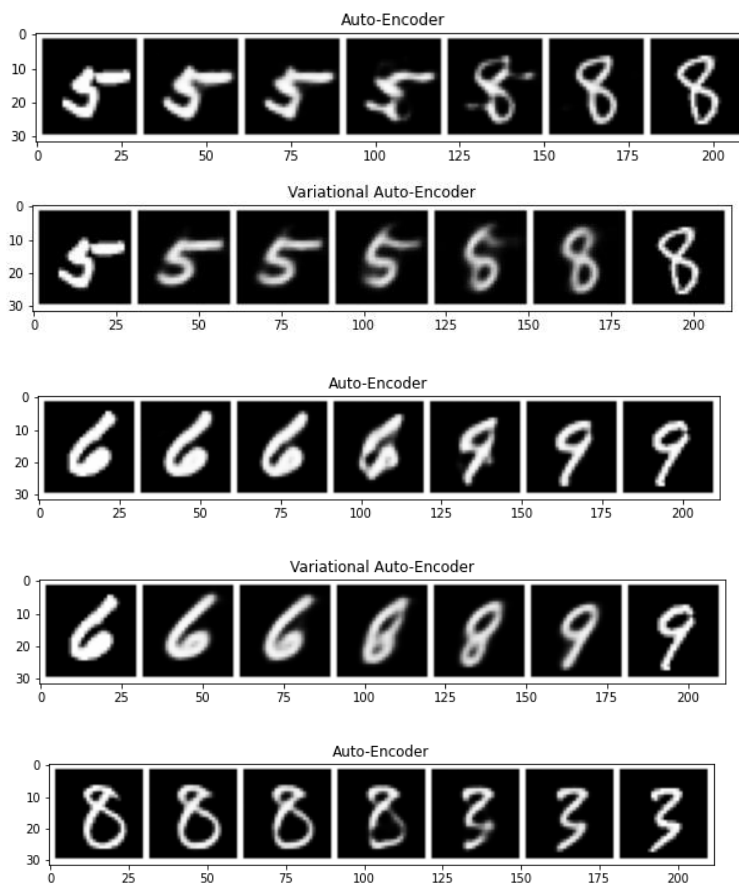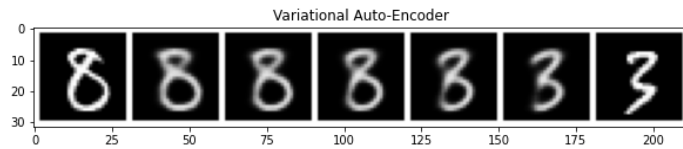
Inline Question: Repeat the interpolation experiment with different start / end labels and multiple samples. Describe your observations! Focus on:

1. How do AE and VAE embedding space interpolations differ?
2. How do you expect these differences to affect the usefulness of the learned representation for downstream learning?
   (please limit your answer to <300 words)

Answer:

1) The embedding space interpolations of a few images of AE during the transition from start label to the end label may not resemble images from training data, while images of VAE during the transition resembles the images from training data. This is happening probably due to constraint on latent code distribution enforced in VAE.
   When doing the interpolation of two given images, the interpolation results of AE may be noisy as not all in between results fall in any distribution. On the other hand, in VAE with the help of KL divergence loss we are forcing the latent variables in to be unit Gaussian. Therefore, the VAE model can generate images that look likes training data.

2) The latent space of the AE model is not observed to be continuous and doesn't allow smooth interpolations between starting and ending label making it difficult to utilize it for learned representation.
   On the contrary, we can utilize the learned representation of the VAE model to do downstream learning, such as generate new images, as the latent space's distribution is known (unit gaussian), and we can generate meaning samples from it.