CSCI 566 Assignment 2

Problem 1

Autoencoders

Describe your observations, why do you think they occur?

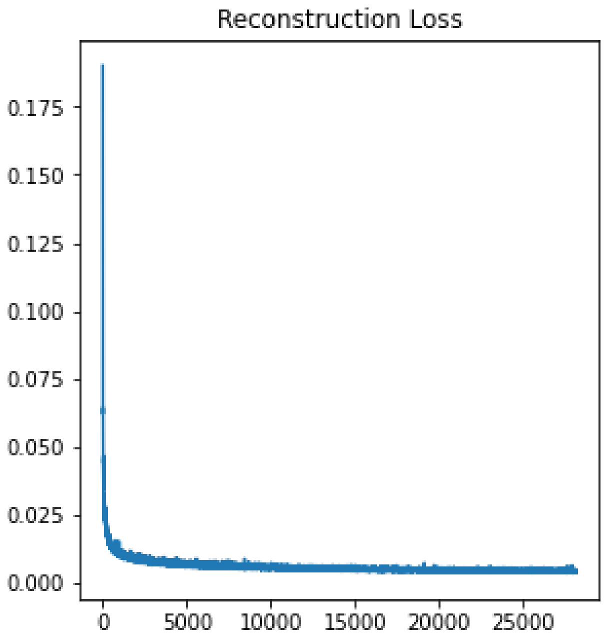
The autoencoders consists of an encoder to convert the input (high dimensional) to encodings or latent representation (low dimensional) and reconstruct the images using the latent representations. The problem with autoencoders is that the latent space that the inputs are converted may not allow easy interpolation. If any sample embedding which was not seen by the model was chosen it would not perform well because the encoded space was not compact following any distribution. The latent space is not regular, it depends on the distribution of data, model architecture and the dimension of latent space. The model does not produce organized latent space during the training process, in order to choose a random point from the latent space and decode it to get new image. The autoencoder can perform well only if the sample embedding is within the encoded space (the inputs that were previously seen). It can perform well for reconstruction but not in generating images from new sample embeddings. Thus, the results obtained are random and not well-defined digits.

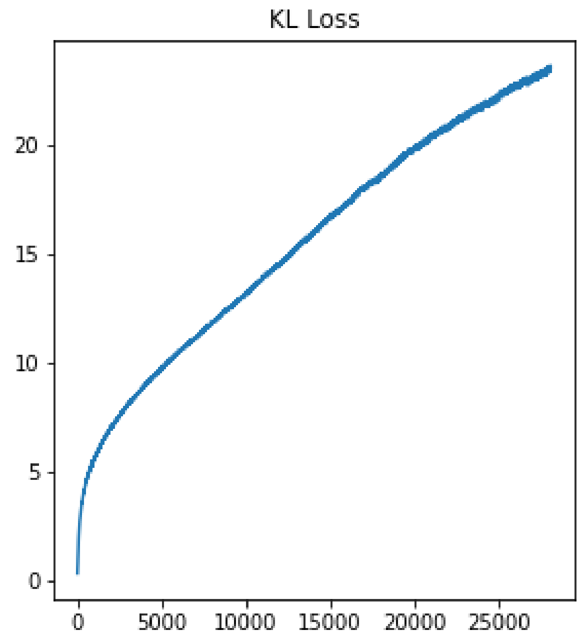
Inline Question: What can you observe when setting $\beta=0$? Explain your observations!

Due to the irregularity of the latent space in autoencoders, some points generate a different decoded output and the other points in the latent space generate a different decoded output. Variational Autoencoders adds a regularization term between the output of the encoder and the prior distribution. It pushes the output of the encoder to be as close as possible to the prior distribution. By setting Beta=0, VAE almost behaves like autoencoders without the regularization term beta on the latent space distribution. Thus, it does not have a good generative process, any sample point after decoding might not produce clear outputs. We can observe that the samples from generative model are not good and the KL loss in the training curves plot, keeps increasing. The reconstruction image is the same as the autoencoder, because by setting the beta term to zero, we are still retaining the reconstruction loss term, which is the same objective function as the autoencoders.

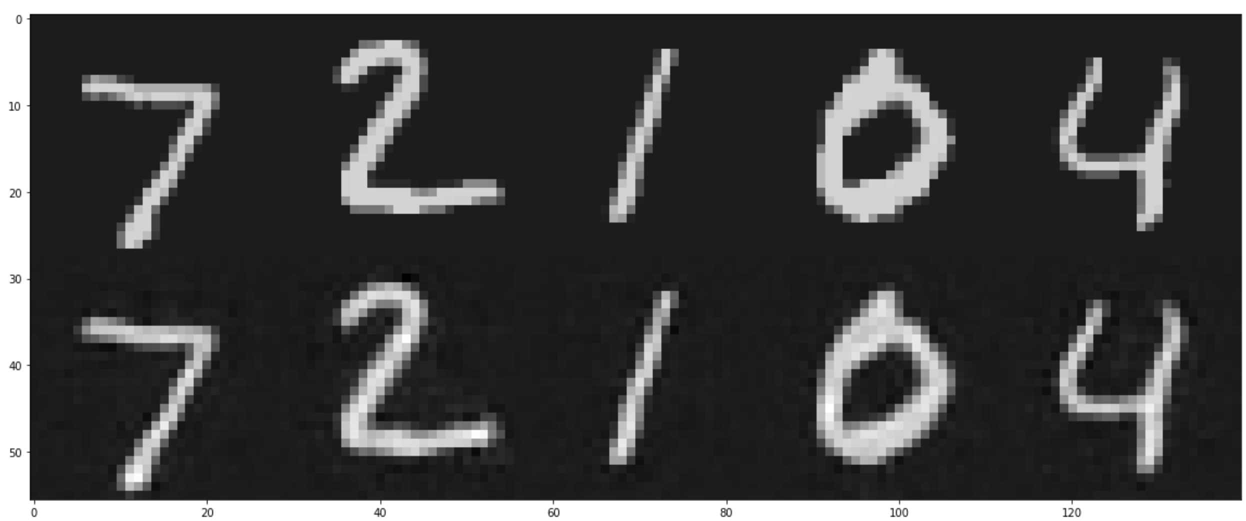
Beta=0

VAE Training Curves

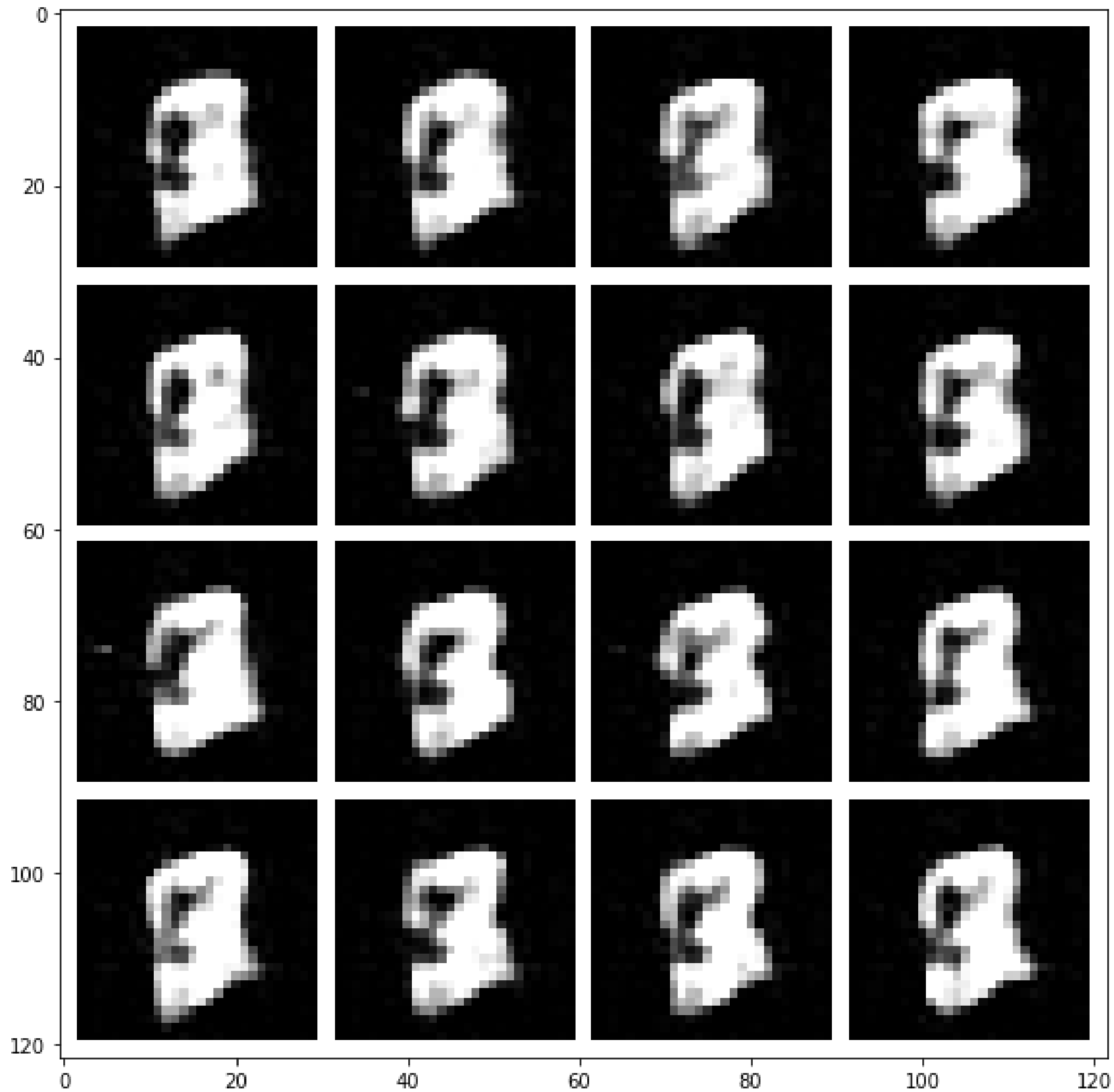




Reconstruction



Samples

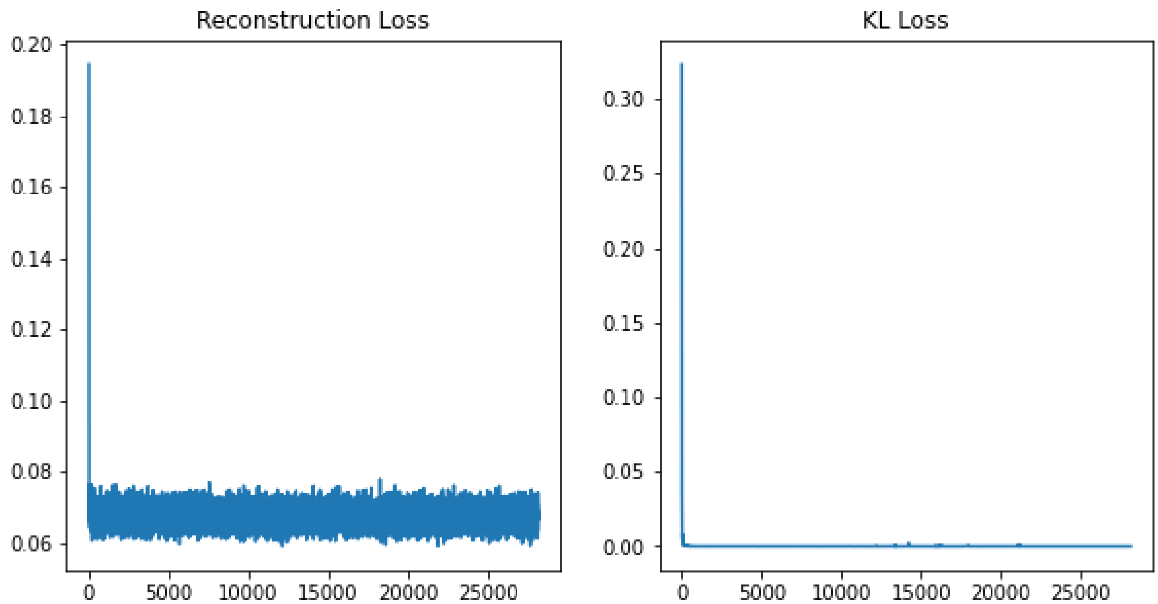


Inline Question: What can you observe when setting β=10? Explain your observations!

The Beta value=10, is very high. We can observe that the reconstructions and the samples are blurred. Higher beta value than 0 applies a stronger constraint to the latent distribution, thus it can generate efficient latent encoding and the representations are disentangled. But if the Beta value is too high like 10, there will be a trade off between reconstruction quality and the disentanglement. We can clearly see that by increasing the beta value to 10, applies a stronger constraint on the latent distribution but it compromises the reconstruction quality. We can see that the KL Loss is so low. By constraining the latent distribution with high beta value and maintaining a very low KL Loss, it compensates the reconstruction

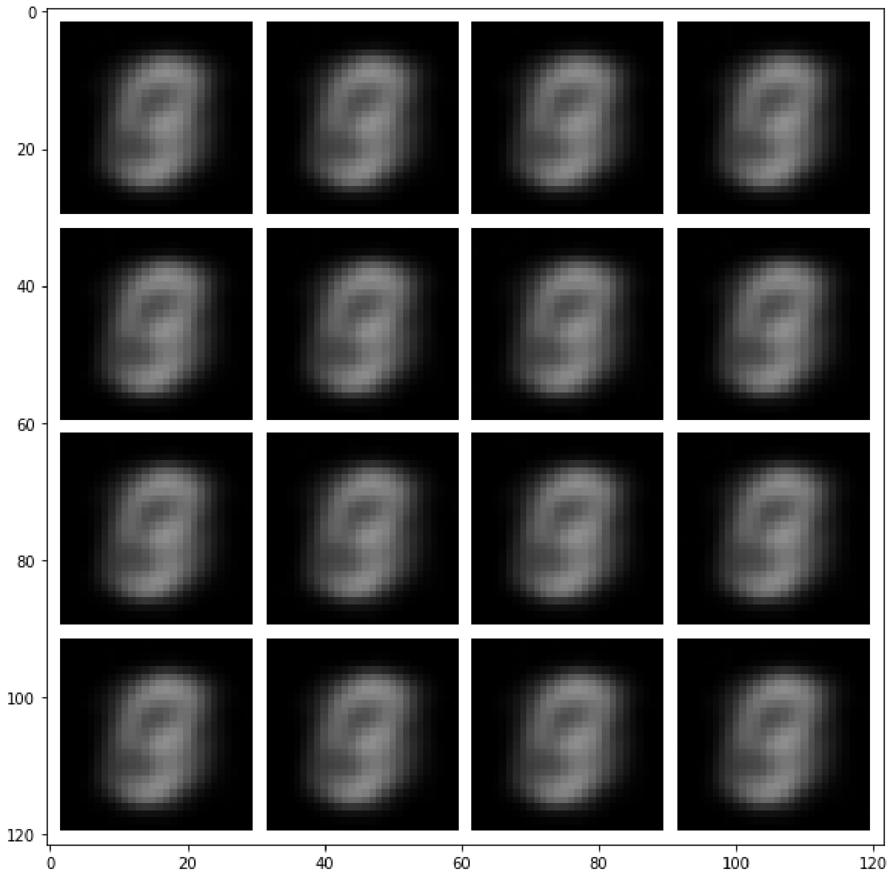
Beta=10

Training Curves

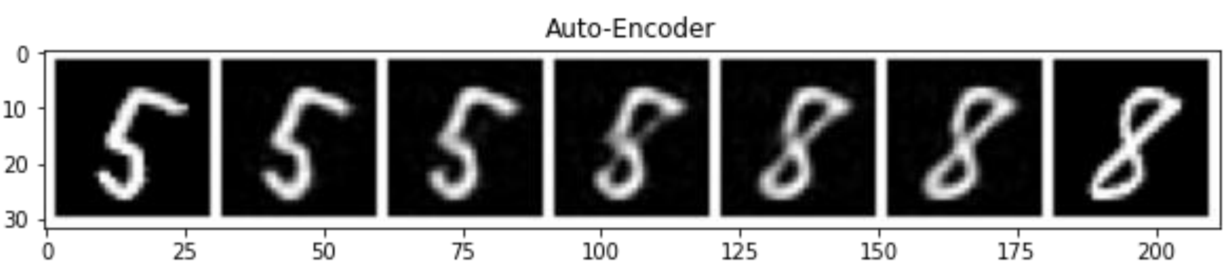


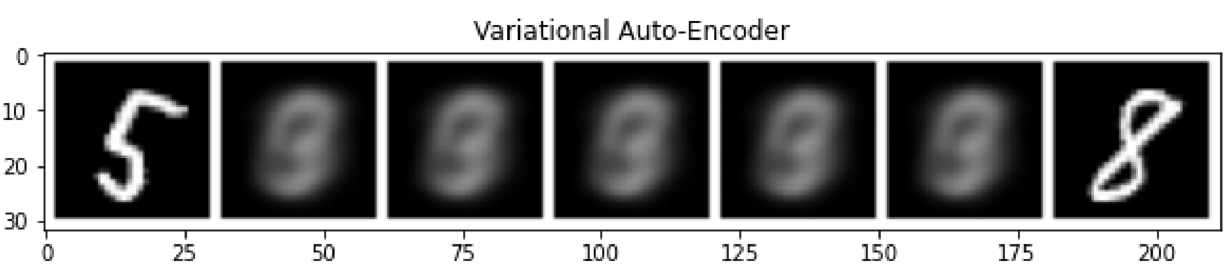
Reconstruction

Samples



Embedding Space Interpolation



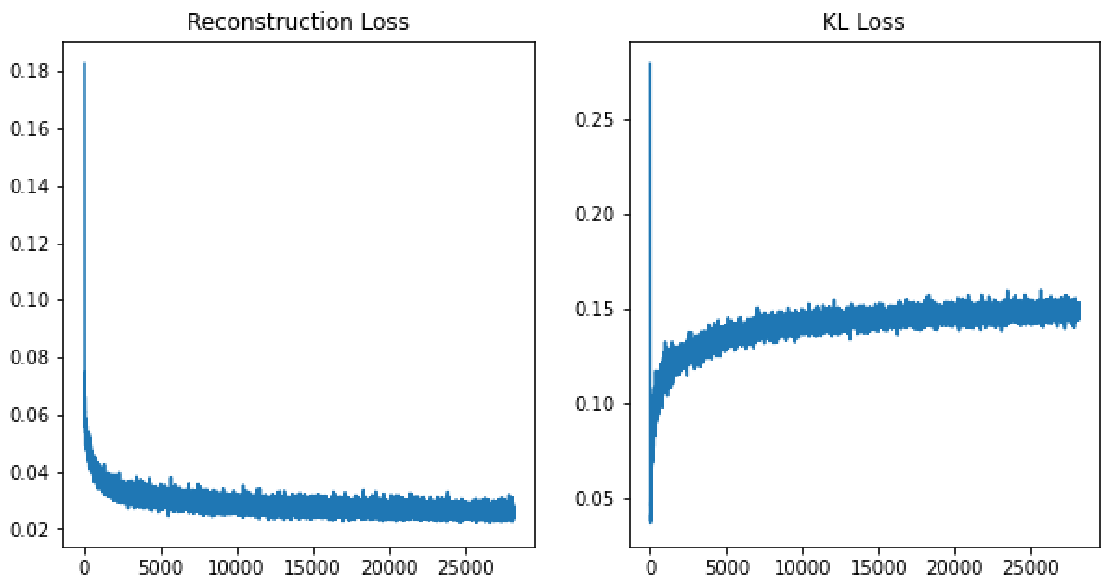


Inline Question: Characterize what properties you would expect for reconstructions (1pt) and samples (2pt) of a well-tuned VAE!

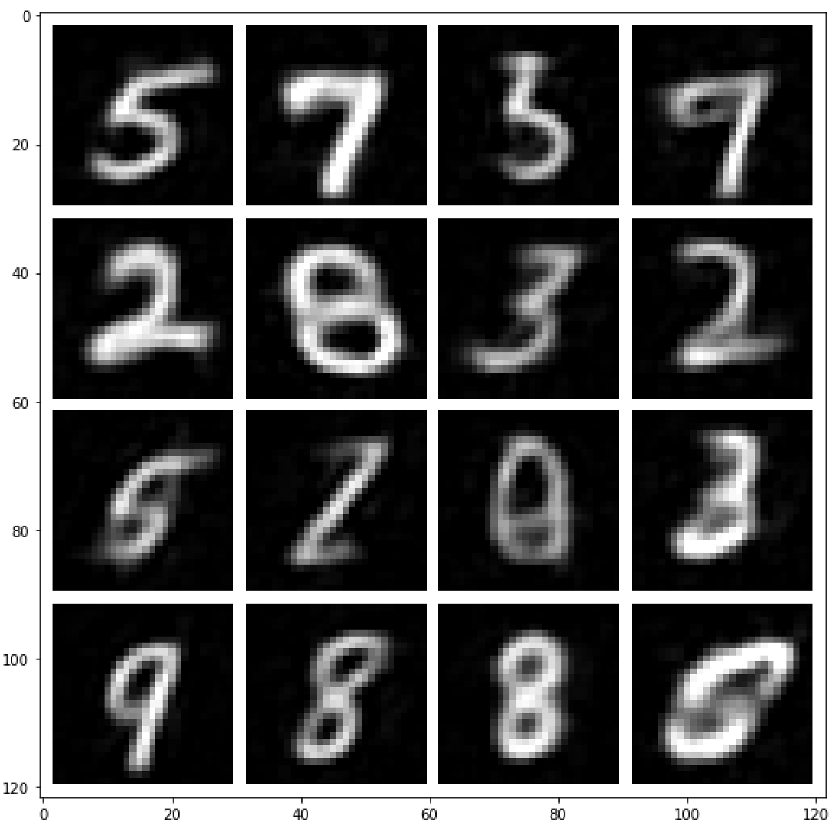
For a well-tuned VAE will have better reconstructions, which will look exactly to the original images. In case of the samples, well-tuned VAE would tackle the problem of the latent space irregularity by making the encoder to return a latent space that follows a prior distribution, and thus it can generate images that looks like the original images. The generated samples are not blurry because the beta value is properly tuned in order to maintain the reconstruction quality and the constraint the latent distribution. Thus, the generated samples are better with a well-tuned VAE, and with lesser or greater beta values both the reconstructed and the generated samples are blurry. We can see that the model performs better for the beta value= 0.1. The KL loss for beta value= 0.1 after few iterations saturates instead of increases unlike with Beta=0, where the KL loss is increasing.

Best Beta value =0.1

Training Curve

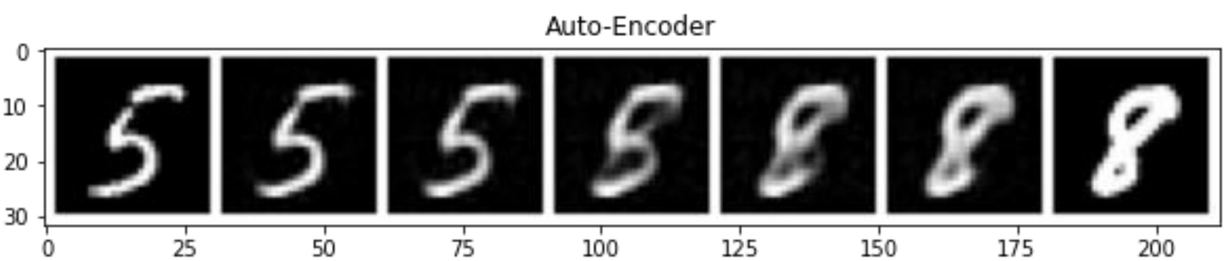


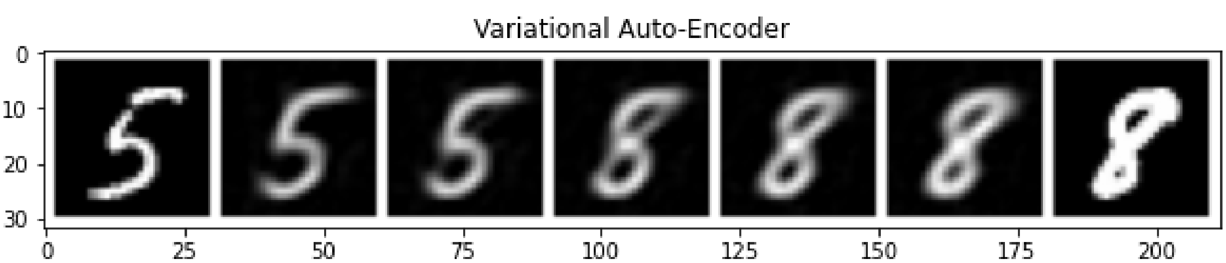
Reconstruction and Generated Samples



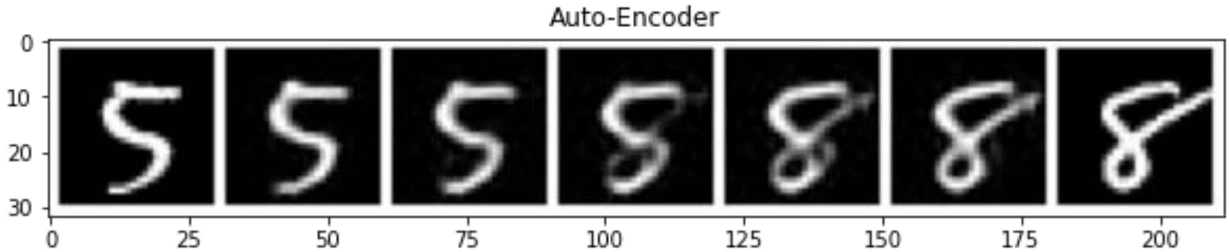
Embedding Space Interpolation

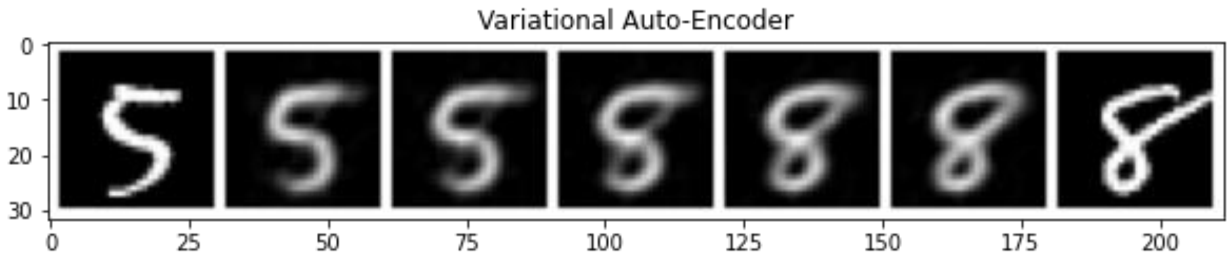
Representation 1



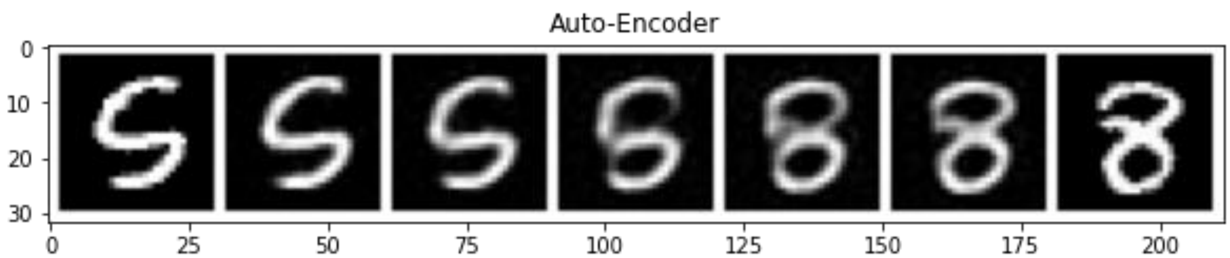


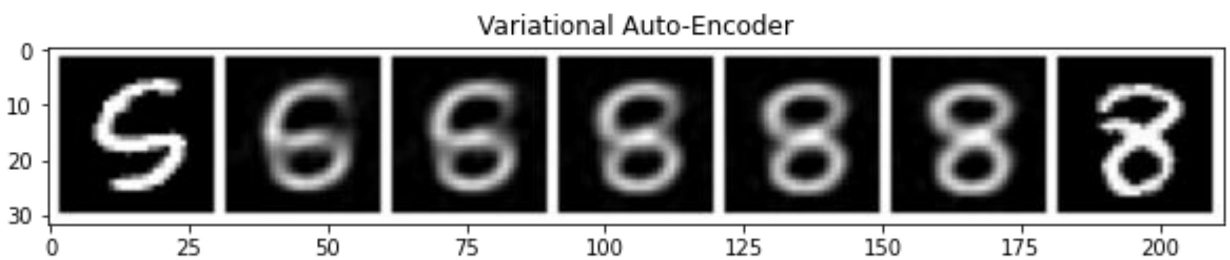
Representation 2





Representation 3





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

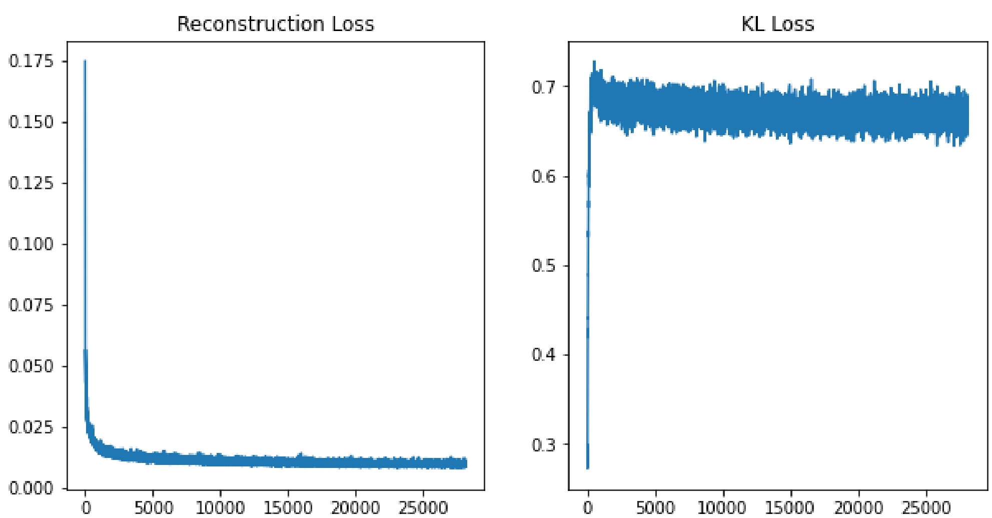
1. **How do AE and VAE embedding space interpolations differ?**
2. **How do you expect these differences to affect the usefulness of the learned representation for downstream learning?**

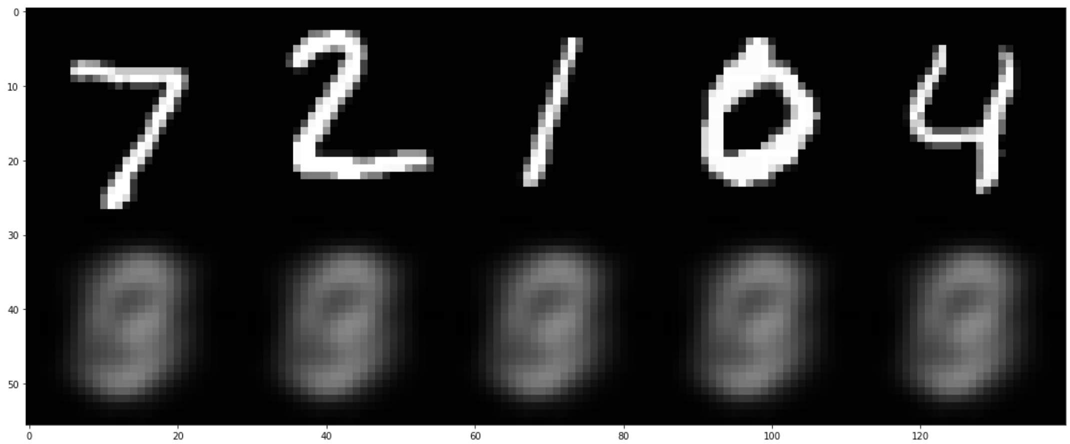
Answer: 1) We can clearly notice that the AE and VAE embedding space interpolations are different. The AE embedding transition from the start image to the end image, but the edges of the digits are not that prominent. The edges are not continuous (broken and there are many broken edges around the digit). But, the VAE embeddings are better with continuous edges and the transition is very smooth from one image to the other.

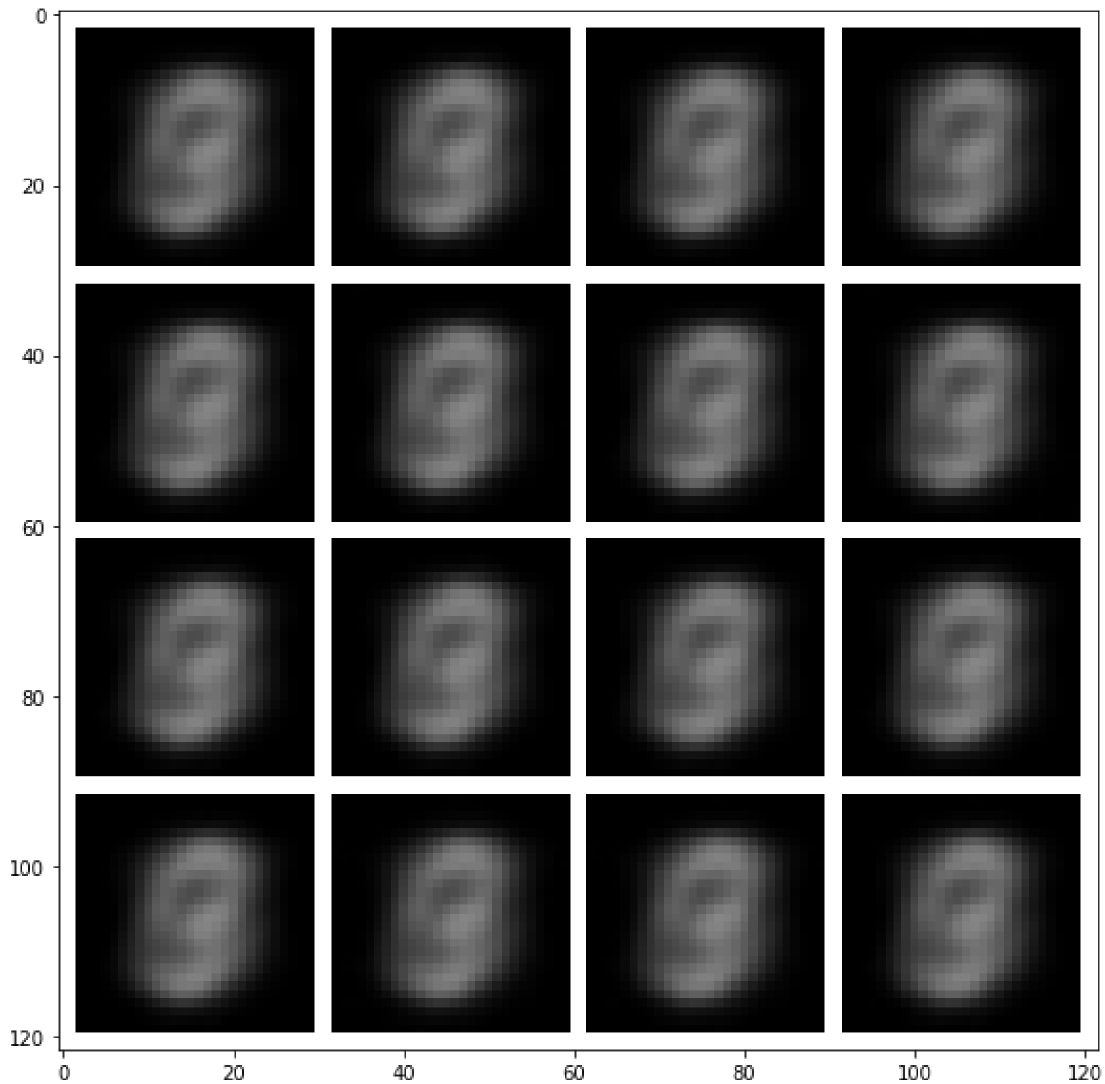
Answer 2) We can see that VAE is performing better than AE. These difference can affect the usefulness of the learned representation for downstream learning, because the representations that are obtained are stable and thus when these representation are used in downstream tasks such as feeding these representation into another model, the output obtained would be more reliable.

Other Results obtained for other beta

Beta=0.01

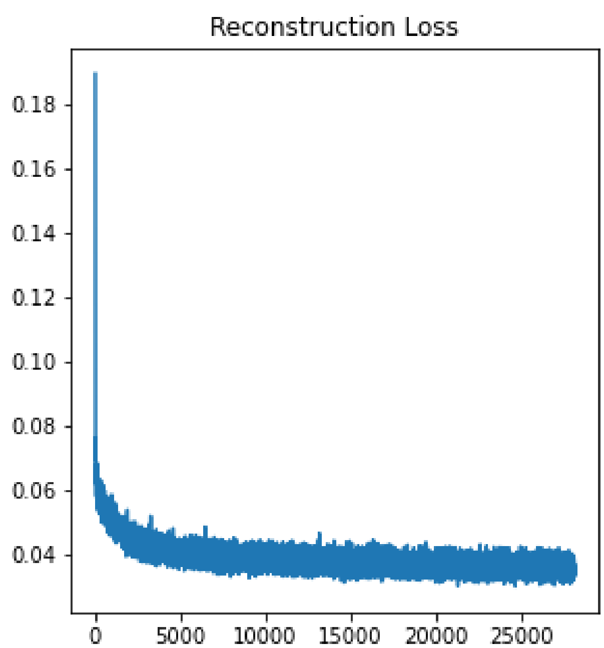



0
10
20
30
40
50
0
20
40
60
80
100
120
7 2 1 0 4

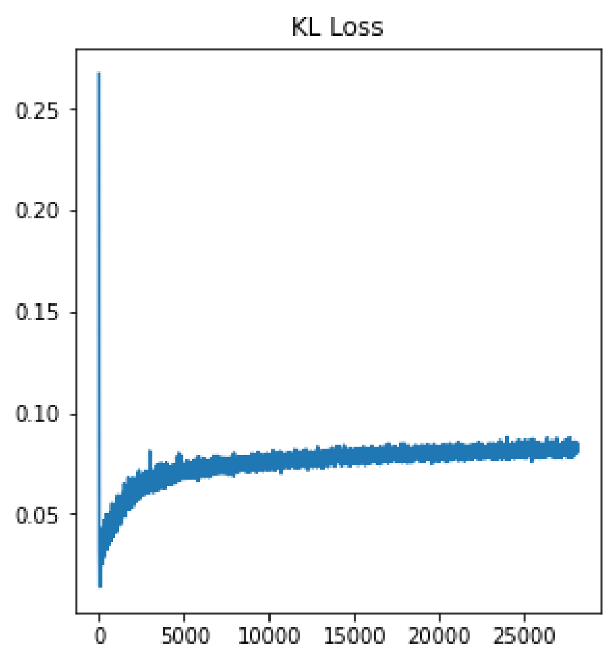

0
20
40
60
80
100
120
0
20
40
60
80
100
120

Beta=0.2

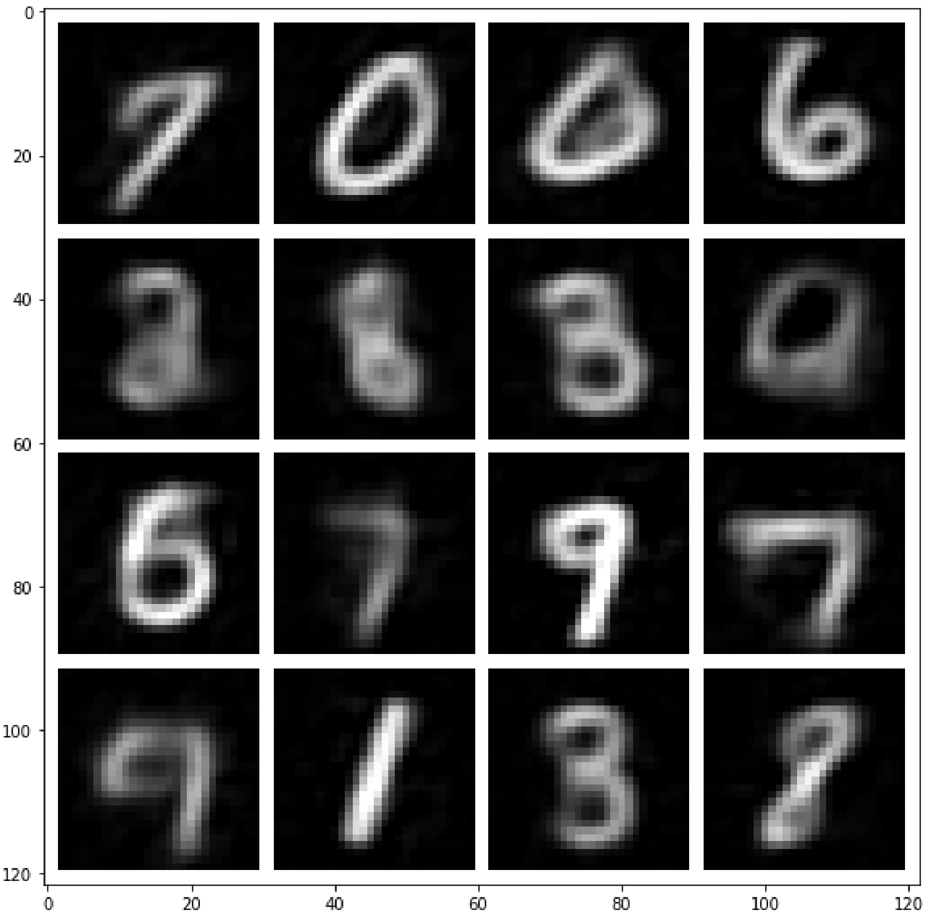
0
10
20
30
40
50
0
20
40
60
80
100
120



Reconstruction Loss
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0
5000
10000
15000
20000
25000



KL Loss
0.25
0.20
0.15
0.10
0.05
0
5000
10000
15000
20000
25000



We can see for very low beta values like 0.01 and higher beta values we get blurry reconstructions and samples. **The Best Beta value obtained is 0.1 (mentioned in the previous section)**