Homework #1

Michael Wacey

For Professor Vijay Raghavan

SEAS 8525: Computer Vision & Generative AI

September 7, 2024

George Washington University

# Question 1

Pick a Generative AI use case for your company or your organization and explain the pros and cons in less than 200 words.

# Response 1

The company that I am currently consulting with sends out workers to merchandise, count and perform any other required activities in retail stores. Clients always want to know the shelf space and quantity on the shelf of their products and competitor products. Having the workers do this is expensive, it is cheaper to have AI process pictures that the workers take. There are many companies that offer software to do this. They train their system on products so that the system can detect the product. With several hundred thousand retail products, new ones being introduced, and existing ones changing their packaging, this is an expensive effort.

Working with two engineers, I came up with a new way to use the ChatCPT generative AI platform to solve this issue. Recently, ChatGPT added the ability to generate text based on pictures. We were able to process a picture of a shelf and generate text that describes all the products on the shelf, all the shelf tags, prices, locations and more. In our testing, we have found it to be over 95% accurate. It requires no ongoing training. This is a huge win for the company. The only con I can think of is the effort that it will take to integrate it into our workflow and address the incorrect responses.

# Question 2

Run Class-1-LatentSpace-PyTorch.ipynb on colab. Either at the end of the notebook or at the beginning add a text cell and in a paragraph or two explain your understand of the code and what you see in the results.

As we discussed in Office Hours, I have included my paragraph here.

# Response 2

I opened the notebook and read the documentation. Based on prior classes it seemed like a typical use of the MNIST data set. In summary, it creates an autoencoder, feeds the MNIST data set to it and encodes the data set and then decodes the encoded data. The training runs ten epochs on the data. It uses MSE for the loss function and Adam for the optimizer.

Looking at the output, the loss is very high – above 0.90. The loss does not even seem to be converging. So, for some reason this autoencoder is not converging on the data. The original representation does not show enough data, it only shows the -1 values for the background. The latent representation has far more zeros than I would have expected.

During my reviews I saw the problems but did not look deep enough to understand the cause. As we discussed during office hours, there is an inconsistency normalizing the data to -1 to +1 but using RELU for non-linearity since only covers 0 to 1. I need to look closer going forward.