ELEC 475 Lab 1 Report

Shenghao Jin (20147024)

September 25, 2023

Table of Contents

Model Details

The model serves as a 4 layer MLP autoencoder using the MNIST dataset. Each image in the dataset is 28x28 pixels. Therefore, an input consists of 784 (28x28) neurons (784x1 dimension) is fed into the encoder, while an encoding layer has 384 (784/2) nodes. A bottleneck layer is where the inputs are further encoded to 8x1 dimension. Correspondingly, the decoding layer has 384 nodes, and the final output has a dimension of 784x1. A ReLU activation function is applied after the two fully connected layers (input to encoding layer and encoding layer to the bottleneck layer) in the encoder, and first fully connected layer in the decoder (the bottleneck layer to the decoding layer) respectively. A sigmoid activation function is applied after the second fully connected layer (decoding layer to the output) in the decoder. [1]

Training Details

Adam is employed as the optimizer, while few optimization hyperparameters are defined as following: initial learning rate is 0.001; the exponential Learning Rate scheduling method is employed, with a gamma of 0.95. One thing worth to note is that both input and output image shall be reshaped to 1x784 dimension tensor to align with the tensor input definition in the model. Loss function used in the training is MSE. Batch size is 2048, bottleneck size is 8, and number of epochs is 50 as per stated in the given command (arguments passed through).

```
==========================================================================================
autoencoderML4Layer                      [1, 784]                  --
├─Linear: 1-1                            [1, 392]                  307,720
├─Linear: 1-2                            [1, 8]                    3,144
├─Linear: 1-3                            [1, 392]                  3,528
├─Linear: 1-4                            [1, 784]                  308,112
==========================================================================================
Total params: 622,504
Trainable params: 622,504
Non-trainable params: 0
Total mult-adds (Units.MEGABYTES): 0.62
==========================================================================================
Input size (MB): 0.00
Forward/backward pass size (MB): 0.01
Params size (MB): 2.49
Estimated Total Size (MB): 2.51
==========================================================================================
```

Figure 1: Training summary

Results

In general, the autoencoder works as expected in terms of its denoising feature and reconstruction of input images. Going through the training dataset of 60,000 images in 50 epochs, the network achieves a loss rate of approximately 0.018. The reconstruction result is shown below as Figure 2. The sample outputs of the denoising feature can be observed in Figure 3 and Figure 4. The denoising output may not

be ideal at certain attempt. Figure 5 is an example of a sort of style transfer of three pairs of images from MNIST dataset. The interpolation works as expected. [1]

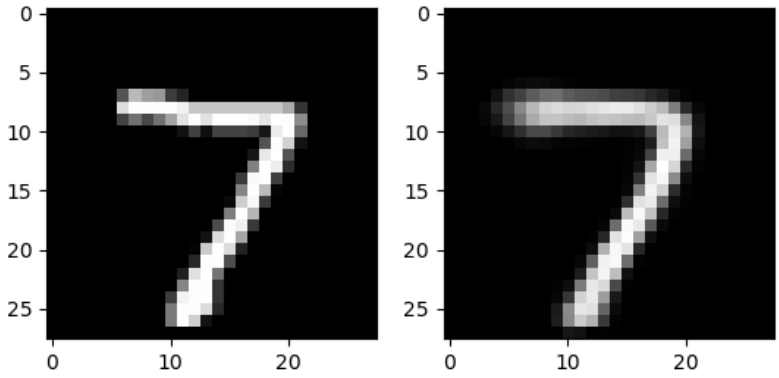

Figure 2: Output of the MLP autoencoder

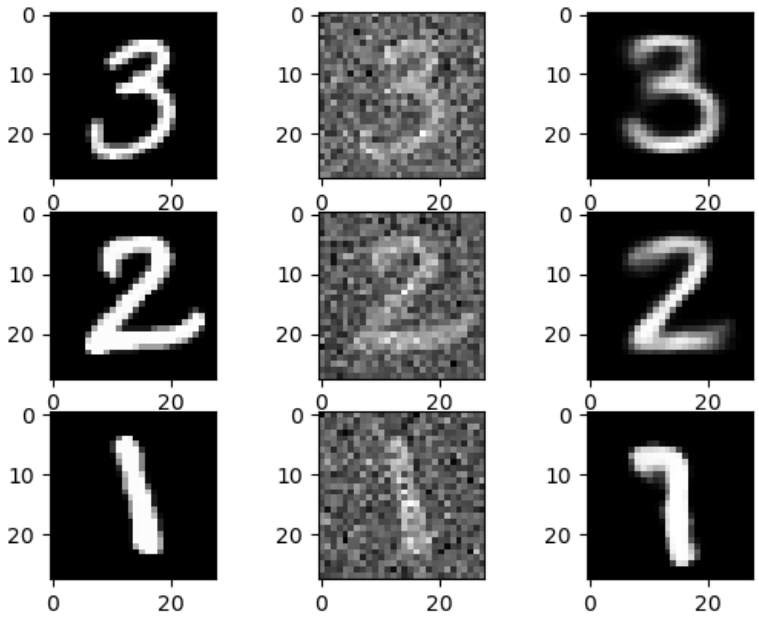

Figure 3: Output of the denoising feature, relative successful one

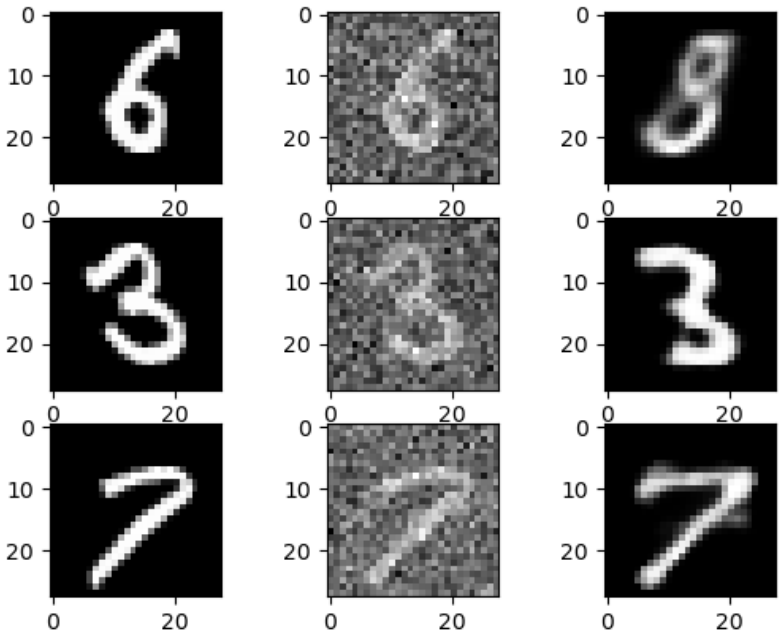

Figure 4: Output of the denoising feature, top row not ideal

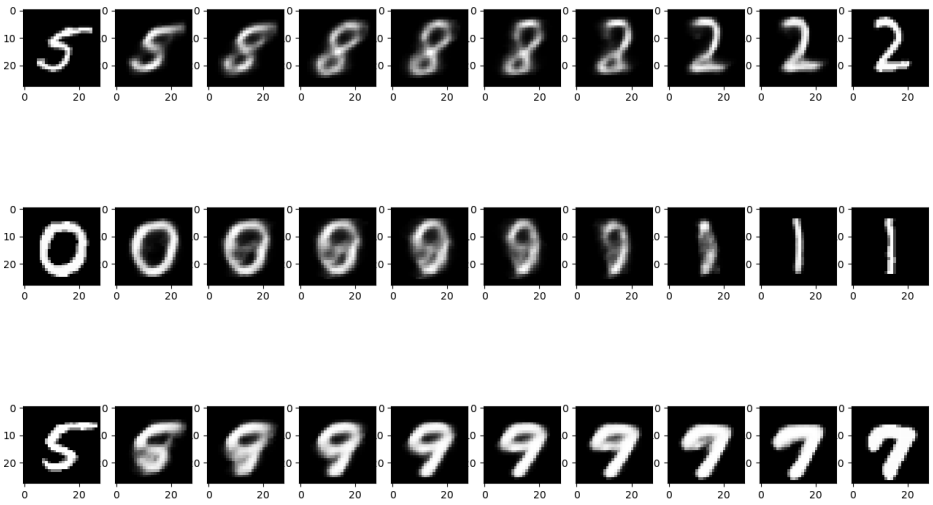

Figure 5: Bottleneck interpolation of three pairs of images from the test set of MNIST

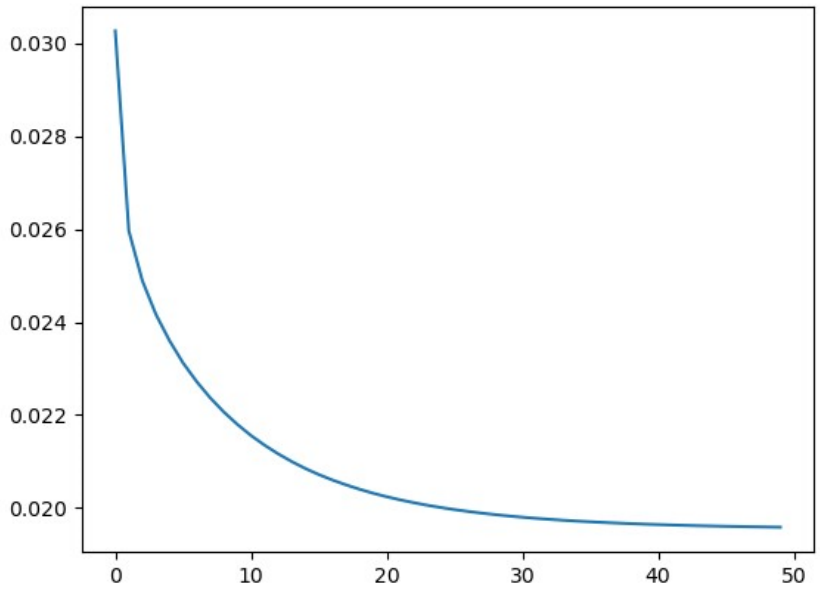

Figure 6: Loss curve of the 4 layer MLP autoencoder training session

The loss curve is relatively smooth thanks to the exponential learning rate scheduling method, which features relatively quick decays at the beginning then slow down. Plus, the training loss curve agrees with the feature of Adam optimizer, fast convergence, and smooth curve. Training loss during epoch 1 is around 0.035, then witnesses a drastic fall to around 0.025 after walking through four epochs. The loss continues to converge but slower also smoother.

References

[1] M. Greenspan, "ELEC 475 Lab 1 - MLP Autoencoder," Kingston, 2023.