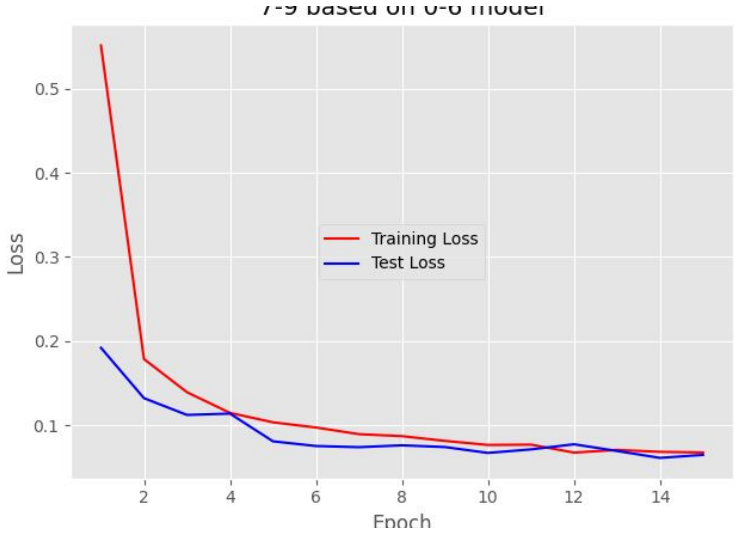
HW2 Report

Deep Learning

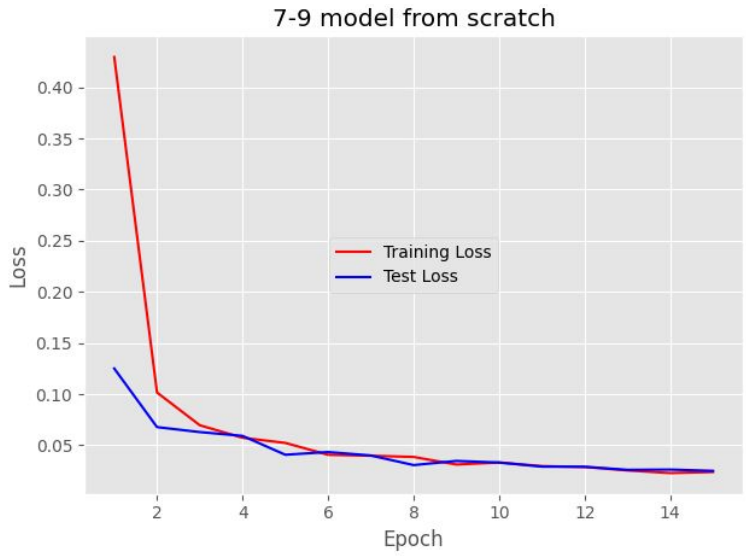
Uri Ariel Yahalom 318261468
Lior Cohen 316409903

1. Question 3 learns a model for digits 0 to 6 from the mnist dataset:
   Model Accuracy = 99.375%
   training and validation loss graph:

2. Question 4 learns a model for digits 7 to 9 from the mnist dataset that is based on a couple of layers from the above presented 0 to 6 model:
   Model Accuracy = 99%
   training and validation loss graph:



3. Question 5 learns a model for digits 7 to 9 from the mnist dataset, just from scratch without any base model.
   Model Accuracy = 99.51%
   training and validation loss graph:



4. When we compare the difference between the model from question 4 and the model from question 5, we can see that model 5 achieves just slightly better results.

5. On the other hand we can see that model 4 learns about 20% percent faster than model 5.

   Hence we conclude that when our network is based on a pretrained model, we can benefit from a faster learning process, due to already trained convolutional layers that makes it easier for the network to start “recognizing” general features in the images from the very beginning.
   However we’re introducing an influence from the features of the original model, it works well if the new model is sharing his low level features with the pretrained model, and also in case we don’t have a lot of data to train our new model compared to the preatrined one.