

Deep Learning Lab

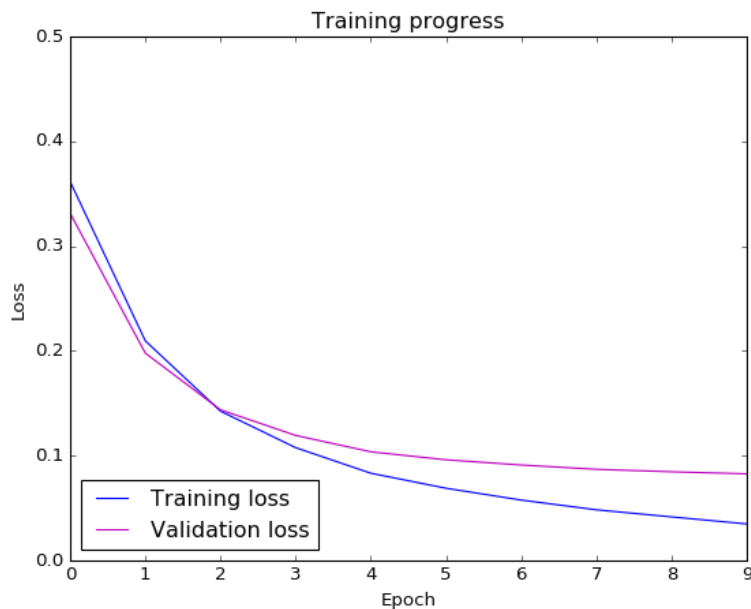
Assignment 1

Name: Aleksandar Petrov

Immatriculation number: 3743542

I have implemented the very basics so that the NN can be trained on the MNIST data set. I had trouble with my gradients but it was all just coding mistakes. I reached a validation error of 2.2% but sadly my NN overfits really strongly. I am using 3 hidden layers with 300, 300, 10 neurons respectively (starting from input to output) the idea I took from some of the results on the MNIST web-site. The first layer is a ReLU and the second one a tanh. My reasoning was that a ReLU would be good to extract the linear features and a tanh is needed to extract some more complex ones. The output is of course a Softmax because our task is in the scope of multiclass classification.

Ideas to get better: Implement Adam and do some regularization on the weights to avoid overfitting. Maybe I can still do that before the submission deadline ;)



On the plot the performance of my NN can be seen. The final test error is 2.57%.