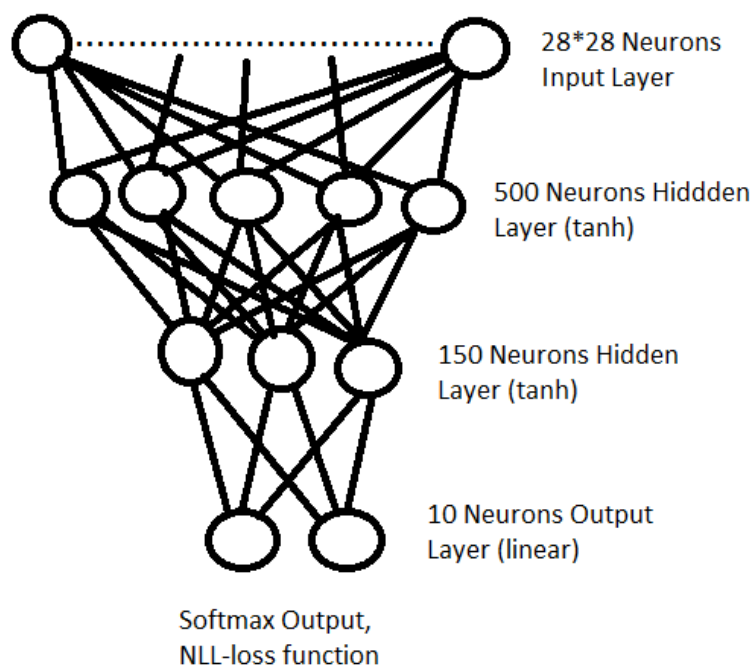


Deep Learning Lab 2018 Exercise 1 Report

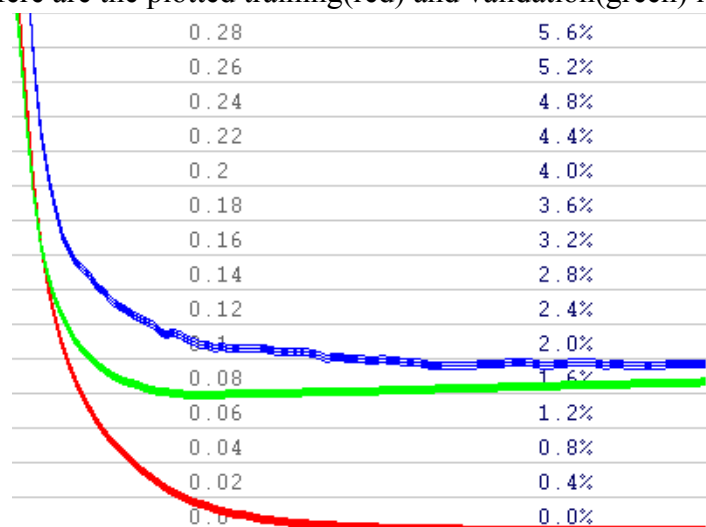
I was able to reach an error of 1.95 % on the validation Set.
The neural network i used has the following structure:



I played around with the number of neurons and different activation functions and i noticed that the results got better when i used more neurons, especially in the first hidden layer. Maybe the error would be lower with even more neurons but I did not add more because the calculation time would be too long. I did not notice a big difference between tanh and sigmoid but they both gave me better results than ReLu.

I used stochastic gradient descent with a batch size of 64 and a learning rate of 0.1. I was not able to get good Results with normal gradient descent, mainly because the learning was so much slower.

Here are the plotted training(red) and validation(green) loss and the error(blue) over 100 epochs:



I think it is interesting that the validation loss starts to rise again while the error on the validation set is still slowly getting smaller.