

## Report – writing our first neural network

The different hyper-parameters:

layer size:

for sizes over 20 the model got fine results. Below was not enough.

For sizes over 180 the model took much longer time to run with not much improvement.

Therefore, I chose: 150, which was better than 20 but not too long to run and check.

learning rate:

for 0.1 or above the model is bad and not really coming closer to the minimum.

0.01 was much better.

I chose 0.005 which was going toward the minimum of the loss, and not being too slow.

activation function:

Relu worked fine and got good result, and the same was true for sigmoid. I chose sigmoid

because for some reason it appeared to be running faster.

Epoch number:

after around 50 epochs the loss of the validation set started becoming worse, so I chose 50.

Most of the different combinations of hyper-parameters yielded a result of about 80-85 percent correctness on the validation set.

Eventually, I received 89 percent using the following combination:

layer size = 150

learning rate = 0.005

activation function = sigmoid

epoch: 50