

# Deep Learning Exercise 1 Report

Kai Haase, 4118916

November 5, 2017

## 1 Overview

As an undergraduate student with no academic experience in ML, the task of the last three weeks proved to not be easy. But by communicating ideas with other students and spending some extra time on the topic, I was able to build a neural network with pleasing results.

Also the code stub and the question times helped a lot in giving a rough idea on which things to work on in which order.

## 2 Results

The final Neural Net was able to classify the images in the validation dataset with an error of about 3% using adagrad. It lead to similar results using stochastic gradient descent and gradient descent, but a lot slower. I also tried implementing a dropout, but it did not prove to work very well. I think it may have to do with the network not being large enough for it to be very effective, or maybe it just needs a lot of time to train.

## 3 Problems and Experiences

Of course the implementation of the neural network, especially of gradient descent, was the biggest part of the task. Just getting the NN to actually feed forward and backward some inputs proved to be very time consuming, as the dimensions of all the inputs, weights, gradients etc. could be very confusing to get right.

Also when using gradient descent and stochastic gradient descent, the error would stay constant for some time until it began to descent. The error at the start would always be 0.8864, and with gradient descent, it stayed there for a really long time.

Here are two pictures the classifier could not label correctly:

Figure 1: Predicted 1, actually was a 6

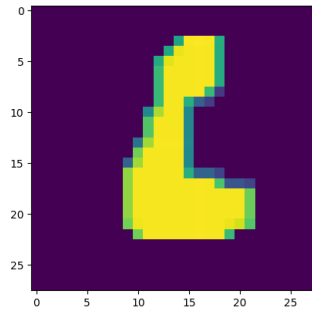
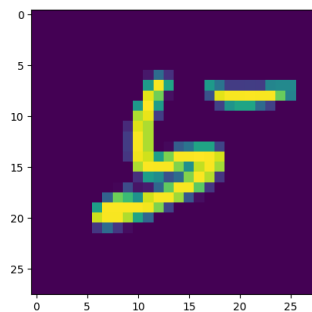


Figure 2: Predicted 4, actually was a 5



## 4 Next things to find out

It would now be interesting to find out how to design a good architecture for a neural network, given a problem. For this task, I chose the number of layers and their size as well as the type of activation function purely based on intuition. So I guess my neural network is not nearly as good as it could be, using the tools we have to work with.

Also it seems like some of the pictures were extremely hard to label, even for humans, because they were just drawn so strangely. Maybe fully connected layers may just not be powerful enough to classify them with human-like or even better performance. My estimation of human performance here would be about 99.3% or so, since there does seem to be a lot of hard pictures in the dataset.