Deep learning - Homework 1

Vito Založnik (63210496)

1 Introduction

In the first homework, we had to implement the basic fully connected neural network. We also had to implement some improvements to it such as adam optimizer, exponential decay schedule and L2 regularization. In the report, I will present the results achieved with different hyperparameter settings.

2 Results

Our first task was to finish $forward\_pass$, $backward\_pass$, and $update\_network$ functions and tune hyperparameters to achieve a classification score of at least 0.48. We achieved this with hidden layers of 100, 100 neurons, and an output layer of 10 neurons. We used 40 epochs, a minibatch of size 20, and a learning rate of 0.02. With those hyperparameters, we achieved a classification score of 0.4934. The biggest improvement was made by lowering the minibatch size and increasing the number of epochs. With a learning rate of 0.01, we achieved a classification score of 0.4816.

2.1 Improvements

We implemented $adam$ optimizer and $exponentialdecay$ learning rate. We tried improving the classification score by tuning these hyperparameters. First, we tried to tune the learning rate on an already selected network with layers with 100, 100, and 10 neurons. With a learning rate of 0.02, we achieved a score of 0.4934 which was also the highest among all hyperparameters combinations.

With adam optimizer we did not achieve better results. We also did not improve the classification score by tuning exponential decay rate on previously selected learning rate and network dimensions hyperparameters.

```
learning rate 0.005:

        sgd: 0.4554
        adam:  0.4421

learning rate 0.01:
        sgd: 0.4816
        adam: 0.3711

learning rate 0.02:
        sgd: 0.4934
        adam: 0.3101

learning rate 0.02:
        decay rate: 0.001:
                sgd: 0.3353
                adam: 0.221


learning rate 0.02:
        decay rate: 0.0001:
                sgd: 0.3758
                adam: 0.2731
```

In the figure above, we can see calculated classification scores with some different hyperparameters.

3 Conclusion

In this homework, we had to evaluate several different training regimes for training a fully connected neural network. We have done so with several different parameters for most of the parameters and represented them in the report. We feel that parameters could be further tuned to achieve better scores. Also time complexity could be improved by selecting different number and size of layers.