Deep Learning Lab 2018

Srdjan Milojevic

Exercise 1

1 Overview of the task

The main goal of the exercise was to implement logic into a pre-composed code. This includes initialization of weights and biases, forward and backward propagation, calculating the gradients of output layers, gradient descent and stochastic gradient descent methods and finally, optimizing hyperparameters.

2 Dropout

One of the additional tasks was to implement dropout. This regularization technique reduces overfitting to the training data. In the realized code, the author has implemented dropout in the last hidden layer with the probability of 0.4. Creation of the dropout mask is achieved with the following line:

```
self.DO = (np.random.rand(*input.shape)<self.dropout_prob)/self.dropout_prob
```

Additionally, it was necessary to multiply (elementwise) the input of the corresponding layer with its mask. Moreover, the same goes for backpropagation.

3 Results and Conclusion

The final results are shown on Figure 1.

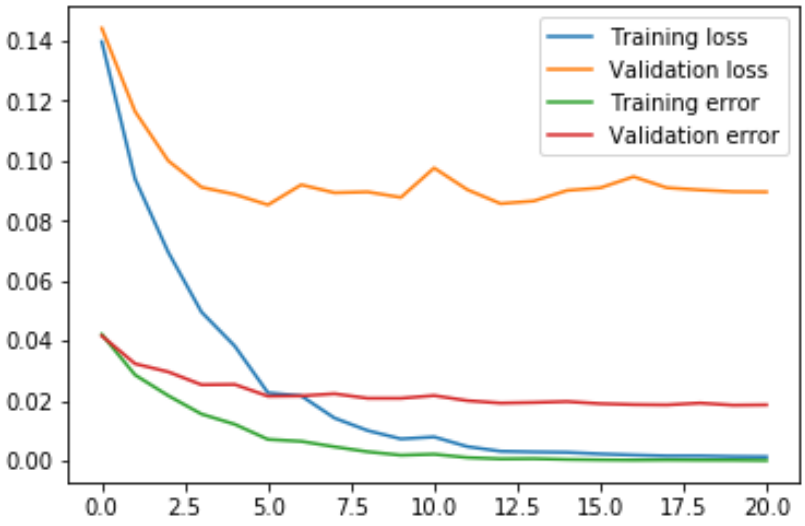

Figure 1: Learning curves of the realized neural network.

The author uses stochastic gradient descent because of its speed. With the mentioned dropout regularization and fine-tuning of other hyperparameters the realized neural network achieved an error rate of XXX on the test set. Another technique that the author uses is the adaptive learning rate, in order to find the minimum of the gradient.

On the Table 3 the author compares the error rate in various cases with and without dropout regularization, and in stochastic gradient descent and normal gradient descent. It is visible that stochastic gradient descent achieves better performance with the same number of training epochs. Clearly, better results are achieved with the dropout in case of SGD because it reduces the overfitting. On the other hand, there is noticeable degradation of the GD network when introduced with dropout.

Learning curves of the realized neural network.

Neural Network Variations		
Dropout	SGD	GD
Enabled	1.780%	37.770%
Disabled	1.890%	26.910%

The final code can be found here.