

30.10.2018

Assignment

The first objective of this exercise was to complete a given code stub, which implements a simple feed-forward-neural network. After that, this network was to be trained and evaluated on the MNIST data set. Using this, the final goal was to investigate the effect of the different parameters of the network.

Implementation

The implementation proved to be more tedious than we originally expected, which was probably mostly due to a recent lack of practice of Python programming. Therefore, we largely stuck to the suggested TODOs and refrained from trying out advanced approaches at this point.

Evaluation

Running the given example network in its default configuration lead to a validation lead to a final test error 0.44% of and validation error of 2.40%. Manipulating the parameters of the this network yielded the following results:

- For different learning rates we found the best results in terms of validation error to be in the range from 0.1 to 0.5. As expected, larger rates have a generally steeper decrease, but got significantly worse at around 0.7.
- Changing the activation functions to *tanh* or *sigmoid* yielded worse results in terms of quality (significantly worse for *sigmoid*) and longer runtimes (around 180s up from around 140s).
- Adding an additional layer before the first one also deteriorated the results, with a runtime of 207s and final validation error of 3.26%. However, additionally increasing the maximum number of epochs to 30 brought the error down to 2.23%, supposedly since more parameters just need more training. On the other hand, this also increased the runtime to 302s.

All in all I cannot say, that I found the silver bullet in terms of parameters for this setting. After some more toying around, I chose a network similar to the original one for my final attempt, with a learning rate of 0.5. The end result was an error of 1.91 on the test set and the learning curve presented in Figure 1.

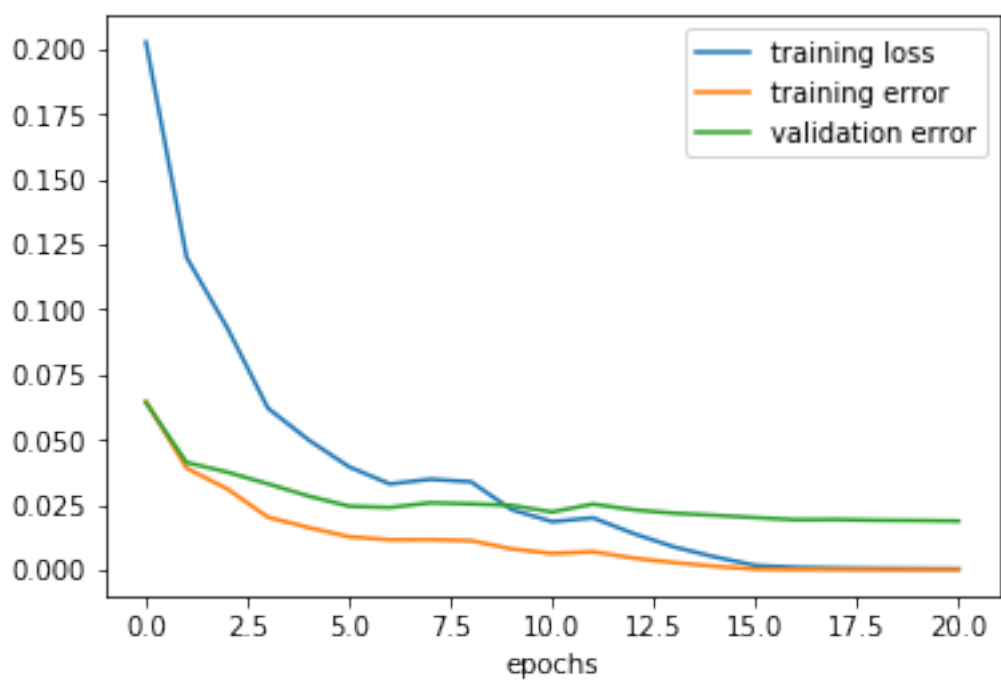


Figure 1: Learning curve of the final neural network