Neural Networks KU WS18

708.076 18W

Exercise Sheet 4

Optimizing Deep Neural Networks [4P]

For this exercise, form groups of two students (see also last paragraph).

Use TensorFlow 1.5 in Python 3 for this task.

Consider the dataset used for Task 2 (isolet). Use the data and load-procedure provided in there.

Optimize cross entropy error of the network, however, all performances should (also) be reported in terms of the misclassification rate (percentage of misclassified examples). For optimization, use the ADAM algorithm and minibatches of size 20. We do not need a validation set for this task. Report erros on the training set and the test set.

Task details [4P]:

a) Train a deep network with 9 hidden layers and 40 neurons per layer. Compare networks with tanh nonlinearities to networks with ReLu nonlinearities. Find a good learning rate. Provide plots where the evolution of the training/test error during training is shown for a number of learning rates. What effects do you observe (in comparison to results of Task 2)?

b) Now, use a ResNet architecture with the same total number of hidden layers and neurons per layer. The first hidden layer should be a standard one with a ReLu nonlinearity. Then add four residual blocks as shown in Figure 1. Let $\mathbf{x}$ be the output of the previous layer. A residual block computes

$$\mathrm{ReLU}\left(\mathbf{x} + W_2\mathrm{ReLU}(W_1\mathbf{x} + \mathbf{b}_1) + \mathbf{b}_2\right), \tag{1}$$

where W_1 and W_2 are weight matrices, $\mathbf{b}_1$ and $\mathbf{b}_2$ are bias vectors, and ReLU is the rectified linear activation function applied component-wise to the argument vector. Find a good learning rate. Provide plots where the evolution of the training/test error during training is shown for a number of learning rates.

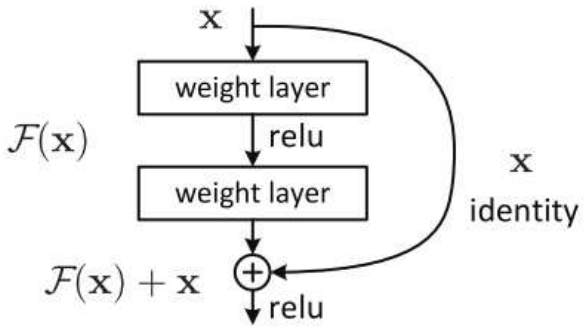

Figure 1: A residual block. $\mathbf{x}$ denotes the output of the previous layers.

Submit the code as a single .py file and a report as PDF in the teach center. Do **not** provide one zip with code and PDF but submit them separately. The code should be executeable under the assumption that the files provided at the homepage are available (unzipped) at the working directory. Both students of a group have to submit their python code and report (it can be identical to the one of the group partner). In the report, indicate your group partner on top of the first page (write "Group partner: ⟨First name⟩, ⟨Last Name⟩, ⟨Matrikelnumber⟩").