Exercise 3 - Optimise your Optimiser

… about Adam and Eve

Learning goals

This exercise is targeted at playing around with different optimisers. Your choice of the optimiser for a specific task.

Thus, the learning goals are:

- … getting to know different optimisers.
- … knowing the differences, advantages and disadvantages of the different optimisers.
- … experimenting with the capacity of the network (hidden dimension) and corresponding dropout.

Deliverables

Upload a PDF file containing your remarks on OLAT.

Format: Ex03_name_surname.ipynb

Deadline

Deadline for Exercise 3 is **12.11.2018, 12:00 (MESZ)**.

Data

Use the code by Görner from Github: https://github.com/martin-gorner/tensorflow-rnn-shakespeare.

Framework

You can transfer the code to Google's Colab in order to let it run on GPUs.

Tasks

1. Use three of the following classifiers for the language modeling task:
    a. SGD
    b. Adam
    c. Adadelta
    d. Adagrad
    e. RMSprop
    f. NAG

    Read Géron pp. 293-302 for an overview. Also consider this blogpost for a good overview.
2. Report the accuracy of the models for each optimiser. Explain why there are differences and where they might come from. Experiment with the size of the hidden layers and the dropout. How do these adjustments influence the training outcome, and how is this related to the optimiser? Görner provides code which lets you display the training curves on the Tensorboard. Consider them in your write-up. Your report should be between 400 and 500 words (without code).