Barbara Hammer, Philip Kenneweg Universität Bielefeld

Deep Learning (SoSe 2024)
1. Sheet

Start: Thursday, 11.04.2024.
End: The worksheets should be solved using Python, in groups of 2-3 people and will be presented in the Tutorials.
Discussion: Thursday, 25.04.2024 in the Tutorials.

Information

The worksheets and necessary toolboxes will be made available in the Lernraum "392221 Deep Learning (V) (SoSe 2024)". Worksheets will usually be released every two weeks on Thursday, and discussed during the exercises on Thursday two weeks later. In order to successfully finish the course, 50% of the available points have to be obtained and each participant has to present his/her results at least once. The Wednesday and Thursday in between the release and discussion of the sheet will be used to discuss the implementation of the various algorithms presented in the lecture, as well as go deeper into the relevant material.

Exercise 1:

(*10 Points*)

You can use code and models which are publicly available. Please provide: short description what you did, how it is done, what is the result. Please be prepared to present the solution in the exercises (best in form of a Jupyter notebook .ipynb).

(a) *(5 Pts.)* Take the Fashion MNIST data set [1] and a suitable model architecture (MLP, Convolutions, ...). Display the effect of the choice of the activation function, i.e., investigate different learning and generalization behavior if sigmoid or a modern activation function (e.g. relu, selu, glu, ...)[2] is used. Compare the behavior of at least three different activation functions.

(b) *(5 Pts.)* Take the same setup and the ADAM optimizer. Perform a small study on the effect of different batch sizes for mini-batch training for the ADAM and the SGD optimizer. Take at least three choices (e.g. 1,10,100) and shortly discuss the differences. Do different batch sizes require further adjustments to the hyperparameters of the optimizers? What is the most compute efficient choice?