Tutorial 3: CNN application

In this tutorial, you will learn to apply a CNN for the classification of the MNIST dataset (we did this before in Tutorial 1 with a FNN, so we can compare) using PyTorch (see Notebook Tutorial_3.ipynb).

(i) Study the code, in particular the structure of the CNN as implemented by PyTorch. How many weights in total have to be optimised in this example?

(ii) In this notebook, we use the Adam variant of the stochastic gradient descend method. How does this differ from the traditional variant of this method?

(iii) Determine the effect of the kernel size in the first convolutional layer on the decrease of the loss function.

(iv) Determine the effect of the padding size in the first convolutional layer on the decrease of the loss function.

(v) Compare the results to that of the FNN used in Tutorial 1 for the classification of the same data. To do this, think of a suitable case for an honest comparison.