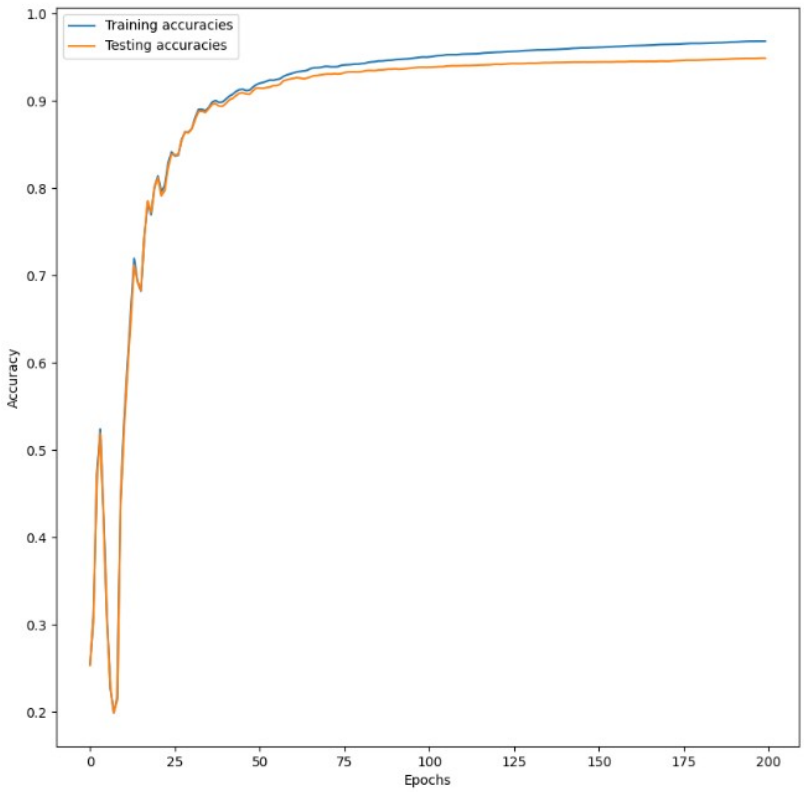
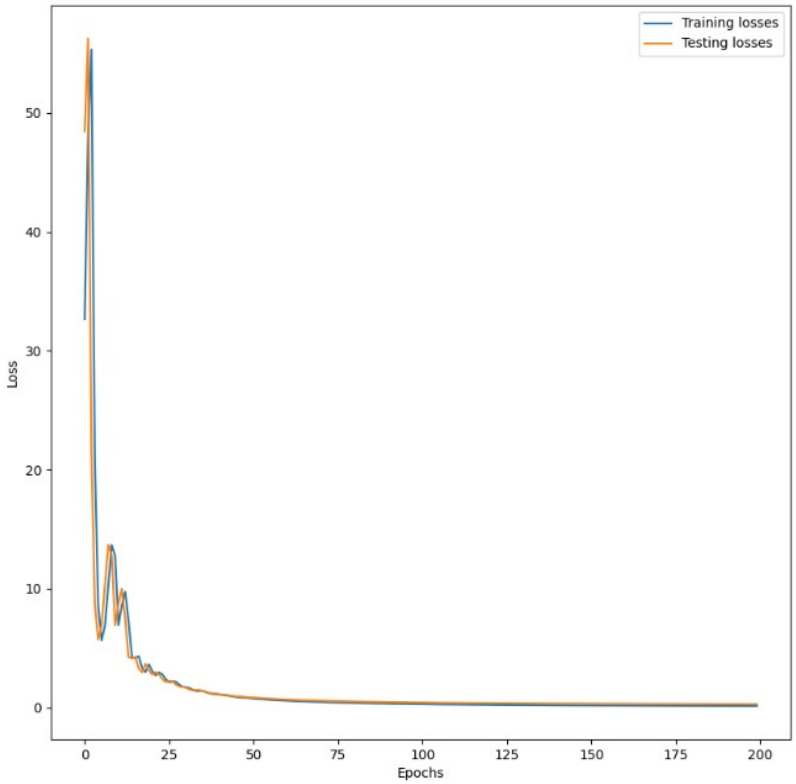
Pattern Recognition

02c CNN

The aim of the task was to implement a Convolutional Neural Network in order to classify a dataset containing 26999 training and 15001 testing instances based on the MNIST dataset, which contains handwritten digits from zero to nine. For this we developed a convolutional neural network consisting of one convolutional and one linear layer.
We then ran the network for several epochs on the training and testing data and tried different learning rates.
We realized that the accuracy rate for both the test and trainig set, after quite some fluctuations in the beginning, started to increase steadily for all consecutive epochs. After around 100 epochs the rate still increased but the improvements were no more significant. So as to still improve the accuracy but not run forever we set the epochs to 200. The same holds true for the loss, but instead of increasing steadily, the loss decreased, as expected. For our network architecture we found out that a learning rate of 0.001 is best. While small increases of the learning rate still lead to increasing accuracies, the improvements were slower and with higher learning rates, the losses exploded.

With a learning rate of 0.001 and running the algo for 200 epochs, we ended up with a test accuracy of 0.948670088660756 and a loss of 0.29194286465644836.




Training losses
Testing losses
Loss
Epochs
0
10
20
30
40
50
0
25
50
75
100
125
150
175
200