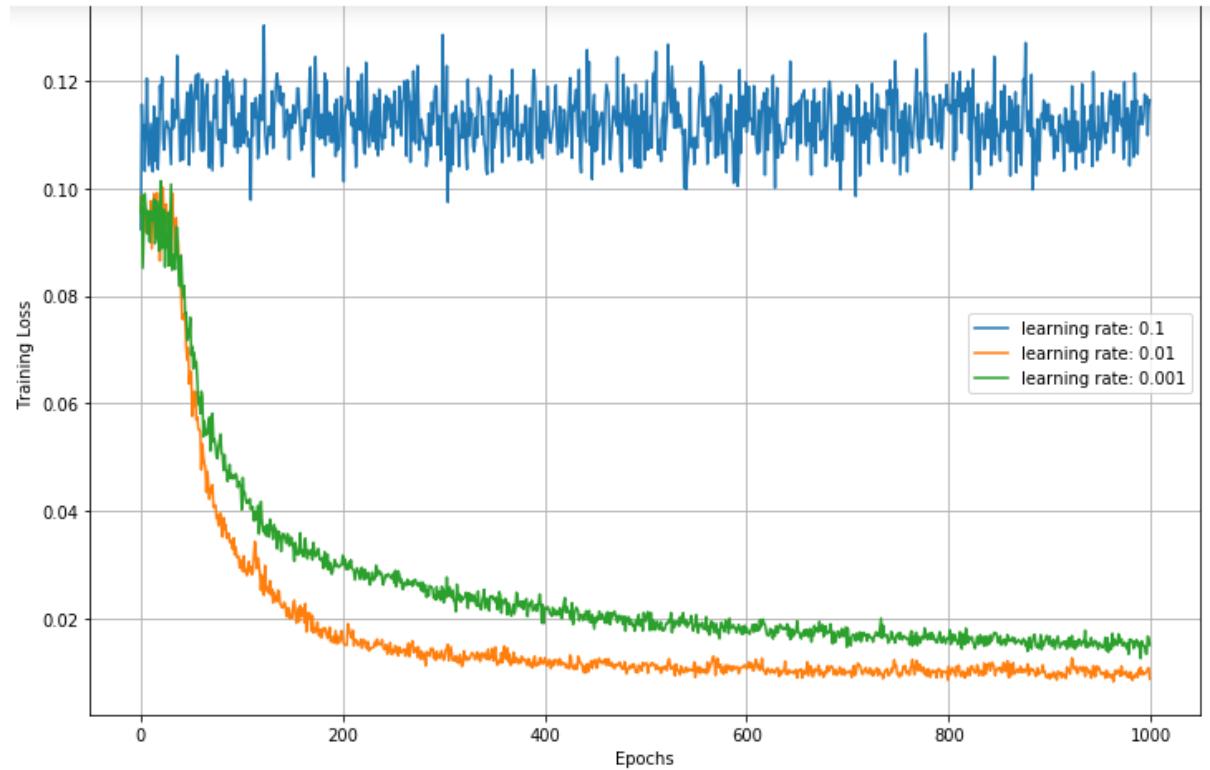


Report Exercise 3

Task 2, learning rates

Because of limited time frame, the model was only tested on a CPU machine and with 1.000 epochs with batch size 64. This did not reach the best possible performance, but showed the behaviour at the different learning rates.



The model did not make any progress at a learning rate of 0.1. Eventually the variance of the gradients was too high, to allow the optimisation process to progress. The best performance was achieved with a learning rate of 0.01. It might be, that after much longer training, the training at the lowest rate of 0.001 might reach the same level.