learning curves: SGD, one hidden layer NN
overparametrized
input dim: 50, batch size: 20
hidden dim: 80
output dim: 1

Legend:
- $\gamma = 0.005$
- $\gamma = 0.001$
- $\gamma = 0.0001$
- $\gamma = 1e\text{-}05$

x-axis: iteration
y-axis: loss