FACULTY OF COMPUTER SCIENCE, HEINRICH-HEINE-UNIVERSITY DÜSSELDORF

Deep Learning

Winter term 25/26 – Exercise Sheet 2

Submission Deadline: Monday, November 3, 2025, 2:00 PM

1. **Gradient descent (3P)**

   a) What is (stochastic) gradient descent? (2P)

   b) What is the difference between full batch and mini-batch GD? (1P)

2. **Backpropagation (7P)** Given is a tiny neural network, with 1 input neuron, 2 hidden neurons and 2 neuron in the output layer. We use the ReLU as activation function, except for the output where we apply a softmax function. The network is trained with mean squared loss. Calculate the gradient for all parameters excluding biases.