



28.02.2022

LAST NAME	First name	Matrikelnb.

Exam Deep Learning, WS21

Group B

You can answer in German or English.

1. Neural Networks (15 Points):

- (a) Describe L2 regularization; Interpreted as MAP inference, to which prior does it correspond? Give a formula for the weight update induced by L2 regularization.
- (b) Consider a single linear neuron for which we optimize the weights using the mean-squared-error. Draw the computational graph for this setup.
- (c) Briefly describe dataset augmentation as a regularization technique. Why is it effective as a regularizer?
- (d) Briefly describe m-fold cross-validation.

2. Modern Architectures (15 Points):

- (a) Describe a convolutional layer in a Convolutional Neural Network (CNN).
- (b) Which statistical insights about images are exploited by CNNs?
- (c) What are the reasons for using pooling layers in CNNs?
- (d) What is the stride parameter in a pooling layer?

3. Gradient Descent (10 Points): You have some one-dimensional training data $\langle x_1, x_2, \dots, x_N \rangle$ where $x_i \in \mathbb{R}$ and corresponding targets $t_1, \dots, t_N \in \mathbb{R}$. You know that the targets depend approximately exponentially on x . Therefore, you want to fit this data with the following model

$$y(x) = a \exp(b \cdot x), \quad (1)$$

with parameters $a, b \in \mathbb{R}$.

You want to minimize the sum-squared error of the fit:

$$E(a, b) = \frac{1}{2} \sum_n (t_n - y(x_n))^2. \quad (2)$$

Derive an update rule for a and b for batch gradient descent on the error (2) of the model.