# Deep Learning

## Theoretical Exercises – Week 8 – Chapter 7

Exercises on the book "Deep Learning" written by Ian Goodfellow,
Yoshua Bengio, and Aaron Courville.
Exercises and solutions by T. Méndez and G. Schuster

FS 2024

# 1 Exercises on Regularization for Deep Learning

1. What is the goal of regularization?

   ***Solution:***
   One goal of machine learning is, that the algorithm generalizes, that is, that it will perform well on data with which it has never been trained. Regularization is any strategy that is designed to improve generalization. Hence, their goal is to reduce the test error, often at the expense of increasing the training error.

2. Describe the following regularization methods in your own words, by explaining the idea behind the method and describing, how it acts as a regularizer.

   (a) Weight decay

   (b) Dataset Augmentation

   (c) Noise Robustness

   ***Solution:***

   (a) **Weight decay**: With weight decay the capacity of the model is limited, by adding a parameter norm penalty, that prevents the weights from growing unnecessarily large. As a result, only those weights become large, which significantly contribute to the reduction of the objective function. The other (unimportant) weights remain small.

   (b) **Dataset Augmentation**: Dataset augmentation improves generalization by training the network on more data. To get more data, artificial training data is created, whereby different methods are available depending on the task at hand.

   - Transforming the inputs of the training set. This is easiest for classification, especially object recognition, since the label does not change, if the images are rotated, moved or scaled.

- Variation of the inputs by injecting noise at the input of the network.

(c) **Noise Robustness**: This method makes the network robust to noise. It encourages the parameters to go to regions in the parameter space, where small perturbations of the weights (or the inputs) have relatively small influence on the output (minima surrounded by flat regions). The noise can be injected at different locations:

- Injecting noise at the output of the hidden layers.
- Adding noise to the weights.
- Injecting noise at the output targets (for maximum likelihood learning schemes with softmax classifiers).