

Processing of Biological Data

Prof. Dr. Volkhard Helms
Winter Semester 2018-2019

Saarland University
Chair for Computational Biology

Exercise Sheet 4

Due: January 10, 2019 23:59

Submission

- You are advised to work in groups of two people. If necessary, we will suggest teammates.
- Submit your solutions on paper in Room 3.03, E2 1 or better send an email with a single PDF attachment to maryam.nazarieh@bioinformatik.uni-saarland.de. Late submissions will not be considered. **In any case, hand in all source code via mail. Please also include your output. Otherwise you will loose points.**
- Do not forget to mention your names/matriculation numbers.
- You are free to use any programming language to solve the problems. The usage of libraries that allow you to circumvent implementing the algorithms asked for will not grant you points.

Exercise 4.1: Basics of Deep learning methods (50 points, 10 points for each subtask.)

- What are the benefits of utilizing neural networks in deep learning?
- How does the back propagation algorithm achieve the targeted output in the context of a deep learning method?
- Formulate linear functions for the values of the nodes in the hidden and output layers as a function of the incoming nodes, see Figure 1.
- Determine the values at hidden and output layers if $x_1 = 1$, $x_2 = 2$, $w_1 = 2$, $w_2 = 4$, $w_3 = 3$, $w_4 = 5$, $w_5 = 1$, $w_6 = 2$, $w_7 = 3$, $w_8 = 3$, see Figure 1.
- Determine the mean squared error if the observed outputs are $y_1 = 58$ and $y_2 = 60$.

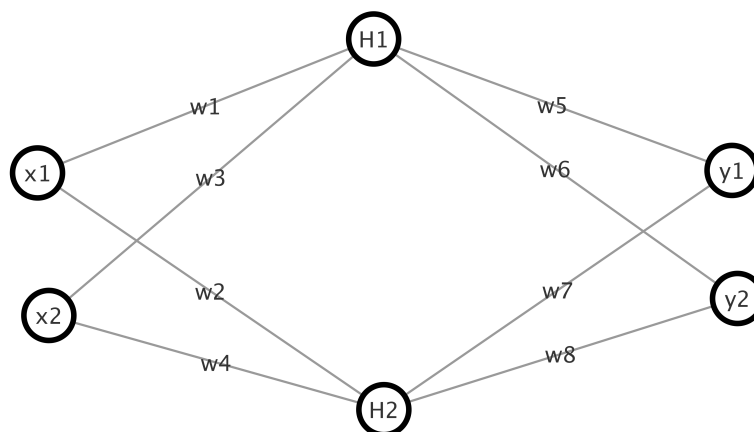


Figure 1: A graphical representation of an example neural network.

Exercise 4.2: TensorFlow (50 points)

TensorFlow is a machine learning framework that was developed by Google technology company. It is used to design, build and train deep learning models. Basically TensorFlow is a flow of multidimensional data arrays that communicate.

Install TensorFlow by following the instructions available for Ubuntu, Windows, macOS, and the Raspberry Pi at <https://www.tensorflow.org/install/>.

Download the code given in the supplementary. The code uses the Fashion MNIST dataset.

- (a) Submit with your solution a plot showing the first 36 images from the training set and display the class name below each image. (10 points)
- (b) Build up two neural network models separately by considering 64 and 256 nodes in the first dense layer. Describe in your solution what parts of the code need to be changed for this and submit the modified code line (s). (20 points)
- (c) Compare the performance of these two models on the test dataset. Report the accuracies in your solution. (10 points)
- (d) Determine and report the confidence of the model for the 1st and 10th images in the test dataset. (10 points)

Have fun!