Task Sheet 1

Nico Schmidt (nico.schmidt@uzh.ch), Mathias Weyland (mathias.weyland@uzh.ch)

Notes

Please remember to write both your name and student ID number at the top of each sheet before it is handed in. You are encouraged to work in groups of two. If you do so, please hand in only one solution and write both names and student IDs on it. Please make an effort to write legibly.

When you hand in the solutions, please staple the sheets together. Alternatively, you can send your solution as one single PDF file by email to mathias.weyland@uzh.ch.

Questions

1. State the "five basics" of neural networks and illustrate them with an example. 1P
2. State three differences between biological neural networks and conventional computers. 1P
3. In artifical neurons, we use an abstract notion of *weights*, *output*, *input* and *connections* between neurons. Figure 1 shows a biological neuron. Point out the entities to which these four abstract terms correspond. 2P

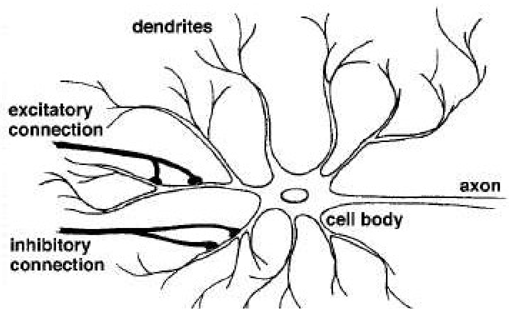
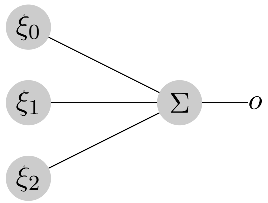


Figure 1: A biological neuron.

4. The NOR function is to be learned with the simple perceptron shown in figure 2. 4P
Complete the table below assuming a learning rate of $\eta = 0.6$, a binary threshold of $\Theta = 0$ and the perceptron learning rule $w_i(t+1) = w_i(t) \pm \eta x_i(t)$. Note that weights are only updated if $o \neq \zeta$.



ξ_1	ξ_2	ζ
0	0	1
0	1	0
1	0	0
1	1	0

Figure 2: Simple perceptron (left) and NOR function (right).

ξ_0	ξ_1	ξ_2	w_0	w_1	w_2	$\sum w_i \xi_i$	o	ζ
1	0	0	-0.4	-0.5	0.5	-0.4	0	1
1	0	1	0.2	-0.5	0.5	0.7	1	0
1	1	0						
1	1	1						
1	0	0						
1	0	1						
1	1	0						
1	1	1						
1	0	0						

5. There are 16 different binary functions with two inputs and one output. Explain why this is. 2P

6. Out of all the above functions, draw the ones that are not linearly separable into coordinate systems. Use the two axes for the inputs and different symbols for the output and explain why they are not linearly separable. 2P

7. Given a perceptron with a weight vector $w^\mathsf{T} = (w_0, w_1, w_2) = (3, 1, 3)$ where node 0 is the bias node with $\xi_0 = 1$, draw a sketch of the decision boundary in $\mathbb{R}^2$ and mark for which area the perceptron outputs 0 and 1, respectively. 2P

8. Consider a simple perceptron with 4 input nodes and 1 output node. $\xi^{(i)}$ is an input 4P

pattern and $\zeta^{(i)}$ is the corresponding desired output. For the following two cases, state whether the patterns are linearly separable and/or linearly independent.

a) $\xi^{(1)} = (1,0,1,0)^\mathsf{T},\ \zeta^{(1)} = 0$
$\xi^{(2)} = (1,0,0,0)^\mathsf{T},\ \zeta^{(1)} = 1$
$\xi^{(3)} = (0,1,0,0)^\mathsf{T},\ \zeta^{(1)} = 0$
$\xi^{(4)} = (0,0,0,0)^\mathsf{T},\ \zeta^{(1)} = 1$

b) $\xi^{(1)} = (1,0,0,0)^\mathsf{T},\ \zeta^{(1)} = 0$
$\xi^{(2)} = (0,0,1,1)^\mathsf{T},\ \zeta^{(1)} = 1$
$\xi^{(3)} = (0,1,0,0)^\mathsf{T},\ \zeta^{(1)} = 0$
$\xi^{(4)} = (0,0,0,0)^\mathsf{T},\ \zeta^{(1)} = 1$

9. For the following statements, tell whether they hold true or not: 4P
   a) Adalines and Perceptrons use the same activation function.
   b) The learning rules for Perceptron, Adalines and MLP can be described by the generalized delta rule. Only the delta differs in the three rules.
   c) Adalines use a threshold, whereas Perceptrons do not.
   d) Perceptrons adjust their weights only in the case of an error, Adalines adjust them always.
   e) For the Perceptron there exists an explicit solution (without learning), if the classes are linearly independent.
   f) Perceptrons and Adalines can only learn linearly separable problems successfully.
   g) By adding more layers, Adalines can learn non-linear problems.
   h) MLPs can learn non-linear problems successfully.

10. Plot the two functions 4P

$$g(x) = \frac{1}{1+\exp(-x)},\ h(x) = \tanh(x)\ ,$$

compute their derivatives and verify the properties

$$\frac{\mathrm{d}g(x)}{\mathrm{d}x} = g(x)\,(1-g(x)) \text{ and } \frac{\mathrm{d}h(x)}{\mathrm{d}x} = 1 - h(x)^2$$

11. A simple perceptron is used to learn patterns. The learning algorithm did not converge after 10^6 iterations. Can you claim that the patterns belong to classes that are not linearly separable? Justify your answer. 2P

12. The decision boundary of a simple perceptron is always a hyperplane. In the case of two inputs, it is a straight line in the input space $\mathbb{R}^2$. Show mathematically why this is (for two inputs). 2P