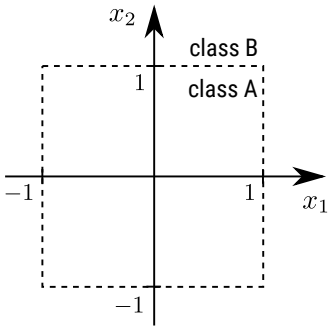
Exercises for the course
Machine Learning 1
Winter semester 2021/22

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Exercise Sheet 6

Exercise 1: Designing a Neural Network (25 P)

We would like to implement a neural network that classifies data points in $\mathbb{R}^2$ according to decision boundary given in the figure below.



We consider as an elementary computation the *threshold neuron* whose relation between inputs $(a_i)_i$ and output a_j is given by

$$z_j = \sum_i a_i w_{ij} + b_j \qquad a_j = 1_{z_j > 0}.$$

(a) *Design* at hand a neural network that takes x_1 and x_2 as input and produces the output "1" if the input belongs to class A, and "0" if the input belongs to class B. *Draw* the neural network model and *write down* the weights w_{ij} and bias b_j of each neuron.

Exercise 2: Backward Propagation (5 + 20 P)

We consider a neural network that takes two inputs x_1 and x_2 and produces an output y based on the following set of computations:

$$z_3 = x_1 \cdot w_{13} + x_2 \cdot w_{23} \qquad z_5 = a_3 \cdot w_{35} + a_4 \cdot w_{45} \qquad y = a_5 + a_6$$
$$a_3 = \tanh(z_3) \qquad a_5 = \tanh(z_5)$$
$$z_4 = x_1 \cdot w_{14} + x_2 \cdot w_{24} \qquad z_6 = a_3 \cdot w_{36} + a_4 \cdot w_{46}$$
$$a_4 = \tanh(z_4) \qquad a_6 = \tanh(z_6)$$

(a) *Draw* the neural network graph associated to this set of computations.

(b) *Write* the set of backward computations that leads to the evaluation of the partial derivative $\partial y / \partial w_{13}$. Your answer should avoid redundant computations. Hint: $\tanh'(t) = 1 - (\tanh(t))^2$.

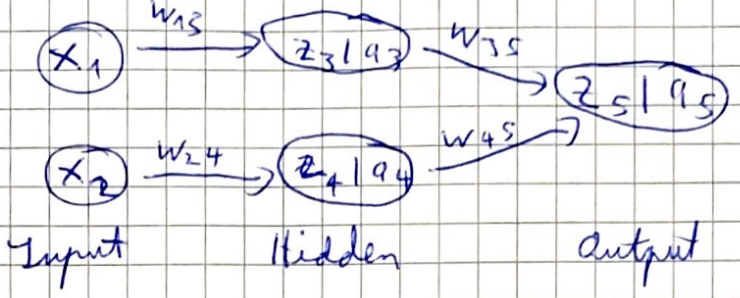
Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.

Sheet 6

1]

We design a neural network with three layers and three units:



We set $w_{13} = w_{24} = 1$, we set $w_{35} = w_{45} = 0,5$, we set all bias terms $b_1 = b_2 = b_3 = b_4 = b_5 = 0$ and we define a threshold function as activation function with $g(x) = \begin{cases} 1 \text{ if } |x| \leq 1 \\ 0 \text{ if } |x| \geq 1 \end{cases}$

Therefore

$$a_3 = g(z_3) = g(w_{13} \cdot x_1 + b_1) = g(x_1) = \begin{cases} 1 \text{ if } |x| \leq 1 \\ 0 \text{ if } |x| \geq 1 \end{cases}$$

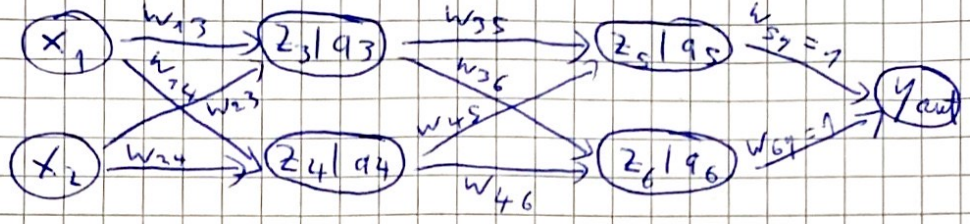
$$a_4 = g(z_4) = g(w_{24} \cdot x_2 + b_2) = g(x_2) = \text{"}$$

$$a_5 = g(z_5) = g(a_3 \cdot w_{35} + a_4 \cdot w_{45} + b) =$$

$$= g(a_3 \cdot 0,5 + a_4 \cdot 0,5) = \begin{cases} 1 \text{ if } |x_1| \text{ and } |x_2| < 1 \\ 0 \text{ if } |x_1| \text{ or } |x_2| \geq 1 \end{cases}$$

2]

a)



b) ~~[illegible]~~

~~[illegible]~~

~~[illegible]~~

$$\frac{\partial y}{\partial w_{13}} = \frac{\partial y}{\partial a_5}\frac{\partial a_5}{\partial z_5}\frac{\partial z_5}{\partial a_3}\frac{\partial a_3}{\partial z_3}\frac{\partial z_3}{\partial w_{13}} + \frac{\partial y}{\partial a_6}\frac{\partial a_6}{\partial z_6}\frac{\partial z_6}{\partial a_3}\frac{\partial a_3}{\partial z_3}\frac{\partial z_3}{\partial w_{13}}$$

$$= \left(\frac{\partial y}{\partial a_5}\frac{\partial a_5}{\partial z_5}\frac{\partial z_5}{\partial a_3} + \frac{\partial y}{\partial a_6}\frac{\partial a_6}{\partial z_6}\frac{\partial z_6}{\partial a_3}\right)\frac{\partial a_3}{\partial z_3}\frac{\partial z_3}{\partial w_{13}}$$

$$= \left[1 \cdot (1 - \tanh^2(z_5)) \cdot w_{35} + 1 \cdot (1 - \tanh^2(z_6)) \cdot w_{36}\right] \cdot (1 - \tanh^2(z_3)) \cdot x_1$$

$$= \left[(1 - a_5^2)\, w_{35} + (1 - a_6^2)\, w_{36}\right](1 - a_3^2)\, x_1$$