Exercises for the course

Machine Learning 1

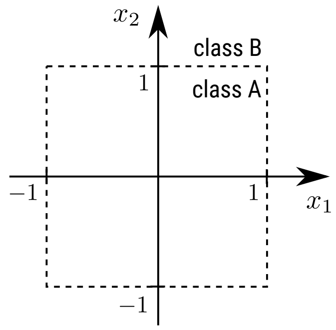
Winter semester 2024/25

Fachgebiet 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 8

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:

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

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

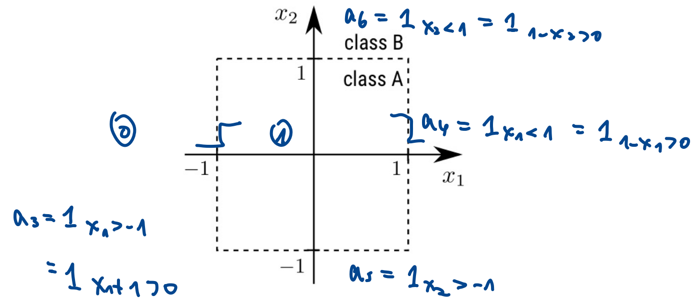
(b) *Write* the set of backward computations that leads to the evaluation of the partial derivative $\partial y / \partial w_{13}$. Your answer should avoid redundant computations. Hint: $\tanh'(t) = 1 - (\tanh(t))^2$.

Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.

Exercise 1: Designing a Neural Network (25 P)

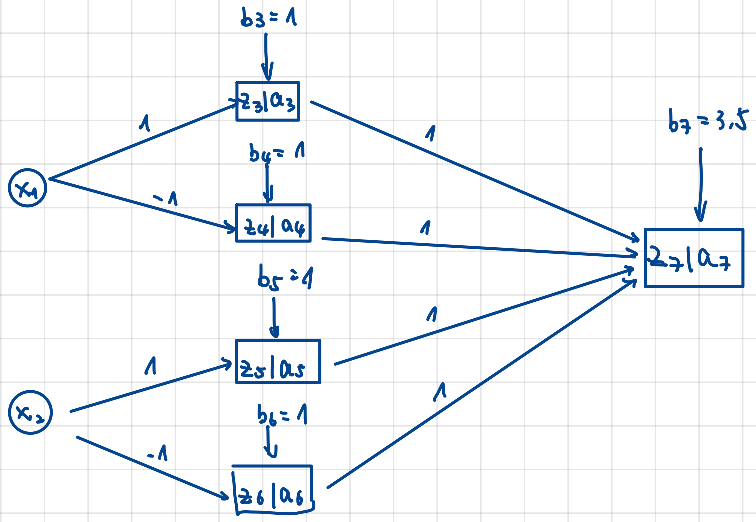
We would like to implement a neural network that classifies data points in $\mathbb{R}^2$ according to decision boundary given in the figure below.



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

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

(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)

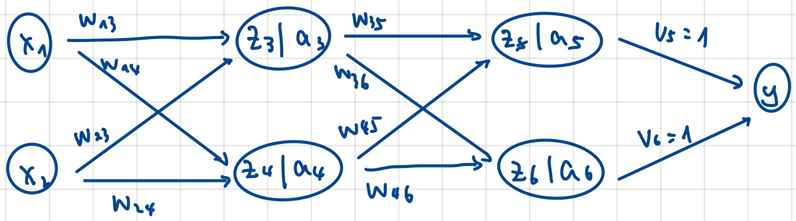
We consider a neural network that takes two inputs x_1 and x_2 and produces an output y based on the following set of computations:

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

(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: $\tanh'(t) = 1 - (\tanh(t))^2$.

a)



b)

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

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

for Programing Part:

$$DY = \frac{\partial \mathcal{E}}{\partial Y} = \frac{\partial}{\partial Y}\left(\sum_{i=1}^{N} \max(0, -YT)\right) = \frac{1}{N}\left(-T \cdot \mathbb{1}_{\{-YT>0\}}\right)$$

$$DZ = \frac{\partial \mathcal{E}}{\partial Y} \cdot \frac{\partial Y}{\partial Z} = DY \cdot V$$

$$DW = \frac{\partial \mathcal{E}}{\partial Z} \cdot \frac{\partial Z}{\partial X}$$

$$= DZ \cdot X$$

$$DB = \frac{\partial \mathcal{E}}{\partial Z} \cdot \frac{\partial Z}{\partial b} = DZ$$

$$DV = \frac{\partial \mathcal{E}}{\partial y} \frac{\partial y}{\partial V} = DY \cdot A$$