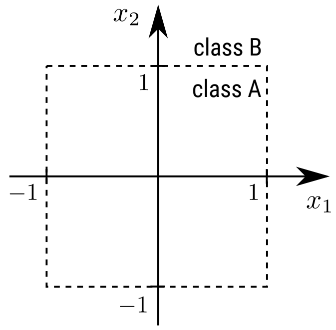
Exercises for the course
Machine Learning for Data Science
Winter Semester 2024/25

G. Montavon
Institute of Computer Science
Department of Mathematics and Computer Science
Freie Universität Berlin

Exercise Sheet 11 (theory part)

Exercise 1: Designing a Neural Network (25 P)

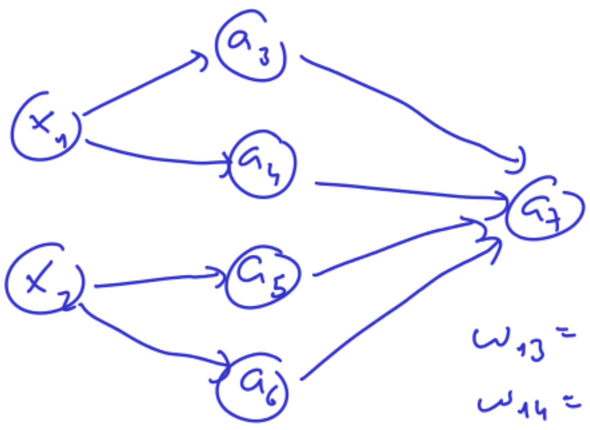
We would like to implement a neural network that classifies data points in $\mathbb{R}^2$ according to decision boundary given in the figure below.



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.



$w_{13} = 1 \quad b_3 = 1$

$w_{14} = -1 \quad b_4 = 1$

$w_{25} = 1 \quad b_5 = 1$

$w_{26} = -1 \quad b_6 = 1$

$w_{37} = w_{47} = w_{57} = w_{67} = 1 \quad b_7 = -3.5$

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:

$$z_3 = x_1 \cdot w_{13} + x_2 \cdot w_{23} \qquad z_5 = a_3 \cdot w_{35} + a_4 \cdot w_{45} \qquad y = a_5 + a_6$$

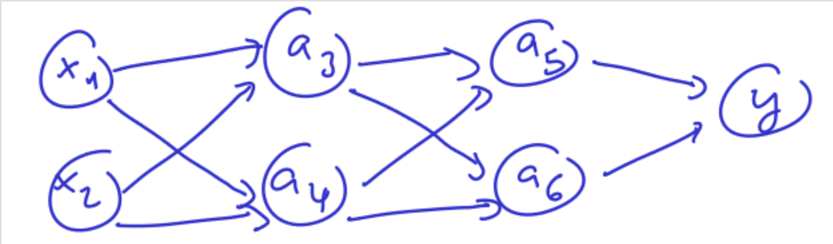
$$a_3 = \tanh(z_3)$$

$$z_4 = x_1 \cdot w_{14} + x_2 \cdot w_{24}$$
$$a_4 = \tanh(z_4)$$

$$a_5 = \tanh(z_5)$$

$$z_6 = a_3 \cdot w_{36} + a_4 \cdot w_{46}$$
$$a_6 = \tanh(z_6)$$

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

$$\delta_6 = \frac{\partial y}{\partial a_6} = 1$$

$$\delta_5 = \frac{\partial y}{\partial a_5} = 1$$

$$\delta_4 = \underbrace{\frac{\partial y}{\partial a_6}}_{\delta_6} \cdot \underbrace{\frac{\partial a_6}{\partial z_6}}_{(1-a_6^2)} \cdot \underbrace{\frac{\partial z_6}{\partial a_4}}_{w_{46}} + \underbrace{\frac{\partial y}{\partial a_5}}_{\delta_5} \cdot \underbrace{\frac{\partial a_5}{\partial z_5}}_{(1-a_5^2)} \cdot \underbrace{\frac{\partial z_5}{\partial a_4}}_{w_{45}}$$

$$\delta_3 = \delta_6 \cdot (1-a_6^2) \cdot w_{36} \qquad \delta_5 \cdot (1-a_5^2) \cdot w_{35}$$

$$\frac{\partial y}{\partial w_{13}} = \underbrace{\frac{\partial y}{\partial a_3}}_{\delta_3} \cdot \underbrace{\frac{\partial a_3}{\partial z_3}}_{(1-a_3^2)} \cdot \underbrace{\frac{\partial z_3}{\partial w_{13}}}_{x_1}$$