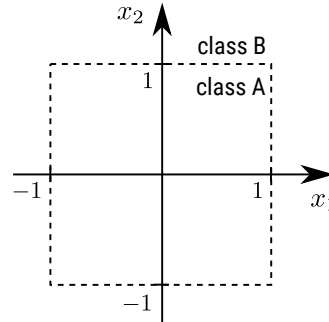


Exercise Sheet 8

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

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

- (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:

$$z_3 = x_1 \cdot w_{13} + x_2 \cdot w_{23}$$

$$a_3 = \tanh(z_3)$$

$$z_4 = x_1 \cdot w_{14} + x_2 \cdot w_{24}$$

$$a_4 = \tanh(z_4)$$

$$z_5 = a_3 \cdot w_{35} + a_4 \cdot w_{45}$$

$$a_5 = \tanh(z_5)$$

$$z_6 = a_3 \cdot w_{36} + a_4 \cdot w_{46}$$

$$a_6 = \tanh(z_6)$$

$$y = a_5 + a_6$$

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

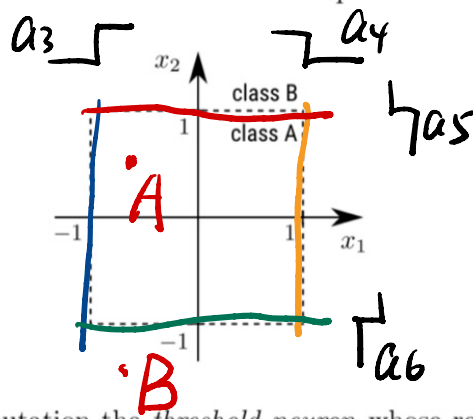
Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.

Exercise Sheet 8

Exercise 1: Designing a Neural Network (25 P)

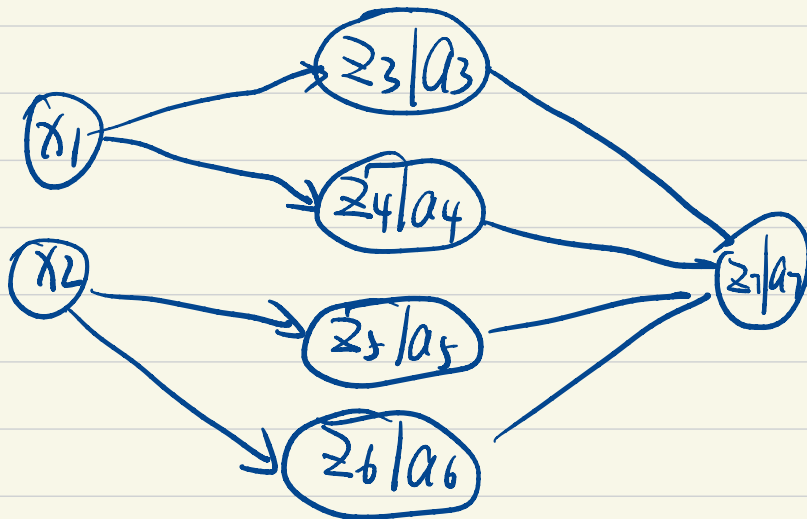
We would like to implement a neural network that classifies data points in \mathbb{R}^2 according to decision boundary given in the figure below.



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

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

- (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.



$$z_3 = w_{13}x_1 + b_3$$

$$0 = (1)(-1) + 1$$

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

$$z_4 = w_{14}x_1 + b_4$$

$$0 = 1 \cdot 1 + (-1)$$

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

$$z_5 = w_{25}x_2 + b_5$$

$$0 = 1 \cdot 1 + (-1)$$

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

$$z_6 = x_2 - w_{26} + b_6 \quad w_{26} = 1$$

$$0 = -1 \cdot 1 + 1 \quad b_6 = 1$$

$$z_7 = w_{37} \cdot a_3 + w_{47} \cdot a_4 + w_{57} \cdot a_5 + w_{67} \cdot a_6 + b_7$$

$$1 - 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + b_7 \geq 0 \quad (\text{case A})$$

$$b_7 \geq -4$$

$$1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 0 + b_7 \leq 0 \quad (\text{case B})$$

$$b_7 \leq -3$$

So $w_{37} = w_{47} = w_{57} = w_{67} = 1 \quad b_7 = -3.3$

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:

$$z_3 = x_1 \cdot w_{13} + x_2 \cdot w_{23}$$

$$a_3 = \tanh(z_3)$$

$$z_4 = x_1 \cdot w_{14} + x_2 \cdot w_{24}$$

$$a_4 = \tanh(z_4)$$

$$z_5 = a_3 \cdot w_{35} + a_4 \cdot w_{45}$$

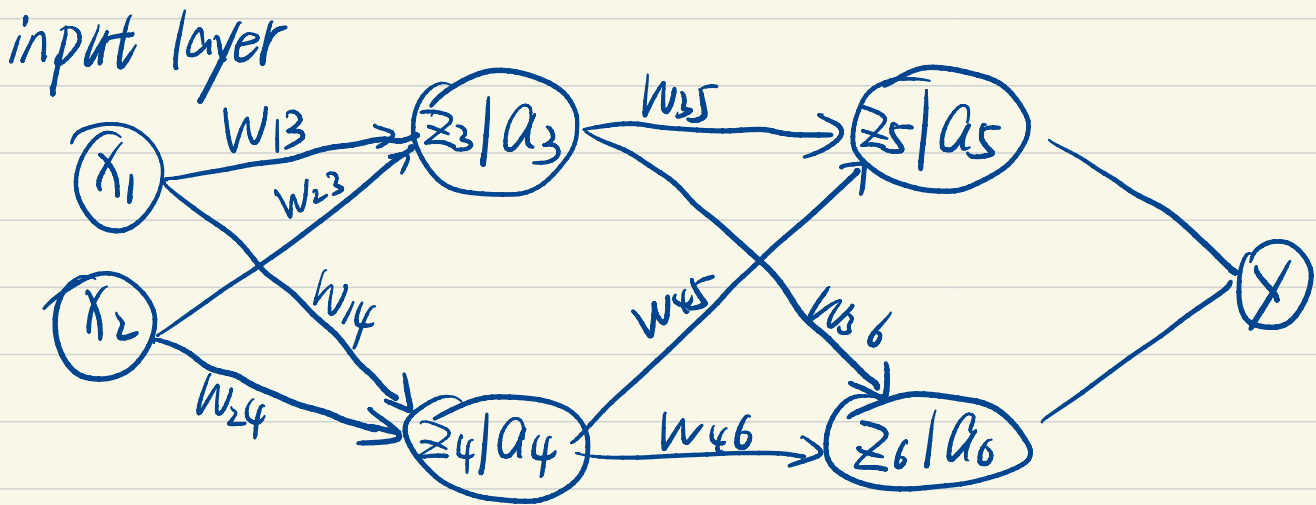
$$a_5 = \tanh(z_5)$$

$$z_6 = a_3 \cdot w_{36} + a_4 \cdot w_{46}$$

$$a_6 = \tanh(z_6)$$

$$y = a_5 + a_6$$

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

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

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