Exercises for the course

Machine Learning 1

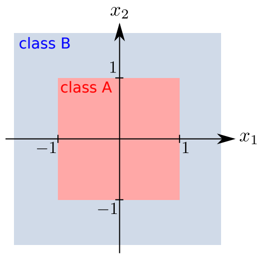
Winter semester 2016/17

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 12

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 the following decision boundary:



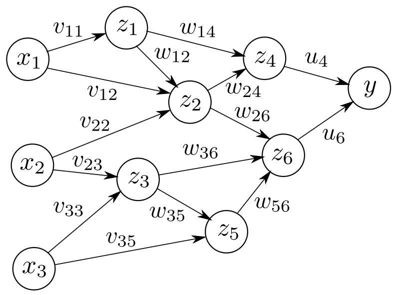
Note that the area for class B stretches to the infinity. We consider as an elementary computation the *threshold neuron* whose relation between inputs $\{z_i\}$ and output z_j is given by

$$z_j = \begin{cases} 1 & \text{if} \quad \sum_i z_i w_{ij} + b_j > 0 \\ 0 & \text{if} \quad \sum_i z_i w_{ij} + b_j \leq 0. \end{cases}$$

In a similar way the XOR problem was solved in the slides, design at hand a neural network that takes x_1 and x_2 as input and produces the output "1" if the input belong to class A, and "0" if the input belongs to class B. *Draw* the neural network model and *write down* its corresponding weight and bias parameters.

Exercise 2: Backward Computations (25 P)

We consider a neural network with the following structure:



The elementary computation of this network is the *sigmoid neuron* defined as:

$$g(a_j) = \frac{e^{a_j}}{1 + e^{a_j}} \qquad \text{where} \quad a_j = \sum_i z_i w_{ij}.$$

Examples of forward computations are:

$$\begin{aligned} z_5 &= g(x_3 \cdot v_{35} + z_3 \cdot w_{35}) \\ z_6 &= g(z_2 \cdot w_{26} + z_3 \cdot w_{36} + z_5 \cdot w_{56}) \\ y &= g(z_4 \cdot u_4 + z_6 \cdot u_6) \end{aligned}$$

Assuming that we have computed the activation of each neuron and the output, and knowing the error gradient $\frac{\partial E}{\partial y}$, *write* the sequence of computations that lead to the evaluation of the error gradient $\frac{\partial E}{\partial w_{12}}$. *Hint:* The derivative of the sigmoid function $g(a_j)$ can be expressed in terms of neuron activations as

$$\frac{\partial}{\partial a_j} \underbrace{g(a_j)}_{z_j} = \underbrace{g(a_j)}_{z_j}(1 - \underbrace{g(a_j)}_{z_j}).$$

Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.