Exercise 3.1: Gradient Descent

(a) Execute four iterations of gradient descent with momentum to find the minimum of the function
\[ f(u) = \frac{u^3}{3} + 50u^2 - 100u - 30. \]
Start with \( u = 20 \), use a learning rate that is set to \( \varepsilon = 0.01 \), and parameter \( \mu \) set to 0.1.

(b) Evaluate the benefit of using a momentum for the task in (a) by comparing your findings to the vanilla gradient descent method.

Exercise 3.2: Forward-Backward-Pass

Examine the multi-layer perceptron given below.

(a) Both neurons use the logistic activation function \( u \mapsto \frac{1}{1 + e^{-u}} \). The network has a single input variable \( x \) and one output variable \( y \). Calculate the output of both neurons and the error made by the MLP when applying a pattern with \( x = 0 \) and target value 0.5.

(b) Calculate the partial derivatives of the error with respect to the weights for the pattern used in task (a).

Exercise 3.3: Multi-Layer Perceptron Architecture

Now, consider the network structure of a multi-layer perceptron with 10 neurons given in Figure 1. Each circle denotes a neuron, the arrows denote connections between neurons.

(a) Which of the neurons are input neurons, which ones are output neurons?
(b) How many layers does this MLP have? Which neurons belong to which layer?

(c) Assume we are applying a pattern to the MLP. Give an order in which the neuron activations $a_i$ can be calculated.

**Exercise 3.4: Multi-Layer Perceptrons**

Which of the functions given in these plots can be implemented by multi-layer perceptrons? The MLP should only contain neurons with logistic activation functions. (*Note:* The weights of the networks must be finite numbers.)