# Problem Set 4 - Gradients and Backpropagation

## DS542 - DL4DS

### Spring, 2025

**Note:** Refer to Chapter 7 in *Understanding Deep Learning.*

## Problem 4.1 (3 points)

Consider the case where we use the logistic sigmoid function as an activation function, defined as:

$$h = \sigma(z) = \frac{1}{1 + e^{-z}}. \tag{1}$$

Compute the derivative $\frac{\partial h}{\partial z}$. What happens to the derivative when the input takes (i) a large positive value and (ii) a large negative value?

## Problem 4.2 (3 points)

Calculate the derivative $\frac{\partial \ell_i}{\partial f[x_i, \phi]}$ for the binary classification loss function:

$$\ell_i = -(1 - y_i) \log[1 - \sigma(f[x_i, \phi])] - y_i \log[\sigma(f[x_i, \phi])], \tag{2}$$

where the function $\sigma(\cdot)$ is the logistic sigmoid, defined as:

$$\sigma(z) = \frac{1}{1 + \exp(-z)}. \tag{3}$$