Metropolia Ammattikorkeakoulu
Mathematics and Methods in Machine Learning and Neural Networks
Mathematics / Home Exercises 4

1. a) Let $f_i(x) = w_i x + b_i$ with $i = 1, 2, 3$, where w_i and b_i are constants. Calculate
$$\frac{d}{dx} f_3(f_2(f_1(x))).$$
b) (*Optional*) Let f_i be as above and let
$$g(x) = \frac{1}{1 + e^{-x}}.$$
The derivative of the function $g(x)$ can be written as
$$g'(x) = g(x)\left(1 - g(x)\right).$$
If we denote
$$z_1 = f_1(x), \quad a_1 = g(z_1), \quad z_2 = f_2(a_1), \quad a_2 = g(z_2)$$
what is the derivative
$$\frac{d}{dx} g(f_2(g(f_1(x))))?$$

2. Let $\alpha = 2$, $X_0 = (2, -1)$, and $X_i = (x_i, y_i)$. Using
$$X_{i+1} = X_i - \alpha \nabla f(X_i)$$
calculate X_1 and X_2 for the functions
a) $f(x, y) = x^2 y + 3y$, b) $f(x, y) = 3x^2 + xy + y^2 - 4y$.

Answer: a) $X_2 = (610, -221)$, b) $X_2 = (206, 27)$

3. (*Python Exercise*) Fit a logistic regression model
$$\hat{y} = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x)}}$$
to the data

x	1.1	4.1	2.7	3.9	7.1	1.3	9.7	0.7	3.4	6.0
y	0	1	0	0	1	0	1	0	1	0

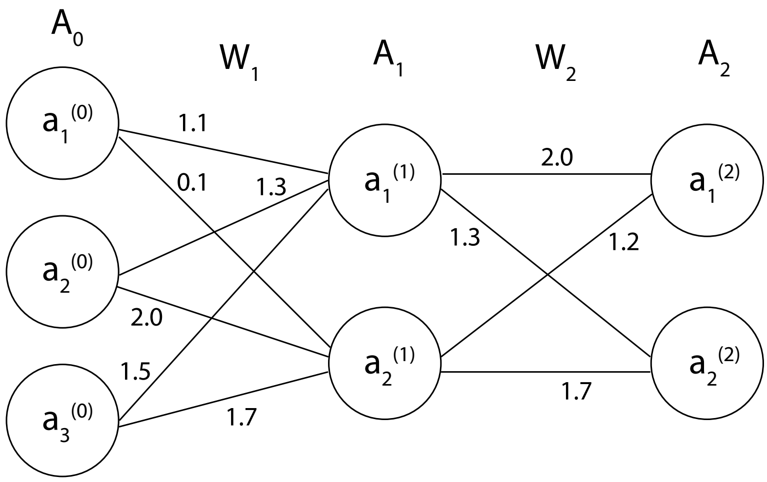
using the gradient descent method.

Answer: $\theta_0 = -3.12097227$, $\theta_1 = 0.66032075$

4. We have the following small network, where the weights are written beside the edges, and there are no bias terms. The values of the cells are stored in matrices A_0, A_1 and A_2, where

$$A_0 = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}, \qquad A_1 = W_1 A_0, \qquad A_2 = W_2 A_1.$$

a) What are the matrices W_1 and W_2? Calculate the matrices A_1 and A_2 by hand or using Python.



b) (*Optional*) If you are using Python, calculate the matrices A_1 and A_2 using the formulae

$$A_1 = g(W_1 A_0) \quad \text{and} \quad A_2 = g(W_2 A_1),$$

where

$$g(x) = \frac{1}{1 + e^{-x}}.$$

Answer: a) $A_2 = \begin{pmatrix} 4.16 \\ 39.18 \end{pmatrix}$, b) $A_2 = \begin{pmatrix} 0.96083404 \\ 0.95257383 \end{pmatrix}$