COMP 4107: Neural Networks – Assignment 1
– Winter 2026 –

Due: January 28, 2026

Group 51
Andrew Wallace - 101210291
Christer Henrysson - 101260693
Amr Altaweel - 101276934

Getting started

Instructions:

To get setup run the following commands on Windows

```
python3 -m venv .venv
source .venv/bin/activate
pip install -r requirements.txt
```

To get setup run the following commands on Linux

```
python3 -m venv .venv
source .venv/Scripts/activate
pip install -r requirements.txt
```

Question 1 [10 marks]

Please see `assignment1.py` for implementation of `artificial_neuron`.

Question 2 [10 marks]

Please see `assignment1.py` for implementation of `gradient_descent`.

Question 3 [10 marks]

Please see `assignment1.py` for implementation of `pytorch_module`.

Question 4 [10 marks]

Please see `assignment1.py` for full implementation details.

(a) Gradient Descent I
We are given the function:

$$f(a,b) = \frac{1}{2}((ax_1 + b - y_1)^2 + (ax_2 + b - y_2)^2) \tag{1}$$

Assuming x_1, x_2, y_1, y_2 are constants, the derivative (gradient) of f is as follows:

$$\nabla f(a,b) = \begin{bmatrix} \frac{\partial f}{\partial a} \\ \frac{\partial f}{\partial b} \end{bmatrix}$$

$$\begin{aligned}
\frac{\partial f}{\partial a} &= \frac{1}{2}\frac{\partial f}{\partial a}(ax_1 + b - y_1)^2 + \frac{1}{2}\frac{\partial f}{\partial a}(ax_2 + b - y_2)^2 \\
\frac{\partial f}{\partial a} &= x_1(ax_1 + b - y_1) + x_2(ax_2 + b - y_2) \\
\frac{\partial f}{\partial b} &= \frac{1}{2}\frac{\partial f}{\partial b}(ax_1 + b - y_1)^2 + \frac{1}{2}\frac{\partial f}{\partial b}(ax_2 + b - y_2)^2 \\
\frac{\partial f}{\partial b} &= (ax_1 + b - y_1) + (ax_2 + b - y_2) \\
\nabla f(a,b) &= \begin{bmatrix} x_1(ax_1 + b - y_1) + x_2(ax_2 + b - y_2) \\ (ax_1 + b - y_1) + (ax_2 + b - y_2) \end{bmatrix}
\end{aligned}$$

Below are the parameters and results from our experiment:

- $\nabla f(a,b) = \begin{bmatrix} x_1(ax_1 + b - y_1) + x_2(ax_2 + b - y_2) \\ (ax_1 + b - y_1) + (ax_2 + b - y_2) \end{bmatrix}$

- $x_1 = 3$
- $x_2 = -2$
- $y_1 = 0.5$
- $y_2 = -0.75$
- Initial guess $= x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Learning rate = alpha $= \alpha = 0.05$
- Value of a at minimum $= 0.2499956502209206$
- Value of b at minimum $= -0.249951760210792$
- Minimum value for $f = 2.2402285941984945 \times 10^{-9}$

(b) Gradient Descent II

We are given the function:

$$f(a, b) = \frac{1}{2}((\mathrm{SiLU}(ax_1 + b) - y_1)^2 + (\mathrm{SiLU}(ax_2 + b) - y_2)^2) \tag{2}$$

Where

$$\mathrm{SiLU}(x) = \frac{x}{1 + \exp(-x)}$$

Assuming x_1, x_2, y_1, y_2 are constants, the derivative (gradient) of f is as follows:

$$\nabla f(a, b) = \begin{bmatrix} \frac{\partial f}{\partial a} \\ \frac{\partial f}{\partial b} \end{bmatrix}$$

We define the following for $i \in \{1, 2\}$:

$$z_i = ax_i + b$$

$$e_i = \mathrm{SiLU}(z_i) - y_i$$

We calculate the partial derivate with respect to a and the partial derivate with respect b as:

$$\frac{\partial f}{\partial a} = e_1 \mathrm{SiLU}'(z_1) x_1 + e_2 \mathrm{SiLU}'(z_2) x_2$$

$$\frac{\partial f}{\partial b} = e_1 \mathrm{SiLU}'(z_1) + e_2 \mathrm{SiLU}'(z_2)$$

So, the gradient of f is

$$\nabla f(a, b) = \begin{bmatrix} e_1 \mathrm{SiLU}'(z_1) x_1 + e_2 \mathrm{SiLU}'(z_2) x_2 \\ e_1 \mathrm{SiLU}'(z_1) + e_2 \mathrm{SiLU}'(z_2) \end{bmatrix}$$

This gives us the following parameters and results of the gradient descent:

- $\nabla f(a, b) = \begin{bmatrix} e_1 \text{SiLU}'(z_1)x_1 + e_2 \text{SiLU}'(z_2)x_2 \\ e_1 \text{SiLU}'(z_1) + e_2 \text{SiLU}'(z_2) \end{bmatrix}$
- $x_1 = 3$
- $x_2 = -2$
- $y_1 = 0.5$
- $y_2 = -0.75$
- Initial guess $= x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Learning rate = alpha $= \alpha = 0.05$
- Value of a at minimum $= 0.4033468264353714$
- Value of b at minimum $= -0.47115909972569386$
- Minimum value for $f = 0.11117286369775675$