

# Submission for Deep Learning Exercise 1

Team: shallow\_learning\_group  
Students: Batuhan Karaca

November 3, 2023

## Pen and Paper task: Forward Propagation

$$X = \begin{bmatrix} x_1^T \\ x_2^T \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$X^{(1)} = \text{ReLU}(XW^{(1)} + \mathbf{1}b^{(1)T})$$

$$= \text{ReLU}\left(\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 2 & 0 \\ -3 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \end{bmatrix}\right)$$

$$= \text{ReLU}\left(\begin{bmatrix} -7 & 4 \\ -13 & 8 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 3 & 0 \end{bmatrix}\right)$$

$$= \text{ReLU}\left(\begin{bmatrix} -4 & 4 \\ -10 & 8 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 0 & 4 \\ 0 & 8 \end{bmatrix}$$

$$\bar{y} = \text{Sigmoid}(X^{(1)}W^{(2)} + \mathbf{1}b^{(2)T})$$

$$= \text{Sigmoid}\left(\begin{bmatrix} 0 & 4 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} (-3)\right)$$

$$= \text{Sigmoid}\left(\begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} -3 \\ -3 \end{bmatrix}\right)$$

$$= \text{Sigmoid}\left(\begin{bmatrix} 1 \\ 5 \end{bmatrix}\right)$$

$$\sim \begin{bmatrix} 0.7311 \\ 0.9933 \end{bmatrix}$$

$$\mathcal{L}(\bar{y}, y) = -y \log \bar{y} - (1 - y) \log(1 - \bar{y})$$

$$= -\begin{bmatrix} 0 \\ 1 \end{bmatrix} \log \begin{bmatrix} 0.7311 \\ 0.9933 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \log \begin{bmatrix} 0.2689 \\ 0.0067 \end{bmatrix}$$

$$= -\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -0.3132 \\ -0.0067 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1.3134 \\ -5.0056 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0.0067 \end{bmatrix} + \begin{bmatrix} 1.3134 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1.3134 \\ 0.0067 \end{bmatrix}$$

$$(1.3134 + 0.0067)/2 \sim 0.66$$

**Question: What is the best accuracy you can achieve in practice using Logistic Regression?**

Because both the logistic regression can create non-linear mappings (due to sigmoid) and XOR function are non-linear, the accuracy should be 100% for a well-trained classifier.