Exercise 02

NAME:- ARNAV SINGHAL
UNI-ID - b0303

1. Hand-Crafted Network

- logical OR map a binary input vector $z \in \{0,1\}^D$ to

$$z \rightarrow f(z) = \begin{cases} 1 & \exists j \text{ such that } z_j = 1 \\ 0 & \text{otherwise i.e } z = 0 \end{cases}$$

Weights: $w = [1, 1]$
Bias: $b = -0.5$
Activation function: step function (Threshold at 0)

Output

$$f(z) = \text{step}(z^* w + b)$$
$$= \text{step}(z[0] * w[0] + z[1]^* w[1] + b)$$
$$= \text{step}(z[0] + z[1] - 0.5)$$

- Masked logical OR: for an arbitrary but fixed binary vector $c \in \{0,1\}^D$ map the input vector $z \in \{0,1\}^D$ to

$$z \rightarrow g(z; c) = \begin{cases} 1 & \exists j \text{ such that } c_j = 1 \text{ and } z_j = 1 \\ 0 & \text{otherwise, i.e } z = 0 \text{ or } z_j = 1 \text{ occurs only at indices where } c_j = 0 \end{cases}$$

weights: $w = [1, 1]$
Bias: $b = -0.5$
Activation function: step function (Threshold at 0)

Output

$$g(z:c) = \text{step}((z^* w + b)^* (1 - c))$$
$$= \text{step}((z[0]^* w[0] + z[1]^* w[1] + b)^* (1 - c))$$
$$= \text{step}((z[0] + z[1] - 0.5)^* (1 - C))$$

• Perfect Match for an arbitrary but fixed binary vector $c \in \{0,1\}^D$ map the input vector

$$z \rightarrow h(z;c) = \begin{cases} 1 & z = c \\ 0 & \text{otherwise} \end{cases}$$

Weights : $w = [1,1]$

Bias : $b = -1.5$

Activation function : Step function (Threshold at 0)

<u>Output</u>

$$h(z;c) = \text{step}\,(z^* w + b - \|c\|^2)$$

$$= \text{Step}\,(z[0]^* w[0] + z[1]^* w[1] + b - \|c\|^2)$$

$$= \text{Step}\,(z[0] + z[1] - 1.5 - \|c\|^2)$$

$\|c\|^2 \rightarrow$ squared euclidean of vector c.

2. Linear Activation Function

To prove that if the activation function ϕ_1 is the identity function for all layers 1, then any network with depth $L>1$ is ~~equ~~ equivalent to a 1-layer neural network, we need to show that the output of the network can be expressed as a linear transformation.

Let's consider, a network with L layers. the output of the first layer can be calculated as :-

$$Z1 = \phi 1(z^{\sim 1}) = \phi 1(X \cdot B1 + b1) = X \cdot B1 + b1$$

since $\phi 1$ is the identity function, Z1 is a linear transformation of the input X.

Assuming that the output of the $(1-1)^{th}$ layer is also a linear transformation.

$$Z1-1 = X \cdot B1-1 + b1-1$$

$$z^{\sim 1} = Z1-1 \cdot B1 + b1 = (X \cdot B1-1 + b1-1) \cdot B1 + b1$$

$$Z1 = \phi_1(z^{\sim 1}) = (X \cdot B1-1 + b1-1) \cdot B1 + b1$$

$$= X \cdot (B1-1 \cdot B1) + (b1-1 \cdot B1 + b1)$$

Z1 can be expressed as linear transformation of input X, with new weights and bias by $(B1-1 \cdot B1)$ and $(b1-1 \cdot B1 + b1)$ respectively.

Iterating process for all layers $(1 \dots L)$

$$Z_L = X \cdot (B_1 \cdot B_2 \dots B_{L-1} \cdot B_L) + (b_1 \cdot B_2 \dots B_{L-1} \cdot B_L + b_2 \cdot B_3 \dots + b_{l-1} \cdot B_L + b_l)$$

Thus, we have demonstrated that if the activation function is the identity function for all layers, any network with depth $L>1$ is equal to a 1-layer neural network.