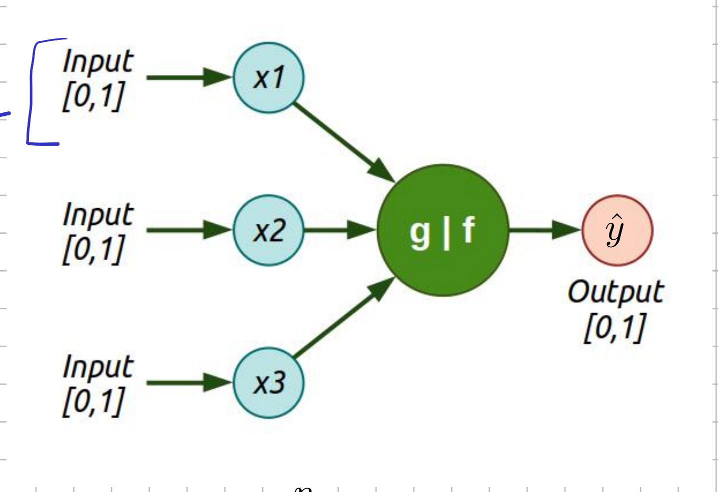
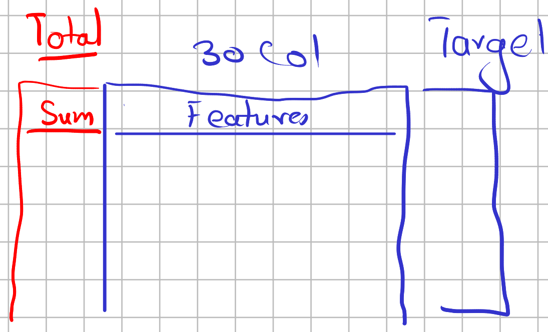
# MP Neurons

0,1




$$z = g(x) = \sum_{i=1}^{n} (x_i)$$

$$\hat{y} = a = f(z)$$

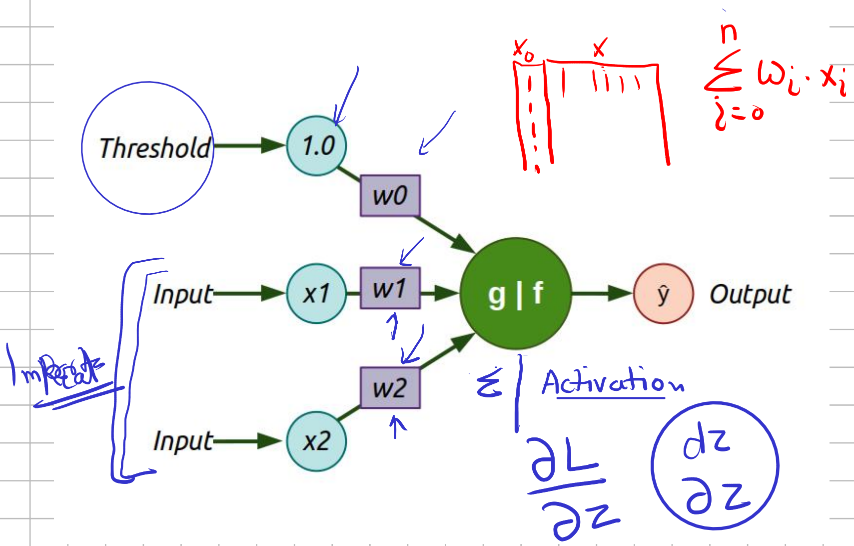
$$\hat{y} = a = 1 \text{ if z } \geq 0 \text{ else } 0$$

Rain – 0-20mm

Normalizing Rent Cost Ground – 500 ⟶ 3000

Home work –

# Perceptrons




## Forward propagation

$$z = g(x) = w_0 + \sum_{i=1}^{n}(w_i * x_i)$$

$$\hat{y} = a = f(z)$$

$$\hat{y} = a = 1 \text{ if z } \geq 0 \text{ else } 0$$

## Back Propagation

$$\partial z = a - y$$

$$\partial w_0 = \partial z$$

$$\partial w_1 = x_1 \circ \partial z$$

$$\partial w_2 = x_2 \circ \partial z$$

$z = \omega_0 + \sum X_i \cdot W_i$

Loss $= -[y.\log(a) + (1-y).\log(1-a)]$

$\frac{-y}{a} + \frac{(1-y)}{1-a}$ — (1)

$\frac{\partial L}{\partial z} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z}$ (2)

$a.(1-a)$

## Updates

Learning Rate [0.1 → 0.001]

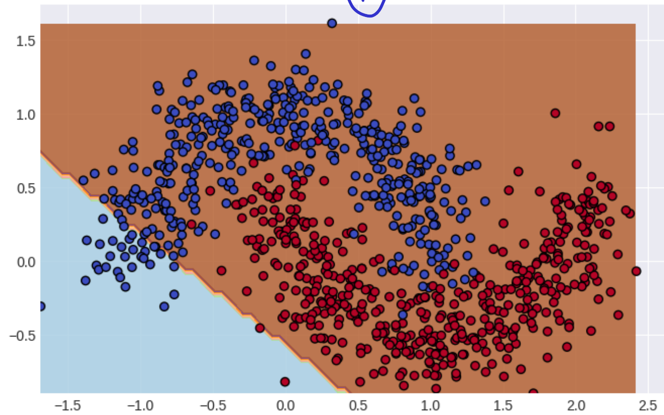
$$w_0 = w_0 - \frac{1}{m}(\alpha \circ \partial w_0)$$

\# of Rows

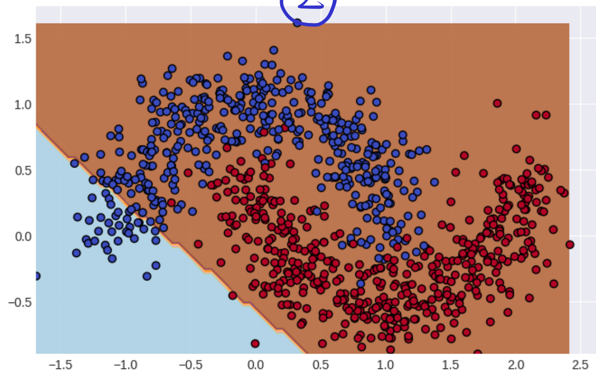
$$w_1 = w_1 - \frac{1}{m}(\alpha \circ \partial w_1)$$

$$w_2 = w_2 - \frac{1}{m}(\alpha \circ \partial w_2)$$

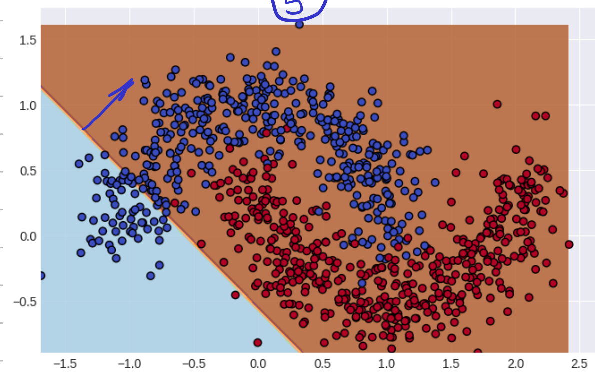
1


1.5
1.0
0.5
0.0
−0.5
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5

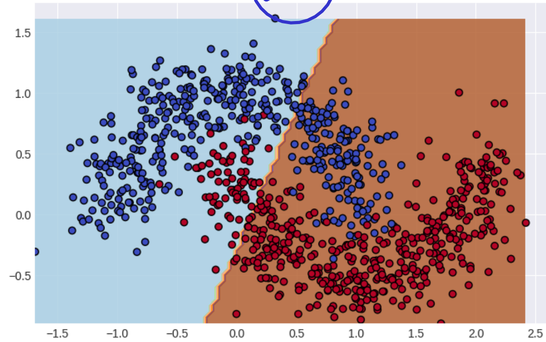
2

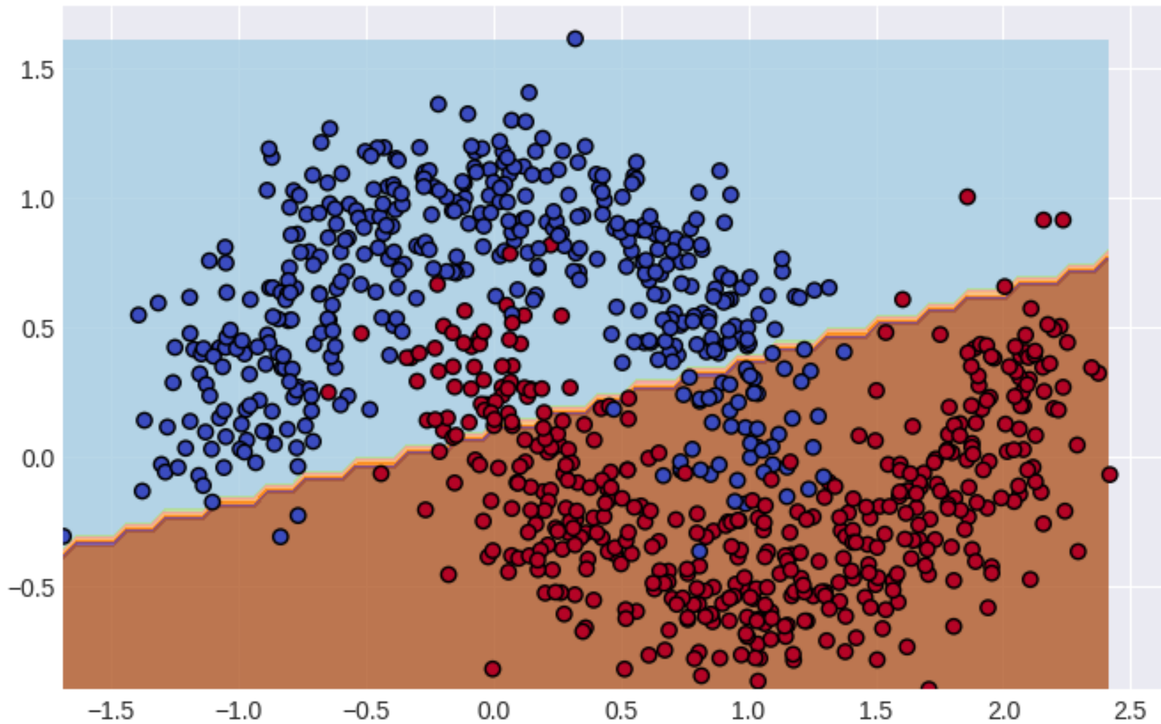

1.5
1.0
0.5
0.0
−0.5
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5

5


1.5
1.0
0.5
0.0
−0.5
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5

50

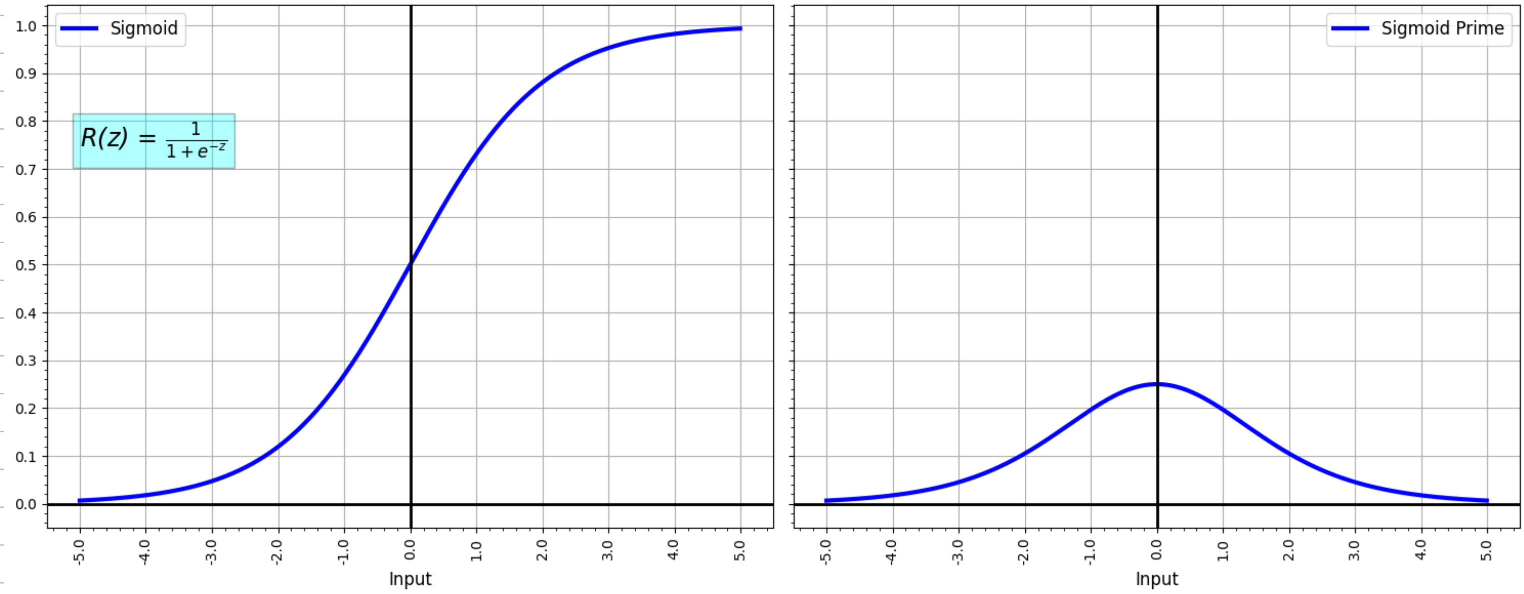

1.5
1.0
0.5
0.0
−0.5
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5

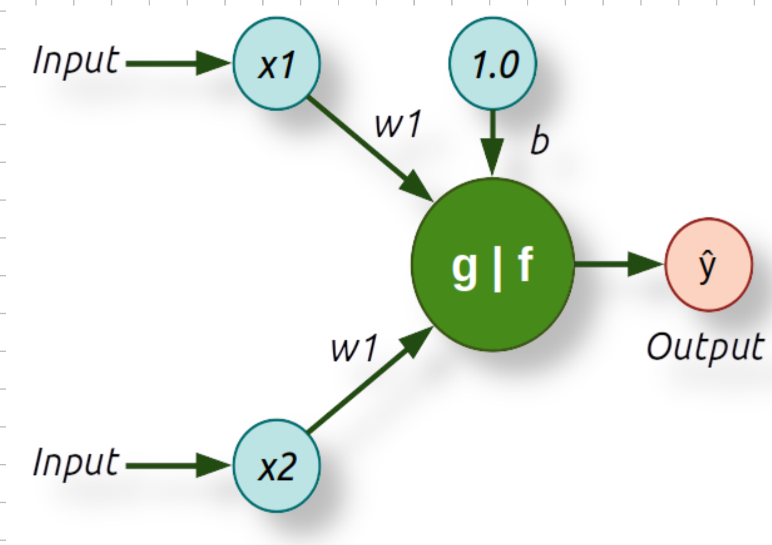

1.5
1.0
0.5
0.0
−0.5
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5 2.0 2.5

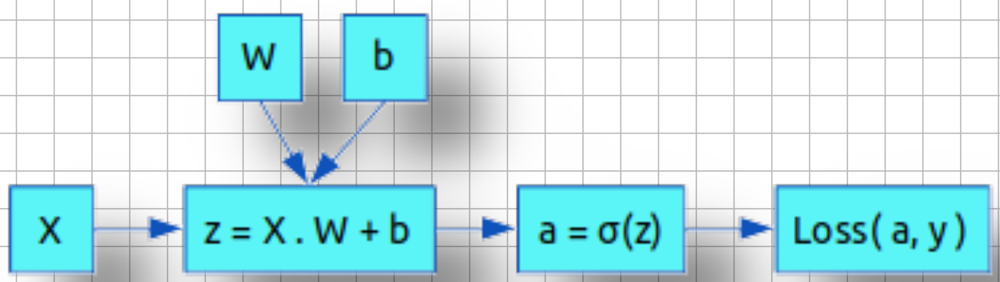
# Assignment A02

Changes :- Sigmoid Activation function

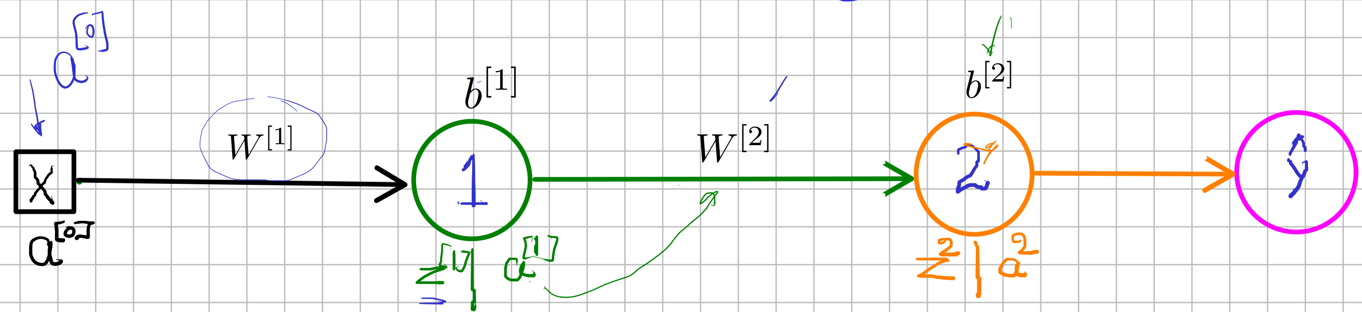
: Loss = row[-1]. Log(a)



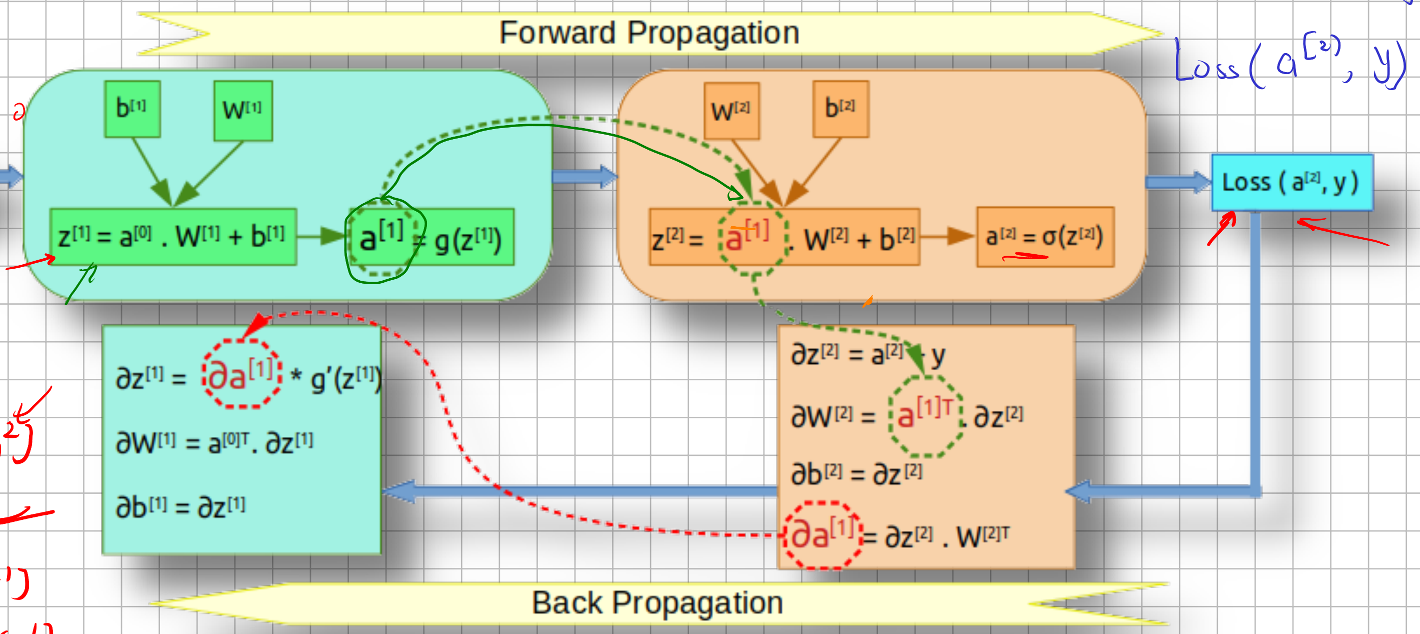



# One Hidden Layer




$z_1 = x \cdot w^{[1]} + b^{[1]}$ $\quad a^{[1]} = \tanh(z^{[1]})$ $\quad z^{[2]} = a^{[1]} \cdot w^{[2]} + b^{[2]}$ $\quad a^{[2]} = \sigma(z^{[2]})$




$\frac{\partial L}{\partial z^{[1]}}$

$L = f_1(a^2)$

$a^2 = f_2(z^2)$

$z^2 = f_3(a^1)$

$a^1 = f_4(z^1)$

$$\frac{\partial L}{\partial z^{[1]}} = \frac{\partial L}{\partial a^{[2]}} \cdot \frac{\partial a^{[2]}}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial a^{[1]}} \cdot \frac{\partial a^{[1]}}{\partial z^{[1]}}$$

$$= \frac{\partial L}{a^{[2]}} \frac{\partial a^{[2]}}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{a^{[1]}} \cdot \frac{a^{[1]}}{\partial z^{[1]}}$$

$dz^{[2]} = a^{[2]} - y$

$\sigma \; db^{[2]} = dz^{[2]}$

$dw^{[2]} = a^{[1]} \cdot dz^{[2]}$

$da^{[1]} = w^{[2]} \cdot dz^{[2]}$

For single perceptron:

$$z = x_1.w_1 + x_2.w_2 + b \quad \text{— } 1$$

$$a = \sigma(z) \quad \text{— } 1$$

$$a = \sigma(x_1.w_1 + x_2.w_2 + b)$$

$$a = \sigma\ \left([x_1, x_2] \circ \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + b\right)$$

For multiple Rows of X:

$$a = \sigma\left(\begin{bmatrix} x_1^{(1)} & x_2^{(1)} \\ x_1^{(2)} & x_2^{(2)} \\ x_1^{(\dots)} & x_2^{(\dots)} \\ x_1^{(m)} & x_2^{(m)} \end{bmatrix} \circ \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + b\right)$$

Scaler

In matrix form it can be represented as:

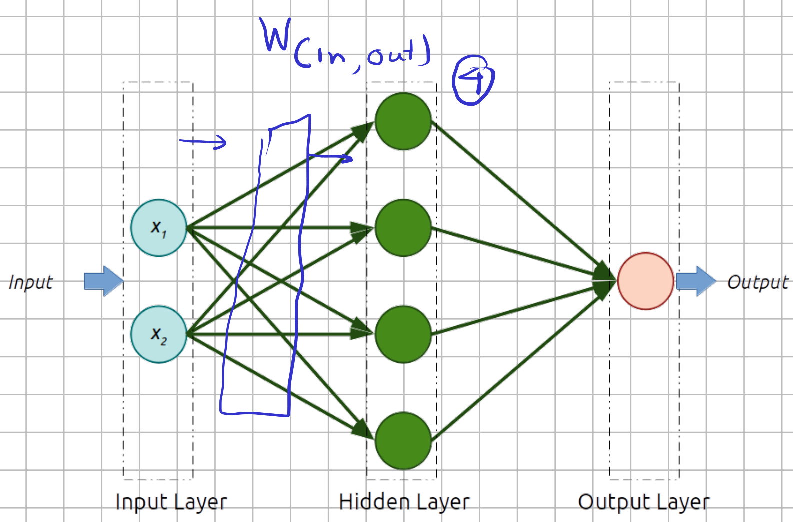
$$a = \sigma\ (X_{shape=(m,2)} \circ W^{[1]}_{shape=(2,1)} + b_{shape=(1,1)})$$

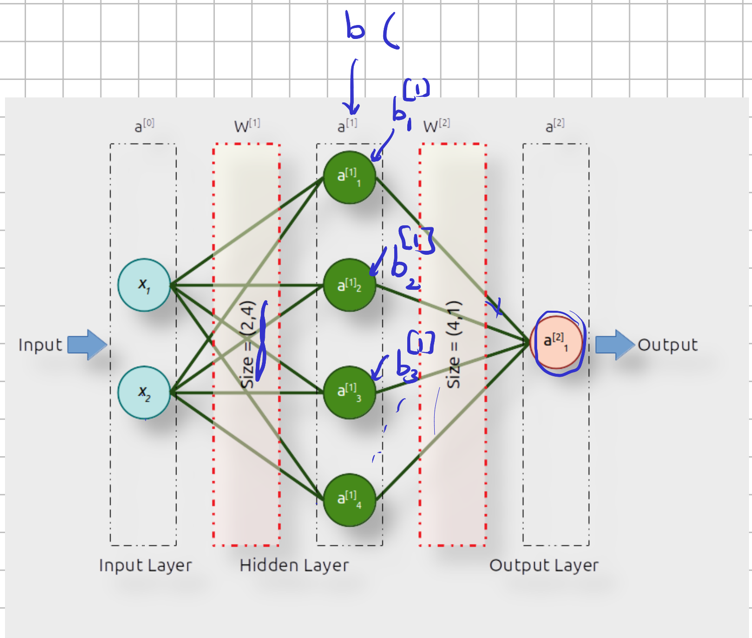
For single row of perceptrons:

$$a^{[1]} = g(X_{shape=(m,2)} \circ W^{[1]}_{shape=(2,1)} + b^{[1]}_{shape=(1,1)}) \quad 1$$

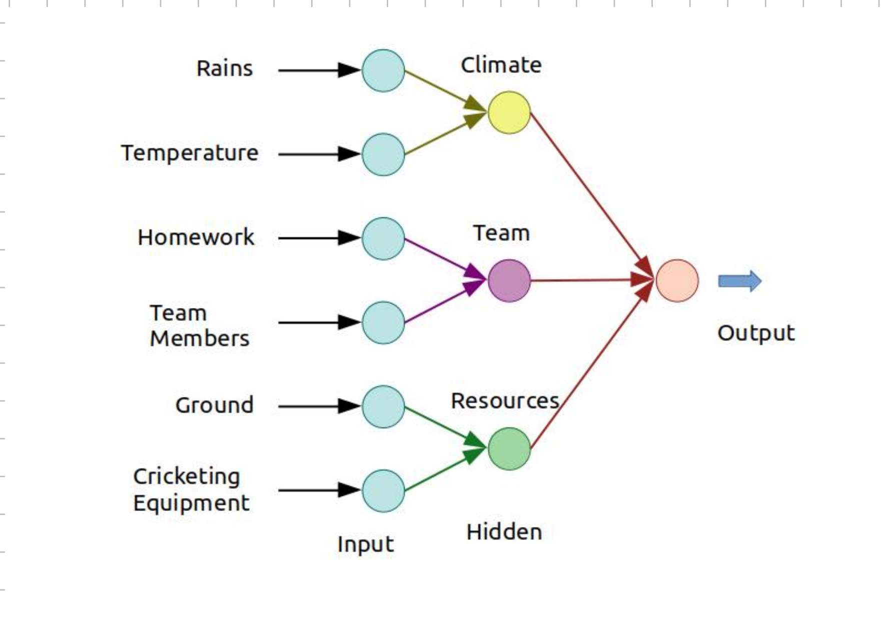
$$a^{[2]} = \sigma(a^{[1]}_{shape=(m,1)} \circ W^{[2]}_{shape=(1,1)} + b^{[2]}_{shape=(1,1)}) \quad 2$$

# Neural Network



# What is this hidden layer?




$$z_1^{[1]} = X.W_1^{[1]} + b_1^{[1]}$$
$$a_1^{[1]} = \tanh(z_1^{[1]})$$

1

$$z_2^{[1]} = X.W_2^{[1]} + b_2^{[1]}$$
$$a_2^{[1]} = \tanh(z_2^{[1]})$$

2

$$z_3^{[1]} = X.W_3^{[1]} + b_3^{[1]}$$
$$a_3^{[1]} = \tanh(z_3^{[1]})$$

3

$$z_4^{[1]} = X.W_4^{[1]} + b_4^{[1]}$$
$$a_4^{[1]} = \tanh(z_4^{[1]})$$

4

Or

$$a^{[1]} = \tanh(X \circ \left[W_1^{[1]}, \quad W_2^{[1]}, \quad W_3^{[1]}, \quad W_4^{[1]}\right] + \left[b_1^{[1]}, \quad b_2^{[1]}, \quad b_3^{[1]}, \quad b_4^{[1]}\right])$$

$$\boxed{a^{[1]} = \tanh(X \circ W^{[1]}) + b^{[1]})}$$

Where:

$$W^{[1]} = \left[W_1^{[1]}, \quad W_2^{[1]}, \quad W_3^{[1]}, \quad W_4^{[1]}\right]$$
$$b^{[1]} = \left[b_1^{[1]}, \quad b_2^{[1]}, \quad b_3^{[1]}, \quad b_4^{[1]}\right]$$

Hence:

$$z^{[1]}_{shape=(m,4)} = X_{shape=(m,2)} \circ W^{[1]}_{shape=(2,4)} + b^{[1]}_{shape=(1,4)}$$

$$a^{[1]}_{shape=(m,4)} = \tanh(z^{[1]})$$

Similarly for second layer.

$$z^{[2]}_{shape=(m,1)} = a^{[1]}_{shape=(m,4)} \circ W^{[2]}_{shape=(4,1)} + b^{[2]}_{shape=(1,1)}$$

$$a^{[2]}_{shape=(m,1)} = \hat{y} = \text{sigmoid}(z^{[2]})$$

Where:

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

## Loss Function

$$L(a_i, y_i) = -\left[y_i.log(a_i) + (1 - y_i).log(1 - a_i)\right]$$

$$J(a, y) = -\frac{1}{m}\sum_{i=1}^{m}\left[y_i.log(a_i) + (1 - y_i).log(1 - a_i)\right]$$

p = 0.6  0.4

H  T

1  0

In case of Binary Classification:

$$L(a_i, y_i) = -y_i.log(a_i)$$

$$J(a, y) = -\frac{1}{m}\sum_{i=1}^{m} y_i.log(a_i)$$

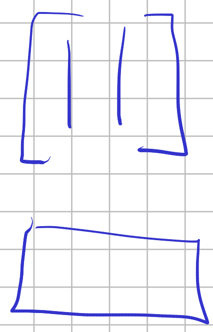
# Backpropagation

$$Loss = f_1(a^{[2]})$$
$$a^{[2]} = f_2(z^{[2]})$$
$$z^{[2]} = f_3(a^{[1]})$$
$$a^{[1]} = f_4(z^{[1]})$$
$$z^{[1]} = f_5(X)$$

Therefore:

$$\boxed{\frac{\partial Loss}{\partial z^{[1]}} = \frac{\partial Loss}{\partial z^{[2]}} \circ \frac{\partial z^{[2]}}{\partial a^{[1]}} \circ \frac{\partial a^{[1]}}{\partial z^{[1]}}}$$

4, 2



(2,

(1, 2)

Back-propagation for all Rows

$$\partial z^{[2]} = a^{[2]} - y$$

$$\partial W^{[2]} = \frac{1}{m} a^{[1]T} \circ \partial z^{[2]}$$

$$\partial b^{[2]} = \frac{1}{m} \text{np.sum}(\partial z^{[2]}, axis = 0, keepdims = True)$$

$$\underline{\partial a^{[1]} = \partial z^{[2]} \circ W^{[2]T}}$$

$$\partial z^{[1]} = \partial a^{[1]} * (1 - tanh(z^{[1]})^2)$$

$$\partial W^{[1]} = \frac{1}{m} X^T \circ \partial z^{[1]}$$

$$\partial b^{[1]} = \frac{1}{m} \text{np.sum}(\partial z^{[1]}, axis = 0, keepdims = True)$$