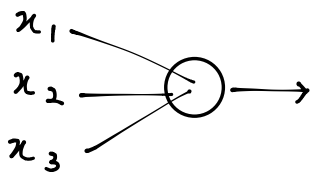
Sid

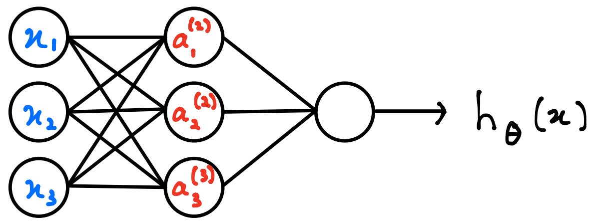
Neuron model: Logistic unit



$$h_\theta(x) = \underbrace{\quad}_{\frac{1}{1+e^{-\theta^T x}}}$$

sigmoid (logistic) activation function

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



Layer 1 (input) Layer 2 (hidden) Layer 3 (output)

a_i^j : "activation" of unit i in layer j

(parameters)

$\theta^{(j)}$: matrix of weights controlling function mapping from layer j to layer j+1

Network w s_j units in layer j & s_{j+1} units in layer j+1

$$\Rightarrow \theta^{(j)} = [s_{j+1} \times (s_j + 1)] \;,\quad \theta^1 = 3 \times 4$$

$$h_\theta(x) = a_1^{(3)} = g(\theta_{10}^{(2)} a_0^{(2)} + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)} + \theta_{13}^{(2)} a_3^{(2)})$$

$$a_1^{(2)} = g(\underline{\theta_{10}^{(1)} x_0 + \theta_{11}^{(1)} x_1 + \theta_{12}^{(1)} x_2 + \theta_{13}^{(1)} x_3}) \rightarrow a_1^{(2)} = g(z_1^{(2)})$$

$$a_2^{(2)} = g(\underline{\theta_{20}^{(1)} x_0 + \theta_{21}^{(1)} x_1 + \theta_{22}^{(1)} x_2 + \theta_{23}^{(1)} x_3}) \rightarrow z_2^{(2)}$$

$$a_3^{(2)} = g(\underline{\theta_{30}^{(1)} x_0 + \theta_{31}^{(1)} x_1 + \theta_{32}^{(1)} x_2 + \theta_{33}^{(1)} x_3}) \rightarrow z_3^{(2)}$$

$$\vec{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad z^{(2)} = \begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \\ z_3^{(2)} \end{bmatrix}$$

$$z^{(2)} = \theta^{(1)} \underline{x} \quad (a^{(1)})$$

$$a^{(2)} = g.(z^{(2)})$$

Add: $a_0^{(2)} = 1$

$$z^{(3)} = \theta^{(2)} a^{(2)}$$

$$h_\theta(x) = a^{(3)} = g(z^{(3)})$$

↑ Foreward Propagation ↑

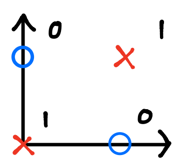
General Vectorized Implementation

$$z^{(j)} = \theta^{(j-1)} a^{(j-1)}$$

$$a^{(j)} = g(z^{(j)})$$

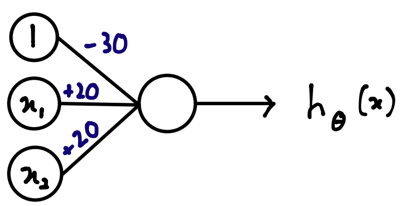
$$z^{(j+1)} = \theta^{(j)} a^{(j)} \implies h_\theta(x) = a^{(j+1)} = g(z^{(j+1)})$$

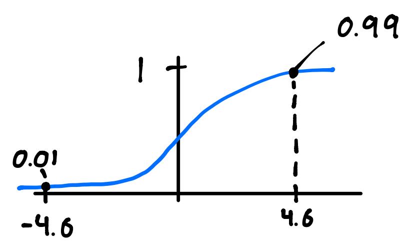
$x_1, x_2 \in \{0, 1\}$



$y = x_1 \text{ XOR } x_2$
(both true or false)

AND : $x_1, x_2 \in \{0, 1\}, y = x_1 \& x_2$





$$h_\theta(x) = g(\underbrace{-30}_{\theta^{(1)}_{10}} + \underbrace{20}_{\theta^{(1)}_{11}} x_1 + \underbrace{20}_{\theta^{(1)}_{12}} x_2)$$

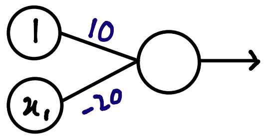
x_1	x_2	$h_\theta(x)$
0	0	$g(-30) \approx 0$
0	1	$g(-10) \approx 0$
1	0	$g(-10) \approx 0$
1	1	$g(10) \approx 1$

And

$$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \to [g(z^{(2)})] \to h_\theta(x)$$

$$\theta^{(1)} = [-30, 20, 20], \quad z^{(2)} = \theta^{(1)} \cdot x^{(1)} = g(-30 + 20x_1 + 20x_2)$$

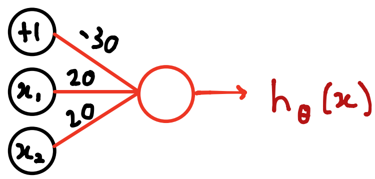
Negation : Not x_1



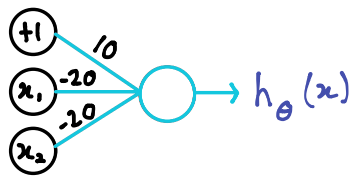
$h_\theta(x) = g(10 - 20x_1)$

x_1	$h_\theta(x)$
0	$g(10) \approx 1$
1	$g(-10) \approx 0$

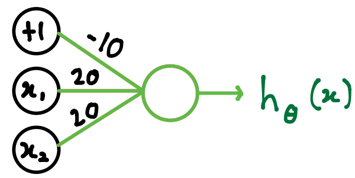
x_1 XNOR x_2 :



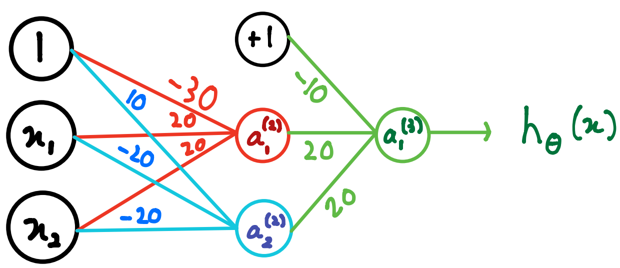
x_1 AND x_2



(NOT x_1) AND (NOT x_2)



x_1 OR x_2



x_1	x_2	$a_1^{(2)}$	$a_2^{(2)}$	$h_\theta(x)$
0	0	0	1	1
0	1	0	0	0
1	0	0	0	0
1	1	1	0	1

AND
$\theta^{(1)} = [-30, 20, 20]$

NOR
$\theta^{(1)} = [10, -20, -20]$

OR
$\theta^{(1)} = [-10, 20, 20]$

$$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} \longrightarrow \begin{bmatrix} a^{(3)} \end{bmatrix} \longrightarrow h_\theta(x)$$

$$\theta^{(1)} = \begin{bmatrix} -30 & 20 & 20 \\ 10 & -20 & -20 \end{bmatrix}$$

$$\theta^{(2)} = \begin{bmatrix} -10 & 20 & 20 \end{bmatrix}$$

$$a^{(2)} = g(\theta^{(1)} \cdot x)$$

$$a^{(3)} = g(\theta^{(2)} \cdot a^{(2)})$$

$$h_\theta(x) = a^{(3)}$$