# ML w4 neural networks

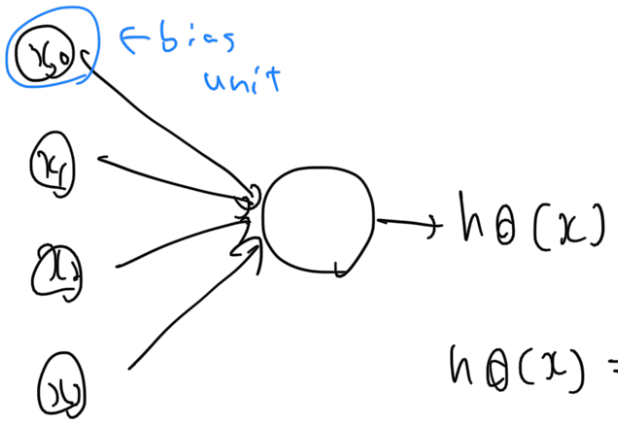
# Non-linear Hypotheses.

polynomial ~~to~~ logistic regression — overfitting
— high cost $(\theta^n)$

⇩

we need non linear hypothesis for machine learning problem.

## Neuron Model : Logistic Unit

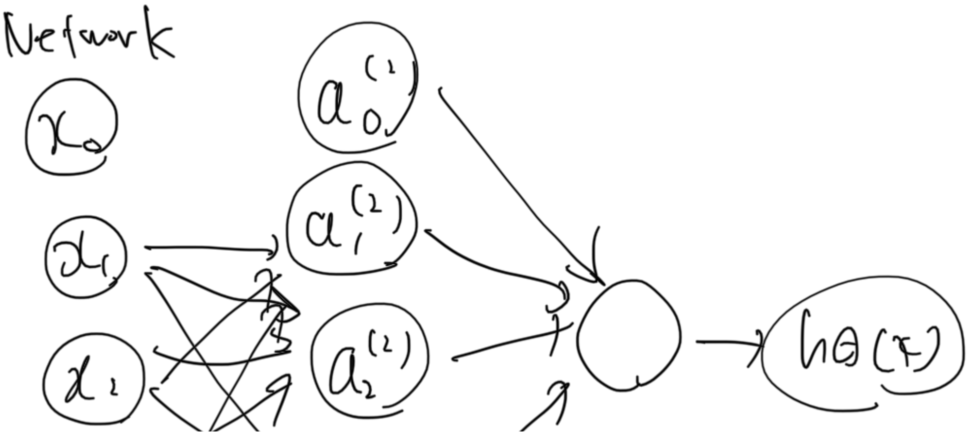



$$h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$$

↑ activation function

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

## Neural Network




layer 1    layer 2    layer 3

Input Layer    hidden layer    output layer

$\uparrow$    $\uparrow$

$a_i^{(j)}$    $\Theta^{(j)}$

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

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

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

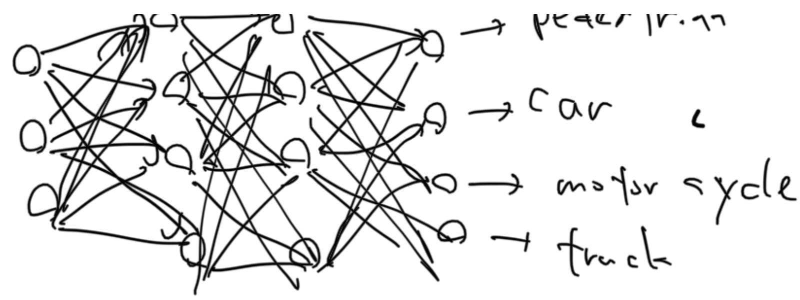
$$\longrightarrow h_\Theta(x) = a_1^{(3)} = g(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13}^{(2)} a_3^{(2)})$$

if network has $s_j$ units in layer $j$, $s_{j+1}$ units in layer $j+1$, then $\Theta^{(j)}$ will be of dimension :

$$\boxed{s_{j+1} \times (s_j + 1)}$$

---

Multi-Class classification




want $h_\Theta(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ $h_\Theta(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ $h_\Theta(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$