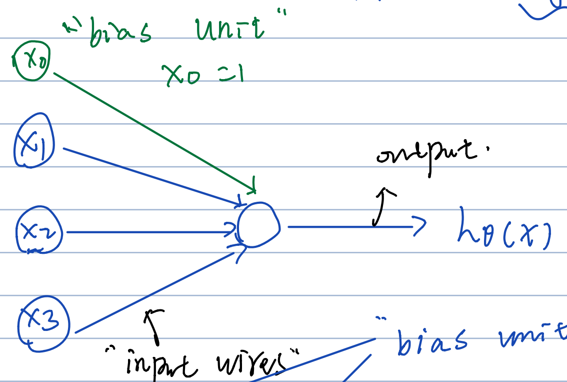
Neuron in your brain:

Input: Dendrite (impulse)

Output: Axon → to other Neurons.

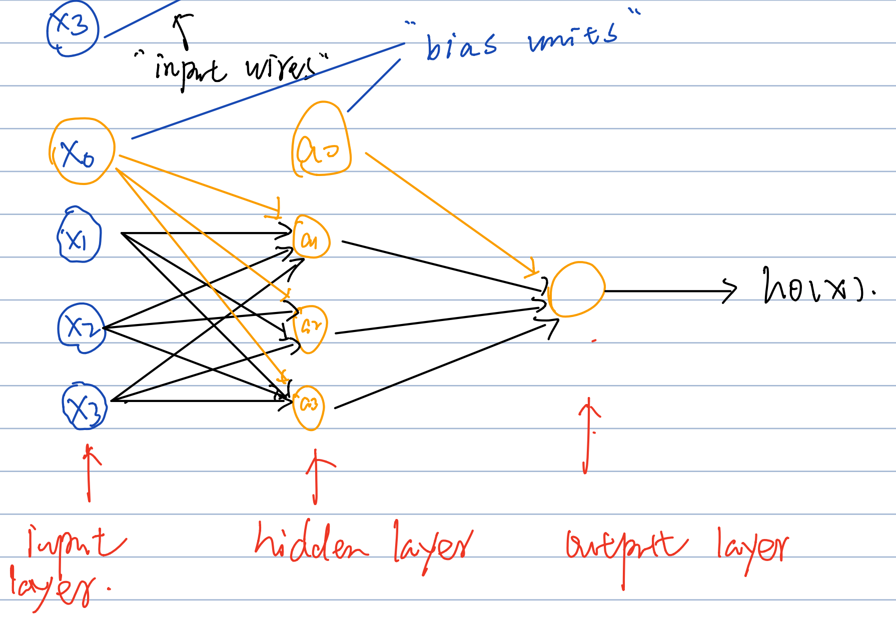
Neuron Model: ANN ← logistic unit.



$$X = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \vdots \\ \theta_{n-1} \end{bmatrix}$$

"weights" (the same as parameters)

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



If network has s_j units in layer j, and s_{j+1} units in layer $j+1$, then $\theta^{(j)}$ will be of dimension $s_{j+1} \times (s_j + 1)$

Model Representation:

$z^{(2)} = \theta^{(1)} \times X$

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

x_1, x_2 is binary

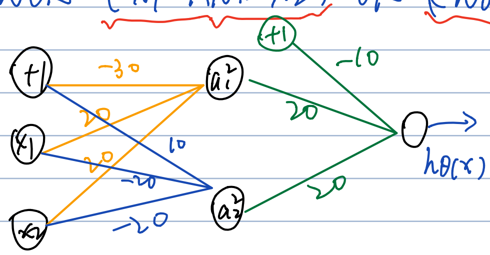
$h_\theta(x) = g(-10 + 20x_1 + 20x_2)$

x_1	x_2	$h_\theta(x)$
0	0	$y \approx 0$
0	1	$y \approx 1$
1	0	$y \approx 1$
1	1	$y \approx 1$

x_1 or x_2.

XNOR 网络:

XNOR = (x_1 AND x_2) OR (NOT x_1 AND NOT x_2).



— x_1 AND x_2

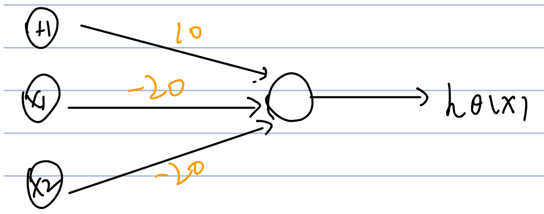
— $\neg x_1$ AND $\neg x_2$

— x_1 OR x_2

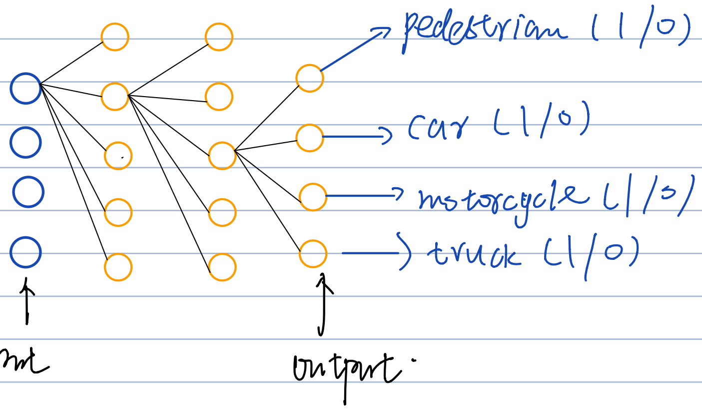
$\neg x_1 \wedge \neg x_2$

x_1	x_2	$h_\theta(x)$
0	0	1
0	1	0
1	0	0
1	1	1

(NOT x_1 AND NOT x_2)

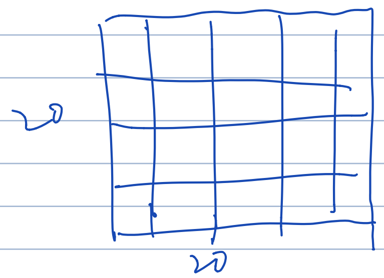


Multiclass classification



输出为四个类别的分类.

$$y^{(i)} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \text{ or } \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \text{ or } \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \text{ or } \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$



$$\longrightarrow \underset{(n+1)\times 1}{x^{(1)}} = \begin{bmatrix} x_0^{(1)} \\ x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_n^{(1)} \end{bmatrix}$$

m个 traing examples

$$\begin{array}{c} \\ \underset{m\times(n+1)}{X} = \end{array} \begin{bmatrix} -(x^{(1)})^T- \\ -(x^{(2)})^T- \\ -(x^{(3)})^T- \\ \vdots \\ -(x^{(m)})^T- \end{bmatrix} = \begin{array}{cccccc} \theta_0 & \theta_1 & \theta_2 & \cdots & \boxed{\theta_n} \end{array}$$

$$= \begin{bmatrix} x_0^{(1)}, & x_1^{(1)}, & x_2^{(1)} & \cdots, & x_n^{(1)} \\ x_0^{(2)}, & x_1^{(2)}, & x_2^{(2)} & \cdots, & x_n^{(2)} \\ \vdots & & & & \\ x_0^{(m)}, & x_1^{(m)}, & x_2^{(m)}, & \cdots, & x_n^{(m)} \end{bmatrix}$$

$m = 5000$

$$\underset{m\times 1}{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix}$$

$$\theta \in \mathbb{R}^{n+1}$$

$$\underset{(n+1)\times 1}{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$$

$m\times(n+1)$ $(n+1)\times 1$

$h_\theta(x) = g(\theta^T x^{(i)}) = g(\underline{X} \cdot \underline{\theta}) \Rightarrow \underline{m \times 1}$ $(n\times 1)(n\times 1)$

$J(\theta) = \frac{1}{m}(-y^{(i)} \cdot \log(h_\theta(x^{(i)})) - (1-y^{(i)})\log(1-h_\theta(x^{(i)}))) + \frac{1}{2m}\sum_{j=1}^{n}\theta_j^2$

Vectorized

$J(\theta) = \frac{1}{m}[\underbrace{-y^T}_{1\times m} \cdot \underbrace{g(\underline{X} \cdot \underline{\theta})}_{m\times 1} - \underbrace{(1-y)^T}_{1\times m} \cdot \underbrace{(1 - g(\underline{X} \cdot \underline{\theta}))}_{m\times 1}] + \frac{1}{2m} \cdot \underbrace{\theta[1:n]^T}_{1\times n} * \underbrace{\theta[1:n]}_{n\times 1}$

grad = grad (un-regularized) + grad (regularized)

$\boxed{\begin{array}{l} \frac{\partial}{\partial\theta_j}J(\theta) = \frac{1}{m}(h_\theta(x^{(i)}) - y^{(i)})\, x_j^{(i)} \leftarrow \text{无正则化项.} \\ \text{temp} = \theta \\ \text{temp}(0) = 0 \leftarrow \theta_0^{(i)} = 0 \\ \text{grad}(\text{正则化}) = \frac{\lambda}{m} \cdot \theta \end{array}}$

$j=0: \theta_0^{(i)} := \theta_0^{(i)} - \alpha[\frac{1}{m}(h_\theta(x_1^{(i)}) - y_1^{(i)}) \cdot x_1^{(i)}]$

$j\neq 0: \theta_j^{(i)} := \theta_j^{(i)} - \alpha[\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)}) \cdot x_j^{(i)} + \frac{\lambda}{m}\sum_{j=1}^{n}\theta_j]$