Neural Networks: Representation
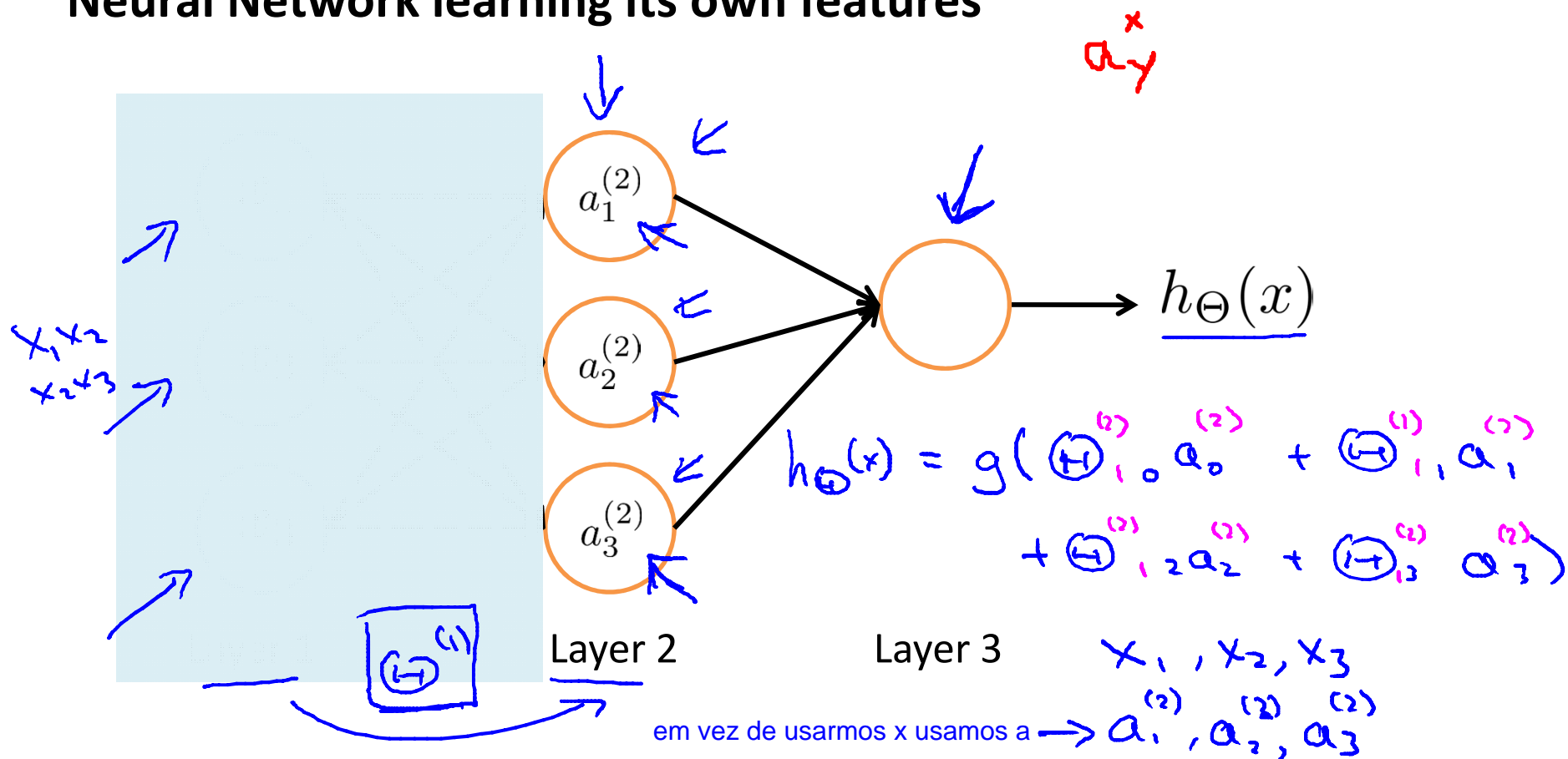
Model representation II

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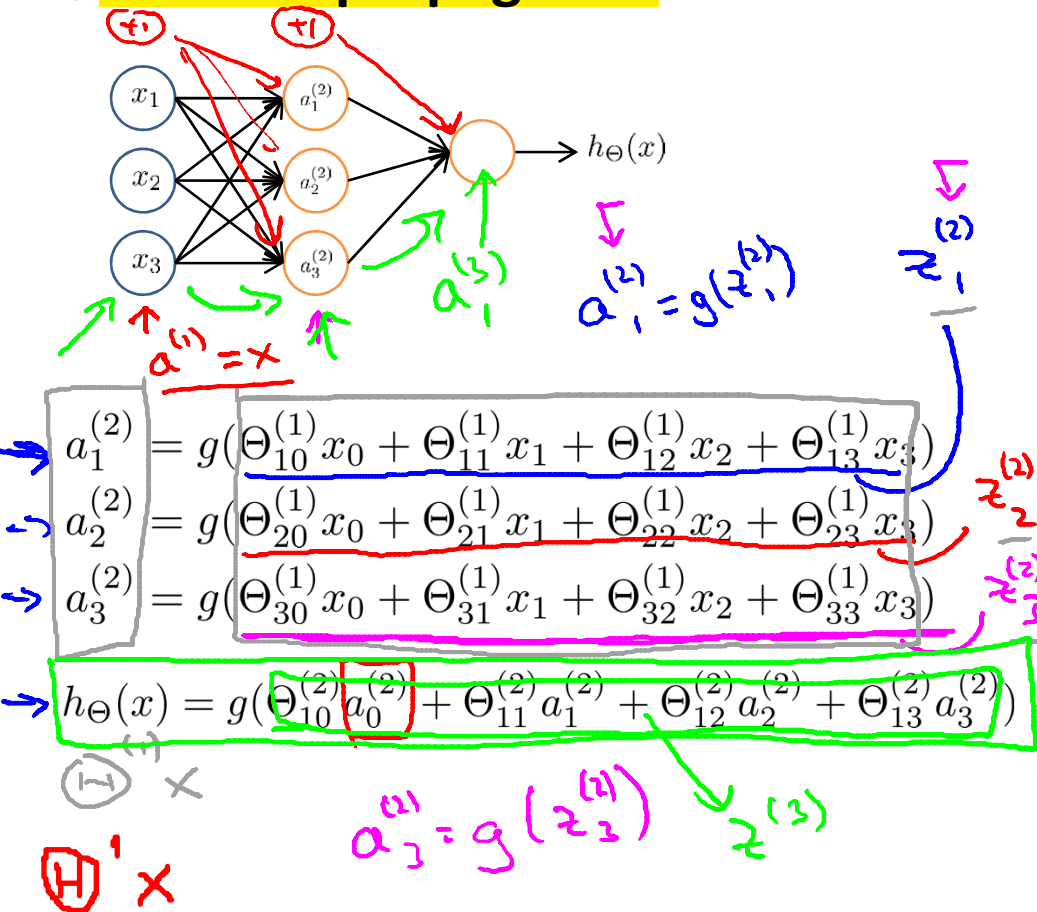
$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad z^{(2)} = \begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \\ z_3^{(2)} \end{bmatrix}$$

$a^{(1)} = x$

$$a_1^{(2)} = g(\Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3)$$
$$a_2^{(2)} = g(\Theta_{20}^{(1)} x_0 + \Theta_{21}^{(1)} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3)$$
$$a_3^{(2)} = g(\Theta_{30}^{(1)} x_0 + \Theta_{31}^{(1)} x_1 + \Theta_{32}^{(1)} x_2 + \Theta_{33}^{(1)} x_3)$$
$$h_\Theta(x) = g(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13}^{(2)} a_3^{(2)})$$

$a_1^{(2)} = g(z_1^{(2)})$

$a_3^{(2)} = g(z_3^{(2)})$

$z^{(2)}$

$z^{(3)}$

$a^{(3)}$

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

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

$\mathbb{R}^3 \qquad \mathbb{R}^3$

Add $a_0^{(2)} = 1.$ $\longrightarrow a^{(2)} \in \mathbb{R}^4$

$a_1^{(2)}$
$a_2^{(2)}$
$a_3^{(2)}$

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

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

Andrew Ng

Neural Network learning its own features

$a - y$

$x_1 x_2$
$x_2 x_3$

$h_\Theta(x)$

Layer 2   Layer 3

$$h_\Theta(x) = g\left(\Theta^{(2)}_{10} a^{(2)}_0 + \Theta^{(1)}_{11} a^{(2)}_1 + \Theta^{(2)}_{12} a^{(2)}_2 + \Theta^{(2)}_{13} a^{(2)}_3 \right)$$

x_1, x_2, x_3

em vez de usarmos x usamos a $\longrightarrow a^{(2)}_1, a^{(2)}_2, a^{(2)}_3$

Other network architectures

como os neuronios estao ligados

$h_\Theta(x)$

Layer 1
Input

Layer 2

Layer 3

Layer 4
Output

Hidden layer