$$l = -y \ln\left(x_1^{(3)}\right) - (1-y)\ln\left(1 - x_1^{(3)}\right)$$

l

$$x_3^{(2)} = \sigma\left(s_1^{(2)}\right) = \frac{1}{1+e^{-s_1^{(2)}}}$$

$x_1^{(3)}$

$$s_1^{(2)} = \mathbf{w_1^{(2)}}^\top \mathbf{x^{(2)}}$$

$s_1^{(2)}$

$\mathbf{w_1^{(2)}}$

We collectively refer to these nodes as $\mathbf{x^{(2)}}$

$$x_1^{(2)} = \sigma\left(s_1^{(1)}\right) = \frac{1}{1+e^{-s_1^{(1)}}}$$

$x_1^{(2)}$ ... $x_m^{(2)}$

$$x_m^{(2)} = \sigma\left(s_m^{(1)}\right) = \frac{1}{1+e^{-s_m^{(1)}}}$$

$$s^{(1)} = \mathbf{w_1^{(1)}}^\top \mathbf{x}$$

$s_1^{(1)}$ ... $s_m^{(1)}$

$$s^{(m)} = \mathbf{w_m^{(1)}}^\top \mathbf{x}$$

$\mathbf{w_1^{(1)}}$ $\mathbf{w_m^{(1)}}$

$\mathbf{x^{(1)}}$