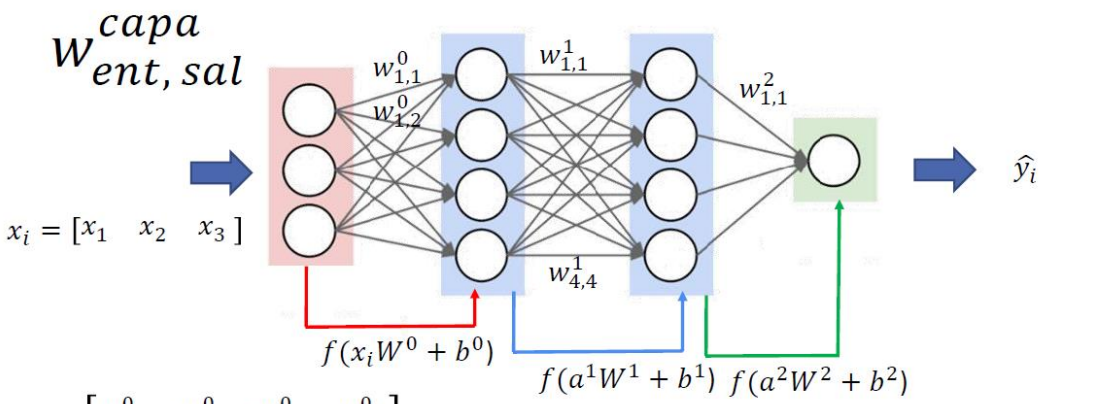
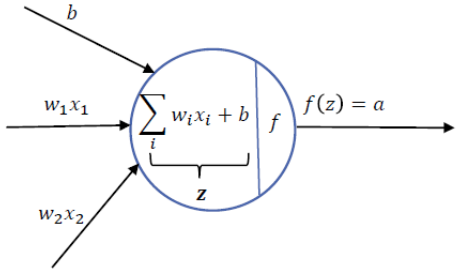
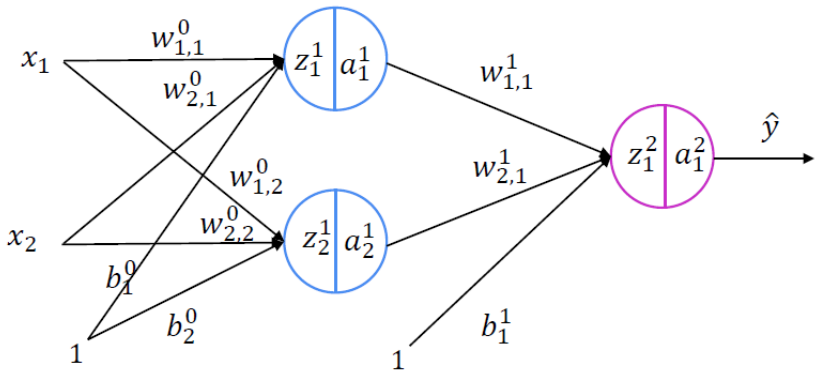
Estructura de una red neuronal



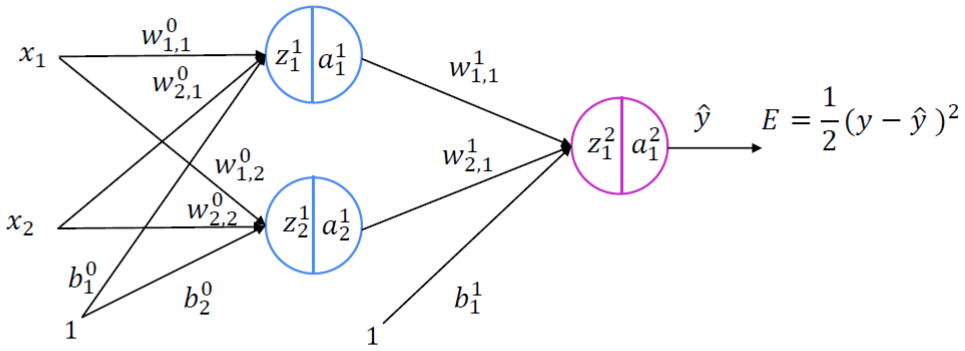
$$W^0 = \begin{bmatrix} w^0_{1,1} & w^0_{1,2} & w^0_{1,3} & w^0_{1,4} \\ w^0_{2,1} & w^0_{2,2} & w^0_{2,3} & w^0_{2,4} \\ w^0_{3,1} & w^0_{3,2} & w^0_{3,3} & w^0_{3,4} \end{bmatrix}_{3\times4} \quad W^1 = \begin{bmatrix} w^1_{1,1} & w^1_{1,2} & w^1_{1,3} & w^1_{1,4} \\ w^1_{2,1} & w^1_{2,2} & w^1_{2,3} & w^1_{2,4} \\ w^1_{3,1} & w^1_{3,2} & w^1_{3,3} & w^1_{3,4} \\ w^1_{4,1} & w^1_{4,2} & w^1_{4,3} & w^1_{4,4} \end{bmatrix}_{4\times4} \quad W^2 = \begin{bmatrix} w^2_{1,1} \\ w^2_{2,1} \\ w^2_{3,1} \\ w^2_{4,1} \end{bmatrix}_{4\times1}$$

Forward and Backward propagation





Forward



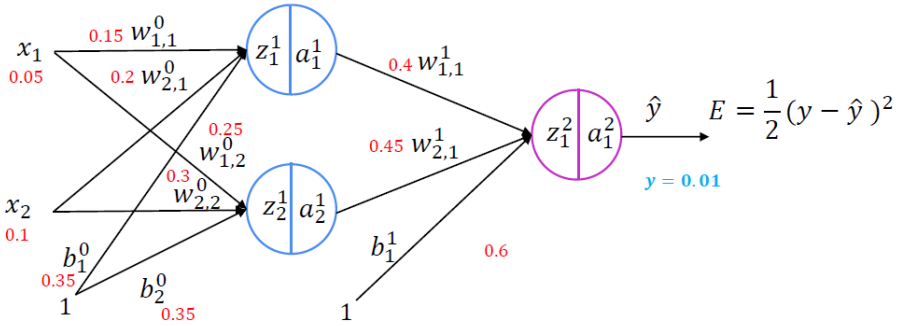
$$z^1_1 = w^0_{1,1}x_1 + w^0_{2,1}x_2 + b^0_1 \qquad z^1_2 = w^0_{1,2}x_1 + w^0_{2,2}x_2 + b^0_2$$

$$a^1_1 = \frac{1}{1 + e^{-z^1_1}} \qquad a^1_2 = \frac{1}{1 + e^{-z^1_2}}$$

$$z^2_1 = w^1_{1,1}a^1_1 + w^1_{2,1}a^1_2 + b^1_1$$

$$a^2_1 = \frac{1}{1 + e^{-z^2_1}} = \hat{y}$$

Forward propagation



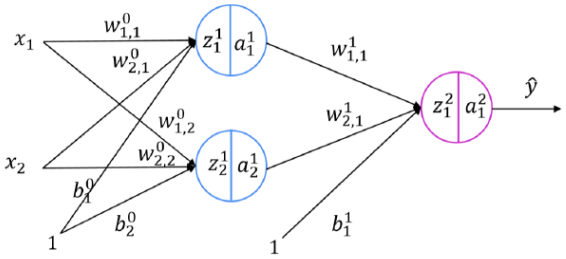
$z^1_1 = w^0_{1,1}x_1 + w^0_{2,1}x_2 + b^0 = 0.15 * 0.05 + 0.2 * 0.1 + 0.35 * 1 = 0.3775$ $\quad a^1_1 = \frac{1}{1 + e^{-z^1_1}} = \frac{1}{1 + e^{-0.3775}} = 0.5933$

$z^1_2 = w^0_{1,2}x_1 + w^0_{2,2}x_2 + b^0 = 0.25 * 0.05 + 0.3 * 0.1 + 0.35 * 1 = 0.3925$ $\quad a^1_2 = \frac{1}{1 + e^{-z^1_2}} = \frac{1}{1 + e^{-0.3925}} = 0.5969$

$z^2_1 = w^1_{1,1}a^1_1 + w^1_{2,1}a^1_2 + b^1 = 0.4 * 0.5933 + 0.45 * 0.5969 + 0.6 * 1 = 1.1059$ $\quad a^2_1 = \frac{1}{1 + e^{-z^2_1}} = = \frac{1}{1 + e^{-1.1059}} = 0.7514$

$$E = \frac{1}{2}(y - \hat{y})^2 = \frac{1}{2}(y - a^2_1)^2 = 0.2748$$

Backpropagation



1. Actualización de pesos etapa 1

$$w^1_{1,1}(t+1) = w^1_{1,1}(t) - \eta \frac{\partial E}{\partial w^1_{1,1}}$$

$$w^1_{2,1}(t+1) = w^1_{2,1}(t) - \eta \frac{\partial E}{\partial w^1_{2,1}}$$

1. Actualización de pesos etapa 0

$$w^0_{1,1}(t+1) = w^0_{1,1}(t) - \eta \frac{\partial E}{\partial w^0_{1,1}}$$

$$w^0_{2,1}(t+1) = w^0_{2,1}(t) - \eta \frac{\partial E}{\partial w^0_{2,1}}$$

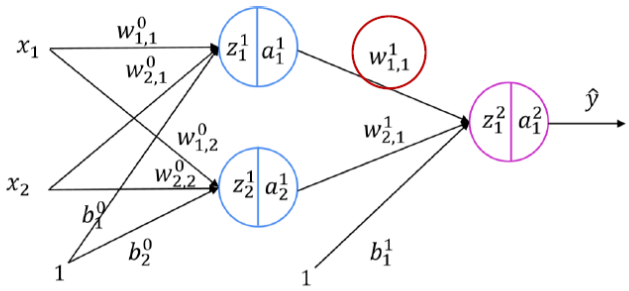
$$w^0_{1,2}(t+1) = w^0_{1,2}(t) - \eta \frac{\partial E}{\partial w^0_{1,2}}$$

$$w^0_{2,2}(t+1) = w^0_{2,2}(t) - \eta \frac{\partial E}{\partial w^0_{2,2}}$$

Backpropagation

$$w^1_{1,1}(t+1) = w^1_{1,1}(t) - \eta \frac{\partial E}{\partial w^1_{1,1}}$$

$$\frac{\partial E}{\partial w^1_{1,1}}$$



$$E = \frac{1}{2}(y - \hat{y})^2 = \frac{1}{2}(y - a^2_1)^2$$

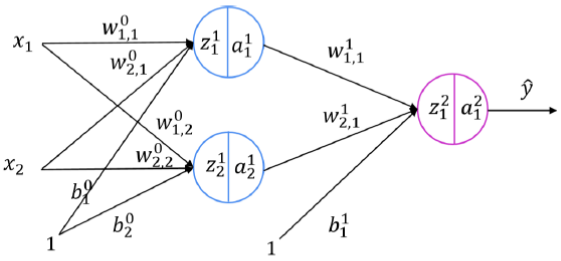
$$\frac{\partial E}{\partial w^1_{1,1}} \longrightarrow a^2_1 = \frac{1}{1+e^{-z^2_1}} \longrightarrow z^2_1 = w^1_{1,1}a^1_1 + w^1_{2,1}a^1_2 + b^1_1$$

Regla de la cadena

$$\frac{\partial E}{\partial w^1_{1,1}} = \frac{\partial E}{\partial a^2_1} * \frac{\partial a^2_1}{\partial z^2_1} * \frac{\partial z^2_1}{\partial w^1_{1,1}}$$

Backpropagation

$$\frac{\partial E}{\partial w_{1,1}^1} = \frac{\partial E}{\partial a_1^2} * \frac{\partial a_1^2}{\partial z_1^2} * \frac{\partial z_1^2}{\partial w_{1,1}^1}$$



$$\frac{\partial E}{\partial a_1^2} \qquad E = \frac{1}{2}(y - a_1^2)^2$$

$$\frac{\partial E}{\partial a_1^2} = \frac{1}{2} * 2 * (-1)(y - a_1^2) = (a_1^2 - y)$$

$$\frac{\partial a_1^2}{\partial z_1^2} \qquad a_1^2 = \frac{1}{1 + e^{-z_1^2}} \qquad \frac{\partial a_1^2}{\partial z_1^2} = \frac{-e^{-z_1^2}}{\left(1 + e^{-z_1^2}\right)^2} = a_1^2(1 - a_1^2)$$

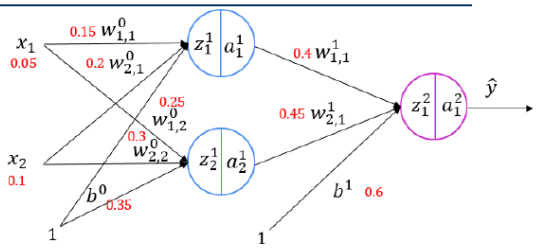
$$\frac{\partial z_1^2}{\partial w_{1,1}^1} \qquad z_1^2 = w_{1,1}^1 a_1^1 + w_{2,1}^1 a_2^1 + b_1^1 \qquad \frac{\partial z_1^2}{\partial w_{1,1}^1} = a_1^1$$

$$\frac{\partial E}{\partial w_{1,1}^1} = (a_1^2 - y) * a_1^2(1 - a_1^2) * a_1^1$$

Backpropagation

$$\frac{\partial E}{\partial w_{1,1}^1} = \frac{\partial E}{\partial a_1^2} * \frac{\partial a_1^2}{\partial z_1^2} * \frac{\partial z_1^2}{\partial w_{1,1}^1}$$

$$\delta_{a_1^2} = \frac{\partial E}{\partial a_1^2} * \frac{\partial a_1^2}{\partial z_1^2}$$



$$\delta_{a_1^2} = (a_1^2 - y) * a_1^2(1 - a_1^2)$$

$$\frac{\partial E}{\partial w_{1,1}^1} = \delta_{a_1^2} * a_1^1$$

$$\delta_{a_1^2} = (0.7514 - 0.01) * 0.7514 * (1 - 0.7514) = 0.1385 \qquad \frac{\partial E}{\partial w_{1,1}^1} = 0.1385 * 0.5933 = 0.0822$$

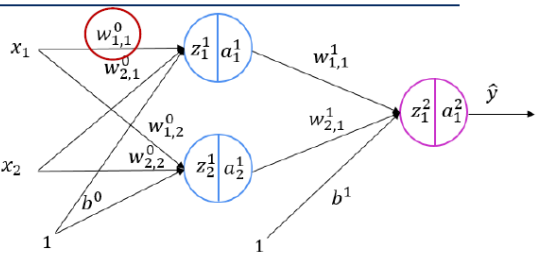
Cuánto un cambio en $\boldsymbol{w_{1,1}^1}$ afecta a $\boldsymbol{E}$

supongamos $\eta = 0.5$ $\qquad w_{1,1}^1(t+1) = w_{1,1}^1(t) - \eta \frac{\partial E}{\partial w_{1,1}^1} = 0.4 - 0.5 * 0.082 = 0.3590$

Backpropagation

$$w_{1,1}^0(t+1) = w_{1,1}^0(t) - \eta \frac{\partial E}{\partial w_{1,1}^0}$$

$$\frac{\partial E}{\partial w_{1,1}^0} \qquad E = \frac{1}{2}(y - \hat{y})^2 = \frac{1}{2}(y - a_1^2)^2$$



$$\frac{\partial E}{\partial w_{1,1}^0} \longrightarrow a_1^2 = \frac{1}{1+e^{-z_1^2}} \longrightarrow z_1^2 = w_{1,1}^1 a_1^1 + w_{2,1}^1 a_2^1 + b^1 \longrightarrow a_1^1 = \frac{1}{1+e^{-z_1^1}}$$

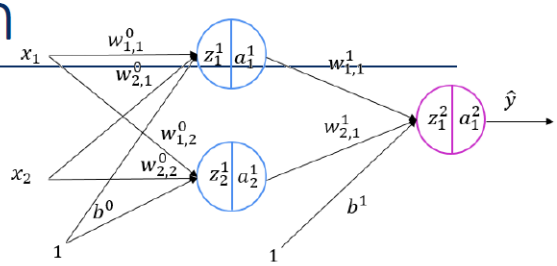
$$z_1^1 = w_{1,1}^0 x_1 + w_{2,1}^0 x_2 + b^0$$

$$\frac{\partial E}{\partial w_{1,1}^0} = \frac{\partial E}{\partial a_1^2} * \frac{\partial a_1^2}{\partial z_1^2} * \frac{\partial z_1^2}{\partial a_1^1} * \frac{\partial a_1^1}{\partial z_1^1} * \frac{\partial z_1^1}{\partial w_{1,1}^0}$$

Backpropagation

$$\frac{\partial E}{\partial w_{1,1}^0} = \frac{\partial E}{\partial a_1^2} * \frac{\partial a_1^2}{\partial z_1^2} * \frac{\partial z_1^2}{\partial a_1^1} * \frac{\partial a_1^1}{\partial z_1^1} * \frac{\partial z_1^1}{\partial w_{1,1}^0}$$

$$\frac{\partial E}{\partial a_1^2} * \frac{\partial a_1^2}{\partial z_1^2} \rightarrow \delta_{a_1^2}$$



$$\frac{\partial z_1^2}{\partial a_1^1} \qquad z_1^2 = w_{1,1}^1 a_1^1 + w_{2,1}^1 a_2^1 + b^1 \qquad \frac{\partial z_1^2}{\partial a_1^1} = w_{1,1}^1$$

$$\frac{\partial a_1^1}{\partial z_1^1} \qquad a_1^1 = \frac{1}{1+e^{-z_1^1}} \qquad \frac{\partial a_1^1}{\partial z_1^1} = a_1^1 * (1 - a_1^1)$$

$$\frac{\partial z_1^1}{\partial w_{1,1}^0} \qquad z_1^1 = w_{1,1}^0 x_1 + w_{2,1}^0 x_2 + b^0 \qquad \frac{\partial z_1^1}{\partial w_{1,1}^0} = x_1$$

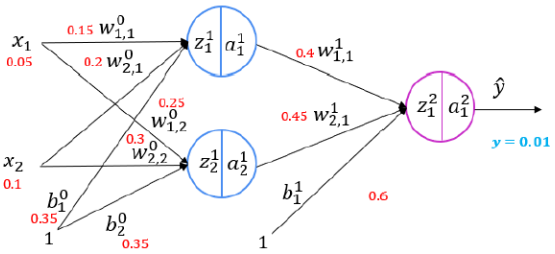
$$\frac{\partial E}{\partial w_{1,1}^0} = \delta_{a_1^2} * \frac{\partial z_1^2}{\partial a_1^1} * \frac{\partial a_1^1}{\partial z_1^1} * \frac{\partial z_1^1}{\partial w_{1,1}^0} = \delta_{a_1^2} * w_{1,1}^1 * a_1^1 * (1 - a_1^1) * x_1$$

Backpropagation

$$\frac{\partial E}{\partial w^0_{1,1}} = \delta_{a^2_1} * w^1_{1,1} * a^1_1 * (1 - a^1_1) * x_1$$

$$\delta_{a^1_1} = \delta_{a^2_1} * w^1_{1,1} * a^1_1 * (1 - a^1_1)$$

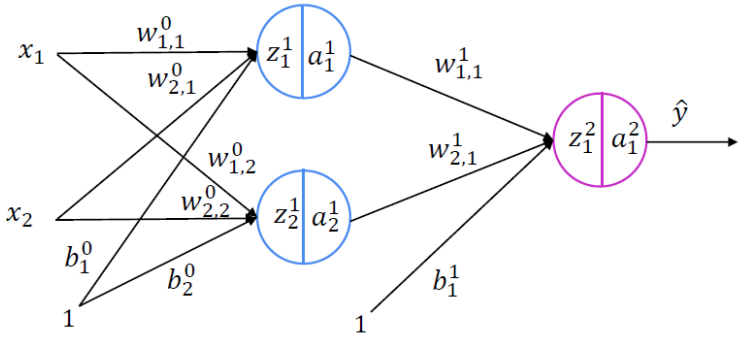
$$\delta_{a^1_1} = 0.1385 * 0.0965$$



$$\frac{\partial E}{\partial w^0_{1,1}} = \delta_{a^2_1} * \delta_{a^1_1} * x_1 = 0.1385 * 0.0965 * 0.05 = 6.6835e-04$$

supongamos $\eta = 0.5$ $w^0_{1,1}(t+1) = w^0_{1,1}(t) - \eta \frac{\partial E}{\partial w^0_{1,1}} = 0.15 - 0.5 * 6.6835e-04 = 0.1497$

Backpropagation



Matricialmente

Capas $l: 0 \ldots . L$ $A^0 = [x_1 \quad x_2 \quad 1]$ $Z^1 = [z^1_1 \quad z^1_2]$ $A^1 = [a^1_1 \quad a^1_2 \quad 1]$

$$W^0 = \begin{bmatrix} w^0_{1,1} & w^0_{1,2} \\ w^0_{2,1} & w^0_{2,2} \\ b^0_1 & b^0_2 \end{bmatrix} \qquad W^1 = \begin{bmatrix} w^1_{1,1} \\ w^1_{2,1} \\ b^1_1 \end{bmatrix}$$

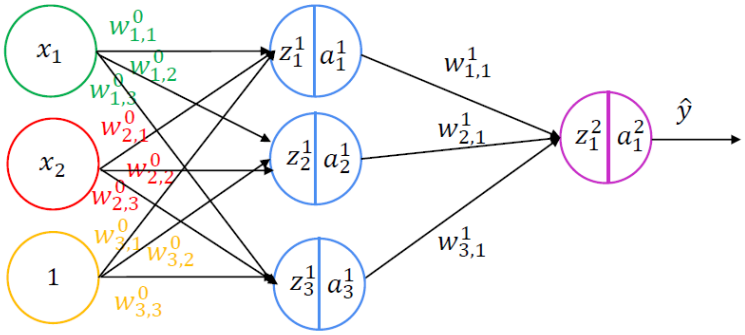
Forward propagation

Matricialmente

Capas $l: 0 \ldots . L$

$$W^0 = \begin{bmatrix} w^0_{1,1} & w^0_{1,2} & w^0_{1,3} \\ w^0_{2,1} & w^0_{2,2} & w^0_{2,3} \\ w^0_{3,1} & w^0_{3,2} & w^0_{3,3} \end{bmatrix}$$

$$W^1 = \begin{bmatrix} w^1_{1,1} \\ w^1_{2,1} \\ w^1_{3,1} \end{bmatrix}$$



Producto matricial

$$A^1 = sigmoide(A^0 * W^0)$$

$$A^0 = [x_1 \quad x_2 \quad 1] \qquad A^1 = [a^1_1 \quad a^1_2 \quad a^1_3]$$

$$A^2 = sigmoide(A^1 * W^1)$$

$$A^2 = [a^2_1]$$

Backpropagation

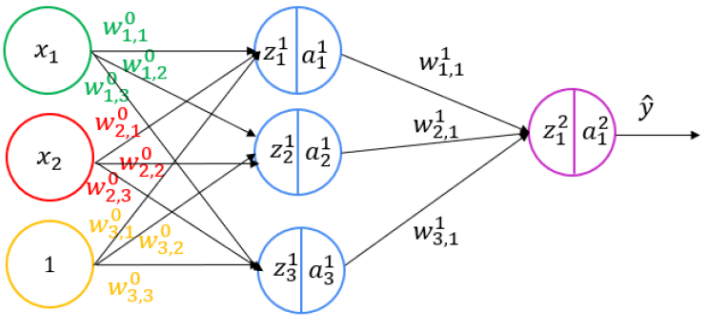
Capa 2

$$w^1_{1,1}(t+1) = w^1_{1,1}(t) - \eta \delta_{a^2_1} a^1_1$$

$$w^1_{2,1}(t+1) = w^1_{2,1}(t) - \eta \delta_{a^2_1} a^1_2$$

$$w^1_{3,1}(t+1) = w^1_{3,1}(t) - \eta \delta_{a^2_1} a^1_3$$

$$\delta_{a^2_1} = (a^2_1 - y) * a^2_1(1 - a^2_1)$$



Capa 1

$w^0_{1,1}(t+1) = w^0_{1,1}(t) - \eta \delta_{a^1_1} x_1$	$w^0_{1,2}(t+1) = w^0_{1,2}(t) - \eta \delta_{a^1_2} x_1$	$w^0_{1,3}(t+1) = w^0_{1,3}(t) - \eta \delta_{a^1_3} x_1$
$w^0_{2,1}(t+1) = w^0_{2,1}(t) - \eta \delta_{a^1_1} x_2$	$w^0_{2,2}(t+1) = w^0_{2,2}(t) - \eta \delta_{a^1_2} x_2$	$w^0_{2,3}(t+1) = w^0_{2,3}(t) - \eta \delta_{a^1_3} x_2$
$w^0_{3,1}(t+1) = w^0_{3,1}(t) - \eta \delta_{a^1_1}$	$w^0_{3,2}(t+1) = w^0_{3,2}(t) - \eta \delta_{a^1_2}$	$w^0_{3,3}(t+1) = w^0_{3,3}(t) - \eta \delta_{a^1_3}$
$\delta_{a^1_1} = \delta_{a^2_1} \cdot w^1_{1,1} \cdot a^1_1 \cdot (1 - a^1_1)$	$\delta_{a^1_2} = \delta_{a^2_1} \cdot w^1_{2,1} \cdot a^1_2 \cdot (1 - a^1_2)$	$\delta_{a^1_3} = \delta_{a^2_1} \cdot w^1_{3,1} \cdot a^1_3 \cdot (1 - a^1_3)$

Backpropagation

Matricialmente

Capas $l: 0 \ldots . L$

$$W^0 = \begin{bmatrix} w^0_{1,1} & w^0_{1,2} & w^0_{1,3} \\ w^0_{2,1} & w^0_{2,2} & w^0_{2,3} \\ w^0_{3,1} & w^0_{3,2} & w^0_{3,3} \end{bmatrix} \quad W^1 = \begin{bmatrix} w^1_{1,1} \\ w^1_{2,1} \\ w^1_{3,1} \end{bmatrix} \quad A^0 = [x_1 \quad x_2 \quad 1]$$

$$A^1 = [a^1_1 \quad a^1_2 \quad a^1_3] \qquad A^2 = [a^2_1]$$

Capa 1

$$\delta_{a^1_1} = \delta_{a^2_1} \cdot w^1_{1,1} \cdot a^1_1 \cdot (1 - a^1_1)$$

$$\delta_{a^1_2} = \delta_{a^2_1} \cdot w^1_{2,1} \cdot a^1_2 \cdot (1 - a^1_2)$$

$$\delta_{a^1_3} = \delta_{a^2_1} \cdot w^1_{3,1} \cdot a^1_3 \cdot (1 - a^1_3)$$

Capa 2

$$\delta_{a^2_1} = (a^2_1 - y) * a^2_1(1 - a^2_1)$$

$$\Delta^2 = \left[\delta_{a^2_1}\right] = (a^2_1 - y) * (A^2)'$$

$$W^1(t+1) = W^1(t) - \eta (A^1)^T * \Delta^2$$

$$\Delta^1 = \left[\delta_{a^1_1} \quad \delta_{a^1_2} \quad \delta_{a^1_3}\right] =$$

$$= \Delta^2 * \begin{bmatrix} w^1_{1,1} \\ w^1_{2,1} \\ w^1_{3,1} \end{bmatrix}^T \cdot [(a^1_1)' \quad (a^1_2)' \quad (a^1_3)']$$

$$\Delta^1 = \Delta^2 * (W^1)^T \qquad \Delta^1 = \Delta^1 \cdot (A^1)'$$

$$W^0(t+1) = W^0(t) - \eta (A^0)^T * \Delta^1$$