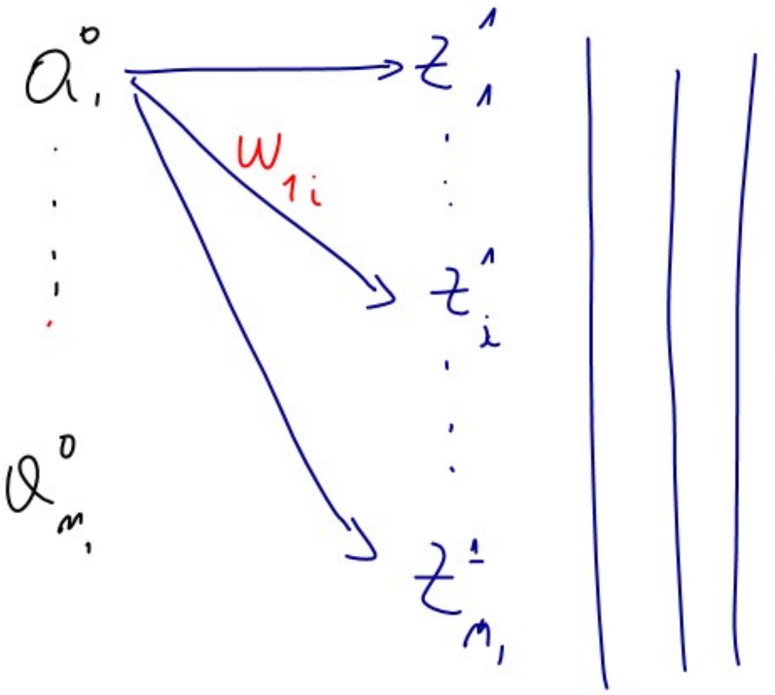


$$z^{l+1}_j = \sum_h w^{l+1}_{hj}\, a^l_h$$

$$z^l_1 \to a^l_1 = f(z^l_1)$$

$$z^l_i \to a^l_i = f(z^l_i) \xrightarrow{w^{l+1}_{ij}} z^{l+1}_j \to a^{l+1}_j = f(z^{l+1}_j)$$

$$z^L_1 \to f(z^L_1) = a^L_1$$

$$z^L_h \to f(z^L_h) = a^L_h$$

$$a^L_{m_1}$$

$$\delta^l_i = \frac{\partial \mathcal{L}}{\partial z^l_i} = \sum_{j=1}^{m_{l+1}} \frac{\partial \mathcal{L}}{\partial z^{l+1}_j}\, \frac{\partial z^{l+1}_j}{\partial a^l_i}\, \frac{\partial a^l_i}{\partial z^l_i}\, f'(z^l_i)$$

δ^{l+1}_j w^{l+1}_{ij}

$$\delta^l_i = \sum_j \delta^{l+1}_j\, w^{l+1}_{ij}\, f'(z^l_i) = f'(z^l_i) \sum_j \delta^{l+1}_j\, w^{l+1}_{ij}$$

TEOREMA DE APROXIMACIÓN UNIVERSAL

$\varphi : K \to \mathbb{R}$ $\quad$ K COMPACTO

φ CONTINUA

$\forall \varepsilon > 0 \ \exists$ RED NEURONAL : (i) 1 CAPA OCULTA

(ii) $f(\cdot)$ ACTIVACIÓN SIGMOIDE.

$$\sup_{z \in K} \left| RED(z) - \varphi(z) \right| < \varepsilon$$

$$\delta_i^\ell = \frac{\partial \mathcal{L}}{\partial z_i^\ell}$$

$$\delta_k^L = \frac{\partial \mathcal{L}}{\partial z_k^L}$$

$$\nabla_W \mathcal{L} = \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial w_{11}^\ell} \\ \frac{\partial \mathcal{L}}{\partial w_{12}^\ell} \\ \vdots \end{pmatrix}$$

$$\frac{\partial \mathcal{L}}{\partial w_{ij}^\ell} = \frac{\partial \mathcal{L}}{\partial z_j^\ell} \frac{\partial z_j^\ell}{\partial w_{ij}^\ell}$$

$$\frac{\partial \mathcal{L}}{\partial z_k^L} = \frac{\partial \mathcal{L}}{\partial a_k^L} \cdot \frac{\partial a_k^L}{\partial z_k^L}$$

$$\boxed{\frac{\partial \mathcal{L}}{\partial w_{ij}^\ell} = \delta_j^\ell \, a_i^{\ell-1}}$$

$$\delta_k^L = (a_k^L - y_k) f'(z_k^L)$$

$$\frac{\partial z_j^\ell}{\partial w_{ij}^\ell} = \frac{\partial \sum_h w_{hj}^\ell \, a_h^{\ell-1}}{\partial w_{ij}^\ell} = a_i^{\ell-1}$$

$w(0)$ # VALOR INICIAL

EPOCH

BCM

$\nabla_w \mathcal{L}_B$

$w \leftarrow w - \lambda \nabla_w \mathcal{L}_B$

$$\mathcal{L} = \sum_{i \in B} h(a_i^L, y_i)$$

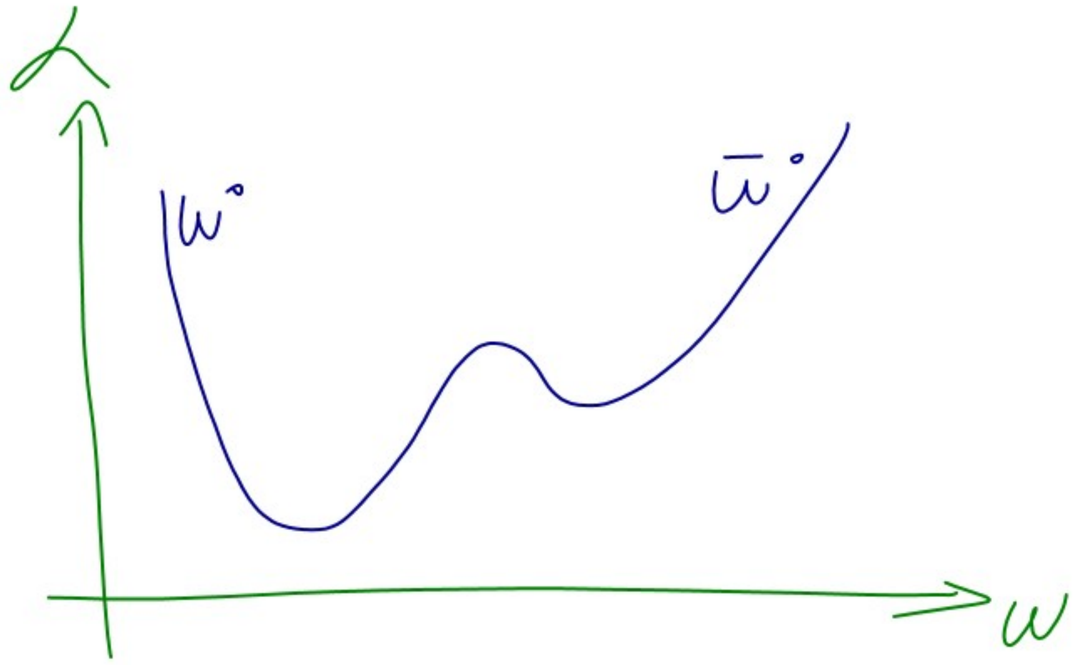
BATCH

BCMUESTRA

$$h(a^L, y) = \frac{1}{2} \sum_i (a_i^L - y_i)^2$$

$$= -\sum_i \sum_k y_{ik} \ln \underbrace{a_{ik}}_{a_k^L(i)}$$

$$h(a^L, y) = \sum_i |a_i^L - y_i|$$


$\mathcal{L}$
ω°
$\bar{\omega}^\circ$
ω

W OBJETIVO $\min_{\{w\}} g_w(x)$

$W(1) \rightarrow g_{w(1)}(x)$ $\longrightarrow$

$\vdots$ $\longrightarrow$ $\hat{W}(j)$

$W(N) \rightarrow g_{w(N)}(x)$ $\longrightarrow$

$j \in$ MEJORES