



Função de Perda:

$$L = \frac{1}{m} \sum_{i=1}^m y_i \ln \hat{y}_i$$

CROSS-ENTROPY

FORMATO CATEGÓRICO
i.e. $y = 3$

$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$$dz^{[2]} = \frac{1}{m} (A^{[2]} - y)$$

$10 \times m$ $10 \times m$ $10 \times m$

$$dW^{[2]} = dz^{[2]} A^{[1]T}$$

10×10 $10 \times m$ $m \times 10$

$$db^{[2]} = \sum dz^{[2]}$$

10×1 10×1

$$dz^{[1]} = W^{[2]T} dz^{[2]} * g'(z^{[1]})$$

$10 \times m$ 10×10 $10 \times m$ $10 \times m$

$$dW^{[1]} = dz^{[1]} X^T$$

10×784 $10 \times m$ $m \times 784$

$$db^{[1]} = \sum dz^{[1]}$$

10×1 10×1

BACK
PROPAGATION

$$A^{[0]} = X \quad (784 \times m)$$

$$z^{[1]} = W^{[1]} A^{[0]} + b^{[1]}$$

$10 \times m$ 10×784 $784 \times m$ 10×1 $\Rightarrow 10 \times m$

$$A^{[1]} = g(z^{[1]}) = \text{ReLU}(z^{[1]})$$

$$z^{[2]} = W^{[2]} A^{[1]} + b^{[2]}$$

$10 \times m$ 10×10 $10 \times m$ 10×1 $\Rightarrow 10 \times m$

$$A^{[2]} = \text{softmax}(z^{[2]})$$

FORWARD
PROPAGATION

$$W^{[1]} = W^{[1]} - \alpha dW^{[1]}$$

$$b^{[1]} = b^{[1]} - \alpha db^{[1]}$$

$$W^{[2]} = W^{[2]} - \alpha dW^{[2]}$$

$$b^{[2]} = b^{[2]} - \alpha db^{[2]}$$

ATUALIZAÇÕES

α É A TAXA DE APRENDIZAGEM