# Ejercicio 1: Backpropagation

Jordi Blasco Lozano

## Enunciado

Calcular el Forward pass y Backward pass para una red neuronal con los siguientes datos:
**Datos iniciales:**

- Entrada: $x = 1$
- Pesos iniciales: $W^{(L-2)} = 0{,}32$, $W^{(L-1)} = 0{,}18$, $W^{(L)} = 0{,}23$
- Bias iniciales: $b^{(L-2)} = 0$, $b^{(L-1)} = 0$, $b^{(L)} = 0$
- Learning rate: $\alpha = 0{,}1$
- Valor objetivo (target): $y = 1$

**Fórmulas utilizadas:**

- Función de coste: $L = C_0 = (\hat{y} - y)^2$
- Función sigmoide: $\sigma(x) = \frac{1}{1+e^{-x}}$
- Derivada de la sigmoide: $\sigma'(x) = \sigma(x)(1 - \sigma(x))$
- Entrada de la neurona: $z^{(l)} = W^{(l)} \cdot a^{(l-1)} + b^{(l)}$
- Activación: $a^{(l)} = \sigma(z^{(l)})$

## 1. Apartado 1: Obtener las salidas $(z, a)$ y la predicción (Forward Pass)

En el forward pass, calculamos las salidas de cada capa desde la entrada hasta la predicción final.

### 1.1. Capa L-2 (Primera capa oculta)

**Paso 1: Calcular $z^{(L-2)}$**

$$z^{(L-2)} = W^{(L-2)} \cdot x + b^{(L-2)} \tag{1}$$
$$z^{(L-2)} = 0{,}32 \cdot 1 + 0 \tag{2}$$
$$z^{(L-2)} = 0{,}32 \tag{3}$$

**Paso 2: Calcular la activación $a^{(L-2)}$**

$$a^{(L-2)} = \sigma(z^{(L-2)}) = \frac{1}{1 + e^{-z^{(L-2)}}} \tag{4}$$
$$a^{(L-2)} = \frac{1}{1 + e^{-0{,}32}} \tag{5}$$
$$a^{(L-2)} = \frac{1}{1 + 0{,}7261} \tag{6}$$
$$a^{(L-2)} = \frac{1}{1{,}7261} \tag{7}$$

$$\boxed{a^{(L-2)} = 0{,}5793}$$

## 1.2. Capa L-1 (Segunda capa oculta)

**Paso 3: Calcular $z^{(L-1)}$**

$$z^{(L-1)} = W^{(L-1)} \cdot a^{(L-2)} + b^{(L-1)} \tag{8}$$
$$z^{(L-1)} = 0{,}18 \cdot 0{,}5793 + 0 \tag{9}$$
$$z^{(L-1)} = 0{,}1043 \tag{10}$$

**Paso 4: Calcular la activación $a^{(L-1)}$**

$$a^{(L-1)} = \sigma(z^{(L-1)}) = \frac{1}{1 + e^{-z^{(L-1)}}} \tag{11}$$
$$a^{(L-1)} = \frac{1}{1 + e^{-0{,}1043}} \tag{12}$$
$$a^{(L-1)} = \frac{1}{1 + 0{,}9009} \tag{13}$$
$$a^{(L-1)} = \frac{1}{1{,}9009} \tag{14}$$

$$\boxed{a^{(L-1)} = 0{,}5260}$$

## 1.3. Capa L (Capa de salida)

**Paso 5: Calcular $z^{(L)}$**

$$z^{(L)} = W^{(L)} \cdot a^{(L-1)} + b^{(L)} \tag{15}$$
$$z^{(L)} = 0{,}23 \cdot 0{,}5260 + 0 \tag{16}$$
$$z^{(L)} = 0{,}1210 \tag{17}$$

**Paso 6: Calcular la predicción $\hat{y} = a^{(L)}$**

$$a^{(L)} = \hat{y} = \sigma(z^{(L)}) = \frac{1}{1 + e^{-z^{(L)}}} \tag{18}$$
$$\hat{y} = \frac{1}{1 + e^{-0{,}1210}} \tag{19}$$
$$\hat{y} = \frac{1}{1 + 0{,}8860} \tag{20}$$
$$\hat{y} = \frac{1}{1{,}8860} \tag{21}$$

$$\boxed{\hat{y} = a^{(L)} = 0{,}5302}$$

## 1.4. Cálculo del Error (Loss)

**Paso 7: Calcular la función de coste**

$$L = C_0 = (\hat{y} - y)^2 \tag{22}$$
$$L = (0{,}5302 - 1)^2 \tag{23}$$
$$L = (-0{,}4698)^2 \tag{24}$$

$$\boxed{L = 0{,}2207}$$

### 1.5. Resumen del Forward Pass

**Valores calculados en el Forward Pass:**

- $z^{(L-2)} = 0{,}32, \quad a^{(L-2)} = 0{,}5793$
- $z^{(L-1)} = 0{,}1043, \quad a^{(L-1)} = 0{,}5260$
- $z^{(L)} = 0{,}1210, \quad \hat{y} = a^{(L)} = 0{,}5302$
- Error (Loss): $L = 0{,}2207$

## 2. Apartado 2: Actualizar $W^{(L)}$ y $b^{(L-1)}$ usando el gradiente (Backward Pass)

En el backward pass, calculamos los gradientes de la función de coste respecto a cada parámetro usando la regla de la cadena, para luego actualizar los pesos.

### 2.1. Gradientes en la Capa L

**Paso 8: Calcular $\frac{\partial L}{\partial a^{(L)}}$**

$$\frac{\partial L}{\partial a^{(L)}} = \frac{\partial}{\partial a^{(L)}}(a^{(L)} - y)^2 \tag{25}$$

$$\frac{\partial L}{\partial a^{(L)}} = 2(a^{(L)} - y) \tag{26}$$

$$\frac{\partial L}{\partial a^{(L)}} = 2(0{,}5302 - 1) \tag{27}$$

$$\frac{\partial L}{\partial a^{(L)}} = 2(-0{,}4698) \tag{28}$$

$$\boxed{\frac{\partial L}{\partial a^{(L)}} = -0{,}9396}$$

**Paso 9: Calcular $\frac{\partial a^{(L)}}{\partial z^{(L)}}$ (derivada de la sigmoide)**

$$\frac{\partial a^{(L)}}{\partial z^{(L)}} = \sigma'(z^{(L)}) = \sigma(z^{(L)}) \cdot (1 - \sigma(z^{(L)})) \tag{29}$$

$$\frac{\partial a^{(L)}}{\partial z^{(L)}} = a^{(L)} \cdot (1 - a^{(L)}) \tag{30}$$

$$\frac{\partial a^{(L)}}{\partial z^{(L)}} = 0{,}5302 \cdot (1 - 0{,}5302) \tag{31}$$

$$\frac{\partial a^{(L)}}{\partial z^{(L)}} = 0{,}5302 \cdot 0{,}4698 \tag{32}$$

$$\boxed{\frac{\partial a^{(L)}}{\partial z^{(L)}} = 0{,}2491}$$

**Paso 10: Calcular $\frac{\partial z^{(L)}}{\partial W^{(L)}}$**

$$z^{(L)} = W^{(L)} \cdot a^{(L-1)} + b^{(L)} \tag{33}$$

$$\frac{\partial z^{(L)}}{\partial W^{(L)}} = a^{(L-1)} \tag{34}$$

$$\boxed{\frac{\partial z^{(L)}}{\partial W^{(L)}} = a^{(L-1)} = 0{,}5260}$$

**Paso 11: Calcular el gradiente $\frac{\partial L}{\partial W^{(L)}}$ usando la regla de la cadena**

$$\frac{\partial L}{\partial W^{(L)}} = \frac{\partial L}{\partial a^{(L)}} \cdot \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial z^{(L)}}{\partial W^{(L)}} \tag{35}$$

$$\frac{\partial L}{\partial W^{(L)}} = (-0{,}9396) \cdot (0{,}2491) \cdot (0{,}5260) \tag{36}$$

$$\frac{\partial L}{\partial W^{(L)}} = -0{,}2340 \cdot 0{,}5260 \tag{37}$$

$$\boxed{\frac{\partial L}{\partial W^{(L)}} = -0{,}1231}$$

## 2.2. Actualización del peso $W^{(L)}$

**Paso 12: Actualizar $W^{(L)}$**

$$W^{(L)}_{nuevo} = W^{(L)} - \alpha \cdot \frac{\partial L}{\partial W^{(L)}} \tag{38}$$

$$W^{(L)}_{nuevo} = 0{,}23 - 0{,}1 \cdot (-0{,}1231) \tag{39}$$

$$W^{(L)}_{nuevo} = 0{,}23 + 0{,}0123 \tag{40}$$

$$\boxed{\textbf{Peso actualizado: } W^{(L)}_{nuevo} = 0{,}2423}$$

## 2.3. Gradientes en la Capa L-1

**Paso 13: Calcular $\frac{\partial z^{(L)}}{\partial a^{(L-1)}}$**

$$\frac{\partial z^{(L)}}{\partial a^{(L-1)}} = W^{(L)} \tag{41}$$

$$\boxed{\frac{\partial z^{(L)}}{\partial a^{(L-1)}} = W^{(L)} = 0{,}23}$$

**Paso 14: Calcular $\frac{\partial L}{\partial a^{(L-1)}}$ (propagación del error)**

$$\frac{\partial L}{\partial a^{(L-1)}} = \frac{\partial L}{\partial a^{(L)}} \cdot \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial z^{(L)}}{\partial a^{(L-1)}} \tag{42}$$

$$\frac{\partial L}{\partial a^{(L-1)}} = (-0{,}9396) \cdot (0{,}2491) \cdot (0{,}23) \tag{43}$$

$$\frac{\partial L}{\partial a^{(L-1)}} = -0{,}2340 \cdot 0{,}23 \tag{44}$$

$$\boxed{\frac{\partial L}{\partial a^{(L-1)}} = -0{,}0538}$$

**Paso 15: Calcular $\frac{\partial a^{(L-1)}}{\partial z^{(L-1)}}$**

$$\frac{\partial a^{(L-1)}}{\partial z^{(L-1)}} = \sigma'(z^{(L-1)}) = a^{(L-1)} \cdot (1 - a^{(L-1)}) \tag{45}$$

$$\frac{\partial a^{(L-1)}}{\partial z^{(L-1)}} = 0{,}5260 \cdot (1 - 0{,}5260) \tag{46}$$

$$\frac{\partial a^{(L-1)}}{\partial z^{(L-1)}} = 0{,}5260 \cdot 0{,}4740 \tag{47}$$

$$\boxed{\frac{\partial a^{(L-1)}}{\partial z^{(L-1)}} = 0{,}2493}$$

**Paso 16: Calcular $\frac{\partial z^{(L-1)}}{\partial b^{(L-1)}}$**

$$z^{(L-1)} = W^{(L-1)} \cdot a^{(L-2)} + b^{(L-1)} \tag{48}$$

$$\frac{\partial z^{(L-1)}}{\partial b^{(L-1)}} = 1 \tag{49}$$

$$\boxed{\frac{\partial z^{(L-1)}}{\partial b^{(L-1)}} = 1}$$

**Paso 17: Calcular el gradiente $\frac{\partial L}{\partial b^{(L-1)}}$**

$$\frac{\partial L}{\partial b^{(L-1)}} = \frac{\partial L}{\partial a^{(L-1)}} \cdot \frac{\partial a^{(L-1)}}{\partial z^{(L-1)}} \cdot \frac{\partial z^{(L-1)}}{\partial b^{(L-1)}} \tag{50}$$

$$\frac{\partial L}{\partial b^{(L-1)}} = (-0{,}0538) \cdot (0{,}2493) \cdot (1) \tag{51}$$

$$\frac{\partial L}{\partial b^{(L-1)}} = -0{,}0134 \tag{52}$$

$$\boxed{\frac{\partial L}{\partial b^{(L-1)}} = -0{,}0134}$$

## 2.4. Actualización del bias $b^{(L-1)}$

**Paso 18: Actualizar $b^{(L-1)}$**

$$b^{(L-1)}_{nuevo} = b^{(L-1)} - \alpha \cdot \frac{\partial L}{\partial b^{(L-1)}} \tag{53}$$

$$b^{(L-1)}_{nuevo} = 0 - 0{,}1 \cdot (-0{,}0134) \tag{54}$$

$$b^{(L-1)}_{nuevo} = 0 + 0{,}00134 \tag{55}$$

**Bias actualizado:** $b^{(L-1)}_{nuevo} = 0{,}00134$

# 3. Resumen Final de Resultados

**Forward Pass - Valores calculados:**

- $z^{(L-2)} = 0{,}32, \quad a^{(L-2)} = 0{,}5793$
- $z^{(L-1)} = 0{,}1043, \quad a^{(L-1)} = 0{,}5260$
- $z^{(L)} = 0{,}1210, \quad \hat{y} = a^{(L)} = 0{,}5302$
- Error (Loss): $L = 0{,}2207$

**Backward Pass - Gradientes calculados:**

- $\frac{\partial L}{\partial W^{(L)}} = -0{,}1231$
- $\frac{\partial L}{\partial b^{(L-1)}} = -0{,}0134$

**Parámetros actualizados (con $\alpha = 0{,}1$):**

- $W^{(L)}_{nuevo} = 0{,}2423$
- $b^{(L-1)}_{nuevo} = 0{,}00134$