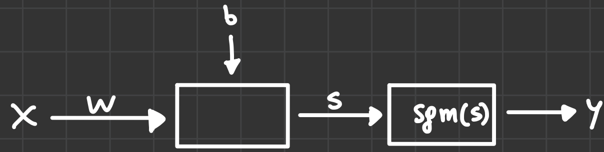
Joaquín Zepeda
20.127.150-9



nota: $y = sgm(s) = \dfrac{1}{1+\exp(-s)}$, donde $s = w \cdot x + b$

Parte a: Derive una regla de aprendizaje por gradiente descendente para el ajuste del peso w y del bias b.

De las clases: $W_{k+1} = W_k + \mu \cdot (-\nabla_k)$

① $\dfrac{\partial y}{\partial b} = \dfrac{\exp(-(wx+b))}{(\exp(-(wx+b))+1)^2} = y^2 \cdot \exp(-(wx+b))$

② $\dfrac{\partial y}{\partial w} = \dfrac{\exp(-(wx+b))}{(\exp(-(wx+b))+1)^2} \cdot x = y^2 \cdot \exp(-(wx+b)) \cdot x$

1) Usando el error cuadrático

$$E(w,b) = (t-y)^2 = \left(t - \frac{1}{1+\exp(-(wx+b))}\right)^2$$

$$\boxed{b_{k+1} = b_k - \mu \cdot \frac{\partial E(w,b)}{\partial b}} \; *$$

$$y = \frac{1}{1+\exp(-(wx+b))}$$

Luego reemplazando:

$$b_{k+1} = b_k - \mu \cdot 2 \cdot (t-y) \cdot \left(-\frac{\partial y}{\partial b}\right)$$

$$b_{k+1} = b_k + \mu \cdot 2 \cdot (t-y) \cdot \frac{\exp(-(wx+b))}{(\exp(-(wx+b))+1)^2}$$

$$b_{k+1} = b_k + 2 \cdot \mu \cdot (t-y) \cdot y^2 \cdot \exp(-(wx+b))$$

ahora para los pesos:

$$\boxed{W_{k+1} = W_k + \mu \cdot (-\nabla_k)} \; *$$

$$W_{k+1} = W_k - \mu \cdot 2 \cdot (t-y) \cdot \left(-\frac{\partial y}{\partial w}\right)$$

$$W_{k+1} = W_k + 2\cdot\mu\cdot(t-y)\cdot\frac{\exp(-(wx+b))}{(\exp(-(wx+b))+1)^2}\cdot x$$

$$\Leftrightarrow W_{k+1} = W + 2\cdot\mu\cdot(t-y)\cdot y^2\cdot\exp(-(w\cdot x+b)\cdot x$$

En Resumen, las reblas para el ajuste del Bias y de los pesos son:

b	$b_{k+1} = b + \mu\cdot 2\cdot\left(t - \frac{1}{1+\exp(-(wx+b))}\right)\cdot\frac{\exp(-(wx+b))}{(\exp(-(wx+b))+1)^2}$
W	$W_{k+1} = W + 2\mu\cdot\left(t - \frac{1}{1+\exp(-(wx+b))}\right)\cdot\frac{\exp(-(wx+b))}{(\exp(-(wx+b))+1)^2}\cdot x$

o De forma más simplificada:

b	$b_{k+1} = b + 2\mu\cdot(t-y)\cdot y^2\cdot\exp(-(wx+b))$
W	$W_{k+1} = W + 2\cdot\mu\cdot(t-y)\cdot y^2\cdot\exp(-(wx+b))\cdot x$

2) Usando la entropía cruzada

$$E(w,b) = -t\cdot\ln(y) - (1-t)\ln(1-y)$$

primero para tener una idea de como es la Derivada, Derivemos con respecto a y:

$$\frac{dE(y)}{dy} = -t\cdot\frac{1}{y}\cdot y' - (1-t)\cdot\frac{1}{(1-y)}\cdot(-y')$$

Con $y = \frac{1}{1+\exp(-(wx+b))}$

$$\boxed{b_{k+1} = b - \mu \cdot \frac{\partial E(w,b)}{\partial b}}$$

$$\boxed{W_{k+1} = W - \mu \cdot (\nabla_k)}$$

Para el bias (b):

$$\frac{\partial E(w,b)}{\partial b} = -t \cdot \frac{1}{y} \cdot \frac{\partial y}{\partial b} - (1-t) \cdot \frac{1}{(1-y)} \cdot \left(\frac{-\partial y}{\partial b} \right)$$

POR 1

$$= -t \cdot \frac{1}{y} \cdot y^2 \cdot exp(-(wx+b))$$

$$+ (1-t) \cdot \frac{1}{(1-y)} \cdot y^2 \cdot exp(-(wx+b))$$

$$= -t \cdot y \cdot exp(-(wx+b)) + \frac{(1-t) \cdot y^2}{(1-y)} \cdot exp(-(wx+b))$$

<u>Para los pesos</u>:

$$\frac{\partial E(w,b)}{\partial w} = -t \cdot \frac{1}{y} \cdot \frac{\partial y}{\partial w} + (1-t) \cdot \frac{1}{(1-y)} \cdot \left(\frac{\partial y}{\partial w} \right)$$

POR 2

$$= -t \cdot \frac{1}{y} \cdot y^2 \cdot exp(-(wx+b)) \cdot x$$

$$+ (1-t) \cdot \frac{1}{(1-y)} \cdot y^2 \cdot exp(-(wx+b)) \cdot x$$

$$= -t \cdot y \cdot exp(-(wx+b)) \cdot x +$$

$$+ (1-t) \cdot \frac{1}{(1-y)} \cdot y^2 \cdot exp(-(wx+b)) \cdot x$$

En Resumen:

$$\boxed{\begin{aligned} b_{k+1} &= b - \mu\cdot\left(-t\cdot y\cdot \exp(-(wx+b)) + \frac{(1-t)\cdot y^2}{(1-y)}\cdot \exp(-(wx+b))\right) \\ w_{k+1} &= w - \mu\cdot\Big(-t\cdot y\cdot \exp(-(wx+b))\cdot x + \\ &\quad + (1-t)\cdot\frac{1}{(1-y)}\cdot y^2\cdot \exp(-(wx+b))\cdot x\Big) \end{aligned}}$$