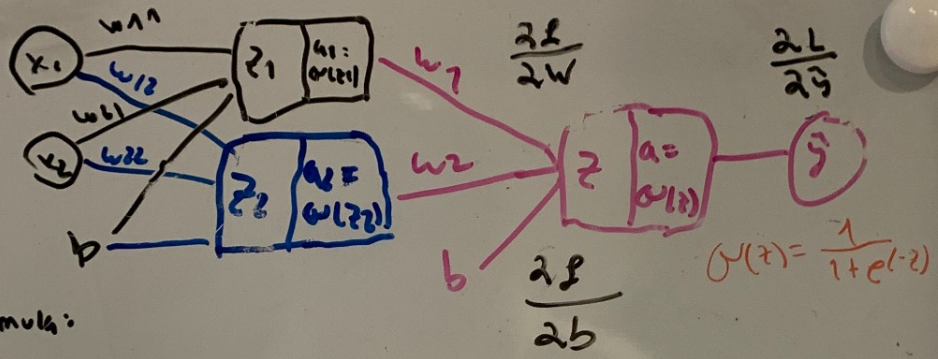
$$\hat{y} = wx + b$$

$$\mathcal{L}(w,b) = -y\ln(\hat{y}) - (1-y)\ln(1-\hat{y})$$



Formula:

$$z_1 = w_{11}\cdot x_1 + w_{21}\cdot x_2 + b$$

$$a_1 = \sigma(z_1)$$

$$z_2 = w_{12}\cdot x_1 + w_{22}\cdot x_2 + b$$

$$a_2 = \sigma(z_2)$$

$$z = w_1\cdot a_1 + w_2\cdot a_2 + b$$

$$y = \sigma(z)$$

Gradient Descent First Layer

$$\frac{\partial f}{\partial w} = \frac{\partial z}{\partial w}\cdot\frac{\partial a}{\partial z}\cdot\frac{\partial L}{\partial a}$$

$$\frac{\partial f}{\partial b} = \frac{\partial z}{\partial b}\cdot\frac{\partial a}{\partial z}\cdot\frac{\partial f}{\partial a}$$

Gradient Descent Second Layer

$$\frac{\partial f}{\partial w} = \frac{\partial z}{\partial w}\cdot\frac{\partial a}{\partial z}\cdot\frac{\partial z}{\partial a}\cdot\frac{\partial a}{\partial z}\cdot\frac{\partial f}{\partial a}$$

$$\frac{\partial f}{\partial b} = \frac{\partial z}{\partial b}\cdot\frac{\partial a}{\partial z}\cdot\frac{\partial z}{\partial a}\cdot\frac{\partial a}{\partial z}\cdot\frac{\partial f}{\partial a}$$

Employee promotion description (3/9/25)

x_1 : input : year of experiences

x_2 : input : Performance score

$\hat{y}$ 70.5 = Promotion $\quad L(w,b) = (-y \log(\hat{y}) - (1-y)\log(1-\hat{y}))$

$\hat{y} = wx + b$

Formula:

$z_1 = w_{11} \cdot x_1 + w_{21} \cdot x_2 + b$

$a_1 = \sigma(z_1)$

$z_2 = w_{12} \cdot x_1 + w_{22} \cdot x_2 + b$

$a_2 = \sigma(z_2)$

$z_f = w_1 * a_1 + w_2 \cdot a_2 + b$

$\rightarrow \hat{y} = \sigma(z_f) \leftarrow$

Partial derivations

Gradient Descent First layer $\quad a = \hat{y}$

$$\frac{\partial f}{\partial w} = \frac{\partial z}{\partial w} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial f}{\partial a}$$

$$\frac{\partial f}{\partial b} \cdot \frac{\partial z}{\partial b} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial f}{\partial a}$$

$$\frac{\partial f}{\partial a} = \frac{-y}{\hat{y}} + \frac{(1-y)}{(1-\hat{y})} = \frac{-(y-\hat{y})}{\hat{y}(1-\hat{y})}$$

$\rightarrow \frac{\partial a}{\partial z} \leftarrow = a(1-a) \qquad \frac{\partial f}{\partial w} = x \cdot \cancel{a(1-a)} \cdot \frac{-(y-\hat{y})}{\cancel{\hat{y}(1-\hat{y})}}$

$\frac{\partial z}{\partial w} = x_n \qquad \frac{\partial f}{\partial w} = -(y-\hat{y})x_n$

$$\frac{\partial f}{\partial b} = -(y-\hat{y}) \cdot 1$$

Back Propasation second layer $\hat{y} = a$

$$\frac{\partial L}{\partial w^{[1]}} = \frac{\partial z}{\partial w} \cdot \frac{\partial a^{[1]}}{\partial z^{[1]}} \cdot \frac{\partial z^{[2]}}{\partial a^{[1]}} \cdot \frac{\partial a^{[2]}}{\partial z^{[2]}} \cdot \frac{\partial L}{\partial a^{[2]}}$$

$$\frac{\partial L}{\partial b^{[1]}} = \frac{\partial z}{\partial b} \cdot \frac{\partial a^{[1]}}{\partial z^{[1]}} \cdot \frac{\partial z^{[2]}}{\partial a^{[1]}} \cdot \frac{\partial a^{[2]}}{\partial z^{[2]}} \cdot \frac{\partial L}{\partial a^{[2]}}$$

$\frac{\partial a}{\partial z} = a_1(1-a) = \hat{y}(1-\hat{y})$ $\quad \frac{\partial z^{[2]}}{\partial a^{[1]}} = w^{[2]}$

$\frac{\partial z}{\partial w} = x_n$ $\quad \frac{\partial z}{\partial b} = 1$

$\frac{\partial L}{\partial a} = \frac{-(y-\hat{y})}{\hat{y}(1-\hat{y})}$

$\frac{\partial L}{\partial w^{[1]}} = x \cdot a_1(1-a_1) \cdot w \cdot a_2(1-a_2) \cdot \frac{-(y-\hat{y})}{\hat{y}(1-\hat{y})}$

$\frac{\partial L}{\partial w^{[1]}} = -x \cdot w \cdot a_1(1-a_1) \cdot (y-\hat{y})$

$\frac{\partial L}{\partial b} = -1 \cdot a_1(1-a_1) \cdot w \cdot a_2(1-a_2) \cdot (y-\hat{y})$

$\frac{\partial L}{\partial b} = -1 \cdot a_1(1-a_1) \cdot w \cdot (y-\hat{y})$

$w = w - \alpha(-x \cdot w \cdot a_1(1-a_1) \cdot (y-\hat{y}))$

$b = b - \alpha(-1 \cdot w \cdot a_1(1-a_1) \cdot (y-\hat{y}))$

α = learning Rate