Basics of Neural Network Programming

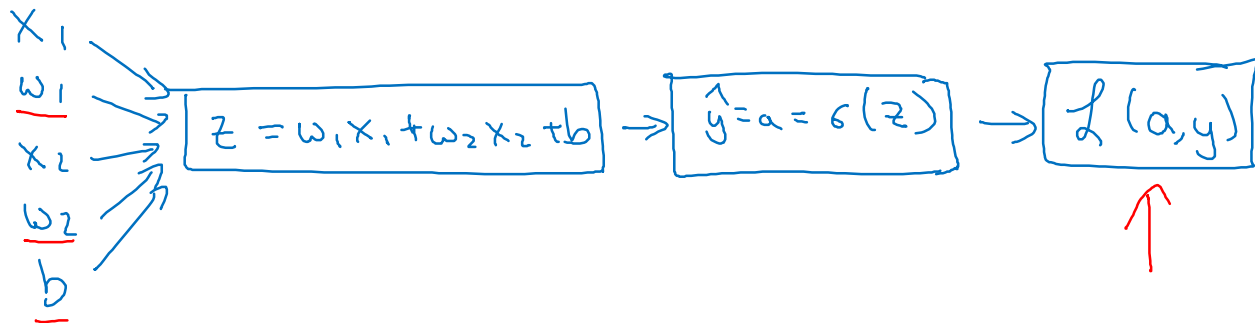## Logistic Regression Gradient descent

Logistic regression recap

$$z = w^T x + b$$

$$\hat{y} = a = \sigma(z)$$

$$\mathcal{L}(a, y) = -(y \log(a) + (1 - y) \log(1 - a))$$

x_1
w_1
x_2
w_2
b

$$z = w_1 x_1 + w_2 x_2 + b \rightarrow \hat{y} = a = \sigma(z) \rightarrow \mathcal{L}(a, y)$$

Logistic regression derivatives

x_1
w_1
x_2
w_2
b

$$z = w_1 x_1 + w_2 x_2 + b$$

$$a = \sigma(z)$$

$$\mathcal{L}(a, y)$$

$$dz = \frac{d\mathcal{L}}{dz} = \frac{d\mathcal{L}(a,y)}{dz}$$

$$"da" = \frac{d\mathcal{L}(a,y)}{da}$$

$$= a - y$$

$$= \frac{d\mathcal{L}}{da} \cdot \frac{da}{dz}$$

$$a(1-a)$$

$$= -\frac{y}{a} + \frac{1-y}{1-a}$$

$$\frac{d\mathcal{L}}{dw_1} = "dw_1" = x_1 \cdot dz.$$

$$dw_2 = x_2 \cdot dz.$$

$$db = dz.$$

$$w_1 := w_1 - \alpha \, dw_1$$
$$w_2 := w_2 - \alpha \, dw_2$$
$$b := b - \alpha \, db.$$

Andrew Ng