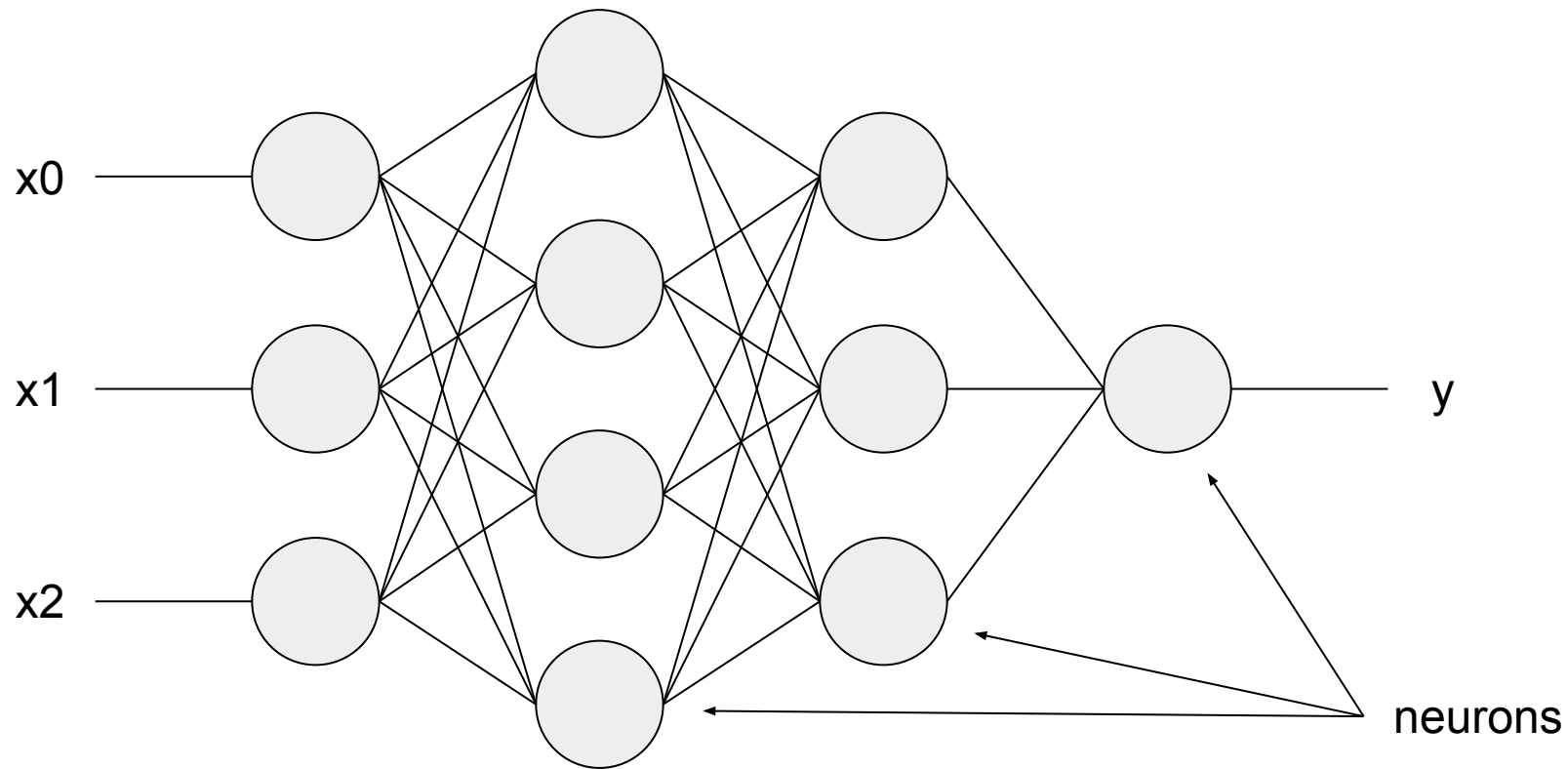


Artificial Neural Network

inputs

layers

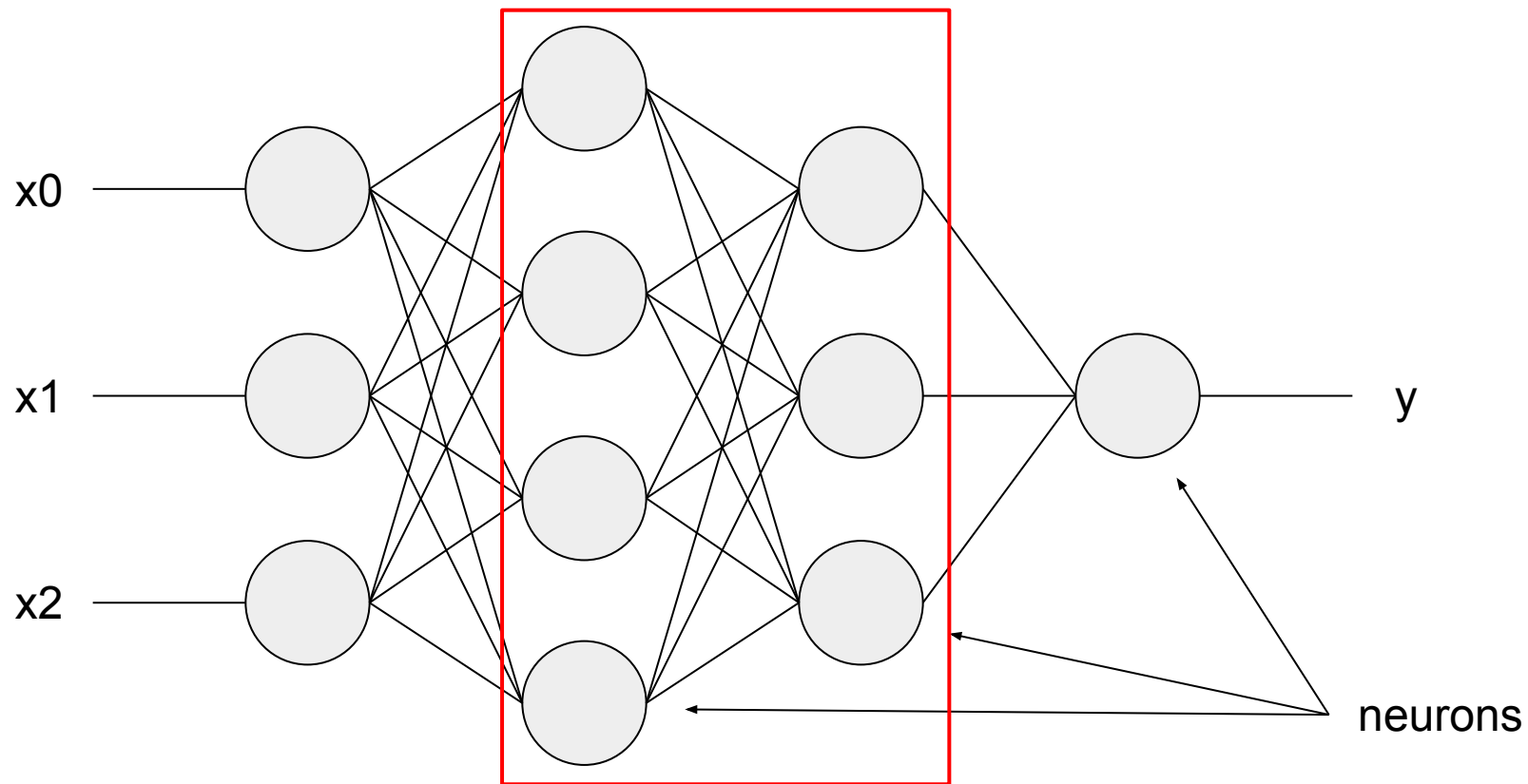
output

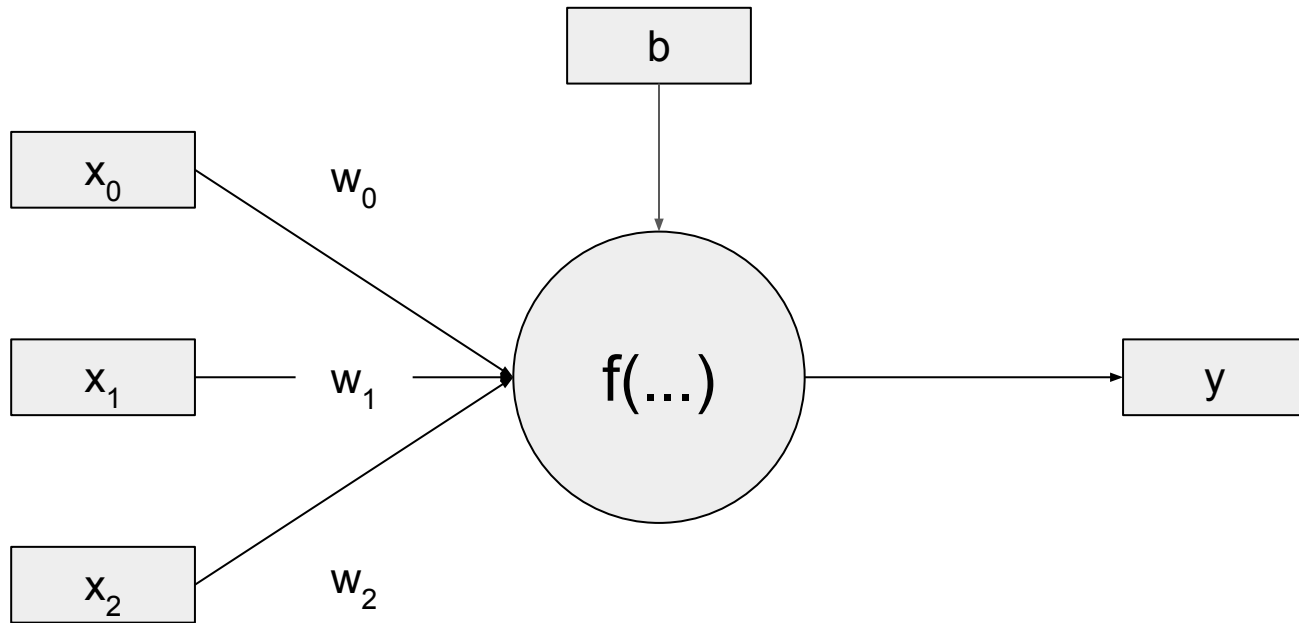


inputs

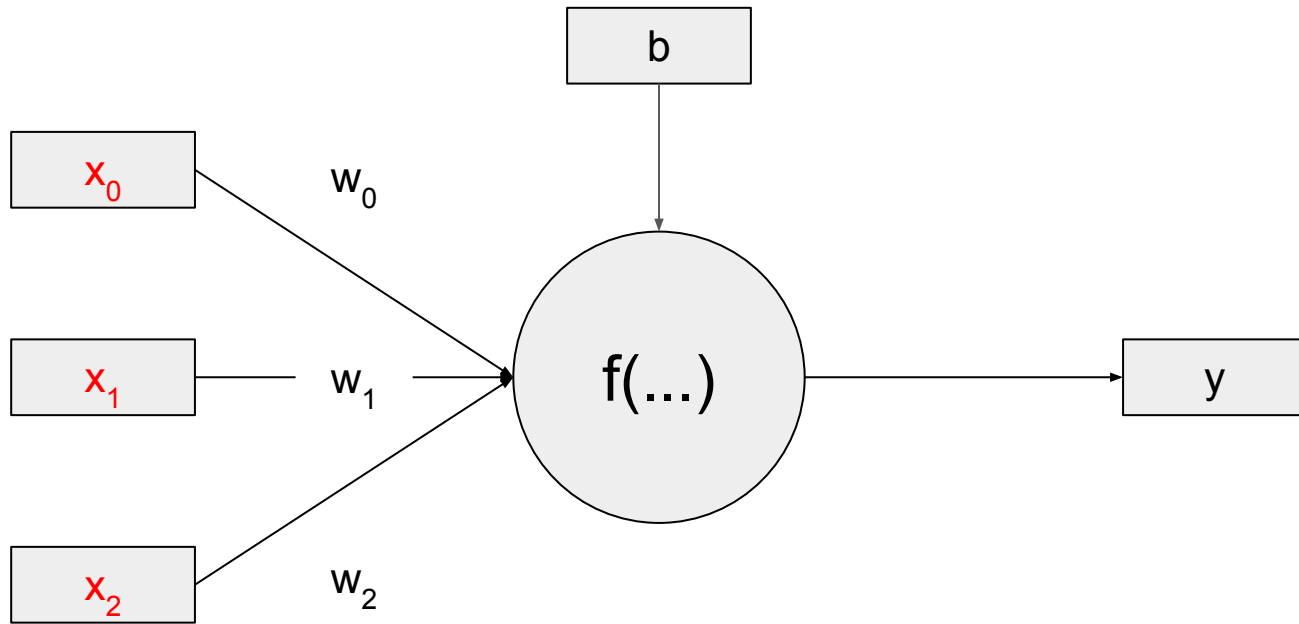
layers

output

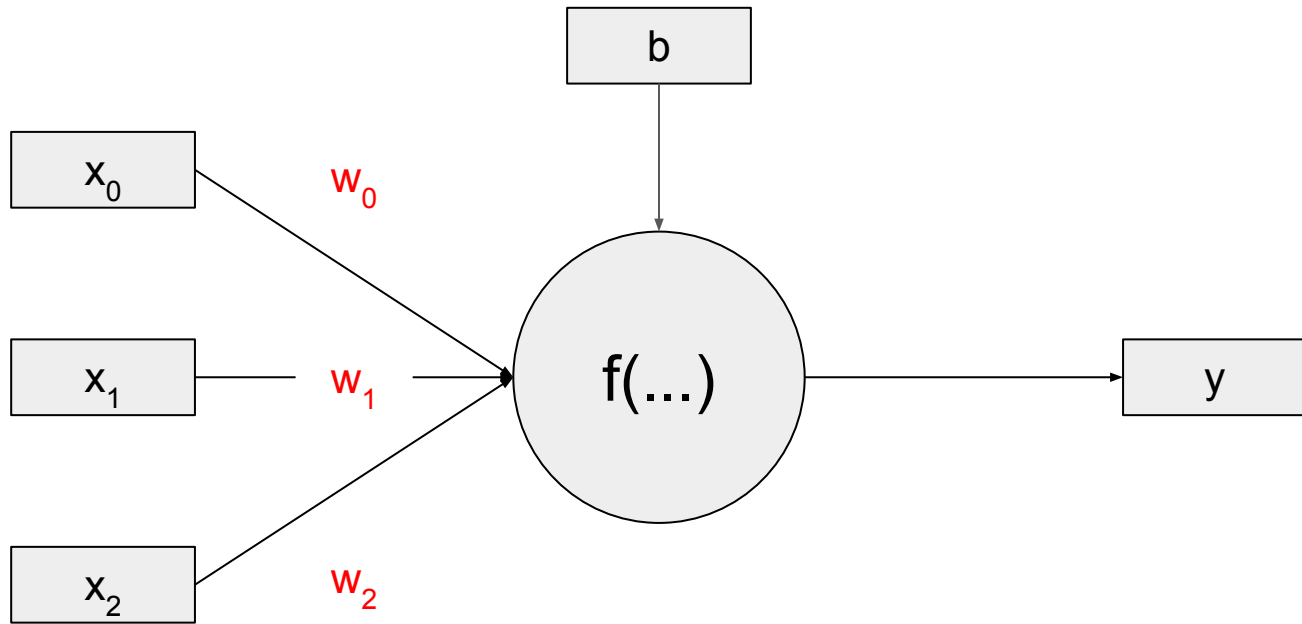




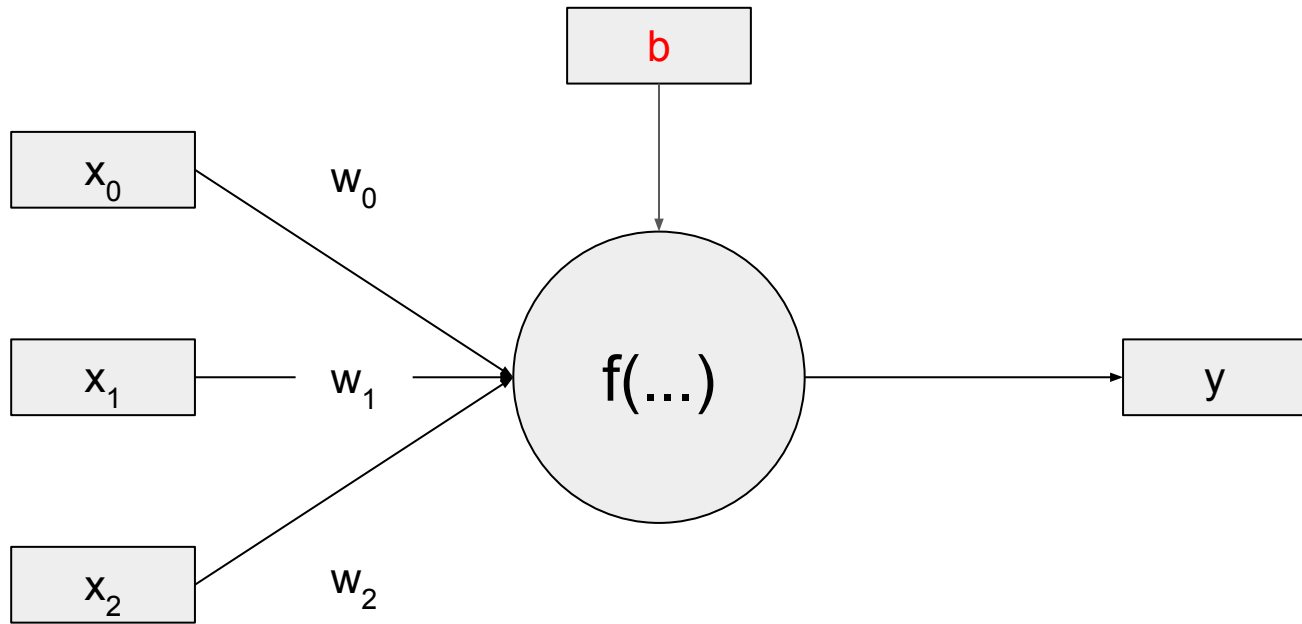
$$y = f(x_0 * w_0 + x_1 * w_1 + x_2 * w_2 + b)$$



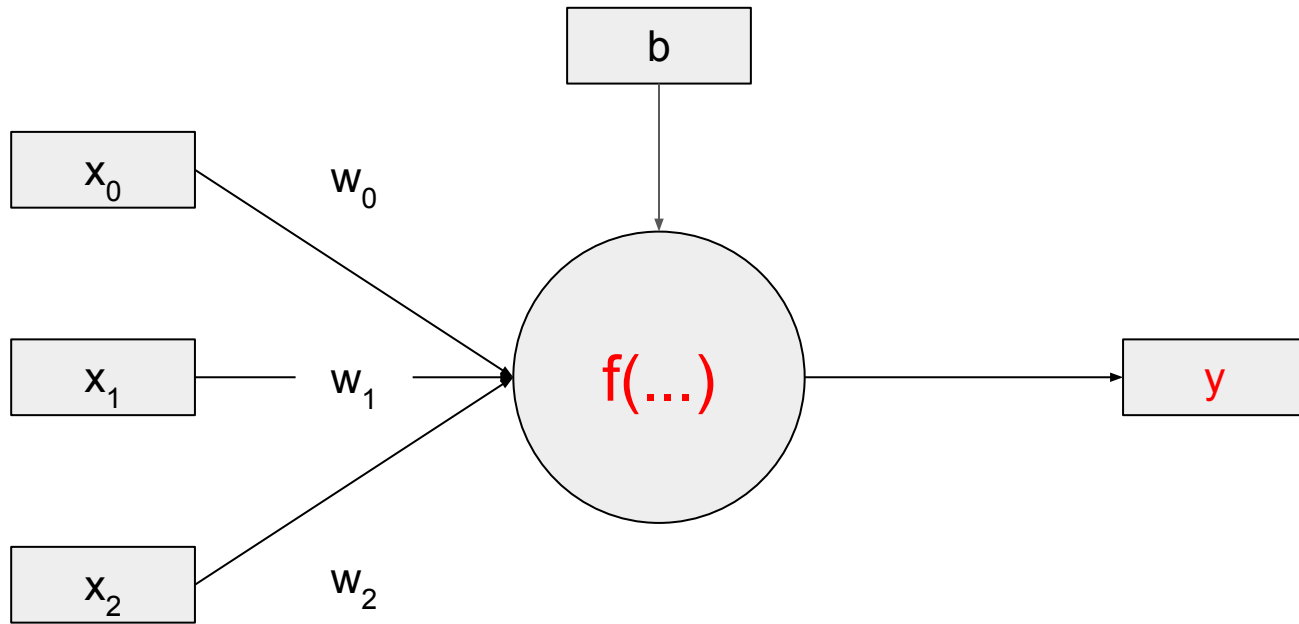
$$y = f(x_0 * w_0 + x_1 * w_1 + x_2 * w_2 + b)$$



$$y = f(x_0 * w_0 + x_1 * w_1 + x_2 * w_2 + b)$$



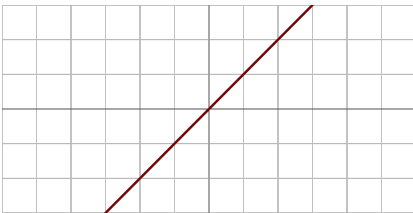
$$y = f(x_0 * w_0 + x_1 * w_1 + x_2 * w_2 + b)$$



$$y = f(x_0 * w_0 + x_1 * w_1 + x_2 * w_2 + b)$$

Identity

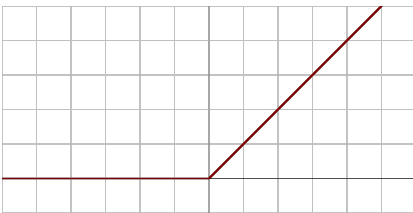
range: $(-\infty, \infty)$



$$f(x) = x$$

Rectified linear unit (ReLU)

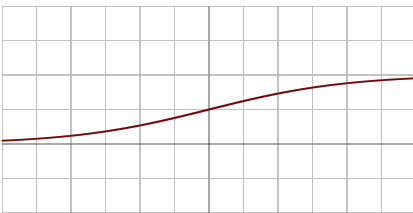
range: $[0, \infty)$



$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$$

Logistic, sigmoid

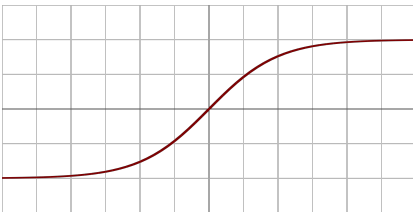
range: $(0, 1)$



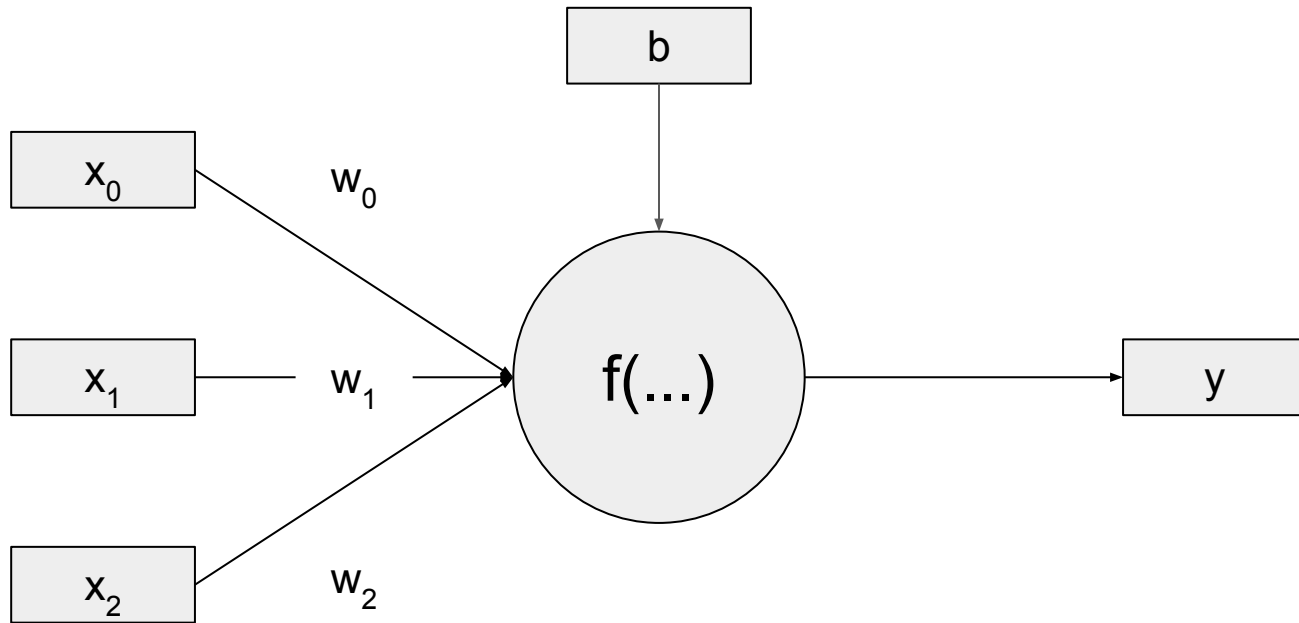
$$f(x) = \frac{1}{1 + e^{-x}}$$

Hyperbolic tangent

range: $(-1, 1)$

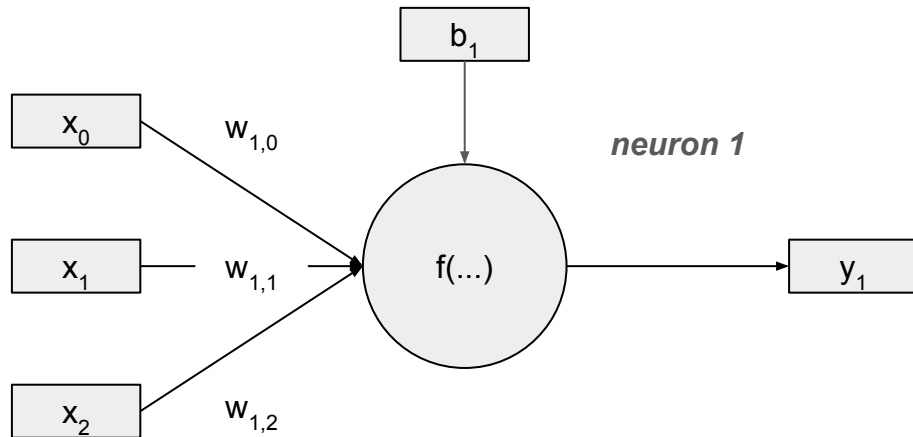
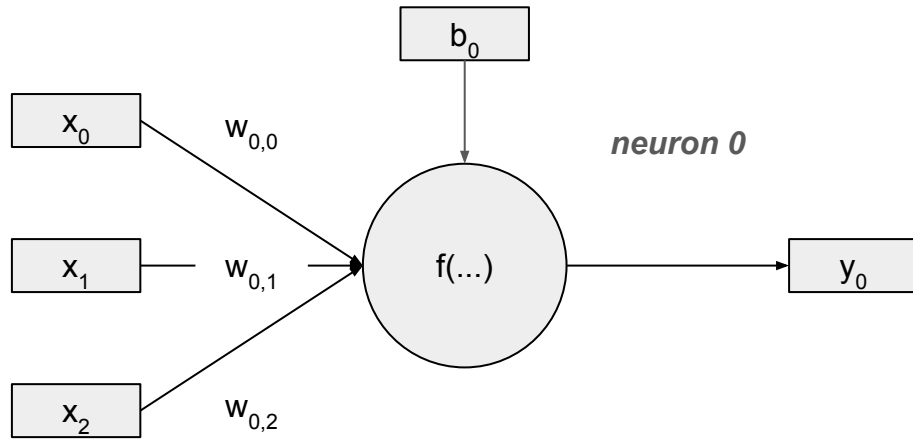


$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



$$Y = f(W \bullet X + b)$$

$$X = [x_0, x_1, x_2], W = [w_0, w_1, w_2]$$



$$\mathbf{Y} = \mathbf{f}(\mathbf{W} \bullet \mathbf{X}^T + \mathbf{B}^T)$$

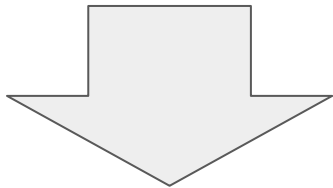
$$\mathbf{Y} = [y_0, y_1]$$

$$\mathbf{X} = [x_0, x_1, x_2]$$

$$\mathbf{W} = \begin{bmatrix} w_{0,0} & w_{0,1} & w_{0,2} \\ w_{1,0} & w_{1,1} & w_{1,2} \end{bmatrix}$$

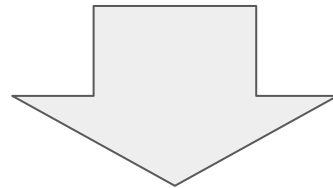
$$\mathbf{B} = [b_0, b_1]$$

$$X = \begin{bmatrix} x_0 & x_1 & x_2 \end{bmatrix}$$



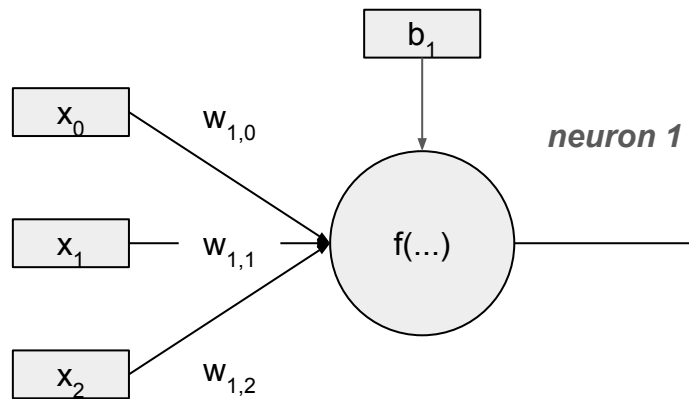
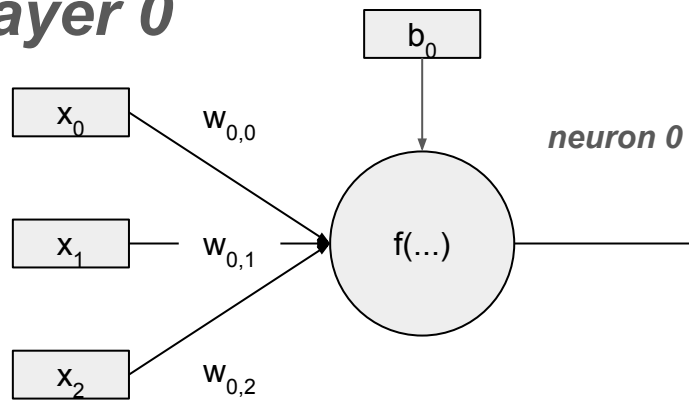
$$X^T = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}$$

$$W = \begin{bmatrix} w_{0,0} & w_{0,1} & w_{0,2} \\ w_{1,0} & w_{1,1} & w_{1,2} \end{bmatrix}$$

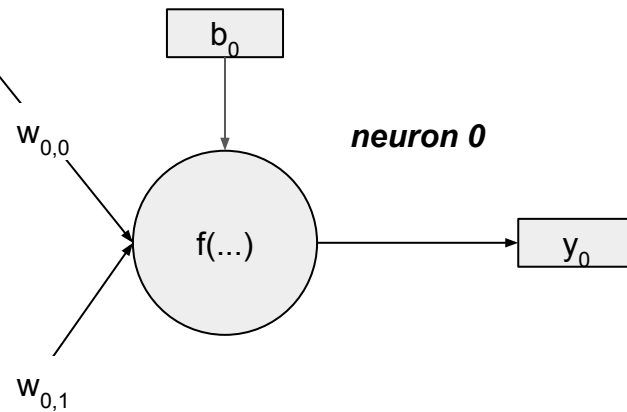


$$W^T = \begin{bmatrix} w_{0,0} & w_{1,0} \\ w_{0,1} & w_{1,1} \\ w_{0,2} & w_{1,2} \end{bmatrix}$$

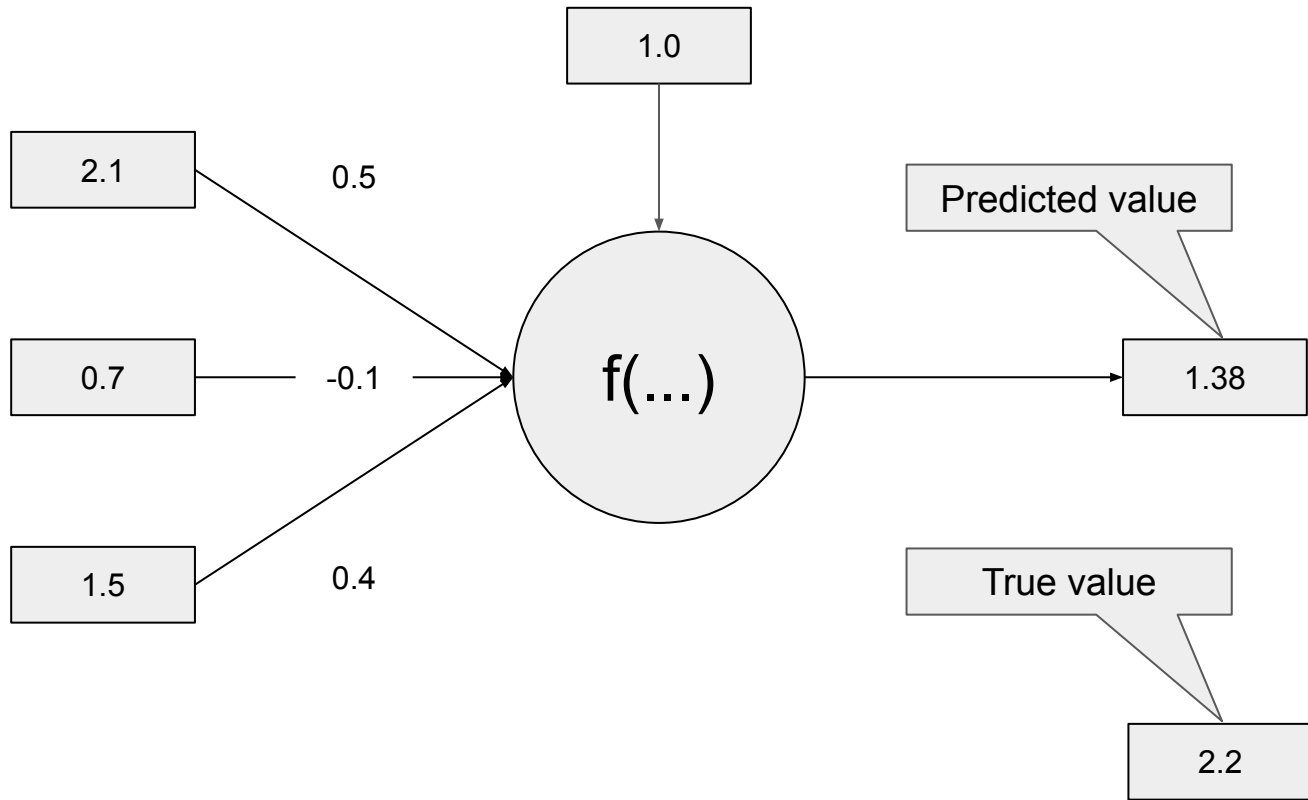
layer 0



layer 1



Forward propagation



$$\text{error} = y^{\text{pred}} - y^{\text{true}} = -0.82$$

$$W_{\text{new}} = W_{\text{old}} - \Delta W$$

$$b_{\text{new}} = b_{\text{old}} - \Delta b$$

$$\Delta W = ??$$

$$\Delta b = ??$$

$$z = W \bullet X + b$$

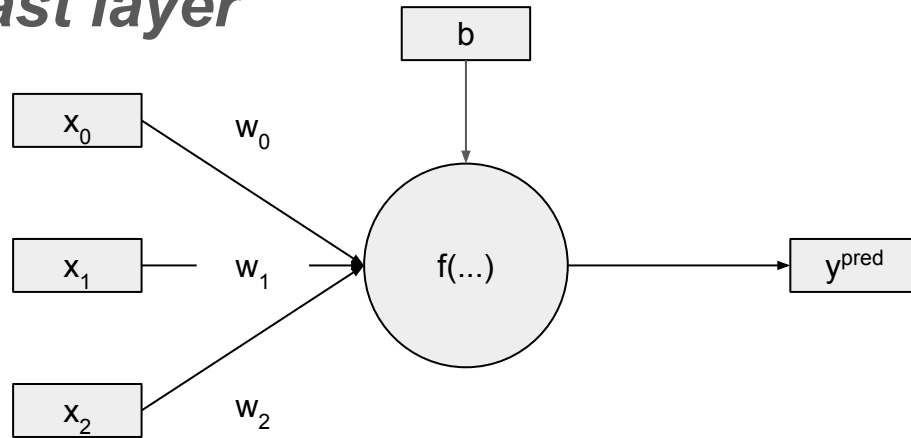
$$y^{\text{pred}} = f(z)$$

$$\delta = (y^{\text{pred}} - y^{\text{true}}) * f'(z)$$

$$\Delta W = LR * \delta * X$$

$$\Delta b = LR * \delta$$

last layer



$$z = W \bullet X + b$$

$$\Delta W = LR * \delta * X$$

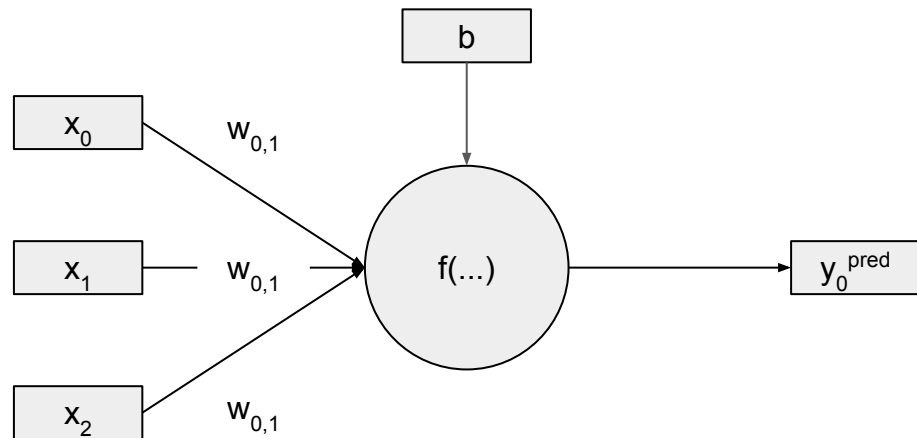
$$y^{\text{pred}} = f(z)$$

$$\Delta b = LR * \delta$$

$$\delta = (y^{\text{pred}} - y^{\text{true}}) * f'(z)$$

$$\delta = \sum_l (\text{delta}_l * W_l) * f'(z)$$

hidden layer (0)



$$z = W \bullet X + b$$

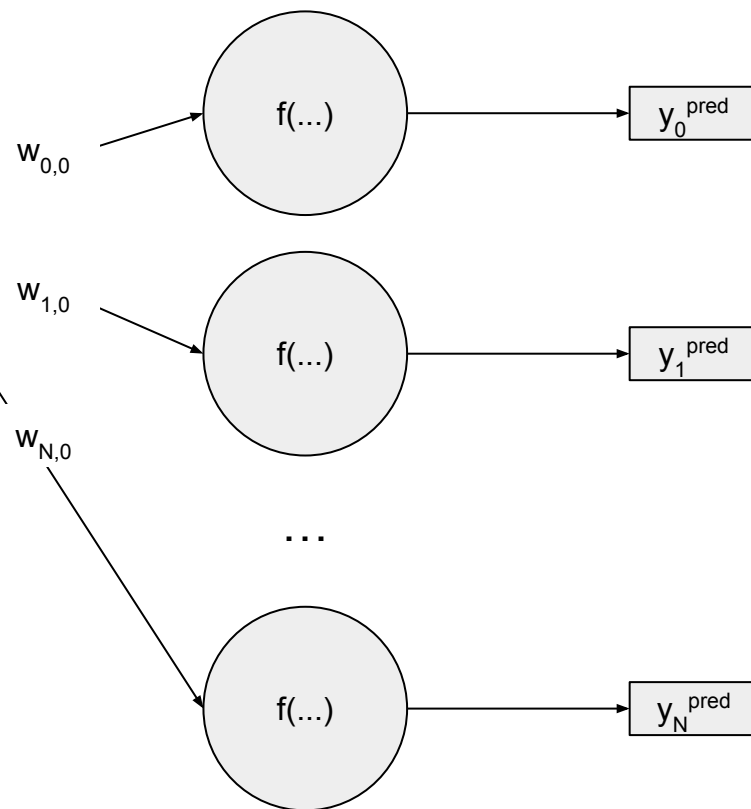
$$\Delta w = LR * \delta * X$$

$$y^{\text{pred}} = f(z)$$

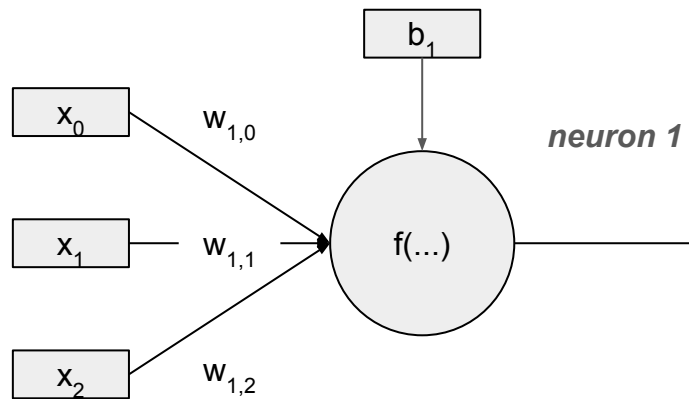
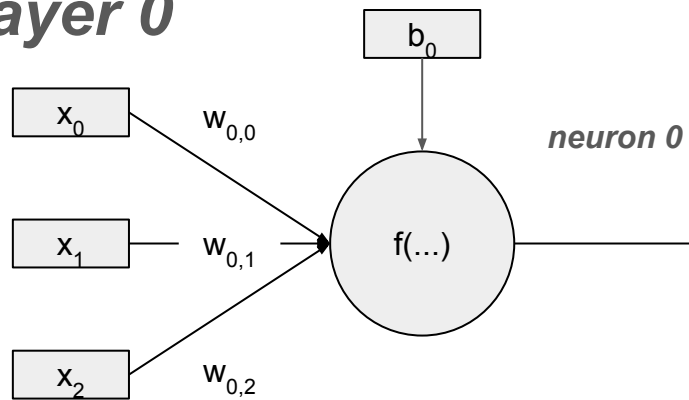
$$\Delta b = LR * \delta$$

$$\delta = \sum_i (\text{delta}_i * W_i) * f'(z)$$

last layer (1)



layer 0



layer 1

