Gradient Computation: Worked Example

Simple network: 2 inputs, 1 hidden neuron, 1 output

$x_1 = 0.5$

$w_1 = 0.4$

$w_3 = 0.8$

h

$\hat{y}$

$y = 1$

$x_2 = 0.3$

$w_2 = 0.6$
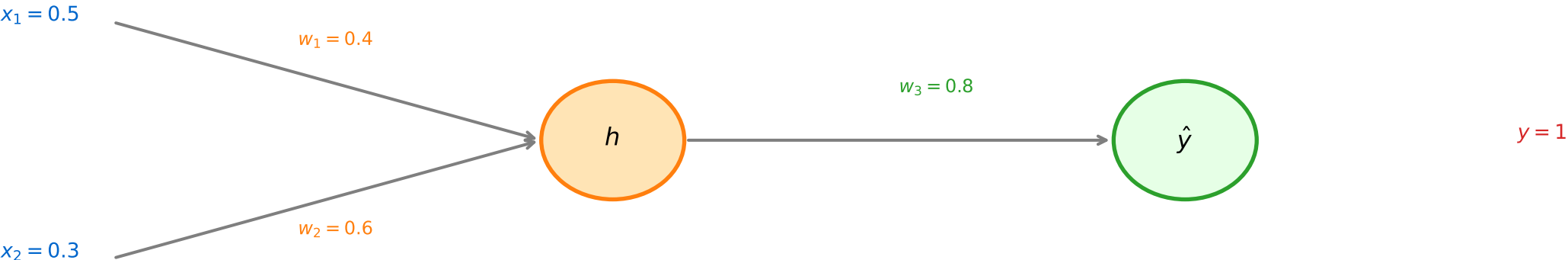
Step-by-Step Calculation

FORWARD:

$z_h = w_1 x_1 + w_2 x_2 = 0.4(0.5) + 0.6(0.3) = 0.38$

$h = \sigma(z_h) = \sigma(0.38) = 0.594$

$z_o = w_3 h = 0.8(0.594) = 0.475$

$\hat{y} = \sigma(z_o) = \sigma(0.475) = 0.617$

BACKWARD:

$L = \frac{1}{2}(y - \hat{y})^2 = 0.073$

$\delta_o = (\hat{y} - y) \cdot \sigma'(z_o) = -0.091$

$\delta_h = \delta_o \cdot w_3 \cdot \sigma'(z_h) = -0.017$

GRADIENTS:

$\frac{\partial L}{\partial w_3} = \delta_o \cdot h = -0.054$

$\frac{\partial L}{\partial w_1} = \delta_h \cdot x_1 = -0.009$

$\frac{\partial L}{\partial w_2} = \delta_h \cdot x_2 = -0.005$

UPDATE ($\eta = 0.1$):

$w_3 \leftarrow 0.8 - 0.1(-0.054) = 0.805$