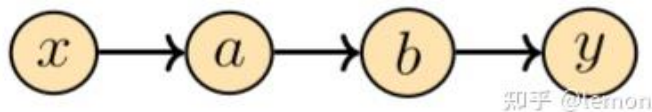


## 参数优化：误差反向传播 (error Back Propagation, BP)

- 神经网络将输入数据 $x$ 经过层层计算, 最终得到预测输出 $y$ 的过程叫做正向传播(Forward Propagation), 这个方向我们也称为数据流动(Data Flow)的方向。



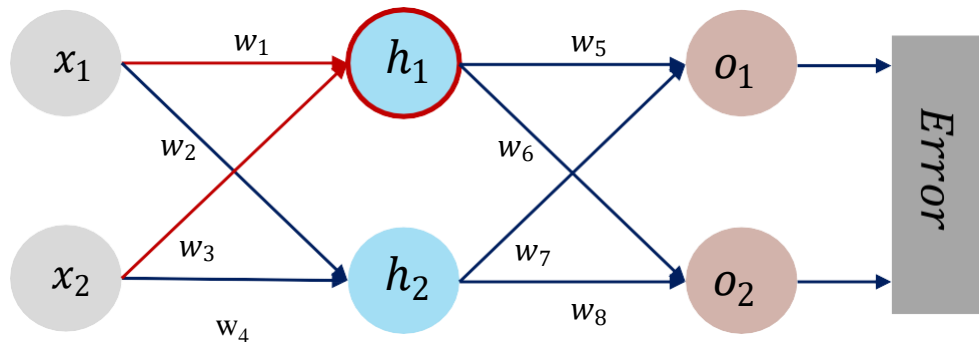
- 如果我们想计算梯度 $\frac{dy}{dx}$ , 则由链式法则得:  $\frac{dy}{dx} = \frac{dy}{db} \frac{db}{da} \frac{da}{dx}$  显然, 计算这个梯度有两种不同的计算方法:

★反向传播误差=反向传播梯度

$$\begin{aligned} \text{反向模式} & \xrightarrow{\quad} \frac{dy}{dx} = \left( \frac{dy}{db} \frac{db}{da} \right) \frac{da}{dx}, \\ \frac{dy}{dx} &= \frac{dy}{db} \left( \frac{db}{da} \frac{da}{dx} \right). \end{aligned}$$

← 正向模式

## 前馈神经网络误差反向传播：



**正向传播：** 隐藏层  $h_1$ 、 $h_2$  和输出层  $o_1$ 、 $o_2$  的激活函数均为 sigmoid 函数。输入输出函数用  $in$ 、 $out$  表示，即  $In_{h_1}$ 、 $In_{h_2}$ 、 $In_{o_1}$ 、 $In_{o_2}$ 、 $Out_{h_1}$ 、 $Out_{h_2}$ 、 $Out_{o_1}$ 、 $Out_{o_2}$ ，它们关系如下 (以  $h_1$  为例)：

$$In_{h_1} = w_1 \times x_1 + w_3 \times x_2; \quad Out_{h_1} = Sigmoid(In_{h_1})$$

**误差函数为：**

$$Error = \frac{1}{2} \sum_{i=1}^2 (o_i - y_i)^2$$

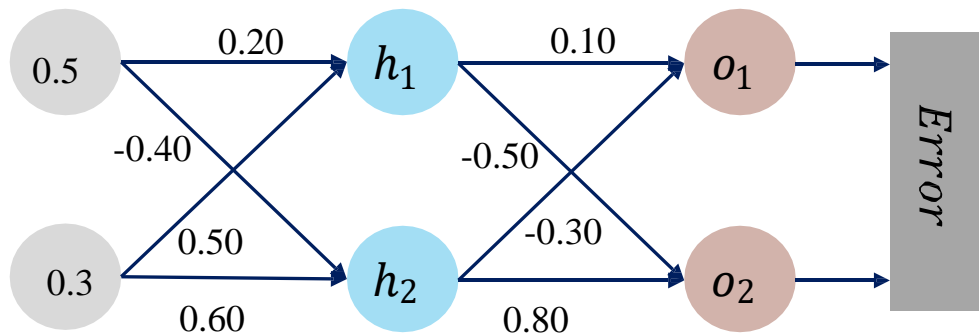
**题目1：** 用梯度计算写出反向传播公式，即写出  $w_1 \sim w_8$  的梯度表达式，如  $\delta_1 = \frac{\partial Error}{\partial w_1} = \dots$ ，并考察  $\delta_i$  之间有无关联。

## 前馈神经网络误差反向传播：

权值更新公式： $w_i' = w_i - \eta \times \delta_i$

↑      ↑      ↑  
原有参数   步长   传递误差

参数初始化：



其中， $Error = 0.5 \times (o_1 - 0.23)^2 + 0.5 \times (o_2 - (-0.07))^2$ ，这里0.23和-0.07是对输入样本数据(0.5, 0.3)的标注信息。

**题目2：**根据权值更新公式和初始化的参数，求出网络的所有变量和权值，即  $h_1$  ,  $h_2$  ,  $o_1$  ,  $o_2$  ,  $Error$  ,  $\delta_i$  ,  $w_1-w_8$  。注： $\eta = 1$ 。

更新后的参数，仅供参考。

