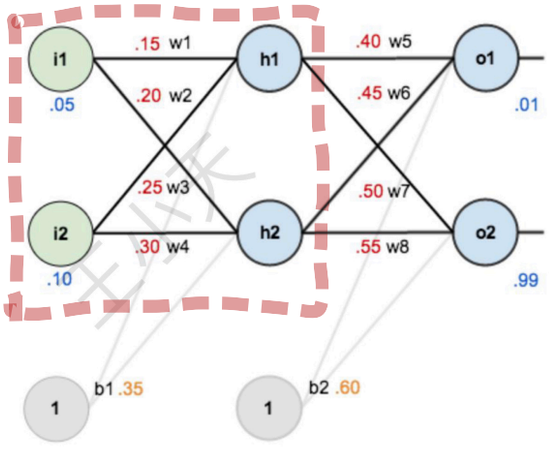
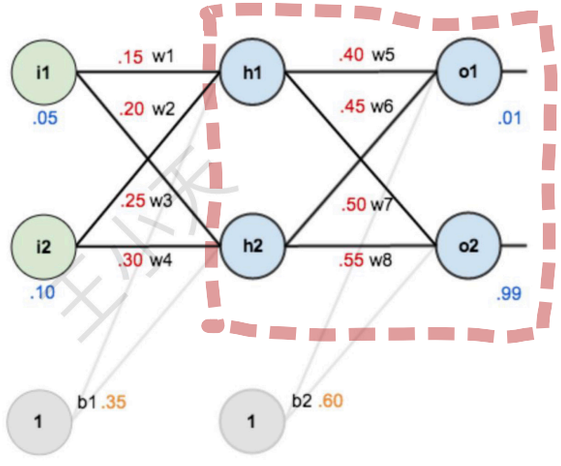
手推神经网络训练过程



激活函数：sigmod

加权输入和：z

神经元输出：a.



Step 1 前向传播

1. 输入层 → 隐藏层

$$z_{h_1} = w_1 \times i_1 + w_2 \times i_2 + b_1 \times 1$$
$$= 0.15 \times 0.05 + 0.2 \times 0.1 + 0.35 \times 1$$
$$= 0.3775$$

$$a_{h_1} = \frac{1}{1+e^{-z_h}}$$
$$= \frac{1}{1+e^{-0.3775}}$$
$$\approx 0.5933$$

同理：$a_{h_2} \approx 0.5969$.

2. 隐藏层 → 输出层

$$z_{O_1} = w_5 \times h_1 + w_6 \times h_2 + b_2 \times 1$$
$$= 0.4 \times 0.5933 + 0.45 \times 0.5969 + 0.6 \times 1$$
$$= 1.1059$$

$$a_{O_1} = \frac{1}{1+e^{-z_{O_1}}}$$

$$= \frac{1}{1+e^{-1.1059}}$$

$$\doteq 0.7514$$

同理：$a_{O_2} \doteq 0.7729$

Step 2. 反向传播

1. 计算预先函数

$$E_{total} = \sum_{i=1}^{n} \frac{1}{n}(target - output)^2$$

$$E_{O_1} = \frac{1}{2}(0.01 - 0.7514)^2 \doteq 0.2748$$

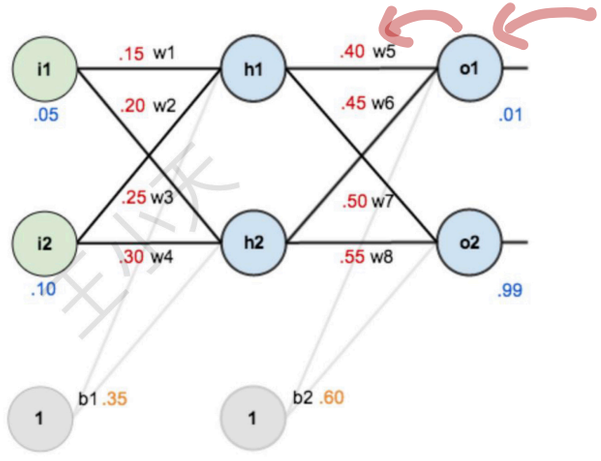
$$E_{O_2} = \frac{1}{2}(0.99 - 0.7729)^2 \doteq 0.0236$$

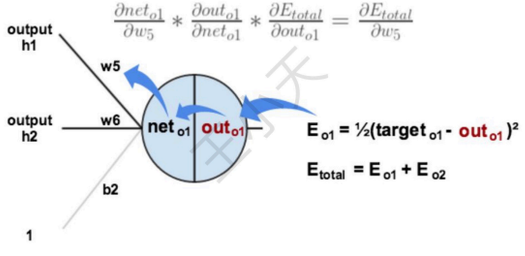
$$E_{total} = E_{O_1} + E_{O_2}$$

$$= 0.2748 + 0.0236$$

$$= 0.2984$$

2. 隐藏层 → 输出层权值更新：





更新 w_5 为例

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial a_{o_1}} \cdot \frac{\partial a_{o_1}}{\partial z_{o_1}} \cdot \frac{\partial z_{o_1}}{\partial w_5}$$

$$E_{total} = \frac{1}{2}(target_{o_1} - a_{o_1})^2 + \frac{1}{2}(target_{o_2} - a_{o_2})^2$$

$$\frac{\partial E_{total}}{\partial a_{o_1}} = a_{o_1} - target_{o_1}$$
$$= 0.7514 - 0.01$$
$$= 0.7414$$

$$\frac{\partial a_{o_1}}{\partial z_{o_1}} = a_{o_1} \times (1 - a_{o_1})$$
$$= 0.7514 \times (1 - 0.7514)$$
$$= 0.1868$$

$$\frac{\partial z_{o_1}}{\partial w_5} = a_{h_1}$$
$$= 0.5933$$

$$\therefore \frac{\partial E_{total}}{\partial w_5} = 0.7414 \times 0.1868 \times 0.5933 \approx 0.0822$$

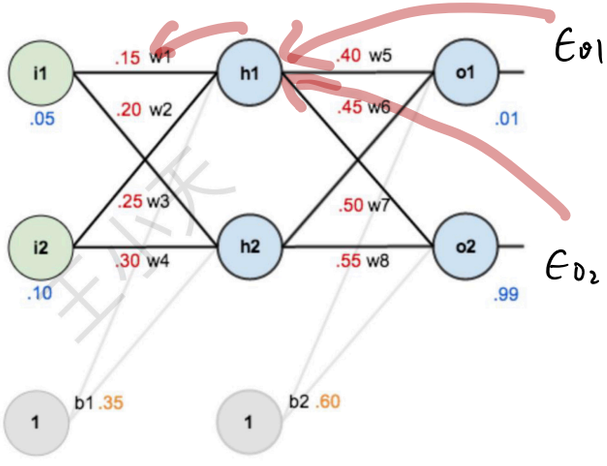
η：学习率

更新 w_5 的值

$$w_5^+ = w_5 - \eta \times \frac{\partial E_{total}}{\partial w_5}$$

$$= 0.4 - 0.5 \times 0.0822$$

$$= 0.3589$$

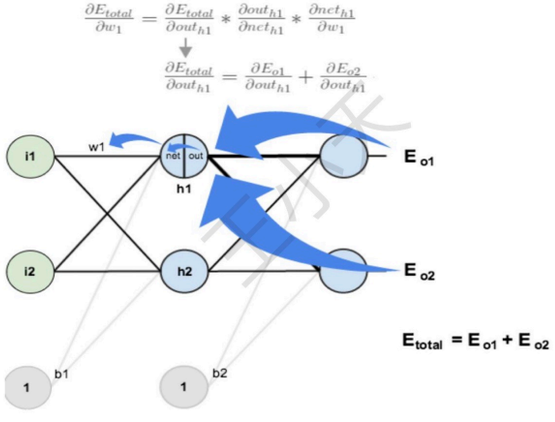


3. 隐藏层 → 隐藏层的权值更新.

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial a_{h_1}} \times \frac{\partial a_{h_1}}{\partial z_{h_1}} \times \frac{\partial z_{h_1}}{\partial w_1}$$

↓

$$\frac{\partial E_{total}}{\partial a_{h_1}} = \frac{\partial E_{o1}}{\partial a_{h_1}} + \frac{\partial E_{o2}}{\partial a_{h_1}}$$



$$\frac{\partial E_{total}}{\partial a_{h_1}}$$

$$\frac{\partial E_{o1}}{\partial a_{h_1}} = \frac{\partial E_{o1}}{\partial a_{o_1}} \times \frac{\partial a_{o_1}}{\partial z_{o1}} \times \frac{\partial z_{o1}}{\partial a_{h_1}}$$

$\doteq 0.7414 \times 0.1868 \times 0.4$

$\doteq 0.0554$

同理 $\frac{\partial E_{O_2}}{\partial a_{h_2}} \doteq -0.0191$

$$\frac{\partial E_{total}}{\partial a_{h_1}} = 0.0554 + (-0.0191)$$
$$= 0.0363$$

$$\frac{\partial a_{h_1}}{\partial z_{h_1}} = a_{h_1}(1 - a_{h_1})$$
$$\doteq 0.5933 \times (1 - 0.5933)$$
$$\doteq 0.2413.$$

$$\frac{\partial z_{h_1}}{\partial w_1} = i_1 = 0.05$$

$$\therefore \frac{\partial E_{total}}{\partial w_1} = 0.0363 \times 0.2413 \times 0.05$$
$$\doteq 0.00044$$

更新 w_1 的值.

$$w_1^+ = w_1 - \eta \times \frac{\partial E_{total}}{\partial w_1}$$

$$= 0.15 - 0.5 \times 0.00044$$

$$= 0.149$$