

过拟合解决方法

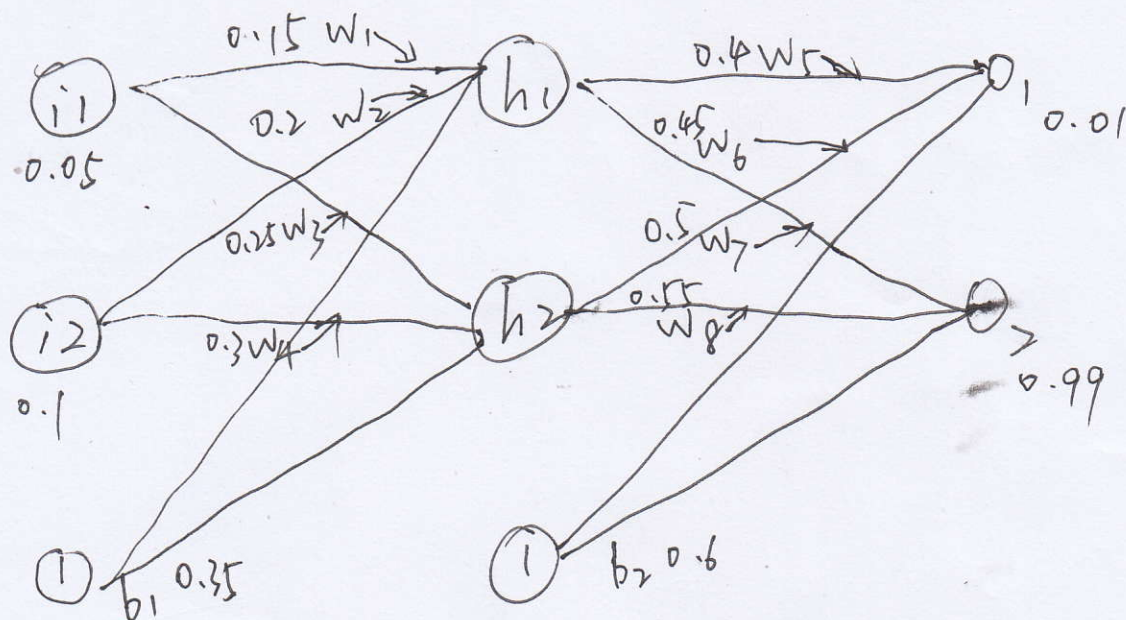
1. 减少特征. (特征提取, 特征选择)

2. Early stopping 连续10次Epoch后准确率不再提高.

3. 更多训练样本

4. 重新清洗数据.

5. Dropout 随机删除隐藏神经元 (0.5, 0.2, 0.1)



正向传播 ① 输入层到隐藏层

$$z_{h1} = i_1 w_1 + i_2 w_2 + b_1 = 0.05 \times 0.15 + 0.1 \times 0.2 + 0.35 = 0.3775$$

$$h_1 \text{ 的输出 } a_{h1} = \text{sigmoid}(z_{h1}) = \frac{1}{1 + e^{-z_{h1}}} = \frac{1}{1 + e^{-0.3775}} = 0.593269992$$

② 隐藏层到输出层

$$\text{同理 } a_{h2} = 0.596884378$$

③ 输出层输入加权和

$$z_{o1} = a_{h1} w_5 + a_{h2} w_6 + b_2 = 0.593269992 \times 0.4 + 0.596884378 \times 0.45 + 0.6 = 1.105905967$$

④ 输出

$$a_{o1} = \frac{1}{1 + e^{-z_{o1}}} = \frac{1}{1 + e^{-1.105905967}} = 0.751365069$$

$$\text{同理 } a_{o2} = 0.772928465$$

损失函数 $E = \sum_{i=1}^n \frac{1}{2} (\text{target} - \text{output})^2 = \sum_{i=1}^n \frac{1}{2} (\text{target}_{o_i} - a_{o_i})^2$

$E_{\text{total}} = E_{o_1} + E_{o_2} \quad E_{o_1} = \frac{1}{2} (0.01 - 0.751365069)^2 = 0.274811083$

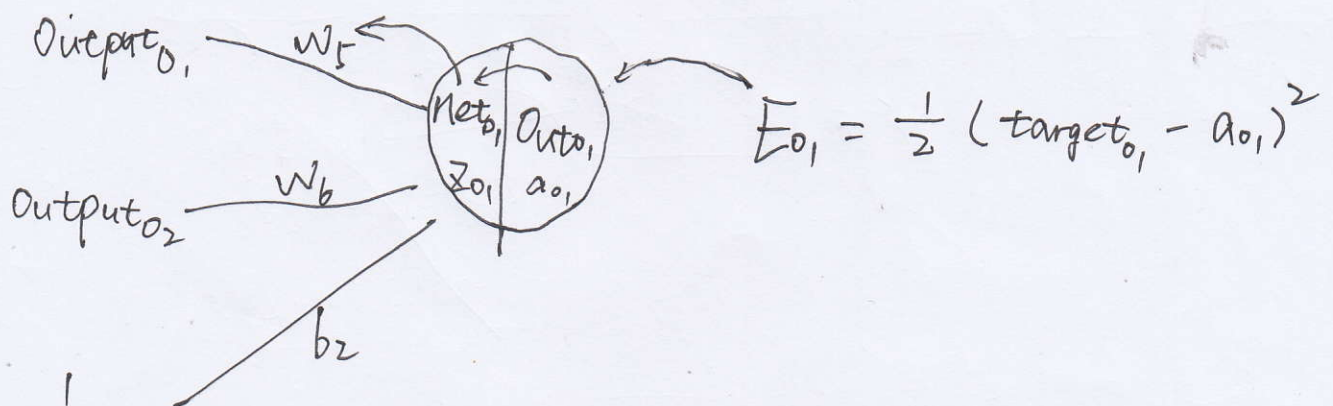
$E_{o_2} = \frac{1}{2} (0.99 - 0.772928465)^2 = 0.023560026$

$E_{\text{total}} = 0.274811083 + 0.023560026 = 0.298371109$

反向传播

隐藏层 \rightarrow 输出层 权重修正 ($w_5 \rightarrow w_8$)

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



$$\frac{\partial E_{\text{total}}}{\partial a_{o_1}} = 2 \times \frac{1}{2} (\text{target}_{o_1} - a_{o_1}) \times (-1) = a_{o_1} - \text{target}_{o_1} = 0.751365069 - 0.01 = 0.741365069$$

$$\frac{\partial a_{o_1}}{\partial z_{o_1}} = a_{o_1} \cdot (1 - a_{o_1}) = 0.751365069 \times (1 - 0.751365069) = 0.186815602$$

$$\frac{\partial z_{o_1}}{\partial w_5} = a_{h_1} = 0.593269992$$

$$\frac{\partial E_{\text{total}}}{\partial w_5} = 0.741365069 \times 0.186815602 \times 0.593269992 = 0.082167041$$

更新 w_5

η 学习率 0.5

$$w_5^+ = w_5 - \eta \frac{\partial E_{total}}{\partial w_5} = 0.4 - 0.5 \times 0.082167041 = 0.35891648$$

同理更新 w_6, w_7, w_8

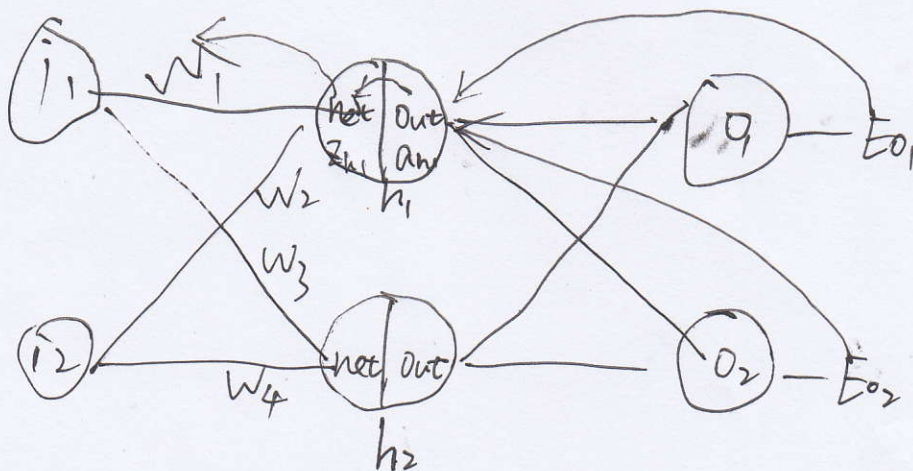
$$w_6^+ = 0.408666186$$

$$w_7^+ = 0.511321270$$

$$w_8^+ = 0.561370121$$

输入层 \rightarrow 隐藏层权重更新 $(w_1 - w_4) \frac{\partial E_{total}}{\partial w_1}$

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial a_{h1}} \times \frac{\partial a_{h1}}{\partial z_{h1}} \times \frac{\partial z_{h1}}{\partial w_1} = \left(\frac{\partial E_{o1}}{\partial a_{h1}} + \frac{\partial E_{o2}}{\partial a_{h1}} \right) \times \frac{\partial a_{h1}}{\partial z_{h1}} \times \frac{\partial z_{h1}}{\partial w_1}$$



$$\frac{\partial E_{o1}}{\partial a_{h1}} = \frac{\partial E_{o1}}{\partial a_{o1}} \cdot \frac{\partial a_{o1}}{\partial z_{o1}} \cdot \frac{\partial z_{o1}}{\partial a_{h1}} = (a_{o1} - target_{o1}) \times a_{o1} \times (1 - a_{o1}) \times w_5$$

$$= 0.741365069 \times 0.18685602 \times 0.4 = 0.055399425$$

$$\frac{\partial E_{o2}}{\partial a_{h1}} = -0.019049119$$

$$\frac{\partial E_{total}}{\partial a_{h1}} = 0.055399425 - 0.019049119 = 0.036350306$$

$$\frac{\partial a_{h1}}{\partial z_{h1}} = a_{h1} (1 - a_{h1}) = 0.2413007086$$

$$\frac{\partial z_{h1}}{\partial w_1} = i_1 = 0.05$$

$$\frac{\partial E_{total}}{\partial w_1} = 0.036350306 \times 0.2413007086 \times 0.05 = 0.000438568$$

更新 W_1

$$W_1^+ = W_1 - \eta \frac{\partial E_{total}}{\partial W_1} = 0.15 - 0.5 \times 0.000438568 = 0.149780716$$

同理 $W_2^+ = 0.19956143$

$$W_3^+ = 0.24975114$$

$$W_4^+ = 0.29950229$$

重新计算, 总误差更新为 0.291027924

不停迭代, 直到满足条件