1 Definições

Assumindo a seguinte terminologia para o desenvolvimento das equações:

- W1 corresponde a matriz de coeficientes entre a 1ª e 2ª camada.
- W2 corresponde a matriz de coeficientes entre a 2ª e 3ª camada.
- W3 corresponde a matriz de coeficientes entre a 1ª e 3ª camada.
- $I_j^{(L)}$ são os vetores cujos elementos denotam a entrada ponderada em relação ao j-ésimo neurônio da camada L.

Os vetores $I_j^{(L)}$ serão definidos abaixo.

$$I_1^{(1)} = \sum_{i=0}^{n} W1_{(1,i)} \cdot x_i \Longrightarrow W1_{1,0} \cdot x_0 + W1_{1,1} \cdot x_1 + \cdots + W1_{1,n} \cdot x_n \tag{1}$$

$$I_2^{(1)} = \sum_{i=0}^{n} W1_{(2,i)} \cdot x_i \Longrightarrow W1_{2,0} \cdot x_0 + W1_{2,1} \cdot x_1 + \cdots + W1_{2,n} \cdot x_n \tag{2}$$

$$I_{N1}^{(1)} = \sum_{i=0}^{n} W1_{(N1,i)} \cdot x_i \Longrightarrow W1_{N1,0} \cdot x_0 + W1_{N1,1} \cdot x_1 + \cdots + W1_{N1,n} \cdot x_n \tag{3}$$

$$I_1^{(2)} = \sum_{i=0}^{n} W2_{(1,i)} \cdot x_i \Longrightarrow W2_{1,0} \cdot x_0 + W2_{1,1} \cdot x_1 + \cdots + W2_{1,n} \cdot x_n \tag{4}$$

$$I_2^{(2)} = \sum_{i=0}^{n} W2_{(2,i)} \cdot x_i \Longrightarrow W2_{2,0} \cdot x_0 + W2_{2,1} \cdot x_1 + \cdots + W2_{2,n} \cdot x_n \tag{5}$$

$$I_{N1}^{(2)} = \sum_{i=0}^{n} W2_{(N1,i)} \cdot x_i \Longrightarrow W2_{N1,0} \cdot x_0 + W2_{N1,1} \cdot x_1 + \cdots + W2_{N1,n} \cdot x_n \tag{6}$$

$$I_1^{(3)} = \sum_{i=0}^{n} W2_{(1,i)} \cdot x_i + \sum_{i=0}^{n} W3_{(1,i)} \cdot x_i \Longrightarrow \sum_{i=0}^{n} X_i \cdot (W3_{(1,i)} + W2_{(1,i)}) \Longrightarrow$$
$$\Longrightarrow (W3_{1,0} + W2_{1,0}) \cdot x_0 + (W3_{1,1} + W2_{1,1}) \cdot x_1 + \cdots + (W3_{1,n} + W2_{1,n}) \cdot x_n$$

2 Ajuste da camada de saída

$$\bigtriangledown E^{(3)} = \frac{\partial E}{\partial W_{ji}^{(3)}} = \frac{\partial E}{\partial Y_j^{(3)}} \cdot \frac{\partial Y_j^{(3)}}{\partial I_j^{(3)}} \cdot \frac{\partial I_j^{(3)}}{\partial W_{ji}^{(3)}} \tag{7}$$

Utilizando-se das definições feitas anteriormente, podemos extrair que:

$$\frac{\partial I_j^{(3)}}{\partial W_{ji}^{(3)}} = Y_j^{(2)} \tag{8}$$

$$\frac{\partial Y_j^{(3)}}{\partial I_j^{(3)}} = g'(I_j^{(3)}) \tag{9}$$

$$\frac{\partial E}{\partial Y_j^{(3)}} = -(d_j - Y_j^{(3)}) \tag{10}$$

Substituindo as equações (8), (9) e (10) em (7), obtemos:

$$\frac{\partial E}{\partial W_{ji}^{(3)}} = -(d_j - Y_j^{(3)}) \cdot g'(I_j^{(3)}) \cdot Y_i^{(2)} \tag{11}$$

Obtém-se:

$$\Delta W_{ji}^{(3)} = -\eta \cdot \frac{\partial E}{\partial W_{ji}^{(3)}} \Longleftrightarrow \Delta W_{ji} = \eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)} \tag{12}$$

3 Ajuste da 1ª camada escondida

$$\bigtriangledown E^{(2)} = \frac{\partial E}{\partial W_{ji}^{(1)}} = \frac{\partial E}{\partial Y_j^{(1)}} \cdot \frac{\partial Y_j^{(1)}}{\partial I_j^{(1)}} \cdot \frac{\partial I_j^{(1)}}{\partial W_{ji}^{(1)}} \tag{13}$$

Utilizando os conceitos definidos anteriormente, obtém-se:

$$\frac{\partial I_j^{(1)}}{\partial W_{ji}^{(1)}} = x_i \tag{14}$$

$$\frac{\partial Y_j^{(1)}}{\partial I_j^{(1)}} = g'(I_j^{(1)}) \tag{15}$$

$$\frac{\partial E}{\partial Y_j^{(1)}} = \sum_{k=1}^{n_2} \frac{\partial E}{\partial I_k^{(2)}} \cdot \frac{\partial I_k^{(2)}}{\partial Y_j^{(1)}} = \sum_{k=1}^{n_2} \frac{\partial E}{\partial I_k^{(2)}} \cdot \frac{\partial (\sum_{k=1}^{n_2} W_{kj}^{(2)} \cdot Y_j^{(1)})}{\partial Y_j^{(1)}} \tag{16}$$

Para o caso específico deste EPC, o valor de n_2 é 1, uma vez que a primeira camada escondida é seguida por uma camada com apenas um neurônio.

A expressão (16) pode ser reescrita da seguinte forma:

$$\frac{\partial E}{\partial Y_j^{(1)}} = \sum_{k=1}^{1} \delta_k^{(2)} \cdot W_{kj}^{(2)} \tag{17}$$

É possível obter a seguinte expressão:

$$\frac{\partial E}{\partial W_{ji}^{(1)}} = -(\sum_{k=1}^{1} \delta_k^{(2)} \cdot W_{kj}^{(2)}) \cdot g'(I_j^{(1)}) \cdot x_i \tag{18}$$

Dessa forma, a regra para o ajuste de pesos da matriz da primeira camada escondida é:

$$\Delta W_{ji}^{(1)} = -\eta \cdot \frac{\partial E}{\partial W_{ji}^{(1)}} \iff \Delta W_{ji}^{(1)} = \eta \cdot \delta_j^{(1)} \cdot x_i \tag{19}$$