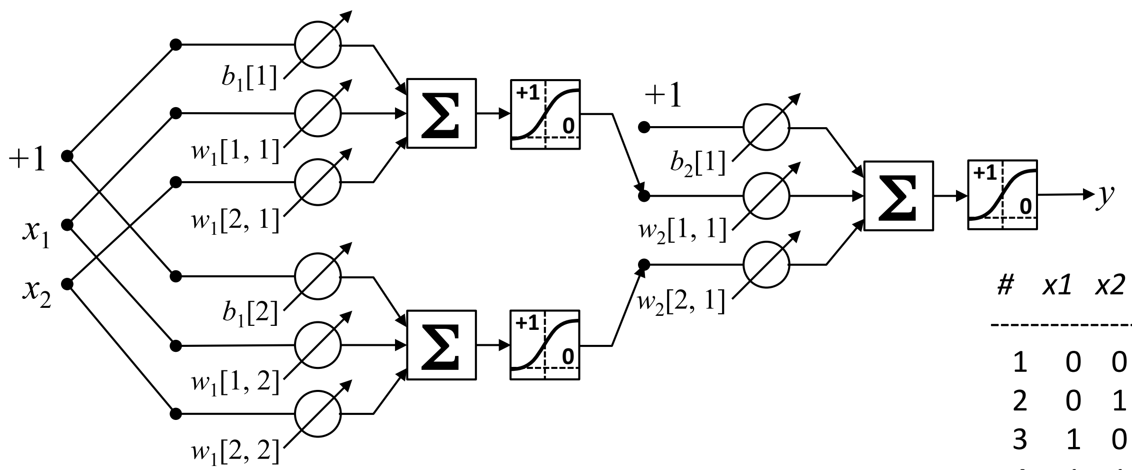
Backpropagation en Perceptrones Multicapa

Juan David Velásquez Henao
jdvelasq@unal.edu.co
Universidad Nacional de Colombia, Sede Medellín
Facultad de Minas
Medellín, Colombia

Haga clic aquí para acceder al repositorio online.

Se desean estimar los pesos sinápticos óptimos de un perceptrón multicapa (MLP) que reproduzca los resultados de la función binaria presentada en la tabla. El MLP es el siguiente:



#	x1	x2	d
1	0	0	1
2	0	1	0
3	1	0	0
4	1	1	1

PESOS SINÁPTICOS INICIALES:

$$b_1 = \begin{bmatrix} -0.5 \\ -0.2 \end{bmatrix}, \quad w_1 = \begin{bmatrix} -0.4 & -0.1 \\ -0.3 & +0.1 \end{bmatrix}$$

$$b_2 = +0.2, \qquad w_2 = \begin{bmatrix} +0.3 \\ +0.4 \end{bmatrix}$$

PROPAGACIÓN DE LA SEÑAL (HACIA DELANTE):

Para la capa oculta:

$$a_1[h] = b_1[h] + \sum_i w_1[i,h]x_i\,, \qquad z_1[h] = \sigma(a_1[h])$$

Para la capa de salida:

$$a_2[o] = b_2[o] + \sum_h w_2[h,o]z_1[h], \quad z_2[o] = \sigma(a_2[o])$$

PROPAGACIÓN DEL ERROR HACIA ATRÁS (BACKPROPAGATION) Y CÁLCULO DEL GRADIENTE:

<u>Capa de salida.</u>

$$e_2[0] = d - z_2[o]$$

$$z_2[o]' = z_2[o] * (1 - z_2[o])$$

$$\delta_2[o] = e_2[0] * z_2[o]'$$

$$g.b_2[o] = \frac{\partial}{\partial b_2[o]} E_p = -\delta_2[o]$$

$$g.w_2[h,o] = \frac{\partial}{\partial w_2[h,o]} E_p = -\delta_2[o] \cdot z_1[h]$$

<u>Capa oculta.</u>

$$e_1[h] = \sum_o \delta_2[o] \cdot w_2[h,o]$$

$$z_1[h]' = z_1[h] * (1 - z_1[h]), \qquad \delta_1[h] = -z_1[h]' \cdot e_1[h]$$

$$g.b_1[h] = \frac{\partial}{\partial b_1[h]} E_p = -\delta_1[h]$$

$$g.w_1[i,h] = \frac{\partial}{\partial w_1[i,h]} E_p = -\delta_1[h] \cdot x_i$$

Para la optimización use $\lambda = 1.0$.

CALCULOS

1. Propague la señal de la capa de entrada a la capa de salida.

#	x_1	x_2	d	$a_1[1]$	$z_1[1]$	$a_1[2]$	$z_1[2]$	$a_2[1]$	$z_2[1]$
1	0	0	1	-0.5000	_____	-0.2000	0.4502	_____	0.6209
2	0	1	0	-0.8000	0.3100	_____	0.4750	0.4830	0.6185
3	1	0	0	-0.9000	0.2891	-0.3000	0.4256	_____	0.6123
4	1	1	1	-1.2000	0.2315	-0.2000	0.4502	0.4495	0.6105

$$\text{MSE} = \frac{1}{2}\sum(d - z_2[1])^2 = 0.5264$$

2. Calcule el gradiente de los pesos sinápticos que llegan a la neurona de salida.

#	$e_2[1]$	$z_2[1]'$	$\delta_2[1]$	$g.b_2[1]$	$g.w_2[1,1]$	$g.w_2[2,1]$
1	_____	0.2354	0.0892	-0.0892	-0.0337	_____
2	-0.6185	0.2360	-0.1460	_____	0.0453	_____
3	-0.6123	_____	-0.1454	0.1454	_____	0.0619
4	0.3895	0.2378	_____	-0.0926	-0.0214	-0.0417
				----------	----------	----------
Corrección neta =>				0.1096	0.0322	0.0494

3. Calcule el gradiente de los pesos sinápticos que llegan a la <u>primera</u> neurona de la capa oculta.

#	$e_1[1]$	$z_1[1]'$	$\delta_1[1]$	$g.b_1[1]$	$g.w_1[1,1]$	$g.w_1[2,1]$
1	0.0268	_____	0.0063	-0.0063	0.0000	_____
2	-0.0438	0.2139	_____	0.0094	_____	0.0094
3	_____	0.2055	-0.0090	0.0090	_____	0.0000
4	0.0278	0.1779	0.0049	_____	-0.0049	_____
				----------	----------	----------
Corrección neta =>				0.0072	0.0041	0.0045

4. Calcule el gradiente de los pesos sinápticos que llegan a la <u>segunda</u> neurona de la capa oculta.

#	$e_1[2]$	$z_1[2]'$	$\delta_1[2]$	$g.b_1[2]$	$g.w_1[1,2]$	$g.w_1[2,2]$
1	_____	0.2475	0.0088	_____	0.0000	_____
2	-0.0584	0.2494	_____	0.0146	0.0000	0.0146
3	-0.0582	0.2445	-0.0142	0.0142	_____	_____
4	0.0370	_____	0.0092	-0.0092	-0.0092	-0.0092
				----------	----------	----------
Corrección neta =>				0.0108	0.0050	0.0054

5. Calcule los nuevos pesos sinápticos.

$$b_1 = \begin{bmatrix}-0.5\\-0.2\end{bmatrix} - \begin{bmatrix}+0.0072\\+0.0108\end{bmatrix} = \begin{bmatrix}-5.0072\\-2.0108\end{bmatrix}$$

$$w_1 = \begin{bmatrix}-0.4 & -0.1\\-0.3 & +0.1\end{bmatrix} - \begin{bmatrix}+0.0041 & +0.0050\\+0.0045 & +0.0054\end{bmatrix}$$

$$= \begin{bmatrix}-0.4041 & -0.1050\\-0.3045 & +0.0946\end{bmatrix}$$

$$b_2 = +0.2 - 0.1096 = 0.0904$$

$$w_2 = \begin{bmatrix}+0.3\\+0.4\end{bmatrix} - \begin{bmatrix}+0.0322\\+0.0494\end{bmatrix} = \begin{bmatrix}+0.2678\\+0.3506\end{bmatrix}$$

6. Se propaga nuevamente la señal de la capa de entrada a la capa de salida y se calcula el error.

#	x_1	x_2	d	$a_1[1]$	$z_1[1]$	$a_1[2]$	$z_1[2]$	$a_2[1]$	$z_2[1]$
1	0	0	1	_____	0.3759	-0.2108	0.4475	_____	0.5861
2	0	1	0	-0.8117	0.3075	-0.1162	0.4710	0.3379	0.5837
3	1	0	0	-0.9113	0.2867	_____	0.4217	0.3150	0.5781
4	1	1	1	-1.2158	0.2287	-0.2212	0.4449	0.3076	0.5763

$$\text{MSE} = 0.5 * \sum(d - z_2[1])^2 = 1.0257$$