Ejemplo de entrenamiento

La Figura siguiente muestra una red neuronal con una entrada $\vec{x} \in \mathbb{R}^2$ (con $D = 2$, sin incluir la neurona con valor unitario), con $\overrightarrow{y^o} \in \mathbb{R}^2$ (con $D = 2$, sin incluir la neurona con valor unitario) y $\overrightarrow{y^s} \in \mathbb{R}^1$ ($K = 1$ denotando la pertenencia a una clase $C = 1$ y la no pertenencia a la misma). A continuación se definen los valores de la entrada y los pesos en la capa oculta y de salida (los cuales se pueden suponer fueron inicializados aleatoriamente): [

$$
\begin{array}{cccc}
x_0 = 1 & t_1 = 0 & W^o_{0,1} = -1 & W^s_{0,1} = -1 \\
x_1 = 0 & & W^o_{0,2} = 0 & W^s_{1,1} = -2 \\
x_2 = 1 & & W^o_{1,1} = 3 & W^s_{2,1} = 3 \\
 & & W^o_{1,2} = -2 & \\
 & & W^o_{2,1} = 1 & \\
 & & W^o_{2,2} = 2 &
\end{array}
$$

]

Suponga que $\alpha = 1$.

Para actualizar los pesos de la red para la iteración $\tau = 1$, realizamos las dos etapas: pasada hacia adelante y retropropagación del error

Pasada hacia adelante: A partir de los datos anteriores, calculamos primero los pesos netos para la capa oculta:

$$\begin{aligned} p_1^o = &\quad W_{0,1}^o\, x_0 + W_{1,1}^o\, x_1 + W_{2,1}^o\, x_2 = -1 * 1 + 3 * 0 + 1 * 1 = 0 \\ p_2^o = &\quad W_{0,2}^o\, x_0 + W_{1,2}^o\, x_1 + W_{2,2}^o\, x_2 = 0 * 1 + -2 * 0 + 2 * 1 = 2 \end{aligned}$$

para posteriormente calcular la salida de cada unidad oculta:

$$\Rightarrow y_1^o = g^o\left(p_1^o\right) = \frac{1}{\left(1+e^0\right)} = 0.5$$

$$\Rightarrow y_2^o = g^o\left(p_2^o\right) = \frac{1}{\left(1+e^{-2}\right)} = 0.8808$$

Respecto a la capa de salida se tiene que:

$$p_1^s = \quad W_{0,1}^s\, y_0^o + W_{1,1}^s\, y_1^o + W_{2,1}^s\, y_2^o = -1 * 1 + -2 * 0.5 + 3 * 0.8808 = 0.6424$$

$$\Rightarrow y_1^s = g^s\left(p_1^s\right) = \frac{1}{\left(1+e^{-0.64}\right)} = 0.6553$$

Pasada hacia atrás y actualización de los pesos: Para la capa de salida, calculamos el delta de la única unidad como:

$$\delta_1^s = \left(y_1^s - t_{1,1}\right)\left(y_1^s\left(1 - y_1^s\right)\right) = (0.6553 - 0)\,(0.6553\ (1 - 0.6553)) = 0.148$$

Con base al cálculo del delta para la capa de salida, se actualizan los pesos en la capa de salida según la ecuación

$$W_{m,k}^{s}(\tau + 1) = W_{m,k}^{s}(\tau) + \triangle W_{m,k}^{s} = W_{m,k}^{s}(\tau) - \alpha\, \delta_k^s\, y_m^o$$

:

$$W_{0,1}^{s}(\tau + 1) = W_{0,1}^{s}(\tau) - 1\ \delta_1^s\, y_0^o = -1 - 0.148 * 1 = -1 - 0.148 * 1 = -1 - 0.148 = -1.148$$

$$W_{1,1}^{s}(\tau + 1) = W_{1,1}^{s}(\tau) - 1\ \delta_1^s\, y_1^o = -2 - 0.148 * 0.5 = -2 - 0.0740 = -2.074$$

$$W_{2,1}^{s}(\tau + 1) = W_{2,1}^{s}(\tau) - 1\ \delta_1^s\, y_2^o = 3 - 0.148 * 0.8808 = 3 - 0.1304 = 2.8696$$

Y para la capa oculta tenemos de forma similar, el delta para las dos unidades:

$$\delta_1^o = \left(\textstyle\sum_{k=1}^{K=1} \delta_k^s\, W_{1,k}^{s}\right)\left(y_1^o\left(1 - y_1^o\right)\right) = (0.148 * -2) * (0.5 * (1 - 0.5)) = -0.0740$$

$$\delta_2^o = \left(\textstyle\sum_{k=1}^{K=1} \delta_k^s\, W_{2,k}^{s}\right)\left(y_2^o\left(1 - y_2^o\right)\right) = (0.148 * 3) * (0.8808 * (1 - 0.8808)) = 0.0466$$

Por lo que los pesos nuevos para la capa oculta según la ecuación

$$W_{d,m}^{o}(\tau + 1) = W_{d,m}^{o}(\tau) + \triangle W_{d,m}^{o} = W_{d,m}^{o}(\tau) - \alpha\, \delta_m^o\, x_d,$$

vienen dados por:

$$W_{0,1}^{o}\,(\tau+1) = W_{0,1}^{o}\,(\tau) - 1\ \delta_1^o\, x_0 = -1 + 1 * 0.074 * 1 = -1 + 0.074 = -0.926$$
$$W_{1,1}^{o}\,(\tau+1) = W_{1,1}^{o}\,(\tau) - 1\ \delta_1^o\, x_1 = 3 + 1 * 0.074 * 0 = 3 + 0 = 3$$
$$W_{2,1}^{o}\,(\tau+1) = W_{2,1}^{o}\,(\tau) - 1\ \delta_1^o\, x_2 = 1 + 1 * 0.074 * 1 = 1 + 0.074 = 1.074$$
$$W_{0,2}^{o}\,(\tau+1) = W_{0,2}^{o}\,(\tau) - 1\ \delta_2^o\, x_0 = 0 - 1 * 0.0466 * 1 = 0 - 0.0466 = -0.0466$$
$$W_{1,2}^{o}\,(\tau+1) = W_{1,2}^{o}\,(\tau) - 1\ \delta_2^o\, x_1 = -2 - 1 * 0.0466 * 0 = -2 - 0 = -2$$
$$W_{2,2}^{o}\,(\tau+1) = W_{2,2}^{o}\,(\tau) - 1\ \delta_2^o\, x_2 = 2 - 1 * 0.0466 * 1 = 2 - 0.0466 = 1.9534$$