Machine Intelligence 1 - Exercise 3: Multilayer Perceptrons and Backpropagation Algorithm

Liu, Zhiwei 387571 | Moon, Chulhyun 392865 | Wenzel, Daniel 365107 | Ozmen, Cengizhan 388011

Pipo, Aiko 390011

November 8, 2017

H3.1 Binary Classification (3 Points)

(a)

We can simply derivate the given function $e^{(\alpha)}$:

$$\begin{aligned}
\frac{\delta e^{(\alpha)}}{\delta y(\underline{x}^{(\alpha)};\underline{w}))} &= -\frac{y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)};\underline{w})} - (1-y_T^{(\alpha)})\left(-\frac{1}{1-y(\underline{x}^{(\alpha)};\underline{w})}\right) \\
&= -\frac{y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)};\underline{w})} - \frac{1-y_T^{(\alpha)}}{1-y(\underline{x}^{(\alpha)};\underline{w})} \\
&= -\frac{y_T^{(\alpha)}(1-y(\underline{x}^{(\alpha)};\underline{w})) - (1-y_T^{(\alpha)})y(\underline{x}^{(\alpha)};\underline{w})}{y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w})} \\
&= -\frac{y_T^{(\alpha)} - y_T^{(\alpha)}y(\underline{x}^{(\alpha)};\underline{w}) - y(\underline{x}^{(\alpha)};\underline{w}) + y_T^{(\alpha)}y(\underline{x}^{(\alpha)};\underline{w})}{y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))} \\
&= -\frac{y_T^{(\alpha)} - y(\underline{x}^{(\alpha)};\underline{w})}{y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))} \\
&= \frac{y(\underline{x}^{(\alpha)};\underline{w}) - y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))}
\end{aligned}$$

(b)

Again, a simple derivation yields:

$$\begin{aligned}
f'(h_1^2) &= -\frac{1}{(1+e^{-h_1^2})^2}\cdot\left(-e^{-h_1^2}\right) \\
&= \frac{1}{1+e^{-h_1^2}}\cdot\frac{e^{-h_1^2}}{1+e^{-h_1^2}} \\
&= f(h_1^2)\cdot\frac{1+e^{-h_1^2}-1}{1+e^{-h_1^2}} \\
&= f(h_1^2)\left(\frac{1+e^{-h_1^2}}{1+e^{-h_1^2}} - \frac{1}{1+e^{-h_1^2}}\right) \\
&= f(h_1^2)(1-f(h_1^2))
\end{aligned}$$

(c)

First, we look at the derivative of $e^{(\alpha)}$ again:

$$\frac{\delta e^{(\alpha)}}{\delta w_{1j}^{21}} = \frac{y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)};\underline{w})}\left(\frac{\delta y(\underline{x}^{(\alpha)};\underline{w})}{\delta w_{1j}^{21}}\right) - \frac{1-y_T^{(\alpha)}}{1-y(\underline{x}^{(\alpha)};\underline{w})}\left(\frac{\delta y(\underline{x}^{(\alpha)};\underline{w})}{\delta w_{1j}^{21}}\right)$$

We now look at the derivative for $y(\underline{x}^{(\alpha)};\underline{w}) = S_1^2 = f_1^2(h_1^2)$:

$$\frac{\delta y(\underline{x}^{(\alpha)};\underline{w})}{\delta w_{1j}^{21}} = \left(\frac{\delta}{\delta h_1^2} f_1^2(h_1^2)\right)\left(\frac{\delta}{\delta w_{1j}^{21}} h_1^2\right)$$

With the knowledge from (b) we can simplify this to

$$\frac{\delta y(\underline{x}^{(\alpha)};\underline{w})}{\delta w_{1j}^{21}} = f_1^2(h_1^2)(1-f_1^2(h_1^2))\left(\frac{\delta}{\delta w_{1j}^{21}} h_1^2\right).$$

Transforming back and adding the explicit form for h_1^2, we obtain

$$\frac{\delta y(\underline{x}^{(\alpha)};\underline{w})}{\delta w_{1j}^{21}} = y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))\left(\frac{\delta}{\delta w_{1j}^{21}} \sum_{i=0}^{N} w_{1i}^{21}\cdot S_i^1\right),$$

and thus

$$\frac{\delta y(\underline{x}^{(\alpha)};\underline{w})}{\delta w_{1j}^{21}} = y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))S_j^1.$$

Now, plugging this back into our original derivation we get

$$\begin{aligned}
\frac{\delta e^{(\alpha)}}{\delta w_{1j}^{21}} &= \frac{y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)};\underline{w})} y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))S_j^1 - \frac{1-y_T^{(\alpha)}}{1-y(\underline{x}^{(\alpha)};\underline{w})} y(\underline{x}^{(\alpha)};\underline{w})(1-y(\underline{x}^{(\alpha)};\underline{w}))S_j^1 \\
&= \left((1-y_T^{(\alpha)})y(\underline{x}^{(\alpha)};\underline{w}) - y_T^{(\alpha)}(1-y(\underline{x}^{(\alpha)};\underline{w}))\right) S_j^1 \\
&= \left(y(\underline{x}^{(\alpha)};\underline{w}) - y_T^{(\alpha)} y(\underline{x}^{(\alpha)};\underline{w}) - y_T^{(\alpha)} + y_T^{(\alpha)} y(\underline{x}^{(\alpha)};\underline{w})\right) S_j^1 \\
&= \left(y(\underline{x}^{(\alpha)};\underline{w}) - y_T^{(\alpha)}\right) S_j^1
\end{aligned}$$