Machine Learning Assignment 95

Elijah Tarr

February 26, 2021

1 Problem 95-1

1.1 Derivative with respect to last edge

Calculating the derivative of the squared sum of errors with respect to the w_{23} edge in the given neural network:

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}w_{23}} &= \frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[\left(y_{\text{predicted}} - y_{\text{actual}}\right)^2\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[\left(a_3 - y_{\text{actual}}\right)^2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[a_3 - y_{\text{actual}}\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[a_3\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[f_3(i_3)\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right) f_3(i_3)\frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[i_3\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right) f_3(i_3)\frac{\mathrm{d}}{\mathrm{d}w_{23}}\left[a_2 * w_{23}\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right) f_3'(i_3) a_2
\end{aligned}$$

To check my work:

$$\begin{aligned}
2\left(a_3 - y_{\text{actual}}\right) f_3'(i_3) a_2 &= 2(5-1)\cdot 9 \cdot 4 \\
&= 288
\end{aligned}$$

1.2 Derivative with respect to second-to-last edge

$$
\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}w_{12}} &= \frac{\mathrm{d}}{\mathrm{d}w_{12}}\left[(y_{\text{predicted}} - y_{\text{actual}})^2\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}w_{12}}\left[(a_3 - y_{\text{actual}})^2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{12}}\left[a_3 - y_{\text{actual}}\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{12}}\left[a_3\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{12}}\left[f_3(i_3)\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)\frac{\mathrm{d}}{\mathrm{d}w_{21}}\left[i_3\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)\frac{\mathrm{d}}{\mathrm{d}w_{21}}\left[w_{23} \cdot a_2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}\frac{\mathrm{d}}{\mathrm{d}w_{21}}\left[a_2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}\frac{\mathrm{d}}{\mathrm{d}w_{21}}\left[f_2(i_2)\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)\frac{\mathrm{d}}{\mathrm{d}w_{21}}\left[i_2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)\frac{\mathrm{d}}{\mathrm{d}w_{21}}\left[w_{21} \cdot a_1\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)a_1
\end{aligned}
$$

To check my work:

$$
\begin{aligned}
2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)a_1 &= 2(5-1) \cdot 9 \cdot 12 \cdot 8 \cdot 3 \\
&= 20736
\end{aligned}
$$

1.3 Derivative with respect to first edge

$$
\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}w_{01}} &= \frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[\left(y_{\text{predicted}} - y_{\text{actual}}\right)^2\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[\left(a_3 - y_{\text{actual}}\right)^2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[a_3 - y_{\text{actual}}\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[a_3\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[f_3(i_3)\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[i_3\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[w_{23}\cdot a_2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[a_2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[f_2(i_2)\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[i_2\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[w_{21}\cdot a_1\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)w_{21}\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[a_1\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)w_{21}\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[f_1(i_1)\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)w_{21}f_1'(i_1)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[i_1\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)w_{21}f_1'(i_1)\frac{\mathrm{d}}{\mathrm{d}w_{01}}\left[a_0\cdot w_{01}\right] \\
&= 2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)w_{21}f_1'(i_1)a_0
\end{aligned}
$$

To check my work:

$$
\begin{aligned}
2\left(a_3 - y_{\text{actual}}\right)f_3'(i_3)w_{23}f_2'(i_2)w_{21}f_1'(i_1)a_0 &= 2(5-1)\cdot 9\cdot 12\cdot 8\cdot 11\cdot 7\cdot 2 \\
&= 1064448
\end{aligned}
$$