# 1. Regression

I'll denote * $O_0^{(2)} = y'$

* $SE(y, y') = C = (y - y')^2$

$$\frac{\partial C}{\partial a_{k0}^{(1)}} = \frac{\partial C}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial a_{k0}^{(1)}}$$

$$\frac{\partial C}{\partial a_{ik}^{(0)}} = \frac{\partial C}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial O_k^{(1)}} \cdot \frac{\partial O_k^{(1)}}{\partial a_{ik}^{(0)}}$$

* $$\frac{\partial C}{\partial a_{01}^{(0)}} = \frac{\partial C}{\partial y'} \cdot \frac{\partial y'}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial a_{01}^{(0)}}$$

$$= \frac{\partial C}{\partial a_{01}^{(0)}} = \frac{\partial (y - y')^2}{\partial y'} \cdot \frac{\partial y'}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial a_{01}^{(0)}}$$

$$O_1^{(1)} = 1 \cdot a_{01}^{(0)} + x_1 \cdot a_{11} + x_2 \cdot a_{21}$$

$$y' = a_{00}^{(1)} + O_1^{(1)} \cdot a_{10}^{(1)} + O_2^{(1)} \cdot a_{20}^{(1)} + O_3^{(1)} \cdot a_{30}^{(1)}$$

$$= \frac{\partial (y-y')^2}{\partial y'} \cdot \frac{\partial \left(a_{00}^{(1)} + O_1^{(1)} \cdot a_{10}^{(1)} + O_2^{(1)} \cdot a_{20}^{(1)} + O_3^{(1)} a_{30}^{(1)}\right)}{\partial O_1^{(1)}}$$

$$+ \frac{\partial \left(a_{01}^{(0)} + x_1 a_{11}^{(0)} + x_2 a_{21}^{(0)}\right)}{\partial a_{01}^{(0)}}$$

$$* \quad \frac{\partial (y-y')^2}{\partial y'} = \boxed{2 \cdot (y-y')}$$

$$* \quad \frac{\partial \left(a_{00} + O_1^{(1)} \cdot a_{10}^{(1)} + O_2^{(1)} \cdot a_{20}^{(1)} + O_3^{(1)} \cdot a_{30}^{(1)}\right)}{\partial O_1^{(1)}} = \boxed{a_{10}^{(1)}}$$

$$* \quad \frac{\partial \left(a_{01}^{(0)} + x_1 a_{11}^{(0)} + x_2 a_{21}^{(0)}\right)}{\partial a_{01}^{(0)}} = \boxed{1}$$

Therefore : for weight rule update,

$$a_{ik}^{(0)} = a_{ik}^{(0)} - a \frac{\partial C}{\partial a_{ik}^{(0)}}$$

**General Rule**

$$\boxed{\frac{\partial C}{\partial a_{ik}^{(0)}} = 2.(y - O_0^{(2)}) + a_{KO}^{(1)} + O_i^{(0)}}$$

$$\boxed{a_{01\,new}^{(0)} = a_{01\,old}^{(0)} - a \cdot \left(2 \cdot (y - O_0^{(2)}) + a_{10}^{(1)} + 1\right)}$$

$$\boxed{a_{02\,new}^{(0)} = a_{02\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{20}^{(1)} + 1\right)}$$

$$\boxed{a_{03\,new}^{(0)} = a_{03\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{30}^{(1)} + 1\right)}$$

$$\boxed{a_{11\,new}^{(0)} = a_{11\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{10}^{(1)} + x_1\right)}$$

$$\boxed{a_{12\,new}^{(0)} = a_{12\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{20}^{(1)} + x_1\right)}$$

$$\boxed{a_{13\,new}^{(0)} = a_{13\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{30}^{(1)} + x_1\right)}$$

$$\boxed{a_{21\,new}^{(0)} = a_{21\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{10}^{(1)} + x_2\right)}$$

$$\boxed{a_{22\,new}^{(0)} = a_{22\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{20}^{(1)} + x_2\right)}$$

$$\boxed{a_{23\,new}^{(0)} = a_{23\,old}^{(0)} - a\cdot\left(2\cdot\left(y - O_0^{(2)}\right) + a_{30}^{(1)} + x_2\right)}$$

# Explonation

For weights that start from $O_0^{(0)}$ there is 1 for those that start from $O_1^{(0)}$ there is $x_1$ and for those that start from $O_2^{(0)}$ there is $x_2$.

The first weight I calculated step-by-step and in a long fashion looking at those calculations it is trivial to find the other weights.

---

Now I will calculate the weights between hidden layer and output layer.

$$\frac{\partial C}{a_{k0}^{(1)}} = \frac{\partial C}{\partial y'} \cdot \frac{\partial y'}{\partial a_{k0}}$$

$$y' = a_{00}^{(1)} + O_1^{(1)} \cdot a_{10}^{(1)} + O_2^{(1)} \cdot a_{20}^{(1)} + O_3^{(1)} \cdot a_{30}^{(1)}$$

$$\frac{\partial C}{\partial y'} = 2 \cdot (y - y')$$

$$\frac{\partial y'}{\partial a_{k0}} = \frac{a_{00}^{(1)} + O_1^{(1)} \cdot a_{10}^{(1)} + O_2^{(1)} \cdot a_{20}^{(1)} + O_3^{(1)} \cdot a_{30}^{(1)}}{\partial a_{k0}}$$

Referring to the general formula:

$$\boxed{a_{00_{new}}^{(1)} = a_{00_{old}}^{(1)} - a.\left(2.\left(y - O_0^{(2)}\right) + 1\right)}$$

$$\boxed{a_{10_{new}}^{(1)} = a_{10_{old}}^{(1)} - a.\left(2.\left(y - O_0^{(2)}\right) + O_1^{(1)}\right)}$$

$$\boxed{a_{20_{new}}^{(1)} = a_{20_{old}}^{(1)} - a.\left(2.\left(y - O_0^{(2)}\right) + O_2^{(1)}\right)}$$

$$\boxed{a_{30_{new}}^{(1)} = a_{30_{old}}^{(1)} - a.\left(2.\left(y - O_0^{(2)}\right) + O_3^{(1)}\right)}$$

We can also plug in:

$$O_1^{(1)} = a_{01}^{(0)} + a_{11}^{(0)}.\,x_1 + a_{21}^{(0)}.\,x_2$$

$$O_2^{(1)} = a_{02}^{(0)} + a_{12}^{(0)}.\,x_1 + a_{22}^{(0)}.\,x_2$$

$$O_3^{(1)} = a_{03}^{(0)} + a_{13}^{(0)}.\,x_1 + a_{23}^{(0)}.\,x_2$$

# 2. Classification

* $CE(f, f') = -\sum_i f_i * \log(f_i')$

* $f$ = ground truth, $f'$ = network output

* Our CE function $= -\sum_{i=0}^{2} f_i * \log(O_i^{(2)}) = C$

$$C = -\left(f_0 \cdot \log(O_0^{(2)}) + f_1 \cdot \log(O_1^{(2)}) + f_2 \cdot \log(O_2^{(2)})\right)$$

$$O_n^{(2)} = a_{0n} + a_{1n} \cdot O_1^{(1)} + a_{2n} \cdot O_2^{(1)} + a_{3n} + O_3^{(1)}$$

$$\frac{\partial C}{\partial O_i^{(2)}} = -\frac{f_i}{O_i^{(2)}}$$

$$\frac{\partial O_n^{(2)}}{\partial a_{kn}^{(1)}} = O_k^{(1)}$$

$$\frac{\partial C}{\partial a_{kn}^{(1)}} = \frac{\partial C}{\partial O_n^{(2)}} \cdot \frac{\partial O_n^{(2)}}{\partial a_{kn}^{(1)}}$$

$$\boxed{= \frac{-f_n}{O_n^{(2)}} \cdot O_k^{(1)}}$$

→ General formula

$$\boxed{a_{00_{new}}^{(1)} = a_{00_{old}}^{(1)} - a \cdot \left( \frac{-f_0}{O_0^{(2)}} \right)}$$

$$\boxed{a_{01_{new}}^{(1)} = a_{01_{old}}^{(1)} - a \cdot \left( \frac{-f_1}{O_1^{(2)}} \right)}$$

$$\boxed{a_{02_{new}}^{(1)} = a_{02_{old}}^{(1)} - a \cdot \left( \frac{-f_2}{O_2^{(2)}} \right)}$$

$$\boxed{a_{03_{new}}^{(1)} = a_{03_{old}}^{(1)} - a \cdot \left( \frac{-f_3}{O_3^{(2)}} \right)}$$

$$\boxed{a_{10_{new}}^{(1)} = a_{10_{old}}^{(1)} - a \cdot \left( \frac{-f_0}{O_0^{(2)}} \cdot O_1^{(1)} \right)}$$

$$\boxed{a_{11_{new}}^{(1)} = a_{11_{old}}^{(1)} - a \cdot \left( \frac{-f_1}{O_1^{(2)}} \cdot O_1^{(1)} \right)}$$

$$\boxed{a_{12_{new}}^{(1)} = a_{12_{old}}^{(1)} - a \cdot \left( \frac{-f_2}{O_2^{(2)}} \cdot O_1^{(1)} \right)}$$

$$\boxed{a_{13_{new}}^{(1)} = a_{13_{old}}^{(1)} - a \cdot \left( \frac{-f_3}{O_3^{(2)}} \cdot O_1^{(1)} \right)}$$

$$\boxed{a^{(1)}_{20\,new} = a^{(1)}_{20\,old} - a\cdot\left(\frac{-f_0}{O_0^{(2)}}\cdot O_2^{(1)}\right)}$$

$$\boxed{a^{(1)}_{21\,new} = a^{(1)}_{21\,old} - a\cdot\left(\frac{-f_1}{O_1^{(2)}}\cdot O_2^{(1)}\right)}$$

$$\boxed{a^{(1)}_{22\,new} = a^{(1)}_{22\,old} - a\cdot\left(\frac{-f_2}{O_2^{(2)}}\cdot O_2^{(1)}\right)}$$

$$\boxed{a^{(1)}_{23\,new} = a^{(1)}_{23\,old} - a\cdot\left(\frac{-f_3}{O_3^{(2)}}\cdot O_2^{(1)}\right)}$$

$$\boxed{a^{(1)}_{30\,new} = a^{(1)}_{30\,old} - a\cdot\left(\frac{-f_0}{O_0^{(2)}}\cdot O_3^{(1)}\right)}$$

$$\boxed{a^{(1)}_{31\,new} = a^{(1)}_{31\,old} - a\cdot\left(\frac{-f_1}{O_1^{(2)}}\cdot O_3^{(1)}\right)}$$

$$\boxed{a^{(1)}_{32\,new} = a^{(1)}_{32\,old} - a\cdot\left(\frac{-f_2}{O_2^{(2)}}\cdot O_3^{(1)}\right)}$$

$$\boxed{a^{(1)}_{33\,new} = a^{(1)}_{33\,old} - a\cdot\left(\frac{-f_3}{O_3^{(2)}}\cdot O_3^{(1)}\right)}$$

We can also plug in:

$$O_1^{(1)} = a_{01}^{(0)} + a_{11}^{(0)} \cdot x_1 + a_{21}^{(0)} \cdot x_2$$

$$O_2^{(1)} = a_{02}^{(0)} + a_{12}^{(0)} \cdot x_1 + a_{22}^{(0)} \cdot x_2$$

$$O_3^{(1)} = a_{03}^{(0)} + a_{13}^{(0)} \cdot x_1 + a_{23}^{(0)} \cdot x_2$$