Calculations for a Neural Network with 2 input neurons, 1 hidden Layer with 3 neurons and an output layer with 2 neurons using SGD and backpropagation.

$$let\ \vec{x} = input$$

$$let\ \vec{y} = actual\ output$$

$$ReLU(\vec{x}) = \ \forall i \in \{a \in \mathbb{N}_0 | \alpha < |x|\}: \begin{cases} x_i = x_i, if\ x_i \geq 0 \\ x_i = 0, if\ x_i < 0 \end{cases}$$

$$ReLU'(\vec{x}) = \forall i \in \{a \in \mathbb{N}_0 | \alpha < |x|\}: \begin{cases} x_i = 1, if\ x_i \geq 0 \\ x_i = 0, if\ x_i < 0 \end{cases}$$

$$\sigma(\vec{x}) = \forall i \in \{a \in \mathbb{N}_0 | \alpha < |x|\}: x_i \ = \frac{1}{1 + e^{-x_i}}$$

$$\sigma'(\vec{x}) = \sigma(\vec{x})(1 - \sigma(\vec{x}))$$

$$C(w, \dots, b, \dots) = \frac{1}{2}\sum_{i=0}^{N}(\hat{y}_i - y_i)^2$$

$$C'(w, \dots, b, \dots) = \hat{y} - y$$

$$z^L = W^L \cdot a^{L-1} + B^L$$

$$g_L(x) = activation\ function\ from\ layer\ L$$

$$a^L = g_L(z^L)$$

$$z^0 = a^0 = \vec{x} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$$

$$W^1 = \begin{pmatrix} w_{00}^1 & w_{01}^1 \\ w_{10}^1 & w_{11}^1 \\ w_{20}^1 & w_{21}^1 \end{pmatrix}$$

$$B^1 = \begin{pmatrix} b_0^1 \\ b_1^1 \\ b_2^1 \end{pmatrix}$$

$$z^1 = W^1 \cdot a^0 + B^1 = \begin{pmatrix} w_{00}^1 & w_{01}^1 \\ w_{10}^1 & w_{11}^1 \\ w_{20}^1 & w_{21}^1 \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} + \begin{pmatrix} b_0^1 \\ b_1^1 \\ b_2^1 \end{pmatrix} = \begin{pmatrix} w_{00}^1 \cdot x_0 + w_{01}^1 \cdot x_1 + b_0^1 \\ w_{10}^1 \cdot x_0 + w_{11}^1 \cdot x_1 + b_1^1 \\ w_{20}^1 \cdot x_0 + w_{21}^1 \cdot x_1 + b_2^1 \end{pmatrix}$$

$$a^1 = ReLU(z^1) = ReLU\left(\begin{pmatrix} w_{00} \cdot x_0 + w_{01} \cdot x_1 + b_0 \\ w_{10} \cdot x_0 + w_{11} \cdot x_1 + b_1 \\ w_{20} \cdot x_0 + w_{21} \cdot x_1 + b_2 \end{pmatrix}\right) = \begin{pmatrix} a_0^1 \\ a_1^1 \\ a_2^1 \end{pmatrix}$$

$$W^2 = \begin{pmatrix} w_{00}^2 & w_{01}^2 & w_{02}^2 \\ w_{10}^2 & w_{11}^2 & w_{12}^2 \end{pmatrix}$$

$$B^2 = \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix}$$

$$z^2 = W^2 \cdot a^1 + B^2 = \begin{pmatrix} w_{00}^2 & w_{01}^2 & w_{02}^2 \\ w_{10}^2 & w_{11}^2 & w_{12}^2 \end{pmatrix} \cdot \begin{pmatrix} a_0^1 \\ a_1^1 \\ a_2^1 \end{pmatrix} + \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix} = \begin{pmatrix} w_{00}^2 \cdot a_0^1 + w_{01}^2 \cdot a_1^1 + w_{02}^2 \cdot a_2^1 + b_0^2 \\ w_{10}^2 \cdot a_0^1 + w_{11}^2 \cdot a_1^1 + w_{12}^2 \cdot a_2^1 + b_1^2 \end{pmatrix}$$

$$a^2 = \hat{y} = \sigma(z^2) = \sigma\left(\begin{pmatrix} w_{00}^2 \cdot a_0^1 + w_{01}^2 \cdot a_1^1 + w_{02}^2 \cdot a_2^1 + b_0^2 \\ w_{10}^2 \cdot a_0^1 + w_{11}^2 \cdot a_1^1 + w_{12}^2 \cdot a_2^1 + b_1^2 \end{pmatrix}\right) = \begin{pmatrix} \hat{y}_0 \\ \hat{y}_1 \end{pmatrix}$$

$$\boldsymbol{W_{new}^L = W_{old}^L - \eta \frac{\partial C}{\partial W^L}}$$

$$\boldsymbol{\frac{\partial C}{\partial B^L} = \frac{\partial C}{\partial z^L}}$$

$$\boldsymbol{\frac{\partial C}{\partial W^L} = \frac{\partial C}{\partial z^L} \cdot \left(a^{L-1}\right)^T}$$

$$\boldsymbol{\frac{\partial C}{\partial z^{L-1}} = \left(W^L\right)^T \cdot \frac{\partial C}{\partial z^L} \circ g_{L-1}'\left(z^{L-1}\right)}$$

$$\begin{aligned}
\frac{\partial C}{\partial W^2} = \frac{\partial C}{\partial z^2} \cdot \begin{pmatrix} a_0^1 \\ a_1^1 \\ a_2^1 \end{pmatrix}^T &= \sigma'(z^2)(a^2 - y) \cdot \begin{pmatrix} a_0^1 & a_1^1 & a_2^1 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{1+e^{-z_0^2}}\left(1 - \frac{1}{1+e^{-z_0^2}}\right) \\ \frac{1}{1+e^{-z_1^2}}\left(1 - \frac{1}{1+e^{-z_1^2}}\right) \end{pmatrix} \circ \begin{pmatrix} a_0^2 - y_0 \\ a_1^2 - y_1 \end{pmatrix} \cdot \begin{pmatrix} a_0^1 & a_1^1 & a_2^1 \end{pmatrix} \\
&= \begin{pmatrix} z_0^2 \\ z_1^2 \end{pmatrix} \circ \begin{pmatrix} a_0^2 - y_0 \\ a_1^2 - y_1 \end{pmatrix} \cdot \begin{pmatrix} a_0^1 & a_1^1 & a_2^1 \end{pmatrix} = \begin{pmatrix} z_0^2 a_0^2 - z_0^2 y_0 \\ z_1^2 a_1^2 - z_1^2 y_1 \end{pmatrix} \cdot \begin{pmatrix} a_0^1 & a_1^1 & a_2^1 \end{pmatrix} \\
&= \begin{pmatrix} a_0^1 z_0^2 a_0^2 - a_0^1 z_0^2 y_0 & a_1^1 z_0^2 a_0^2 - a_1^1 z_0^2 y_0 & a_2^1 z_0^2 a_0^2 - a_2^1 z_0^2 y_0 \\ a_0^1 z_1^2 a_1^2 - a_0^1 z_1^2 y_1 & a_1^1 z_1^2 a_1^2 - a_1^1 z_1^2 y_1 & a_2^1 z_1^2 a_1^2 - a_2^1 z_1^2 y_1 \end{pmatrix} \\
&= \begin{pmatrix} d_{00}^2 & d_{01}^2 & d_{02}^2 \\ d_{10}^2 & d_{11}^2 & d_{12}^2 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
W_{new}^2 = \begin{pmatrix} w_{00}^2 & w_{01}^2 & w_{02}^2 \\ w_{10}^2 & w_{11}^2 & w_{12}^2 \end{pmatrix} - \eta \frac{\partial C}{\partial W^2} &= \begin{pmatrix} w_{00}^2 & w_{01}^2 & w_{02}^2 \\ w_{10}^2 & w_{11}^2 & w_{12}^2 \end{pmatrix} - \eta \begin{pmatrix} d_{00}^2 & d_{01}^2 & d_{02}^2 \\ d_{10}^2 & d_{11}^2 & d_{12}^2 \end{pmatrix} \\
&= \begin{pmatrix} w_{00}^2 & w_{01}^2 & w_{02}^2 \\ w_{10}^2 & w_{11}^2 & w_{12}^2 \end{pmatrix} - \begin{pmatrix} \eta d_{00}^2 & \eta d_{01}^2 & \eta d_{02}^2 \\ \eta d_{10}^2 & \eta d_{11}^2 & \eta d_{12}^2 \end{pmatrix} \\
&= \begin{pmatrix} w_{00}^2 - \eta d_{00}^2 & w_{01}^2 - \eta d_{01}^2 & w_{02}^2 - \eta d_{02}^2 \\ w_{10}^2 - \eta d_{10}^2 & w_{11}^2 - \eta d_{11}^2 & w_{12}^2 - \eta d_{12}^2 \end{pmatrix} = \begin{pmatrix} w_{00}^2 & w_{01}^2 & w_{02}^2 \\ w_{10}^2 & w_{11}^2 & w_{12}^2 \end{pmatrix}_{new}
\end{aligned}$$

$$\begin{aligned}
B_{new}^2 = \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix} - \eta \frac{\partial C}{\partial z^2} &= \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix} - \eta \sigma'(z^2)(a^2 - y) \\
&= \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix} - \eta \begin{pmatrix} \frac{1}{1+e^{-z_0^2}}\left(1 - \frac{1}{1+e^{-z_0^2}}\right) \\ \frac{1}{1+e^{-z_1^2}}\left(1 - \frac{1}{1+e^{-z_1^2}}\right) \end{pmatrix} \circ \begin{pmatrix} a_0^2 - y_0 \\ a_1^2 - y_1 \end{pmatrix} \\
&= \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix} - \eta \begin{pmatrix} z_0^2 \\ z_1^2 \end{pmatrix} \circ \begin{pmatrix} a_0^2 - y_0 \\ a_1^2 - y_1 \end{pmatrix} = \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix} - \eta \begin{pmatrix} z_0^2 a_0^2 - z_0^2 y_0 \\ z_1^2 a_1^2 - z_1^2 y_1 \end{pmatrix} = \begin{pmatrix} b_0^2 \\ b_1^2 \end{pmatrix}_{new}
\end{aligned}$$

$$\frac{\partial C}{\partial W^1} = \frac{\partial C}{\partial z^1} \cdot \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}^T = \left(\frac{\partial a^1}{\partial z^1} \frac{\partial z^2}{\partial a^1} \frac{\partial C}{\partial z^2} \right) \cdot \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}^T = (W^2)^T \cdot \frac{\partial C}{\partial z^2} \circ g_1'(z^1) \cdot \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}^T$$

$$= \begin{pmatrix} w_{00}^2 & w_{10}^2 \\ w_{01}^2 & w_{11}^2 \\ w_{02}^2 & w_{12}^2 \end{pmatrix} \cdot \begin{pmatrix} d_0 \\ d_1 \end{pmatrix} \circ \begin{pmatrix} r_0 \\ r_1 \\ r_2 \end{pmatrix} \cdot \begin{pmatrix} x_0 & x_1 \end{pmatrix} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} \circ \begin{pmatrix} r_0 \\ r_1 \\ r_2 \end{pmatrix} \cdot \begin{pmatrix} x_0 & x_1 \end{pmatrix}$$

$$= \begin{pmatrix} p_0 r_0 \\ p_1 r_1 \\ p_2 r_2 \end{pmatrix} \cdot \begin{pmatrix} x_0 & x_1 \end{pmatrix} = \begin{pmatrix} p_0 r_0 x_0 & p_0 r_0 x_1 \\ p_1 r_1 x_0 & p_1 r_1 x_1 \\ p_2 r_2 x_0 & p_2 r_2 x_1 \end{pmatrix} = \begin{pmatrix} d_{00}^1 & d_{01}^1 \\ d_{10}^1 & d_{11}^1 \\ d_{20}^1 & d_{21}^1 \end{pmatrix}$$

$$W_{new}^1 = \begin{pmatrix} w_{00}^1 & w_{01}^1 \\ w_{10}^1 & w_{11}^1 \\ w_{20}^1 & w_{21}^1 \end{pmatrix} - \eta \frac{\partial C}{\partial W^1} = \begin{pmatrix} w_{00}^1 & w_{01}^1 \\ w_{10}^1 & w_{11}^1 \\ w_{20}^1 & w_{21}^1 \end{pmatrix} - \eta \begin{pmatrix} d_{00}^1 & d_{01}^1 \\ d_{10}^1 & d_{11}^1 \\ d_{20}^1 & d_{21}^1 \end{pmatrix} = \begin{pmatrix} w_{00}^1 & w_{01}^1 \\ w_{10}^1 & w_{11}^1 \\ w_{20}^1 & w_{21}^1 \end{pmatrix}_{new}$$

$$B_{new}^1 = \begin{pmatrix} b_0^1 \\ b_1^1 \\ b_2^1 \end{pmatrix} - \eta \frac{\partial C}{\partial z^1} = \begin{pmatrix} b_0^1 \\ b_1^1 \\ b_2^1 \end{pmatrix} - \eta \begin{pmatrix} p_0 r_0 \\ p_1 r_1 \\ p_2 r_2 \end{pmatrix} = \begin{pmatrix} b_0^1 \\ b_1^1 \\ b_2^1 \end{pmatrix}_{new}$$