For backpropagation, we need to perform differentiation of the activation functions, for this we had 3 activation functions:

Sigmoid:

$$\text{Sigmoid} = 1/(1+e^{-x})$$

$$\frac{d}{dx} = \frac{d}{dx}(1/1+e^{-x}) = (1+e^{-x})^{-1}$$

$$= -(1+e^{-x})^{-2}\frac{d}{dx}(1+e^{-x})$$

$$= -\frac{1}{(1+e^{-x})^2}\frac{d}{dx}(1+e^{-x})$$

$$= -\frac{1}{(1+e^{-x})^2}e^{-x}\frac{d}{dx}(-x)$$

$$= \frac{-1}{(1+e^{-x})^2}e^{-x}(-1) = \frac{e^{-x}}{(1+e^{-x})^2}$$

Adding $+1-1$ to the numerator yields:

$$\frac{d}{dx} = \frac{e^{-x}+1-1}{(1+e^{-x})^2} = \left(\frac{1}{1+e^{-x}}\right)\left(\frac{e^{-x}+1-1}{(1+e^{-x})}\right)$$

$$= \left(\frac{1}{1+e^{-x}}\right)\left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right)$$

$$= \left(\frac{1}{1+e^{-x}}\right)\left(1 - \frac{1}{1+e^{-x}}\right)$$

We know $\frac{1}{1+e^{-x}}$ is the sigmoid function $\sigma(z)$

$$\therefore \frac{d}{dx} = \sigma(z)(1-\sigma(z))$$

$\therefore$ differentiation for sigmoid yields:

$$\frac{d}{dx} = \text{sigmoid}(1-\text{sigmoid})$$

Relu

$$\text{Relu} = \max(0,x)$$

if $x>0$, $\frac{\partial(\text{relu})}{\partial x} = 1$

Otherwise, $\frac{\partial(\text{relu})}{\partial x} = 0$

Softmax

$$\text{Softmax} = e^{x_a}\Big/\sum_{q=1}^{n} e^{x_q}$$

$$= e^{x_1}/(e^{x_1}+e^{x_2}+e^{x_3})$$

$$\frac{\partial(\text{softmax})}{\partial x_1} = \frac{\partial}{\partial x}(e^{x_1})(e^{x_1}+e^{x_2}+e^{x_3})^{-1}$$

Using chain rule from calculus:

$$\frac{\partial}{\partial x_1} = (e^{x_1})\frac{\partial}{\partial x}(e^{x_1}+e^{x_2}+e^{x_3})^{-1} + \frac{\partial}{\partial x}(e^{x_1})(e^{x_1}+e^{x_2}+e^{x_3})^{-1}$$

$$= -(e^{x_1})(e^{x_1}+e^{x_2}+e^{x_3})^{-2}(e^{x_1})(1) + (e^{x_1})(1)(e^{x_1}+e^{x_2}+e^{x_3})^{-1}$$

$$= -\frac{(e^{x_1})^2}{(e^{x_1}+e^{x_2}+e^{x_3})^2} + \frac{e^{x_1}}{e^{x_1}+e^{x_2}+e^{x_3}}$$

Multiplying the right side by $(e^{x_1}+e^{x_2}+e^{x_3})$ yields:

$$= \frac{-(e^{x_1})^2}{(e^{x_1}+e^{x_2}+e^{x_3})} + \frac{e^{x_1}(e^{x_1}+e^{x_2}+e^{x_3})}{(e^{x_1}+e^{x_2}+e^{x_3})(e^{x_1}+e^{x_2}+e^{x_3})}$$

$$= \frac{(-e^{x_1})^2+(e^{x_1})^2+e^{x_1}(e^{x_2}+e^{x_3})}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$$

$$= \frac{\cancel{(-e^{x_1})^2}+\cancel{(e^{x_1})^2}+e^{x_1}(e^{x_2}+e^{x_3})}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$$

$$\frac{\partial(\text{softmax})}{\partial x_1} = \frac{e^{x_1}(e^{x_2}+e^{x_3})}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$$