3 Layer Classification MLP Differentials with Sigmoid as Activation Function

$\frac{dC}{dW^{[3]}}$ Calculation:

$$\frac{dC}{dW^{[3]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial W^{[3]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial W^{[3]}} = \frac{d}{dW^{[3]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = A^{[2]}$$

$$\frac{dC}{dW^{[3]}} = (A^{[3]} - Y)A^{[2]}$$

$\frac{dC}{db^{[3]}}$ Calculation:

$$\frac{dC}{db^{[3]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial b^{[3]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial b^{[3]}} = \frac{d}{db^{[3]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = 1$$

$$\frac{dC}{db^{[3]}} = (A^{[3]} - Y)$$

$\frac{dC}{dW^{[2]}}$ Calculation:

$$\frac{dC}{dW^{[2]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial A^{[2]}} \frac{\partial A^{[2]}}{\partial Z^{[2]}} \frac{\partial Z^{[2]}}{\partial W^{[2]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(sigmoid(Z^{[2]}\right) = A^{[2]}(1 - A^{[2]})$$

$$\frac{\partial Z^{[2]}}{\partial W^{[2]}} = \frac{d}{dW^{[2]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = A^{[1]}$$

$$\frac{dC}{dW^{[2]}} = (A^{[3]} - Y)\, W^{[3]}A^{[2]}(1 - A^{[2]})A^{[1]}$$

$\frac{dC}{db^{[2]}}$ Calculation:

$$\frac{dC}{db^{[2]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial A^{[2]}} \frac{\partial A^{[2]}}{\partial Z^{[2]}} \frac{\partial Z^{[2]}}{\partial b^{[2]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(sigmoid(Z^{[2]}\right) = A^{[2]}(1 - A^{[2]})$$

$$\frac{\partial Z^{[2]}}{\partial b^{[2]}} = \frac{d}{db^{[2]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = 1$$

$$\frac{dC}{db^{[2]}} = (A^{[3]} - Y)\, W^{[3]}A^{[2]}(1 - A^{[2]})$$

$\frac{dC}{dW^{[1]}}$ Calculation:

$$\frac{dC}{dW^{[1]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial A^{[2]}} \frac{\partial A^{[2]}}{\partial Z^{[2]}} \frac{\partial Z^{[2]}}{\partial A^{[1]}} \frac{\partial A^{[1]}}{\partial Z^{[1]}} \frac{\partial Z^{[1]}}{\partial W^{[1]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(sigmoid(Z^{[2]}\right) = A^{[2]}(1 - A^{[2]})$$

$$\frac{\partial Z^{[2]}}{\partial A^{[1]}} = \frac{d}{dA^{[1]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = W^{[2]}$$

$$\frac{\partial A^{[1]}}{\partial Z^{[1]}} = \frac{d}{dZ^{[1]}}\left(sigmoid(Z^{[1]}\right) = A^{[1]}(1 - A^{[1]})$$

$$\frac{\partial Z^{[1]}}{\partial W^{[1]}} = \frac{d}{dW^{[1]}}\left(W^{[1]}A^{[0]} + b^{[1]}\right) = A^{[0]} = X$$

$$\frac{dC}{dW^{[1]}} = (A^{[3]} - Y)\, W^{[3]} A^{[2]}(1 - A^{[2]})\, W^{[2]} A^{[1]}(1 - A^{[1]})\, X$$

$\frac{dC}{db^{[1]}}$ Calculation:

$$\frac{dC}{db^{[1]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial A^{[2]}} \frac{\partial A^{[2]}}{\partial Z^{[2]}} \frac{\partial Z^{[2]}}{\partial A^{[1]}} \frac{\partial A^{[1]}}{\partial Z^{[1]}} \frac{\partial Z^{[1]}}{\partial b^{[1]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(sigmoid(Z^{[2]}\right) = A^{[2]}(1 - A^{[2]})$$

$$\frac{\partial Z^{[2]}}{\partial A^{[1]}} = \frac{d}{dA^{[1]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = W^{[2]}$$

$$\frac{\partial A^{[1]}}{\partial Z^{[1]}} = \frac{d}{dZ^{[1]}}\left(sigmoid(Z^{[1]}\right) = A^{[1]}(1 - A^{[1]})$$

$$\frac{\partial Z^{[1]}}{\partial b^{[1]}} = \frac{d}{db^{[1]}}\left(W^{[1]}A^{[0]} + b^{[1]}\right) = 1$$

$$\frac{dC}{db^{[1]}} = (A^{[3]} - Y)\, W^{[3]} A^{[2]}(1 - A^{[2]})\, W^{[2]} A^{[1]}(1 - A^{[1]})$$

3 Layer Classification MLP Differentials with Hyperbolic Tangent as Activation Function

$\frac{dC}{dW^{[3]}}$ Calculation:

$$\frac{dC}{dW^{[3]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial W^{[3]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial W^{[3]}} = \frac{d}{dW^{[3]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = A^{[2]}$$

$$\frac{dC}{dW^{[3]}} = (A^{[3]} - Y)A^{[2]}$$

$\frac{dC}{db^{[3]}}$ Calculation:

$$\frac{dC}{db^{[3]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial b^{[3]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial b^{[3]}} = \frac{d}{db^{[3]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = 1$$

$$\frac{dC}{db^{[3]}} = (A^{[3]} - Y)$$

$\frac{dC}{dW^{[2]}}$ Calculation:

$$\frac{dC}{dW^{[2]}} = \frac{\partial C}{\partial Z^{[3]}}\frac{\partial Z^{[3]}}{\partial A^{[2]}}\frac{\partial A^{[2]}}{\partial Z^{[2]}}\frac{\partial Z^{[2]}}{\partial W^{[2]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(tanh(Z^{[2]}\right) = (1 - \tanh^2(Z^{[2]}))$$

$$\frac{\partial Z^{[2]}}{\partial W^{[2]}} = \frac{d}{dW^{[2]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = A^{[1]}$$

$$\frac{dC}{dW^{[2]}} = (A^{[3]} - Y)\, W^{[3]}(1 - \tanh^2(Z^{[2]}))A^{[1]}$$

$\frac{dC}{db^{[2]}}$ Calculation:

$$\frac{dC}{db^{[2]}} = \frac{\partial C}{\partial Z^{[3]}}\frac{\partial Z^{[3]}}{\partial A^{[2]}}\frac{\partial A^{[2]}}{\partial Z^{[2]}}\frac{\partial Z^{[2]}}{\partial b^{[2]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(tanh(Z^{[2]}\right) = (1 - \tanh^2(Z^{[2]}))$$

$$\frac{\partial Z^{[2]}}{\partial b^{[2]}} = \frac{d}{db^{[2]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = 1$$

$$\frac{dC}{db^{[2]}} = (A^{[3]} - Y)\, W^{[3]}(1 - \tanh^2(Z^{[2]}))$$

$\frac{dC}{dW^{[1]}}$ Calculation:

$$\frac{dC}{dW^{[1]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial A^{[2]}} \frac{\partial A^{[2]}}{\partial Z^{[2]}} \frac{\partial Z^{[2]}}{\partial A^{[1]}} \frac{\partial A^{[1]}}{\partial Z^{[1]}} \frac{\partial Z^{[1]}}{\partial W^{[1]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(tanh(Z^{[2]}\right) = (1 - \tanh^2(Z^{[2]}))$$

$$\frac{\partial Z^{[2]}}{\partial A^{[1]}} = \frac{d}{dA^{[1]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = W^{[2]}$$

$$\frac{\partial A^{[1]}}{\partial Z^{[1]}} = \frac{d}{dZ^{[1]}}\left(tanh(Z^{[1]}\right) = (1 - \tanh^2(Z^{[1]}))$$

$$\frac{\partial Z^{[1]}}{\partial W^{[1]}} = \frac{d}{dW^{[1]}}\left(W^{[1]}A^{[0]} + b^{[1]}\right) = A^{[0]} = X$$

$$\frac{dC}{dW^{[1]}} = (A^{[3]} - Y)\, W^{[3]}(1 - \tanh^2(Z^{[2]}))W^{[2]}(1 - \tanh^2(Z^{[1]}))X$$

$\frac{dC}{db^{[1]}}$ Calculation:

$$\frac{dC}{db^{[1]}} = \frac{\partial C}{\partial Z^{[3]}} \frac{\partial Z^{[3]}}{\partial A^{[2]}} \frac{\partial A^{[2]}}{\partial Z^{[2]}} \frac{\partial Z^{[2]}}{\partial A^{[1]}} \frac{\partial A^{[1]}}{\partial Z^{[1]}} \frac{\partial Z^{[1]}}{\partial b^{[1]}}$$

$$\frac{\partial C}{\partial Z^{[3]}} = (A^{[3]} - Y)$$

$$\frac{\partial Z^{[3]}}{\partial A^{[2]}} = \frac{d}{dA^{[2]}}\left(W^{[3]}A^{[2]} + b^{[3]}\right) = W^{[3]}$$

$$\frac{\partial A^{[2]}}{\partial Z^{[2]}} = \frac{d}{dZ^{[2]}}\left(tanh(Z^{[2]}\right) = (1 - \tanh^2(Z^{[2]}))$$

$$\frac{\partial Z^{[2]}}{\partial A^{[1]}} = \frac{d}{dA^{[1]}}\left(W^{[2]}A^{[1]} + b^{[2]}\right) = W^{[2]}$$

$$\frac{\partial A^{[1]}}{\partial Z^{[1]}} = \frac{d}{dZ^{[1]}}\left(tanh(Z^{[1]}\right) = (1 - \tanh^2(Z^{[1]}))$$

$$\frac{\partial Z^{[1]}}{\partial b^{[1]}} = \frac{d}{db^{[1]}}\left(W^{[1]}A^{[0]} + b^{[1]}\right) = 1$$

$$\frac{dC}{db^{[1]}} = (A^{[3]} - Y)\, W^{[3]}A^{[2]}(1 - \tanh^2(Z^{[2]}))\, W^{[2]}(1 - \tanh^2(Z^{[1]}))$$