INITIAL FORMULAS.

$cost_{avg} = \frac{1}{n} \cdot \sum_{s=0}^{n-1} cost_s$

$cost_s = \sum_{j=0}^{n_L - 1} cost(a_j^L, y_j)$

$a_j^L = \sigma(z_j^L)$

$z_j^L = b_j^L + \sum_{k=0}^{n_{(L-1)}-1} w_{jk}^L \cdot a_k^{L-1}$

DERIVATIVES.

COST CHANGE WITH RESPECT TO A:

$\frac{\partial cost_{avg}}{\partial A} = \frac{1}{n} \cdot \sum_{s=0}^{n-1} \frac{\partial cost_s}{\partial A}$

$\frac{\partial cost_s}{\partial A} = \sum_{j=0}^{n_L - 1} \frac{\partial cost(a_j^L, y_j)}{\partial A}$

SAMPLE'S COST CHANGE WITH RESPECT TO WEIGHTS:

$\frac{\partial cost(a_j^L, y_j)}{\partial w_{ji}^L} = \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial a_j^L}{\partial w_{ji}^L}$

$\frac{\partial a_j^L}{\partial w_{ji}^L} = \frac{\partial \sigma(z_j^L)}{\partial w_{ji}^L} = \frac{\partial \sigma(z_j^L)}{\partial z_j^L} \cdot \frac{\partial z_j^L}{\partial w_{ji}^L}$

$\frac{\partial z_j^L}{\partial w_{ji}^L} = \frac{\partial b_j^L}{\partial w_{ji}^L} + \sum_{k=0}^{n_{(L-1)}-1} \frac{\partial w_{jk}^L \cdot a_k^{L-1}}{\partial w_{ji}^L} = a_i^{L-1}$

$\frac{\partial cost(a_j^L, y_j)}{\partial w_{ji}^L} = \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial \sigma(z_j^L)}{\partial z_j^L} \cdot a_i^{L-1}$

SAMPLE'S COST CHANGE WITH RESPECT TO BIASES:

$\frac{\partial cost(a_j^L, y_j)}{\partial b_j^L} = \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial \sigma(z_j^L)}{\partial z_j^L} \cdot \frac{\partial z_j^L}{\partial b_j^L}$

$\frac{\partial z_j^L}{\partial b_j^L} = \frac{\partial b_j^L}{\partial b_j^L} + \sum_{k=0}^{n_{(L-1)}-1} \frac{\partial w_{jk}^L \cdot a_k^{L-1}}{\partial b_j^L} = 1$

$\frac{\partial cost(a_j^L, y_j)}{\partial b_j^L} = \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial \sigma(z_j^L)}{\partial z_j^L}$

SAMPLE'S COST CHANGE WITH RESPECT TO THE PREVIOUS LAYER (INPUT):

$\frac{\partial cost(a_j^L, y_j)}{\partial a_i^{L-1}} = \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial \sigma(z_j^L)}{\partial z_j^L} \cdot \frac{\partial z_j^L}{\partial a_i^{L-1}}$

$\frac{\partial z_j^L}{\partial a_i^{L-1}} = \frac{\partial b_j^L}{\partial a_i^{L-1}} + \sum_{k=0}^{n_{(L-1)}-1} \frac{\partial w_{jk}^L \cdot a_k^{L-1}}{\partial a_i^{L-1}} = w_{ji}^L$

$\frac{\partial cost(a_j^L, y_j)}{\partial a_i^{L-1}} = \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial \sigma(z_j^L)}{\partial z_j^L} \cdot w_{ji}^L$

COST CHANGE WITH RESPECT TO WEIGHTS, BIASES AND THE PREVIOUS LAYER:

$\frac{\partial cost_s}{\partial w_{ik}^L} = \sum_{j=0}^{n_L-1} \frac{\partial cost(a_j^L, y_j)}{\partial w_{ik}^L} = \frac{\partial cost(a_i^L, y_i)}{\partial a_i^L} \cdot \frac{\partial \sigma(z_i^L)}{\partial z_i^L} \cdot a_k^{L-1}$

$\frac{\partial cost_s}{\partial b_i^L} = \sum_{j=0}^{n_L-1} \frac{\partial cost(a_j^L, y_j)}{\partial b_i^L} = \frac{\partial cost(a_i^L, y_i)}{\partial a_i^L} \cdot \frac{\partial \sigma(z_i^L)}{\partial z_i^L}$

$\frac{\partial cost_s}{\partial a_k^{L-1}} = \sum_{j=0}^{n_L-1} \frac{\partial cost(a_j^L, y_j)}{\partial a_k^{L-1}} = \sum_{j=0}^{n_L-1} \frac{\partial cost(a_j^L, y_j)}{\partial a_j^L} \cdot \frac{\partial \sigma(z_j^L)}{\partial z_j^L} \cdot w_{jk}^L$