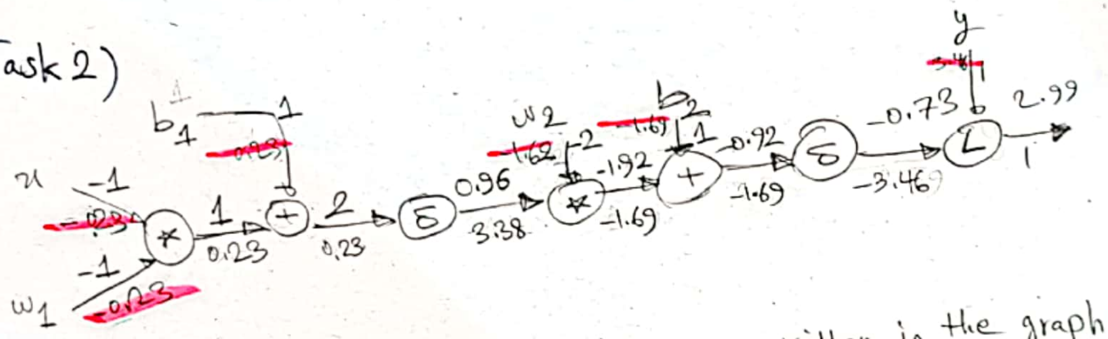
Amir hossein Eyvazkhani
1747696
Ex8

Task 1)

$$\frac{\partial \sigma_1(x)}{\partial x} = \frac{-(-e^{-x})}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}}\left(1-\frac{1}{1+e^{-x}}\right) = \sigma_1(x)\left(1-\sigma_1(x)\right)$$

$$\frac{\partial \sigma_2(x)}{\partial x} = \frac{(e^{x}+e^{-x})^2-(e^{x}-e^{-x})^2}{(e^{x}+e^{-x})^2} = 1-\left(\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}\right)^2 = 1-\sigma_2^2(x)$$

Task 2)



The forward pass values are written in the graph

$$Loss = (f(x)-y)^2 = (-0.73-1)^2 = 2.99$$

$$\frac{\partial L}{\partial L} = 1 \;,\; \frac{\partial L}{\partial y} = -2(f(x)-y) \;,\; \frac{\partial L}{\partial \sigma} = 2(f(x)-y)$$

$$\frac{\partial \sigma}{\partial z} = 1-\sigma^2(z)$$

The backward pass is also denoted on the graph.
The upstream gradient * local gradient = down stream gredient
is used for calculations based on the slides.
The final gredient with respect to each variable is highlighted

$$\eta = 1 \Rightarrow \begin{cases} w_1' = w_1 - \eta\frac{\partial L}{\partial w_1} = -1-(-0.23) = -0.77 \\ b_1' = b_1 - \eta\frac{\partial L}{\partial b_1} = 1-(0.23) = 0.77 \\ w_2' = w_2 - \eta\frac{\partial L}{\partial w_2} = -2-(-1.62) = -0.38 \\ b_2' = b_2 - \eta\frac{\partial L}{\partial b_2} = 1-(-1.69) = 2.69 \end{cases}$$