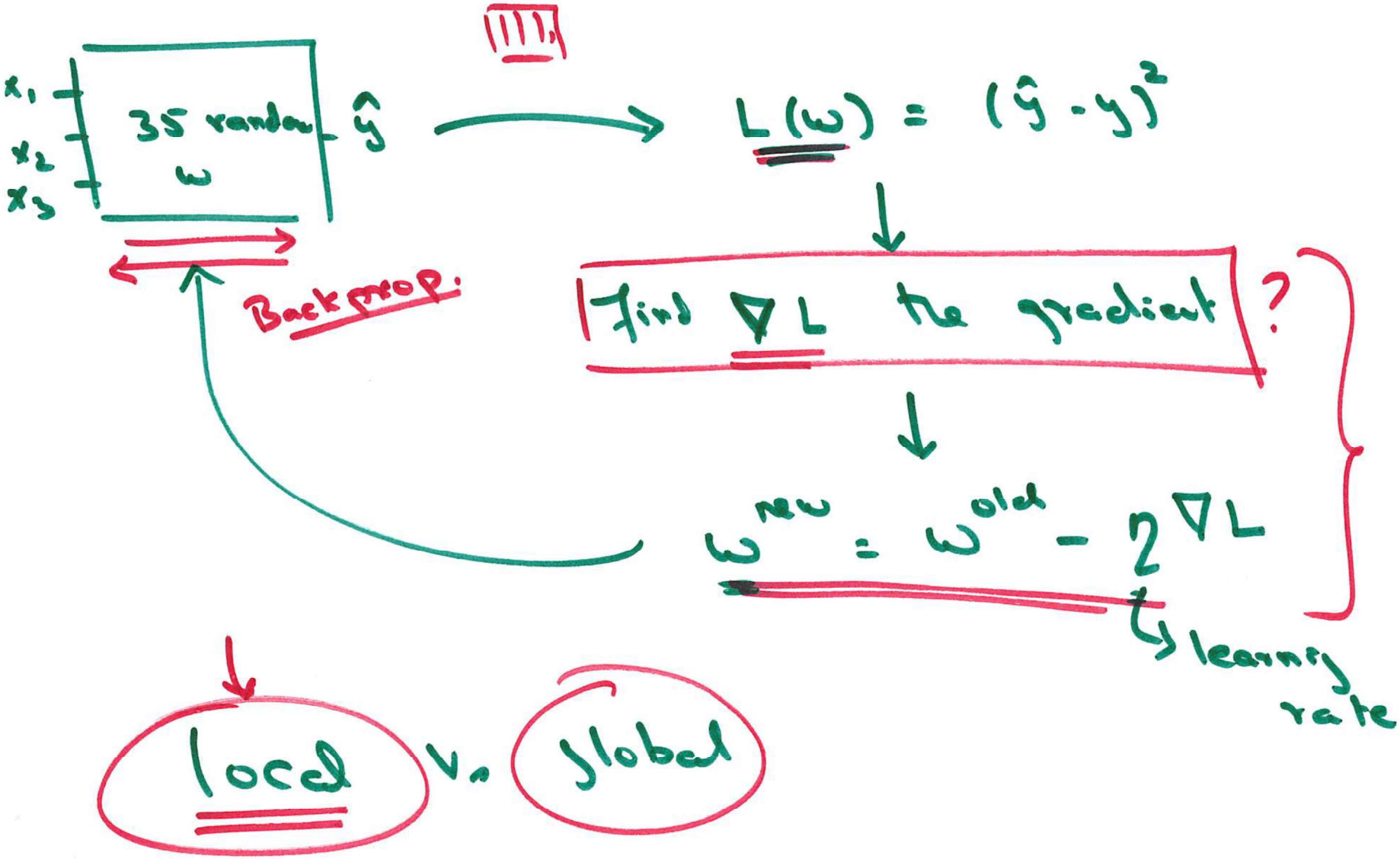
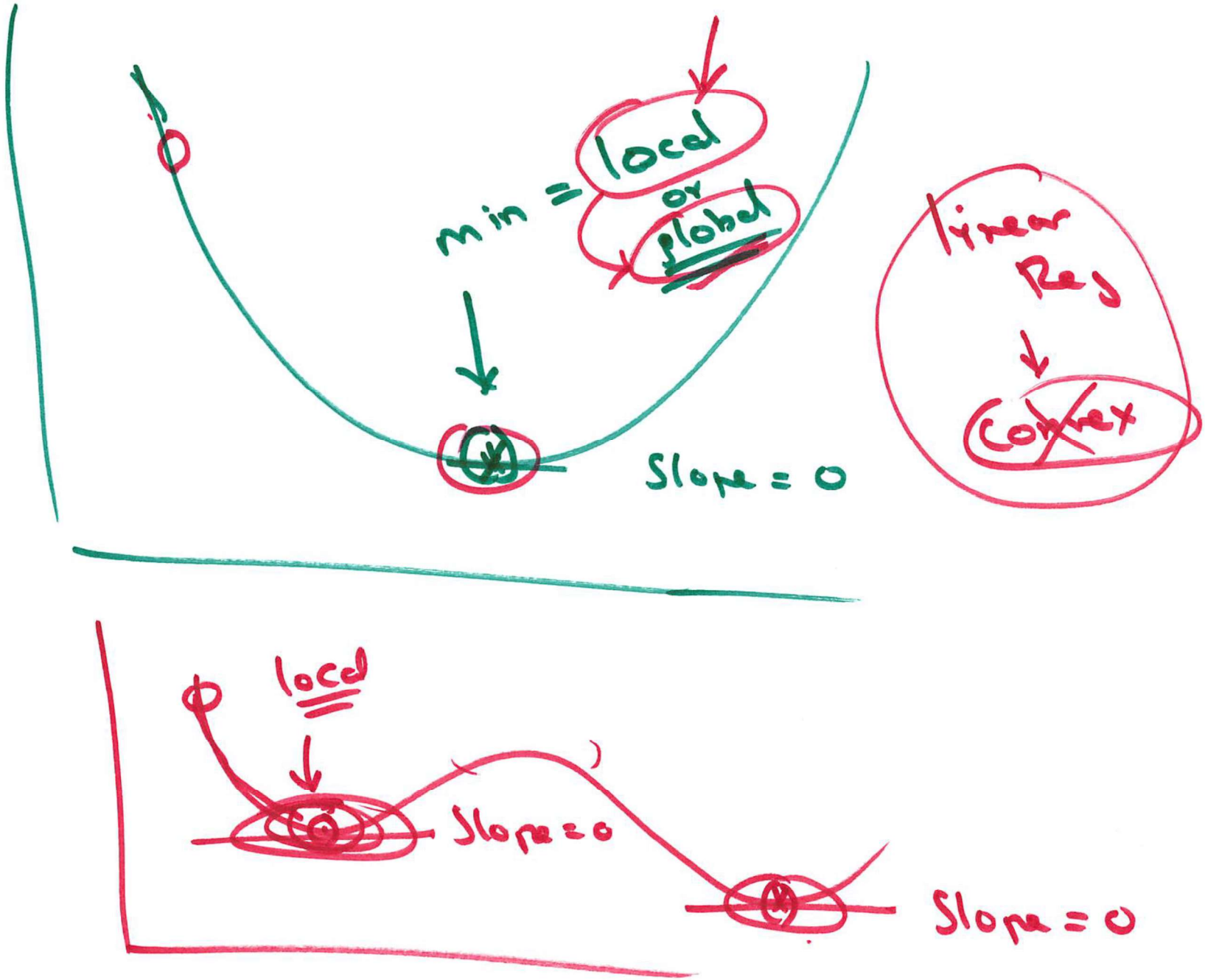
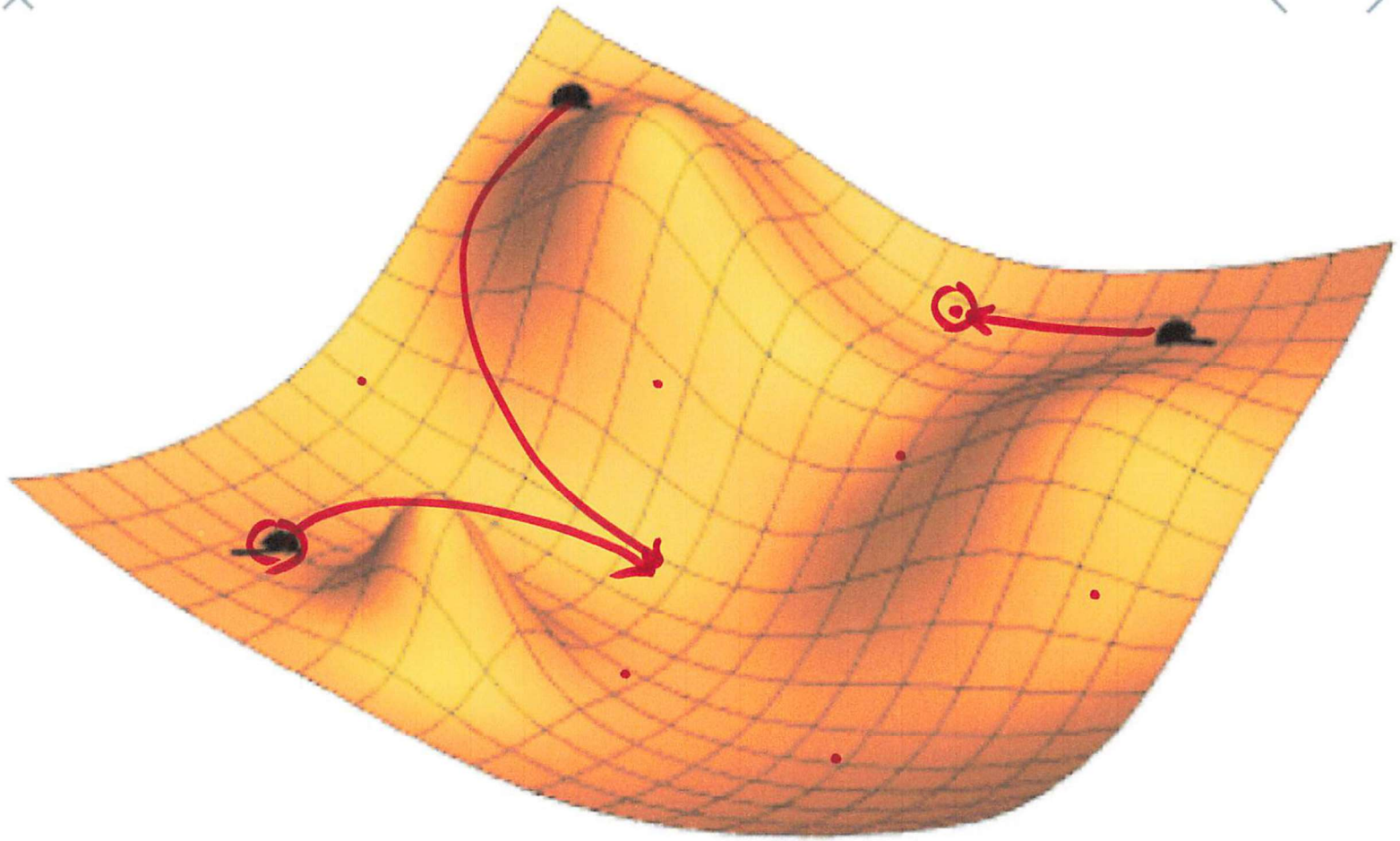
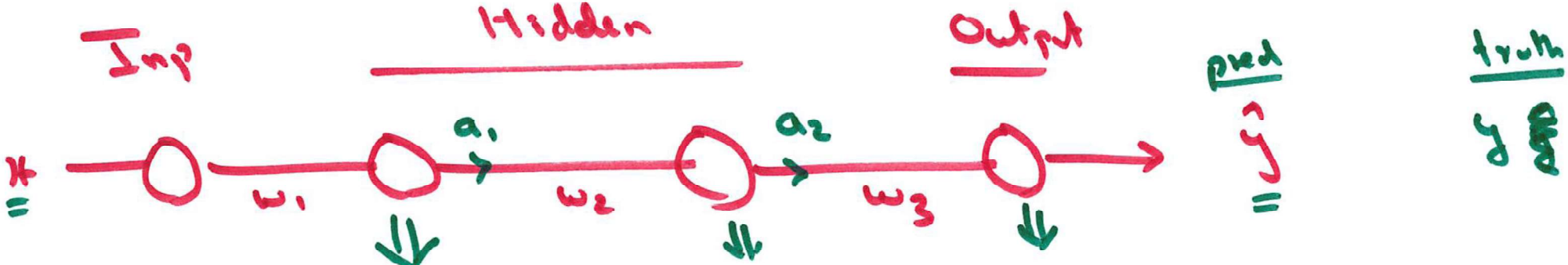




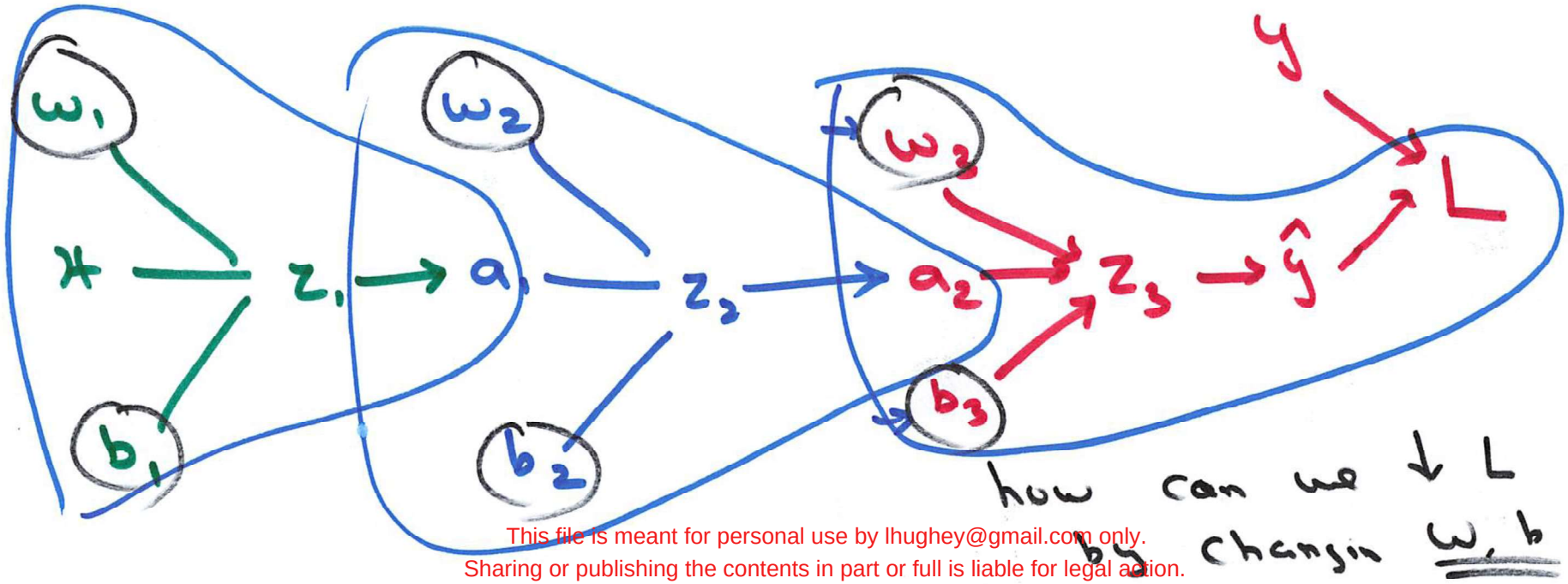




$$z_1 = x w_1 + b_1 \qquad z_2 = a_1 w_2 + b_2 \qquad z_3 = a_2 w_3 + b_3$$
$$a_1 = \sigma(z_1) \qquad a_2 = \sigma(z_2) \qquad \hat{y} = \sigma(z_3)$$

$$L(w) = \Sigma(\hat{y} - y)^2$$

∇L ?



how can we ↓ L
by changin w, b

Slope of L with respect to ω_3 $= \dfrac{\partial L}{\partial \omega_3}$

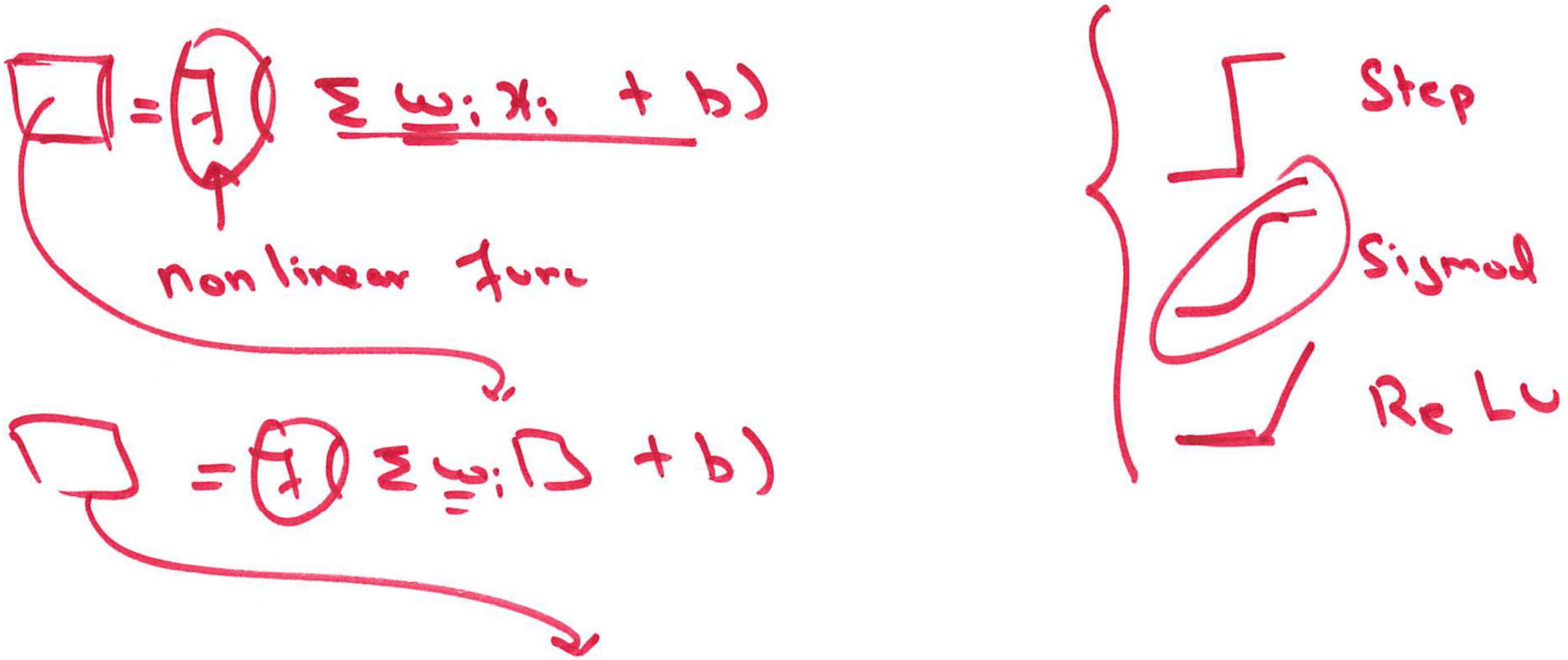
$$\frac{\partial L}{\partial \omega_3} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_3} \cdot \frac{\partial z_3}{\partial \omega_3}$$

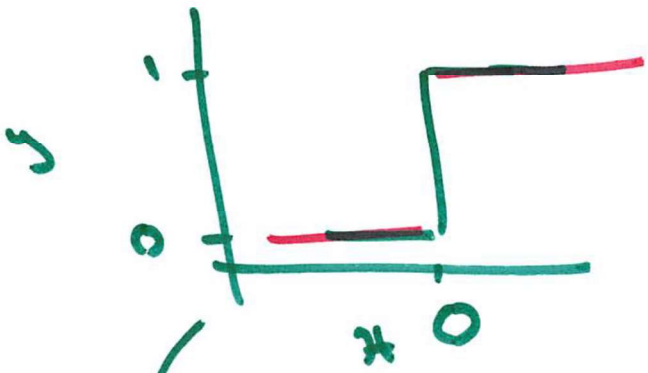
$\dfrac{\partial L}{\partial \hat{y}} \rightarrow 2(\hat{y} - y)$

$\dfrac{\partial \hat{y}}{\partial z_3} \rightarrow \sigma'(z_3)$

$\dfrac{\partial z_3}{\partial \omega_3} \rightarrow a_2$

$$\frac{\partial a_2}{\partial \omega_2}$$

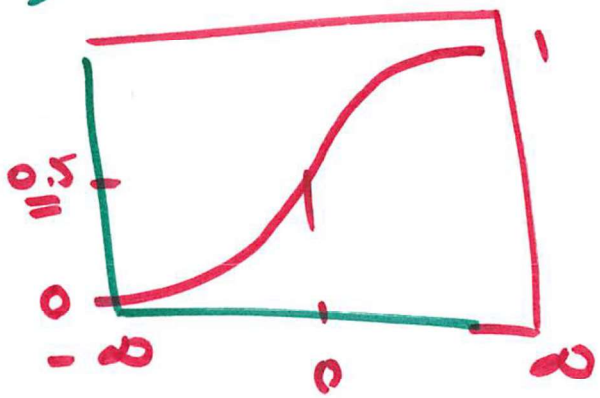




$$y = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

$y = c$ $\quad$ $y = f(x)$

$m = \omega = \omega_0 - 2(\text{slope of } L)$



Sigmoid → 0 to 1 $\quad \sigma'(x)$

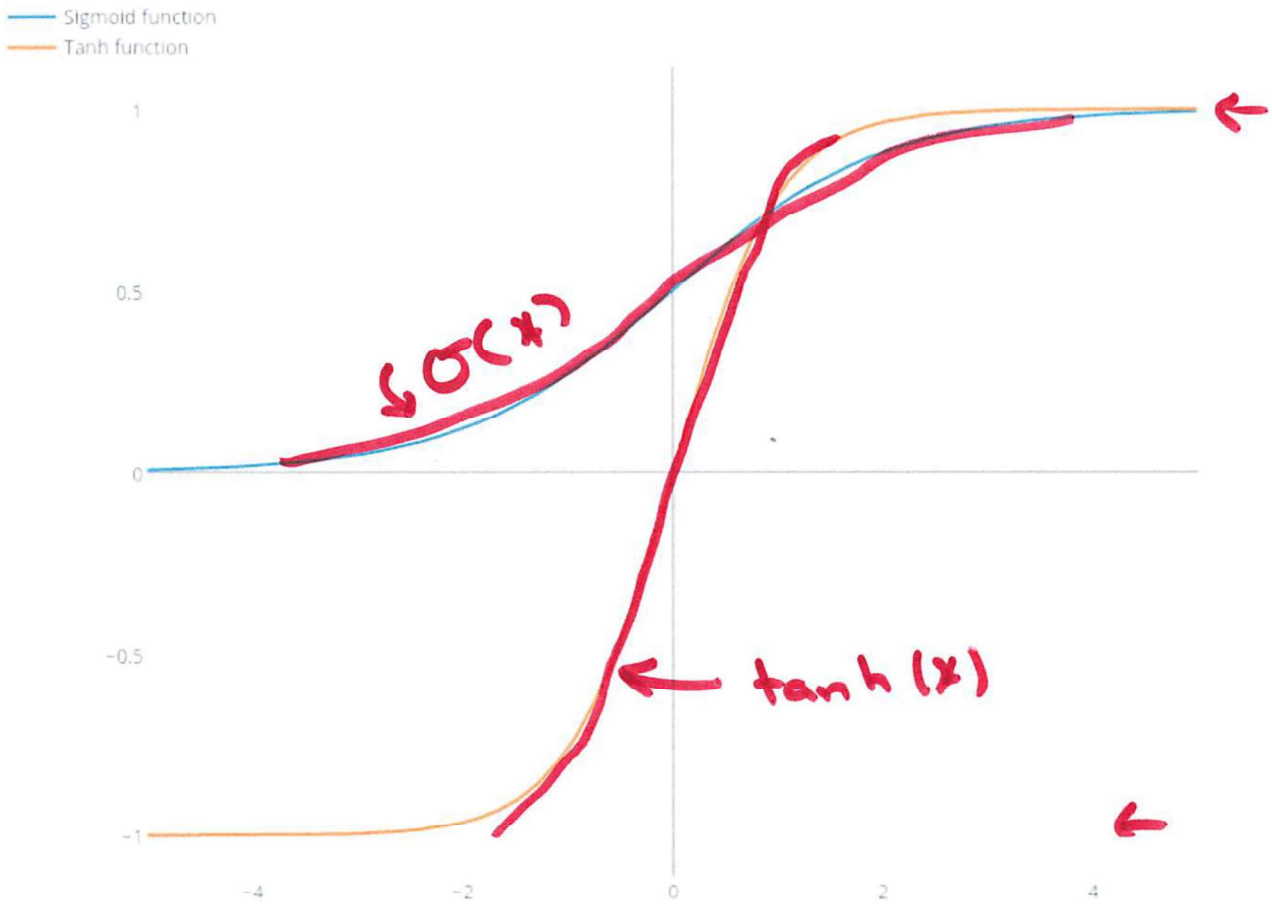
$$y = \sigma(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}$$

$2\sigma(x) - 1 \rightarrow -1$ to 1

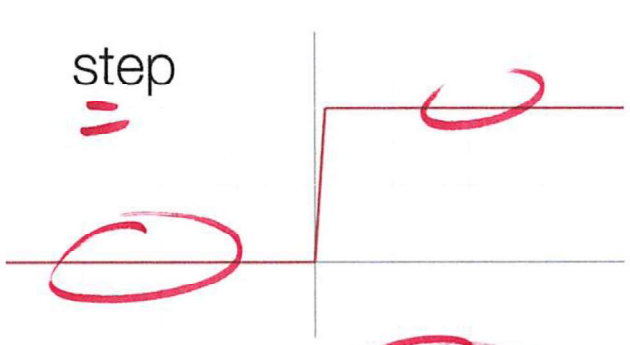
$$y = \tanh(x) = 2\sigma(2x) - 1 \rightarrow -1 \text{ to } 1$$

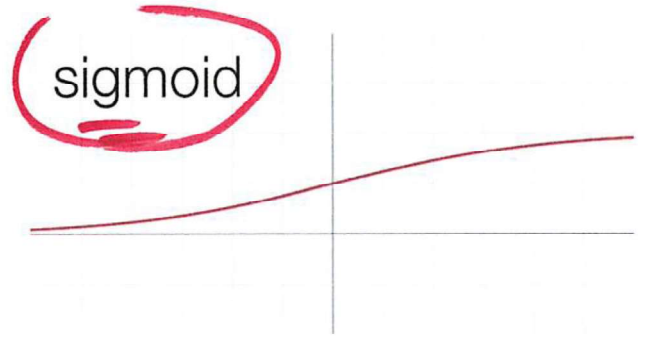
$$\hookrightarrow \frac{e^x - e^{-x}}{e^x + e^{-x}} \equiv \frac{e^{2x}-1}{e^{2x}+1}$$

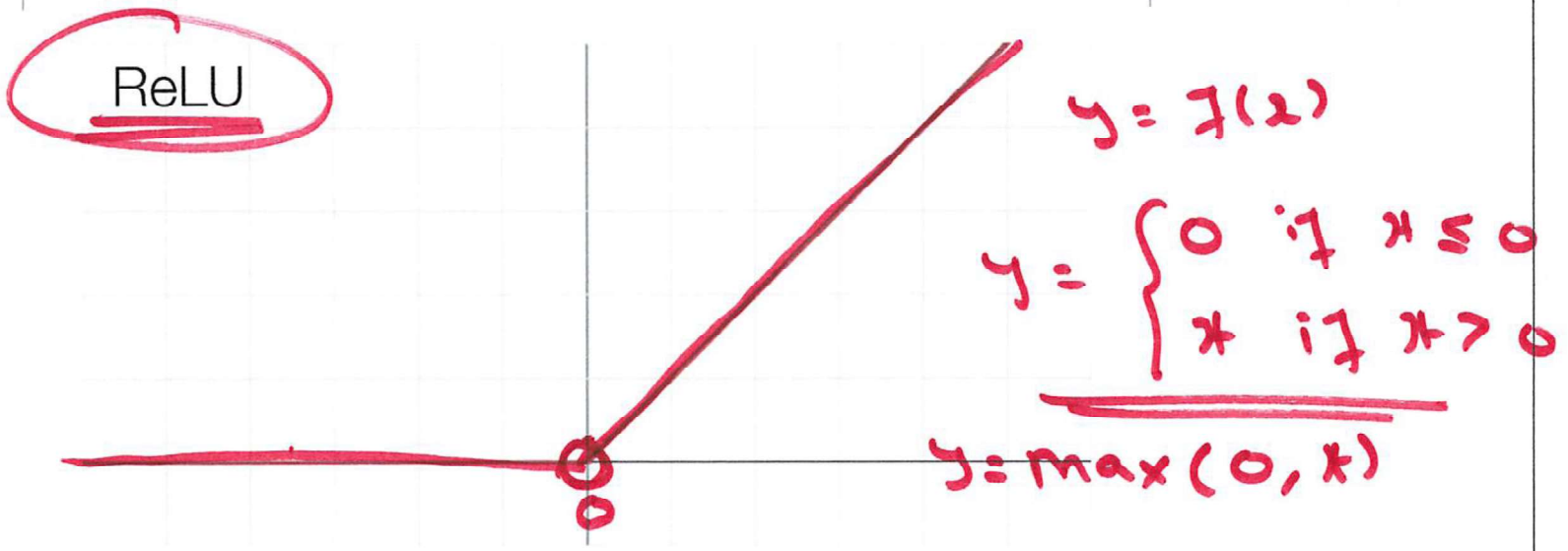
$$\tanh(x) = 2\sigma(2x) - 1$$

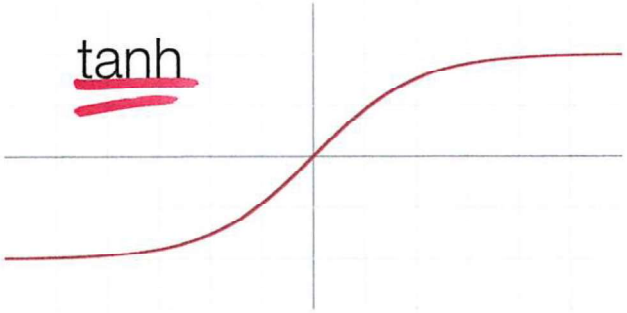


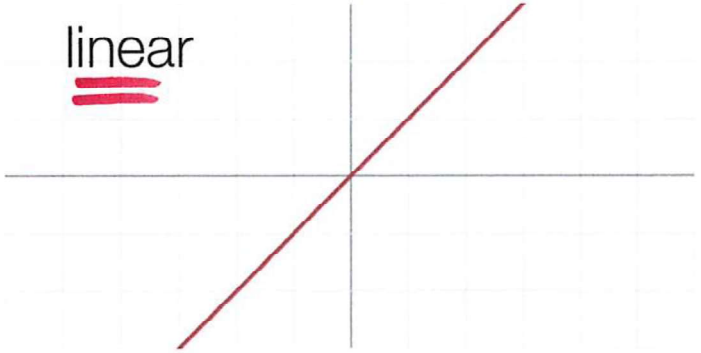
Activation Functions



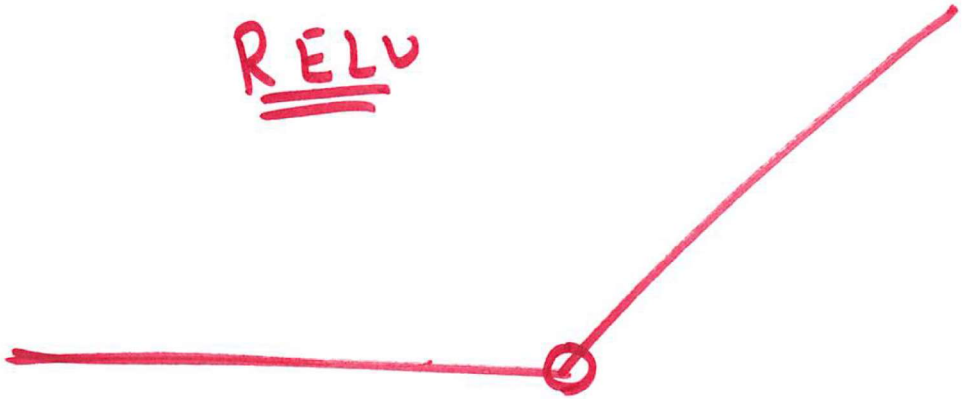


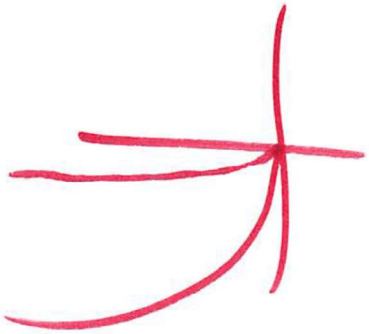




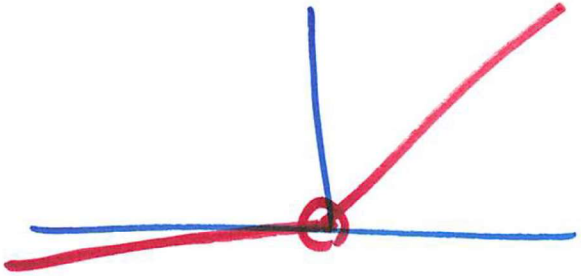


RELU



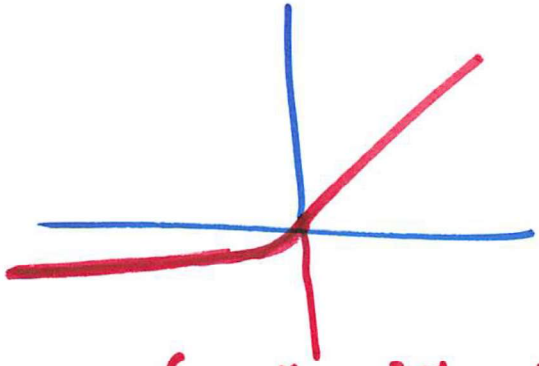


Leaky RELU

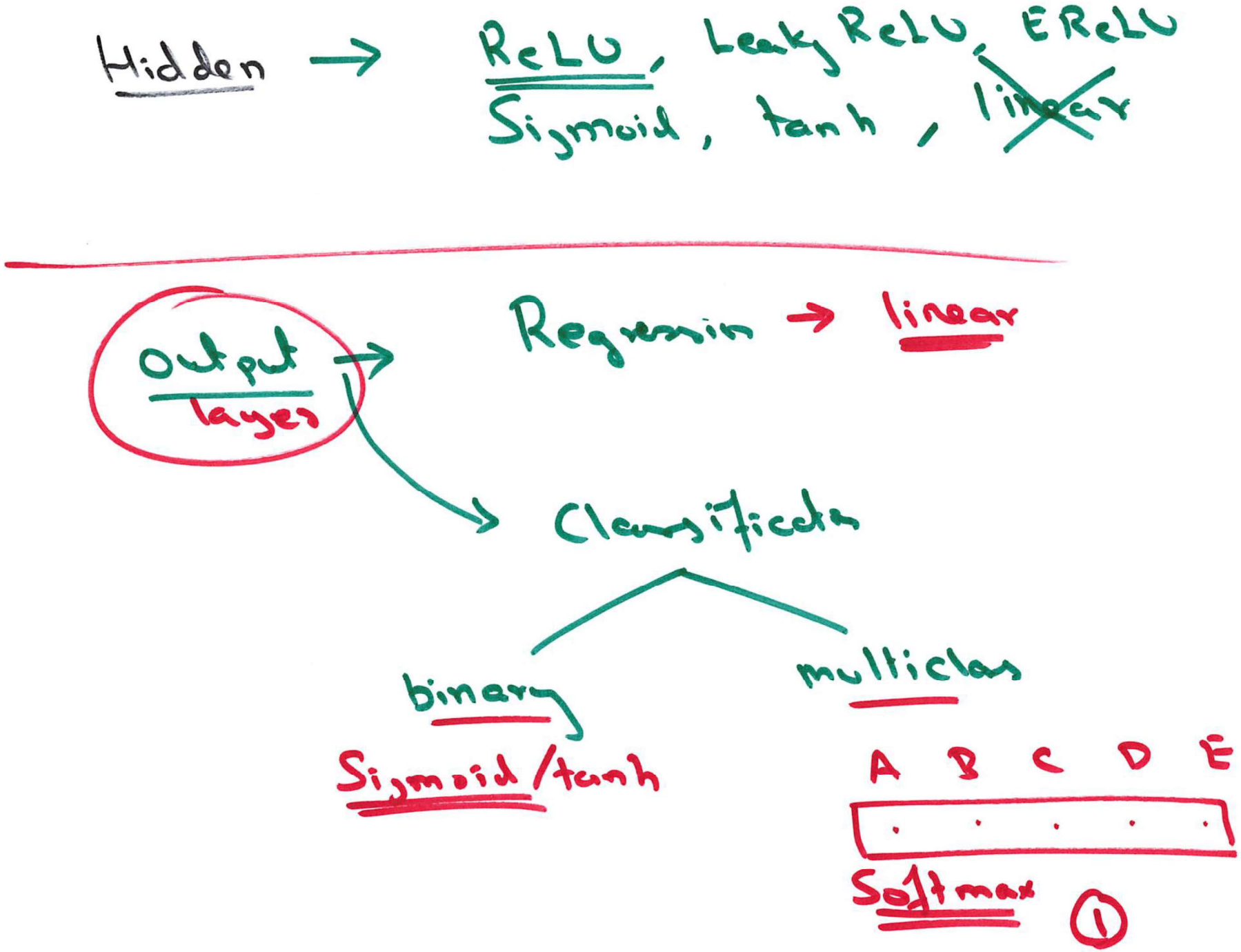


$$y = \begin{cases} x & \text{if } x \geq 0 \\ 0.001x & \text{if } x < 0 \end{cases}$$

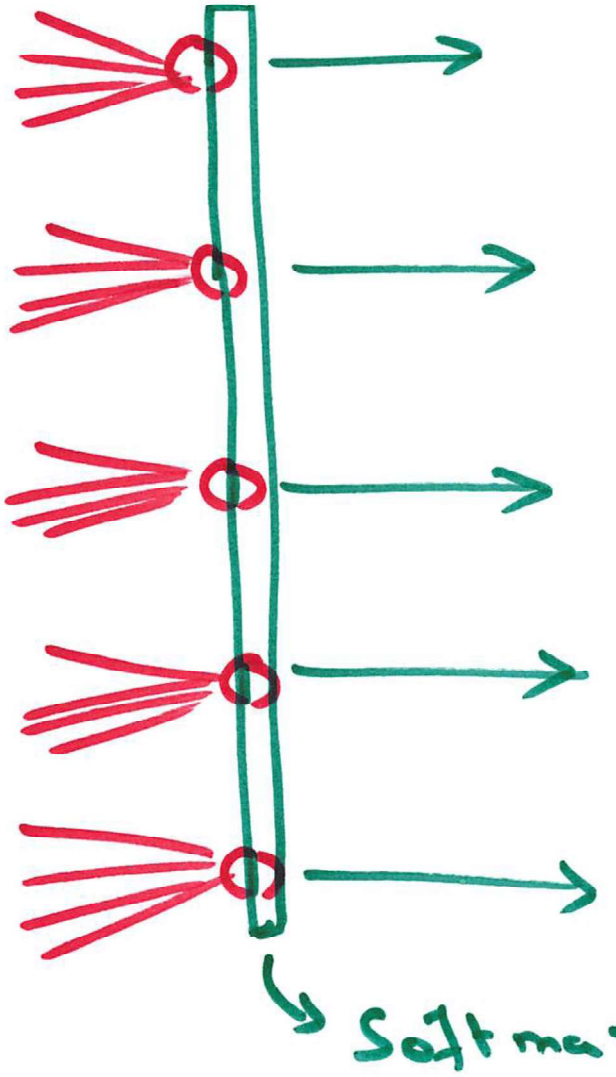
ERELU



$$y = \begin{cases} x & \text{if } x \geq 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases}$$



$\mathcal{T}(\ldots) \rightarrow$



↳ Softmax activation

$z_1 = (\Sigma w x + b)$

$a_1 = f(z_1)$