



$$\frac{\partial \mathcal{L}}{\partial w_1} = \frac{2}{B} (\hat{y} - y) \cdot o_1$$

$$\frac{\partial \mathcal{L}}{\partial w_2} = \frac{2}{B} (\hat{y} - y) \cdot w_2^T \cdot I[h_1 > 0] \cdot x$$