

$$\frac{\partial a}{\partial w_2} = aw_2$$
  $\nabla_{\vec{w}} a = (\vec{w}^T \vec{x} - y) \vec{x}^T$   $\frac{\partial \sigma}{\partial a} = (1 - \sigma(a)) \sigma(a)$