$$J(\alpha; x_k, y_k) = \frac{1}{2} \| h_\alpha(x_k) - y_k \|^2$$

↑ output of NN.

$\| \cdot \|^2$ is the "two norm" i.e.

$$\sum_{i=1}^{n} a_i^2$$

n_ℓ = #layers, incl. input and output.

s_ℓ: #nodes in layer ℓ

$$J(\alpha) = \left[\frac{1}{n} \sum_{k=1}^{n} J(\alpha; x_k, y_k) \right] + \frac{\lambda}{2} \sum_{\ell=1}^{n_\ell - 1} \sum_{i=1}^{s_\ell} \sum_{j=1}^{s_{\ell+1}} \left[\alpha_{ji}^{(\ell)} \right]^2$$