Notation

Layer l:

- output of layer l : $\mathbf{y}_{(l)}$
- input of layer l : $\mathbf{y}_{(l-1)}$
- layer 0 (*the input layer*)
  - $\mathbf{y}_{(0)} = \mathbf{x}$
- layer L (output layer)
  - $\hat{y} = \mathbf{y}$
    $= \mathbf{y}_{(L)}$
- layer $(L+1)$ (*"Loss layer"*)
  - $\mathbf{y}_{(L+1)}$
    $= \mathcal{L}$