



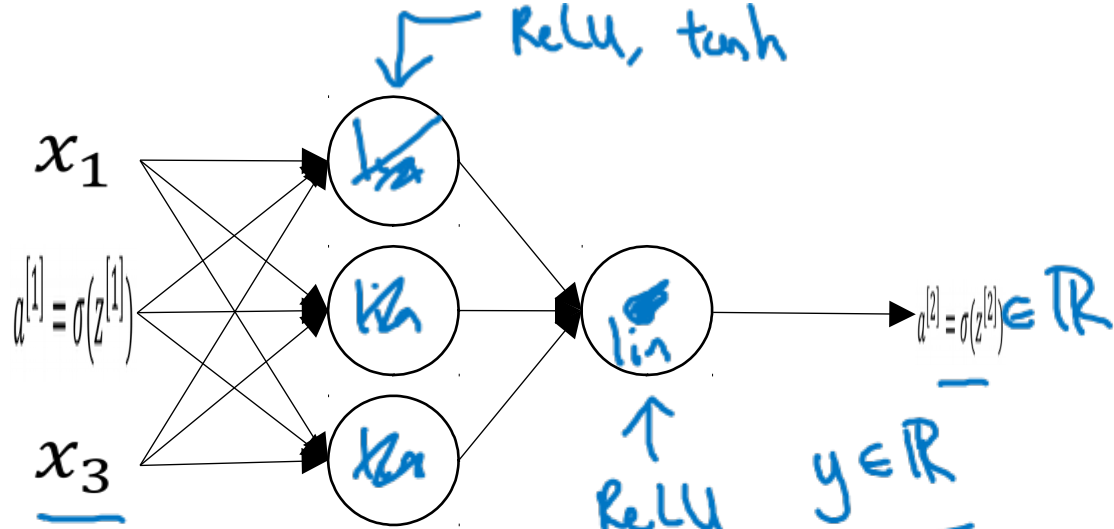
**deeplearning.ai**

# One hidden layer Neural Network

---

Why do you  
need non-linear  
activation functions?

# Activation function



Given  $x$ :

$$\begin{aligned}
 &\rightarrow z^{[1]} = W^{[1]}x + b^{[1]} \\
 &\rightarrow a^{[1]} = \cancel{g^{[1]}(z^{[1]})} \quad z^{[1]} \\
 &\rightarrow z^{[2]} = W^{[2]}a^{[1]} + b^{[2]} \\
 &\rightarrow a^{[2]} = \cancel{g^{[2]}(z^{[2]})} \quad z^{[2]}
 \end{aligned}$$

$g(z) = z$   
 "linear activation function"

$$\begin{aligned}
 a^{[1]} = z^{[1]} &= W^{[1]}x + b^{[1]} \\
 a^{[2]} = z^{[2]} &= W^{[2]}a^{[1]} + b^{[2]}
 \end{aligned}$$

$$a^{[2]} = W^{[2]} \left( \underbrace{W^{[1]}x + b^{[1]}}_{a^{[1]}} \right) + b^{[2]}$$

$$\begin{aligned}
 &= \underbrace{(W^{[2]}W^{[1]})}_w x + \underbrace{(W^{[2]}b^{[1]} + b^{[2]})}_b \\
 &= \underline{w'x + b'} \\
 &g(z) = z
 \end{aligned}$$