



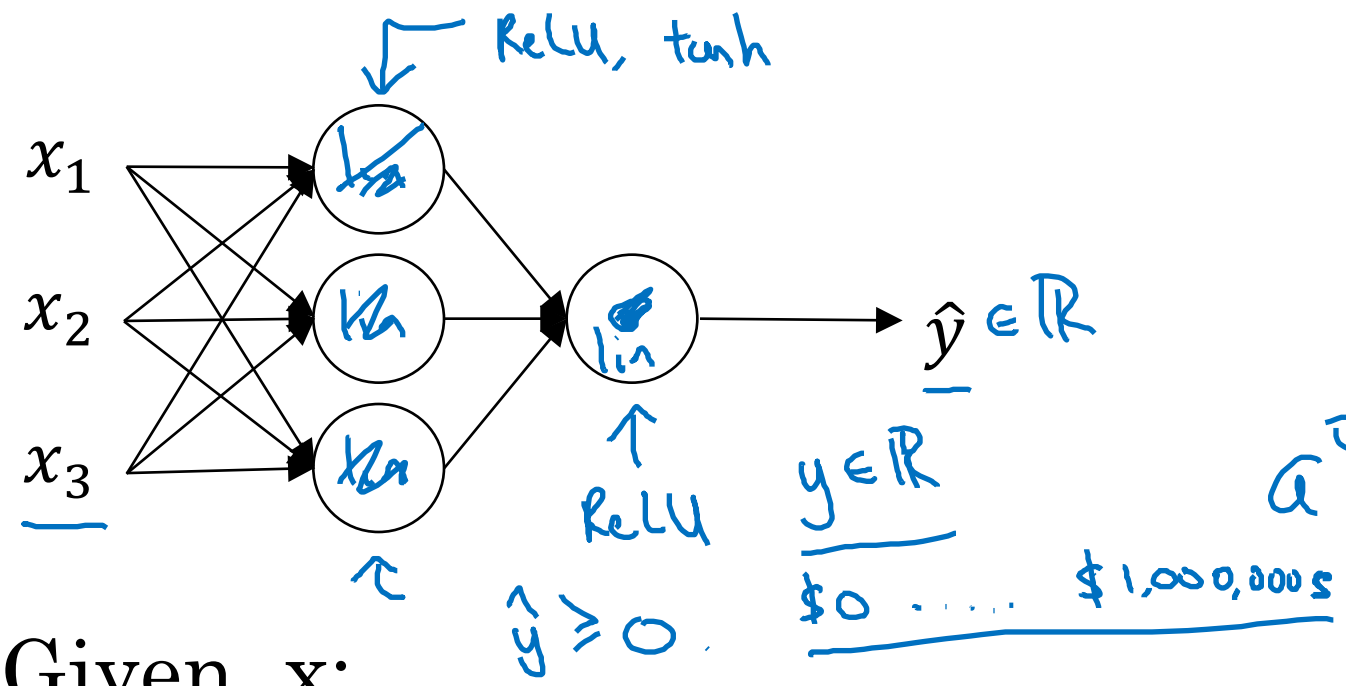
deeplearning.ai

# One hidden layer Neural Network

---

Why do you  
need non-linear  
activation functions?

# Activation function



Given  $x$ :

- $z^{[1]} = W^{[1]}x + b^{[1]}$
- $a^{[1]} = \cancel{g^{[1]}(z^{[1]})} z^{[1]}$
- $z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}$
- $a^{[2]} = \cancel{g^{[2]}(z^{[2]})} z^{[2]}$

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

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