

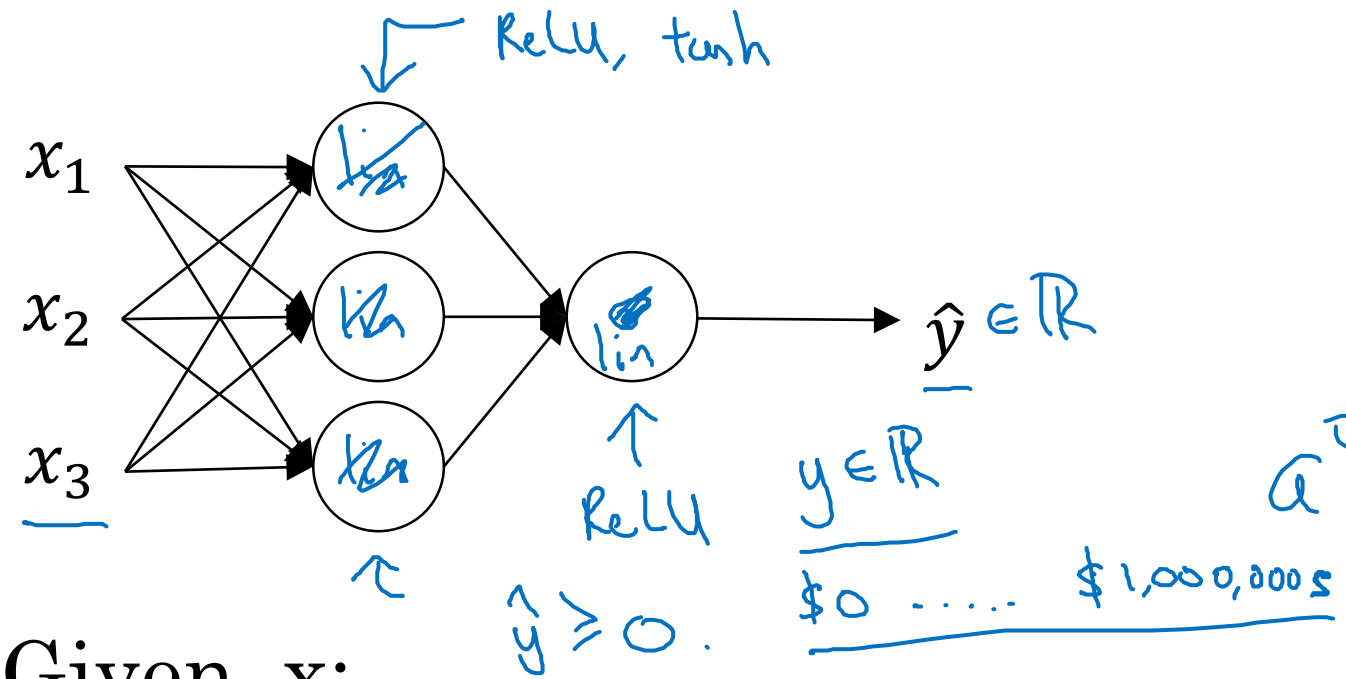


deeplearning.ai

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"

$$a^{[1]} = z^{[1]} = W^{[1]}x + b^{[1]}$$

$$a^{[2]} = z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}$$

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

$a^{[1]}$

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

w' b'

$$= w'x + b'$$

$$g(z) = z$$