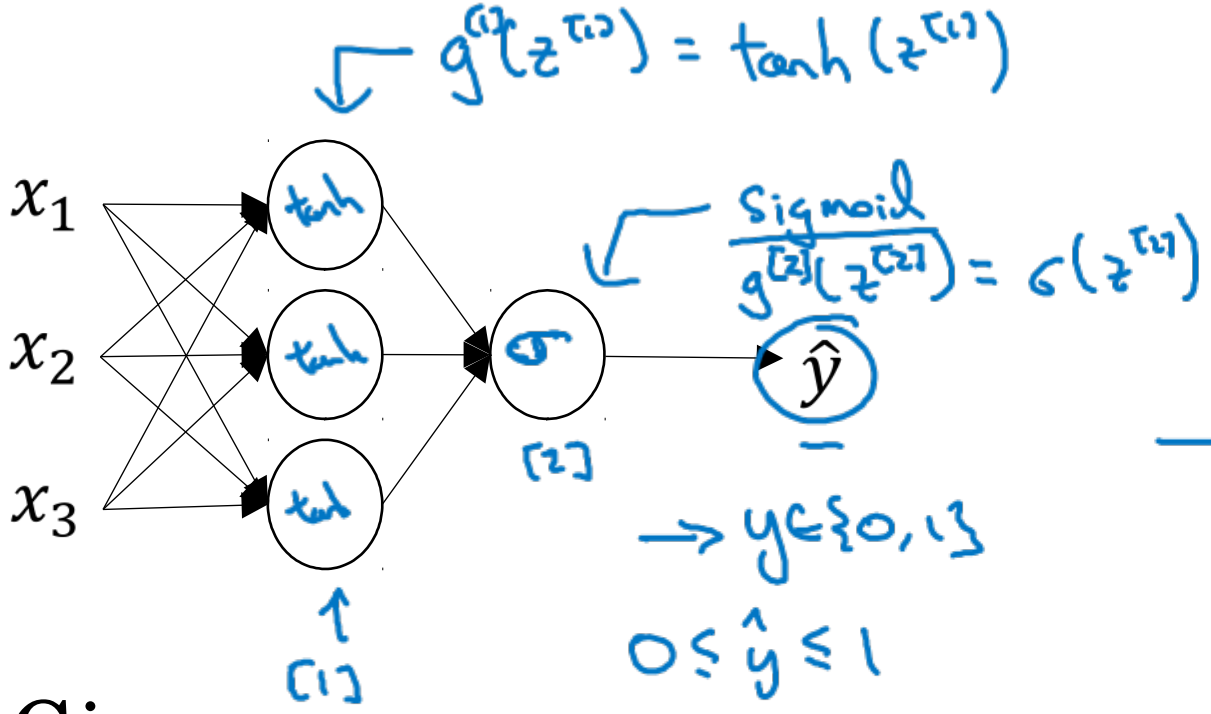
deeplearning.ai

One hidden layer Neural Network

Activation functions

Activation functions



$g^{[1]}(z^{[1]}) = \tanh(z^{[1]})$

Sigmoid
$g^{[2]}(z^{[2]}) = \sigma(z^{[1]})$

$\rightarrow y \in \{0, 1\}$

$0 \le \hat{y} \le 1$

$-1 \quad 1$

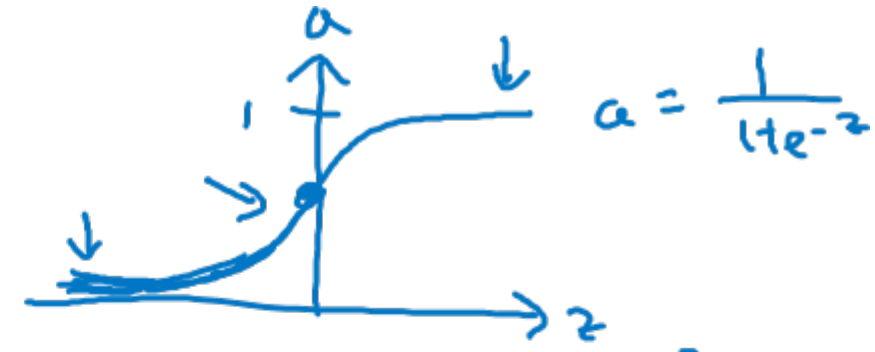
Given x:

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

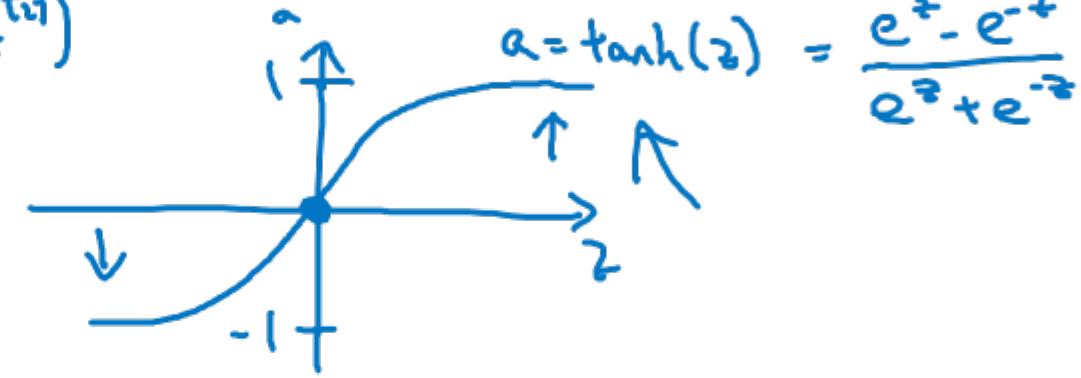
$\rightarrow$ $g^{[1]}(z^{[1]})$

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

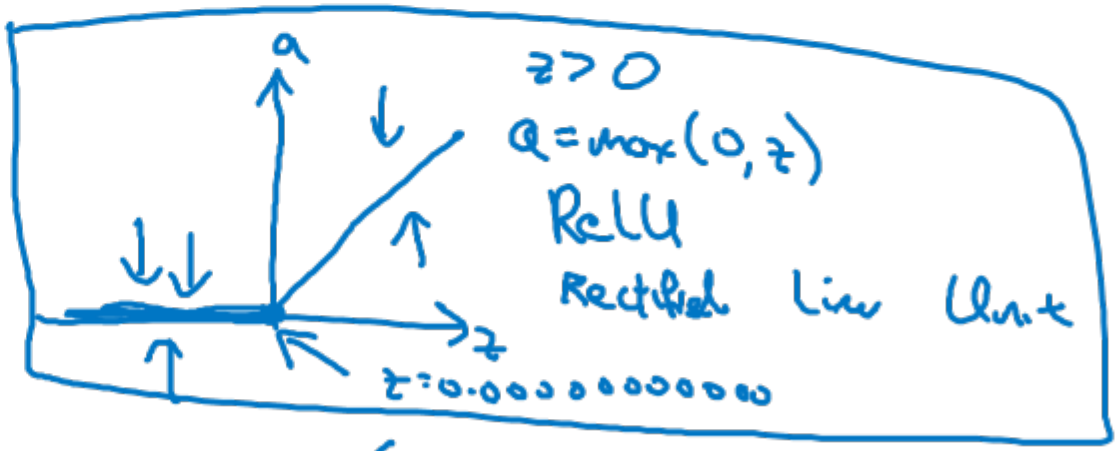
$\rightarrow$ $g^{[2]}(z^{[2]})$



$a = \frac{1}{1+e^{-z}}$



$a = \tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$



$z > 0$

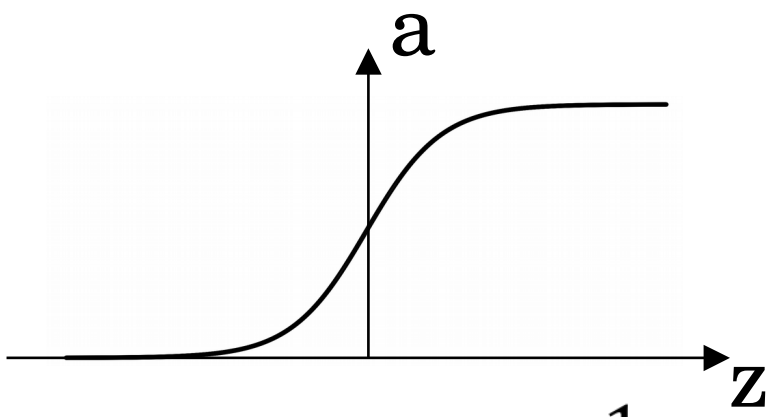
$a = \max(0, z)$

ReLU

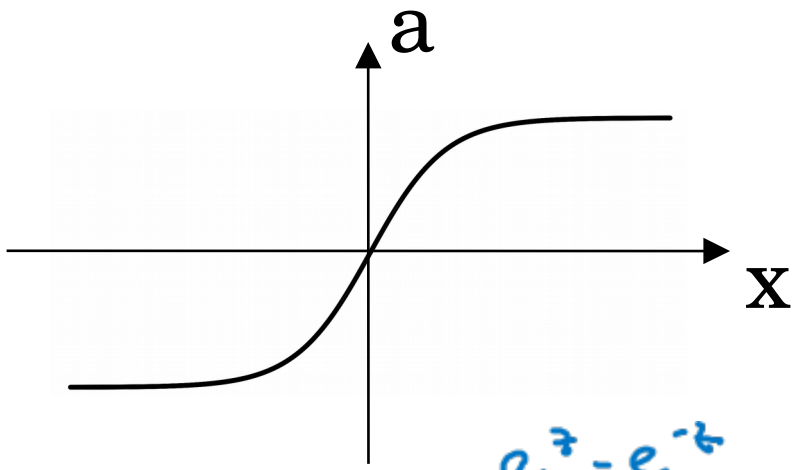
Rectified Linear Unit

Leaky ReLU

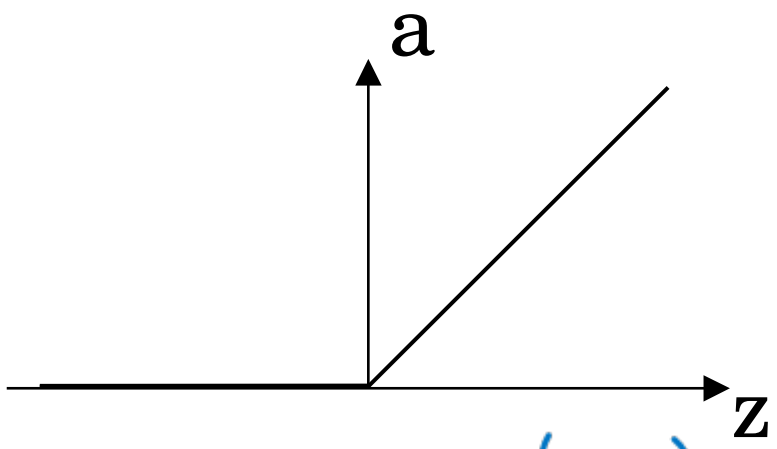
Pros and cons of activation functions



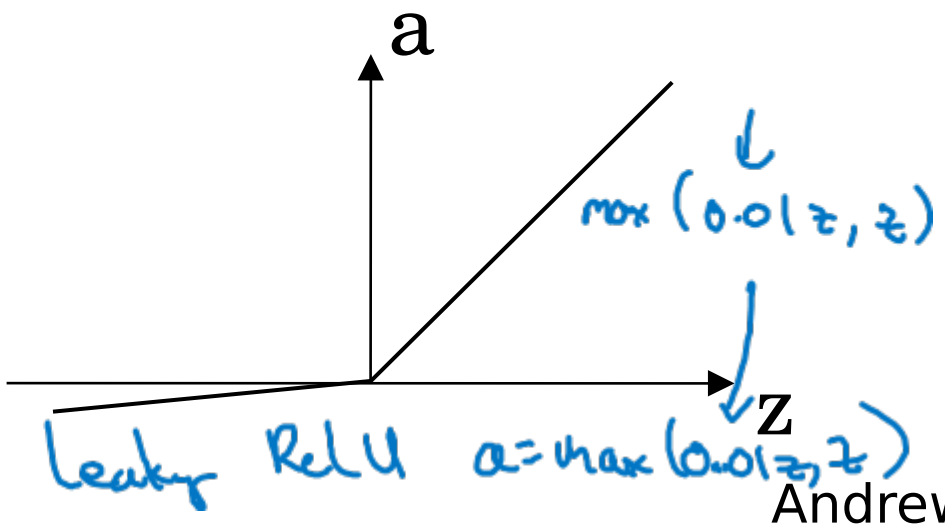
sigmoid: $a = \frac{1}{1+e^{-z}}$



tanh: $a = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$



ReLU $a = \max(0, z)$



Leaky ReLU $a = \max(0.01z, z)$