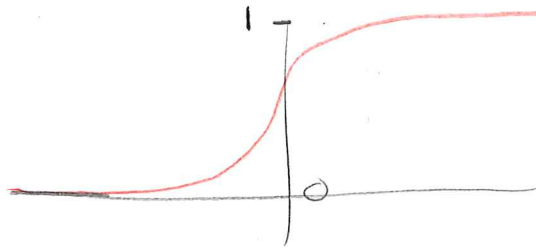


# Activation Functions

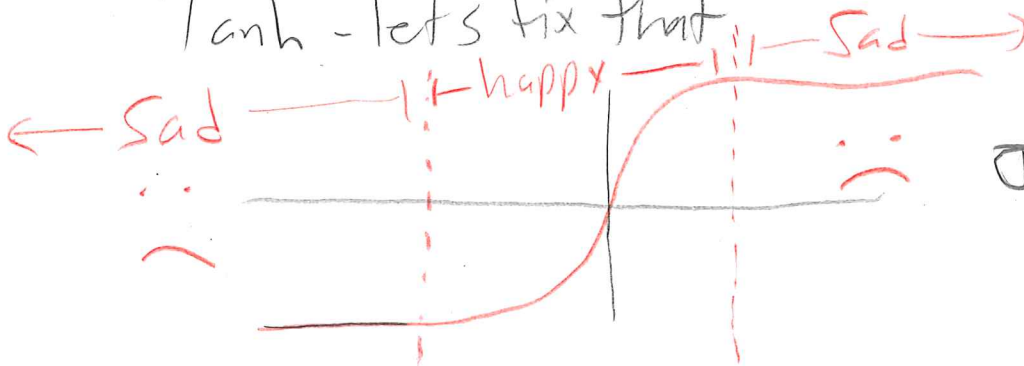
Sigmoid



$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- + seems like biological thing - neuron fires if inputs are strong enough
- not centered on zero (it's 0 to 1)  
so if our scores check the sign we have to learn a bias to get back to a y/N regime every time.

Tanh - let's fix that

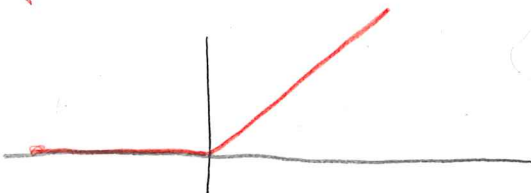


$$\sigma(x) = \tanh(x)$$

- Vanishing gradients

- $\frac{\partial \sigma}{\partial x} \approx 0$  unless you're in a very narrow range of  $x$  can't learn!

ReLU - let's fix that



$$\sigma(x) = \max(0, x)$$