**Activation Functions**

**Sigmoid**

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

- Seems like a 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**

- Vanishing gradients

\[ \sigma(x) = \tanh(x) \]

\[ \frac{d\sigma}{dx} = 0 \text{ unless you're in a very narrow range of } x \]

**ReLU - let's fix that**

\[ \sigma(x) = \max(0, x) \]