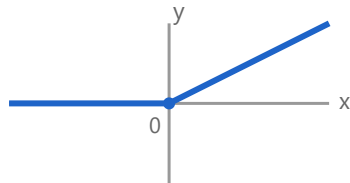


## The ReLU Revolution

### DEFINITION

$$\text{ReLU}(x) = \max(0, x)$$



### 🏆 HISTORICAL IMPACT

Enabled training of **AlexNet (2012)**

First deep CNN to win **ImageNet**

→ Sparked the deep learning revolution

✓ Prevents vanishing gradients (derivative = 1 for  $x > 0$ )

✓ Computational efficiency (no exponential calculations)

✓ Sparse activation (~50% neurons output zero)

✓ Unchanged gradient flow through active neurons

**LIMITATIONS:** Dead ReLU problem · Non-zero centered · Unbounded output