Random Initialization and Breaking Symmetry

✗ Zero Initialization

0 0 0

All neurons identical
Symmetry problem

↓

✓ Random Initialization

0.02 -0.01 0.03

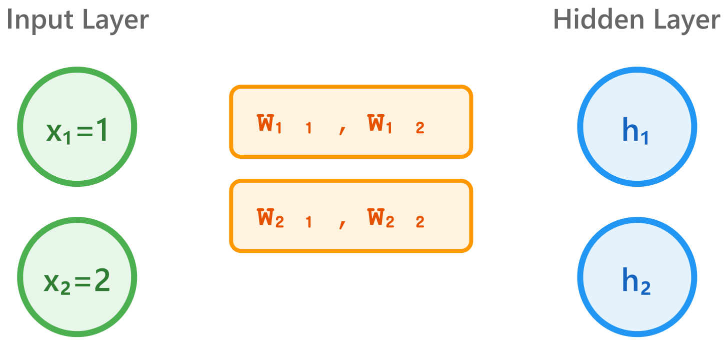
Each neuron unique
Symmetry broken!

Common Distributions
`N(0, 0.01) or U(-0.01, 0.01)`

Key Benefits

- **Breaks symmetry** between neurons with small random values
- Each neuron learns **different features** independently
- Enables **diverse representation** learning across layers
- **Scale matters:** Too large → saturation, Too small → vanishing gradients
- Foundation for **all modern methods**

Practical Example: Why Zero Initialization Fails



❌ Zero Initialization

$W_{11} = W_{12} = W_{21} = W_{22} = 0$

$h_1 = W_{11} \cdot x_1 + W_{12} \cdot x_2$
$= 0 \cdot 1 + 0 \cdot 2 = \mathbf{0}$

$h_2 = W_{21} \cdot x_1 + W_{22} \cdot x_2$
$= 0 \cdot 1 + 0 \cdot 2 = \mathbf{0}$

$h_1 = h_2 = 0$ (Identical!)

Gradient Update:

$\partial L/\partial W_{11} = \partial L/\partial W_{12}$ (same)
$\partial L/\partial W_{21} = \partial L/\partial W_{22}$ (same)

W_{11} and W_{12} stay identical
W_{21} and W_{22} stay identical

⚠ Symmetry Never Broken

✅ Random Initialization

$W_{11}=0.5$, $W_{12}=-0.3$

$W_{21}=0.2$, $W_{22}=0.4$

$h_1 = W_{11} \cdot x_1 + W_{12} \cdot x_2$
$= 0.5 \cdot 1 + (-0.3) \cdot 2$
$= 0.5 - 0.6 = \mathbf{-0.1}$

$h_2 = W_{21} \cdot x_1 + W_{22} \cdot x_2$
$= 0.2 \cdot 1 + 0.4 \cdot 2$
$= 0.2 + 0.8 = \mathbf{1.0}$

$h_1 = -0.1 \neq h_2 = 1.0$ (Different!)

Gradient Update:

$\partial L/\partial W_{11} \neq \partial L/\partial W_{12}$ (different)
$\partial L/\partial W_{21} \neq \partial L/\partial W_{22}$ (different)

Each weight learns
unique features independently

✓ Symmetry Broken Successfully