

Gradient Vanishing/Exploding Problems

↓ Vanishing Gradients

L4

L3

L2

L1

Gradients become **extremely small** in early layers, preventing effective learning in deep networks.

↑ Exploding Gradients

L1

L2

L3

L4

Gradients grow **exponentially large** during backpropagation, causing training instability.

Key Causes & Effects

Cause: Repeated multiplication of small/large values during backpropagation

Susceptible activations: Sigmoid/tanh particularly vulnerable

Impact: Prevents effective learning in deep networks

Consequence: Slow convergence or complete training failure

✓ Solution

Proper initialization helps maintain gradient magnitude throughout the network