Vanishing Gradient Problem in RNNs

Legend:
- Vanilla RNN ($\lambda=0.9$)
- Vanilla RNN ($\lambda=0.7$)
- LSTM ($\lambda=0.98$)

LSTM maintains gradient flow

Gradients vanish exponentially

Gradient ≈ 0

X-axis: Timesteps

Y-axis: Gradient Magnitude