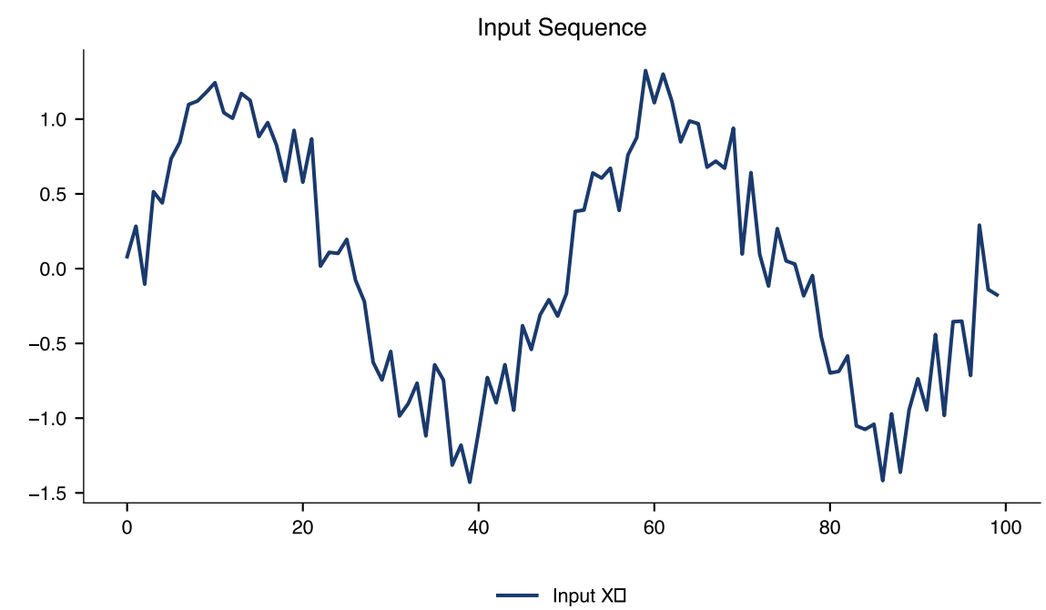

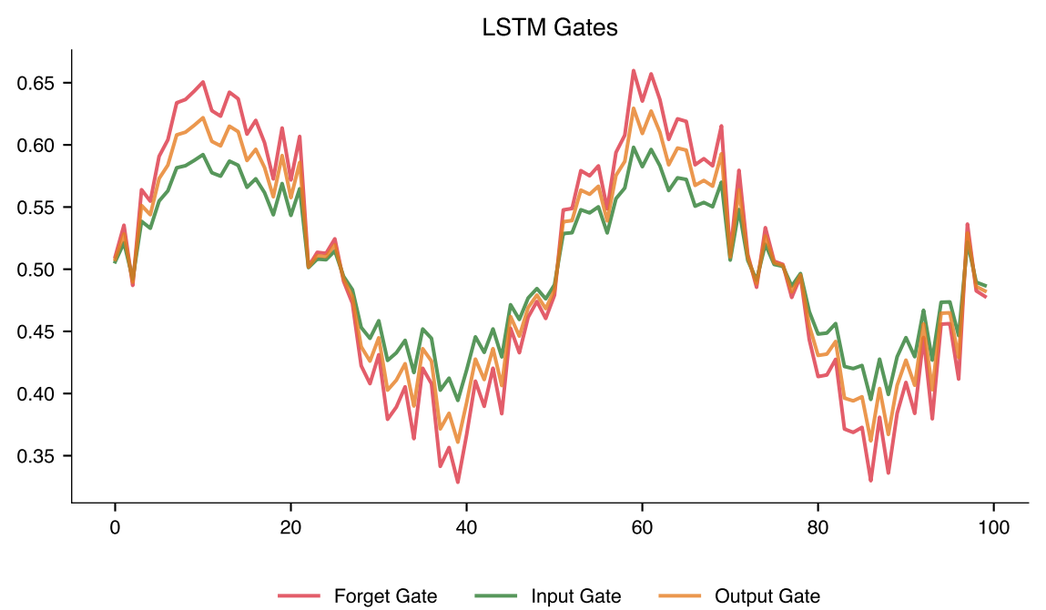

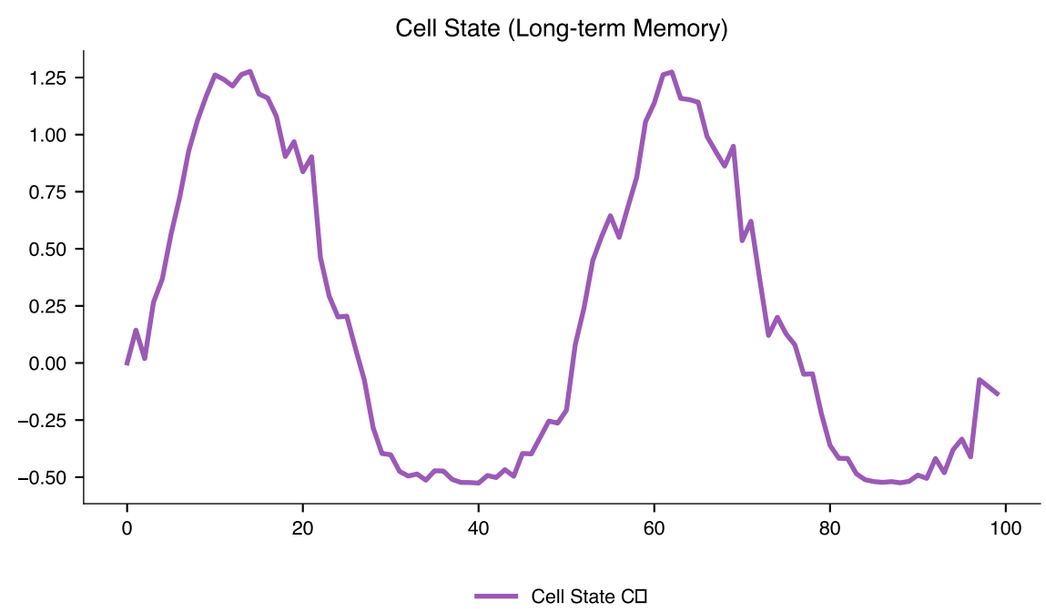

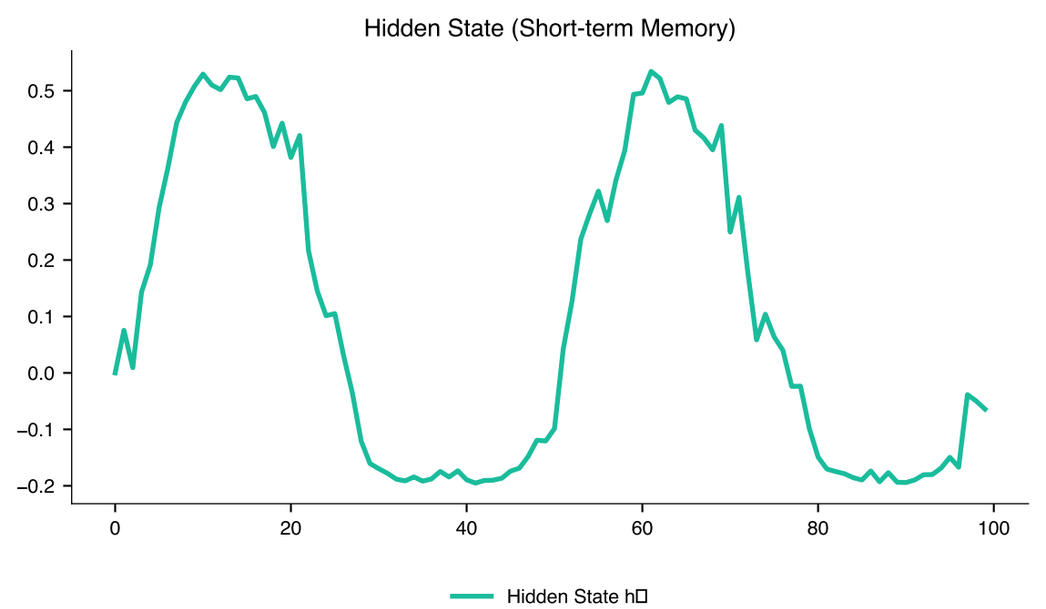

LSTM Cell Architecture

$$\text{Forget Gate: } f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$\text{Input Gate: } i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\text{Cell State: } C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_c \cdot [h_{t-1}, x_t])$$

$$\text{Output Gate: } o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$\text{Hidden State: } h_t = o_t \odot \tanh(C_t)$$

LSTM Advantages for Time Series

- □ Captures long-term dependencies
- □ Handles variable-length sequences
- □ Mitigates vanishing gradient problem
- □ Learns complex nonlinear patterns
- □ Automatic feature learning