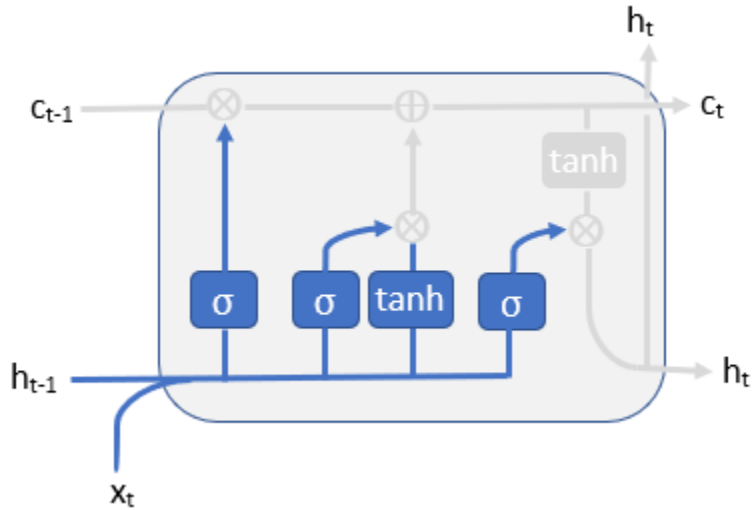# LSTM Forward Pass

LSTM Forward Pass

$$f_t = \sigma(W_f \cdot x_t + U_f \cdot h_{t-1} + b_i)$$

$$i_t = \sigma(W_i \cdot x_t + U_i \cdot h_{t-1} + b_i)$$
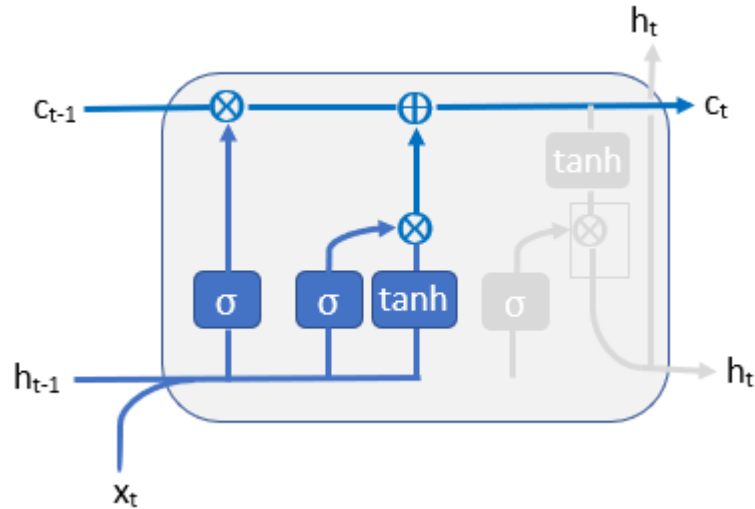
$$o_t = \sigma(W_o \cdot x_t + U_o \cdot h_{t-1} + b_o)$$

$$\widetilde{c}_t = tanh(W_c \cdot x_t + U_c \cdot h_{t-1} + b_c)$$

$$\boldsymbol{z_t} = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} W_c & U_c \\ W_f & U_f \\ W_i & U_i \\ W_o & U_o \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{x_t} \\ \boldsymbol{h_{t-1}} \end{bmatrix} + \begin{bmatrix} b_c \\ b_f \\ b_i \\ b_o \end{bmatrix} = W \cdot \boldsymbol{I_t} + \boldsymbol{b} \qquad h_t = tanh(c_t) \otimes o_t$$

LSTM Forward Pass

$$f_t = \sigma(W_f \cdot x_t + U_f \cdot h_{t-1} + b_i)$$

$$i_t = \sigma(W_i \cdot x_t + U_i \cdot h_{t-1} + b_i)$$

$$o_t = \sigma(W_o \cdot x_t + U_o \cdot h_{t-1} + b_o)$$

$$\tilde{c}_t = tanh(W_c \cdot x_t + U_c \cdot h_{t-1} + b_c)$$

$$c_t = c_{t-1} \otimes f_t \oplus i_t \otimes \tilde{c}_t$$

$$c_t = c_{t-1} \otimes Z_f \oplus Z_i \otimes Z_{\tilde{c}}$$

$$Z_t = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} W_c & U_c \\ W_f & U_f \\ W_i & U_i \\ W_o & U_o \end{bmatrix} \cdot \begin{bmatrix} x_t \\ h_{t-1} \end{bmatrix} + \begin{bmatrix} b_c \\ b_f \\ b_i \\ b_o \end{bmatrix} = W \cdot I_t + b$$

LSTM Forward Pass

$$f_t = \sigma(W_f \cdot x_t + U_f \cdot h_{t-1} + b_i)$$

$$i_t = \sigma(W_i \cdot x_t + U_i \cdot h_{t-1} + b_i)$$

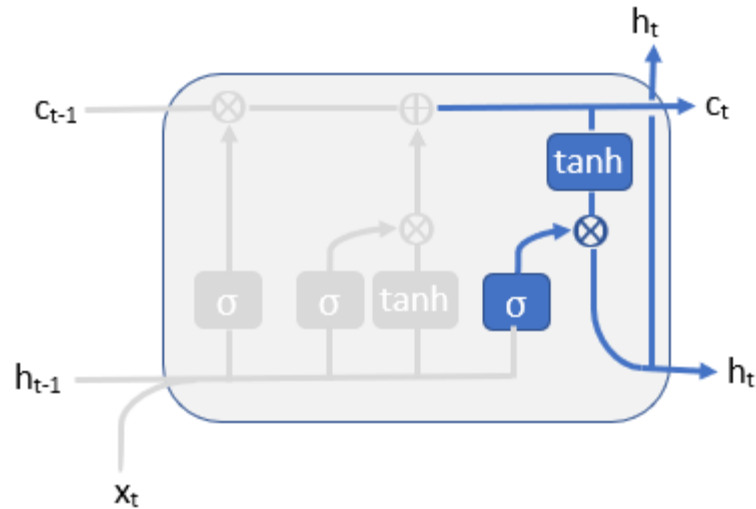$$o_t = \sigma(W_o \cdot x_t + U_o \cdot h_{t-1} + b_o)$$

$$\widetilde{c}_t = tanh(W_c \cdot x_t + U_c \cdot h_{t-1} + b_c)$$

$$Z_t = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} W_c & U_c \\ W_f & U_f \\ W_i & U_i \\ W_o & U_o \end{bmatrix} \cdot \begin{bmatrix} x_t \\ h_{t-1} \end{bmatrix} + \begin{bmatrix} b_c \\ b_f \\ b_i \\ b_o \end{bmatrix} = W \cdot I_t + b$$

$$h_t = tanh(c_t) \otimes o_t$$

$$h_t = tanh(c_t) \otimes Z_o$$

LSTM Forward Pass

$$x_t = [1, 2, 1] \qquad h_{t-1} = [0.3, 0.4] \qquad c_{t-1} = [0.1, 0.7]$$

LSTM Forward Pass

$$x_t = [1, 2, 1] \qquad h_{t-1} = [0.3, 0.4] \qquad c_{t-1} = [0.1, 0.7]$$

$$W_{\tilde{c}} = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_f = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_i = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_o = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix}$$

$$U_{\tilde{c}} = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_f = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_i = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_o = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix}$$

LSTM Forward Pass

$$x_t = [1, 2, 1] \qquad h_{t-1} = [0.3, 0.4] \qquad c_{t-1} = [0.1, 0.7]$$

$$W_{\tilde{c}} = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_f = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_i = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_o = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix}$$

$$U_{\tilde{c}} = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_f = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_i = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_o = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix}$$

$$Z_t = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} W_c & U_c \\ W_f & U_f \\ W_i & U_i \\ W_o & U_o \end{bmatrix} \cdot \begin{bmatrix} x_t \\ h_{t-1} \end{bmatrix} + \begin{bmatrix} b_c \\ b_f \\ b_i \\ b_o \end{bmatrix}$$

$$Z_t = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \end{bmatrix} \cdot \begin{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.99 \\ 0.97 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \end{bmatrix}$$

LSTM Forward Pass

$$x_t = [1, 2, 1] \qquad h_{t-1} = [0.3, 0.4] \qquad c_{t-1} = [0.1, 0.7]$$

$$W_{\tilde{c}} = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_f = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_i = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_o = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix}$$

$$U_{\tilde{c}} = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_f = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_i = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_o = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix}$$

$$Z_t = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} & \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \end{bmatrix} \cdot \begin{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.99 \\ 0.97 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \end{bmatrix}$$

$$c_t = c_{t-1} \otimes Z_f \oplus Z_i \otimes Z_{\tilde{c}}$$

$$c_t = [0.1, 0.7] \otimes [0.93, 0.89] \oplus [0.93, 0.89] \otimes [0.99, 0.97] = [0.09, 0.58] \oplus [0.92, 0.86] = [1.01, 1.44]$$

LSTM Forward Pass

$$x_t = [1, 2, 1] \qquad h_{t-1} = [0.3, 0.4] \qquad c_{t-1} = [0.1, 0.7]$$

$$W_{\tilde{c}} = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_f = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_i = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix} \qquad W_o = \begin{bmatrix} 0.2 & 0.1 \\ 0.5 & 0.3 \\ 0.3 & 0.1 \end{bmatrix}$$

$$U_{\tilde{c}} = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_f = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_i = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix} \qquad U_o = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.3 \end{bmatrix}$$

$$z_t = \begin{bmatrix} \tilde{c}_t \\ f_t \\ i_t \\ o_t \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \\ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 \\ 0.5 & 0.3 \end{bmatrix} \end{bmatrix} \cdot \begin{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.99 \\ 0.97 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \\ \begin{bmatrix} 0.93 \\ 0.89 \end{bmatrix} \end{bmatrix}$$

$$c_t = [1.01, 1.44]$$

$$h_t = tanh(c_t) \otimes Z_o$$

$$h_t = tanh([1.01, 1.44]) \otimes [0.93, 0.89] = [0.77, 0.89] \otimes [0.93, 0.89] = [0.72, 64]$$