

LSTM (Long Short Term Memory)

Cell State

$$(C_t \in R^h, \tilde{C}_t \in (-1,1)^h)$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t$$

Forget Gate ($f_t \in (0,1)^h$)

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

Input Gate ($i_t \in (0,1)^h$)

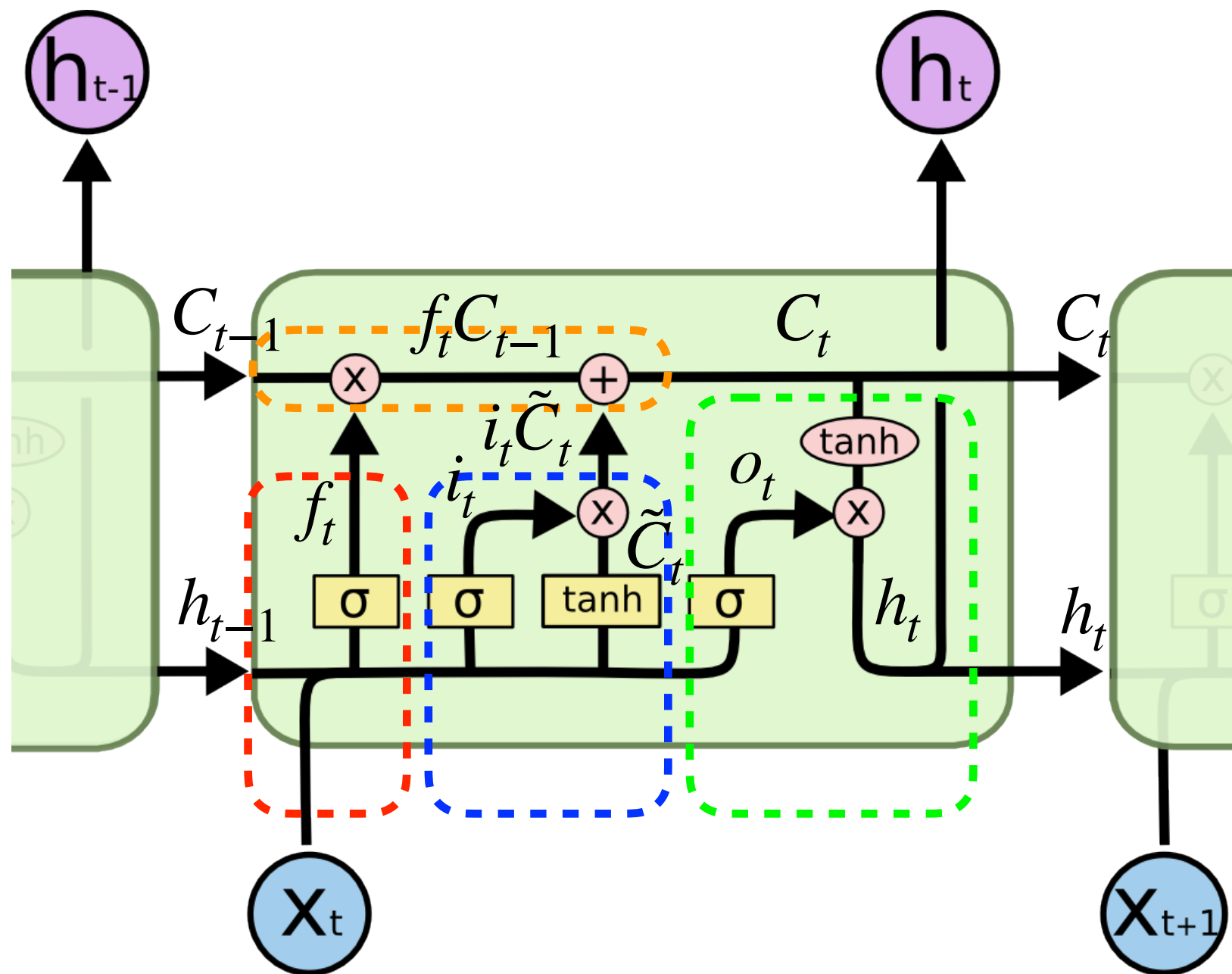
$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i)$$

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

Output Gate ($o_t \in (0,1)^h$)

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

$$h_t = o_t \tanh(C_t)$$



Input vector: $x_t \in R^d$

Hidden layer: $h_t \in (-1,1)^h$

h : # of hidden units d : # of input features

$\sigma(\cdot)$: Sigmoid activation

$\tanh$ can be other activation

LSTM (Long Short Term Memory)

