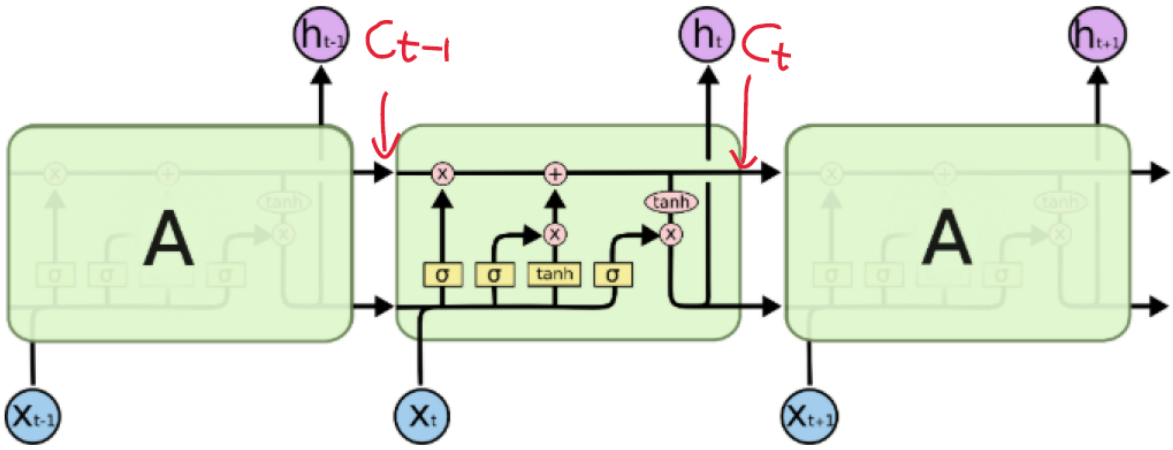
PART 1 LSTM and GRU

note: * denotes the element-wise product of two matrices

LSTM



$$
\begin{aligned}
f_t &= \sigma(W_0 x_t + U_0 h_{t-1} + b_0) \\
i_t &= \sigma(W_1 x_t + U_1 h_{t-1} + b_1) \\
o_t &= \sigma(W_2 x_t + U_2 h_{t-1} + b_2) \\
c_t &= f_t * c_{t-1} + i_t * tanh(W_3 x_t + U_3 h_{t-1} + b_3) \\
h_t &= tanh(c_t) * o_t
\end{aligned}
$$

GRU

$$
\begin{aligned}
z_t &= \sigma(W_0 x_t + U_0 h_{t-1} + b_0) \\
r_t &= \sigma(W_1 x_t + U_1 h_{t-1} + b_1) \\
h_t^{'} &= tanh(W_2 x_t + U_2 (r_t * h_t - 1) + b_2) \\
h_t &= (1 - z_t) * h_{t-1} + z_t * h_t^{'}
\end{aligned}
$$