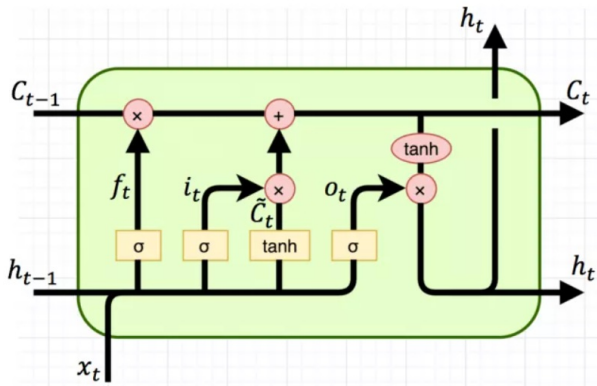

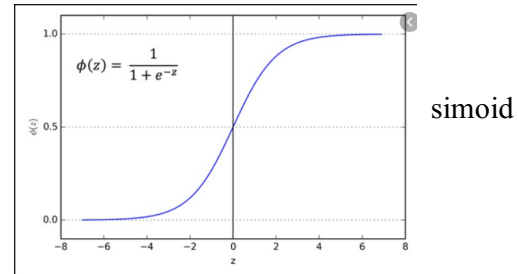


LSTM (Long Short-Term Memory)



 =
sigmoid



C = Memory cell (เก็บความจำระยะยาว)

C_{t-1} = ความจำระยะยาว Time Step(t-1)

C_t = ความจำระยะยาวที่ Time Step(t)

h = ความจำระยะสั้น

h_t = ความจำระยะสั้นที่ Time Step(t)

h_{t-1} = ความจำระยะสั้นที่ Time Step(t-1)

X_t = Input ที่ Time Step (t)

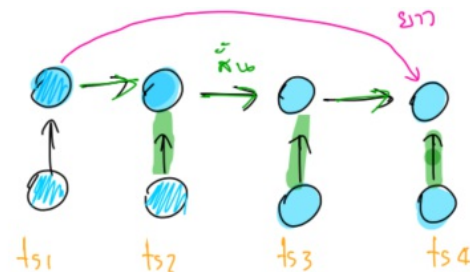
W = Weight ; U = Weight ก่อนหน้า

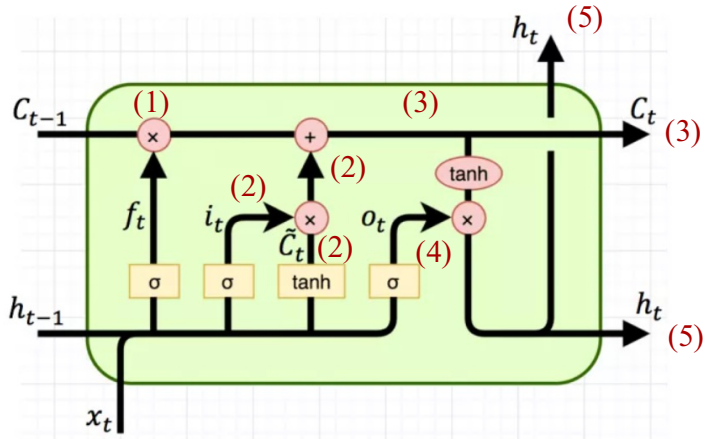
1) ความจำระยะยาว หายไปได้ (ไม่ได้ใช้นาน)

2) ความจำระยะสั้นที่สำคัญ -->
กลายเป็นความจำระยะยาวได้

3) ลืมความจำระยะยาว (ใหม่) = เกิดจาก (1) + (2)

4) ความจำระยะสั้น (ใหม่) => มีผลต่อความจำระยะยาว





$$1) f_t = \text{sigmoid} (X_t * W_f + h_{t-1} * U_f + B_f) \quad f_t = \sigma(x_t U^f + h_{t-1} W^f)$$

--> $C_{t-1} \otimes f_t$; Ex 100% \otimes 0.8 = 80% ข้อมูลที่เหลือส่งไป
(ลืมความจำระยะยาว)

$$2) C_t = \tanh (X_t * W_{ct} + h_{t-1} * U_{ct} + B_{ct}) \quad \tilde{C}_t = \tanh (x_t U^g + h_{t-1} W^g)$$

$$i_t = \text{sigmoid} (X_t * W_i + h_{t-1} * U_i + B_i) \quad i_t = \sigma(x_t U^i + h_{t-1} W^i)$$

EX -> $C_t = 100\%$ \otimes ได้ 20% ที่เหลือ
(ลืมความจำระยะสั้น)

$$3) C_3 = C_{t-1} + (C_t \otimes i_t) \quad C_t = \sigma(f_t * C_{t-1} + i_t * \tilde{C}_t)$$

$$4) O_t = \text{sigmoid}(X_t * W_o + h_{t-1} * U_o + B_o) \quad o_t = \sigma(x_t U^o + h_{t-1} W^o)$$

$$5) h_t = \tanh(C_3 * O_t) \quad h_t = \tanh(C_t) * o_t$$

	TS_1		TS_2		TS_3	
	H	L	H	L	H	L
1	1.135	1.132	1.138	1.133	1.14	1.135
2	1.137	1.132	1.136	1.132	1.137	1.133
3	1.137	1.132	1.137	1.132	1.136	1.132
4	1.136	1.128	1.137	1.132	1.137	1.132

$$W_i = \begin{bmatrix} 0.28 & -1.39 & 1.42 \\ 0.49 & 0.13 & 0.67 \end{bmatrix} \qquad U_i = \begin{bmatrix} 1.13 & 0.51 & 0.82 \\ 0.47 & -0.31 & -1.4 \\ 0.23 & 77 & 1.3 \end{bmatrix}$$

$$B_f = [0.14 \quad -0.67 \quad 0.31]$$

$$W_{ct} = \begin{bmatrix} 0.56 & 0.12 & -0.32 \\ 0.79 & -0.13 & 0.83 \end{bmatrix} \qquad U_{ct} = \begin{bmatrix} -1.31 & 0.18 & 0.65 \\ 0.14 & 0.81 & 0.21 \\ 0.75 & 0.43 & 0.86 \end{bmatrix}$$

$$B_{ct} = [0.35 \quad -0.43 \quad 0.75]$$

$$W_o = \begin{bmatrix} 0.14 & 0.28 & 0.51 \\ 0.65 & -0.13 & -1.43 \end{bmatrix} \qquad U_o = \begin{bmatrix} -0.44 & 0.34 & 0.84 \\ 0.21 & 1.34 & 0.23 \\ -0.34 & 0.65 & 0.84 \end{bmatrix}$$

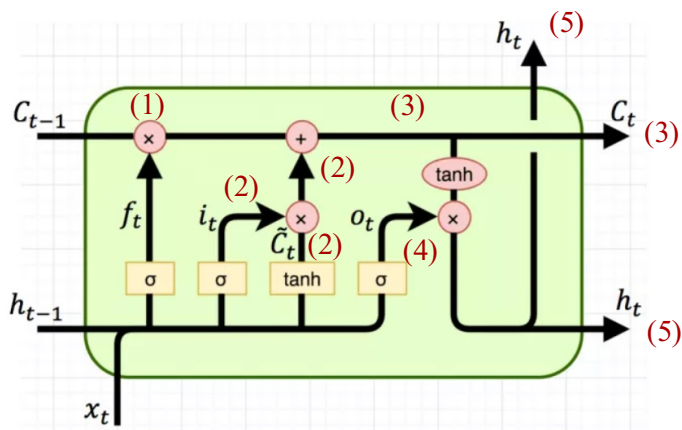
$$B_o = [0.21 \quad 0.81 \quad 0.33]$$

	TS_1	
	H	L
1	1.135	1.132
2	1.137	1.132
3	1.137	1.132
4	1.136	1.128

$$X_{TS_1} = \begin{bmatrix} 1.135 & 1.132 \\ 1.137 & 1.132 \\ 1.137 & 1.132 \\ 1.136 & 1.128 \end{bmatrix}$$

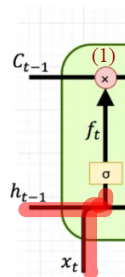
$$H1_{TS_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad C1_{TS_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad U_f = \begin{bmatrix} 0.34 & 0.44 & 0.42 \\ 0.69 & 0.18 & 0.59 \\ 0.21 & -1.39 & 1.34 \end{bmatrix}$$

$$W_f = \begin{bmatrix} 0.23 & -0.71 & 0.42 \\ 1.74 & 0.34 & 0.67 \end{bmatrix} \qquad B_f = [0.14 \quad -0.67 \quad 0.31]$$



$$1) f_t = \text{sigmoid} (X_t * W_f + h_{t-1} * U_f + B_f) \quad f_t = \sigma(x_t U^f + h_{t-1} W^f)$$

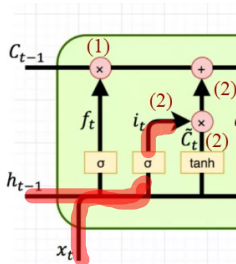
--> $C_{t-1} \otimes f_t$; Ex 100% $\otimes 0.8 = 80\%$ ข้อมูลที่เหลือส่งไป
(ลืมความจำระยะยาว)



$$F1_Ts_1 = \sigma (X_TS_1 * W_f + H1_TS_0 * U_f + B_f)$$

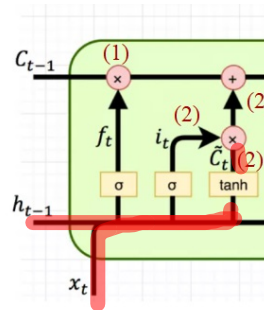
$$i_t = \sigma(x_t U^i + h_{t-1} W^i)$$

$$2) i_t = \text{sigmoid} (X_t * W_i + h_{t-1} * U_i + B_i)$$



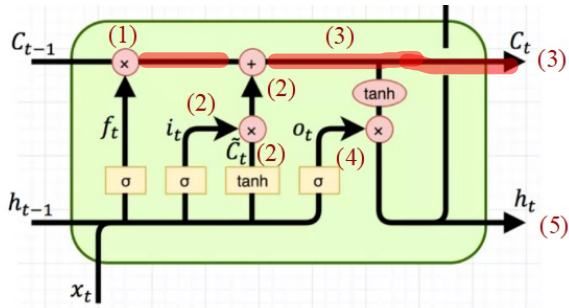
$$I1_Ts_1 = \sigma (X_TS_1 * W_i + H1_TS_0 * U_i + B_i)$$

$$2.2) C_t = \tanh (X_t * W_c + h_{t-1} * U_c + B_c) \quad \tilde{C}_t = \tanh (x_t U^g + h_{t-1} W^g)$$



$$C1_Ts_1 = \tanh (X_TS_1 * W_c + H1_TS_0 * U_c + B_c)$$

$$3) C_3 = C_{t-1} + (C_t \otimes i_t) \quad C_t = \sigma(f_t * C_{t-1} + i_t * \tilde{C}_t)$$

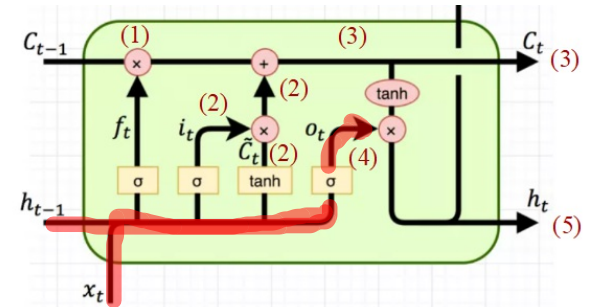


$$C1_TS_1 = (F1_TS_1) + (T1_TS_1 * CT1_TS_1)$$

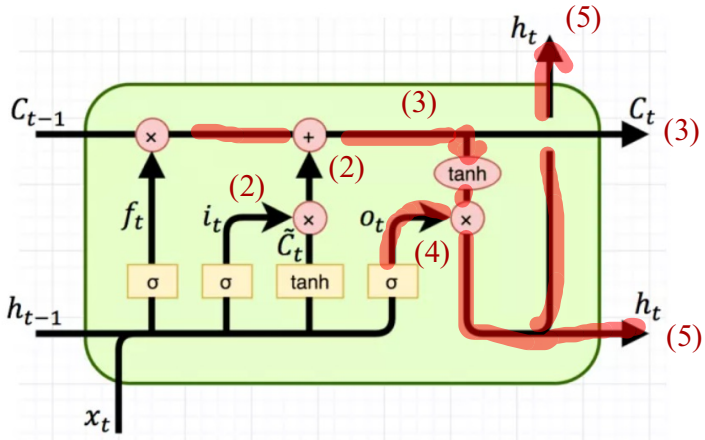
< C = ความจำระยะยาว >

$$4) O_t = \text{sigmoid}(X_t * W_o + h_{t-1} * U_o + B_o) \quad o_t = \sigma(x_t U^o + h_{t-1} W^o)$$

$$O1_TS_1 = \sigma (X_TS_1 * W_o + H1_TS_0 * U_o + B_o)$$



$$5) h_t = \tanh(C_t * O_t) \quad h_t = \tanh(C_t) * o_t$$



$$H1_Ts_1 = \tanh (C1_TS_1) * O1_TS_1$$

< h = ความจำระยะสั้น >

$$X_{TS_2} = \begin{bmatrix} 1.138 & 1.133 \\ 1.136 & 1.132 \\ 1.137 & 1.132 \\ 1.137 & 1.132 \end{bmatrix}$$

$$F1_{Ts_2} = \phi (X_{TS_2} * W_f + H1_{Ts_1} * U_f + B_f)$$

$$I1_{Ts_2} = \phi (X_{TS_2} * W_i + H1_{Ts_1} * U_i + B_i)$$

$$C1_{TS_2} = \tanh (X_{TS_2} * W_{ct} + H1_{TS_1} * U_{ct} + B_{ct})$$

$$C1_{TS_2} = (F1_{TS_1} * CT1_{TS_1}) + (T1_{TS_1} * CT1_{TS_2})$$

$$O1_{TS_2} = \phi (X_{TS_2} * W_o + H1_{TS_1} * U_o + B_o)$$

< C = ความจำระยะยาว >

$$H1_{Ts_2} = \tanh (C1_{TS_2}) * O1_{TS_2}$$

< h = ความจำระยะสั้น >