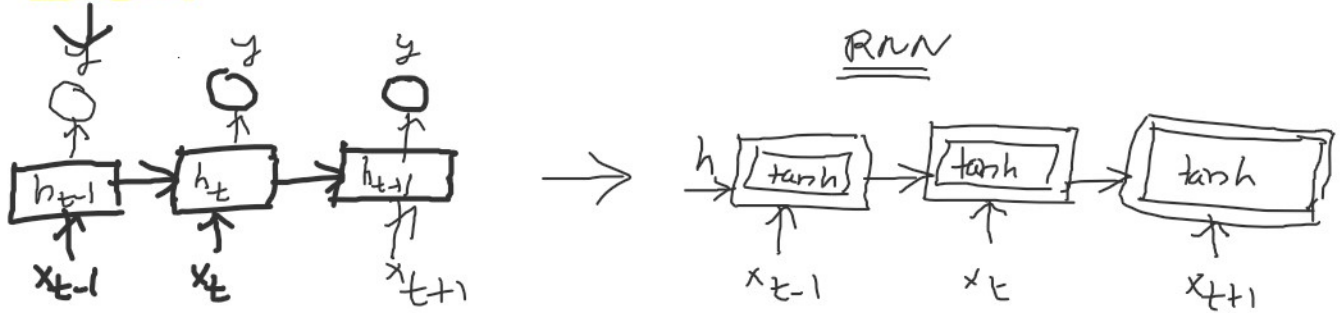
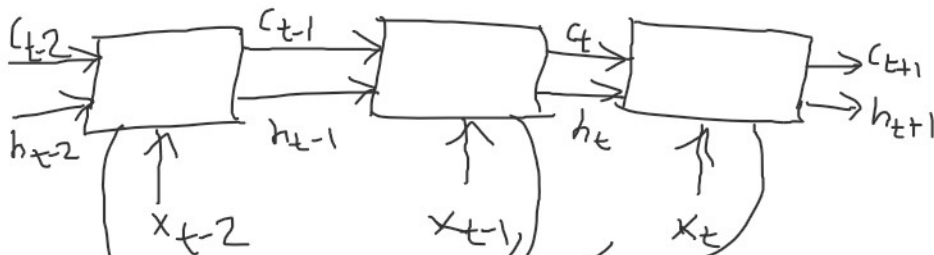


Long Short term Memory

RNN

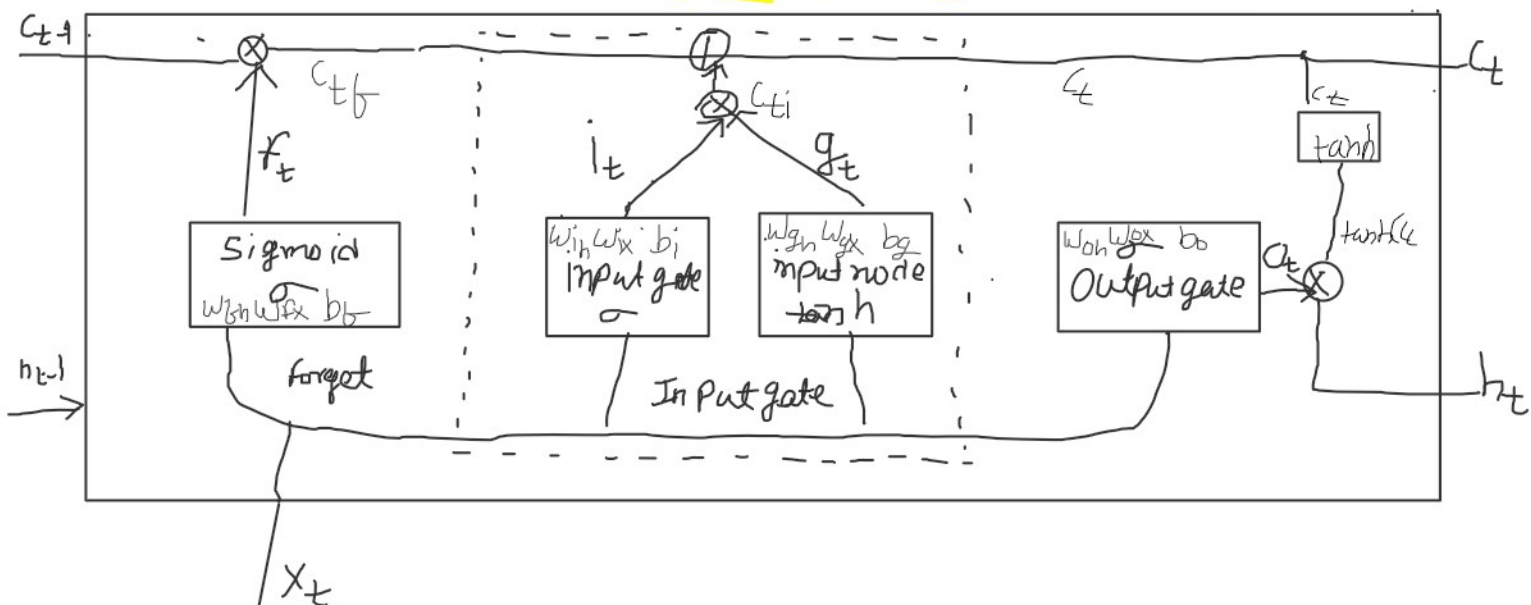


LSTM



- LSTM cell contains
- a simple RNN cell
 - cell state (Long Term Memory)
 - input gate
 - forget gate
 - output gate

One LSTM cell



x_t

Forget State

$$f_t = \sigma \left[(w_{fh} \times h_{t-1}) + (w_{fx} \times x_t) + b_f \right]$$

$$C_{tf} = C_{t-1} \times f_t, \text{ Same dimension for dot product}$$

The main purpose of this forget state is to forget those information which are out of context.

$$\text{e.g.} \quad [1, 4, 2] \cdot [1, 0, 1] \\ = [1, 0, 2]$$

We can see that the information which is on the 1st index is forgotten.

Input State

$$i_t = \sigma \left[(w_{ih} \times h_{t-1}) + (w_{ix} \times x_t) + b_i \right]$$

$$g_t = \tanh \left[(w_{gh} \times h_{t-1}) + (w_{gx} \times x_t) + b_g \right]$$

$$C_{ti} = i_t \times g_t$$

$$C_t = C_{ti} + C_{tf}$$

Output State

$$o_t = \sigma \left[(w_{oh} \times h_{t-1}) + (w_{ox} \times x_{t-1}) + b_o \right]$$

$$h_t = \tanh(c_t) \times O_t$$

Four weight = $\begin{bmatrix} w_f \\ w_i \\ w_g \\ w_o \end{bmatrix}$ } Four bias = $\begin{bmatrix} b_f \\ b_i \\ b_g \\ b_o \end{bmatrix}$

Link = colab.research.google.com