


LSTM cell \rightarrow why? because RNN often forget something

$h(t)$: short-term

$c(t)$: long-term

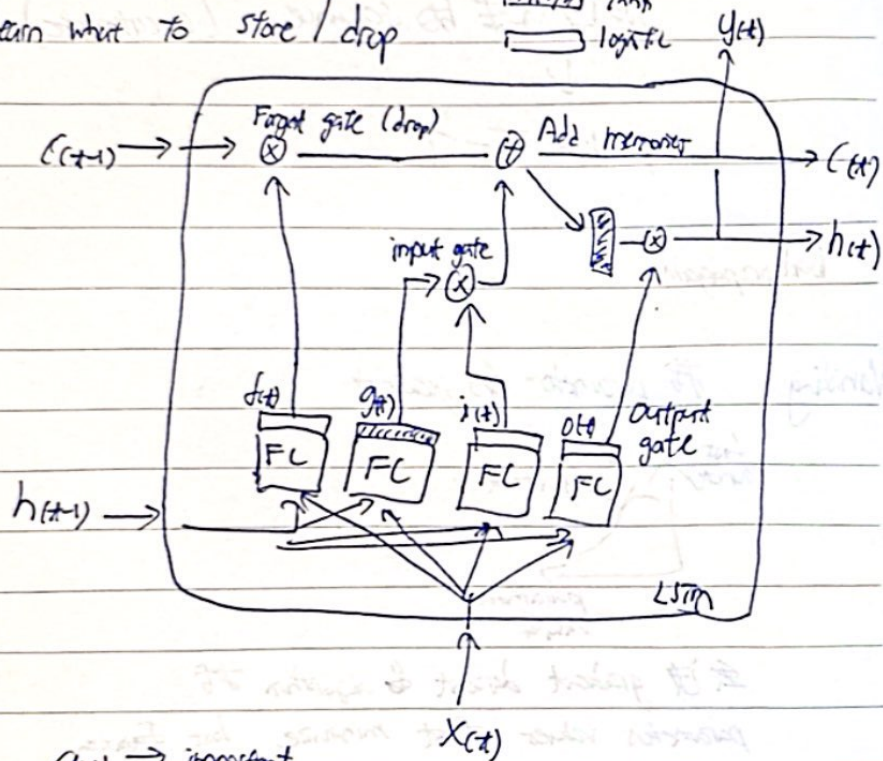
\otimes multiplication (element-wise)

\oplus addition

 tanh

 logistic

Learn what to store / drop



g_{in} \rightarrow important

analyze (short-term and input)

most important are stored in long-term

gate controller: output (0-1), multiplication

0 \rightarrow close gate

1 \rightarrow open gate

$$c(t) = c(t-1) \otimes f(t) + i(t) \otimes g(t)$$

$$y(t) = h(t) = O(t) \otimes \tanh(c(t))$$

control activation threshold \uparrow

$$g(t) = \tanh(W_{xg}^T x(t) + W_{hg}^T h(t-1) + b_g)$$

$$f(t) = \sigma(W_{xf}^T x(t) + W_{hf}^T h(t-1) + b_f)$$

$$O(t) = \sigma(W_{xo}^T x(t) + W_{ho}^T h(t-1) + b_o)$$

$$i(t) = \sigma(W_{xi}^T x(t) + W_{hi}^T h(t-1) + b_i)$$