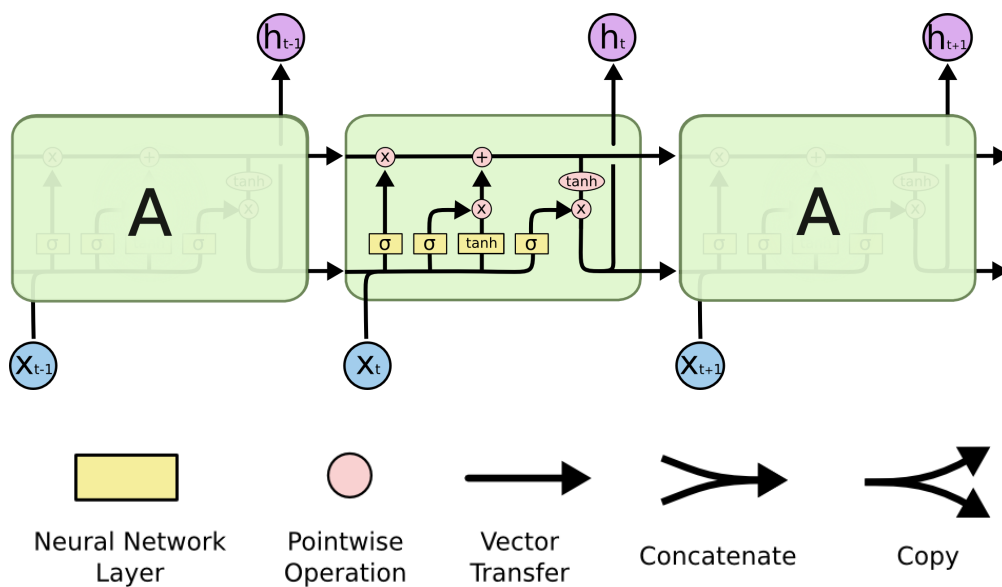


LSTM算法简单梳理

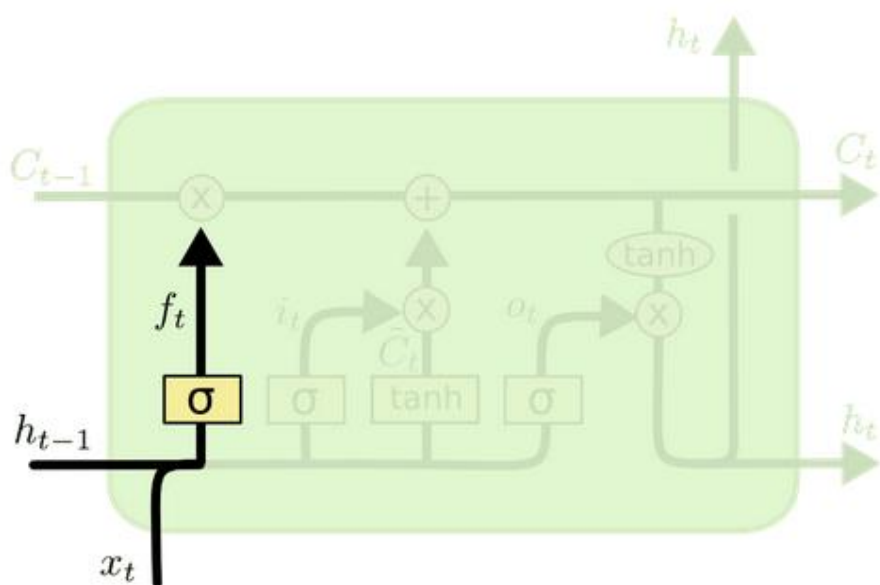
杨航锋

1 LSTM 框架结构



h_t : 当前序列的隐藏状态、 x_t : 当前序列的输入数据、 C_t : 当前序列的细胞状态、 σ : *sigmoid* 激活函数、 \tanh : *tanh* 激活函数。

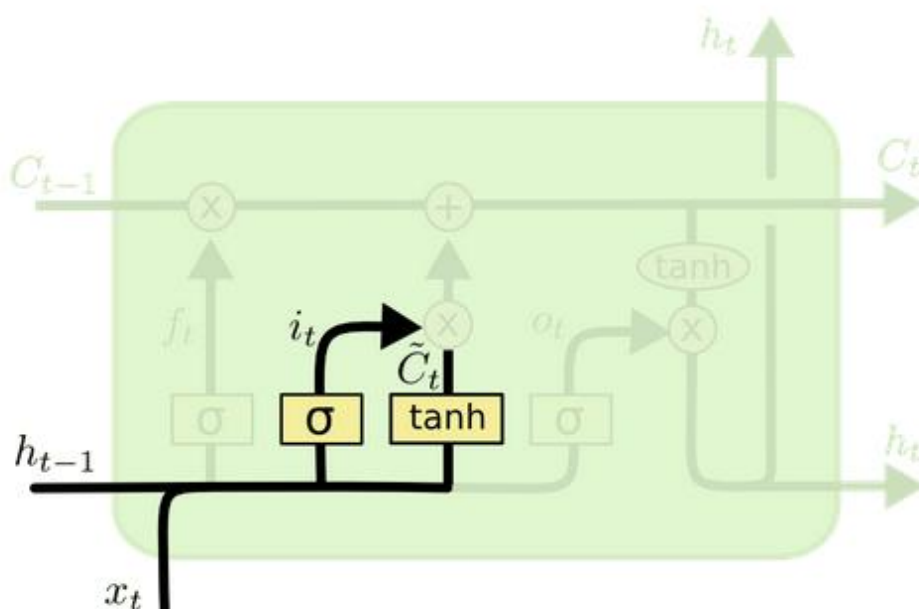
2 LSTM 之遗忘门



遗忘门是控制是否遗忘的，在 *LSTM* 中即以一定的概率控制是否遗忘上一层的细胞状态。图中输入的有前一序列的隐藏状态 h_{t-1} 和当前序列的输入数据 x_t ，通过一个 *sigmoid* 激活函数得到遗忘门的输出 f_t 。因为 *sigmoid* 函数的取值在 $[0, 1]$ 之间，所以 f_t 表示的是遗忘前一序列细胞状态的概率，数学表达式为

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

3 *LSTM* 之输入门

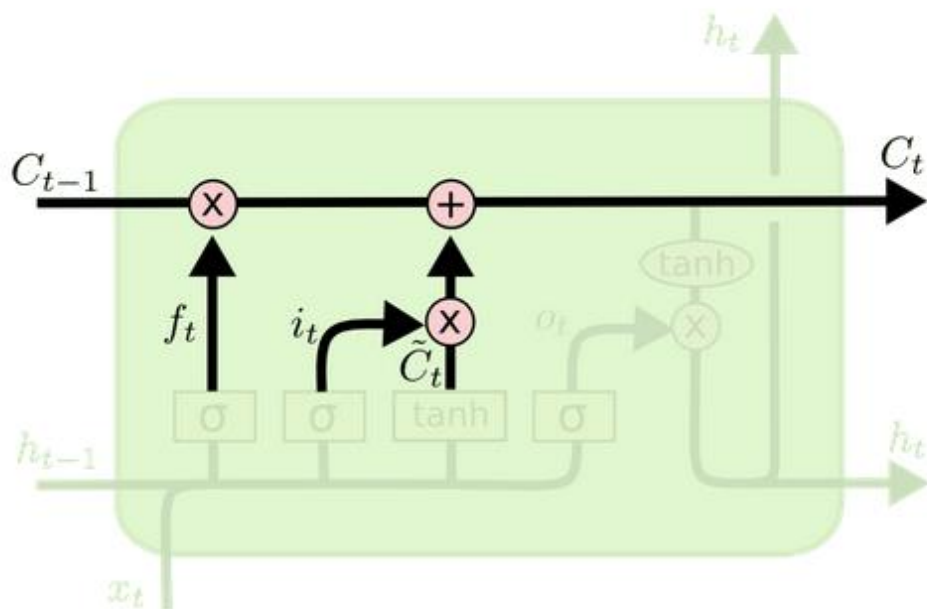


输入门是用来决定哪些数据是需要更新的，由 *sigmoid* 层决定；然后，一个 *tanh* 层为新的候选值创建一个向量 \tilde{C}_t ，这些值能够加入到当前细胞状态中，数学表达式为

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

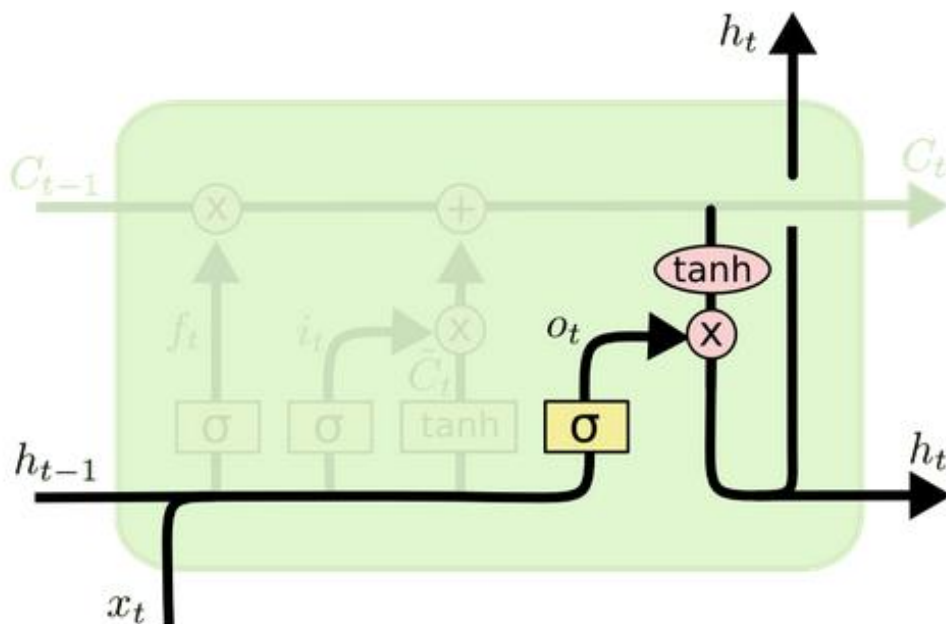
4 *LSTM* 之细胞状态更新



前面的遗忘门和输入门的结果都会作用于细胞状态 C_t ，在决定需要遗忘和需要加入的记忆之后，就可以更新前一序列的细胞状态 C_{t-1} 到当前细胞状态 C_t 了，前一序列的细胞状态 C_{t-1} 乘以遗忘门的输出 f_t 表示决定遗忘的信息， $i_t \odot \tilde{C}_t$ 表示新的记忆信息，数学表达式为

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

5 LSTM 之输出门



在得到当前序列的细胞状态 C_t 后，就可以计算当前序列的输出隐藏状态 h_t 了，隐藏状态 h_t 的更新由两部分组成，第一部分是 o_t ，它由前一序列的隐藏状态 h_{t-1} 和当前序列的输入数据 x_t 通过激活函数 *sigmoid* 得到，第二部分由当前序列的细胞状态 C_t 经过 *tanh* 激活函数后的结果组成，数学表达式为

$$\begin{aligned}o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\h_t &= o_t \odot \tanh(C_t)\end{aligned}$$

参考资料

[Understanding LSTM Networks](#)