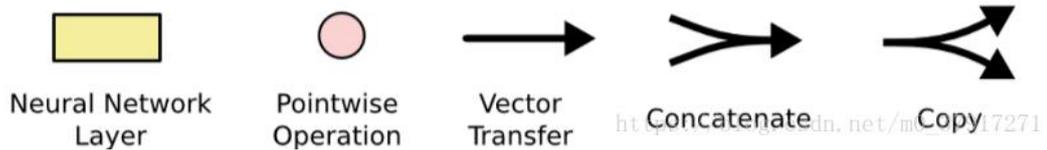
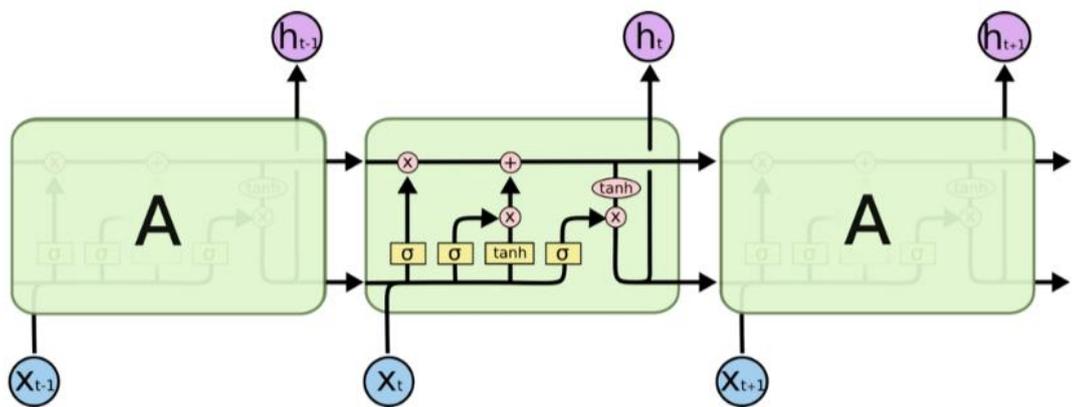


LSTM 算法(Long Short Term Memory, 长短期记忆网络)



$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

一、长期记忆、短期记忆和三个门

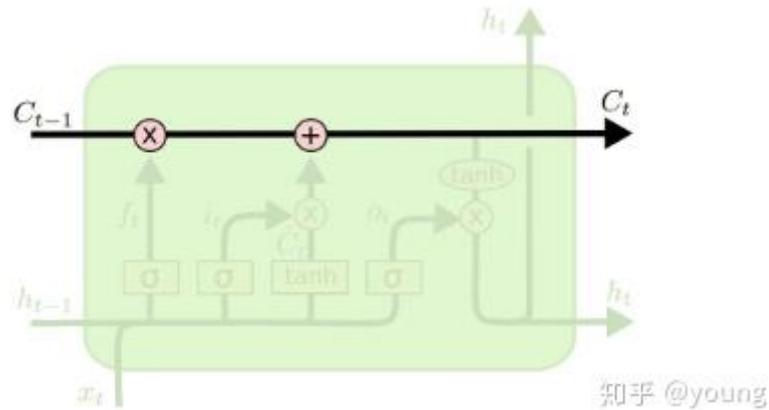


图 1

长期记忆 C_t ，也叫细胞状态，用于完成序列学习的工作，如图 1 中最上方的一条长线，并在线上连接了三道门。

短期记忆 h_t ，事件信息 x_t ，由门来控制添加到长期记忆中的部分，从而达到选择性通过的结果。

LSTM 算法是在 RNN 算法的基础上，在每个单元模块中添加了三个重复的神经网络层，分别为：遗忘门、输入门和输出门。

二、遗忘门

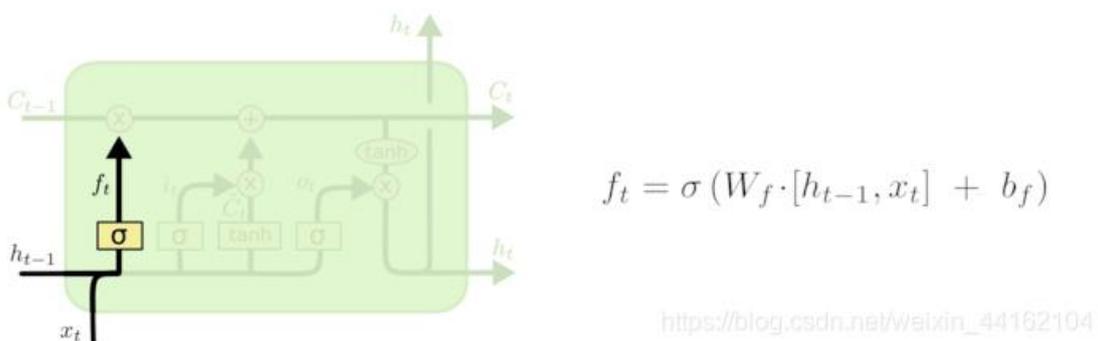
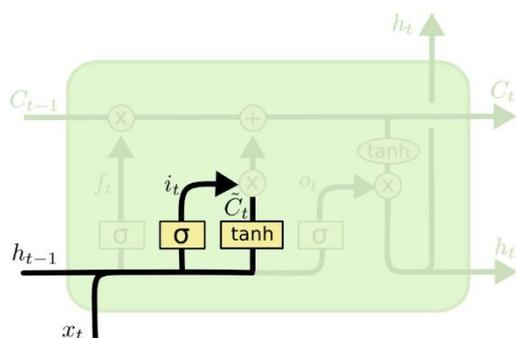


图 2

LSTM 中的第一道门，决定了在细胞状态中我们需要舍弃多少信息。

如图 2，读入 h_{t-1} 和 x_t 后经过 σ 函数整合为一个 0~1 之间的数，来告诉 C_{t-1} 中的数字需要保留多少，“0”是全部舍弃，“1”是全部保留。

三、输入门



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

https://blog.csdn.net/weixin_44162104

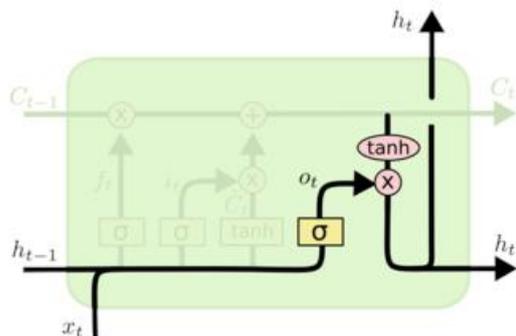
图 3

输入门用于计算新的细胞状态 C_t 。

输入门（如图 3）由两步构成，第一步是先计算一个叫做“Input Gate Layer”的 sigmoid 层 i_t 和一个 tanh 层 \tilde{C}_t 来决定哪些信息需要更新。

然后将 $i_t \times \tilde{C}_t$ 加到 C_{t-1} 中，更新为 C_t 。

四、输出门



$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t * \tanh(C_t)$$

https://blog.csdn.net/weixin_44162104

图 4

输出门用于计算该细胞单元的短期记忆值 h_t 。

h_t 由一个 sigmoid 层和一个 tanh 层相乘而来，其中 tanh 层完全由当前细胞状态 C_t 处理得到。