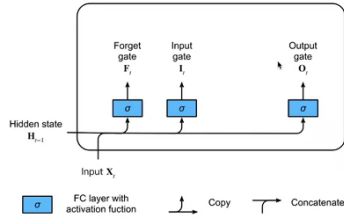


LSTM

- 长短期记忆网络：用于解决隐变量模型存在的长期信息保存和短期输入缺失的问题
- 门控记忆元（单元）：使用三种门控制
 - 忘记门：用于重置单元内容，将值朝0减少
 - 输入门（记忆门）：决定何时将数据读入单元，是否忽略掉输入数据
 - 输出门：从单元中输出条目，决定是否使用隐状态



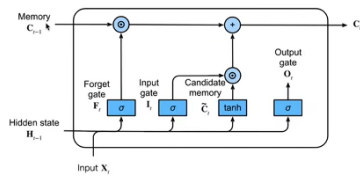
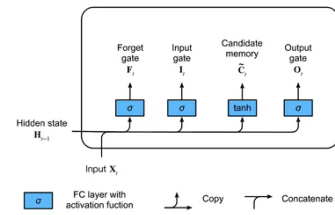
$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f)$$

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i)$$

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o)$$

候选记忆单元

$$\tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c)$$



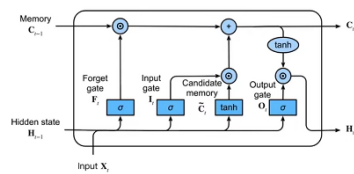
记忆单元

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t$$

F 用于控制多大程度接收过去信息(忘掉前面)

I 用于控制多大程度接收当前输入信息(忘掉当前)

$$H_t = O_t \odot \tanh(C_t)$$



隐状态

$$H_t = O_t \odot \tanh(C_t)$$

O_t 用于控制当前状态是否重置

(即若 $O_t \rightarrow 0$, 既不用当前输入, 又不用过去记忆)