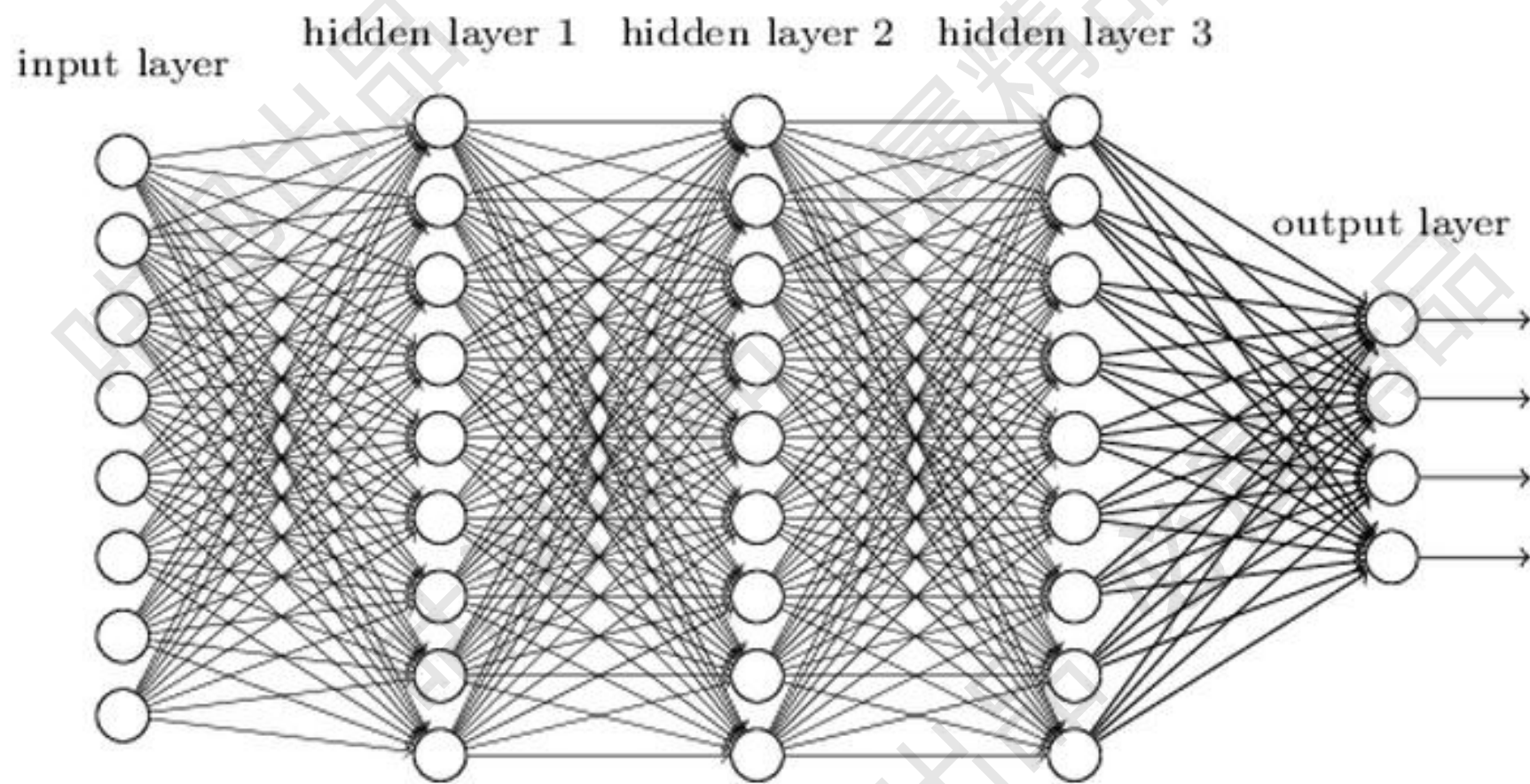
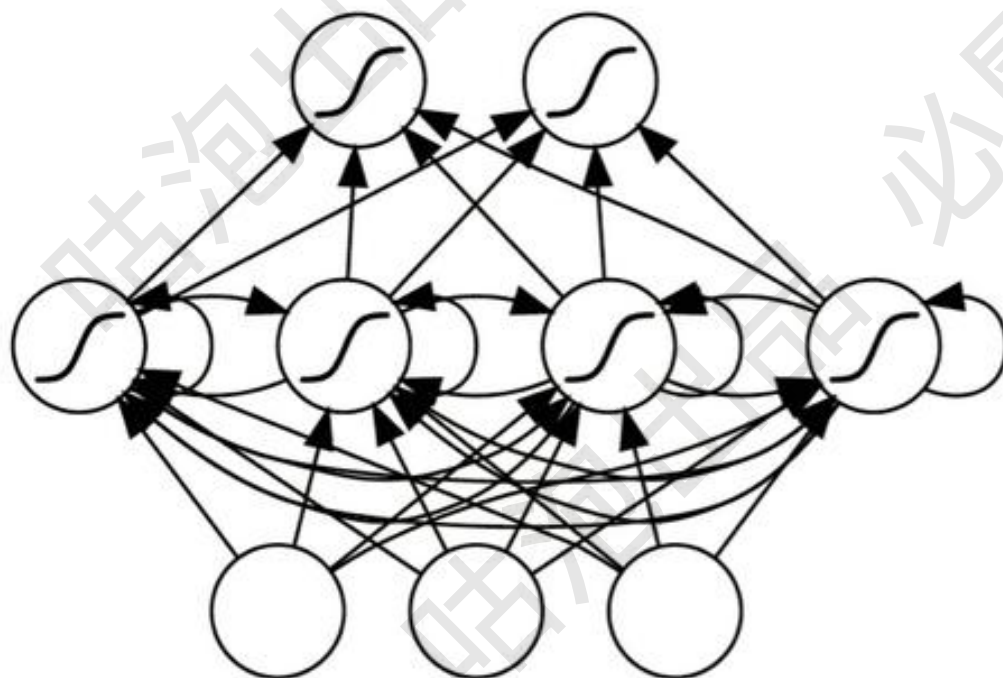


RNN



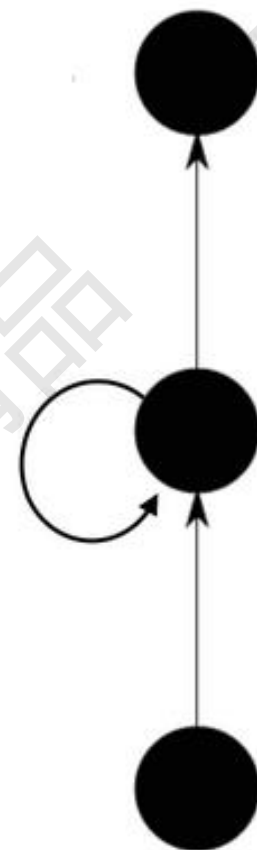
RNN



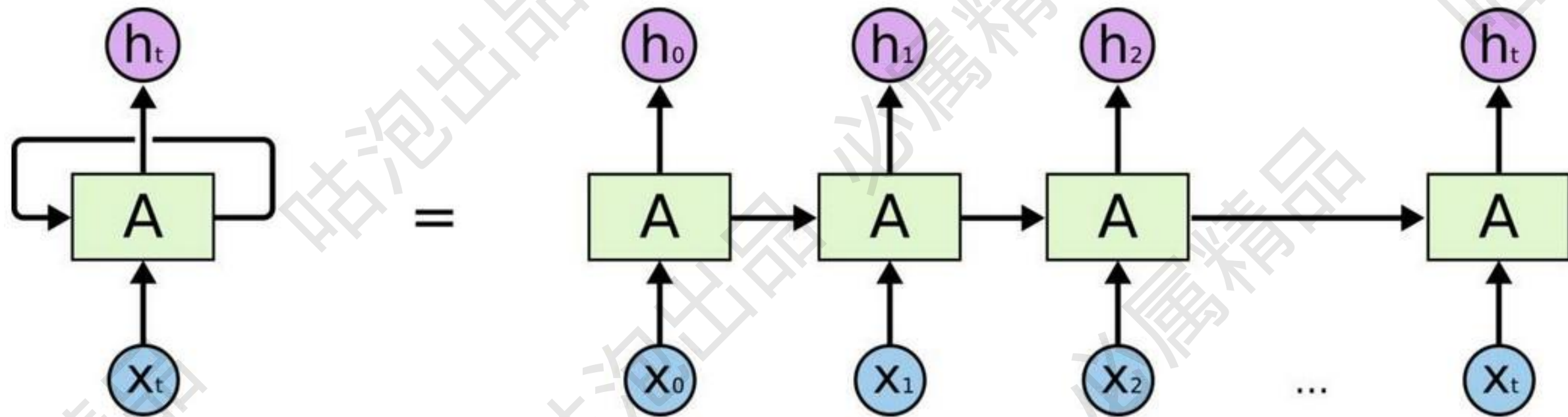
Output Layer

Hidden Layer

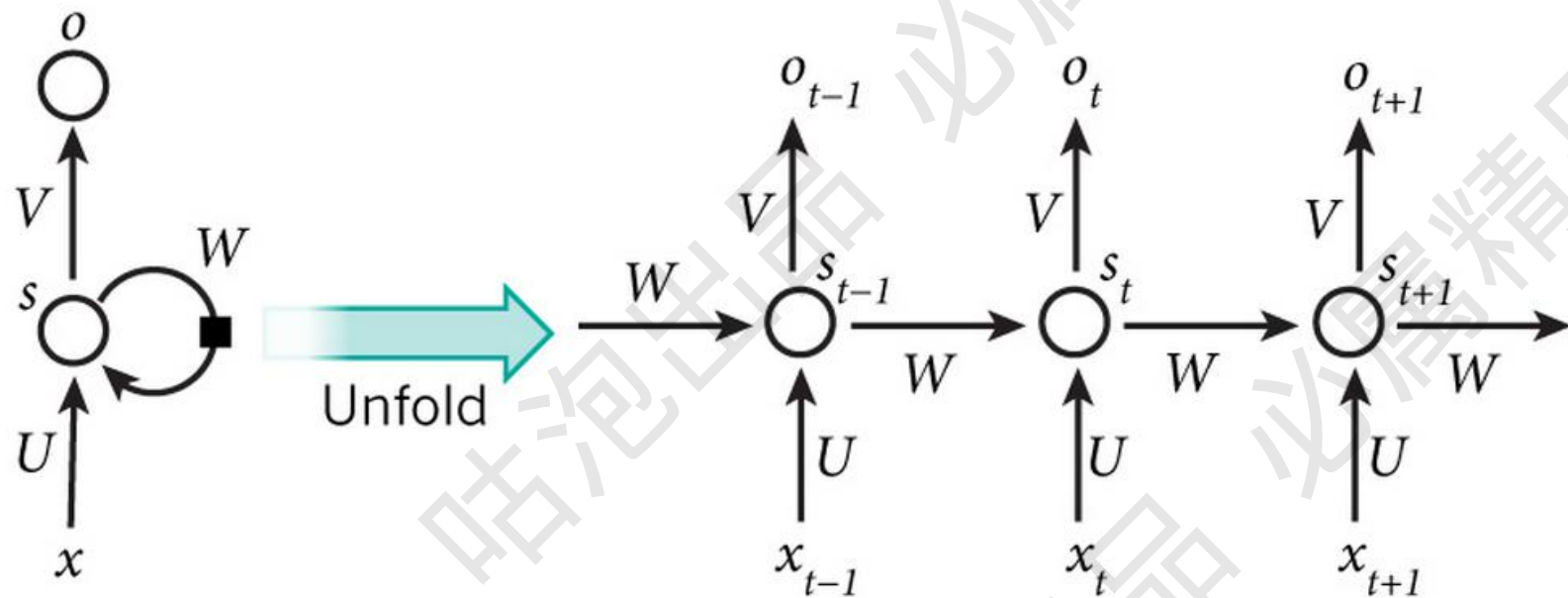
Input Layer



RNN



RNN



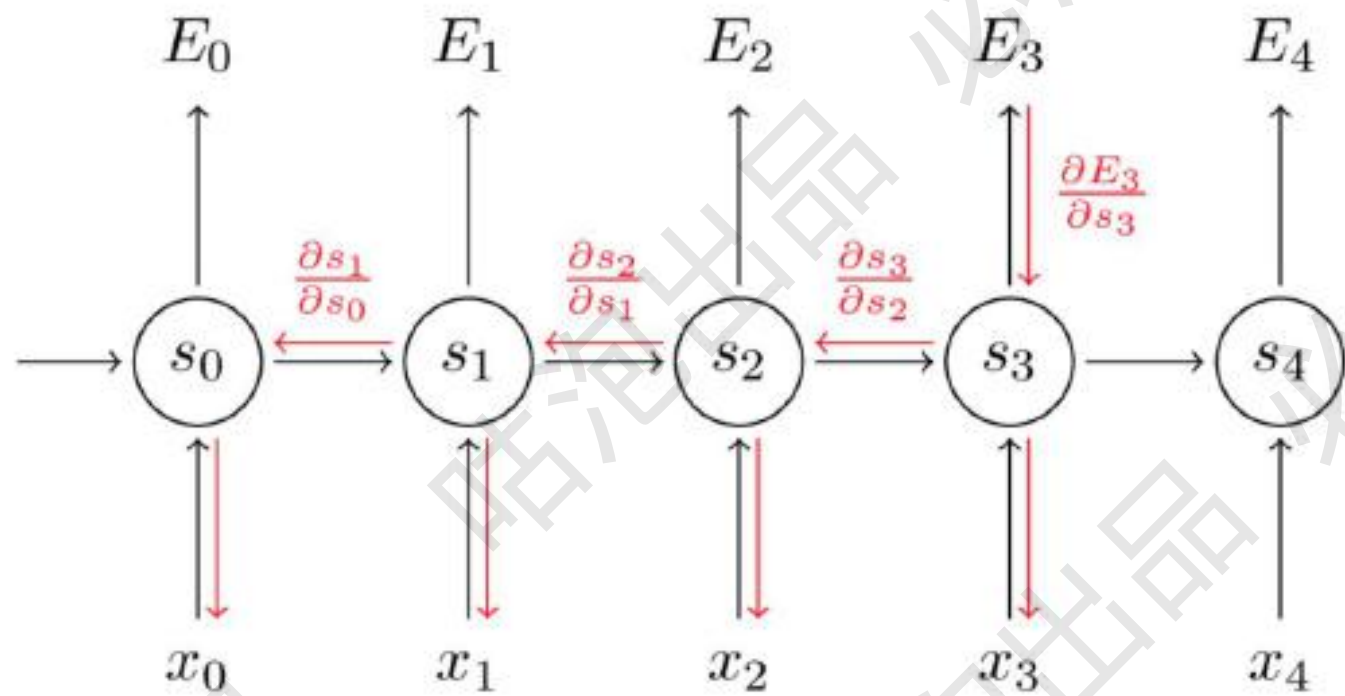
x_t 表示第 $t, t=1, 2, 3 \dots$ 步(step)的输入

s_t 为隐藏层的第 t 步的状态，它是网络的记忆单元。

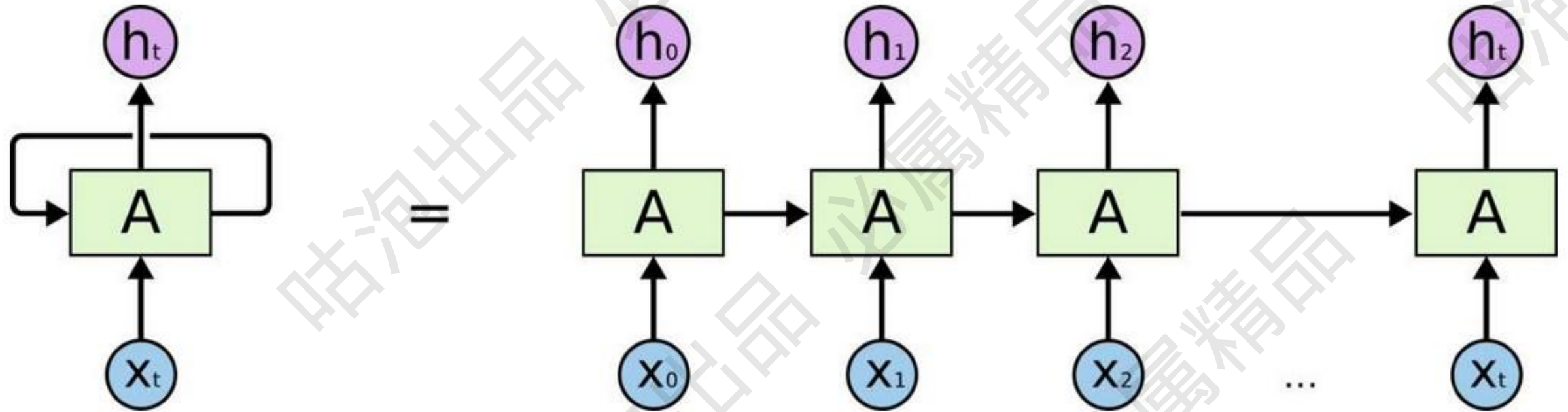
$s_t = f(Ux_t + Ws_{t-1})$ ，其中 f 一般是非线性的激活函数

o_t 是第 t 步的输出，如下个单词的向量表示 $\text{softmax}(Vs_t)$

RNN



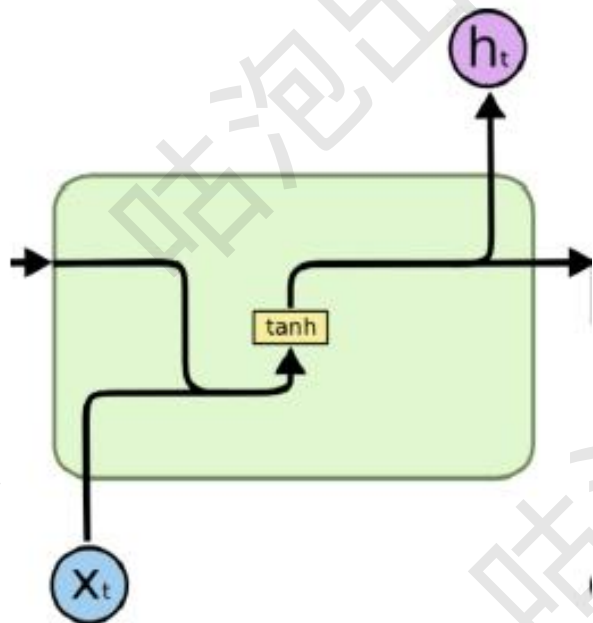
RNN



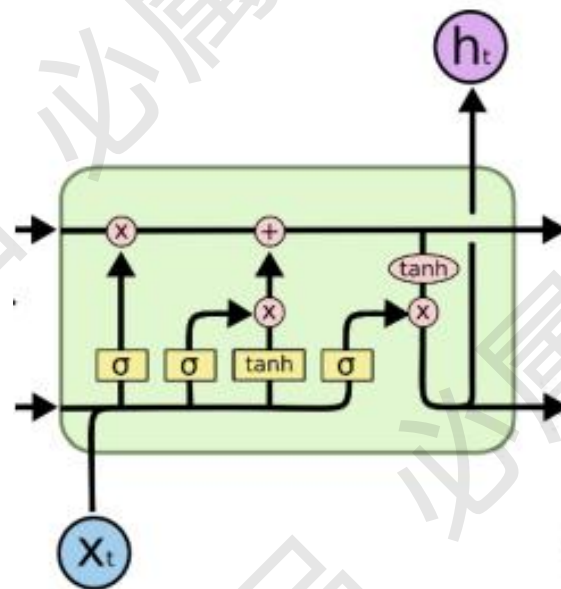
I am Chines, I Love China

LSTM

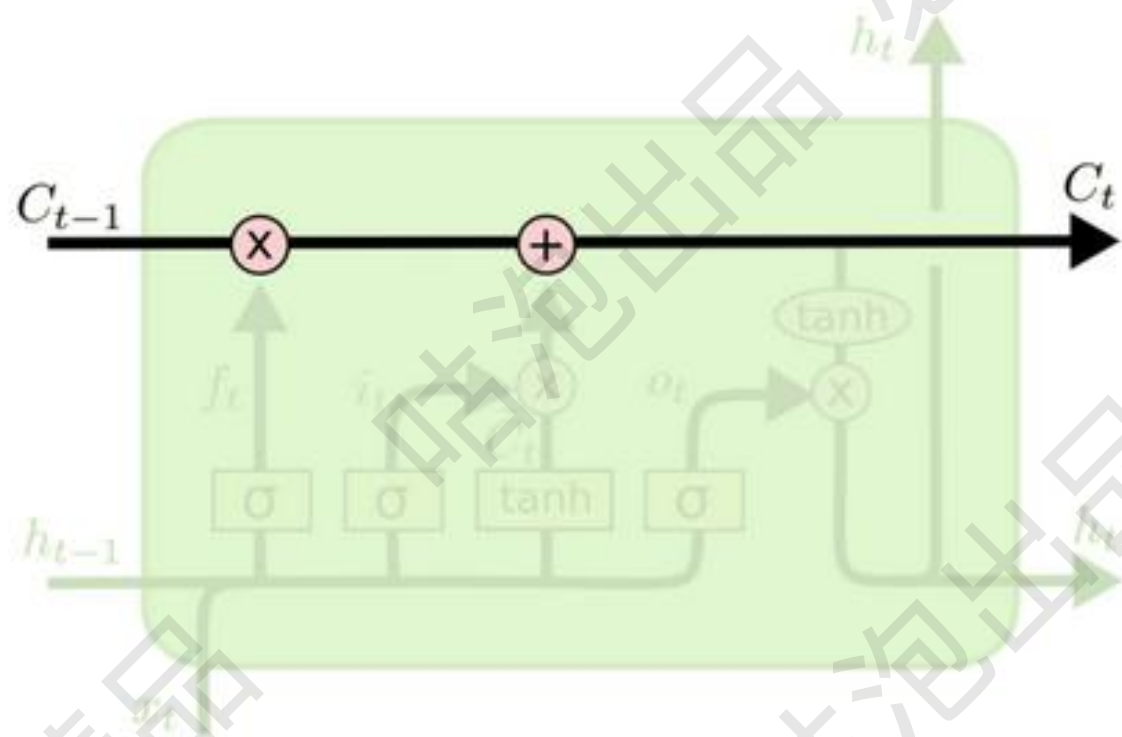
RNN



LSTM



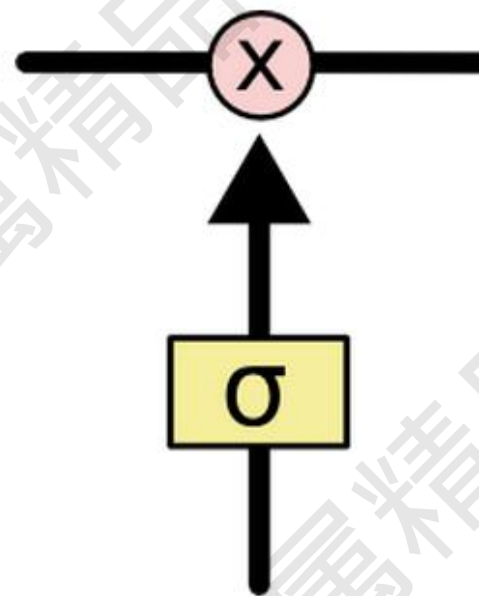
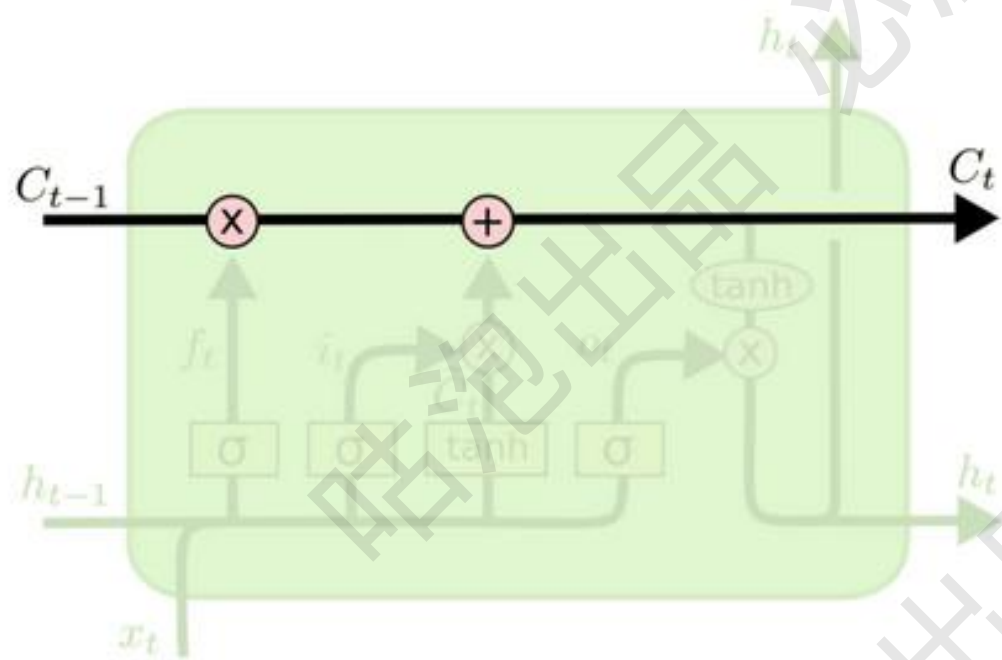
LSTM



C: 控制参数

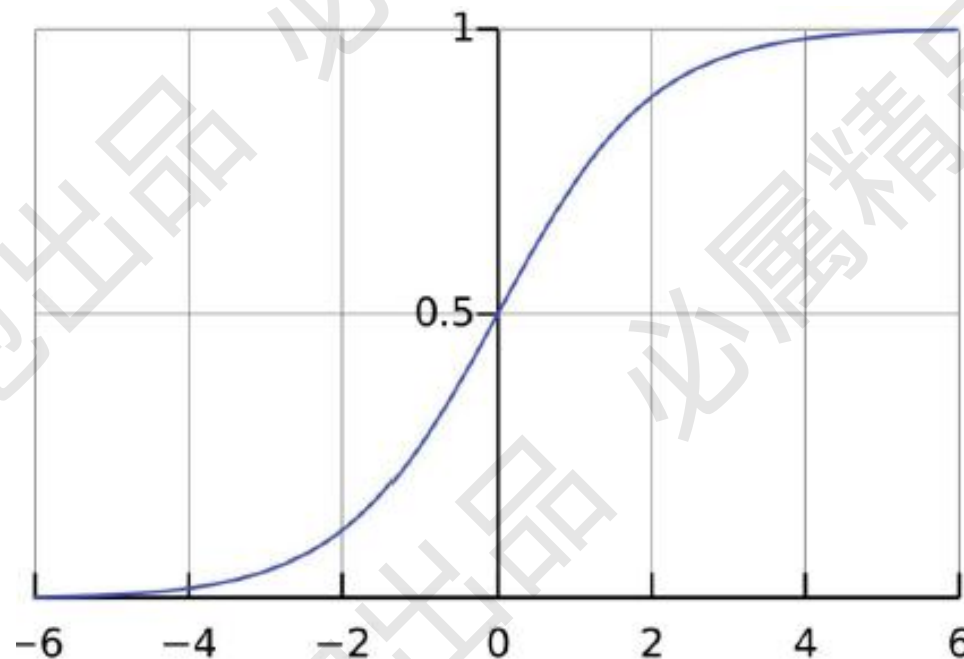
决定什么样的信息会被保留什么样的会被遗忘

LSTM



门是一种让信息选择式通过的方法
sigmoid 神经网络层和一乘法操作

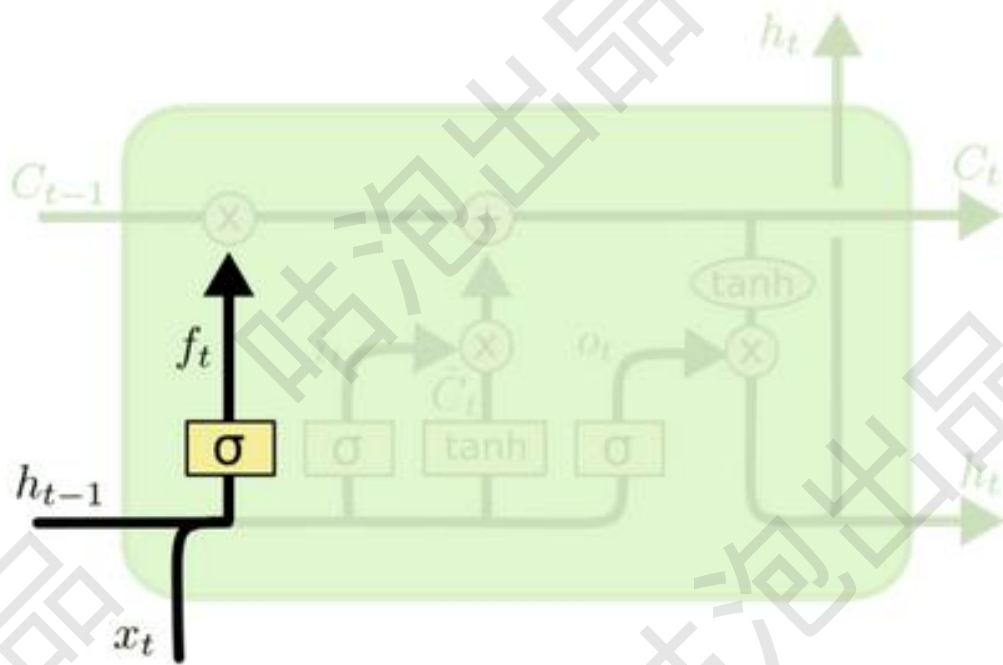
LSTM



Sigmoid函数

Sigmoid 层输出 0 到 1 之间的数值，描述每个部分有多少量可以通过。 0 代表“不许任何量通过”， 1 就指“允许任意量通过”！

LSTM

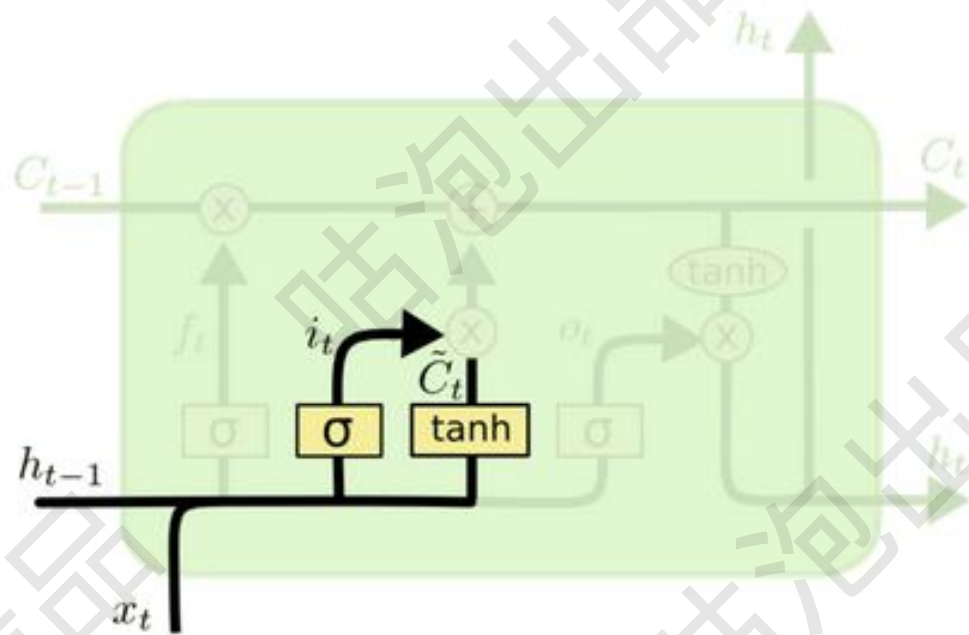


$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

f_t 与 C_{t-1} 计算决定丢弃什么信息

决定丢弃信息

LSTM



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

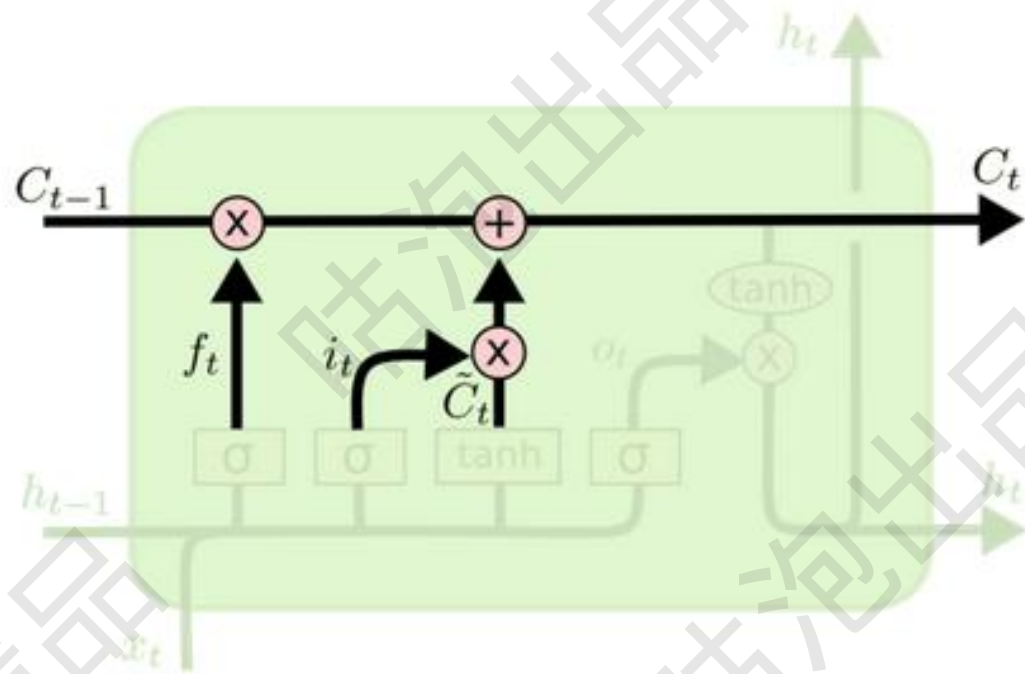
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

i_t 要保留下来的新信息

C_t 新数据形成的控制参数

确定更新的信息

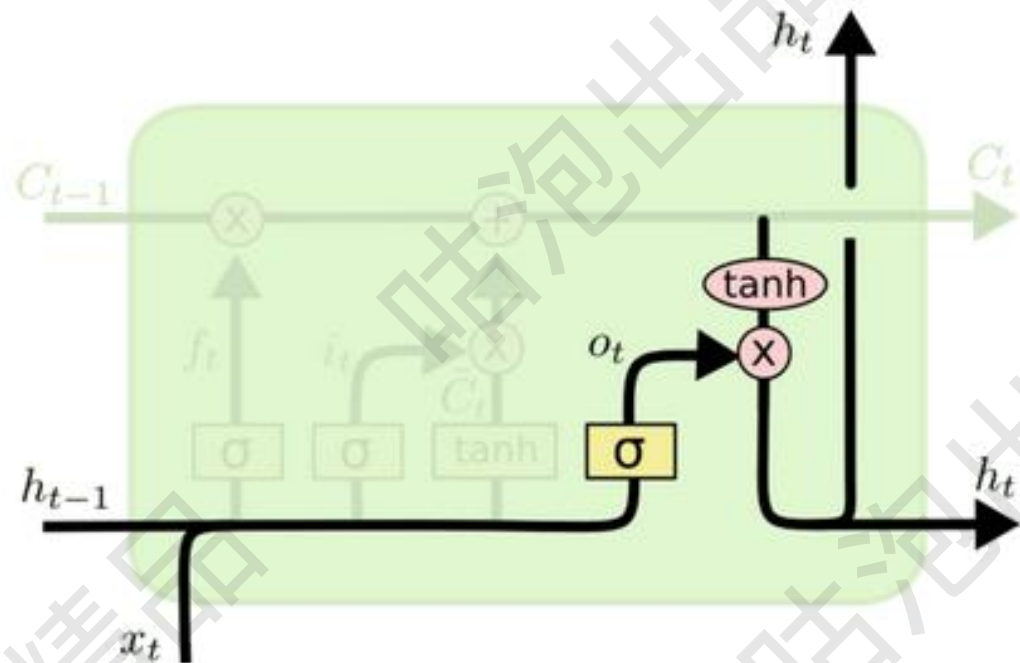
LSTM



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

更新细胞状态

LSTM



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

$$h_t = o_t * \tanh(C_t)$$

利用新的控制参数产生输出

输出信息

LSTM

