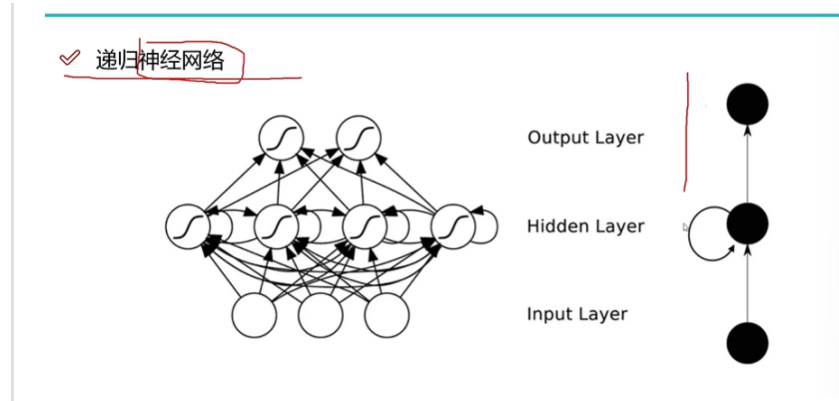
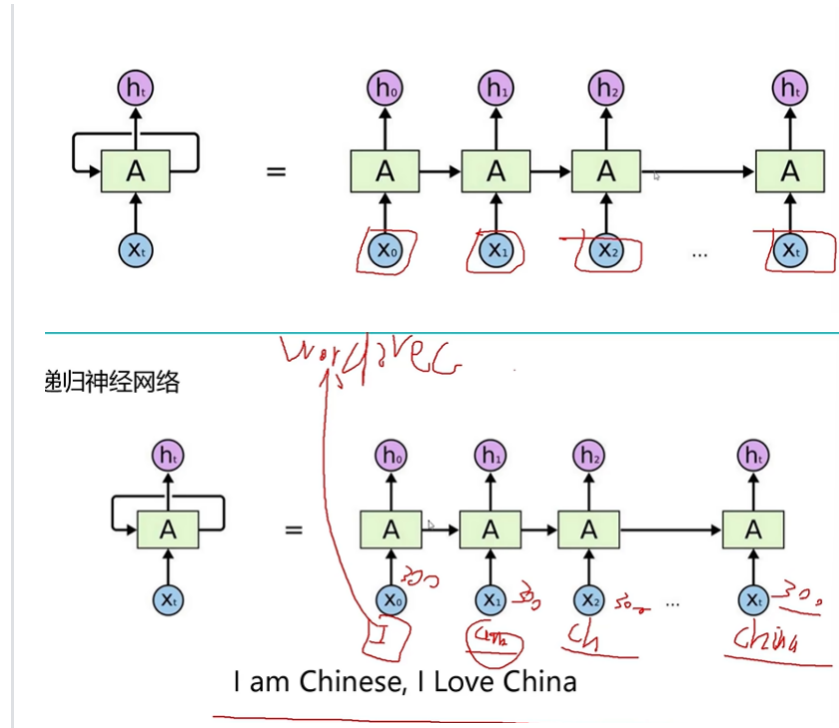


# 递归神经网络

- LSTM（在传统神经网络基础上进行改进得到）（前一个时刻训练出来的特征也会对后一个造成影响，如下图 第二个黑圆输入有两个箭头）（可以考虑时间序列的关系，他会学习前后的关系）（CNN-计算机视觉CV，RNN-自然语言处理NLP）



- （例如得到数据，没有显示出时间序列特征，但是可以自己进行拆分， $x_0$ 、 $x_1$ 、 $x_3$ 。h代表中间结果，最终只选择一个结果，选择 $h_t$ {总结了前面}）（单词编码-word2vec，把一个词转化为向量）

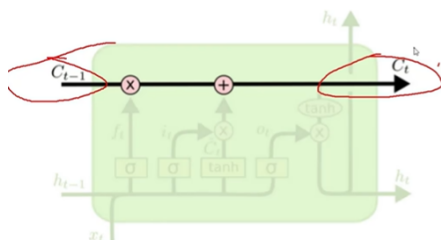
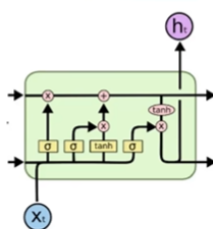
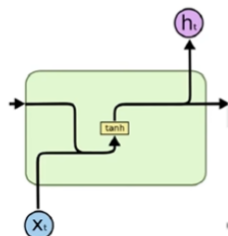


- LSTM（RNN会把之前所有的东西全部记下来，但是不是所有结果不一定有用，不精）（LSTM会让他烧鸡一点，控制参数，门单元-sigmoid。结果要跟前一轮结果比较，看要丢弃什么）（说白了，LSTM是在RNN的基础上进行改进的，进行了信息过滤，加上了控制参数）

## ✓ LSTM网络

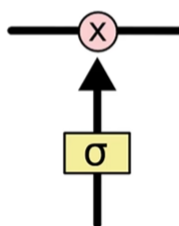
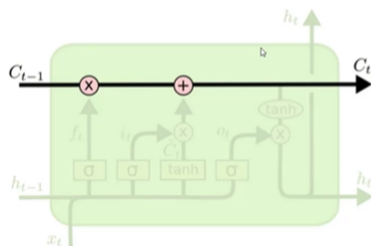
RNN

LSTM

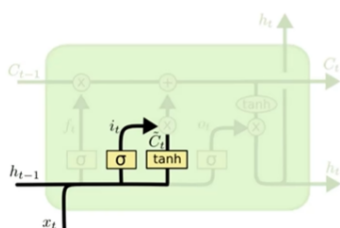


### C: 控制参数

决定什么样的信息会被保留什么样的会被遗忘



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



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

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

$i_t$  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

- 词向量模型-Word2Vec (改变位置不会改变结果, 但是语言是有顺序的, 所有要考虑顺序; 有些词比较相近, 在向量层面的表达也要比较相似)