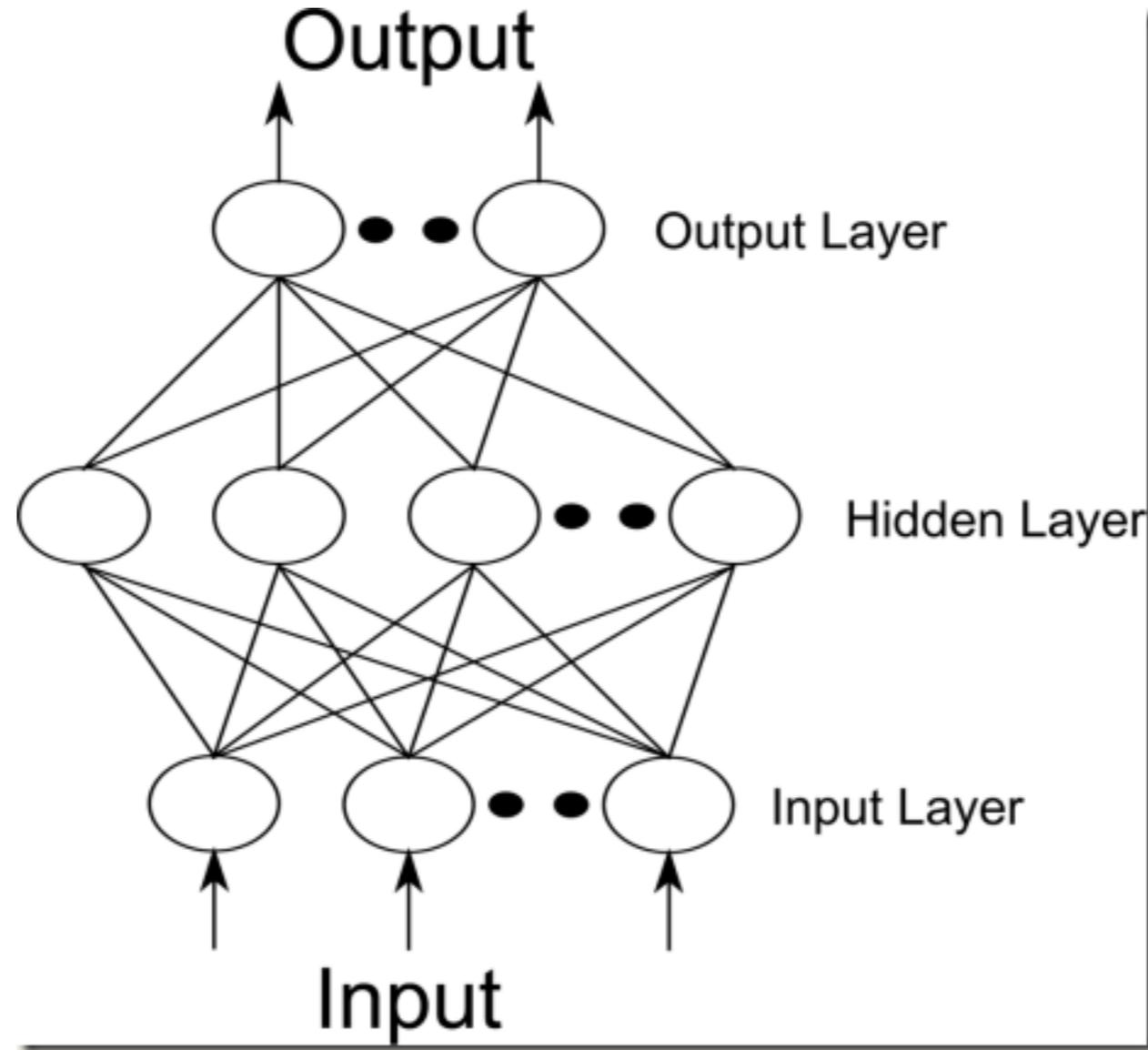
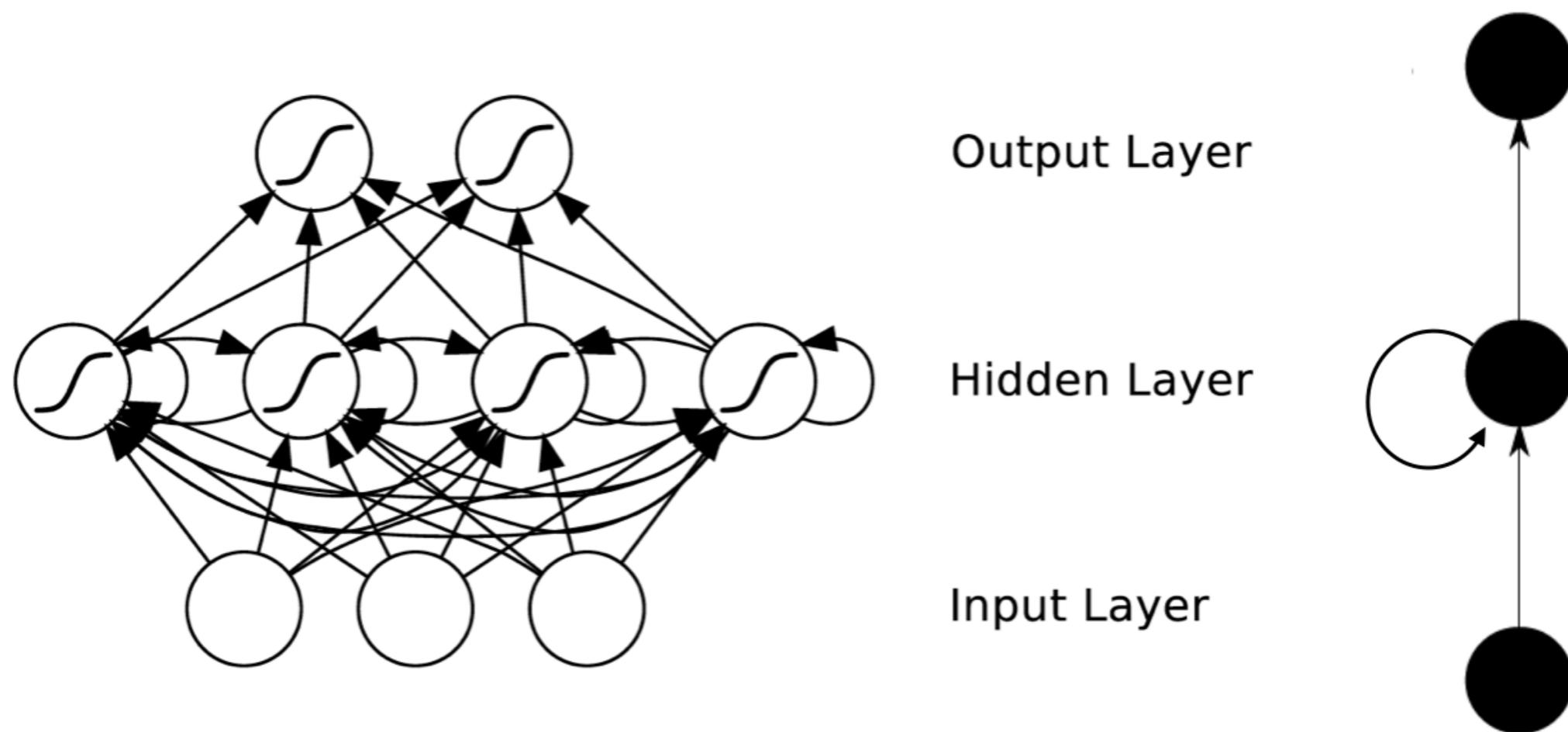


Recurrent Neural Networks

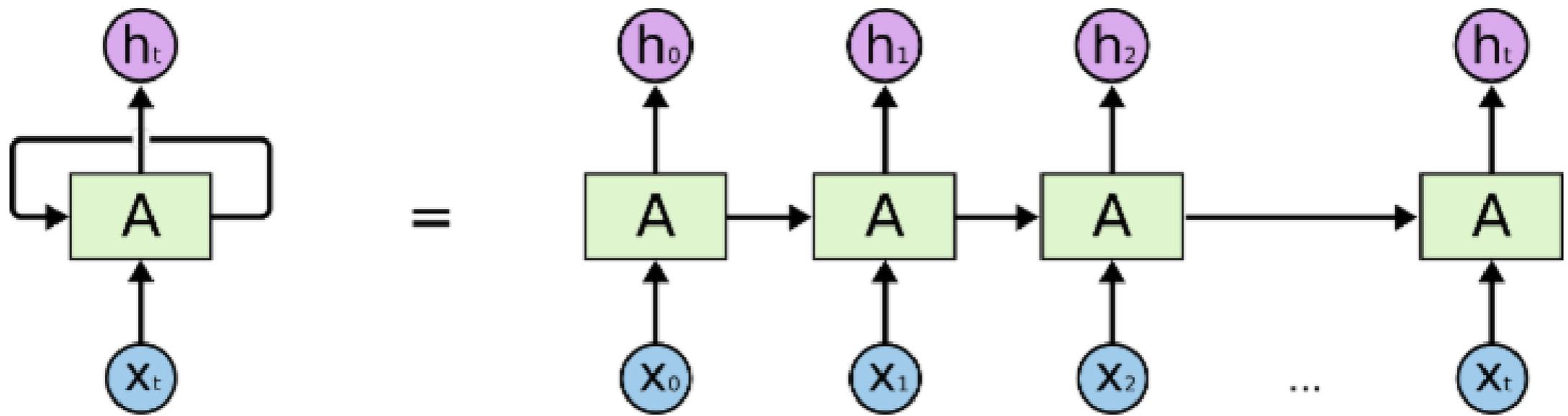


每层之间的节点是无连接的

无法解决序列问题，即每一个输出不是独立的，而是与前面的输出有关

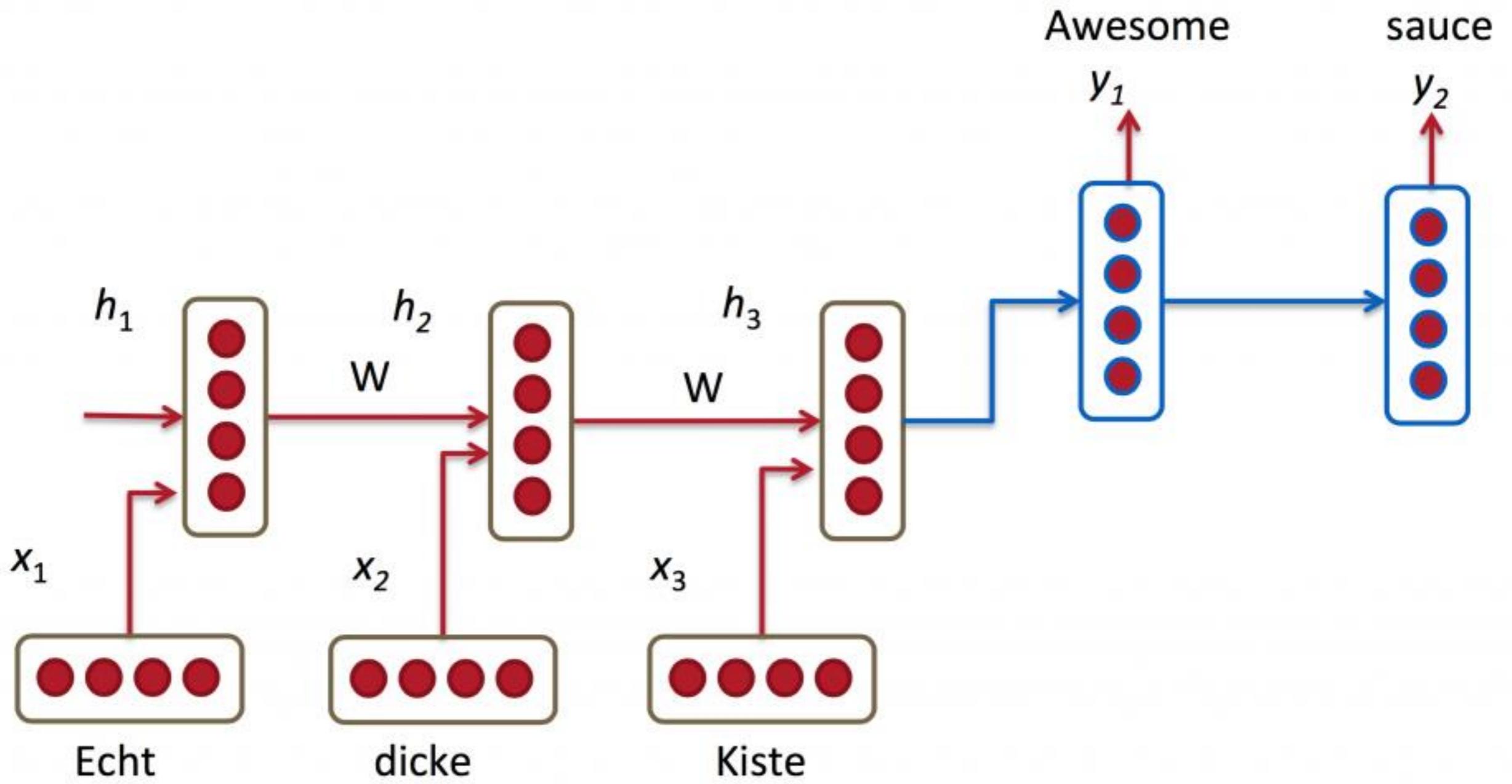


RNN之所以称为循环神经网络，即一个序列当前的输出与前面的输出也有关。具体的表现形式为网络会对前面的信息进行记忆并应用于当前输出的计算中，即隐藏层之间的节点不再无连接而是有连接的，并且隐藏层的输入不仅包括输入层的输出还包括上一时刻隐藏层的输出。



RNN对于序列的处理过程

词向量表达、语句合法性检查、词性标注



RNN情感分析过程

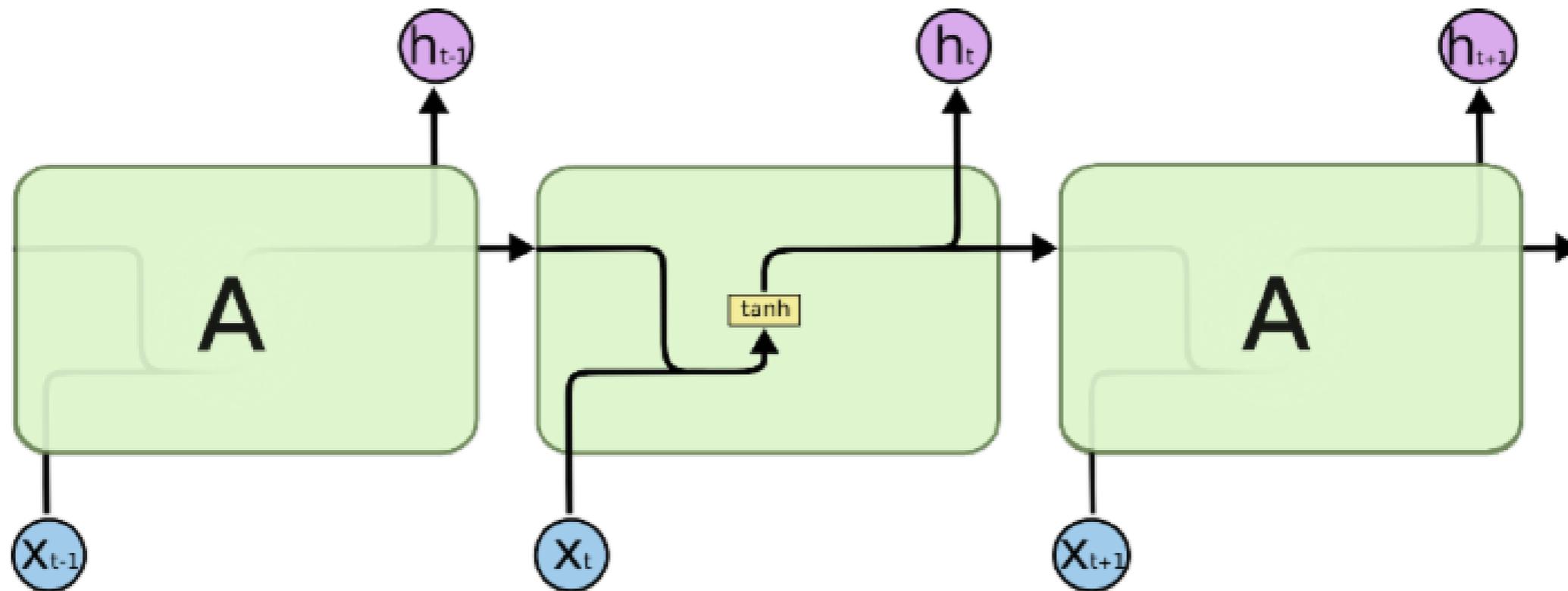
如何训练RNN

BP误差反向传播算法的改进 BPTT

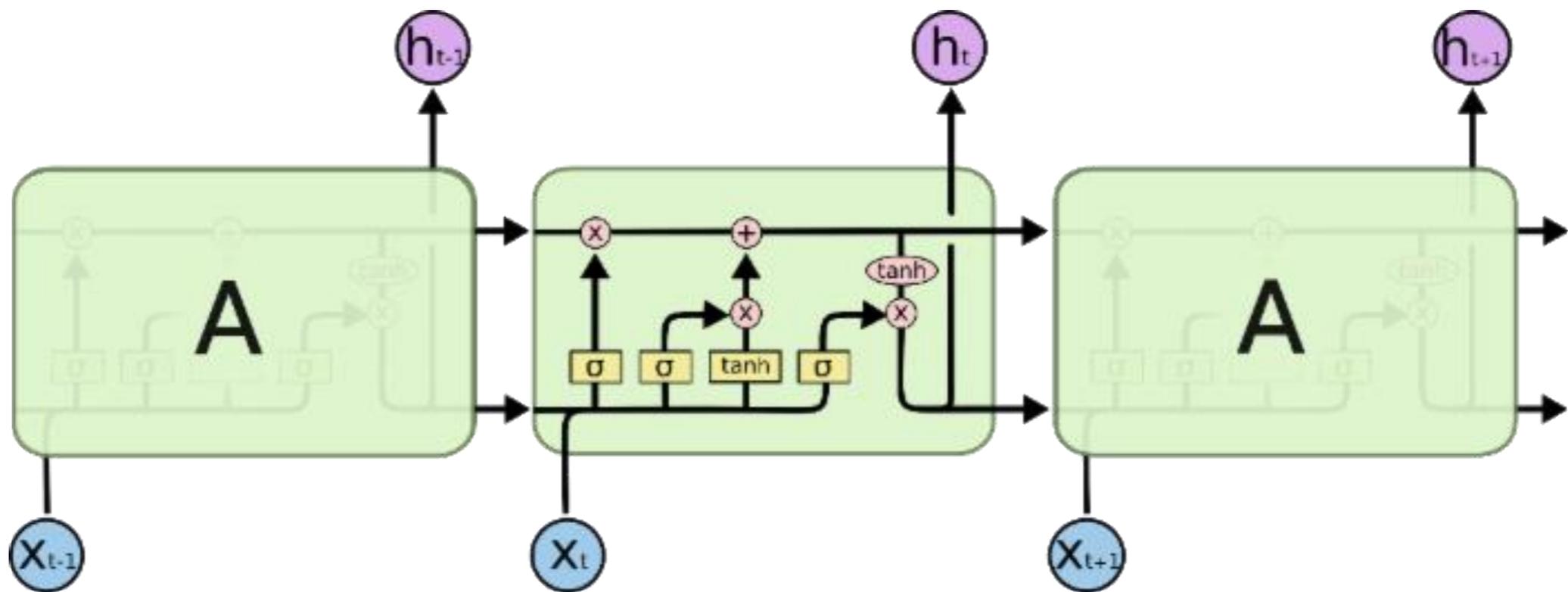
缺点: BPTT无法解决长时依赖问题, 即如果与很长一段序列有关, 则可能造成梯度爆炸问题

LSTM 网络

LSTM 通过刻意的设计来避免长期依赖问题

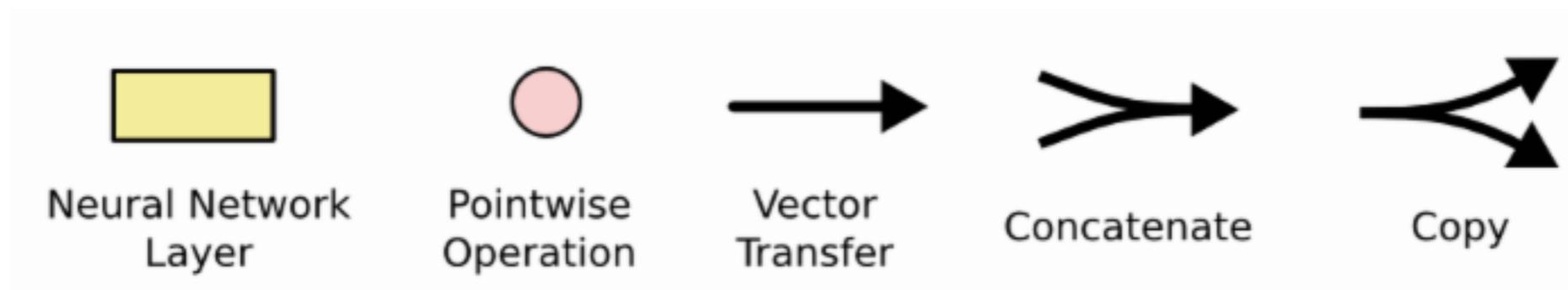


普通RNN网络

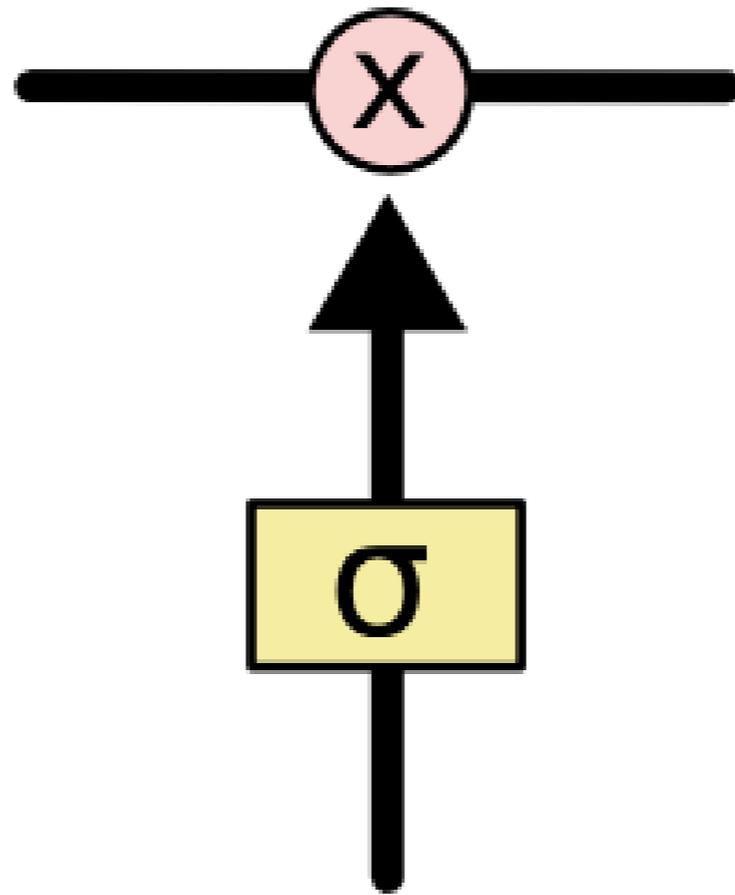


LSTM

不同于 单一神经网络层，这里是有四个，以一种非常特殊的方式进行交互。



在上面的图例中，每一条黑线传输着一整个向量，从一个节点的输出到其他节点的输入。粉色的圈代表 **pointwise** 的操作，诸如向量的和，而黄色的矩阵就是学习到的神经网络层。合在一起的线表示向量的连接，分开的线表示内容被复制，然后分发到不同的位置。

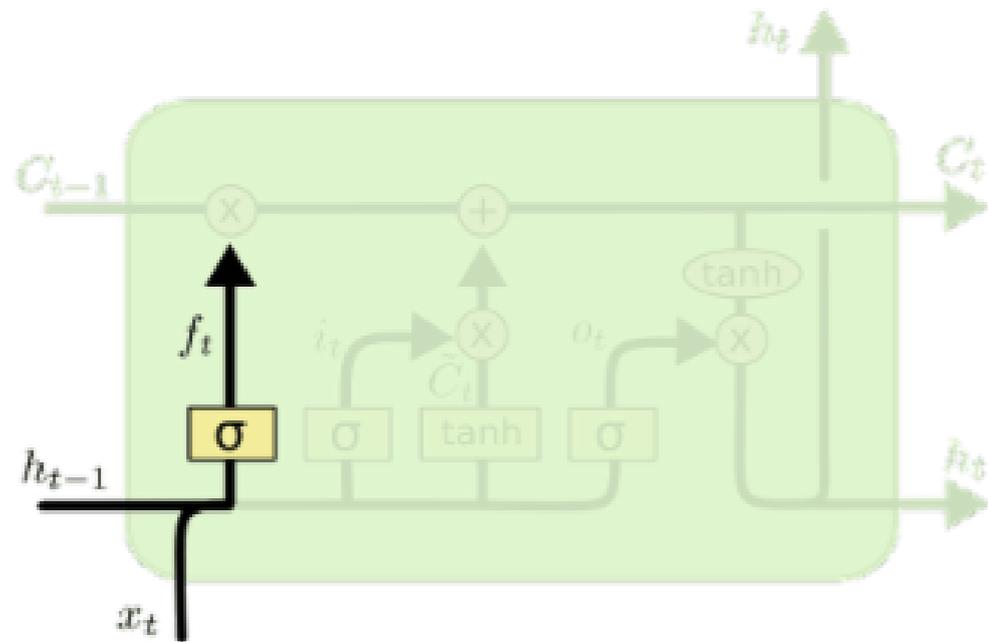


LSTM 有通过精心设计的称作为“门”的结构来去除或者增加信息到细胞状态的能力。

门是一种让信息选择式通过的方法。

他们包含一个sigmoid神经网络层和一个pointwise 乘法操作。

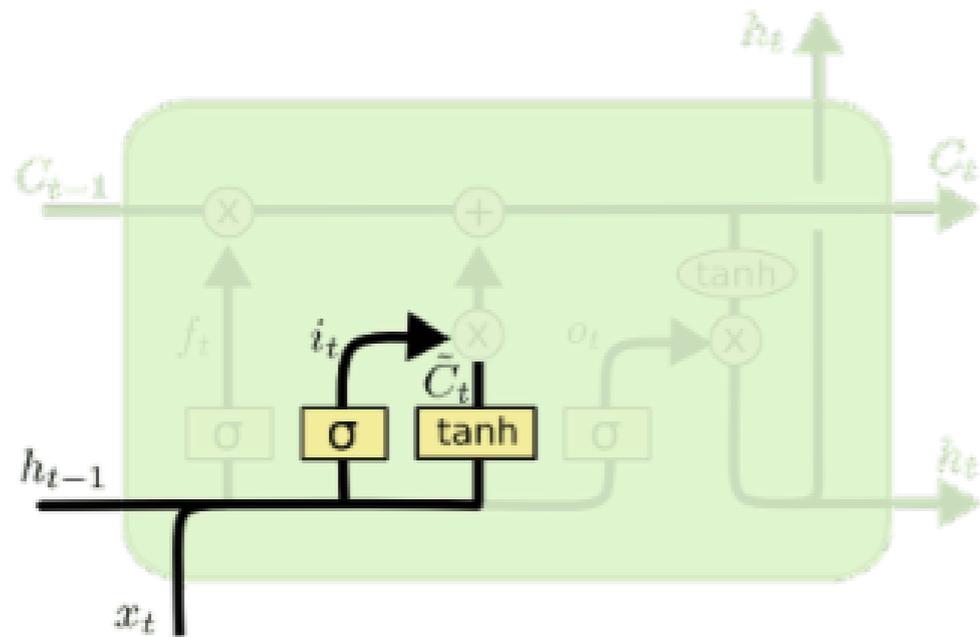
通过sigmoid函数来评判每个部分有多少量可以通过，可能为部分量，也可能为任意量



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

决定丢弃信息

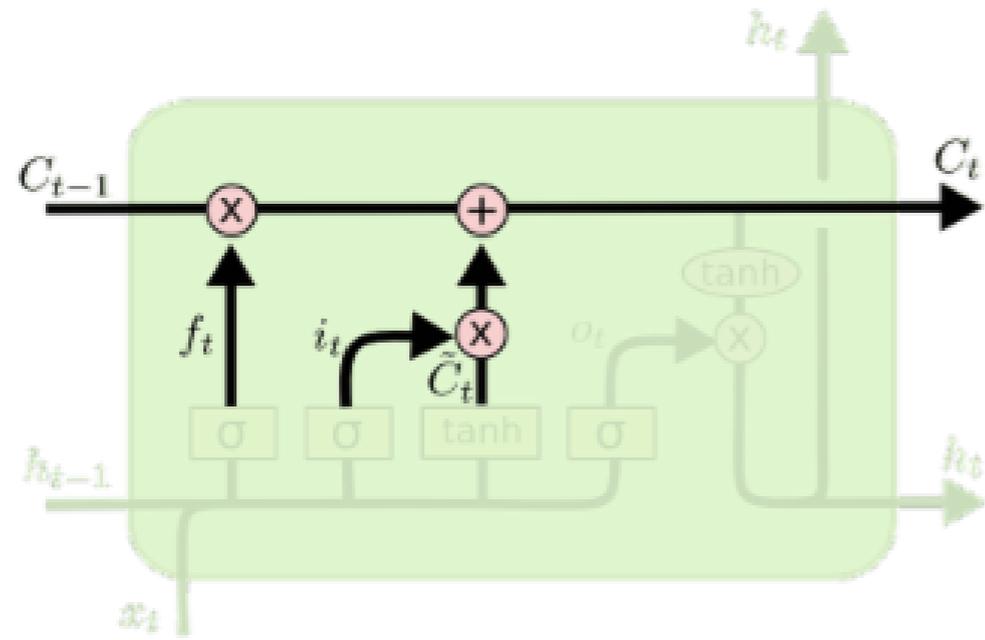
- e. g. 在这个问题中，细胞状态可能包含当前主语的分类，因此正确的代词可以被选择出来。当我们看到新的代词，我们希望忘记旧的代词。



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

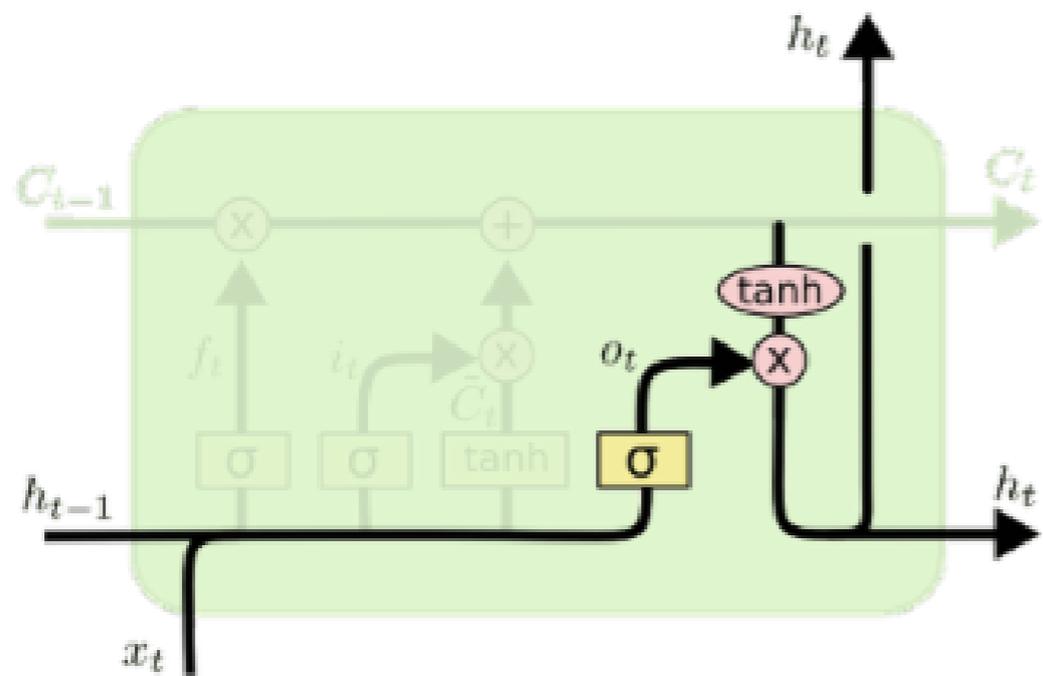
$$\tilde{C}_t = \sigma(W_C \cdot [h_{t-1}, x_t] + b_C)$$

确定更新的信息



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

更新细胞状态



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

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

输出信息

Thanks