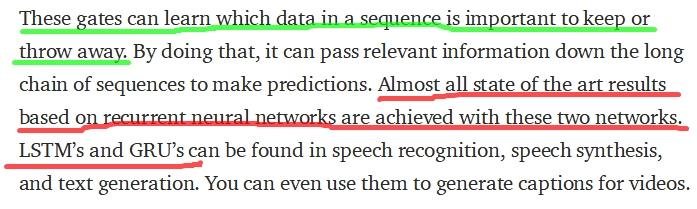
**LSTM原理**

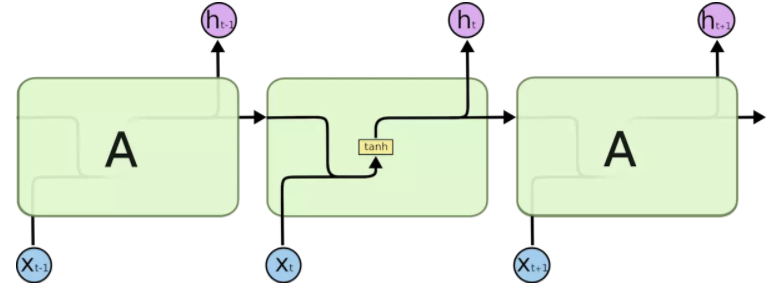
1. **LSTM**

LSTM (long short term memory)，即长短期记忆网络，是一种RNN特殊类型，其目的是解决RNN中长时间依赖问题，其核心是通过内置门控达到调节信息流的目的。引用参考文献[3]中一段话说明LSTM重要性

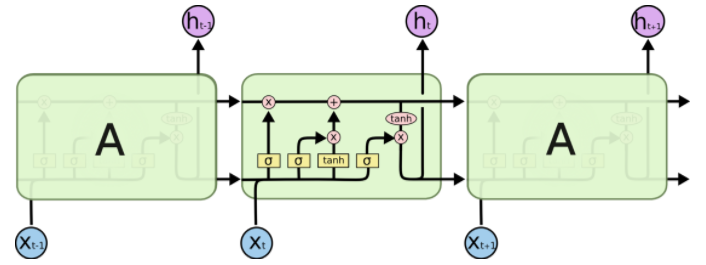


1. **LSTM结构图**

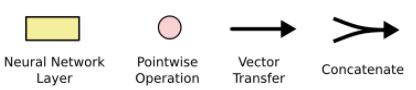
所有的RNN 都具有一种重复神经网络模块的链式形式。在标准的 RNN 中，这个重复的模块只有一个非常简单的结构，例如一个 tanh层，一般结构如下



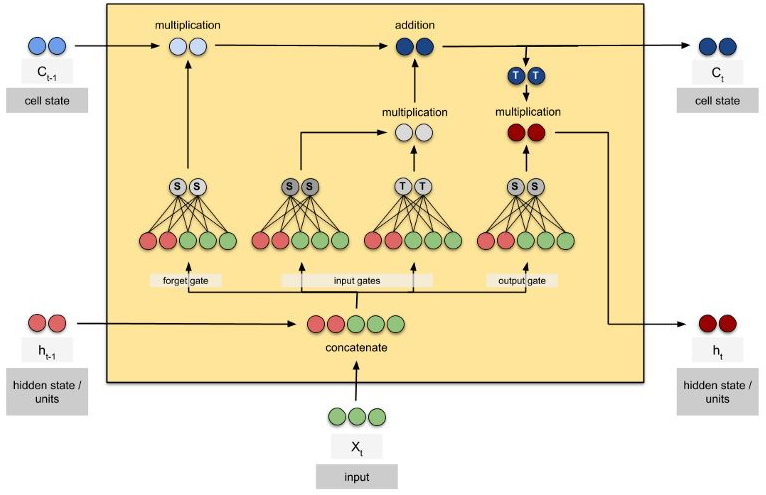
而LSTM区别在于设计了不同的重复模块，这个重复模块内部并非单一神经网络层，而是有四层，且不同层间以特殊的方式进行交互，结构如下



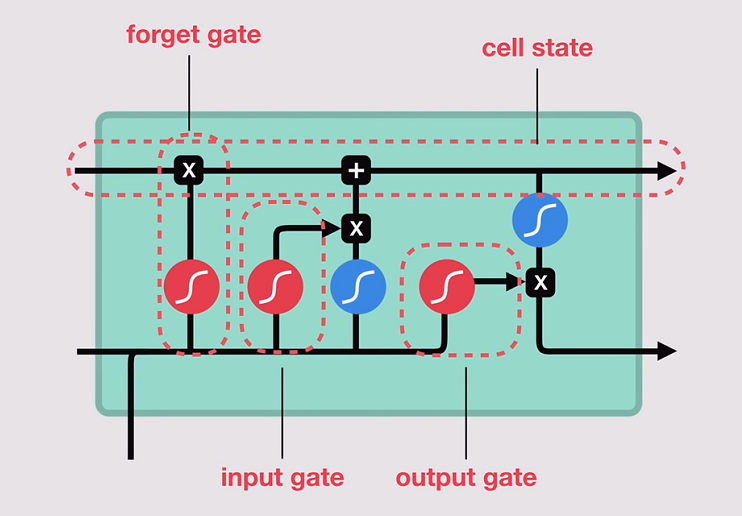
其中



节点图表示如下

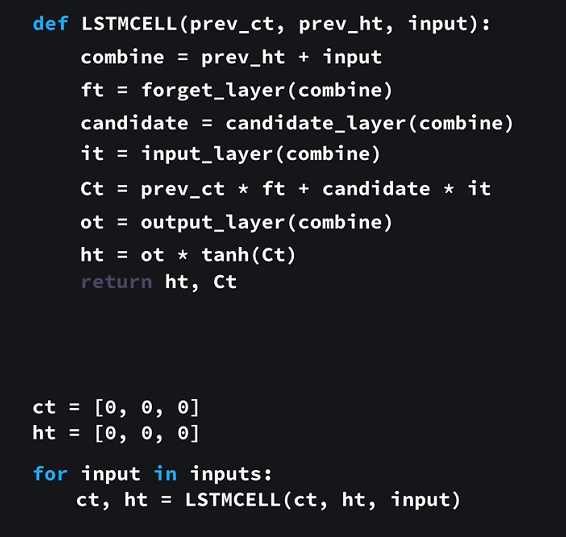


门控示意图

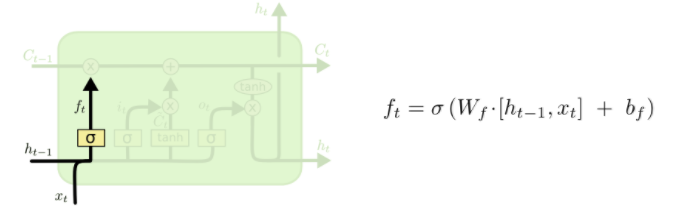


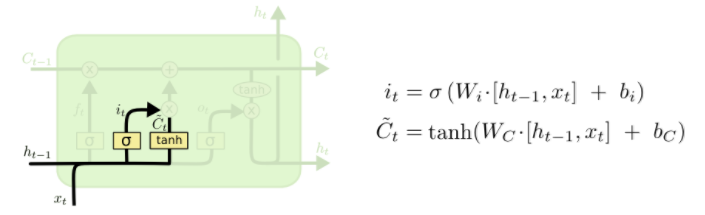


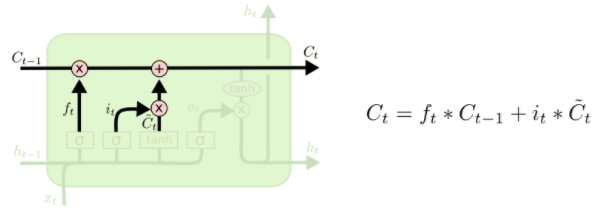
伪码

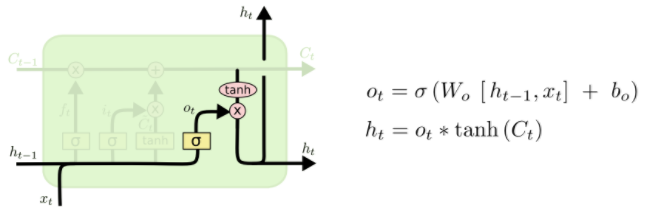


1. **前向传播函数方程**



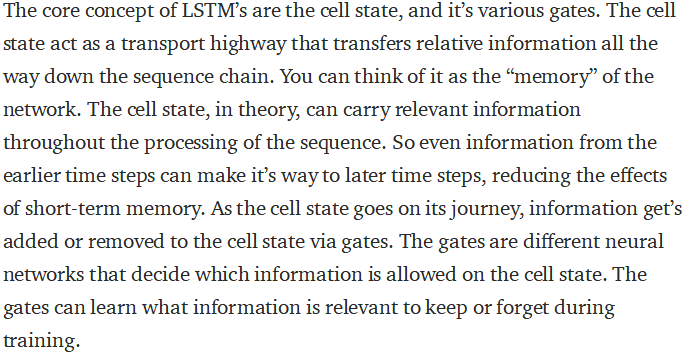






为sigmoid函数，为tanh函数。

1. **核心思想—门控原理**



1. 遗忘门 (Forget gate)

利用sigmoid函数取值范围为[0,1]，越接近0遗忘程度越大，与先前单元信息做点积，sigmoid为0时遗忘先前单元中对应部分的信息。

1. 输入门 (Input Gate)

先前单元中隐藏层信息与当前单元输入信息结合，该结合信息经过激活函数为tanh网络将其映射到[-1,1]，同时该结合信息经过激活函数sigmoid将其映射到[0,1]，前述两者映射后的信息做点积，得到调节后的输入信息，将调节后信息加入当前单元状态。

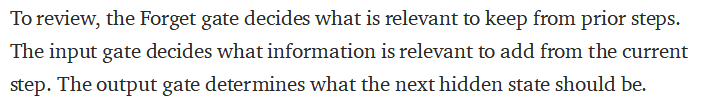
1. 单元状态 (Cell State)

该单元的信息传递纽带或者枢纽，同时包含了先前单元的信息，含有记忆功能。

1. 输出门 (Output Gate)

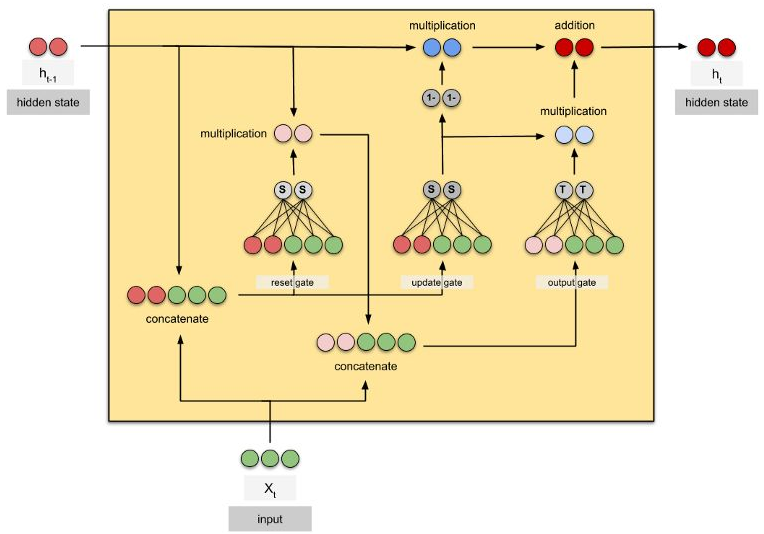
通过sigmoid函数调节该单元输出到下一个单元隐藏层的信息，以及该单元预测输出的信息。

总之：

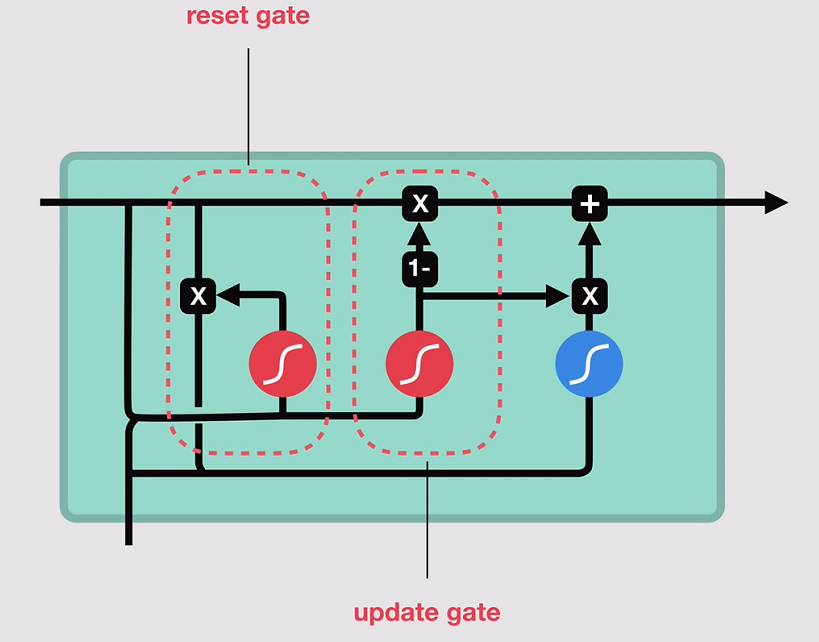


**附录1**

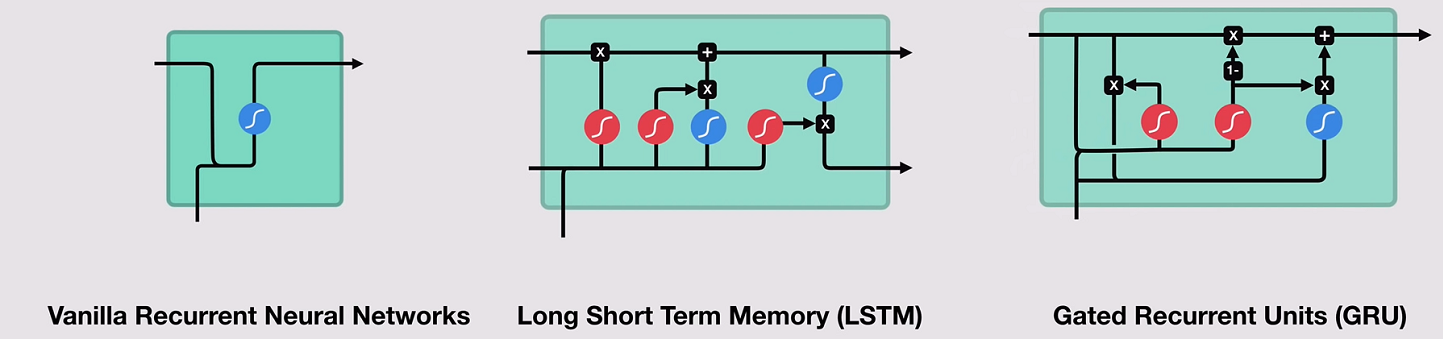
GRU节点图



GRU示意图



Vanilla RNN 、LSTM与GRU



**参考文献**

1. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
2. <https://www.lizenghai.com/archives/24286.html>
3. <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>