RNN循环神经网络和LSTM网络

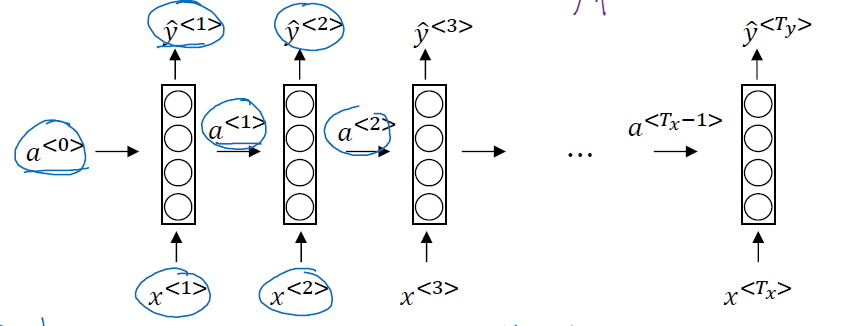
(参考网易云课堂吴恩达课程)

**一、RNN运用场景**

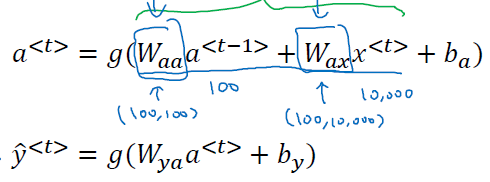
用来处理序列数据，处理文本，最常见的如自然语言处理。RNN每一层隐藏层会输出一个a值，将其保存下来并传递给下一个隐藏层，因此可以充分利用好序列中每一个元素的信息。

**二、RNN原理**

1.前向传播过程

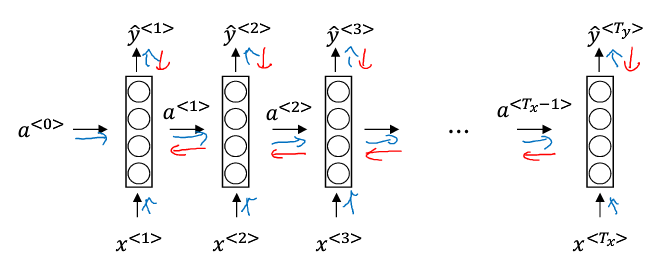


初始化a<0>，传入x<1>，计算出a<1>，利用a<1>计算出y<1>,计算方向为逆时针

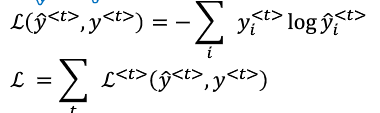


参数说明：g为激活函数，第一行常用tanh函数，第二行常用sigmoid函数，b为偏置项，W为权重

2.反向传播



反向传播过程如红色箭头标识，从y<Ty>传播至a<Tx-1>，采用交叉熵函数，作梯度下降：



L函数作为交叉熵误差函数，梯度下降，修改权重W和偏置项b

**三、GRU和LTSM**

在理论上，RNN 绝对可以处理这样的 长期依赖 问题。人们可以仔细挑选参数来解决这类问题中的最初级形式，但在实践中，RNN 肯定不能够成功学习到这些知识。Bengio, et al. (1994)等人对该问题进行了深入的研究，他们发现一些使训练 RNN 变得非常困难的相当根本的原因。

LSTM 由Hochreiter & Schmidhuber (1997)提出，并在近期被Alex Graves进行了改良和推广。在很多问题，LSTM 都取得相当巨大的成功，并得到了广泛的使用。

GRU：the Gated Recurrent Unit门控循环单元

LTSM：the Long Short-Term Memory长短期记忆网络

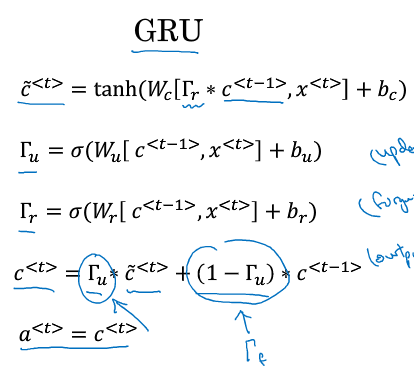
提出原因：为了使模型能够长期记忆之前的元素信息。

举例：

The cat, already ate …, was full.

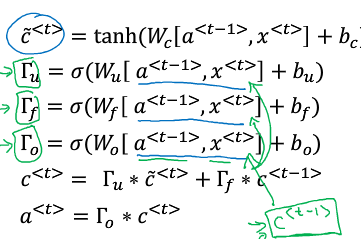
The cats, already ate …, were full.

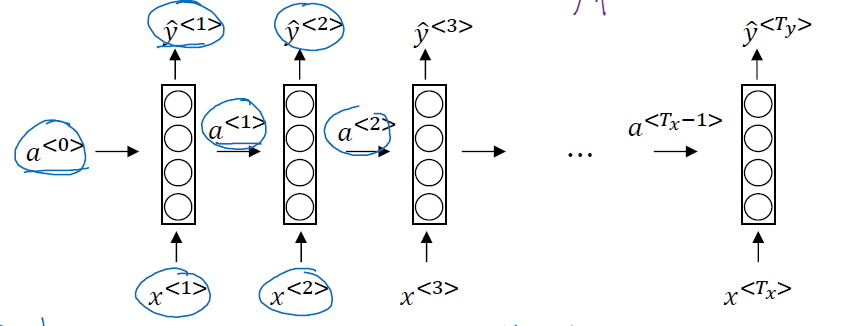
主谓语可能被间隔的太远，我们需要用GRU或LTSM来长远记忆信息。

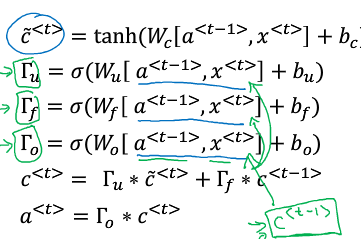


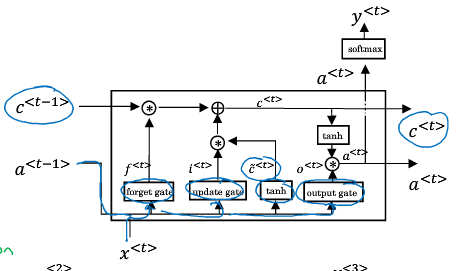
GRU符号介绍：c称为memory cell, c~<t>记为替代值，sigma是sigmoid函数，gamma称作门函数











LSTM中用到3个门函数，u: update更新门，f: forget 遗忘门, o: output 输出门

https://blog.csdn.net/sinat\_24070543/article/details/75113014