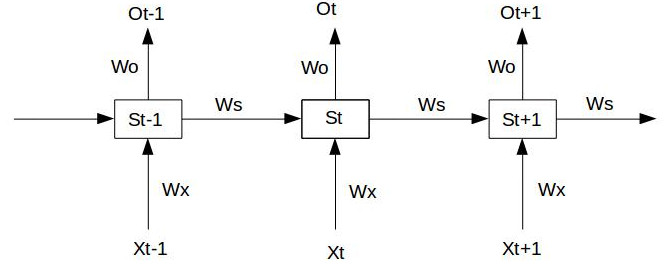
**Thinking1 常用的文本分类方法都有哪些**

主要分为两大类：

1. 基于传统机器学习的文本分类
   * 传统机器学习的文本分类通常提取tfidf或者词袋特征，然后给LR模型进行训练；这里模型有很多，比如朴素贝叶斯，KNN方法，决策树，支持向量机，GBDT/XGBOOST等
2. 基于深度学习的文本分类
   * 主要采用CNN、RNN、LSTM、GRU、Attention等

**Thinking2 RNN为什么会出现梯度消失**

经典的RNN结构如下图所示：



假设我们的时间序列只有三段，S(t)的值，不仅取决于x(t)，还取决于S(t-1)：



RNN在每个阶段都共享一个参数W，隐藏层激活函数通常采用tanh。

RNN反向传播中，需要计算U,V,W等参数的梯度

针对W梯度计算（T时刻）：



针对部分，展开可以得到：

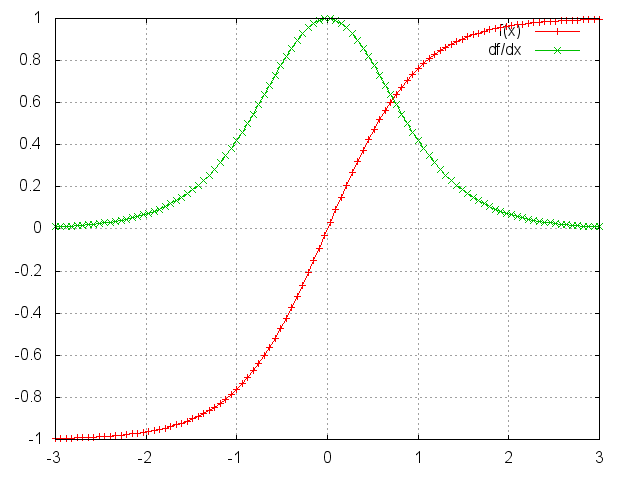


代入，求得任意时刻w的梯度：



因为

所以



w的梯度（T时刻）：





因为tanh’ <= 1，训练过程大部分情况下tanh的导数是小于1的。

如果Ws也在[0,1]之间，当T很大的时候， 接近于0 => 梯度消失

当Ws很大时，可以导致，当T很大时接近于无穷大 => 梯度爆炸