——RNN递归神经网络

自然语言处理离不开RNN

BP和CNN输入和输出相互独立，没有上下文联系。

但是实际中上下文是有关联的。预测视频中下一帧的内容，输出依赖前面的输入，必须要有记忆能力。RNN解决了BP和CNN不能记忆的作用。

有纵向输出和横向输出，横向输出给下一时刻用，纵向输出是给分类用。

U、V、W是权重；s代表细胞的状态，o代表纵向的输出，h代表横向的输出。

对于每一个时刻t的RNN网络，网络的输出ot和实际的标签都会差生一定的误差et（预测值和实际值之间的误差），误差的损失函数，可以实交叉熵也可以是平方误差。

V可以直接求导计算，W、U是不能直接计算，可以链式求导法则。

梯度求导的一部分环节是一个指数模型，x的n次方，若x<1则会出现梯度消失，当x>1，就会出现梯度爆炸。

使用梯度截断的方法克服梯度下降，当超过阈值c时，或者小于阈值-c时，设置成c或者-c。梯度变化非常大，参数变化很快，梯度变化小，参数几乎没有变化。

——LSTM是RNN的代表，RNN中介绍到对于长依赖的问题没有办法解决，可能会产生梯度消失和梯度爆炸问题，LSTM特别适合解决长依赖问题。

Pointwise operation对应的元素相乘或者相加

Concatenate合并。一个神经元有四个激活函数。三个输入两个输出。

[h, x]这个表示串联

信息增强门，正的增强，负的减弱。