RNN & BPTT

大家下午好，我是林广和。今天我我将讲一个模型递归神经网络（RNN）以及它的训练方法—BPTT。（Click）

大家都知道，传统的前馈神经网络所有的输入之间是相互独立的。而在实际中，预测一个句子的下一个词，如果能有上下文的帮助，会有更好的效果。（待补充RNN的Recurrent）而RNN正好具备这一特性，能充分利用上下文的信息，具有很好的预测效果。理论上我们能利用任意长度的序列信息，但在实践过程中只能局限在有限的几步，（待补充）。

我们能从这张图中看到在RNN也是由输入x，隐藏层s和输出层o构成，将其展开来，可以明显看到,在第t步，xt和上一个步的状态输入经过非线性的激活函数后得到了当前的状态，再经由softmax获得当前的输出ot。

这里需要注意几点：RNN中U、V、W在所有过程中均采用同一套参数，从而减少了参数学习的数量；上面的输入和输出并不是必须的，可以根据实际的需要进行调整。比如，当你对句子进行情感分类时，只有在输入整个句子后才能判断情感是正向或是负向。

RNN主要的应用于语言模型和生成文本、机器翻译、语音识别和生成图像描述。

我们来举一个例子：假定需要一个RNN模型来预测来下一个字符，一个可能的字符集为｛‘h’、‘e’、‘l’、‘o’｝，样本序列为“hello”，每一步输入一个字符，编码形式为one-hot，就如图中输入层所示，只有一个是1，其余为0；最后经过softmax得到输出。在输入‘h’时，我们可以看到输出中‘h’、‘e’、‘l’、‘o’分别被赋值为1.0,2.2,-3.0,4.0，而期望的输出是’e’，那么就需要应用反向传导算法进行参数调整，使得代表‘e’的绿色（正确部分）上增加，而其他字母（错误部分）对应的位置下降。其他也是如此。