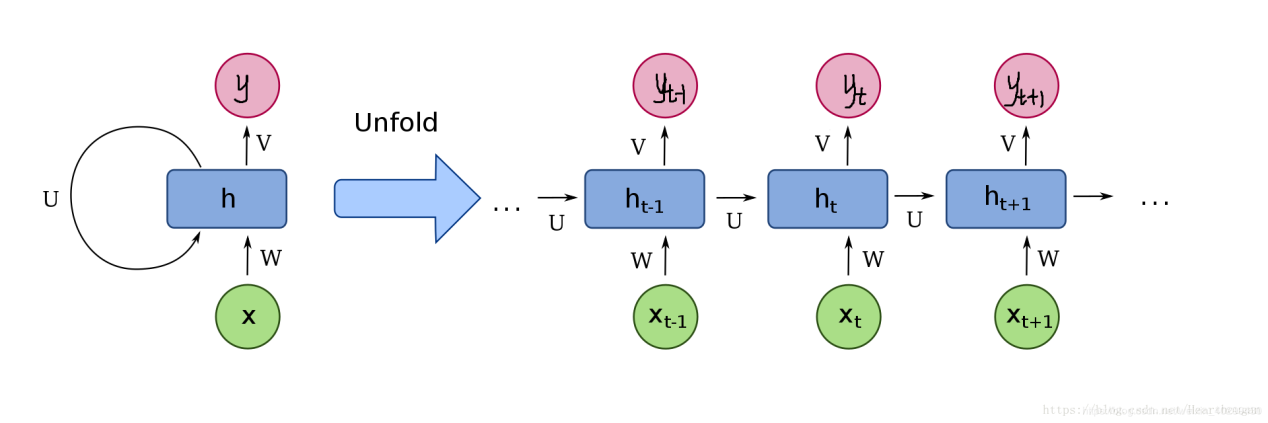
**Task5 循环和递归神经网络 (2 days)**

* RNN的结构。循环神经网络的提出背景、优缺点。着重学习RNN的反向传播、RNN出现的问题（梯度问题、长期依赖问题）、BPTT算法。
* 双向RNN
* 递归神经网络
* LSTM、GRU的结构、提出背景、优缺点。
* 针对梯度消失（LSTM等其他门控RNN）、梯度爆炸（梯度截断）的解决方案。
* Memory Network（自选）
* Text-RNN的原理。
* 利用Text-RNN模型来进行文本分类。
* Recurrent Convolutional Neural Networks（RCNN）原理。
* 利用RCNN模型来进行文本分类。

循环神经网络（Recurrent Neural Network，RNN）DNN以及CNN在对样本提取特征的时候，样本与样本之间是独立的，而有些情况是无法把每个输入的样本都看作是独立的，比如NLP中的此行标注问题，ASR中每个音素都和前一个音素是相关的，这类问题可以看做一种带有时序序列的问题，无法将样本看做是相互独立的，因此单纯的DNN和CNN解决这类问题就比较棘手。此时RNN就是一种解决这类问题很好的模型。



由上图可以看出，RNN的结构是一个重复的过程，且权重W,U,V是共享的，这也是借鉴了CNN的思想，可以减少参数量，从而减少计算的复杂度。第t时刻隐藏层的输出需要t-1时刻的隐藏层的输出，RNN以此来实现信息的传递。

这是一个标准的RNN结构图，图中每个箭头代表做一次变换，也就是说箭头连接带有权值。左侧是折叠起来的样子，右侧是展开的样子，左侧中h旁边的箭头代表此结构中的“循环“体现在隐层。

在展开结构中我们可以观察到，在标准的RNN结构中，隐层的神经元之间也是带有权值的。也就是说，随着序列的不断推进，前面的隐层将会影响后面的隐层。图中O代表输出，y代表样本给出的确定值，L代表损失函数，我们可以看到，“损失“也是随着序列的推荐而不断积累的。

除上述特点之外，标准RNN的还有以下特点：

1、权值共享，图中的W全是相同的，U和V也一样。

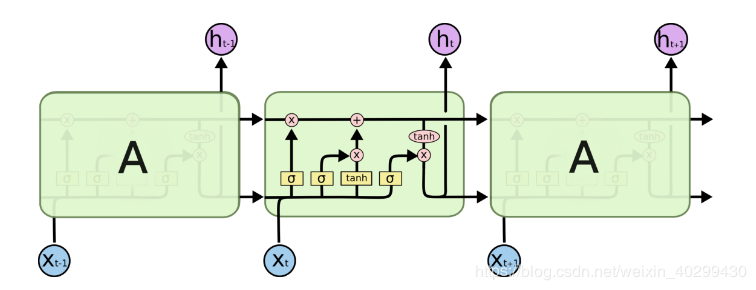
2、每一个输入值都只与它本身的那条路线建立权连接，不会和别的神经元连接。

二、递归神经网络

递归神经网络（recursive neural network）是具有树状阶层结构且网络节点按其连接顺序对输入信息进行递归的人工神经网络（Artificial Neural Network, ANN），是深度学习（deep learning）算法之一。

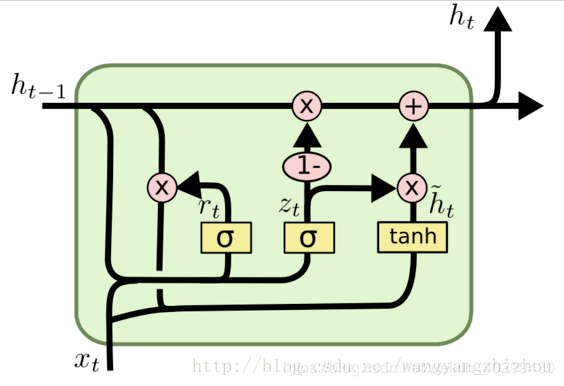
三、LSTM、GRU

long short term memory，即我们所称呼的LSTM，是为了解决长期以来问题而专门设计出来的，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RNN中，这个重复的结构模块只有一个非常简单的结构，例如一个tanh层。LSTM的结构如下图：



GRU是LSTM网络的一种效果很好的变体，它较LSTM网络的结构更加简单，而且效果也很好，因此也是当前非常流形的一种网络。GRU既然是LSTM的变体，因此也是可以解决RNN网络中的长依赖问题。

在LSTM中引入了三个门函数：输入门、遗忘门和输出门来控制输入值、记忆值和输出值。而在GRU模型中只有两个门：分别是更新门和重置门。具体结构如下图所示：



图中的zt和rt分别表示更新门和重置门。更新门用于控制前一时刻的状态信息被带入到当前状态中的程度，更新门的值越大说明前一时刻的状态信息带入越多。重置门控制前一状态有多少信息被写入到当前的候选集 htht 上，重置门越小，前一状态的信息被写入的越少。

四、Text-RNN的原理

CNN的核心点在于可以捕获信息的局部相关性，具体到文本分类任务中可以利用CNN来提取句子中类似N-Gram的关键信息。

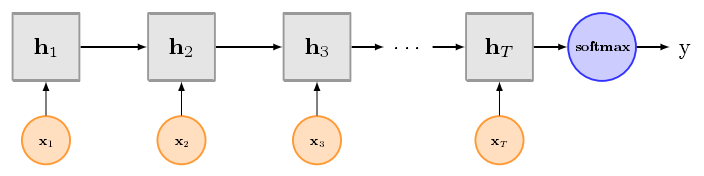
1]一维卷积：使用不同尺寸的kernel\_size来模拟语言模型中的N-Gram，提取句子中的信息。即TextCNN中的卷积用的是一维卷积，通过不同kernel\_size的滤波器获取不同宽度的视野。

[2]词向量：static的方式采用预训练的词向量，训练过程不更新词向量，本质就是迁移学习，主要用于数据量比较小的情况。not-static的方式是在训练过程中更新词向量。推荐的方式是not-static的fine-tunning方式，它是以预训练的词向量进行初始化，训练过程中调整词向量。在工程实践中，通常使用字嵌入的方式也能得到非常不错的效果，这样就避免了中文分词。

[3]最大池化：TextCNN中的池化保留的是Top-1最大信息，但是可能保留Top-K最大信息更有意义。比如，在情感分析场景中，“我觉得这个地方景色还不错，但是人也实在太多了”，这句话前半部分表达的情感是正向的，后半部分表达的情感是负向的，显然保留Top-K最大信息能够很好的捕获这类信息。

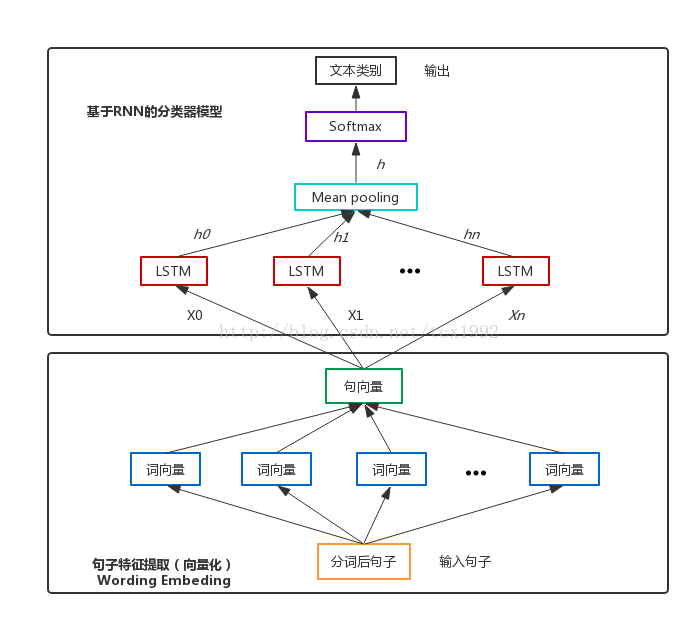
2、TextCNN原理

TextCNN擅长捕获更短的序列信息，但是TextRNN擅长捕获更长的序列信息。具体到文本分类任务中，BiLSTM从某种意义上可以理解为可以捕获变长且双向的N-Gram信息。

将CNN和RNN用在文本分类中都能取得显著的效果，但是有一个不错的地方就是可解释性不好，特别是去分析错误案例的时候，而注意力机制[Attention]能够很好的给出每个词对结果的贡献程度，已经成为Seq2Seq模型的标配，实际上文本分类也可以理解为一种特殊的Seq2Seq模型。因此，注意力机制的引入，可以在某种程度上提高深度学习文本分类模型的可解释性。

五、利用Text-RNN模型来进行文本分类

参考文献： https://blog.csdn.net/tcx1992/article/details/78194384

基于LSTM（Long-Short Term Memory，长短时记忆人工神经网络，RNN的一种）搭建一个文本意图分类的深度学习模型（基于Python3和Tensorflow1.2），其结构图如下：

六．RCNN

RCNN（Regions with CNN features）是将CNN方法应用到目标检测问题上的一个里程碑，由年轻有为的RBG大神提出，借助CNN良好的特征提取和分类性能，通过RegionProposal方法实现目标检测问题的转化。

        算法可以分为四步：

        1）候选区域选择

        Region Proposal是一类传统的区域提取方法，可以看作不同宽高的滑动窗口，通过窗口滑动获得潜在的目标图像，关于Proposal大家可以看下SelectiveSearch，一般Candidate选项为2k个即可，这里不再详述；

        根据Proposal提取的目标图像进行归一化，作为CNN的标准输入。

        2）CNN特征提取

        标准CNN过程，根据输入进行卷积/池化等操作，得到固定维度的输出；

        3）分类与边界回归

        实际包含两个子步骤，一是对上一步的输出向量进行分类（需要根据特征训练分类器）；二是通过边界回归（bounding-box regression) 得到精确的目标区域，由于实际目标会产生多个子区域，旨在对完成分类的前景目标进行精确的定位与合并，避免多个检出。