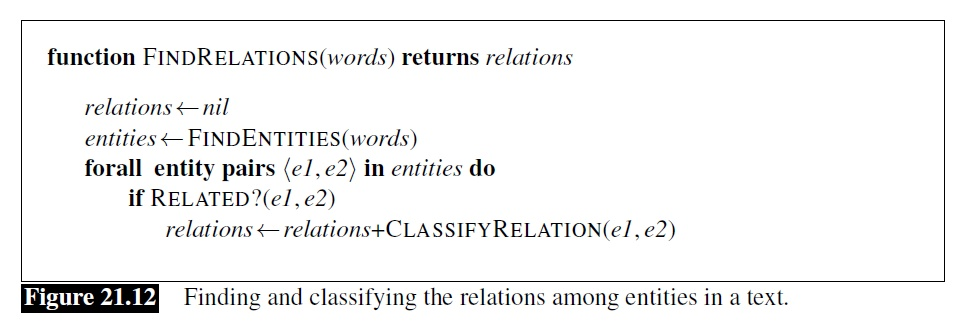
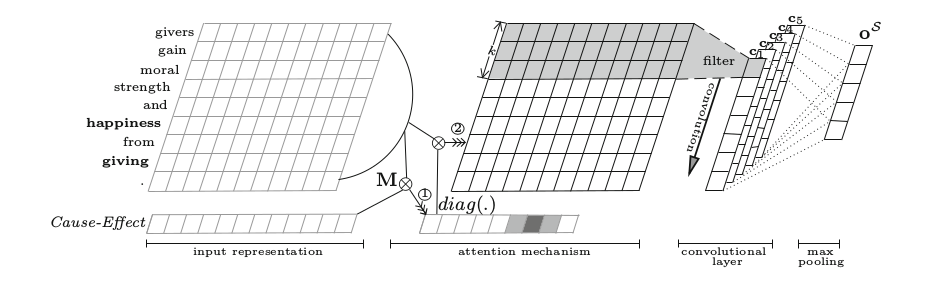
信息抽取预研：

信息抽取的思路是：给出一句话，抽取可能实体，对每对实体进行关系分类。

模型：

1）Relation+Classification+via+Target-Concerntrated+Attention+CNNs

模型结构：

这个模型大致分为四步：词向量表示、利用att转换词向量表示作为cov输入、

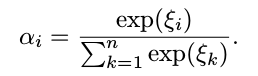
pooling、计算得分。其创新点在输入层，也就是卷积的输入时做了注意力机制。

第一步：输入词向量表示，包含了word embedding vector和position vector(，有两个position vector，对应两个实体词的相对位置)

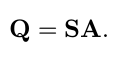
第二步:计算注意力权重，Ｍ是需要学出的,¨ξi表示第ｉ个词向量的att大小



输入层表示：１）分配每个词向量权重：



　　　　　２）转换为输入矩阵：其中Ａ是α的对角矩阵，Ｓ是词向量矩阵



第三步：

在卷积之后，进行max pooling,每个meature map进行max pooling后只有一个值，所以转换之后变为固定的一列向量Ｏ。

第四步：

计算得分并训练：其中Ｕ是降维矩阵需要训练，将Ｏ转化为ｒ（关系）长度



损失函数为：



ｒ－和不正确关系距离，r+是和正确关系的距离，其实就是ｍａｒｇｉｎ

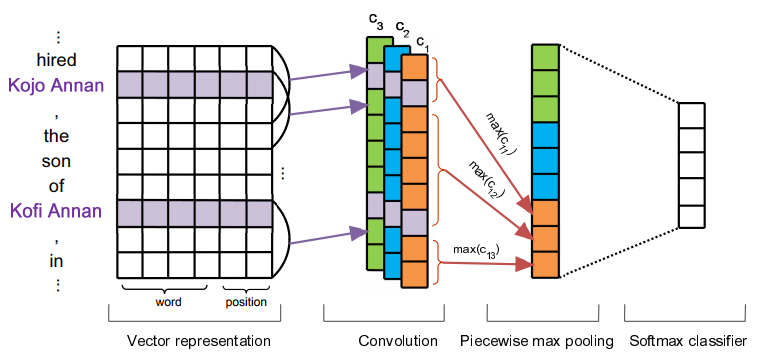
效果：数据集　SemEval 2010　Task 8 dataset , 87.3％

代码地址：https://github.com/htw2012/nlp\_relation\_extraction

2）Distant+Supervision+for+Relation+Extraction+via+Piecewise+Convolutional+Neural+Networks

这是利用远监督实现的信息抽取模型。（看完模型结构后不知道题目中的远程监督体现在哪里？？？感觉也可以用在数据量大的监督学习上）

模型结构：该文的创新点在于pooling,词向量表示和上一篇文章类似，卷积也是正常的卷积。在对每个卷积后的feature map进行pooling时，是根据实体的位置来进行pooling的，分成了３段。举例来说，下图橙色的feature map被两个实体（淡紫色）分成了上中下三段，对每一段进行了pooling操作



效果：数据集　http://iesl.cs.umass.edu/riedel/ecml/　效果７８．３％

代码地址：<https://github.com/smilelhh/ds_pcnns>

3）Relation+Classification+via+Multi-Level+Attention+CNNs

模型结构：论文的创新点在输入层和pooling层都引入注意力机制，相当与综合了上两篇文章。

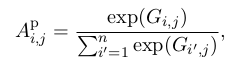
１）在输入层加入att机制，这里有两个att矩阵，每个关键词有一个att矩阵。对实体１有个Ａ１对角矩阵（att矩阵与上文类似），对实体２有个Ａ２对角矩阵。输入时两个对角矩阵累加和平均再跟输入矩阵相乘（为什么要有两个，既然是累加和平均，一个也可以进行解决了，可能这两个矩阵算出的方式不一样，还不太明白？？？）

２）对于卷积的话也是正常的卷积，对于pooling的话，引入了pooling att。在进行卷积后得到Ｒ\*，会有个计算pooling att的操作，如下面两行公式：

公式１）计算pooling att矩阵（这里的ＷＬ是如何来的？？？，其第二维度应该是和关系类别数相同）；公式２）用于怎样计算pooling，

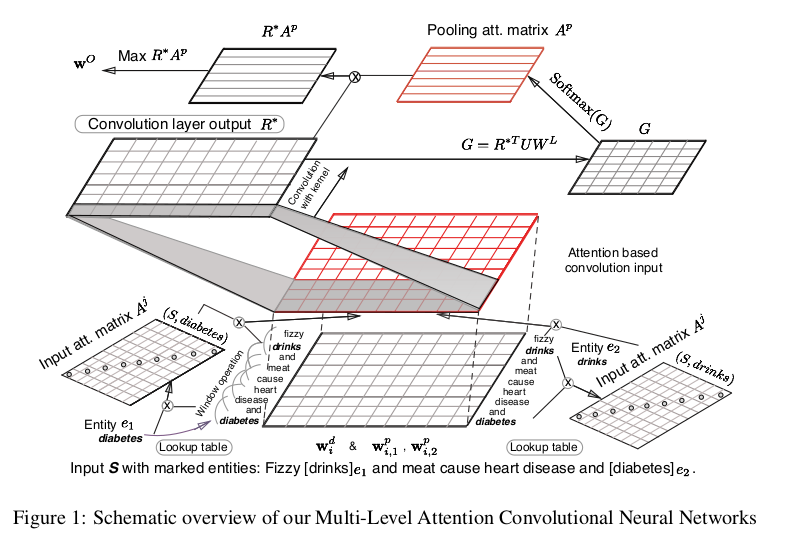
１）





２）





效果：数据集　SemEval-2010　Task 8 dataset，结果８７．５％

代码地址：

https://github.com/lawlietAi/relation-classification-via-attention-model

ｐｓ:看某一篇博客上说是效果最好的模型(参考https://www.jianshu.com/p/11821ce9905d)