

阅读理解相关论文

黄焱晖

2019 年 3 月 6 日

0.1 阅读论文: Zero-Shot Relation Extraction via Reading Comprehension

动机: 利用阅读理解的方法来完成关系抽取的任务。有以下三个优点: (1) 可以将关系抽取任务用最新的阅读理解模型来完成; (2) 对于训练集中未出现的关系, 在测试集中能够抽取出来

思路: 以往关系抽取的方法都是首先通过识别出语料中的待抽取关系的两个实体, 然后通过分类的方法, 得到两个实体之间的关系。作者的方法是将关系抽取看作是一个问答的过程, 对于三元组($x, \text{relation}, y$)将关系映射成问题 $q(x)$, 其中 $q(x)$ 的答案是 y , 例如下图所示:

关键点: 对于本篇论文来说, 关键的地方在于常规的阅读理解模型, 默认答案是存在于语料中的, 但是关系抽取转化为的阅读理解, 需要考虑的是

Relation	Question Template
$educated_at(x, y)$	Where did x graduate from?
	In which university did x study?
	What is x 's alma mater?
$occupation(x, y)$	What did x do for a living?
	What is x 's job?
	What is the profession of x ?
$spouse(x, y)$	Who is x 's spouse?
	Who did x marry?
	Who is x married to?

图 1: 关系映射成问题示意图

语料中是否存在答案。需要对现有的模型进行改进。

解决方案：作者利用了经典的阅读理解模型BiDAF模型（下一篇介绍的内容），并在最后的输出层进行了细微的改造，使这个模型能够处理没有答案的情况，BiDAF模型输出的是两个序列，长度等于阅读语料，这两个序列指示了答案的开始位置和答案的结束位置，这两个序列之间两两相乘取概率最大的值作为答案的序列，给每个序列增加一个参数，长度变为阅读语料长度加1，没有答案的概率等于两个序列最后一个参数相乘

总结：该论文的最大突破使用阅读理解的方式去完成关系抽取任务，关键点在于修改阅读理解模型，使之能够处理没有答案的情况，可以将这个方案用在知识补全上面。

0.2 阅读论文：BI-DIRECTIONAL ATTENTION FLOW FOR MACHINE COMPREHENSION

该论文发表在2017ICLR上，但是大体的思路和模型还是和二维匹配模型一致。

动机：以往的阅读理解attention模型都是单向的，但在该论文中，作者提出了一种双向attention的模型，能够更好地得到语料的表示。从而提升模型的效果

思路：论文得到阅读语料和问题的词向量和字符向量，并且将字符向量和词向量结合起来，得到阅读语料和问题的表示，对词两两相乘，得到一个二维相似度矩阵，从这个二维相似度矩阵得到两个角度的阅读语料的表示，并将这个表示喂入模型，得到输出，模型图如下所示：

总结：该论文最大的突破是从两个角度去得到阅读理解的表示，利用了双向的attention，使得模型的效果更好，下一步的目标是改动模型，使模型能够处理没有答案的情况。

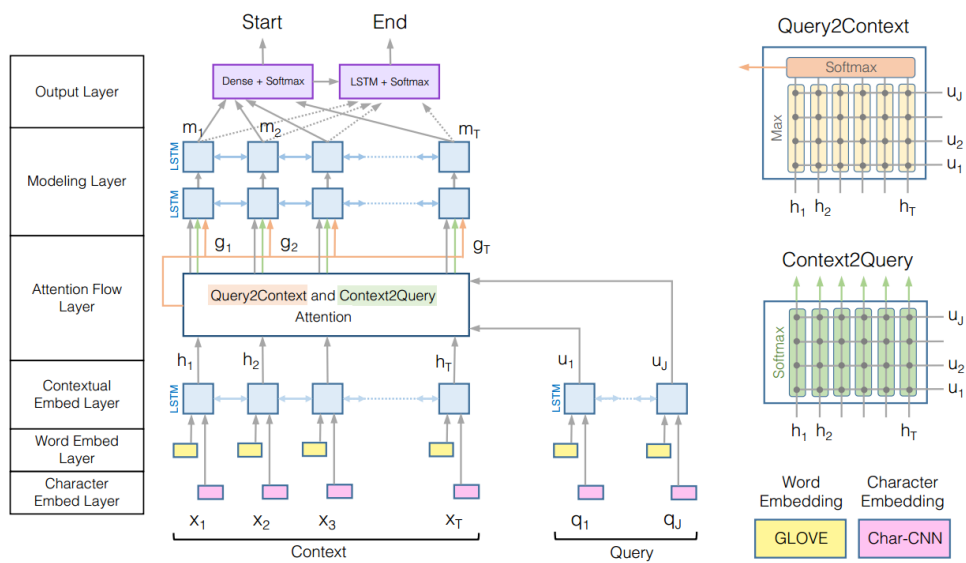


图 2: BiDAF模型

0.3 运行已有的BiDAF模型，测试结果为EM值为64.5%，F1值为75%，相比较原文中的结果下降了2%左右。pytorch模型撰写注释