|  |  |
| --- | --- |
| Inline Text Wrapping Picture |  |

硕士研究生学位论文阶段报告

学 号: 2018140521

姓 名: 施 晴

学 院: 计算机学院

专业(领域): 计算机技术

研究方向: 大数据技术与智能信息处理

导师姓名: 于艳华

北京邮电大学

2020年9月24日

|  |  |  |  |
| --- | --- | --- | --- |
| 论文题目 | 基于深度学习的事件抽取系统设计与实现 | | |
| 论文类型 | 应用研究 | 选题来源 | 其他 |
| 开题日期 | 2019-11-01 | 是否开题题目 | 是 |
| 论文开始日期 | 2019-11-01 | 报告日期 | 2020-09-24 |
| 报告地点 | 教三楼912 | 报告时间 | 上午8:30 |
| **一、研究内容简介**（包括：选题背景、研究内容、关键技术、论文计划、论文进度及目标，不少于：5000）  **1、选题背景**  近年来，随着互联网信息技术的高速发展和科技的进步，人类社会迈入了大数据时代。在网络平台上产生大量的无结构化文本信息，为了精准、快速地从海量的无结构化的数据中筛选、获取可用信息，信息抽取研究已经成为自然语言处理领域的热点分支。信息抽取是指从非结构化的数据抽取出用户需要的信息，并进行处理，形成结构化的数据供用户浏览，从而将互联网变成一个巨大的知识库，以便于用户能够轻松的找到需要的信息或者为自然语言处理领域的其他任务或应用提供数据支持。  信息抽取其主要包括命名实体识别、关系抽取和事件抽取这三个部分。命名实体识别的目标是抽取出文本中的实体信息，比如人名、地点、组织结构等。关系抽取的目标是抽取出文本中实体与实体之间存在的二元语义关系。事件抽取的目标由自动内容抽取测评会议（Automatic Content Extraction，ACE）定义[1]，事件抽取定义为从非结构化的自然语言文本中抽取特定的事件结构信息，其主要包含两个核心步骤：事件触发词抽取和事件元素的抽取[2]。其中事件触发词的抽取的目标是识别出触发规定事件类型的词，即最能体现事件发生的词，并对其进行分类。事件元素抽取的目标是抽取出事件发生的时间、地点、参与时间的人物等关键信息元素。事件触发词和事件元素两者可以后对一个事件的完整描述。  事件抽取有着重要的理论研究和实际应用价值。在理论研究方面，事件抽取的研究对于自动问答[3]、信息检索[4]、阅读理解[5]、自动文摘[6-7]等自然语言处理任务的进展有着非常大的推动作用。在实际应用方面，事件抽取已经在商业化的搜索系统以及舆情分析等方面有着广泛的应用。  近些年来，深度学习的发展使得基于神经网络的方法得到广泛应用。神经网络模型可以避免传统机器学习对于特征工程的依赖，除了能够获取上下文语义信息外，还能通过自动组合和抽取来获取深层抽象特征。目前，神经网络模型在自然语言处理众多任务中都已经取得突破性进展。在事件抽取领域也有研究者开始采用基于深度学习的神经网络模型。  目前的深度学习方面主要研究方法有：Chen等2015年[8]提出基于DMCNN(动态多池化)的模型用于检测单句子中包含的多事件，将句子动态分割成两部分，对每部分分别预测触发词，从而提高了单句子多事件的准确率值。Feng等2016年[9]提出利用BiLSTM双向长短期记忆网络进行事件识别任务。Liu J等在2018年[10]提出了一种新颖的多语言方法-被称为门控多语言注意(GMLATT)框架，用于解决数据稀缺和单语歧义的问题，通过上下文关注机制利用多语言数据的一致性信息解决数据稀疏问题，利用门控式跨语言模型解决词语的歧义性问题。  在联合抽取模型方面，Nguyen等2016年[11]使用循环神经网络来抽取句子语义特征，并对句子的依存树信息进行编码得到句子的结构信息从而进行事件识别和论元识别的联合抽取。Liu X等2018年[12]通过使用基于注意力机制的图卷积模型对句子中的语法树信息进行编码并输入到模型中，进行事件识别和论元识别的联合抽取。Nguyen等2018年[13]提出一种基于语法依赖树的图卷积神经网络进行事件检测。并提出了一种新的基于实体提及的聚合卷积向量的池方法。实验证明了该方法的优异性。然而这些方法目前在图注意力的计算方式上都没有更加全面的考虑，模型的效果还有待提高。  基于上述总结，本文提出了基于图注意力神经网络模型，通过在图模型中多方面考虑注意力的影响因素来提高事件抽取的效果，并进一步通过数据增强的办法来提高在ACE2005数据集上事件抽取的性能。  **2、研究内容**  本文的研究内容是利用深度学习神经网络模型提高事件抽取的性能。对句子中事件触发词的识别和分类可以建模成对每个单词的分类任务进行研究。如下图1所示的英文语料中的句子，这句话的意思是“数十名伊拉克平民和士兵在目击者所说的美国炮兵部队的攻击中丧生”，对于事件识别任务我们需要识别出一下内容：首先killed触发了Die事件，然后barrage触发了attack事件，最后其他词没有触发事件。对于论元角色识别中，识别触发词-实体对之间的关系，如需要识别出killed的触发的死亡事件的受害者、地点等论元。    图1 事件抽取示例  本文主要从事件抽取神经网络模型和半监督学习进行数据增强以及系统可视化展示这3个方面来进行事件抽取相关研究。  **（1）针对事件抽取算法的研究**  过去基于事件抽取的神经网络模型大部分基于序列结构模型如循环神经网络RNN/LSTM等[11][14-15]，对于长文本数据，还可能会丢失长距离依赖关系，造成重要语义丢失的问题。为解决上述问题，在我们的模型中主要采用以下几个步骤解决，首先通过使用斯坦福的自然语言处理工具[16]进行句子语法依赖弧的标注，其次将序列结构数据通过语法依赖弧转为图结构数据，通过语法依赖弧建立远距离词之间的语义联系，如图1中的killed（死亡）和barrage（攻击）之间的相对距离在序列结构为6次跳转，但是通过语法弧标注之后killed可以通过nmod-arc（复合名词修饰）弧跳转到witnesses，然后通过acl-arc弧跳转到*called*词，再通过xcomp-arc（x从句补语）弧跳转到barrage词，即可通过3次跳转就可以实现信息流动，提高信息的流动。最终使用图神经网络对图结构数据中的节点执行分类任务，完成触发词的识别和分类。  同时针对现存图神经网络的缺点如缺少注意力机制[13]、注意力考虑不全面和特征选取不丰富[17-18]等创新性的提出了一种基于多头图注意力神经网络模型解决事件抽取任务。通过一定的权重系数将图结构中邻居信息融合到中心节点中，使中心节点拥有更加丰富的语义信息。这个权重参数不仅考虑到中心节点在整个句子中的语义权重信息，还考虑到了中心节点和邻居节点的相关性权重信息来共同决定最终的权重系数。不仅如此，我们还通过设置级联多头图注意卷积来获取更加丰富的语义信息，同时为了让原语义信息在图注意网络中的信息不流失，通过使用残差网络[19]的形式来进行使得原语义信息无损的向前传播。该机制较之前的图神经网络，在利用邻居节点的特征聚合更新中心节点的过程中，更加全面的考虑了中心节点对其邻居节点赋予的权重参数问题，大大提高了事件检测的精确率，召回率以及F1值。  **（2）针对数据增强算法的研究**  目前广泛采用的弱监督方法充分利用了大量的原始数据，特别是一些具体的信息提取工作，探索了弱监督自动标注事件抽取训练数据的方法，这种弱监督方法可以有效地推广到实际的事件抽取应用中，而不需要大量的劳动。  弱监督方法虽然取得了很好的结果，但仍然存在一些严重的问题：如会受到数据中不可避免的噪声的影响，并且现有的弱监督ED模型采用复杂的预定义规则和不完全的知识库[20],[21]来自动获取数据，导致数据的自动标注覆盖率低、主题偏差大等问题。  基于上述问题为了构建一个覆盖范围更广的大规模数据集，减少主题偏差、避免采用复杂的预定义规则和繁重的语义组件分析工具包，采用了一个简单的基于触发器的潜在实例发现策略。我们采用如下假设：如果一个给定的单词作为已知事件实例中的触发词，那么提到这个单词的所有实例也可能表示了一个事件。与复杂的规则相比，该策略在单词、触发器和事件类型之间的相关性方面限制较少。因此，该策略可以在不需要任何人工设计的情况下，获得覆盖更多主题和实例的候选集[22]。  进一步为了过滤一些噪声数据，训练生成对抗网络[23]，它不仅可以从候选集中提取出信息实例，将自动标注的数据集分为可靠集和不可靠集，对于而且可以提高ED模型在诸如远程监控等噪声场景下的性能。  在实际的本文研究的事件抽取领域，由于ACE2005数据集是可用于事件抽取任务的已经标注好的数据集，并且有成熟的评价体系来验证模型的有效性，因此本文主要基于ACE2005数据集进行事件抽取任务的研究。ACE2005数据集主要来源于广播新闻，广播对话，博客等平台。由于该数据集的数据量小，并且存在类型不均衡等问题，会导致容易导致神经网络训练的过拟合，模型的泛化能力不够，而在测试数据上得到的效果不好。  针对上述问题，我们可以通过使用上述方法来扩展训练数据，以此来提高在该数据集中测试数据集上的效果。主要通过将同种类型的、无标注的、非结构化的文本语料如纽约时报的新闻语料等，首先利用自定义的触发词发现策略进行自动标注，生成不可信任的标注数据集。然后利用生成对抗网络进行数据集的筛选，过滤噪声大的数据，最后将筛选出来的相对信任标注数据作为增强的训练数据，并与原ACE2005的训练数据集共同输入到多头图注意神经网络中进行模型的训练。通过扩充训练数据的数据量来提高模型参数的泛化能力，实现在ACE2005测试数据集上的效果提升。  **（3）针对事件抽取系统的研究**  针对深度学习神经网络可解释性差的问题，本文实现了可视化的事件抽取系统。对于输入的序列结构的句子，首先可以利用斯坦福自然语言处理工具将转换之后的图结构数据进行展示，其次还可以在多头图注意力神经网络中节点之间权重分配进行可视化的展示，从而增强用户对模型的理解，最后对事件抽取的结果进行展示。   1. **关键技术**   该模块主要针对研究内容的三个部分内容采用的关键技术进行描述。  **（1）多头图注意力神经网络**  在事件抽取算法中主要技术点为多头图注意力神经网络抽取模型。该模型用于解决事件检测问题。权重图神经网络算法主要遵从以下几步：  1）生成具有上下文语义信息的特征向量：将序列结构的句子中每个单词的字符特征、词向量特征、词性特征、实体特征、位置特征输入到编码层中得到相应的特征向量，并将这些特征向量合并，输入到双向循环神经网络中，得到具有上下文的语义信息的特征向量。  2) 数据结构转换：利用斯坦福自然语言处理工具对句子进行语法依赖树的标注，将语料库中序列结构数据转为图结构数据，序列结构的每个单词即为图结构的每个节点。  3）图注意力网络模型计算过程：此过程是本模型的重点技术点，将2）过程生成的节点的特征输入到图模型中，利用权重图注意力神经网络的计算公式，通过将邻居节点的特征进行聚合来更新中心节点的特征。  4）节点分类：将图模型的每个节点的特征输入到分类器中进行分类。  5）输出：根据节点分类结果输出每个节点的触发词类别。  该模型的整体架构图如图2所示：    图2权重图注意力神经网络模型  该权重图神经网络主要包括3个模块，词向量表达模块，权重图神经网络计算模块，触发词分类模块，下面将详细描述各个模块的具体细节。  **词向量表达模块-获取具有上下文的词语语义信息**  在单词编码模块，句子中的每个单词被转换为特征向量，其中包括：   * 的词义特征：我们使用BERT预训练的向量[24]进行单词词义的特征表示。 * 的字符特征：为了得到更加丰富的单词语义特征，如时态、词缀等，我们通过将字符进行编码，并输入到不同卷积核的卷积神经网络中，使用最大池进行抽取单词中特征最大的部分，作为单词向量的一部分特征表示。 * 的词性特征：通过查找随机初始化的词性矩阵进行特征表示。 * 的位置特征：通过使用句子中的每个单词的绝对位置进行编码，并输入到随机初始化的编码层中进行位置向量的特征表示。 * 的实体特征：我们使用BIO标注格式对识别类型标签，我们使用语料库标注的所有的实体标签，这可能会产生重复的覆盖，我们求和所有可能的实体类型特征。 * 经过上述操作把每个单词转换为特征向量，把输入句子转为序列结构的特征向量，在输入图卷积网络之前，通过将特征向量输入到BiLSTM神经网络层，使每个词向量都具有上下文语义特征，并将具有上下文语义特征的向量用于下一阶段多头图注意力网络的输入。   **权重图神经网络计算模块（重点技术）-根据邻居节点特征更新中心节点特征**  经过BiLSTM神经网络，句子向量转换成。对于句子来说，为得到更加抽象的特征，将BiLSTM网络模型的输出再输入线性层中，如下：   |  |  | | --- | --- | |  | (1) |   其中线性层的权重矩阵参数，偏置项参数  通过使用依赖树的边将句子的序列结构转为无向图结构,是图中的节点，是边集合，在中每个是每个单词在图中的节点表示，每个边代表了存在一条有向语法依赖边从指向，为了使信息沿着依赖边反向传播，当存在从到的有向边时，我们也添加了从到的有向边，即，对于所有的节点，我们添加了自环边即到的边，即,对于所有的。为了简便我们不考虑不同类型的依赖边的权重大小，都将其视为同样的权重1，但是为了说明边的不同指向关系，我们设置3种类型的边,包含了自环的边[25]，如下公式所示。   |  |  | | --- | --- | | *前向:*  *反向:*  *自环:*  *from to,*  *from to,*  *from to,* | (2) |   对于图中的任何一个节点来说，需要利用的邻居节点来更新，对于不同的邻居节点，对其赋予的权重参数不仅需要考虑到当前节点在句子中的权重，其计算方式如下：   |  |  | | --- | --- | |  | (3) |   其中线性层的权重矩阵参数，偏置项参数，为非线性激活函数，为第一个公式的线性神经网络的输出向量。  还需要考虑到当前节点和其邻居节点之间的相关系数权重来共同决定利用邻居节点的权重值。为了得到各个节点更加抽象的特征，至少需要经过一次线性变换得到更加高维的特征。为了计算简便，我们设置共享参数的权重矩阵 这个矩阵会应用到每个节点中，其中每个节点和邻居节点计算相关系数公式如下，其中的维度变化为：   |  |  | | --- | --- | |  | (4) |   其中 为权重矩阵，为非线性激活函数，为第一个公式的线性神经网络的输出向量。  由于上述公式在实际操作中需要将两个节点的维度进行并列操作，这会占用大量资源，在实际优化中，我们将上述公式修改如下,即通过加法运算来计算节点和邻居节点的相关系数：   |  |  | | --- | --- | |  | (5) |   其中为第一个公式的线性神经网络的输出向量。实验证明两个达到的效果是相同的，但是下面的公式可以占用更少的计算资源。  为了规格化节点在不同类型边对应的每个邻居节点之间的权重参数，对于节点的每种类型边的每个邻居节点,其归一化的注意力权重如下所示：   |  |  | | --- | --- | |  | (6) |   综合上述公式，最终节点赋予给的注意力分数为：   |  |  | | --- | --- | |  | (7) |   得到了注意力分数，就会得到节点的更新公式   |  |  | | --- | --- | |  | (8) |   其中 是非线性激活函数,上述公式计算的是经过了一次图神经网络，即单独一个头的结果。  为了得到更加丰富的语义表示，我们采用级联的多头图注意网络组织方式，如图3所示。与之前的多头组织方式不同[26][27]。在本组织方式中，下一个头的输入除了包括前一个头的输出之外还包括了前一个头的输入，通过这种方式能够得到更加丰富的语义表示，我们定义为第k个头的输入，为第k-1个头的输出，公式表示如下：   |  |  | | --- | --- | |  | (9) |     图3 模型的多头组织方式，下一个Head的输入不仅包括前一个Head的输入，还包括前一个Head的输出，从而得到更加丰富的语义表示。  同时，我们初始化，则第k个头的输出表示为：   |  |  | | --- | --- | |  | (10) |   其中权重矩阵 是非线性激活函数，最后将多个头的输出结果进行联合：   |  |  | | --- | --- | |  | (11) |   其中||代表连接操作。为了进一步进行信息的聚合，我们再次采用图注意力网络进行信息操作，得到最终的节点更新结果。   |  |  | | --- | --- | |  | (12) |   其中权重矩阵 是非线性激活函数,为根据向量计算得到的注意力分数。  为了防止在图注意网络中原语义信息节点的丢失，采用了类似残差网络的结构,公式表示如下：   |  |  | | --- | --- | |  | (13) |   其中为上一个公式的输出向量，为LSTM神经网络的输出向量。  此时完成了一层多头图注意力网络模型的计算，即每个节点都获得了其一阶邻居节点的语义信息，为了使节点语义更加丰富，我们考虑加入其阶邻居节点的语义，即使用层多头图注意网络模型，这样每个节点就会有阶邻居的节点的语义信息，但是有时候两个单词之间的距离可能会小于，因此我们采用高速神经网络单元[28]来防止信息的过渡传播，表示如下：   |  |  | | --- | --- | |  | (14) | |  | (15) |   其中权重矩阵，偏置项,代表了BERT词向量，，是元素级别的乘法操作，是非线性激活函数。  最终第二层权重神经网络的输入应该为而不是。然而并不是多头图注意网络模型的层数越多越好，我们发现加入第三层多头图注意力网络模型，模型的性能反而下降，这可能是因为信息流动太多，导致每个节点的语义都会彼此接近，不利于后面节点的分类。因此本模型采用两层。  **触发词分类模块**  我们将多头图神经网络的输出输入到线性分类器中进行分类，我们发现将多头图神经网络的输出和BERT词向量相加，可以提高在触发词分类阶段的性能，因此最终分类器表示如下：   |  |  | | --- | --- | |  | (16) |   其中权重矩阵,偏置项为 层的WGAT模型的输出向量，n是使用BIO标注的触发词类别标签的数量。  **（2）生成对抗网络**  在半监督场景中，利用生成对抗网络进行数据增强的过程主要包含以下几个步骤：  1）设计一个判别器（判断给定实例的触发词和事件是否正确）和一个生成器（从原始数据中选择最易混淆的实例来欺骗判别器）。  2）预训练判别器：利用少量的标记数据对编码器和判别器进行预训练，使其在一定程度上获得检测事件触发器和识别事件类型的能力。  3）构造候选集：基于我们的实例发现策略，以标记数据中的触发词作为启发式种子，构造一个大规模的潜在候选集。并使用预先训练的编码器和判别器，为候选集合中的所有实例自动标记触发器和事件类型，以建立有噪声的大规模数据。  3）对抗训练：以少量标记数据为可靠集R，以大量自动标号数据为不可靠集U，对编码器、判别器、生成器进行优化，共同进行对抗训练。在训练过程中，生成器可以提供大量的潜在噪声数据来增强判别器，判别器可以影响生成器选择那些迷惑性更大的数据。当生成器和判别器经过一定的训练达到平衡时，判别器可以提高对噪声的抵抗能力，更好地对事件进行分类，生成器可以有效地为判别器选择信息实例。最后将平衡时由生成器从不可靠集U中挑选的并被判别器正确标记的实例从U调整到R，并保存下来作为增强数据使用。  4）重新测试：将增强的训练数据和ACE2005测试数据重新输入到前面的多头图注意力神经网络中，重新进行训练，模型收敛后得到测试数据上的效果，查看此时在训练集上的效果是否有提升。  生成对抗网络的整体框架图如图4所示：    图4 生成对抗网络框架图  其主要分为3部分，第一部分为Encoder编码层，Generator生成器以及Discriminator判别器组成：  编码器用于将句子中词语编码成特征向量。  判别器检测数据集中对每个实例给定事件触发器和事件类型匹配的概率值，当给定一个有噪声的实例时，鉴别器也要抵抗噪声，并明确指出没有触发器和事件，即匹配值为0。  生成器用于从不可靠的数据集U中选择实例，以尽可能地混淆判别器。下面我们对每部分技术点进行详细阐述。  对于每个实例我们可显式地表示其标记其触发词和事件类型。对于每个实例在生成对抗网络中被认为是不可信的，即存在一定的错误标记的概率。因此我们设计了一个判别器来判断给定的实例能否代表其标记的事件类型，其目的是使条件概率和最大化。生成器的目标是训练以从中选择最具有迷惑性的实例提供给判别器，即通过选择实例。训练过程是一个对抗性的min-max博弈。   |  |  | | --- | --- | |  | (17) | |  | (18) |   其中是可靠数据分布，生成器根据概率分布从不可靠的数据中抽取对抗性实例，虽然和是冲突的，但是不可靠数据对和都有副作用。当生成器和判别器经过充分训练达到平衡时，生成器倾向于选择那些概率比更高的数据实例，而判别器由于增强了对噪音的抵抗能力，能更好地分类事件，同时也可以更好的从不可靠数据中筛选出来较为可靠的数据。  **Encoder编码器**  在单词编码模块，句子中的每个单词被转换为特征向量，其中包括：   * 的词义特征：我们使用BERT预训练的向量进行单词词义的特征表示。 * 的词性特征：通过查找随机初始化的词性矩阵进行特征表示。 * 的实体特征：我们使用BIO标注格式对识别类型标签，我们使用语料库标注的所有的实体标签，这可能会产生重复的覆盖，我们求和所有可能的实体类型特征。 * 经过上述操作把每个单词转换为特征向量，把输入句子转为序列结构的特征向量。   **Generator生成器**  生成器的目标是从自动标注的数据集中选择最有迷惑性的实例来欺骗判别器。设计生成器以优化概率分布为目标来选择实例。生成器计算所有R中的混淆分数，以评估他们的迷惑性，并进一步计算混淆概率,如下所示：   |  |  | | --- | --- | |  | (19) | |  | (20) |   其中是实例x的编码表示，和是分离超平面的参数。  我们认为由判别器计算出来的实例的分数越高，该实例就越具有迷惑性，因为它们更容易欺骗判别器做出错误的决定。因此我们期望优化的生成器可以更加关注那些最具有迷惑性的实例。因此，给定一个实例及其不可靠的标记触发器和事件类型，损失函数表示为如下：   |  |  | | --- | --- | |  | (21) |   其中由判别器计算。在优化生成器时，我们将计算组件作为参数进行更新，损失函数对应于。  在标记为NA的中可能存在一些实例，这些实例总是被错误地预测到其他一些事件中。因此我们使用所有可行事件的平均得分来代替公式中的。   |  |  | | --- | --- | |  | (22) |   代表了事件类型的集合。  **Discriminator判别器**  给定一个实例和它标注的触发词和事件类型,判别器的任务是判断给定的实例和标注的触发词和事件类型是否符合。对实例编码为之后，实现如下的判别器。   |  |  | | --- | --- | |  | (23) | |  | (24) |   其中是事件类型的编码表示。  一个最优化的判别器会给R中给那些实例分配高分，同时不信任那些在U中的实例标签，即分配低分，因此损失函数设计为   |  |  | | --- | --- | |  | (25) |   在对判别器进行优化时，由于R 和U中可能存在大量的实例，直接作为参数计算和是非常耗时的，并且频繁地遍历R和U的整个数据集也变得困难。为了提高训练效率，对R和U的子集进行采样以逼近基本概率分布，并将一个新的损失函数形式化为如下所示，对此函数进行优化。   |  |  | | --- | --- | |  | (26) | |  | (27) |   其中和是从R和U采样的子集，是公式（7）的近似。   |  |  | | --- | --- | |  | (28) |   是一个超参数，它控制概率分布的锐度，以避免权重集中在某些特定实例上。最终的全局优化函数为：   |  |  | | --- | --- | |  | (29) |   其中，是权衡系数。在对抗性训练中的和是交替优化的，并且我们在的优化过程中加入了对的学习，这样避免了对的额外调整。  **（3）基于深度学习的事件抽取系统研究**    图5 事件抽取系统框架图  实现了上述算法，本文搭建了基于深度学习的事件抽取系统，如图5所示，对深度学习模型运行过程中的注意力参数和运行结果进行可视化的展示，用户可以输入一句话，系统加载预训练的模型进行事件抽取，并最终在页面上显示该句子的语法弧，该句子中触发词分类的结果，以及事件类型分类的结果，和注意力矩阵的可视化展示。  **4、论文计划**  表1 论文计划   |  |  |  | | --- | --- | --- | | 序号 | 时间 | 研究内容 | | 1 | 2019.9－2019.12 | 进行前期调研，确定研究方向与目标, 制订实施方案，论文开题 | | 2 | 2020.1－2020.3 | 学习Pytorch框架，筛选出效果比较好的论文并对相关模型进行复现 | | 3 | 2020.4－2020.7 | 事件抽取模型设计及调参，并验证效果 | | 4 | 2020.8－2020.11 | 生成对抗网络模型设计及调参，并验证效果 | | 5 | 2020.12－2021.1 | 优化模型，系统设计并搭建可视化系统，并发表专利 | | 6 | 2021.2－2021.3 | 系统测试和论文初稿撰写 | | 7 | 2021.4－2021.6 | 完成论文撰写，完成答辩材料的准备 |   **5、论文进度及目标**  截止目前为止，前3项已经顺利完成，并完成了一篇专利的撰写，第4项工作进行了50%，完成了开题计划中应该完成的工作。  本文旨在使用深度学习神经网络的技术和传统的自然语言处理的方法相结合，完成事件抽取算法实现和事件抽取系统的设计开发。针对传统序列结构的神经网络中存在长依赖消失的问题，本文提出使用图神经网络来实现事件抽取算法实现。针对图神经网络中中心节点对邻居节点的注意力权重计算考虑不全面、特征不丰富的问题，本文提出了基于权重图神经网络技术，并且使用多头图神经网络保证特征的足够丰富，同时也防止了原语义消失的问题。同时由于BERT在众多自然语言处理任务中都表现出很强的效果，本文的词向量选择的也是Bert词向量。针对事件抽取领域ACE2005数据集存在的类型不均匀，数据量少的问题，文本提出使用触发词发现策略生成大规模不可靠数据集，并利用生成对抗网络进行数据筛选，并作为数据增强数据使用，提高多头图注意力网络的模型的泛化能力，从而提高在ACE2005测试集上效果的提升。结合以上两种算法，本文设计实现了基于深度学习的事件抽取系统，主要对事件抽取的注意力权重进行可视化和对事件抽取的结果进行可视化的展示。  论文整体架构及内容如下：  第一章为绪论，主要介绍本论文的研究背景及意义，并通过阅读大量文献归纳国内外研究发展现状，进而提出本论文的研究内容和和组织架构。  第二章为详细介绍事件抽取算法，包括不同的事件抽取任务和算法，评价指标等。  第三章为相关技术介绍，主要介绍本论文需要用到的相关技术，包括基于Pytorch的组件研究、BERT模型的研究、事件抽取算法的研究、生成对抗网络算法的研究。  第四章为基于BERT的多头图注意力神经网络模型研究，以及实验设置和模型效果展示。  第五章为生成对抗网络算法研究，以及实验设置和模型效果展示。  第六章为基于深度学习的事件抽取系统，包括系统需求分析，数据展示方式，各模块设计和系统展示。  第七章为总结与展望，客观地总结在论文的写作、算法实现和系统设计方面主要做的工作，并给出论文的写作、算法实现和系统设计和测试环节中发现的不足之处，并对以后的改进进行展望。  目前，本文已完成对基于深度学习的事件抽取算法的调研，完成了多头图注意力模型搭建和调优，实验结果分析和系统设计部分，完成了生成对抗网络的设计实现工作，还在进一步的调优和效果验证阶段。  接下来的时间里，将进一步优化生成对抗网络和完成事件抽取系统的实现工作；预计在2021年3月前完成论文初稿。并广泛征求老师和同学的建议，对论文进行修改，使论文的整体结构更加合理，论述更加专业化，在2021年5月前完成论文的修改工作，产出最终的毕业论文，并完成毕业答辩各项准备工作。 | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **二、论文进展情况**（包括：报告工作计划、实际进展情况）  **1、工作计划**  本论文的主要工作包括：研究多头图注意力网络模型在事件抽取领域的应用和生成对抗网络在半监督场景下进行数据增强的应用。针对传统序列结构神经网络存在长依赖消失的问题，本文提出了使用图神经网络进行解决。针对常规的图神经网络注意力机制考虑不全面、特征考虑不丰富等问题，本文提出基于权重图神经网络，多方面考虑图结构数据中中心节点对邻居节点赋予的权重问题，并且为了获取丰富的特征采用多头图注意力网络，进行模型的设计。针对事件抽取领域训练数据量少，类型不均匀等问题，本文提出使用触发词发现策略，构造大规模不可靠候选集，并利用生成对抗网络进行筛选，最终选取较为可靠的数据集作为增强数据使用，并重新输入到多头图注意力神经网络中重新训练，通过提高模型泛化能力，提高在ACE2005测试数据集上的效果。融合以上两种算法，本文实现了基于深度学习的事件抽取系统，系统可以接受用户输入的句子，并对神经网络的注意力参数进行可视化展示，对运行结果包括句子的触发词分类结果以及事件类型预测的结果等进行展示。具体工作计划和安排如表2所示：  表2具体工作计划和安排   |  |  |  | | --- | --- | --- | | **时间** | **工作内容** | **预期成果** | | 2019.09－2019.12 | 进行前期调研，确定研究方向与目标, 制订实施方案，论文开题 | 确定研究方向，了解现有方法的不足，初步确定解决方案 | | 2020.01－2020.03 | 学习Pytorch框架，筛选出效果比较好的论文并对相关模型进行复现 | 复现模型，验证效果，发现不足 | | 2020.04－2020.07 | 事件抽取模型设计及调参，并验证效果，并发表专利 | 完成事件抽取模型相关实验，并发表专利 | | 2020.08－2020.11 | 生成对抗网络模型设计及调参，并验证效果 | 完成生成对抗网络模型相关实验 | | 2020.12－2021.01 | 优化模型，系统设计并搭建可视化系统， | Django搭建系统，并发表一篇专利 | | 2021.02－2021.03 | 系统测试和论文初稿撰写 | 完成系统测试和论文初稿撰写 | | 2021.04－2021.06 | 完成论文撰写，完成答辩材料的准备 | 完成毕业论文撰写 |   **2、实际进展情况**  目前，实际进展情况如表3所示：  表3 实际进展情况   |  |  |  | | --- | --- | --- | | **时间** | **工作成果** | **完成情况** | | 2019.09－2019.12 | 提出并确定研究方向 | **已完成** | | 2020.01－2020.03 | 复现模型，验证效果，发现不足 | **已完成** | | 2020.04－2020.07 | 完成事件抽取模型相关实验，并发表一篇专利 | **已完成** | | 2020.08－2020.11 | 完成生成对抗网络模型相关实验 | 部分完成 | | 2020.12－2021.01 | Django搭建系统，完成系统测试 | 未完成 | | 2021.02－2021.03 | 论文初稿撰写 | 未完成 | | 2021.04－2021.06 | 完成毕业论文撰写 | 未完成 | |
| **三、工作成果**（目前已完成学位论文工作的内容、取得的阶段性成果和主要创新点。包括所完成的理论和实验研究以及所获得的结论；已发表的与学位论文相关的学术论文等,字数不少于：1500)  **1、已完成学位论文工作的内容**  本文旨在使用深度学习和传统自然语言处理任务相结合的方法进行事件抽取任务的研究。  在技术研究方面：本文对BERT模型，图注意神经网络，生成对抗网络等进行深入研究，提出了基于BERT和多头图注意力神经网络模型，并实现了基于Pytorch的模型搭建。  在研究内容方面：为了解决传统序列结构长依赖消失的问题，本文提出使用图神经网络进行事件抽取。为解决图神经网络中注意力考虑不全面，特征不丰富等问题，本文提出在图神经网络中计算权重注意力，并使用多头图注意力网络，提高特征的丰富性。同时针对训练数据量少，类型不均匀等问题提出使用触发词发现策略构建大规模数据候选集，并利用生成对抗网络进行数据筛选，将筛选出的数据作为增强数据使用，与原有的ACE2005训练数据集合并，重新输入到多头图注意力神经网络中，通过提高模型的泛化能力，来提高模型在ACE2005测试数据集上的预测效果。  **已完成的工作如下：**   1. 环境搭建，数据处理如下表所示   表4环境配置和选取的数据集   |  |  | | --- | --- | | 环境 | 版本号 | | Pytorch | 1. 0 | | CUDA | 9.0 | | GPU | 1080Ti，11G | | 数据集 | 语言 | | ACE2005 | 英文 | | NewYorkNews | 英文 |  1. 完成BERT算法、图注意神经网络、生成对抗网络的研究 2. 完成基于BERT的多头图注意力神经网络模型的搭建和训练 3. 完成生成对抗网络的预训练工作。   **2、取得的阶段性成果**  **（1）数据处理和评估方法研究**  **①数据集描述和处理：**  本文采用在事件抽取领域广泛使用的ACE2005数据集，该数据集包含完整的英语、阿拉伯语和汉语训练数据，用于2005年自动内容提取(ACE)技术评估。语料库由多种类型的数据组成包括实体、关系和事件，这些数据由语言数据联盟(LDC)标注，数据标注以xml文件的格式进行存储，源数据对训练集、测试集等是没有划分的，默认采用Li Q [29]的数据集划分方式，即以40篇新闻专线文章作为测试集，从不同类型的文章中随机挑选30篇文档作为开发集，剩下的529篇文章作为训练集。  预处理数据的过程包括读取xml格式文件，使用斯坦福自然语言处理工具进行句子划分等操作，最终源数据的数据集统计结果如下表所示：  表5 ACE2005英文数据集统计结果   |  |  |  |  | | --- | --- | --- | --- | |  | Train | dev | Test | | Documents | 529 | 30 | 40 | | Sentences | 14724 | 875 | 713 | | Triggers | 4312 | 492 | 422 | | Arguments | 7811 | 933 | 892 | | Entity Mentions | 53045 | 4050 | 4226 |   **②评估方法选择：**  同以前的研究工作一样[15][17][29-31]，当一个预测出的事件触发词在文章中的位置、对应的事件类型和子事件类型都和人工标注的答案一致时，即认定预测出的事件触发词分类（Trigger Classification）。  最后以精度(P)、召回率(R)和F1分数(F1-score)作为评价指标，计算公式为：   |  |  | | --- | --- | |  | (30) | |  | (31) | |  | (32) |   **（2）基于多头图注意力网络模型的事件抽取任务实验结果**  本文实现了基于Pytorch的图注意力网络模型的事件抽取系统。采用stanfordcorenlp自然语言处理工具进行语料的句子分割、语法依赖树的标注。在词向量表示模块，我们设计BERT词向量的维度为768维，词性向量和实体类型的编码向量维度为50维，位置编码为10维，字符编码维度为50维。在图神经网络模块，我们采用一层双向LSTM，输出维度为768，我们使用2层图注意力神经网络，每层网络内部有3个头实现丰富特征的获取。同时设置句子的固定长度为50，通过对短句子进行词填充，对于长句子进行剪裁实现。学习率设置为1e-5。模型结果如下表所示：  表6.模型效果对比表   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Method | Trigger Identification(%) | | | Trigger Classification(%) | | | | | | P | R | F1 | P | R | | F | | | Cross-event statistical model[30] | N/A | | | 68.7 | | 68.9 | | 68.8 | | Joint local+  glocal features[29] | 76.8 | 65.0 | 70.4 | 73.7 | 62.3 | | 67.5 | | | DMCNN[8] | 80.4 | 67.7 | 73.5 | 75.6 | 63.6 | | 69.1 | | | JRNN[11] | 68.5 | 75.7 | 71.9 | 66.0 | 73.0 | | 69.3 | | | DMCNN-DS[21] | 79.7 | 69.6 | 74.3 | 75.7 | 66.0 | | 70.5 | | | db-RNN[15] | N/A | | | 74.1 | 69.8 | | 71.9 | | | GCN[13] | N/A | | | **77.9** | 68.6 | | 73.1 | | | JMEE[17] | 80.2 | 72.1 | 75.9 | 76.3 | 71.3 | | 73.7 | | | DyGIE++[26] | N/A | N/A | 76.5 | N/A | N/A | | 73.6 | | | **WGAT(ours)** | **80.5** | **76.8** | **78.6** | 77.2 | **73.7** | | **75.4** | |   可以看到我们的模型在触发词识别模块，效果要好于上述所列的所有的模型，比所列的最好的模型F1值要高1.9%。在触发词分类模块我们的效果好于上述所列的所有的模型，比所列最好的模型F1值要高1.8%。这说明我们的模型能够更好的捕获单词之间的注意力和能够抽取更好的抽象特征进行分类，证明了我们模型的有效性。  除了整体上对数据效果进行评估比较外，我们也按照之前的研究将验证我们的模型在单事件和多事件上的效果。其中单事件指的是一个句子中只有一个触发词，那多事件就对应于一个句子有两个及两个以上的触发词。我们将测试数据按照上述规则划分两部分1/1和1/N对应于上述所说的单事件和多事件，最终模型效果的展示如下表所示：  表 7 模型在1/1和1/N数据上的F1值对比   |  |  |  |  | | --- | --- | --- | --- | | **Method** | **1/1** | **1/N** | **ALL** | | DMCNN[8] | 74.3 | 50.9 | 69.1 | | JRNN[11] | 75.6 | 64.8 | 69.3 | | JMEE[17] | 75.2 | **72.7** | 73.7 | | WGAT(ours) | **80.8** | 71.0 | **75.4** |   可以看到虽然我们的模型在1/N上效果不是很好，比JMEE模型低1.7%，但是我们在单事件上事件抽取的效果远远超过于其他的模型，比目前最好的模型效果高5.6%，这验证我们的模型设计的有效性和正确性。  **（3）基于生成对抗网络模型的数据增强部分实验结果**  ①纽约时报数据集的统计  表8.纽约时报数据集统计   |  |  | | --- | --- | | 总数据量sentences | 选取数据量sentences | | 566630 | 3000 |   ②预训练判别器  表9. 预训练判别器在ACE2005测试数据集上的效果   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | Method | Trigger Identification | | | Trigger Classification | | | | P | R | F1 | P | R | F1 | | 判别器 | 67.3 | 62.3 | 64.7 | 59.8 | 55.3 | 57.5 |   ③对抗训练  对抗过程中数据筛选的阈值设置为0.6，即10次中有6次及以上判别器分类结果和伪标签相同我们即认为这条数据是较为可靠的数据。  **3、已发表的与学位论文相关的学术论文等**  1.提交一篇专利《一种基于权重图注意力神经网络的事件检测方法》，  专利号：202010246814.9  **4、主要创新点**  **（1）基于多头图注意神经网络的事件抽取模型**  本文将权重图注意力神经网络技术引入事件抽取领域，并创新性的提出一种注意力计算机制和多头图神经网络组织方式。过去基于事件抽取的模型大部分基于序列结构模型，这可能会丢失长距离依赖关系，在我们的模型中,通过使用图神经网络将序列结构通过语法依赖树的边转为图结构数据，并通过一定的权重将邻居信息融合到中心节点中，使中心节点拥有更加丰富的语义信息，这个权重参数不仅考虑到中心节点在整个句子中的语义权重信息，还考虑到了中心节点和邻居节点的相关性权重信息来共同决定最终的权重参数。不仅如此，我们还通过设置级联多头图注意卷积来获取更加丰富的语义信息，同时为了让原语义信息在图注意网络中的信息不流失，通过使用残差网络的形式来进行使得原语义信息无损的向前传播。该机制较之前的图神经网络，在利用邻居节点的特征聚合更新中心节点的过程中，更加全面的考虑了中心节点对其邻居节点赋予的权重参数问题，大大提高了事件检测的精确率，召回率以及F1值。  **（2）半监督场景下基于生成对抗网络模型的数据增强方法**  本文将弱监督的简单自训练方法引入事件抽取领域，并使用生成对抗网络进行数据筛选实现数据增强。事件抽取领域虽然前面已有大量的研究，但是大部分都是基于ACE2005数据集进行研究，但是该数据集存在一定的局限性，如事件类型不均匀，数据量少等的问题。有许多非结构化的文本语料由于没有标注而不能使用。针对此问题，我们提出使用触发词发现策略，生成大量的不可靠数据候选集，并利用生成对抗网络进行数据的筛选，将筛选出来的较为可靠的数据作为增强数据输入到多头图注意力神经网络中，通过增加训练数据集的数量，提到模型的泛化能力，从而提高模型在ACE2005测试集上的效果。  **（3）基于深度学习的事件抽取系统**  本文实现了用于事件抽取的系统。目前大部分基于深度学习的事件抽取系统都只停留在算法实现方面，并没有真正的实用性系统。针对此问题，我们简单的设计了一个可视化可交互的事件抽取系统。系统可以接受用户输入的句子，利用训练好的模型进行预测，并将预测结果包括触发词的识别和触发的事件类型可视化的展示。 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **四、下一步工作计划及进度安排**（内容、时间及工作量估计，字数不少于：150）  预计到2020年11月完成生成对抗网络算法的开发工作，到2021年1月底完成事件抽取系统的开发工作，并完成系统优化测试，提高系统性能；到2021年3月底，完成论文初稿的撰写工作；到2021年5月底完成优化论文细节，准备完答辩所需材料。  表10 下一步计划及进度安排   |  |  |  | | --- | --- | --- | | 时间 | 研究内容 | 预期效果 | | 2020.9－2020.11 | 完成生成对抗网络模型相关实验 | 生成对抗网络收敛，生成的增强数据在事件抽取模型上带来在ACE2005测试集上效果的提升 | | 2020.12－2021.1 | Django搭建系统，完成系统测试 | 完成系统搭建所有工作 | | 2021.2－2021.3 | 论文初稿撰写 | 撰写至少3万字的论文 | | 2021.4－2021.5 | 完成毕业论文撰写 | 完成论文，准备好答辩材料 | |
| **五、论文及整改方案**（论文后期工作存在的困难、问题及整改方案，字数不少于：200）  **1、论文后期工作存在的困难和问题**  1）通过实验发现，生成对抗网络训练比较困难，难以收敛，调参量大，算法效果有待提高，需要进一步的探索该目前算法的问题。  2）事件抽取系统的可视化展示比较少，系统的功能可根据需要进行进一步的功能开发，系统的性能也是需要考虑的因素。  **2、整改方案**  1）通过查阅相关文献，进一步调试生成对抗网络的算法。并通过实验对不同参数进行验证。  2）设计事件抽取系统的功能，美化系统界面，提升系统响应速度。增加如用户登录，错误报告等功能，提升系统的性能和用户体验。 |

|  |
| --- |
| **参考文献**  [1] Zhao Qi,Liu Jianhua,Feng Haoran. An ACE View of the Development Tendency of Information Extraction Technology. New Technology of Library and Information Service, DOI：10.11925/infotech.1003-3513.2008.03.04  [2] David Ahn. 2006. The stages of event extraction. In Proceedings of ACL Workshop on Annotating and Reasoning about Time and Events, pages 1–8  [3] Hui Yang, Tat-Seng Chua, Shuguang Wang, and ChunKeat Koh. 2003. Structured use of external knowledge for event-based open domain question answering. In Proceedings of SIGIR, pages 33–40.  [4] P Basile, A Caputo, G Semeraro, and L Siciliani. 2014. Extending an information retrieval system through time event extraction. In Proceedings of DART, pages 36–47  [5] Jonathan Berant, Vivek Srikumar, Pei-Chun Chen, Abby Vander Linden, Brittany Harding, Brad Huang, Peter Clark, and Christopher D. Manning. Modeling biological processes for reading comprehension. In Proceedings of the 2014 Conferenceon Empirical Methods in Natural Language Processing, pages 1499–1510. Association for Computational Linguistics, October 2014.  [6]Naomi Daniel, Dragomir Radev, and Timothy Allison. Sub-event based multi-document summarization. In Proceedings of Human Language Technologies:The 2003 Annual Conference of the North American Chapter of the Association for Computational Linguisticson on Text summarization workshop, pages 9–16.Association for Computational Linguistics, 2003.  [7] Elena Filatova and Vasileios Hatzivassiloglou. Event-based extractive summarization. In Proceedings of Annual Conference of the Association for Computational Linguistics Workshop on Summarization. Association for Computational Linguistics, 2004.  [8] Chen Y, Xu L, Liu K, et al. Event extraction via dynamic multi-pooling convolutional neural networks[C]// Proc of the 53rd ACL and the 7th IJCNLP, Volume 1: Long Papers. Beijing, China: ACL, 2015: 167-176  [9] Feng X, Qin B, Liu T. A language-independent neural network for event detection[J]. Science China Information Sciences, 2018, 61(9): 092106  [10] Liu J, Chen Y, Liu K, et al. Event detection via gated multilingual attention mechanism[C]//Thirty-Second AAAI Conference on Artificial Intelligence. 2018.  [11] Nguyen T H, Cho K, Grishman R. Joint event extraction via recurrent neural networks[C]// Proc of the 2016 NAACL: Human Language Technologies. San Diego, California: ACL, 2016: 300-309  [12] Liu X, Luo Z, Huang H. Jointly multiple events extraction via attention-based graph information aggregation[J]. arXiv preprint arXiv:1809.09078, 2018.  [13]Nguyen T H, Grishman R. Graph convolutional networks with argument-aware pooling for event detection[C]//Thirty-Second AAAI Conference on Artificial Intelligence. 2018.  [14] Yubo Chen, Shulin Liu, Shizhu He, Kang Liu and Jun Zhao. Event Extraction via Bidirectional Long Short-Term Memory Tensor Neural Networks. In Proceedings of the 15th China National Conference on Computational Linguistics and the 4th International Symposium on Natural Language Processing based on Naturally Annotated Big Data (CCL’2016 & NLP-NABD’2016), pages 190-203.Long Paper. 2016. (Best Paper Award)  [15] Lei Sha, Feng Qian, Baobao Chang, and Zhifang Sui, Jointly extracting event triggers and arguments by dependency-bridge RNN and tensor-based argument interaction, in Proc. AAAI, Louisiana, USA, 2018, pp. 5916-–5923.  [16] Qi P , Zhang Y , Zhang Y , et al. Stanza: A Python Natural Language Processing Toolkit for Many Human Languages[J]. 2020.  [17] Xiao Liu, Zhunchen Luo, Heyan Huang, Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation, in Proc. EMNLP, Brussels, Belgium, 2018, pp. 1247--1256.  [18] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[J]. arXiv preprint arXiv:1710.10903, 2017.  [19] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.  [20]Chen, Yubo, et al. Automatically labeled data generation for large scale event extraction. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vol. 1. 2017  [21] Liu S, Chen Y, He S, et al. Leveraging framenet to improve automatic event detection[C]// Proc of the 54th ACL, Volume 1: Long Papers. Berlin, Germany: ACL, 2016: 2134-2143  [22] Wang X, Han X, Liu Z, et al. Adversarial Training for Weakly Supervised Event Detection[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). 2019: 998-1008.  [23] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in neural information processing systems. 2014: 2672-2680.  [24] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.  [25] Diego Marcheggiani and Ivan Titov, Encoding sentences with graph convolutional networks for semantic role labeling,'' in Proc. EMNLP, Copenhagen, Denmark, 2017, pp. 1506–-1515.  [26] David Wadden, Ulme Wennberg, Yi Luan, Hannaneh Hajishirzi, Entity, Relation, and Event Extraction with Contextualized Span Representations, in Proc. EMNLP-IJCNLP, Hong Kong, China, 2019, pp. 5783--5788.  [27] Zhang Y, Guo Z, Lu W. Attention guided graph convolutional networks for relation extraction[J]. arXiv preprint arXiv:1906.07510, 2019.  [28] Rupesh Kumar Srivastava, Klaus Greff, and Jurgen Schmidhuber, Training very deep networks, in Proc. NIPS, Montreal, Quebec, Canada, 2015, pp. 2377–-2385.  [29] Li Q, Ji H, Huang L. Joint event extraction via structured prediction with global features[C]//Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2013: 73-82.  [30] Shasha Liao and Ralph Grishman. 2010. Using document level cross-event inference to improve event extraction. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL '10). Association for Computational Linguistics, Stroudsburg, PA, USA, 789-797.  [31] Shulin Liu, Kang Liu, Shizhu He, and Jun Zhao, A probabilistic soft logic based approach to exploiting latent and global information in event classification, in Proc. AAAI, Arizona, USA, 2016, pp. 2993–-2999. |
| |  |  |  |  | | --- | --- | --- | --- | | 姓 名 | 职 称 | 职务 | 工 作 单 位 | | 吴斌 | 教授 | 组长 | 北京邮电大学 | | 石川 | 教授 | 成员 | 北京邮电大学 | | 王柏 | 教授 | 成员 | 北京邮电大学 | | 于艳华 | 副教授 | 成员 | 北京邮电大学 |   **评审小组** |

|  |
| --- |
| **导师评语** |
| 导师：  日期： 年 月 日 |
| **阶段报告小组意见：** |
| 负责人：  日期： 年 月 日 |
| **学院意见：** |
| 负责人：  日期： 年 月 日 （签章） |