

# 基于图卷积神经网络的中文事件时序关系分类

陈飞宇, 张茹

(北京邮电大学网络空间安全学院, 北京 100876)

**摘要:** 事件时序关系任务中, 中文数据集相对于英文数据集较为匮乏, 导致中文文本的事件时序关系分类研究相对滞后。本文在中文突发事件语料库 CEC 上进行了事件时序关系分类的探究。传统的深度学习方法擅长处理欧式空间中的数据, 事件之间可能存在种属、依赖等多维特征, 这在非欧空间中更能体现句子含义。为此, 本文提出基于图卷积神经网络的中文事件时序关系分类方法, 将一维的文本语句构建成多维的图结构数据, 在卷积神经网络层中进行训练, 完成分类工作。在 CEC 数据集上的实验表明, 本文方法的性能优于其他基线任务, 取得了更好的实验效果。

**关键词:** 事件时序关系; 图卷积神经网络; 依存句法; 非欧空间

**中图分类号:** TP391.1

## Classification of chinese event sequence relations based on graph convolution neural network

Chen Feiyu, Zhang Ru

(Cyberspace Security School, Beijing University of Posts and Telecommunications, Beijing)

**Abstract:** In the task of event temporal relationship, the Chinese dataset is relatively scarce compared with the English dataset, resulting in the relatively lagging research of event temporal relationship classification of Chinese text. In this paper, the classification of event temporal relations is explored on the Chinese emergency corpus CEC. The traditional deep learning method is good at processing data in European space, and there may be multi-dimensional characteristics such as species and dependency between events, which can better reflect the meaning of sentences in non-European space. For this reason, this paper proposes a Chinese event temporal relationship classification method based on graph convolution neural network, which constructs one-dimensional text statements into multidimensional graph structure data, and trains in the convolution neural network layer to complete the classification work. Experiments on CEC data sets show that the performance of this method is better than other baseline tasks, and better experimental results are achieved.

**Key words:** Event temporal relationship; Graph convolution neural network; Dependency syntax; Non-European space

## 0 引言

目前, 大多数事件时序关系分类任务都是基于英文数据集的, 中文事件时序关系研究的起步要晚于英文, 并且缺乏中文语料资源, 这使得对于中文文本的事件时序关系分类研究工作相对较少<sup>[1]</sup>。Cheng 等人<sup>[2]</sup>根据 TimeML 语料库, 创建了一个中文时序关系语料库, 并利用机器学习算法在此语料库上进行了实验。中文文本相对于英文文本, 其语义和语法更加灵活复杂。英文在词性方面的划分和使用更加严谨, 这使得网络模型在英文语料库上更容易提取出更加准确的语义信息和句法结构。本课题采用中文语料库 CEC 进行研究, 探究中文时

**作者简介:** 陈飞宇(1997-),男,硕士研究生,主要研究方向:自然语言处理, 事件时序关系

**通信联系人:** 张茹, 女, 教授, 主要研究方向: 自然语言处理, 信息隐藏, 图像取证. E-mail: zhangru@bupt.edu.cn

序关系的分类方法。

近年来,事件时序关系分类方法主要分为两种类型,一种是利用深度学习神经网络具有的无监督学习能力和复杂的推导网络,提取文本特征后不断迭代训练,对结果进行拟合。另一种是利用外部知识库的先验知识,辅助自身模型进行预测。本文针对第一种方式,探寻深度学习神经网络方法在中文事件时序关系分类任务中的方法,提出基于图卷积神经网络的中文事件时序关系分类模型。

传统上,针对文本分类的模型一直侧重与单词嵌入的有效性和用于文档嵌入的聚合单词嵌入。这些词嵌入可以是无监督的预训练嵌入(例如 Word2Vec 或 Glove),然后将其输入分类器中。最近,诸如 CNN 和 RNN 的深度学习模型已经成为有用的文本编码器。在这两种情况下,文本表示都是从单词嵌入中学习的。对于欧几里得空间的文本数据,尤其是序列化的短文本,传统的深度学习网络模型(如 LSTM, CNN 等)已经取得了不错的成果。但是,对于长文本结构化的句子,为了更好地捕捉长距离词语间的语义信息,将其转化为图结构化数据更为合适。图结构化数据属于非欧几里得空间数据结构,传统神经网络模型难以胜任,所以研究人员发明了图卷积神经网络。图卷积神经网络在构建语料库图时,将文档和单词视为节点,将文档间或单词间的关系视为节点间的连线。

本文首先将 CEC 数据集进行预处理,将 XML 格式数据转化为纯文本数据。随后利用图构建方法,将 CEC 数据集中的文本语句映射为图数据,作为图卷积神经网络(GCN)的输入。然后使用 GCN 和池化网络学习节点邻居和节点本身的表现信息,在交叉熵损失函数的优化后,使用 Softmax 函数来预测标签。

## 1 相关工作

对于英文文本事件时序关系分类工作的研究,要远远超过中文文本。这是由于英文是全世界应用范围最广的语言,语义表达清晰,句法结构较为固定,对于人和计算机而言,都容易理解。而中文词汇量巨大,语句表达灵活,语义信息丰富,大大提高了机器进行文本理解的难度。另外,由于自然语言处理领域的研究工作,国外向来是领先于国内的,这就导致研究人员过多集中在英文领域。自然语言处理研究工作的发展离不开数据集,数据集的丰富程度直接反映了当前任务的进展,主流数据集(如 ACE、TimeBank 等)大多数属于英文数据集。虽然 ACE 数据集中包含中文部分,但是占总体比例并不高,且标注质量较低。郑新等人<sup>[1]</sup>。进行实验使用的便是 ACE 中文语料库,率先对中文事件时序关系进行了探究。在此基础上,黄一龙等人<sup>[3]</sup>发现因果关系和时序关系之间存在这关联性,提出事件因果与时序关系识别的联合推理模型。李婧<sup>[4]</sup>提出融合事件句、事件属性和事件间的联合路径,从这三个维度入手,利用 BERT 挖掘语义信息,利用事件属性挖掘时序差异,利用最短路径表示事件间的关联性,在中文语料库上取得了较好的性能。

依存句法分析,无论是在英文文本还是在中文文本研究中,都是一项重要的句法分析方法。依存句法分析能够充分挖掘文本语义信息,捕捉长距离词语关系,直观显示文本上下文结构。李良毅等人<sup>[5]</sup>引入依存注意力机制,利用双路依存机制,提取依存路径,并使用 Bi-LSTM 模型进行时序关系分类。戴倩雯等人<sup>[6]</sup>从篇章角度出发,利用依存路径信息和修辞关系构建神经网络来识别事件时序关系。依存句法分析形成的依存句法树是典型的非欧几里得空间数据结构,传统的深度学习方法显然无法应对。针对这一问题,本文引入了图卷积神经网络。

图卷积神经网络从谱图理论和空间方法上发展而来,能够以非欧空间的图数据作为输入,克服平移不变性缺失带来的图卷积算子定义困难。图卷积神经网络已经在许多领域得到应用,不同的应用具有不同的数据结构,所以图节点和边代表的含义也不同。引文网络将论文视为节点,引用关系视为连边。推荐系统根据用户和商品间的关系将其视为矩阵补全或链接预测任务。交通网络是一个典型的图网络结构,路口或地点和路线构成的拓扑结构刚好是自然的图数据。生物化学中的蛋白质、DNA、化学分子都是复杂的网络拓扑结构,将一个蛋白质分子视为一个图,微小分子视为图节点,键或相互作用视为节点间的边。将图片视为节点,图片间的相似度视为边,可以定义一个全连接图,利用图卷积神经网络可以在少样本、主动学习或半监督任务上发挥作用。依存句法树、文档共词矩阵、抽象含义表达是自然语言处理领域常见的图数据。

大量研究表明,图卷积神经网络在解决自然语言处理任务上发挥了很大作用,在一定程度上提升了任务的结果。TextGCN 模型<sup>[7]</sup>将单词和文本视为节点,构建无向加权图,以实现多分类任务。TensorGCN 模型<sup>[8]</sup>基于长短期记忆网络和词语的句法依赖。Text-level-GCN 模型<sup>[9]</sup>将每一篇文章映射成一个有向图,设计全局的边权矩阵和特征矩阵,解决了高内存问题,且泛化能力较强。

本文在中文突发事件数据集 CEC 上探究中文事件时序关系的分类方法,将 CEC 语料作为生数据,利用 Stanford-Core-NLP 工具对 CEC 数据集文本进行依存句法分析,得到预处理后的语句,然后通过图构建方法将其转化为图数据矩阵作为图卷积神经网络模型的输入,经过训练、迭代和激活后进行事件时序关系的分类工作。

## 2 基于图像分解与层间空间注意力传递的多层人脸伪造检测框架

### 2.1 整体架构

本文对基于图卷积神经网络的中文事件时序关系分类方法进行详细介绍,具体模型框架如图 1 所示。具体可分为预处理层、图构建层、图卷积层、激活函数层。其中预处理层负责对 CEC 数据集生语料进行分析,使用 Stanford-Core-NLP 工具进行分词、命名实体识别、依存句法分析等工作,将处理后的语料保存,作为实验的标准数据。图构建层负责将预处理后的数据映射成图数据矩阵,将事件视为节点,包括事件属性、依存关系等特征,将事件间的时序关系视为连边,构成有向图。经过图卷积后,使用交叉熵损失函数进行优化,最后利用 Softmax 激活函数给出事件时序关系的可能性。

### 2.2 预处理层

依存句法在许多自然语言处理任务中都发挥了重要的作用,为网络模型充分学习文本特征并表现良好性能提供了巨大帮助,并且对显示特征依赖更小。根据其他自然语言处理任务的思想,本文将依存句法引入中文事件时序关系分类任务中来。给定一个包含事件对的语句,对该语句进行依存句法分析,以此捕捉长距离的上下文信息。

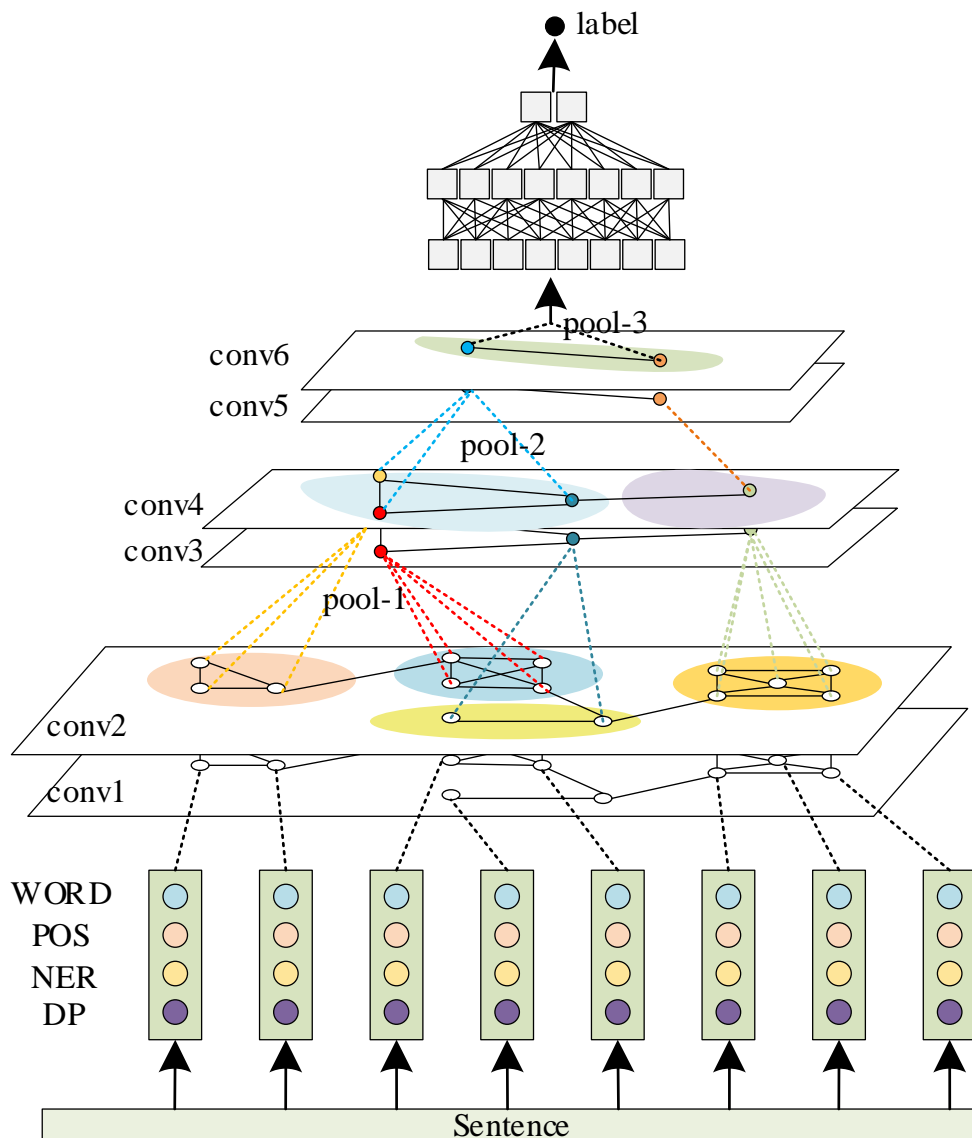


图1 模型框架图

本文使用 Stanford-Core-NLP 工具进行语句分析，主要的分析工作有分词分析、命名实体识别和依存句法分析。这些工作可归纳为两类特征，分别为句法特征和词法特征。其中词法特征包括单词（Word）、词性（Part of speech, POS）、命名实体识别（NER），句法特征则指的是依存句法（Dependency Parsing, DP）。对于输入语句的每一个词语，利用 Glove 预训练词向量进行查表映射，并将词法特征和句法特征进行随机初始化映射，完成特征的实值向量化映射。于是，就可以得到每个词语的向量表示：

$$x_i = x_w \oplus x_p \oplus x_n \oplus x_d \quad (1)$$

其中  $x_w$ ,  $x_p$ ,  $x_n$ ,  $x_d$  分别为分词向量、词性向量、命名实体识别向量和依存关系向量。最终得到句子的矩阵表示：

$$S_i = \{x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}\} \quad (2)$$

### 2.3 图构建层

图卷积神经网络是卷积神经网络用于编码图的改编。一张包含 $n$ 个节点的图可以用邻接矩阵  $A$  表示, 该矩阵的大小为 $n \times n$ 。当节点 $i$ 与节点 $j$ 之间存在一条边时, 邻接矩阵  $A$  中对应的元素为 1, 记为 $A_{ij} = 1$ 。在  $L$  层的 GCN 中, 如果将第 1 层节点 $i$ 的输入向量表示为

130  $h_i^{(l-1)}$ , 第 1 层节点 $i$ 的输出向量表示为 $h_i^l$ , 则图卷积操作可写成:

$$h_i^{(l)} = \sigma(\sum_{j=1}^n A_{ij} W^{(l)} h_j^{(l-1)} + b^{(l)}) \quad (3)$$

其中 $W^{(l)}$ 表示线性变换,  $b^{(l)}$ 表示偏置项,  $\sigma$ 表示非线性激活函数(如 ReLU)。通俗地说, 图卷积过程中, 每个节点会获取与其相连的其他节点的信息, 并将这些信息合并为一个汇总结果。

135 为了建模依赖树, 我们通过将每棵树转换为其对应的邻接矩阵  $A$  来适应图卷积操作。如果在节点 $i$ 和节点 $j$ 之间存在一条依赖边, 则 $A_{ij} = 1$ 。然而在(3)中简单地应用图卷积操作可能导致节点表示具有巨大差异, 因为一个标记词的度变化很大。这可能会是我们的句子表示偏向于支持高频次节点, 而不管节点中携带的信息。此外,  $h_i^{(l-1)}$ 中的信息永远不会转移到 $h_i^l$ 中, 因为节点永远不会在依赖树中连接到自己。本文通过图卷积中的激活归一化来解决这些问题, 然后通过其通过非线性传递, 并向图中的每个节点添加自循环。

$$h_i^{(l)} = \sigma(\sum_{j=1}^n \tilde{A}_{ij} W^{(l)} h_j^{(l-1)} / d_i + b^{(l)}) \quad (4)$$

其中 $\tilde{A} = A + I$ ,  $I$ 是 $n \times n$ 单位矩阵,  $d_i = \sum_{j=1}^n A_{ij}$ 是结果图中标记词 $i$ 的度。

### 2.4 图卷积层

145 在  $L$  层上叠加(4)的操作, 得到了一个深度图卷积神经网络, 设 $h_1^0, h_2^0, \dots, h_n^0$ 为输入词向量, 用 $h_1^L, h_2^L, \dots, h_n^L$ 作为输出单词表示。该网络中的所有操作都可以通过矩阵乘法有效地实现, 使其成为在示例上批量计算和在 GPU 上运行的理想选择。此外, 标记之间的信息传播是并行进行的, 并行运行时不依赖与依存句法树的深度。

150 首先定义关系分类的任务。令(2)表示一个句子, 其中 $x_i$ 是第 $i$ 个符号。关系分类的目标是预测两个事件词之间存在的关系。在词向量上应用  $L$  层 GCN 后, 我们获得了每个标记词的隐藏表示, 这些标记词直接接收到依存句法树中相邻标记词不超过  $L$  条边的影响。

为了使用这些单词表示进行关系分类, 本文首先获得如下所示的句子表示:

$$h_{sent} = f(h^L) = f(GCN(h^0)) \quad (5)$$

155 其中 $h^L$ 表示 GCN 第  $L$  层的全体隐藏表示,  $f: \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ 是一个最大池化函数, 从  $n$  个输出向量映射到句子向量。在依存句法树中接近事件标记的信息通常是时序关系分类的核心, 因此从 $h^L$ 中可以得到两个事件主体表示 $h_s$ 和 $h_o$ 。以 $h_s$ 和 $h_o$ 为核心, 利用一阶图卷积神经网络和池化算子相结合的方法进行训练, 将多个子图池化后形成一个新的节点, 并且保证新的节点处于核心节点 $h_s$ 和 $h_o$ 的周围。在池化过程中, 节点特征和局部结构应该保证得到充分利用。谱池化可以通过谱聚类方法将图分为不重叠的子图, 且可以控制划分后的子图个数和池化比

例。通过 6 层图卷积和 3 层谱池化操作后，最终得到两个核心节点。这两个节点是以  $h_s$  和  $h_o$  为核心不断卷积训练后得到的，包含了其前一层生成该节点的子图的整体特征。

## 2.5 激活函数层

受到关于实体之间关系学习工作的启发，本文通过连接句子中的两个事件节点，并输入到前馈神经网络（FFNN）来获取用于分类的最终表示。

$$h_{final} = FFNN([h_s; h_o]) \quad (6)$$

然后将  $h_{final}$  这一表示送入线性激活函数 Softmax 中进行关系概率分布的计算，本文方法经过神经元后可以输出一个  $\mathbb{R}^4$  维的向量，表示为：

$$p_i = \frac{e^{a_i}}{\sum_{k=1}^n e^{a_k}} \quad (7)$$

其中， $p_i$  为第  $i$  个神经元经过 Softmax 计算得到的值， $a_i$  表示第  $i$  个神经元的输出值。为了衡量模型的优劣，更好地拟合数据，本文选用了交叉熵损失函数进行优化，损失函数公式如下所示：

$$\mathcal{L} = -\sum_i^n p_i \log a_i \quad (8)$$

## 3 实验与分析

### 3.1 实验数据集

本文使用的中文突发事件数据集 CEC 是由上海大学语义智能实验室所构建，从互联网上收集的关于突发事件的新闻报道，共分为火灾、地震、恐怖袭击、交通事故和食物中毒五种类型。以国务院颁布的《国家突发公共事件总体应急预案》为依据对突发事件进行分类，对生语料进行预处理、分析、标注和一致性检验等工作，并将标注后的结果保存成数据集，共计 332 篇文章。标注的部分效果如图 2 所示。

CEC 数据集采用可扩展标记语言（XML）的方式进行标注和存储，每个节点都有这对应的标注属性和要素。表 1 展示了标签和其对应的含义。本文主要关注其中的 Denoter、Time，其他要素用来作为帮助理解语义的特征。不同关系用 relType 表示，其中 eRelation 是事件之间的关系，本文将与时序关系有关的部分从所有关系中抽取出来，作为实验的数据集。从总体上看，标注后的事件要素统计如表 1 所示。从表 2 可以看出，CEC 数据集虽然规模比常用的数据集小，但对事件及事件要素的标注非常全面，对事件间的关系也做了标注，尤其是包含本课题所需要的时间要素和时序关系。

```
<Sentence>
  <Event eid="e2">
    <Time tid="t2" type="relTime">昨天凌晨4点半左右</Time>,
    <Location lid="l2">八达岭高速出京方向距康庄收费站1.5公里处</Location>,
    <Object sid="s2">5辆大货车</Object>发生
    <Denoter did="d2" type="emergency">追尾</Denoter>
  </Event>, 致
  <Event eid="e3">
    <Participant sid="s3">3人</Participant>
    <Denoter did="d3" type="stateChange">死亡</Denoter>
  </Event>
  <Event eid="e4">
    <Participant sid="s4">2人</Participant>
    <Denoter did="d4" type="stateChange">受伤</Denoter>
  </Event>。
</Sentence>
```

图 1 CEC 数据库标注示例

表 1 标签与其含义对应表

标签	含义
Denoter	事件触发词
Time	时间
Object	对象
Location	地点
Participant	参与者

表 2 CEC 各项数据统计表<sup>[10]</sup>

类型	文本	句子	事件	触发词	对象	环境	时间
地震	63	401	1002	1002	939	372	348
火灾	75	433	1216	1216	1107	363	358
交通事故	85	514	1802	1802	1565	459	435
恐怖袭击	48	324	823	823	774	204	256
食物中毒	61	392	1111	1111	1121	245	365
总计	332	2064	5954	5954	5506	1643	1762

3.2 基线方法

由于目前中文事件时序关系分类任务中还没有出现标准的评测体系，缺乏中文事件时序关系分类任务专用的语料库，并且不同研究者对于不同语料的处理方式有所不同，各数据集的标注体系也有所差别，这些都会对事件时序关系分类的结果产生一定的影响。因此无法保证完全客观地对比各个方法的优劣，本文试图在一定程度上展现某个方法的性能，将现有方法与本文方法进行比较，所选取的基线方法如下：

**事件时序关系分类器<sup>[1]</sup>**：该基线方法是首个关注中文事件时序关系的方法，提出了一些特征和定义了一个简单的分类器模型，对中文文本的时序关系做了初步探究；

**事件因果与时序关系联合模型<sup>[3]</sup>**：该基线方法认为事件之间的关系并不是孤立的，提出使用因果关系和时序关系相联合的方式进行推理，并在与基线方法（1）同样的中文数据集上进行了事件时序关系识别的工作；

**融合多维度信息方法<sup>[4]</sup>**：该基线方法融合事件句、事件要素属性和事件之间的联合路径这三个维度信息，在中文时序关系数据集上进行实验，可以较好地处理同句或跨句的事件时序关系识别工作；

多特征和规则约束方法<sup>[11]</sup>：该基线方法根据事件要素、触发词相似度、因果标志词和特殊词等多个特征进行分类，然后利用连接词、时间信息和事件类型等语义规则进行优化，并在 CEC 语料库上进行了实验。

### 3.3 实验设置

本实验共分为两个阶段。第一个阶段是对 CEC 数据集中的生语料进行预处理，使用 Stanford-Core-NLP 工具进行分词、词性标注和依存句法分析等工作。在使用工具进行分析后，获取事件词的位置，以便后续对其进行掩盖，将分词、词性和句法等特征封装成 JSON 数据格式进行存储。在事件时序关系上，本文实验使用关系集合{BEFORE, AFTER, EQAUL, VAGUE}。

第二个阶段是预测阶段，将处理好的数据输入到网络模型中进行实验。使用 Glove 预训练模型进行查表映射，得到句子的向量表示。实验中将词嵌入特征维度设置为 300，词性特征维度和命名实体特征维度均设置为 30，隐藏层维度设置为 200。输入时的 dropout rate 设置为 0.5，图卷积神经网络层中的 dropout rate 设置为 0.5。将单词随机设置为“UNK”的速率设置为 0.04。训练中的学习率设置为 0.3。在图卷积过程中，每层图结构的卷积核都会生成一个与原来相同大小的图，节点数量无法变少，为此本文使用最大池化函数，将其进行降维，在六层卷积层中的每两层之间设置池化层。同时为了防止过拟合，池化时使用 L2 正则化，惩罚参数设置为 0.003。本文方法的训练和验证过程中的平均损失随迭代次数的变化情况如图 4 和图 5 所示，由此可知，模型具有较快的收敛速度，无论是训练实验还是验证实验，在迭代次数 100 次以内都达到收敛状态并趋于稳定。批处理大小和训练批次可以根据内存容量和显卡算力进行调整，本实验将批处理大小设置为 50，进行了 100 轮实验。

在对文本语句进行依存句法分析时，发现依存句法树提供了丰富的结构，可以在关系分类中加以利用，但大多数与时序关系相关的信息通常包含在以两个事件的最低公共祖先为根的子树中，删除这个范围之外的标记有助于从句子中消除不相关的信息来加速关系分类。因此，本文将 GCN 与树剪枝策略结合起来，进一步提高性能。然而，过于激进的剪枝（例如只保留依存路径）会导致关键信息的丢失，反而影响模型的健壮性。例如，当模型只查看事件之间的依存路径时，图 3 中的“营救过程中，民警积极与被困司机交谈，稳定其紧张情绪。”就会只保留依存路径“营救→过程→交谈”，这样就缺少了表示时序关系的关键特征“中”。

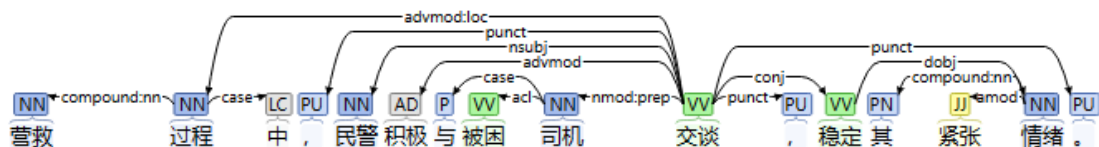
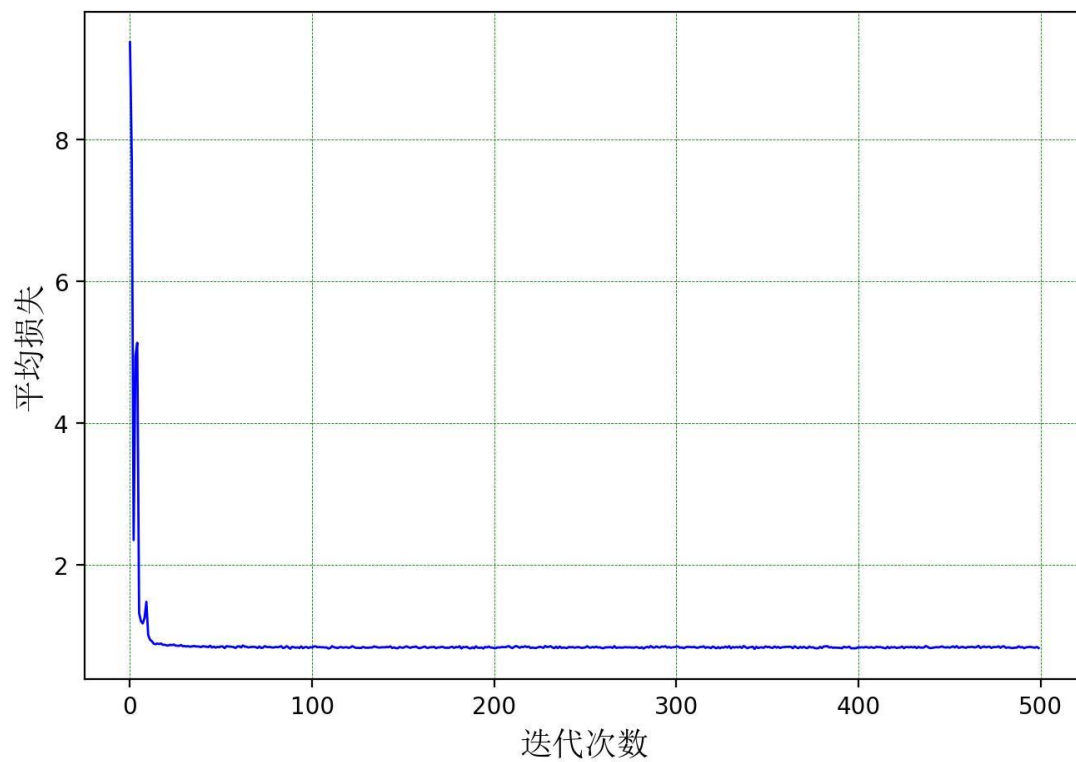


图 3 依存句法线性表示示意图

在这些观察的启发下，本文提出了以路径为中心的剪枝，从依存路径中合并信息。这是通过在最低公共祖先为根的子树中包含距离为  $K$  的标记词来实现的。 $K = 0$ ，对应于将依存句法树向下修剪到路径， $K = 1$  保留直接连接到路径的所有节点， $K = \infty$  保留整个最低公共祖先为根的子树。我们将这种修剪策略与我们的图卷积神经网络模型结合起来，直接将修剪过的树提供给图卷积层。实验结果如图 5 所示， $K = 1$  的修剪在包含相关信息(例如，随后和期间)和尽可能将不相关的内容排除在最终修剪的树之间实现了最佳平衡。



240

图 4 训练损失随迭代次数的变化曲线图

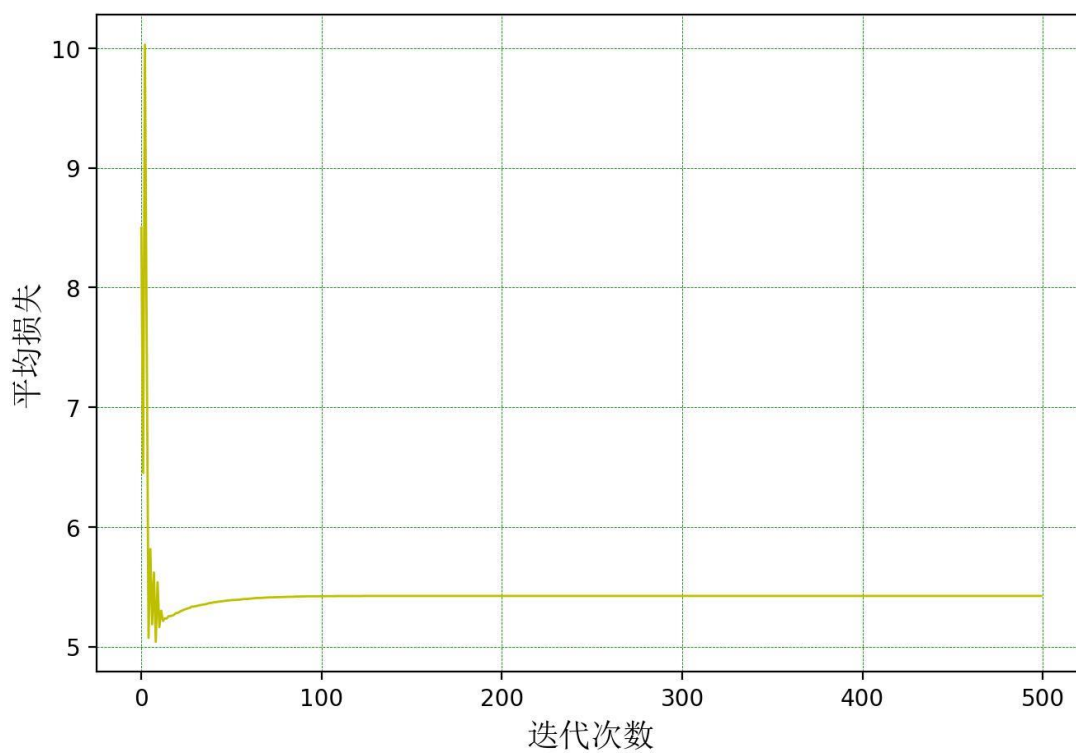


图 5 验证损失随迭代次数的变化曲线图

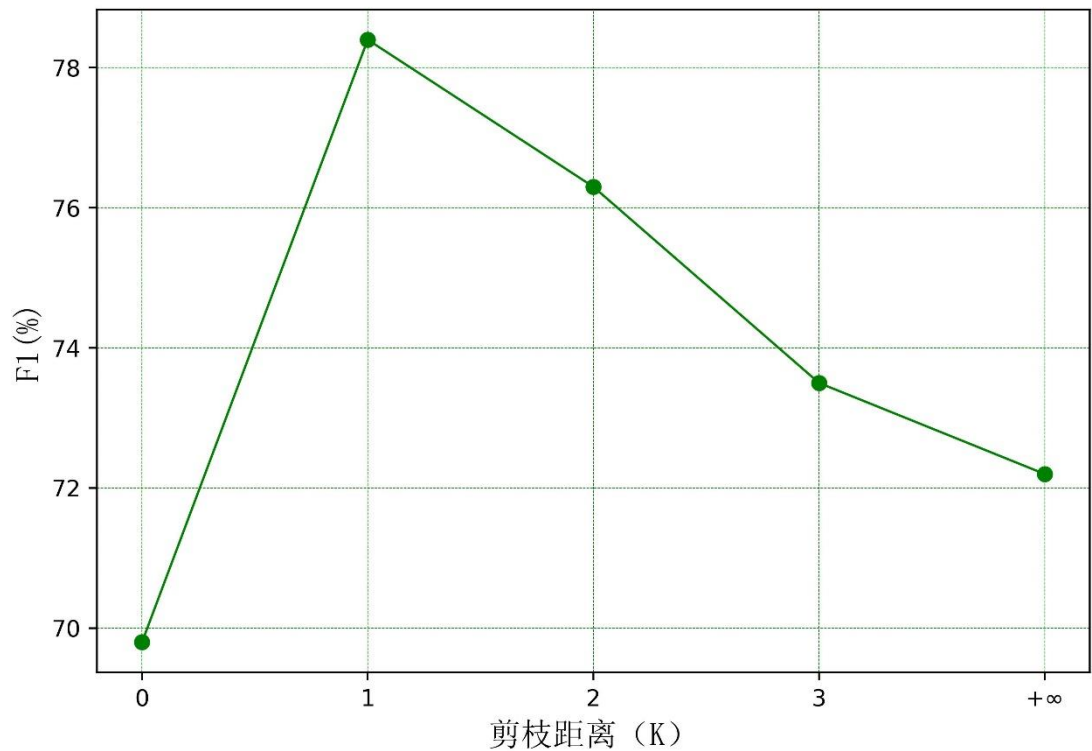


图 6 F1 值随剪枝距离 K 的变化图表

3.4 结果分析

由表 3 可知，相较于基线方法，本文方法在各项评价指标上均有所提升，这表明本文提出的基于图卷积神经网络的事件时序关系分类方法是有效的。事件时序关系分类器只是提出了部分句子特征，使用简单的分类器进行分类，属于传统的机器学习方法，对于四分类问题在精准率和召回率上均超过了 50%，表明机器学习在时序关系分类任务中发挥了一定的作用。事件因果与时序关系联合模型通过结合因果关系和时序关系进行联合推理，因果关系中多隐含了时序关系的先后关系，对事件时序关系分类有一定的帮助，所以在精准率上较事件时序关系分类器提升了将近 10 个百分点。融合多维度信息方法主要在事件句、事件属性和事件的联合路径入手，比较侧重文本的句法结构，仍旧会偏重事件间的先后关系，导致召回率的下滑。多特征和规则约束方法更强调构建更多的特征进行训练，然后使用规则约束优化模型，从而提高模型性能。本文提出的图卷积神经网络具有较强的空间处理能力，图结构更能体现词法和句法特征，并且深度学习方法复杂的网络结构削弱了对数据集质量的依赖，根据数据不平衡的特点，选用了微平均计算评估指标，最后表明本文方法在中文事件时序关系分类任务上取得了较好的表现。

3.5 消融实验

表 3 消融实验结果

模型	P	R	F1
A（基准方法）	79.2	80.1	78.4
B（A-词性）	68.6	65.2	66.8
C（A-依存句法）	60.7	58.3	59.1
D（A-词性-依存句法）	33.5	36.7	34.8

本文方法主要分为两个阶段，第一个阶段对实验数据使用 Stanford-Core-NLP 工具进行分析和处理，第二个阶段利用得到的数据构建图数据并进行图卷积训练。在第一阶段中，本文进行了分词、词性分析和依存句法分析。为了验证词性和依存句法分析在实验中的作用，以本文方法为基准方法进行了消融实验，实验结果如表 3 所示。

B 方法是将本方法中的词性分析去除，C 方法是去除依存句法分析，而 D 方法则是将词性和依存句法同时去除。从表中可以看出 B 方法与基准方法相比发生大幅下降，降幅均超过 10%。C 方法各指标下降程度更大，说明了依存句法和词性对事件时序关系有着很大的作用，并且相比之下依存句法的影响更大一些。D 方法的实验性能低于 A 的一半，去除词性和依存句法特征后，仍能进行预测，证明了图卷积神经网络在时序关系预测上的可行性和有效性。另外，也反映出词性和依存关系等特征在事件时序关系分类中的重要性。

## 4 结论

本文提出了基于图卷积神经网络的中文事件时序分类方法，使用中文突发事件数据集 CEC 作为实验语料，首先将其预处理，使用 Stanford-Core-NLP 工具对其进行分词、词性标注和依存句法分析，并保存为 XML 格式。利用预处理后的数据特征构建图数据矩阵，作为方法模型的输入，形成多个子图，经过图卷积将周围节点的特征进行融合，利用池化函数将多个子图整合成一个新的节点。最后，将训练得到的特征值输入到 Softmax 激活函数中进行时序关系类别预测。经过实验表明，本文提出的方法在所有评价指标上都比基线方法表现更好，展现出更好的分类性能，说明了该方法的有效性和先进性。

## [参考文献] (References)

- [1] 郑新, 李培峰, 朱巧明. 中文事件时序关系的标注和分类方法[J]. 计算机科学, 2015, 42(7): 276-279.
- [2] Cheng Y, Asahara M, Matsumoto Y. Constructing a temporal relation tagged corpus of chinese based on dependency structure analysis[C]//14th International Symposium on Temporal Representation and Reasoning (TIME'07). IEEE, 2007: 59-69.
- [3] 黄一龙, 李培峰, 朱巧明. 事件因果与时序关系识别的联合推理模型[J]. 计算机科学, 2018, 45(06): 204-207
- [4] 李婧, 李培峰, 朱巧明. 融合多维度信息的中文事件时序关系识别方法[J]. 山西大学学报(自然科学版), 2021, 44(03): 411-419
- [5] 李良毅. 融入事件依存路径的事件时序关系识别[J]. 通信技术, 2022, 55(05): 581-587.
- [6] 戴倩雯, 张龙印, 孔芳. 融合依存关系和篇章修辞关系的事件时序关系识别[J]. 模式识别与人工智能, 2019, 32(12): 1100-1106.
- [7] Yao L, Mao C, Luo Y. Graph convolutional networks for text classification[C]//Proceedings of the AAAI conference on artificial intelligence. 2019, 33(01): 7370-7377.
- [8] Liu X, You X, Zhang X, et al. Tensor graph convolutional networks for text classification[C]//Proceedings of the AAAI conference on artificial intelligence. 2020, 34(05): 8409-8416.
- [9] Huang L, Ma D, Li S, et al. Text level graph neural network for text classification[J]. arXiv preprint arXiv:1910.02356, 2019.
- [10] 陈红. 面向突发事件的时序关系识别及其应用研究[D]. 安徽理工大学, 2021.
- [11] 廖涛, 陈红, 方贤进. 基于多特征和规则约束的事件时序关系识别[J]. 阜阳师范学院学报(自然科学版), 2020.