

基于事件模式及类型的事件检测模型



代 翔*

(中国电子科技集团公司第十研究所 成都 610036)

【摘要】针对触发词定义标准模糊、语料标注成本高等问题,提出一种基于事件模式及类型的事件检测深度学习模型(PTNN)。首先基于实体的语法及语义特征获取潜在论元;其次将潜在论元抽象为角色,结合语法、语义、角色特征构建嵌入表示,增强输入对事件模式的体现;最后利用 Bi-LSTM 和基于事件类型的注意力机制,完成事件及类型判定。模型在不识别触发词的前提下,通过强化事件模式特征实现事件检测,避免了触发词标注困难的问题,证明了事件模式在神经网络上对事件检测的积极作用,将同类方法的最优效果提升了 3%,且达到了基于触发词的检测效果。

关键词 注意力机制; 事件检测; 事件模式; 长短时网络; 潜在论元
中图分类号 TP391.1 **文献标志码** A **doi**:10.12178/1001-0548.2021377

Event Detection Model Based on Event Pattern and Type Bias

DAI Xiang*

(The 10th Research Institute of China Electronics Technology Group Corporation Chengdu 610036)

Abstract To address the problems of vague criteria for trigger word definition and the high cost of corpus annotation, a deep learning model for event detection called pattern and type based neural network (PTNN) is proposed. First, potential theorems are obtained based on entities' syntactic and semantic features. Then, the potential theorems are abstracted as roles. The embedding representation of PTNN is constructed by combining syntactic, semantic, and role features to enhance the representation of event patterns. Last, event detection and type determination are accomplished by using Bi-LSTM (bidirectional long short-term memory) with an event type-based attention mechanism. The model achieves event detection by enhancing event pattern features instead of identifying trigger words, thus avoiding the challenging problem of trigger word annotation. Such an approach demonstrates the positive effect of event patterns for event detection on neural networks. Experiments demonstrate that it improves the state-of-the-art of event detection by 3%.

Key words attention; event detection; event pattern; LSTM; potential argument

事件是重要的信息表现形式,具有典型的类型及模式特征。事件模式体现事件的结构化特征,由事件的角色及其关系构成。不同类型事件在模式特征上有各自的个性特征。事件检测任务识别事件及其类型,主流研究均基于触发词特征完成检测。方法需专业人员对训练语料标注触发词,标注成本高且部分领域难以形成有效的标注标准。

为解决上述问题,本文提出面向事件模式及类型特征的事件检测深度学习模型(pattern and type bias neural network, PTNN)。模型不检测触发词,通过实体的语法、语义特征获取潜在论元,判断其

对应角色进行抽象表示,凸显事件模式信息;融合潜在论元的语法、语义及角色特征构建嵌入表示,利用深度学习模型学习事件模式特征及文本语义,并结合事件类型注意力机制实现事件检测。

主要工作如下:

- 1) 以实体类型、依存关系、词性定义论元特征约束,基于论元特征约束抽取潜在论元;
- 2) 以角色抽象潜在论元并进行表示替换,结合语法、语义、角色特征构建词嵌入表示,凸显输入的事件模式信息;
- 3) 使用双向长短时网络(Bi-long short-term

收稿日期: 2021-12-13; 修回日期: 2022-02-22

基金项目: 国家自然科学基金(U19A2078); 四川省科技计划(2020YFG0009)

作者简介: 代翔(1983-), 博士, 高级工程师, 主要从事智能信息处理及数据挖掘方面的研究。

*通信作者: 代翔, E-mail: 18208182867@qq.com

memory, Bi-LSTM) 及类型注意力机制构建深度学习模型, 根据输入的词、句级特征, 学习事件模式及特征类型, 实现事件检测。

1 相关工作

事件检测的主流方法为面向触发词的检测, 包括基于特征的方法和基于表示的方法。基于特征的方法以触发词统计特性为依据, 定义特征建模触发词。常用特征包括词法、句法特征、篇章信息及外部知识等^[1-8]。特征构建耗时耗力, 效果一般且不稳定。基于表示的方法为目前主流方法, 其利用标注触发词的训练语料, 学习输入的高维特征实现触发词识别^[9-14]。虽避免了人工构建特征, 但语料需专业人员标注, 具有触发词标准难以制定、标注成本高等问题, 且同一词语可触发多类事件的问题未很好解决。

而事件论元、模式等特征同样体现事件本质, 因此许多研究从这些特征入手。部分研究从事件实例出发, 利用同类事件在实体组成、语句结构上的相似性进行事件检测。文献 [15] 基于相同事件拥有相似实体的假设, 根据实体元素对文本聚类, 得到若干事件簇实现面向开放域的事件检测。方法避免了触发词标注, 但无法给出事件簇对应的具体类型, 得到的结果是若干关键词, 无法用于后续任务。文献 [16] 以实体和其类型构建“槽值对”, 基于同类事件具有相似“槽值对”集的假设分析事件类型与槽值对的关联。方法一定程度上实现了类型判定, 但“槽值对”不能全面体现事件框架特征, 对部分事件区分度较低。部分研究根据不同类型事件在实体、模式上的差别, 预先定义事件表示框架实现事件检测。文献 [17] 结合事件的 5W1H 分析法, 定义 7 种突发事件类型的事件框架, 抽取突发事件用于公共安全预警; 相似研究在科技、金融等领域也取得了一定成果^[18-19]。该类方法结合远程监督可自动生成标注数据, 解决了语料标注困难的问题。但事件框架表示结构复杂, 严重依赖专家知识, 且不同领域和任务需要不同的表示框架, 重复定义工作量大。文献 [20] 提出了基于实体的 TBNNAM 模型。其以实体为原始特征构建输入, 使用 LSTM 与基于类型的注意力完成事件识别。但该方法仅利用了实体信息, 没有考虑事件模式特征, 且只利用了单向语义信息。文献 [21] 结合 Bi-LSTM 与基于类型的多层多头注意力, 基于词向量的成长文本的向量表示, 识别电影剧本是否包含指

定类型事件。但其关注目标为文档级事件, 弱化了细粒度事件的识别和判定。

2 基于事件模式及类型的事件检测

本文提出一种基于事件模式及类型的事件检测模型 PTNN。模型基于非触发词特征检测事件, 通过学习句中实体的语义、语法及角色信息, 获取隐含的事件语义、结构特征, 利用基于类型的注意力机制强化对事件类型和模式的学习。

模型分为潜在论元获取、角色抽象及词替换、深度学习检测 3 部分。潜在论元获取从实体类型、依存关系、词性 3 个角度构建论元特征约束, 获取潜在论元。角色抽象及词替换将具体的潜在论元抽象为角色进行表示替换, 增强输入对事件模式的表示。深度学习检测融合各特征构建嵌入表示, 使用 Bi-LSTM 与基于类型的注意力从词、句两维度提取事件的类型、模式特征。模型原理如图 1 所示。

获取句 $S = \{w_1, w_2, \dots, w_k\}$ 中实体的语法、语义特征, 包括实体类型序列 E 、依存关系集 D 、词性序列 P 。根据论元特征约束获取潜在论元集 A :

$$A = \{a_1, a_2, \dots, w_l\} \quad a_i \in S, 1 \leq i \leq l \leq k$$

根据潜在论元的语法、语义特征, 将其抽象为不同角色 R 并进行表示替换, 增强输入对事件模式的表示, 得到处理过的序列:

$$S' = \{w'_1, w'_2, \dots, w'_k\}$$

以 X_w 表示词向量、 X_p 表示词性向量、 X_e 表示实体类型向量, 综合潜在论元各特征构建 S' 的向量序列 $X_{S'}$, 与待判断事件类型 T 组成输入对:

$$\langle X_{S'}, T \rangle$$

$X_{S'}$ 为 Bi-LSTM 各时间步输入, T 用于从模型参数中获取对应词、句级事件类型向量 T_{loc} 、 T_{glo} 。结合 Bi-LSTM 的隐层信息与事件类型向量, 获取词级特征 X_{loc} 与句级特征 X_{glo} , 综合两种特征实现事件检测及类型判别。

2.1 潜在论元获取

事件论元包含事件角色和部分属性, 体现了事件模式并隐含部分类型信息。时间、地点、参与者是最具代表性的论元。其中时间、地点在模式上不体现事件类型信息, 参与者则随事件类型的不同有着不同的实体类型, 如判决事件的参考者有被告、法官, 比赛事件的参与者有裁判、运动队。以句 “ S_1 : Police arrested four suspects in Shenyang

yesterday.” 为例，police 和 suspects 是参与者论元，Shenyang 和 yesterday 分别为地点和时间论元。

事件论元基于事件存在。面向一般文本句时无

法确定事件存在，故提出潜在论元概念，根据词的语法、语义特征约束抽取潜在论元，帮助事件检测。

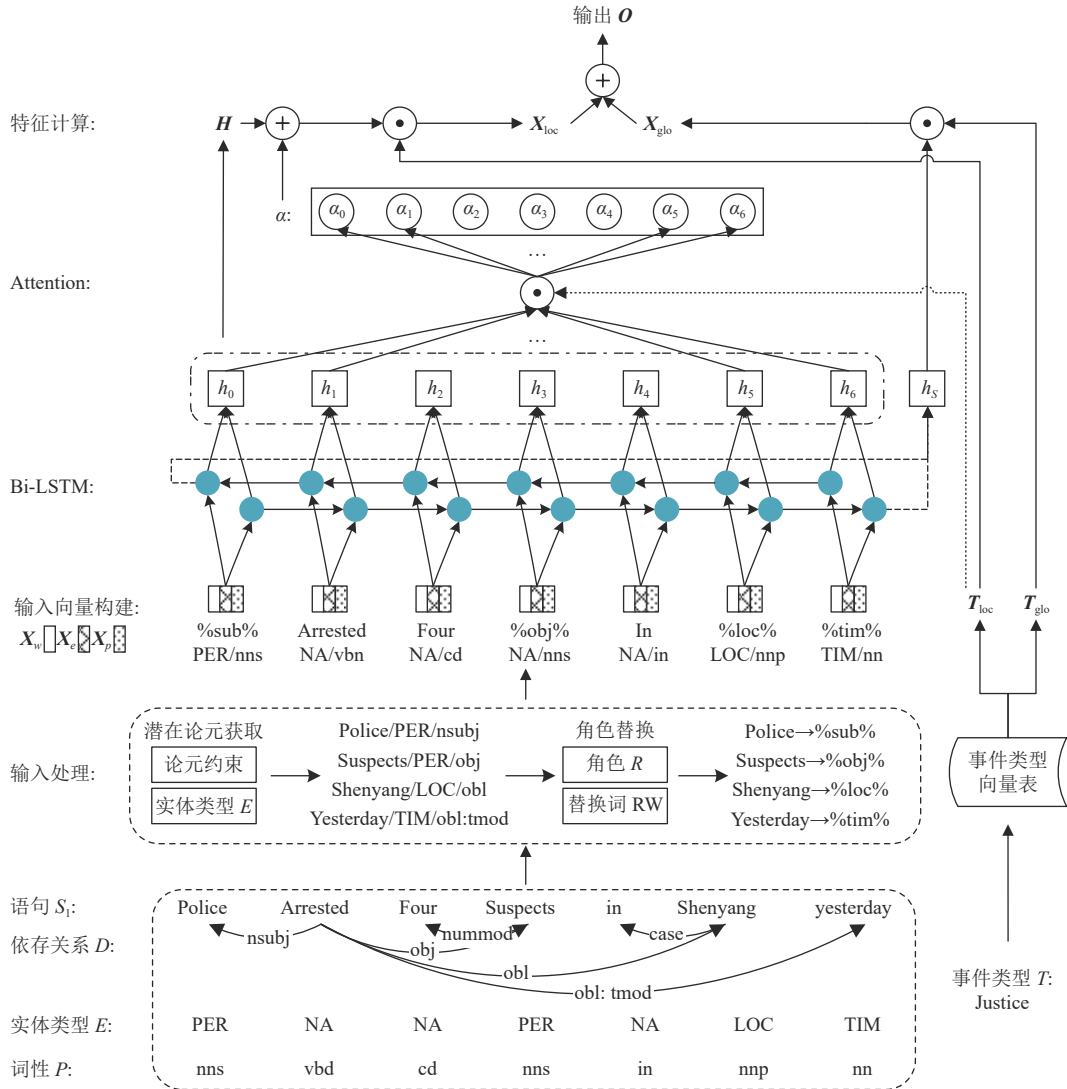


图 1 PTNN 模型

2.1.1 潜在论元

定义 1 潜在论元文本句 $S = \{w_1, w_2, \dots, w_k\}$ 中具有时间、地点及参与者角色属性的实体为 S 的潜在论元，表示为 S_A 。

若实体 w_i 在语义、语法特征上满足论元特征约束，则 $w_i \in S_A$ 语义特征包括实体类型及词性，语法特征为词在句中对应的依存关系，具体约束如下。

1) 实体类型约束 δ_e

三类论元中，时间、地点的实体类型固定，分别以 TIM 和 LOC 表示，参与者的实体类型与事件类型相关，表 1 为 ACE2005 数据集定义的部分类型。

表 1 不同类事件的参与者对应的实体类型

事件类型	参与者可对应实体类型
Life-marry	PER
Life-injure	PER、ORG、GPE、WEA、VEH、SUB
Business-start-org	PER、ORG、GPE

以 $T = \{t_1, t_2, \dots, t_n\}$ 表示指定的事件类型，对于 t_i 类事件，其参与者对应的实体类型集为 $E_{t_i} = \{e_{t_i1}, e_{t_i2}, \dots, e_{t_im}\}$ ，潜在论元应满足的实体类型约束 δ_e 为：

$$\delta_e = E_{t_1} \cup E_{t_2} \cup \dots \cup E_{t_n} \cup \{TIM, LOC\}$$

常见实体类型包括人名 PER、国家 NAT、机构 ORG 等。面向具体领域或特定事件类型时，潜在论元对应的实体类型对应增减。

2) 依存关系约束 δ_d

考虑句子“Tom’s girlfriend Lisa bought a new car.”，Tom 符合“购买”事件的实体类型约束，但其没有和事件动作发生关联，不符合论元的语法特征。

事件的本质是动作发生或状态改变，论元与谓词、论元与论元间存在事实关联，并通过词的依存体现。依存关系常表示为三元组 $\{w_i, w_j, r_k\}$, $w_i, w_j \in S$, r_k 为依存关系类型，关系由 w_i 指向 w_j , w_i 为头实体， w_j 为尾实体。

参与者为事件主客体或运用的工具、物品等。主客体直接发出或承受具体行为，依存关系中在表主谓、动宾关系的依存弧；工具、物品等通过主客体与动作关联，代表主客体的状态，通过状语类或动宾关系体现；时间、地点以状语类关系体现。故潜在论元依存句法约束 δ_d 为：

$$\delta_d = \{\text{主谓依存, 宾语类依存, 状语类依存}\}$$

3) 词性约束 δ_p

参与者体现为事件主客体或运用的工具、物品等，词性为名词或代词；时间为数字或特定词语如

“today、month”，词性为数词或名词；地点表现为名词。潜在论元的词性约束 δ_p 为：

$$\delta_p = \{\text{名词, 代词, 数词}\}$$

2.1.2 抽取算法

算法 1 潜在论元抽取算法

输入：文本句 S ；输出：潜在论元集 S_A ；
 $S_A = \emptyset$

获取句 $S = \{w_1, w_2, \dots, w_k\}$ 的实体类型序列 $E = \{e_1, e_2, \dots, e_k\}$

for $i=1$ to k :

if $e_i \in \delta_e$ then 添加 $w_i \rightarrow S_{Ae}$

end for

获取句 $S = \{w_1, w_2, \dots, w_k\}$ 的依存关系集 $D = \{s_{d1}, s_{d2}, \dots, s_{dk}\}$

for $i=1$ to k :

if 三元组 s_{di} 的 $r \in \delta_d$ then 添加 $w_i \rightarrow S_{Ad}$

end for

$S_A = S_{Ad} \cap S_{Ae}$

获取句 $S = \{w_1, w_2, \dots, w_k\}$ 的词性序列 $P = \{w_{p1}, w_{p2}, \dots, w_{pk}\}$

for $w_i \in S_A$

if $w_{pi} \in \delta_p$ then 保留 w_i

else 从 S_A 中删除 w_i

end for

以 S_1 为例，各信息如图 2 所示。

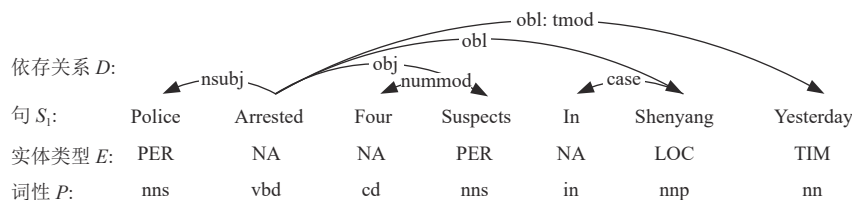


图 2 S_1 语义语法分析

S_1 的实体类型序列 E 为：

$$E = \{\text{PER, NA, NA, PER, NA, LOC, TIM}\}$$

根据实体类型约束 δ_A ，选出满足条件的实体：

$$\{\text{police, suspects, Shenyang, yesterday}\}$$

以 Stanford NLP 工具提取依存关系，根据依存关系约束 δ_d 得满足条件的三元组有：

$$\left\{ \begin{array}{l} (\text{arrested, police, nsubj}) \\ (\text{arrested, suspects, obj}) \\ (\text{arrested, Shenyang, obl}) \\ (\text{arrested, yesterday, obl : t mod}) \end{array} \right\}$$

nsubj 为名词主语，属主谓类依存；obj 为动词

宾语，属宾语类依存；obl 表名词充当附属物，属状语类依存；obl:tmod 则在前者基础上进一步指定为时间修饰语。定位对应尾实体得到满足依存约束的实体集：

$$\{\text{police, suspects, Shenyang, yesterday}\}$$

对两实体集取交集并以词性进行过滤，得到潜在论元集 S_A ：

$$\{\text{police, suspects, Shenyang, yesterday}\}$$

2.2 角色抽象及替换

潜在论元通过其存在和角色体现事件模式，但其具体值体现的是具体信息而非模式信息。如参与

者叫“张三”还是“李四”、发生时间是“今天”还是“昨天”并不决定事件是否存在，存在这些角色才是判定事件模式的要素。同时人名、时间等词是静态词向量“未登录词”问题的主要来源。为有效提取事件模式特征，提出基于角色的词替换。将潜在论元按角色替换为对应抽象表示，凸显事件模式的同时缓解“未登录词”问题。

根据论元的实体类型及依存关系进行角色分类，以 r_i 表示不同角色，形成角色集 $\delta_R = \{r_1, r_2, \dots, r_o\}$ ， $o \in \mathbb{R}$ 。为每个角色 r_i 定义对应抽象表示 r_{w_i} ，形成抽象实体集 $\delta_{RW} = \{r_{w_1}, r_{w_2}, \dots, r_{w_o}\}$ ， $o \in \mathbb{R}$ ， r_i 与 r_{w_i} 一一对应。

根据潜在论元的实体类型及存在的依存关系，判断潜在论元对应的角色，判别规则如表 2 所示。

表 2 最后一行涵盖范围大、具体含义不定，需根据具体事件领域、语料标准进行替换，替换基本规则如下。

1) 需保证分词时被识别为一个词，且不可与已有词、常规中文词相同；

- 2) 格式应与已定义抽象词保持一致；
- 3) 抽象词应能显式体现类别含义。

表 2 潜在论元角色判别规则

实体类型	依存关系	角色 r	抽象词 r_w
人、组织机构	主谓类	施动者	%sub%
	表被动的主谓类	受动者	%obj%
	动宾类	受动者	%sub%
时间	状语类	时间	%time%
地名、国家名	状语类	地点	%loc%
其他实体	状语类、动宾类	其他	-

算法 2 潜在论元替换算法

for $i=1$ to $\text{len}(S_A)$

 取 $w_i \in S_A$

 获得潜在论元 w_i 的实体类型 E ，存在的依存关系 D

 查表 2，判断角色 r_i

 以 r_i 对应的抽象词 r_{w_i} 替换词 w_i

end for

替换示例如图 3 所示。

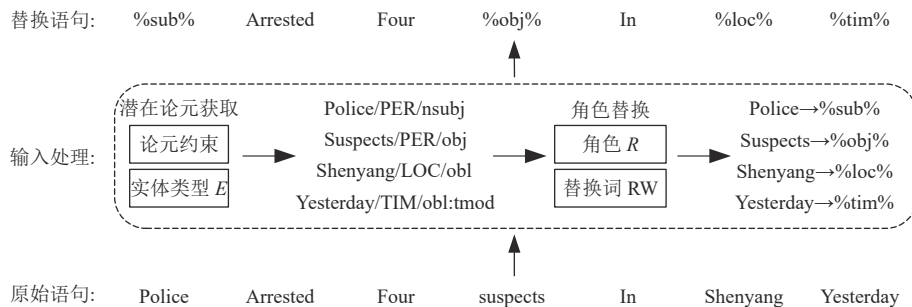


图 3 潜在论元角色替换

2.3 基于 Bi-LSTM 与注意力的深度学习检测

建立基于 Bi-LSTM 与类型注意力的深度学习模型，将经角色抽象后的句子 S' 的语义、语法、模式信息抽象为高维特征实现检测，原理如图 1 所示。

2.3.1 输入构建

使用神经网络处理文本需对文字编码。“事件类型”向量通过训练得到稳定表示，预测阶段值固定不变，故采用静态词向量。

Word2Vec 是最常用的静态词向量获取方法，其中 Skip-gram 方法对生僻词效果更好，选用此方法基于大规模文本训练生成词向量表，查询向量表即可获得词向量表示 X_w 。

使用 Word2Vec 训练得到的词向量包含一定语

义信息，由于对潜在论元进行了角色抽象，替换后的词表示对应的词向量，同时隐含了一定的事件模式信息。但上述信息未完全覆盖潜在论元信息，未覆盖的信息通过词性及实体类型表达。PTNN 模型的输入在 X_w 上，增加词性 X_p 和实体类型 X_e 来强化表示效果。词 w_i 的嵌入表示为：

$$x_i = X_w \oplus X_e \oplus X_p \quad (1)$$

式中， \oplus 表连接操作，词性使用 Stanford NLP 工具获取，每个词性对应一条向量，训练初随机初始化，通过训练得到其稳定表示 X_p 。实体类型向量的处理方式与词性向量相同，表示为 X_e 。

2.3.2 事件类型向量

同类事件有相似类型特征，这种特征除通过触发词表征，在事件的特定属性论元及论元的实体类

型中也有反映。如某些事件必须拥有多个参与者，某些事件的实体有类型限制等，如表1所示。PTNN模型在不识别触发词的前提下，通过学习这类信息获取事件类型特征。

模型通过词级类型向量 T_{loc} 和句级类型向量 T_{glo} 表示事件类型。 T_{loc} 学习事件句中词实体特征和事件类型的隐含联系， T_{glo} 学习事件句整体特征和事件类型的隐含联系。具体地，随机初始各事件类型向量 T_{loc} 和 T_{glo} ，使用标记事件类型的训练语料对其进行训练，迭代至 T_{loc} 和 T_{glo} 稳定。

2.3.3 注意力机制

注意力机制可以使神经网络更关注所需信息。在不利用触发词特征进行事件检测的情况下，使用注意力机制能协助获取句子的类型特征。

T_{loc} 建模词级类型信息，将其作为注意力机制的“查询向量” Q ，计算各时间步输出与 Q 的关系，建模类型信息。 t 时刻注意力分数为：

$$\alpha_t = \frac{\exp(\mathbf{h}_t \mathbf{T}_{loc}^T)}{\sum_i \exp(\mathbf{h}_i \mathbf{T}_{loc}^T)} \quad (2)$$

式中， \mathbf{h}_t 为Bi-LSTM在 t 时刻的隐层向量，当前输入包含类型信息越多则 α_t 越大，其信息在最终特征向量中的占比越大，进而实现事件类型识别。

2.3.4 Bi-LSTM层

Bi-LSTM模型能获取长时依赖并兼顾上下文信息，PTNN将其作为特征提取器，获取输入的词级特征和整体语义两个维度的信息。

对词级特征，拼接同位置的正、反向LSTM隐层向量作为当前时间的特征表示 \mathbf{h}_t ，形成词级特征矩阵 \mathbf{H} ：

$$\mathbf{h}_t = \vec{h}_t \oplus \overleftarrow{h}_{n-t} \quad (3)$$

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n] \quad (4)$$

式中， t 表示时间步； n 表示输入句长度。计算各时间步的词特征对 T_{loc} 的注意力值 α_t ，得到注意力向量 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]$ 。

综合隐层向量 \mathbf{H} 、词级注意力向量 α 与词级类型向量 T_{loc} 得到词级特征 X_{loc} ：

$$X_{loc} = (\alpha^T \mathbf{H}) T_{loc} \quad (5)$$

对于句级特征，拼接正、反向LSTM的最后一个单元的隐层输出作为输入的整体向量表示：

$$\mathbf{h}_s = \vec{h}_n \oplus \overleftarrow{h}_n \quad (6)$$

综合句表示 \mathbf{h}_s 、句级类型向量 T_{glo} 得到句级特征 X_{glo} ：

$$X_{glo} = \mathbf{h}_s \mathbf{T}_{glo}^T \quad (7)$$

X_{loc} 与 X_{glo} 分别描述了输入在词、句级上的事件特征，对两部分特征按权加和作为输入句对某类事件的特征得分，即：

$$O = \sigma(\gamma X_{loc} + (1 - \gamma) X_{glo}) \quad (8)$$

式中， γ 为词句信息比例系数，用于调节最终输出中词级信息与句级信息的占比。

损失函数方面，任务本质是判断输入 S 是否包含 T 类型事件，不同 T 对应的向量空间不同，不能以输出直接构筑二元交叉熵作为损失，故选取均方损失(mean square loss, MSE)，同时引入 L_2 正则避免过拟合。由于正负类样本不均衡，加大对正例损失的惩罚，损失函数为：

$$L_\theta = (o(x) - y)^2 (1 + y\beta) + \rho \|\theta\|^2 \quad (9)$$

式中， β 为惩罚因子， y 为0时其保持沉默， y 为1时对当前样本的损失进行惩罚； ρ 为 L_2 正则系数； θ 为所有参与反向传播的参数集合。预测结果 \tilde{y} 与模型的输出 $o(x)$ 的对应关系为：

$$\tilde{y} = \begin{cases} 0 & o(x) < 0.5 \\ 1 & \text{其他} \end{cases} \quad (10)$$

3 实验结果及分析

3.1 实验设置

模型基于Pytorch框架，使用python3.6进行编码。词性及依存关系使用Stanford CoreNLP工具包获取。

采用ACE2005英文数据的timex2norm文件夹作为语料。其包含599篇标注语料，定义了33类事件及35类论元。从所有类别中随机选取30篇作为验证集，从nw文件夹(通讯社新闻)随机选取40篇作为测试集，其余529篇为训练集。以句为基本单位对语料进行处理，同句内多个同类事件记录一个标签，不同类事件记录多个标签。详细统计如表3所示。

表3 语料详细信息

数据集	文档数	语句数	事件句数	事件数
训练集	529	12855	2929	3376
验证集	30	826	322	392
测试集	40	634	278	347
总计	599	14365	3529	4117

词向量基于 ACE2005 语料原始文本及 1-billion-word-language-modeling-benchmark-r13 混合训练生成, 维度为 200 维; 词性向量及实体类型向量均为 50 维。 L_2 正则系数为 1×10^{-5} , epoch 为 25, batch_size 设置为 100。

3.2 评价方法

评价指标采用精确率 (precision, P)、召回率 (recall, R) 及 F_1 值 (F_1 -score, F_1), 计算公式分别为:

$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN}, F_1 = \frac{2PR}{P + R}$$

式中, TP 表示识别对的正例数; FN 表示识别错的正例数; FP 表示识别错的负例数。

3.3 实验论证

3.3.1 基线模型

由于不使用触发词, 故采用文献 [20] 定义的基线模型及文献 [20] 提出的 TBNNAM 作横向对比, 如表 4 所示。

表 4 基线模型

基线名称	输入	编码器	句子表示
MC-LSTM _{avg}	<Sentence>	LSTM	隐层均值
MC-LSTM _{last}	<Sentence>	LSTM	LSTM输出
BC-LSTM _{avg}	<Sentence, Type>	LSTM	隐层均值
BC-LSTM _{last}	<Sentence, Type>	LSTM	LSTM输出
TBNNAM	<Sentence, Type>	LSTM	LSTM输出

MC-表示执行多分类任务, 输出所有类型中概率最高的一个; BC-表示执行二分类任务, 判定待测句是否包含指定类型事件。

由于 PTNN 模型存在几个可控变量, 为验证不同条件下的效果, 定义如下子模型:

M_1 : 不进行潜在论元识别与词替换;

M_2 : 进行潜在论元识别, 不进行词替换;

M_3 : 进行潜在论元识别, 进行词替换 (即 PTNN)。

3.3.2 词句信息比例系数 γ 对结果的影响

信息比例系数即参数 γ 决定了输出中词、句级评分占比。不同信息占比会导致训练侧重方向的变化, 为此在验证集上考察模型 M_3 在不同 γ 值下的表现, 实验结果如图 4 所示。可知, 模型在 γ 取 0.25 时效果最佳。

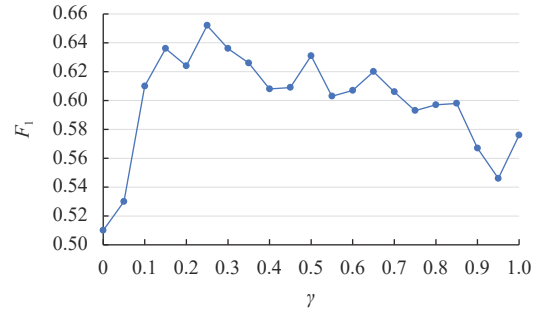


图 4 不同 γ 值对验证集的效果

3.3.3 基线模型与 M_1 的效果对比

M_1 为不使用潜在论元识别与词替换的子模型, 通过其与基线模型对比, 证明双向信息的有效性。实验在测试集上进行, 取 5 轮结果的均值以减小随机性的影响, 结果如表 5 所示。

表 5 基线模型与 M_1 效果对比

方法	P	R	F_1
MC-LSTM _{avg}	0.552	0.421	0.478
MC-LSTM _{last}	0.561	0.416	0.477
BC-LSTM _{avg}	0.563	0.597	0.579
BC-LSTM _{last}	0.589	0.615	0.602
TBNNAM	0.707	0.602	0.65
子模型 M_1	0.708	0.636	0.67

结果表明 M_1 取得了比文献 [20] 中基线模型更好的效果。证明在依靠实体信息进行检测时, 双向信息比单向信息效果更好。

3.3.4 潜在论元识别与类替换效果实验

为验证潜在论元及类替换的有效性, 在测试集上采用相同的参数, 比较 M_1 , M_2 , M_3 的效果, 实验结果如表 6 所示。

表 6 各子模型效果对比

方法	P	R	F_1
TBNNAM	0.707	0.602	0.65
子模型 M_1	0.708	0.636	0.67
子模型 M_2	0.724	0.647	0.683
子模型 M_3 (PTNN)	0.72	0.654	0.685
Nguyen's CNN	0.718	0.664	0.690
Liu's PSL	0.756	0.636	0.691
DS-DMCNN	0.757	0.660	0.705

表 6 后 3 行为基于触发词方法在相同测试集上取得的效果^[12,14-15]。

由结果可知, 潜在论元识别和词替换两个步骤逐步提升了检测效果, 原因在于通过精简实体信息, 减少了对事件无关实体信息的关注, 提升了对事件模式的表达效果; 同时词替换操作使得在表达

剩余实体时更多的关注了其类别信息, 避免了无意义的语义分散, 提升了后续特征提取的效果。

本文方法最终达到了使用触发词特征达到的检测水平。验证了基于事件模式特征的 PTNN 模型对事件检测任务的有效性。

4 结束语

实验证明, PTNN 模型在不依靠触发词的情况下, 仅依靠事件模式信息和潜在论元信息达到了触发词模型的检测水平, 同时其表现优于同类模型。验证了模式信息对事件检测的意义, 并验证了其在深度学习模型中的可用性。

后续拟进一步探究如何更好的表示事件模式特征, 同时探究如何进行事件模式与类型的联合识别。

参 考 文 献

- [1] AHN D. The stages of event extraction[C]//Proceedings of the Workshop on Annotating and Reasoning about Time and Events. Sydney: Association for Computational Linguistics, 2006: 1-8.
- [2] JI H, GRISHMAN R. Refining event extraction through cross-document inference[C]//Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics. Columbus: Association for Computational Linguistics, 2008: 254-262.
- [3] CHEN Z, JI H. Language specific issue and feature exploration in Chinese event extraction[C]//Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers. [S.l.]: Association for Computational Linguistics, 2009: 209-212.
- [4] LIAO S, GRISHMAN R. Using document level cross-event inference to improve event extraction[C]//Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. [S.l.]: Association for Computational Linguistics, 2010: 789-797.
- [5] 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). Sofia: Association for Computational Linguistics, 2013: 73-82.
- [6] 张建恒, 黄蔚, 胡国超. 基于 LDA 模型和 AP 聚类的主题事件抽取技术[J]. *计算机与现代化*, 2017(12): 77-81.
ZHANG J H, HUANG W, HU G C. Topic event extraction technology based on LDA model and AP clustering method[J]. *Computer and Modernization*, 2017(12): 77-81.
- [7] 高源, 席耀一, 李弼程. 基于依存句法分析与分类器融合的触发词抽取方法[J]. *计算机应用研究*, 2016(5): 1407-1410.
GAO Y, XI Y Y, LI B C. Trigger extraction algorithm based on dependency parsing and classifier fusion[J]. *Application Research of Computers*, 2016(5): 1407-1410.
- [8] 万齐智, 万常选, 胡蓉, 等. 基于句法语义依存分析的中文金融事件抽取[J]. *计算机学报*, 2021, 44(3): 508-530.
WAN Q Z, WAN C X, HU R, et al. Chinese financial event extraction base on syntactic and semantic dependency parsing[J]. *Chinese Journal of Computers*, 2021, 44(3): 508-530.
- [9] CHEN Y, LIU S, HE S, et al. Event extraction via bidirectional long short-term memory tensor neural networks[M]//Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. [S.l.]: Springer, 2016: 190-203.
- [10] WU Y, ZHANG J. Chinese event extraction based on Attention and semantic features: A bidirectional circular neural network[J]. *Future Internet*, 2018, 10(10): 95.
- [11] LIN H, LU Y, HAN X, et al. Nugget proposal networks for Chinese event detection[EB/OL]. [2021-10-11]. <https://arxiv.org/pdf/1805.00249.pdf>.
- [12] FENG X, QIN B, LIU T. A language-independent neural network for event detection[J]. *Science China Information Sciences*, 2018, 61(9): 1-12.
- [13] 田梓函, 李欣. 基于 BERT-CRF 模型的中文事件检测方法研究[J]. *计算机工程与应用*, 2021, 57(11): 135-139.
TIAN Z H, LI X. Research on Chinese event detection method based on BERT-CRF model[J]. *Computer Engineering and Applications*, 2021, 57(11): 135-139.
- [14] 陈安南, 叶岩宁, 王畅畅, 等. 基于 BERT-DGCNN 的中文事件抽取方法研究[J]. *计算机科学与技术*, 2021, 11: 1572.
CHEN A N, YE Y N, WANG C C, et al. Research on Chinese event extraction method based on BERT-DGCNN[J]. *Computer Science and Application*, 2021, 11: 1572.
- [15] ZHOU D, ZHANG X, HE Y. Event extraction from Twitter using non-parametric Bayesian mixture model with word embeddings[C]//Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. [S.l.]: Association for Computational Linguistics, 2017: 808-817.
- [16] YUAN Q, REN X, HE W, et al. Open-schema event profiling for massive news corpora[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. [S.l.]: Association for Computing Machinery, 2018: 587-596.
- [17] PETRONI F, RAMAN N, NUGENT T, et al. An extensible event extraction system with cross-media event resolution[EB/OL]. [2021-11-10]. <https://aclanthology.org/E17-1076.pdf>.
- [18] 刘振. 基于网络科技信息的事件抽取研究[J]. *情报科学*, 2018, 36(9): 115-117.
- [19] LIU Z. Research on event extraction from networks scientific information[J]. *Information Science*, 2018, 36(9): 115-117.
- [19] YANG H, CHEN Y, LIU K, et al. Defee: A document-level chinese financial event extraction system based on automatically labeled training data[C]//Proceedings of ACL 2018, System Demonstrations. Melbourne: Association for Computational Linguistics, 2018: 50-55.
- [20] LIU S, LI Y, ZHANG F, et al. Event detection without triggers[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). Minnesota: Association for Computational Linguistics, 2019: 735-744.
- [21] XU W, ZHANG W, WANG D. Event detection without trigger words on movie scripts[C]//2020 International Conference on Image, Video Processing and Artificial Intelligence. [S.l.]: SPIE, 2020, 11584: 115841G.