



低资源事件抽取关键技术研究

报告人：全美涵（2017310729）

导 师：许斌

日 期：2022年05月21日



内容大纲

- ◎ 研究背景和意义
- ◎ 基于多模态增强的事件检测方法
- ◎ 基于外部知识注入的事件检测方法
- ◎ 基于自标注的事件要素抽取方法
- ◎ 事件抽取在NewsMiner系统中的应用
- ◎ 总结



内容大纲

- ◎ 研究背景和意义
- ◎ 基于多模态增强的事件检测方法
- ◎ 基于外部知识注入的事件检测方法
- ◎ 基于自标注的事件要素抽取方法
- ◎ 事件抽取在NewsMiner系统中的应用
- ◎ 总结



研究背景

- 事件反映了世界的动态变化，是人们认知世界的主要手段，也是互联网信息的基本组成单元。

“整个宇宙就是由各种事件互连接、互相包涵而形成的有机系统”

——英国数学家、哲学家和教育理论家怀特海

Process and Reality(MIT press2004)

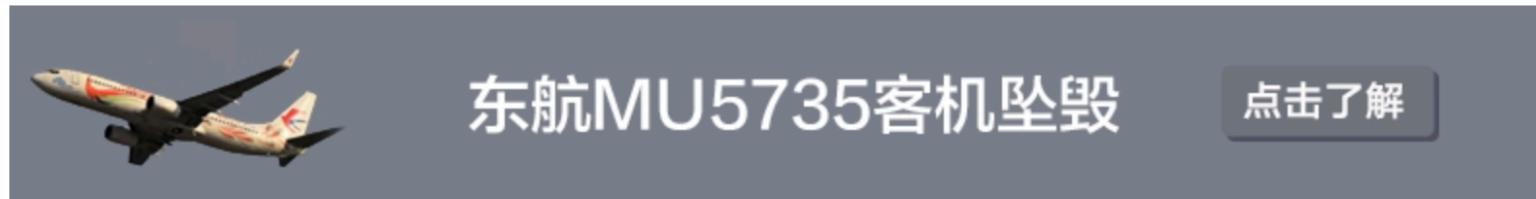


- 人们天然有获取事件信息的诉求



研究背景

- 事件信息需要简练、精准并且涵盖全面



中文名	MU5735	机 龄	6.8年
所属公司	中国东方航空集团有限公司	事故词条	3·21东航客机事故
注册号	B1791	机 型	B737-800NG

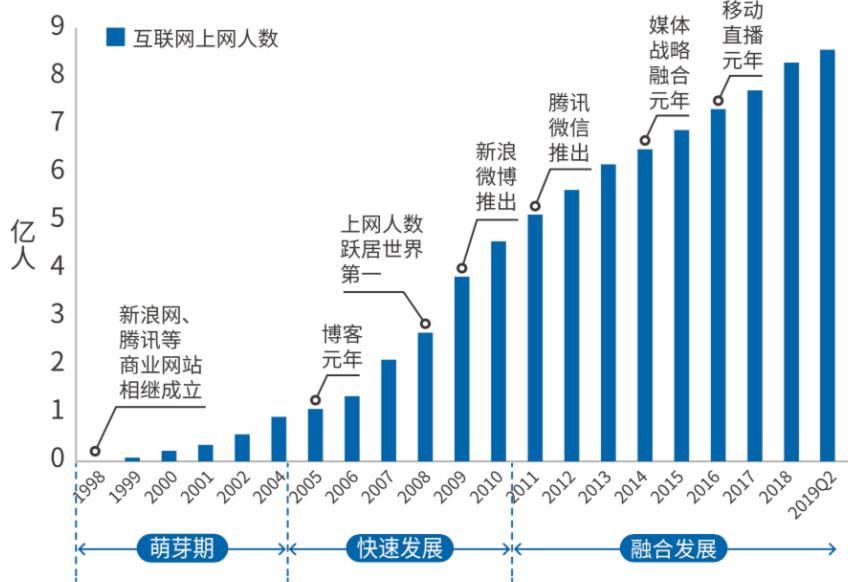
目录	1 客机信息	<ul style="list-style-type: none">▪ 机组人员▪ 事故哀悼▪ 黑匣子▪ 遥感影像▪ 调查进展	<ul style="list-style-type: none">▪ 新闻发布会▪ 波音中国▪ 中国东航▪ 美国联邦航空管理局▪ 美国国家运输安全委员会▪ 印度▪ 中兴华会计师事务所	<ul style="list-style-type: none">▪ 中国民航报▪ 中国银保监会▪ 梧州市社会组织
	2 发生事故	<ul style="list-style-type: none">▪ 事故背景▪ 事故经过▪ 救援情况▪ 官方工作▪ 家属安置		
	3 事故回应	<ul style="list-style-type: none">▪ 民航局		
	4 造传谣处置			
	5 事故理赔			





研究背景

- 近年来，用户倾向于从“网络媒体”中获取事件信息
- 海量的新闻面临信息过载问题，难以仅凭人力整理



任务定义

◎ 事件抽取技术应运而生

- ◆ 事件抽取是自然语言处理领域一个重要的研究方向
 - 事件检测：识别事件触发词并判断事件类型
 - 事件要素抽取：抽取参与事件要素

社会民主党人肖尔茨于12月08日当选德国总理

↓

事件触发词	当选 (触发“就职”事件)	
事件要素	就职者	肖尔茨
	职位	德国总理
	时间	12月08日

图1 事件抽取任务示例

7 ¹<http://www.ldc.upenn.edu/Projects/ACE/>



研究意义

丰富现有知识图谱

- ◆ 多关注于以实体为核心的静态知识
- ◆ 缺乏以事件为核心的动态知识

名称	创建时间	数据来源	数据规模
OpenCyc ^①	1984	专家知识	23 万实体, 未定义事件
WordNet ^②	1985	专家知识	15 万实体, 未定义事件
YAGO ^③	2007	WordNet+Wikipedia	459 万实体, 未定义事件
DBpedia ^④	2007	Wikipedia+ 专家知识	1694 万实体, 8 万事件
Freebase ^⑤	2008	Wikipedia+ 领域知识 + 集体智慧	5872 万实体, 2 万事件





研究意义

◎ 支撑语义搜索

- ◆ 事件抽取能够帮助搜索引擎给出更加精确的回答

Google 汶川地震死了多少人 X |

↓

5·12汶川地震伤亡人数：
69227人遇难，374643人受伤，17923人失踪
5·12汶川地震（2008 Sichuan earthquake），发生于北京时间
(UTC+8)2008年5月12日(星期一)14时28分04秒，根据中华人民共
和国地震局的数据，此次地震的面波震级 里氏震级... [详情>>](#)

来自百度百科



研究意义

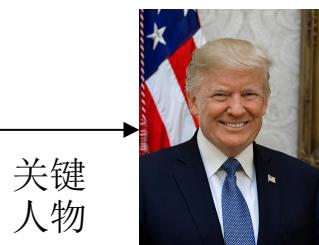
◎ 支撑事件因果分析

- ◆ 事件抽取能够为事件因果关系分析提供更多的语义证据

2021年美国国会大厦遭冲击事件



2021年美国国会大厦遭冲击事件（英语：2021 Storming of United States Capitol）是指2021年1月6日，时任美国总统唐纳德·川普的支持者暴力闯入美国国会大厦的骚乱事件。当天，2000至2500名^[32]示威者在特朗普的号召下冲进了美国国会大厦，扰乱正在进行计票（英语：2021 United States Electoral College vote count）及认证2020年美国总统选举结果以宣告乔·拜登正式获胜的美国国会联席会议^{[4][42]}。国会大厦建筑群（英语：United States Capitol Complex）被封锁，其中的立法者和工作人员被紧急疏散，之后被骚乱人群占领并破坏长达数个小时。一系列的冲突导致5人死亡，超过140人受伤^[43]。



关键人物

关键人物

导致

第二次唐纳德·特朗普弹劾案



美国众议院投票弹劾议案

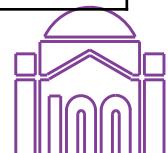
2021年1月13日，美国众议院通过弹劾议案，正式以煽动叛乱（incitement of insurrection; sedition）的罪行弹劾美国第45任总统唐纳德·特朗普。特朗普是美国历史上第一位被弹劾两次的总统（也是美国历史上第一位被弹劾两次的官员）^[1]。此次弹劾是因为特朗普尝试推翻2020年美国总统选举结果、曾向佐治亚州州务卿施压造票以及一周前（1月6日）涉入煽动造成2021年冲击美国国会大厦事件的行为^[2]，这也是总统史上最快弹劾成案的例子。



研究现状

◎ 传统模式识别阶段和机器学习阶段

方法	核心思路	代表工作	缺点
模式识别	句子解析 利用自然语言处理工具， 获得句法词法信息	Ellen et al. ^[1] Yangarber et al. ^[2]	维护精细的事件模式费时费力
	词典构建 利用最大熵模型或专家知识领域词典		
	事件模式设计 从语法树或正则表达式中总结规则		
机器学习	特征工程 词法特征，句法特征，触发词特征 和上下文特征	Li et al. ^[3] Björne et al. ^[4]	面临误差传递，手工构建特征，迁移性差，泛化性不强
	分类算法 支持向量机 (SVM)、最大熵 (ME)、 朴素贝叶斯 (NB)、条件随机场 (CRF)、整数逻辑规划 (ILP)、层次 凝聚聚类 (HAC)		



研究现状

◎ 深度学习阶段

方法	核心思路		代表工作	缺点
基于语义增强的事件抽取方法	基于篇章语义	引入篇章上下文信息，为事件抽取提供更多的语义证据。	Duan et al. [5] Zhao et al. [6]	1、局限于文本模态消歧，没有有效利用多模态信息 2、缺乏任务相关的常识知识监督 3、没有充分利用已有的训练语料
	基于句法语义	引入句法依存图信息，拉近关键信息的距离。	Sha et al. [7] Cui et al. [8]	
基于数据增强的事件抽取方法	基于多任务学习	引入任务相关语料，联合训练	Lu et al. [9] Yang et al. [10]	1、局限于文本模态消歧，没有有效利用多模态信息 2、缺乏任务相关的常识知识监督 3、没有充分利用已有的训练语料
	基于半监督学习	从大规模未标注语料中(通常是新闻中)挖掘更多的训练数据	Ferguson et al. [11] Huang et al. [12]	
	基于远程监督学习	利用外部知识库来扩展事件抽取的训练数据	Liu et al. [13] Reschke et al. [14]	



存在的挑战

- ◎ 深度学习模型需要大规模、高质量标注
- ◎ 获取大规模的标注数据不现实
 - ◆ 标注数据费时费力，还需要大量财力支撑
 - ◆ 新的事件类型不断出现，无法对每个新类都大量的标注

名称	时间	组织者	数据规模
MUC-4	1987-1997	DARPA	1个事件类， 1,700篇 英文新闻
ACE	2000-2010	NIST	33个事件类， 599篇 英文新闻
KBP	2015-2017	NIST	18个事件类， 505篇 英文新闻
CEC	2009	Shanghai University	5个事件类， 332篇 中文新闻
M ² E ²	2020	University of Illinois at Urbana-Champaign,	8个事件类， 245篇 多模态新闻



存在的挑战

◎ 应对措施

◆ 研究数据稀缺的低资源场景下的事件抽取

“深度学习模型应该用较少的数据得出准确的结论”

——谷歌大脑的联合创始人吴恩达



“希望人工智能系统更加灵活、更加健壮，不需要输入海量的原始数据”

——Meta AI研究中心主任Joelle Pineau





存在的挑战

- 低资源场景下现有的深度学习事件抽取方法表现不佳。

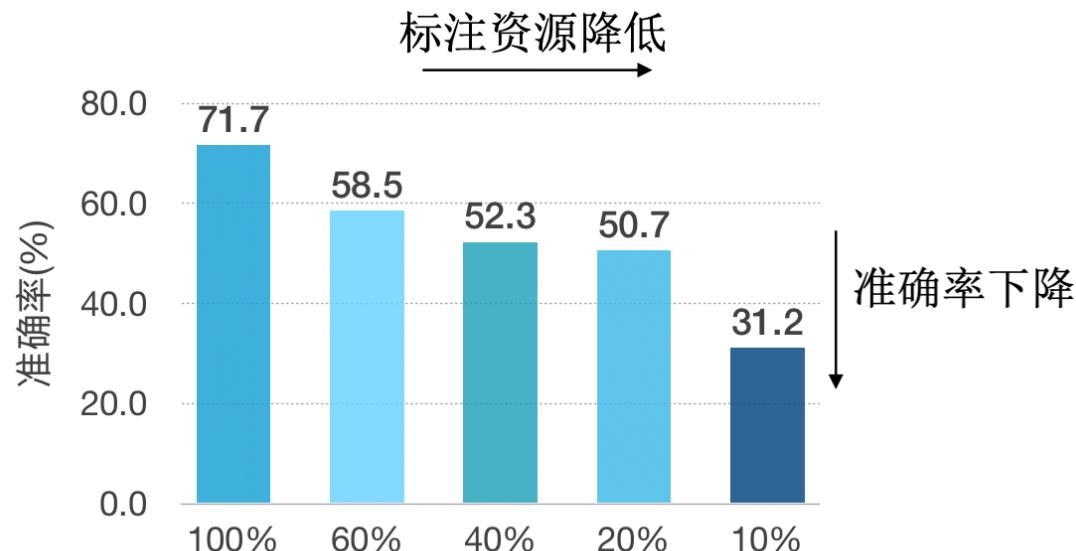


图3 DMBERT在ACE2005数据集上的表现

- 仅有 10% 标注数据的低资源场景下，准确率只有 31.2%，
- 远低于其在标注充足场景下的性能表现(准确率 71.7%)



存在的挑战

◎ 事件类型消歧难

- ◆ 事件触发词的歧义性很大，经常面对同词触发不同事件类型的情况
- ◆ 在低资源场景下，现有的事件抽取方法容易过度拟合训练集中出现的触发词

S1: Ford **confront**_{Meet} members in council chamber

S2: Police **confront**_{Attack} protesters hurling stones

图4 事件类型消歧难



存在的挑战

◎ 事件召回难

- ◆ 由于自然语言的灵活性，一个事件往往可以被多个事件触发词触发
- ◆ 容易过拟合训练数据，只能召回常见的事件触发词触发的事件，而无法召回不常见的事件触发词触发的事件

S1: Iraq terrorist **fired**_{Attack} towards our position

S2: A man was **hacked**_{Attack} to death by the criminal

图5 事件召回难



存在的挑战

◎ 事件要素识别难

- ◆ 自然语言非常灵活，可以用多种说法描述参与一个事件的事件要素的发生
- ◆ 训练语料单一，只能识别固定的几种说法

“<Place> is the birthplace of <People>”



“<People> is born in <Place>”

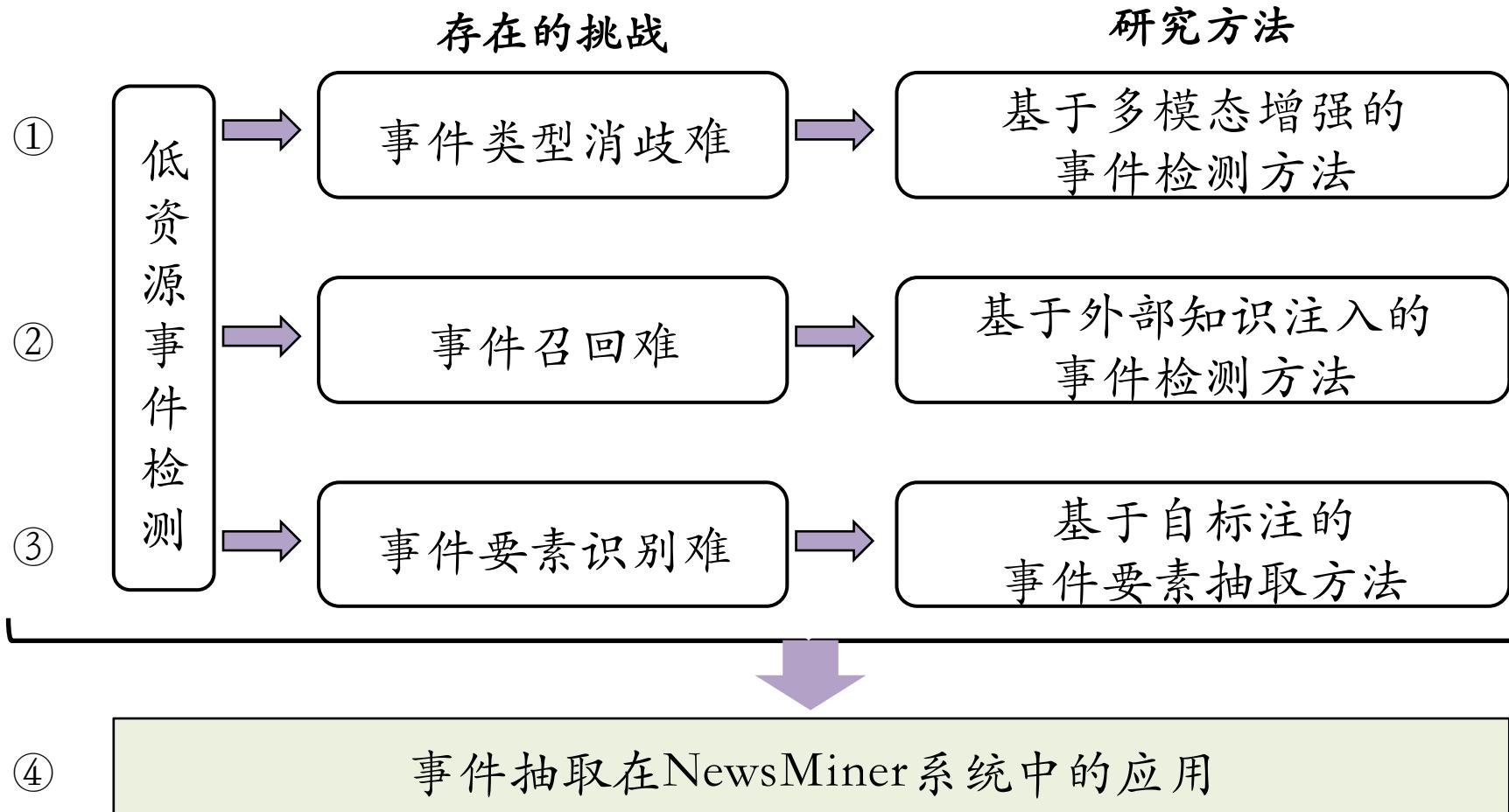


图6 事件要素识别难





研究框架



内容大纲

- ◎ 研究背景和意义
- ◎ 基于多模态增强的事件检测方法
- ◎ 基于外部知识注入的事件检测方法
- ◎ 基于自标注的事件要素抽取方法
- ◎ 事件抽取在NewsMiner系统中的应用
- ◎ 总结



基于多模态增强的事件检测方法

- ◎ 低资源场景下，事件检测面临消歧难的问题
- ◎ 图片信息在事件类型消歧中扮演着重要的角色
 - ◆ 提供难以用文本描述的信息，如场景、人物动作辅助消歧
 - ◆ 反映文本的核心事件，为事件类型定调



S1: Ford confront Meet
members in council chamber

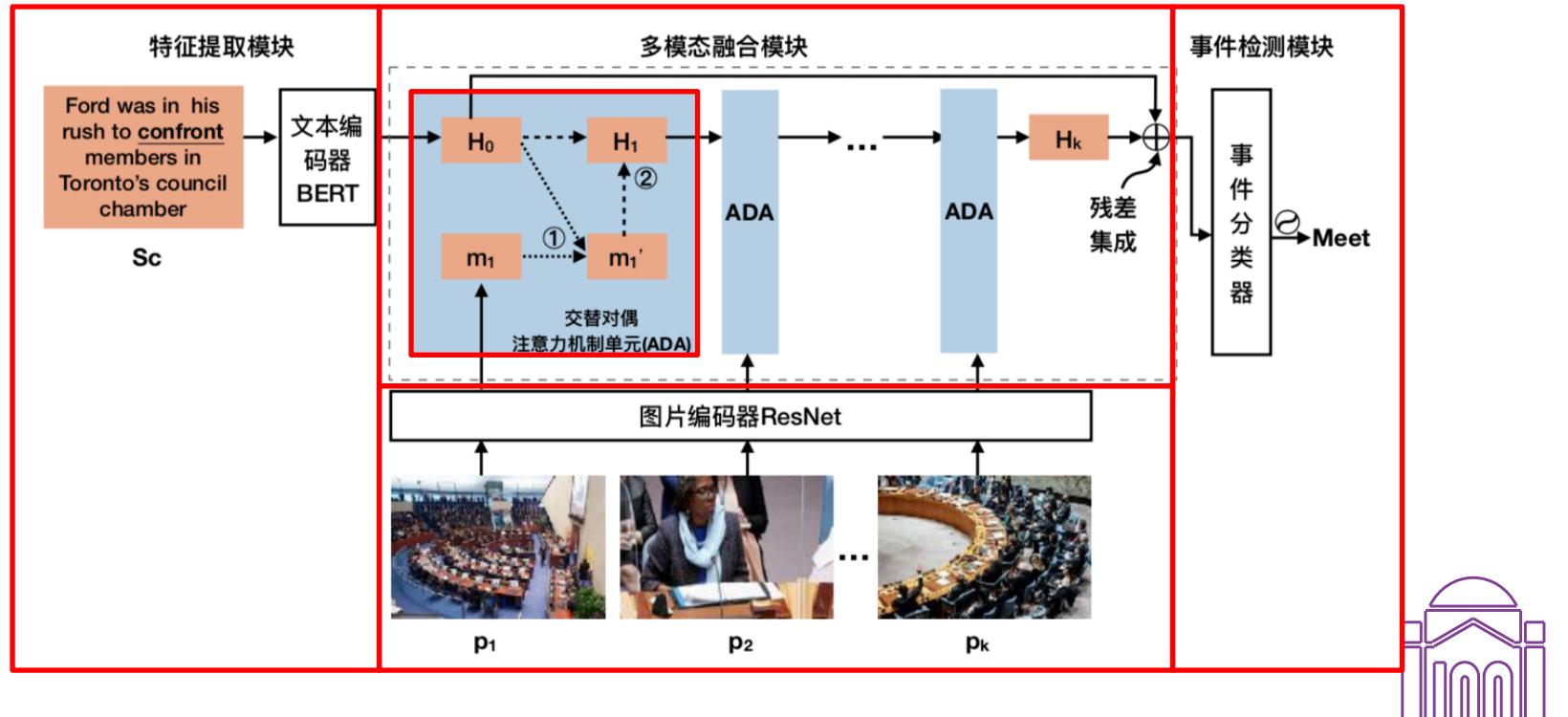
S2: Police confront Attack
protesters hurling stones



DRMM模型-多模态融合

◎ 如何利用图像信息消歧?

- ◆ 提出一个循环对偶多模态事件检测模型 (DRMM)
- ◆ 给定新闻中的句子 $S = \langle w_1, w_2, \dots, w_n \rangle$ 及其该新闻的多张配图 $\Delta = \{p_1, p_2, \dots, p_k\}$, 事件检测旨在最大化概率 $\frac{1}{n} \sum_{i=1}^n P(y_i | w_i, \Delta)$



DRMM模型-多模态融合

◎ 交替对偶注意力机制 (ADA)

- ◆ 分辨图像中有用的信息，剔除无关信息的干扰

- ◎ 根据文本语义筛选图

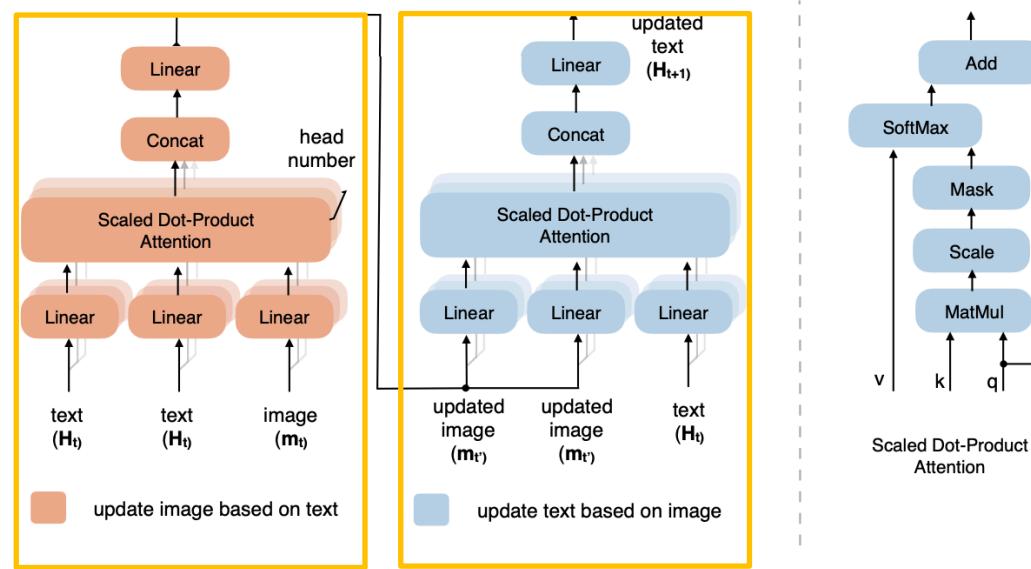
片重点区域

$$m_t' = \Omega(m_t, H_t)$$

- ◎ 根据图片语义筛选文

本语义

$$H_{t+1} = \Omega(m_t', H_t)$$



实验设置

◎ 数据集

- ◆ ACE2005语料库包含8个事件类型33个事件子类型，共599篇数据。
- ◆ 按照 529/30/40 的比例划分训练、验证和测试集(Chen et al. 2015)。
- ◆ 图像语料库包含2815张新闻配图

◎ 评价指标

- ◆ Precision、Recall以及Micro-F1值

◎ 基线模型

- ◆ 融合图片信息：VAD
- ◆ 融合上下文信息：DLRNN, GMLATT, HBTNGMA
- ◆ 融合更多数据信息、：AD-DMBERT

统计名目	数目
图片总数	2815
每篇文章平均图片数	4.7
每篇文章最多图片数	6
每篇文章最少图片数	3

表1 图像数据集统计数据



实验结果

◎ 整体实验效果

Method	Precision	Recall	F1
VAD	75.1	64.3	69.3
DLRNN	77.2	64.9	70.5
ANN-FN	77.6	65.2	70.7
GMLATT	78.9	66.9	72.4
HBTNGMA	77.9	69.1	73.3
AD-DMBERT	77.9	72.5	75.1
DRMM(Our)	77.9	74.8	76.3

- ◆ 多模态比单一模态（DLRNN、AD-DMBERT等）性能要好。
- ◆ 融入图片信息也要慎重，VAD加入没有和事件直接相关的实体图片信息，效果就并不好。

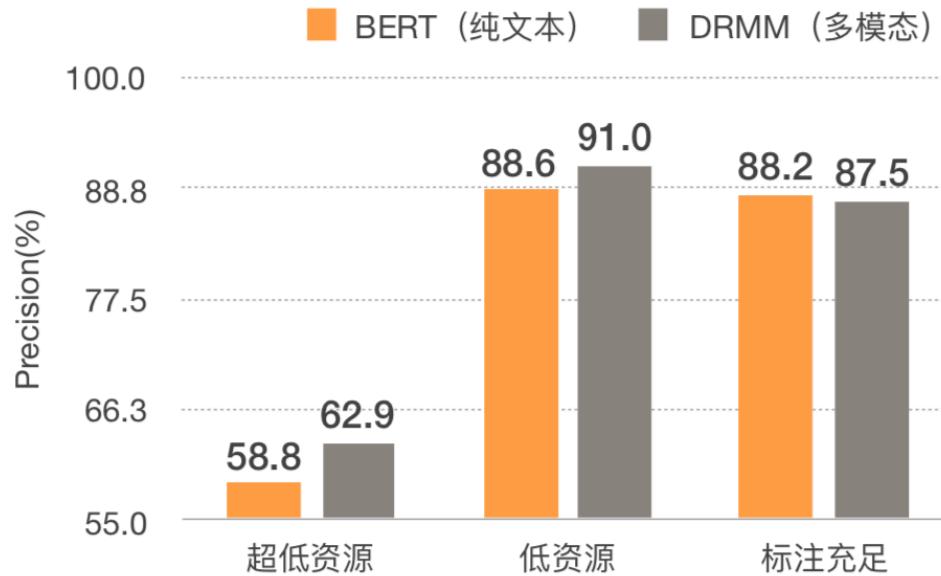




实验结果

◎ 低资源场景下的表现

- ◆ 超低资源：训练数据不超过1条
- ◆ 低资源：训练数据不超过10条
- ◆ 标注充分：训练数据大于100条



小结

- 我们提出了一种新颖的多模态融合事件检测方法，解决了低资源场景下事件消歧难的问题
- 构建图像数据集，为后续多模态事件检测的研究提供数据支撑。
- 实验证明融合图像模态信息对低资源场景下事件消歧有显著的效果。
- 相关成果：**Meihan Tong, Shuai Wang, Yixin Cao, Bin Xu, Juaizi Li, Lei Hou, Tat-Seng Chua. Image Enhanced Event Detection in News Articles. AAAI 2020**



内容大纲

- ◎ 研究背景和意义
- ◎ 基于多模态增强的事件检测方法
- ◎ 基于外部知识注入的事件检测方法
- ◎ 基于自标注的事件要素抽取方法
- ◎ 事件抽取在NewsMiner系统中的应用
- ◎ 总结



基于外部知识注入的事件检测方法

- 低资源场景下，事件召回难
- 仅能识别训练常见的触发词，无法识别训练未见触发词
- 开放域触发词知识能帮助模型从语义角度识别事件触发词，提升事件召回的可能性。

S1: Now we 're **hearing** the **boom** of Iraqi guns as they **fire**_{Attack} towards our positions .

S2: Troops were **trying** to **break** up **stone-throwing** protests , but did not **use** live **fire**_{Attack}.

S3: A man was **hacked**_{Attack} to death by the criminal

S4: The **intifada**_{Attack} **exploded** in September 2000

“训练中常见”的触发词

“训练中未见”的触发词

■ “开放域触发词知识”识别的触发词

EKD模型

如何融合开放域触发词知识？

- ◆ 提出一个丰富知识蒸馏模型（EKD）

事件检测模块

有监督学习

S_5 : Troops were trying to break up stone -throwing protests, but not use live fire.

特征提取

事件分类

... Attack ...

知识蒸馏模块

弱监督学习

S_6^+ A man was hacked to death by the criminal

知识获取

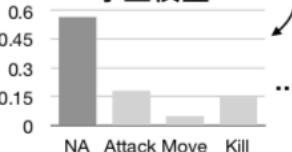
S_6 A man was hacked to death by the criminal

知识融合

教师模型



学生模型



引导学生
模型结果

... 0.023 ...

◆ 联合学习 $J(\theta) = \frac{J_L(\theta)}{\text{有监督}} + \frac{\lambda * J_T(\theta)}{\text{弱监督}}$





EKD模型-事件检测模块

- 给定句子 $S = \langle w_1, w_2, \dots, w_n \rangle$, 事件分类旨在最大化分类概率 $P(Y_i | S_i)$, 其中 Y_i 属于预定义的事件类型集合

◆ 特征提取

$$H = BERT(S)$$

◆ 事件分类

$$p(Y_i | S_i, \theta) = \sum_{j=1}^n \frac{\exp(O_{ijc})}{\sum_{c=1}^C \exp(O_{ijc})} / n$$

◆ 训练目标

$$J_L(\theta) = - \sum_{i=1}^{N_L} \log p(Y_i | S_i, \theta)$$



EKD模型-知识蒸馏模块

◎ 知识收集

- ◆ 借助WordNet语义知识库，从synset的粒度判定某个词是否是事件触发词

A man was hacked to death by the criminal

Attack v1

◎ 知识嵌入

- ◆ [B-Tri]，[E-Tri]：开放域触发词的起始位置和终止位置
- ◆ 采用掩码预训练（Mask-Language Model），获得这两个符号的向量表示

S+: A man was [B-Tri] hacked [E-Tri] to death by the criminal



EKD模型-知识蒸馏模块

◎ 知识融合：教师-学生模型

- ◆ 教师模型注入知识(S^+)，学生不注入知识(S^-)
- ◆ 学生达到和教师一样的事件检测认知水平

$$p(Y|S^+\theta) = p(Y|S^-, \theta)$$

- ◆ 训练目标

$$\begin{aligned} J_T(\theta) &= \mathbf{KL}(p(Y|S^+, \theta) || p(Y|S^-, \theta)) \\ &= \sum_{k=1}^{N_L+N_U} p(Y_{(k)}|S_{(k)}^+, \theta) \frac{p(Y_{(k)}|S_{(k)}^+, \theta)}{p(Y_{(k)}|S_{(k)}^-, \theta)} \end{aligned}$$



实验设置

◎ 数据集

- ◆ ACE2005语料库包含8个事件类型33个事件子类型，共599篇数据。
- ◆ 按照 529/30/40 的比例划分训练、验证和测试集(Chen et al. 2015)。

◎ 评价指标

- ◆ Precision、Recall以及Micro-F1值

◎ 基线模型

- ◆ 纯数据驱动： DMCNN, DLRNN
- ◆ 融合事件要素信息： ANN-S2
- ◆ 融合跨语言信息： GMLATT
- ◆ 融合句法树信息： GCN-ED
- ◆ 融合实体和通用语义信息： Lu's DISTILL, TS-DISTILL





实验结果

◎ 整体实验效果

Method	Precision	Recall	F1
DMCNN	75.6	63.6	69.1
DLRNN	77.2	64.9	70.5
ANN-S2	78.0	66.3	71.7
GMLATT	78.9	66.9	72.4
GCN-ED	77.9	68.8	73.1
Lu's DISTILL	76.3	71.9	74.0
TS-DISTILL	76.8	72.9	74.8
AD-DMBERT	77.9	72.5	75.1
DRMM	77.9	74.8	76.3
EKD (Ours)	79.1	78.0	78.6

◎ 知识注入效果

◆ 在测试时需要外部知识的注入吗？

Test Set	P	R	F
without knowledge	78.8	78.1	78.4
with knowledge	79.1	78.0	78.6

◆ 不需要，知识已经蒸馏到模型中



实验结果

◎ 低资源下性能对比

Methods	超低资源			低资源			标注充足		
	P	R	F	P	R	F	P	R	F
DMBERT	66.7	45.9	54.4	74.4	70.7	72.5	84.8	83.5	84.1
DGBTERT	76.5	42.6	54.7	75.7	70.1	72.8	85.9	83.8	84.3
BOOTSTRAP	73.7	45.9	56.6	76.0	71.3	73.6	90.6	83.5	86.9
EKD (ours)	79.0	52.0	62.7	80.8	72.4	76.4	92.5	82.2	87.1

- ◎ 性能提升：超低资源>低资源>标注充分
- ◎ 我们所提知识蒸馏模型>半监督/远程监督模型>监督模型



实验结果

◎ 案例分析

Sentence	GT	Prediction	
		S	S^+
<i>S1: Mr. Caste leaves at 5 A.M. for a train trek to manhattan and does not return util 6 P.M.</i>	Transport	O	Transport
<i>S2: Militants in the region escalate their attacks in the weeks leading up to the inauguration of Nigeria's president.</i>	Start-Position	O	Start-Position
<i>S3: Mr.Mason, who will be president of CBS radio, said that it would play to radio's strengths in delivering local news.</i>	Start-Position	O	Start-Position

- ◆ 开放域触发词确实帮助模型识别罕见的事件触发词，提升事件召回率。 (CBS radio)



小结

- 提出了一个基于WordNet的词义消歧算法，高效的获取开放域触发词知识。
- 提出了一个丰富知识蒸馏模型，同时在标注数据和大规模未标注上汲取开放域触发词知识。
- 实验表明，我们的方法超过了九个同样获得知识增强的强大基线。
- 相关成果：Meihan Tong, Bin Xu, Shuai Wang, Yixin Cao, Lei Hou, Juanzi Li and Jun Xie . Improving Event Detection via Open-domain Trigger Knowledge. **ACL 2020.**



内容大纲

- ◎ 研究背景和意义
- ◎ 基于多模态增强的事件检测方法
- ◎ 基于外部知识注入的事件检测方法
- ◎ 基于自标注的事件要素抽取方法
- ◎ 事件抽取在NewsMiner系统中的应用
- ◎ 总结



基于自标注的事件要素抽取方法

- 低资源场景下训练数据稀缺、单一，事件要素识别难
- 一个直接的解决方法是增加训练数据
- 如何在没有人工标注的情况下增加训练数据呢？
 - ◆ 从训练数据中挖掘更多的任务相关类

● 预定义类 ● 任务相关类

S_1 : Emeneya was born in Paris and died in local hospital

O_1 O_1 O_3

S_2 : Newton is a polymath. He was born in Lincolnshire.

O_2 O_1

S_3 : The professor from the city studies mathematics

O_2 O_3 O_1

出生事件：出生者、出生地



存在的挑战

- 挖掘任务相关类具有挑战性

- 挑战1：区分任务相关类和噪音

● 预定义类 ● 任务相关类

S_1 : Emeneya was born in Paris and died in local hospital 噪音

S_2 : Newton is a polymath. He was born in Lincolnshire.

S_3 : The professor from the city studies mathematics

- 挑战2：缺乏任务相关类的数据标注

- 既没有标注样例
- 也没有类元数据描述



MUCO模型-任务相关类挖掘

◎ 如何挖掘任务相关类?

- ◆ 借助原型网络和预定义类弱监督信号挖掘
- ◆ 工作原理：原型网络迁移性能
 - 如果A类和B类语义相关，原型网络在A类上训练后，能快速识别B类
 - 本质原因是在A类上预训练之后，B类上即使不训练，实例也有空间聚集的趋势

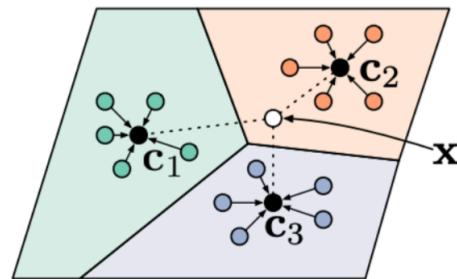


图7 原型网络原理图

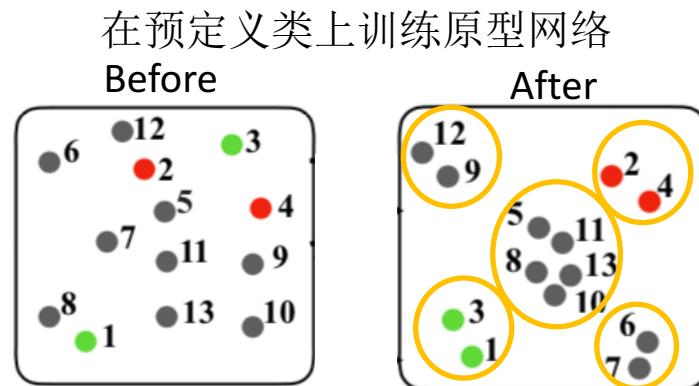


图8 MUCO原理图

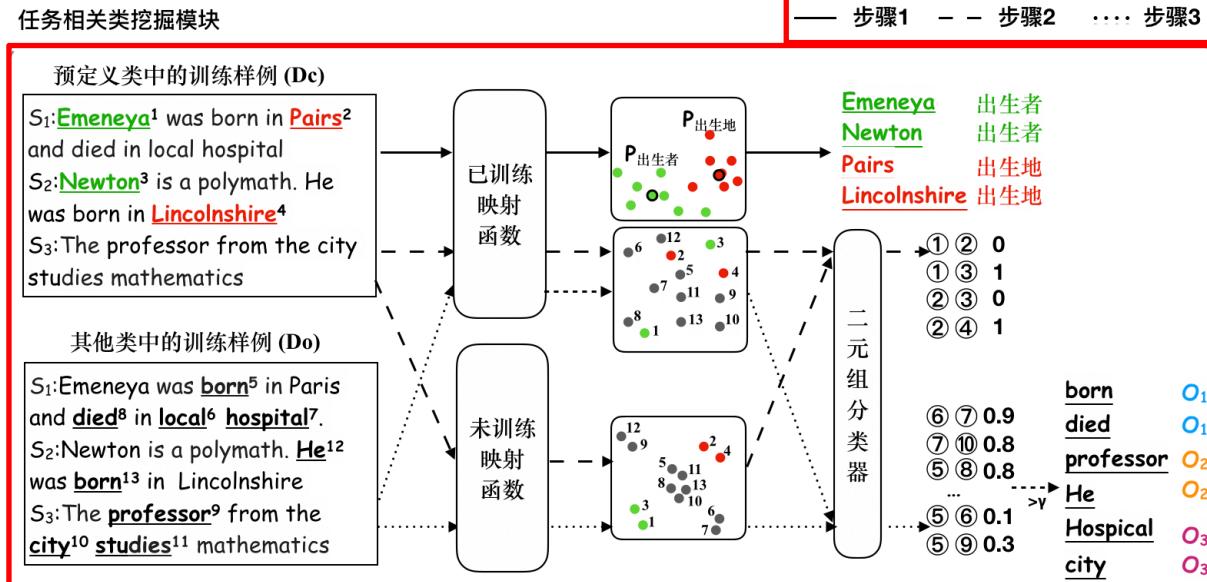




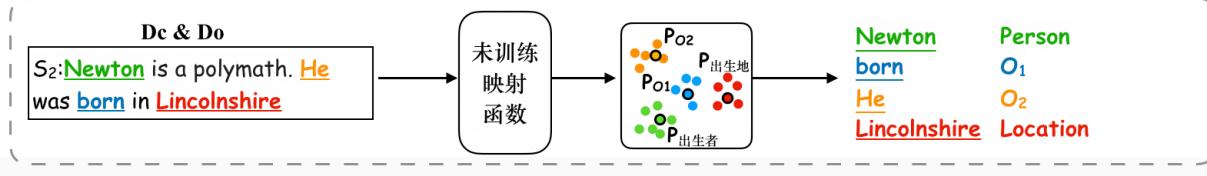
MUCO模型

如何挖掘任务相关类?

◆ 提出一个任务相关类自标注模型 (MUCO)



联合学习模块





MUCO模型-任务相关类挖掘

◎ 第一步：在预定义类上训练原型网络

◆ 原型表示

$$d(x, p_y) = -f_\theta(x)^T p_y$$

◆ 原型学习

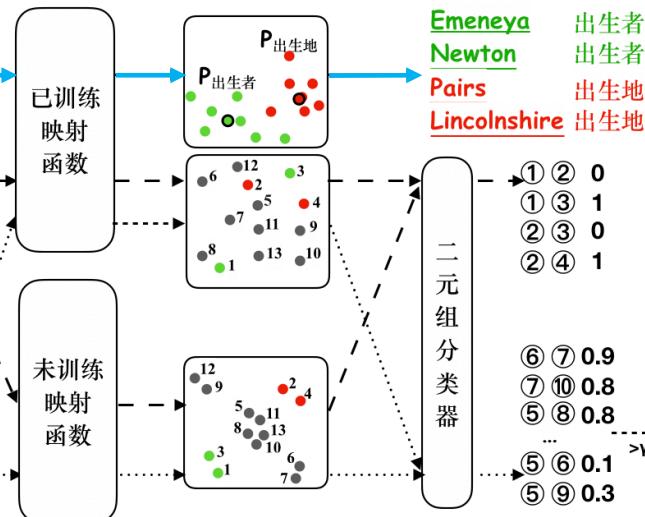
$$L(\theta_1) = -\log \frac{\exp(-d(x, p_y))}{\sum_{p_c \in P_c} \exp(-d(x, p_c))}$$

预定义类中的训练样例 (Dc)

S₁: Emeneya¹ was born in Pairs² and died in local hospital
 S₂: Newton³ is a polymath. He was born in Lincolnshire⁴
 S₃: The professor from the city studies mathematics

其他类中的训练样例 (Do)

S₁: Emeneya was born⁵ in Paris and died⁶ in local⁶ hospital⁷.
 S₂: Newton is a polymath. He¹² was born¹³ in Lincolnshire
 S₃: The professor⁹ from the city¹⁰ studies¹¹ mathematics



MUCO模型-任务相关类挖掘

◎ 第二步：训练二元组分类器

- ◆ 判断预定义类中任意两个点是否属于同一类

$$L(\theta_2) = \frac{1}{N^2} \sum_i^N \sum_j^N (-y_{ij} * \log(b_{ij}) + (1 - y_{ij}) * \log(1 - b_{ij}))$$

- ◆ 空间聚集性特征

$$b_{ij} = W([h_i; h_j; \tilde{h}_i; \tilde{h}_j; |h_i - h_j|; |\tilde{h}_i - \tilde{h}_j|; |h_i - \tilde{h}_i|; |h_j - \tilde{h}_j|]) + b$$

预定义类中的训练样例 (Dc)

S₁: Emeneya¹ was born in Pairs²
and died in local hospital
S₂: Newton³ is a polymath. He
was born in Lincolnshire⁴
S₃: The professor from the city
studies mathematics

已训练
映射
函数

P_{出生地}

Emeneya

出生者

Newton

出生者

Pairs

出生地

Lincolnshire

出生地

①	②	0
①	③	1
②	③	0
②	④	1

其他类中的训练样例 (Do)

S₁: Emeneya was born⁵ in Paris
and died⁸ in local⁶ hospital⁷.
S₂: Newton is a polymath. He¹²
was born¹³ in Lincolnshire
S₃: The professor⁹ from the
city¹⁰ studies¹¹ mathematics

未训练
映射
函数

P_{出生地}

Emeneya

出生者

Newton

出生者

Pairs

出生地

Lincolnshire

出生地

二元组
分类器

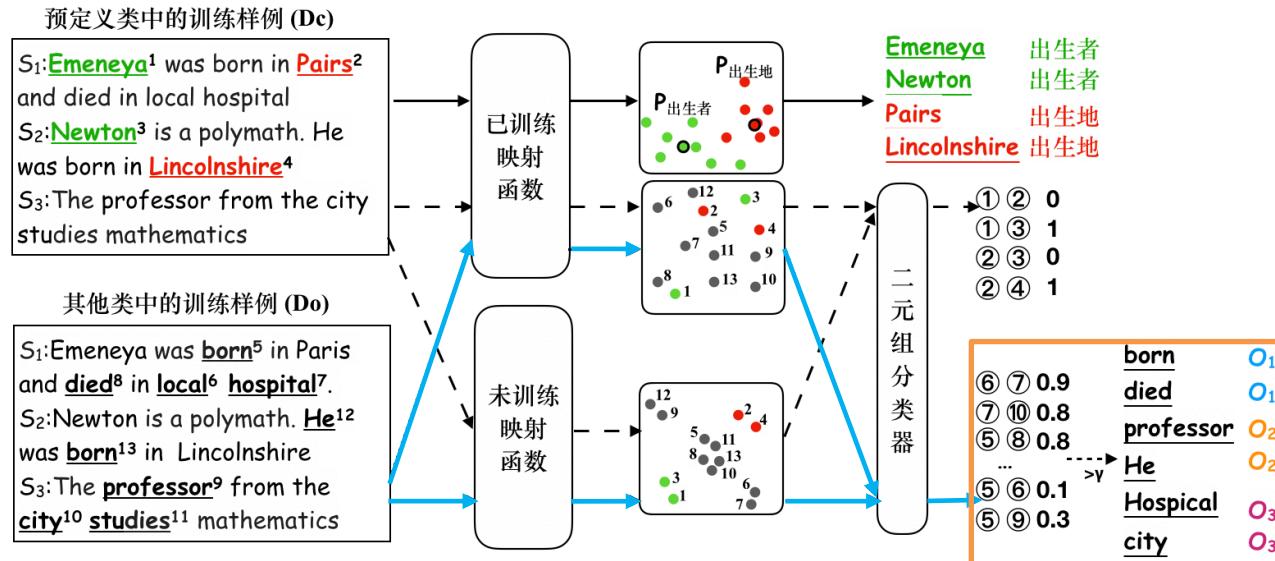
⑥	⑦	0.9
⑦	⑩	0.8
⑤	⑧	0.8
...		
⑤	⑥	0.1
⑤	⑨	0.3

born	O ₁
died	O ₁
professor	O ₂
...	
He	O ₂
Hospital	O ₃
city	O ₃

MUCO模型-任务相关类挖掘

◎ 第三步：利用二元组分类器预测

◆ 判断其他类中的任意两个点是否属于同一个任务相关类



◆ 风险控制机制：软标签

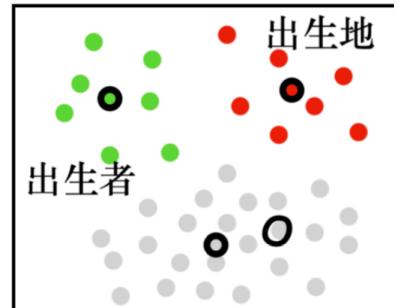


MUCO模型-联合学习

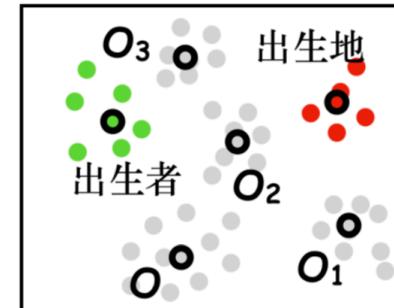
◎ 如何利用任务相关类?

- ◆ 将任务相关类 P_o 和预定义类 P_c 进行联合学习。

$$L(\theta_3) = -\log \frac{\exp(-d(x, p_y))}{\sum_{p \in \{P_c \cup P_o\}} \exp(-d(x, p))}$$



传统原型学习



MUCO原型学习



实验设置

◎ 数据集

- ◆ MUC-4：包含1700篇文章，5个事件要素角色
- ◆ 采用小样本5-shot的实验设定方法，即每个小样本类只有5个标注实例
 - 将“恐怖分子”，“恐怖分子所属机构”，“受害者”为基类
 - “攻击目标”和“攻击方式”为小样本类，

◎ 评测指标

- ◆ Precision、Recall以及Micro-F1值

◎ 基线模型

- ◆ 有监督方法：
 - BERT，Prototypical network
- ◆ 小样本学习方法：
 - Warm Prototypical Network，MAML，L-TapNet+CDT



实验结果

◎ 低资源场景下性能对比

Methods	精准匹配			模糊匹配		
	P	R	F	P	R	F
BERT	84.8	75.1	77.4	86.3	77.2	81.5
PN	81.1	73.6	76.0	83.2	74.7	78.7
L-TapNet+CDT	83.2	72.8	77.7	85.9	76.7	81.0
WPN	86.9	80.0	81.8	87.3	82.4	84.8
MAML	84.6	78.9	81.6	87.6	81.9	84.7
MUCO (ours)	87.5	81.5	83.2	90.5	82.5	86.3

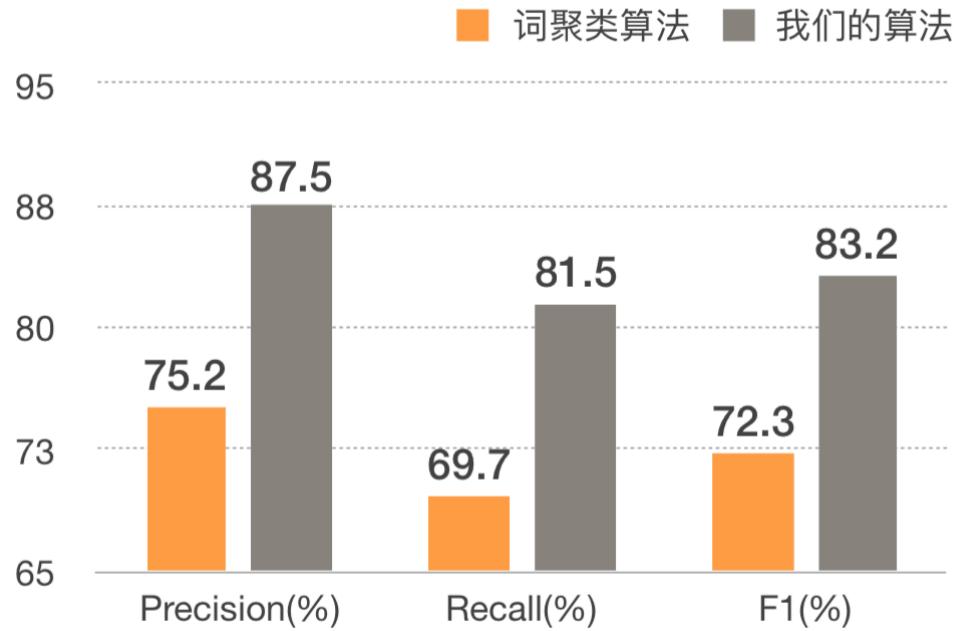




实验结果

◎ 和其他“任务相关类挖掘”算法对比

- ◆ 词聚类基线算法，对其他类中的词进行Kmeans聚类



小结

- 提出了一个新颖的基于自标注的事件要素抽取模型，通过挖掘任务相关类联合学习的方式，提升低资源事件要素抽取性能
- 提出了一个二元组分类器任务相关类挖掘方法，在没有人工标注的情况下，利用预定义类的弱监督信号来查找任务相关类。
- 与五个最先进的基线相比，我们的方法在低资源的实验设置下达到了最优的性能。
- 相关成果：Meihan Tong, Shuai Wang, Bin Xu, Yixin Cao, Minghui Liu, Lei Hou, Juanzi Li . Learning from Miscellaneous Other-Class Words for Few-shot Named Entity Recognition. **ACL 2021**.



内容大纲

- ◎ 研究背景和意义
- ◎ 基于多模态增强的事件检测方法
- ◎ 基于外部知识注入的事件检测方法
- ◎ 基于自标注的事件要素抽取方法
- ◎ 事件抽取在NewsMiner系统中的应用
- ◎ 总结



NewsMiner 系统简介

◎ NewsMiner 简介

- ◆ 面向事件的新闻分析挖掘和搜索系统
- ◆ 由清华大学知识工程实验室发布
- ◆ 致力于利用计算机技术自动化地组织和管理新闻事件

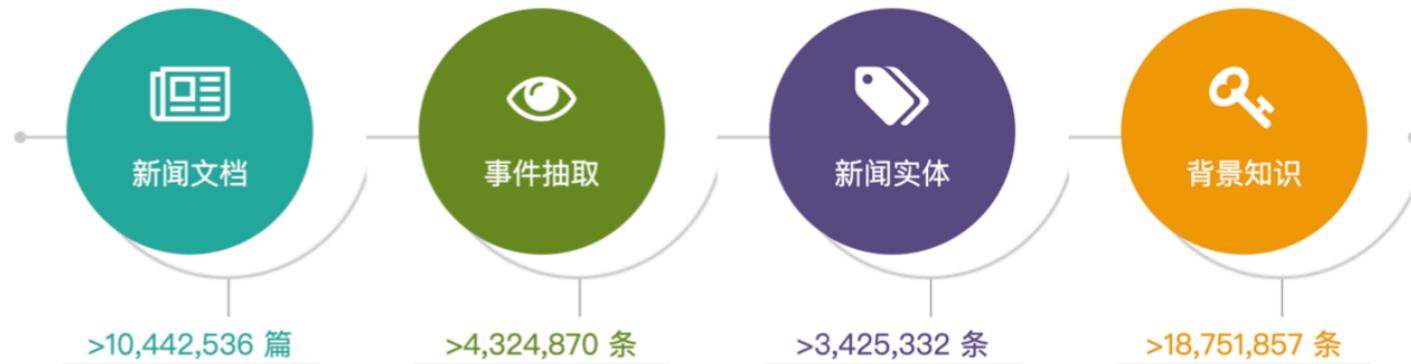


图9 NewsMiner 系统中新闻数量统计



NewsMiner 系统简介

◎ NewsMiner现有的语义分析工具

印尼东努沙登加拉省附近海域发生7.5级地震

2021-12-14 14:14:52 新闻 华龙网 日本气象厅 13条新闻 事件溯源

印尼东努沙登加拉省附近海域发生7.5级地震

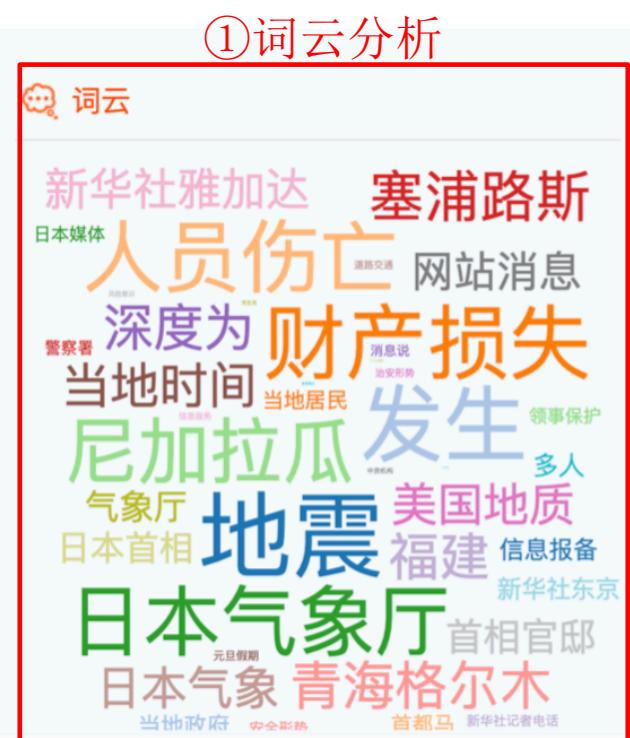
新华社雅加达12月14日电印度尼西亚东努沙登加拉省北部海域14日上午发生7.5级地震，目前尚无人员伤亡和财产损失报告。印尼地震监测部门已发布海啸预警。据印尼气象、气候和地球物理局网站消息，地震发生新华社雅加达12月14日电印度尼西亚东努沙登加拉省北部海域14日上午发生7.5级地震，目前尚无人员伤亡和财产损失报告。印尼...



②关键词抽取

57

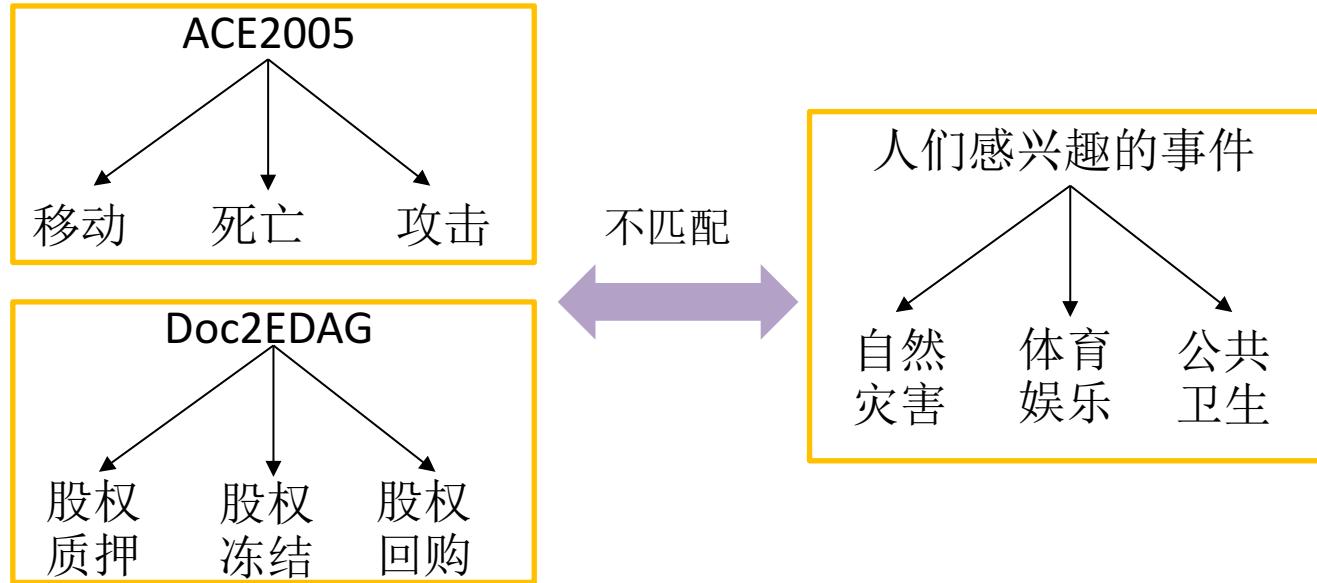
◆ 缺乏事件粒度上语义分析



NewsMiner应用-存在的挑战

- 现有事件分类本体不符合实际应用需求

- ◆ 太过于泛化
- ◆ 太过聚焦到某个特定领域
- ◆ 无法关注到用户感兴趣的事件



NewsMiner应用-构建本体

◎ 事件类型本体构建

◆ 软硬新闻理论

大类	小类
自然灾害	水灾, 旱灾, 火灾, 虫灾, 地震, 饥荒, 海啸 泥石流, 暴风雪, 火山喷发
人为灾害	空难, 沉船, 火车碰撞, 车祸, 银行抢劫 集体中毒, 气体爆炸, 矿井坍塌
经济事件	组织罚款, 组织合并, 组织成立, 组织倒闭, 经济危机, 经济援助
外交事件	外交谈判, 外交访问, 签署协议, 撕毁协议, 加入组织, 退出组织
政治事件	选举, 就职, 辞职, 政府政策变化
暴力冲突事件	罢工, 暴动, 政权更迭, 武装冲突, 军事演习, 抗议活动
公共卫生事件	环境污染, 疾病爆发
科技进步事件	卫星发射, 记录突破, 考古发现, 日食月食
公众人物事件	结婚事件, 离婚事件, 生病事件, 康复事件, 死亡事件, 调查事件 起诉事件, 逮捕事件, 判刑事件, 释放事件, 演讲事件
体育娱乐事件	体育比赛, 颁奖典礼



NewsMiner应用-构建本体

◎ 事件要素本体构建

◆ 更细粒度的事件要素

NAF飞机失事：军方接管现场，AIB开始调查

[1]周五晚上，尼日利亚空军飞机在卡杜纳国际机场坠毁，大量的士兵们封锁了现场。

[2]载有陆军参谋长、易卜拉欣·阿塔希鲁中将和其他高级军官的飞机在机场的跑道附近坠毁，机上11人全部遇难。

[3]卡杜纳机场经理阿米娜萨拉米在谈到这一事件时告诉频道电视台，军方当局已经接管了坠机现场。[4]她补充说，在事故调查局(AIB)的官员到达之前，他们已经阻止民航官员进入坠机现场。[5]她解释说，这架军用飞机最初计划降落在曼多地区的尼日利亚空军基地，但后来由于天气条件恶劣而备降到了卡杜纳机场，该飞机仅服役了三年，未超过生产厂规定的寿命期。

[6]军方说，这起空难事件发生的原因是飞机试图在恶劣天气下着陆。[7]穆罕默杜·布哈里总统说，缺乏严格的飞行前检查也是造成这次灾难的原因。[8]军方预计下周会召开新闻发布会对相关情况做出进一步的说明。[9]善后处理工作正在有序开展，考虑到此次坠机为军方行动，对遇难者家属的抚恤金会有相应的提升。

事件类型: 空难 (Air Crash)

事件要素:

时间
地点
伤亡情况
幸存者
事故调查员
事故原因
乘客
机组人员
承运机构
飞机编号
服役年限
起飞地点
计划降落地点
备用降落地点



NewsMiner应用-构建数据集

◎ DocEE众包标注

英文DocEE和现有数据集对比表

数据集	事件类型	事件要素角色	篇章数	词汇数	句子数	事件要素实例数
ACE2005	33	35	599	290k	15,789	9,590
KBP2016	18	20	169	94k	5,295	7,919
KBP2017	18	20	167	86k	4,839	10,929
MUC-4	4	5	1,700	495k	21,928	2,641
WikiEvents	50	59	246	190k	8,544	5,536
RAMS	139	65	9,124	957k	34,536	21237
DocEE(en)	59	356	27,485	16,268k	749,568	180,528

中文DocEE和现有数据集对比表

数据集	事件类型	事件要素角色	篇章数	词汇数	句子数	事件要素实例数
CEC	5	5	332	106k	1,385	5,954
DuEE	65	121	-	710k	19,640	41,520
DocEE(zh)	59	356	35,250	30,230k	130,000	218,550



NewsMiner应用-实验设置

◎ 数据集

- ◆ 中文DocEE 中选取了 5 个类，每个类挑选 10 条数据作为训练集，剩下的数据按照 1:1 分为验证集和测试集

◎ 中文事件抽取基线模型

- ◆ Rich-C: 字符+篇章特征
- ◆ CMS: 中文词形+义原信息
- ◆ C-BiLSTM: N-gram+序列信息
- ◆ BCNN: 注意力机制+语义特征融合
- ◆ HNN: RNN+CNN
- ◆ CAEE: BERT+图卷积网络融合句法结构

事件类型	训练集	验证集	测试集
空难	10	112	110
暴风雪	10	102	100
船难	10	98	100
颁奖典礼	10	98	110
集体中毒	10	101	108



NewsMiner应用-实验结果

◎ 所提方法在NewsMiner应用实验结果

Methods	EM			HM		
	P	R	F	P	R	F
Rich-C	12.3	10.4	11.3	17.3	14.5	15.8
CMS	13.2	11.6	12.3	18.2	13.8	15.7
C-BiLSTM	17.8	12.4	14.6	19.3	12.5	15.2
BCNN	23.4	16.2	19.1	25.3	16.6	20.0
HNN	24.1	17.2	20.1	26.2	18.6	28.0
CAEE	48.4	42.1	45.0	50.3	42.7	46.2
DRMM(多模态方法)	48.9	46.7	47.7	53.1	47.2	48.6
EKD (知识注入方法)	47.6	47.3	47.4	52.8	48.5	50.6
MUCO (自标注方法)	48.3	48.5	48.5	57.3	49.2	52.9

- ◎ 相关成果： Meihan Tong, Bin Xu, Shuai Wang, Meihuan Han, Yixin Cao, Jiangqi Zhu, Siyu Chen, Lei Hou, Juanzi Li. DocEE: A Large-Scale and Fine-grained Benchmark for Document-level Event Extraction. **NAACL 2022.**



NewsMiner应用-应用展示

◎ 应用展示

巴基斯坦南部地震造成至少15人死亡

① 2021-10-07 08:25:00 国内 新浪新闻 巴基斯坦 1条新闻 事件溯源

巴基斯坦南部地震造成至少15人死亡

原标题：巴基斯坦南部地震造成至少15人死亡 海外网10月7日电 据俄罗斯卫星通讯社报道，巴基斯坦救灾官员表示，当地时间7日凌晨，巴基斯坦南部发生地震，造成至少15人死亡。 中国地震台网正 原标题：巴基斯坦南部地震造成至少15人死亡 海外网10月7日电 据俄罗斯卫星通讯社报道，巴基斯坦救灾官员表示，当地...

地点 巴基斯坦 中国地震台网
热词 巴基斯坦 中国地震 当地时间

事件信息

巴基斯坦南部地震造成至少15人死亡

Type	Value
受影响地区	巴基斯坦南部、巴基斯坦
人员伤亡	15人死亡
日期	10月7日、7日凌晨、10月07日06时01分
震级	5.8级

- ◆ 围绕事件组织关键事实点，直击事件
- ◆ 事件要素详细全面，如“受影响地区”，“人员伤亡”



内容大纲

- 研究背景和意义
- 基于多模态增强的事件检测方法
- 基于外部知识注入的事件检测方法
- 基于自标注的事件要素抽取方法
- 事件抽取在NewsMiner系统中的应用
- 总结



创新点总结

◎ 三个方法+一个应用示范

- ◆ 针对低资源场景下事件类型消歧难的问题，提出了一个基于多模态融合的事件检测方法。
- ◆ 针对低资源场景下事件召回难的问题，提出基于外部知识注入的事件检测方法。
- ◆ 针对低资源场景下事件要素识别难的问题，提出基于自标注的事件要素抽取方法。
- ◆ 在新闻挖掘系统 NewsMiner 上展开示范应用。



未来展望

- 面向弱监督的事件抽取研究
 - ◆ 自监督学习
 - ◆ 零样本学习
- 面向常识约束的事件抽取研究
 - ◆ 整数逻辑规划
 - ◆ 符号学习
- 面向跨篇章的事件抽取研究
 - ◆ 更大的数据集
 - ◆ 更灵活的本体设计



科研成果

◎ 论文发表情况

- ◆ **Meihan Tong**, Shuai Wang, Bin Xu1, Yixin Cao, Minghui Liu, Lei Hou, Juanzi Li . Learning from Miscellaneous Other-Class Words for Few-shot Named Entity Recognition. ACL 2021. 【CCF A类长文】
- ◆ **Meihan Tong**, Bin Xu, Shuai Wang, Yixin Cao, Lei Hou, Juanzi Li and Jun Xie . Improving Event Detection via Open-domain Trigger Knowledge. ACL 2020. 【CCF A类长文】
- ◆ **Meihan Tong**, Shuai Wang, Yixin Cao, Bin Xu, Juaizi Li, Lei Hou, Tat-Seng Chua. Image Enhanced Event Detection in News Articles. AAAI 2020 【CCF A类长文】
- ◆ **Meihan Tong**, Bin Xu, Shuai Wang, Yixin Cao, Lei Hou, Juanzi Li and Jun Xie . Improving Event Detection via Open-domain Trigger Knowledge. NAACL 2022. 【TH-CPL B类长文】
- ◆ **Meihan Tong**, Bin Xu, Shuai Wang, Lei Hou, and Juaizi Li. Improving Low-Resource Chinese Event Detection with Multi-Task Learning. KSEM 2020. 【CCF C类长文】
- ◆ **Meihan Tong**, Bin Xu, Lei Hou, Juanzi Li, Shuai Wang. Leveraging Multi-head Attention Mechanism to Improve Event Detection. CCL2019.
- ◆ Shuai Wang, Lei Hou, **Meihan Tong**. Unsupervised Cross-Lingual Sentence Representation Learning via Linguistic Isomorphism. KSEM 2019.
- ◆ 唐望径, 许斌, 全美涵, 韩美奂, 王黎明, 钟琦。知识图谱增强的科普文本分类方法,《计算机应用》2021
- ◆ 杨玉基, 许斌, 胡家威, 全美涵, 张鹏, 郑莉。一种准确高效的领域知识图谱构建方法,《软件学报》2018,29(10):0.



科研成果

◎ 专利授权情况

专利名称	发明人	申请状态	专利号
一种多模态事件检测方法及装置	许斌; 全美涵; 李涓子; 侯磊	已授权	ZL 2020 1 0076960.1
事件检测深模型的构建方法、装置、电子设备及存储介质	许斌; 全美涵; 李涓子; 侯磊	已授权	ZL 2020 1 0548917.0
用于命名实体识别的模型训练方法、识别方法及装置	许斌; 全美涵; 李涓子; 侯磊	已授权	ZL 2021 1 0621275.7



参考文献 (1/8)

1. Riloff E, et al. Automatically constructing a dictionary for information extraction tasks[C]// Proceedings of the 11th National Conference on Artificial Intelligence. Citeseer, 1993: 2-1.
2. Yangarber R, Grishman R, Tapanainen P, et al. Automatic acquisition of domain knowledge for information extraction[C]//Proceedings of the 18th International Conference on Computational Linguistics. 2000.
3. 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. 2013: 73-82.
4. Björne J, Salakoski T. Generalizing biomedical event extraction[C]//Proceedings of BioNLP Shared Task 2011 Workshop. 2011: 183-191.
5. Duan S, He R, Zhao W. Exploiting document level information to improve event detection via recurrent neural networks[C]//Proceedings of the 8th International Joint Conference on Natural Language Processing. 2017: 352-361.
6. Zhao Y, Jin X, Wang Y, et al. Document embedding enhanced event detection with hierarchical and supervised attention[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018: 414-419.



参考文献 (2/8)

7. Sha L, Qian F, Chang B, et al. Jointly extracting event triggers and arguments by dependency-bridge rnn and tensor-based argument interaction[C]//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. 2018.
8. Cui S, Yu B, Liu T, et al. Edge-enhanced graph convolution networks for event detection with syntactic relation[J]. arXiv preprint arXiv:2002.10757, 2020.
9. Lu W, Nguyen T H. Similar but not the same: Word sense disambiguation improves event
10. Lu W, Nguyen T H. Similar but not the same: Word sense disambiguation improves event detection via neural representation matching[C/OL]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2018: 4822-4828.
<https://aclanthology.org/D18-1517>. DOI: 10.18653/v1/D18-1517.
11. Nguyen T M, Nguyen T H. One for all: Neural joint modeling of entities and events[C]// Proceedings of the 2019 AAAI Conference on Artificial Intelligence. 2019: 6851-6858.



参考文献 (3/8)

12. Ferguson J, Lockard C, Weld D, et al. Semi-supervised event extraction with paraphrase clusters [C/OL]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2018: 359-364. <https://www.aclweb.org/anthology/N18-2058>. DOI: 10.18653/v1/N18-2058.
13. Huang R, Riloff E. Bootstrapped training of event extraction classifiers[C/OL]//Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2012: 286-295. <https://www.aclweb.org/anthology/E12-1029>.
14. Liu S, Chen Y, He S, et al. Leveraging framenet to improve automatic event detection[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. 2016: 2134-2143.
15. Reschke K, Jankowiak M, Surdeanu M, et al. Event extraction using distant supervision[C]// Proceedings of the 9th International Conference on Language Resources and Evaluation. 2014: 4527-4531.
16. Li, Qi et al. Joint event extraction via structured prediction with global features. ACL 2013
18. Linguistic Data Consortium. "ACE (Automatic Content Extraction) Arabic Annotation Guidelines for Entities." (2005).



参考文献 (4/8)

19. Ahn, David. "The stages of event extraction." Proceedings of the Workshop on Annotating and Reasoning about Time and Events. 2006.
20. Hogenboom, Frederik, et al. "An overview of event extraction from text." DeRiVE@ ISWC. 2011.
21. Riloff, Ellen. "Automatically generating extraction patterns from untagged text." Proceedings of the national conference on artificial intelligence. 1996.
22. Ritter, Alan, Oren Etzioni, and Sam Clark. "Open domain event extraction from twitter." Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012.
23. Sekine, Satoshi. "On-demand information extraction." Proceedings of the COLING/ACL on Main conference poster sessions. Association for Computational Linguistics, 2006.
24. Huang, Lifu, et al. "Liberal event extraction and event schema induction." Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016.
25. Buyko, Ekaterina, et al. "Event extraction from trimmed dependency graphs." Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing: Shared Task. Association for Computational Linguistics, 2009.



参考文献 (5/8)

26. S. J. Conlon, A. S. Abrahams, L. L. Simmons, Terrorism Information Extraction from Online Reports, *Journal of Computer Information Systems* 55 (3) (2015) 20–28.
27. Wang, Wei. Event detection and extraction from news articles. Diss. Virginia Tech, 2018.
28. Danilova, Vera, Mikhail Alexandrov, and Xavier Blanco. "A survey of multilingual event extraction from text." *International Conference on Applications of Natural Language to Data Bases/Information Systems*. Springer, Cham, 2014.
29. J. Borsje, F. Hogenboom, F. Frasincar. Semi-Automatic Financial Events Discovery Based on Lexico-Semantic Patterns, *International Journal of Web Engineering and Technology* 6 (2) (2010) 115–140.
30. Zhao, Yan-yan, et al. "Research on Chinese event extraction." *Journal of Chinese Information Processing* 22.1 (2008): 3-8.
31. Hogenboom, Frederik, et al. "A survey of event extraction methods from text for decision support systems." *Decision Support Systems* 85 (2016): 12-22.
32. Wei, Chih-Ping, and Yen-Hsien Lee. "Event detection from online news documents for supporting environmental scanning." *Decision Support Systems* 36.4 (2004): 385-401.
33. Araki J, Mitamura T. Open-domain event detection using distant supervision[C]//Proceedings of the 27th International Conference on Computational Linguistics. 2018: 878-891



参考文献 (6/8)

34. Nguyen, Thien Huu, and Ralph Grishman. "Graph convolutional networks with argument-aware pooling for event detection." Thirty-second AAAI conference on artificial intelligence. 2018.
35. Feng, Xiaocheng, Bing Qin, and Ting Liu. "A language-independent neural network for event detection." Science China Information Sciences 61.9 (2018): 092106.
36. Liu, Shulin, et al. "Exploiting argument information to improve event detection via supervised attention mechanisms." Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017.
37. 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). 2017.
38. Liu, Xiao, Zhunchen Luo, and Heyan Huang. "Jointly multiple events extraction via attention-based graph information aggregation." arXiv preprint arXiv:1809.09078 (2018).
39. Chen, Yubo, et al. "Collective event detection via a hierarchical and bias tagging networks with gated multi-level attention mechanisms." Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018.
40. Yang H, Chen Y, Liu K, et al. Dcfee: A document-level chinese financial event extraction system based on automatically labeled training data[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018: 50-55.



参考文献 (7/8)

41. Veyseh, Amir Pouran Ben, Thien Huu Nguyen, and Dejing Dou. "Graph based Neural Networks for Event Factuality Prediction using Syntactic and Semantic Structures." arXiv preprint arXiv:1907.03227 (2019).
42. Yang, Sen, et al. "Exploring pre-trained language models for event extraction and generation." Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019.
43. Huang, Lifu, et al. "Zero-shot transfer learning for event extraction." arXiv preprint arXiv:1707.01066 (2017).
44. Hong, Yu, et al. "Self-regulation: Employing a generative adversarial network to improve event detection." Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018.
45. Duan, Shaoyang, Ruifang He, and Wenli Zhao. "Exploiting document level information to improve event detection via recurrent neural networks." IJCAI. 2017.
46. Deng, Shumin, et al. "Meta-Learning with Dynamic-Memory-Based Prototypical Network for Few-Shot Event Detection." WSDM. 2020.
47. Lu Y, Lin H, Xu J, et al. Text2event: Controllable sequence-to-structure generation for end-to-end event extraction[J]. arXiv preprint arXiv:2106.09232, 2021.
48. Du X, Rush A M, Cardie C. Template filling with generative transformers[C]//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics. 2021: 909-914.



参考文献 (8/8)

49. Li, Wei, et al. "Joint event extraction based on hierarchical event schemas from framenet." *IEEE Access* 7 (2019): 25001-25015.
50. Lu, Yaojie, et al. "Distilling Discrimination and Generalization Knowledge for Event Detection via Delta-Representation Learning." *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019.
51. Tong, Meihan, et al. "Image Enhanced Event Detection in News Articles." *AAAI*. 2020.
52. Tong, Meihan, et al. "Leveraging Multi-head Attention Mechanism to Improve Event Detection." *China National Conference on Chinese Computational Linguistics*. Springer, Cham, 2019.
53. Ngo, Nghia Trung, Tuan Ngo Nguyen, and Thien Huu Nguyen. "Learning to Select Important Context Words for Event Detection." *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, Cham, 2020.
54. Araki, Jun, and Teruko Mitamura. "Open-domain event detection using distant supervision." *Proceedings of the 27th International Conference on Computational Linguistics*. 2018.





谢谢各位评审老师！