



基于序列标注的事件联合抽取方法

王晓浪¹, 邓蔚^{1,2}, 胡峰¹, 邓维斌¹, 张清华¹

(1. 重庆邮电大学 计算智能重庆市重点实验室, 重庆 400065; 2. 西南财经大学 统计研究中心, 成都 611130)

摘要: 事件抽取是自然语言处理领域的重要研究方向。传统的事件类型抽取系统采用分类方式, 无法解决跨句子的事件角色和事件类型匹配问题。为了解决该问题, 提出一种基于序列标注的事件联合抽取模型, 结合卷积神经网络(convolutional neural networks, CNN) 与长短期记忆网络(long short-term memory, LSTM) 提取全局特征和局部特征; 并在浅层 LSTM 层共享参数实现联合抽取, 以序列标注方式抽取事件论元并匹配事件类型。实验结果表明, 模型能有效提取司法领域的文档事件信息。

关键词: 事件抽取; 联合抽取; 序列标注; 司法领域; 争议焦点

中图分类号: TP183

文献标志码: A

文章编号: 1673-825X(2020) 05-0884-07

Joint event extraction based on sequence annotation

WANG Xiaolang¹, DENG Wei^{1,2}, HU Feng¹, DENG Weibin¹, ZHANG Qinghua¹

(1. Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications,

Chongqing 400065, P. R. China;

2. Center of Statistical Research, Southwestern University of Finance and Economics, Chengdu 611130, P. R. China)

Abstract: Event extraction is an important research in the field of natural language processing. In traditional event extraction systems, event types as a classification task cannot be matched with event arguments. In this paper, a joint event extraction model based on sequence annotation is proposed to solve event matching across sentences. Global and local features are extracted by convolutional neural networks (CNN) and long short-term memory (LSTM). The joint extraction is realized by sharing parameters in the shallow LSTM layer, event types and arguments are extracted and matched by sequence annotation. The experimental results in the judicial field corpus show the effectiveness of proposed model.

Keywords: event extraction; joint extraction; sequence annotation; judicial; focus of dispute

0 引言

作为信息抽取的一个子任务, 事件抽取(event extraction, EE) 旨在从非结构化数据中快速获取重要的事件信息, 形成结构化形式, 并应用到其他下游任务, 例如信息检索、智能问答、阅读理解等任务中。

自动内容抽取(automatic content extraction, ACE) 将事件抽取分为 2 个子任务: ①事件类型的识别, 该任务关注触发词的识别, 通过触发词判断事件类型; ②事件论元的抽取, 该任务关注命名实体和实体在事件中的角色信息。事件抽取相关术语如下。

事件句: 文中描述事件的句子, 通常由一个触发

收稿日期: 2020-06-22 修订日期: 2020-07-22 通讯作者: 胡峰 hufeng@cqupt.edu.cn

基金项目: 国家重点研发计划(2018YFC0832100); 国家自然科学基金(61876201)

Foundation Items: The National Key Research and Development Program of China (2018YFC0832100); The National Natural Science Foundation of China (618762001)

© 1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

词和任意数量的事件论元组成。

事件触发词:能清晰表示事件发生的词语,区分事件类型,是事件句中的关键词。

事件论元:事件的参与者及属性,如事件发生的时间、地点和人物等等,通常都是命名实体。

事件角色:命名实体在事件中扮演的具体角色,如婚姻案件中的人物角色有抚养者和被抚养者。

目前,事件抽取主要存在2种方式:①管道方式,先识别触发词并确定事件类型,再识别事件论元;②联合抽取方法,采用端到端的模型,同时完成触发词识别和论元识别任务。

事件抽取最初基于模板匹配,根据规则定义抽取模板。文献[1]提出基于语言模型去获取事件信息。文献[2]提出基于模板的方式抽取新闻事件。随着神经网络方法在自然语言处理领域的应用越来越成熟,事件抽取也都使用神经网络方式实现。文献[3]提出利用动态多池卷积神经网络(dynamic multi-pooling convolutional neural networks, DMCNN)以管道方式实现事件抽取,并引入词向量^[4]和改进的卷积神经网络(convolutional neural networks, CNN),自动学习词汇和句子级别的特征。文献[5]提出带注意力机制的循环神经网络模型,提高了事件检测的效率。文献[6]提出convolution-BLSTM神经网络抽取中文事件,并对比了基于词和基于字的2种向量嵌入方法。文献[7]将递归神经网络应用到中文事件检测中。文献[8]提出使用生成式对抗网络(generative adversarial networks, GAN)生成一些虚假的特征信息,提高模型的学习效果,虚假特征指在语义上类似但并不相关的特征。文献[9]提出一种基于先验知识和自注意力机制的序列标注模型,将触发词作为先验知识控制端到端模型的学习过程,提高了模型的可解释性。神经网络虽然提高了管道方式的效率,但管道方式容易产生级联错误,且处于下游的任务无法将信息反馈至上游任务,辅助上游任务的识别。

联合抽取不仅能解决错误级联问题,还能有效捕捉触发词和事件论元间的相互依赖关系。文献[10]提出一种基于递归神经网络的联合方法,将4个子任务做成端到端模型,挖掘出事件触发词和事件角色的内在依赖关系。文献[11]针对句子论元分散问题,提出从文档层面出发进行事件抽取,并使用远程监督^[12]方法,扩展了训练集。文献[13]提出一种基于树结构长短期记忆网络(tree-structured

long short-term memory networks, TREE-LSTM)和门控循环单元网络(gated recurrent unit, GRU)的模型, TREE-LSTM网络将事件触发词的依存句法信息添加到了模型中。文献[14]将事件抽取看作序列标注任务,构建了基于条件随机场(conditional random field, CRF)的事件抽取联合模型。文献[15]引入句法依存信息增强信息流,并基于注意力机制构建图信息,实现联合抽取。

已有的事件联合抽取工作都基于ACE 2005数据集,该数据集仅在句子范围内标记事件信息,即只包含句子级事件。只研究句子级的事件抽取会忽略2个关键的问题:论元分散和多事件句。论元分散指事件中的论元分散在多个句子中,只研究句子级事件抽取,会造成事件论元丢失;多事件句指一个句子中包含2个或2个以上的事件,将句子只分为一个事件类型,事件论元必定会匹配错误。研究文档级事件抽取时,事件类型与事件角色难以直接匹配。

序列标注任务,是一种基于字级别的多分类任务,基于字级别的事件类型抽取,可直接匹配事件论元,且构建端到端的联合抽取可同时确定事件类型与事件论元,避免了误差传递。综上,提出采用联合抽取和序列标注的方法提取事件角色和事件类型。总的来说,本文的贡献在于以下2点:

- 1) 提出了一种基于序列标注的联合抽取方法,匹配事件类型与论元,提高了事件抽取性能;
- 2) 研究文档级事件抽取,解决多事件句与事件论元分散问题。

1 事件联合抽取模型

在基于序列标注的联合抽取模型中,事件类型和事件论元被定义为2次序列标注任务,通过长短期记忆网络(long short-term memory networks, LSTM)对上下文进行建模,逐词迭代得全局的特征向量, CNN从词向量序列中自动发现局部特征,并在LSTM层共享参数实现联合抽取。图1描述了事件抽取模型的总体架构,包括以下4个部分:①词向量表示;②联合抽取;③事件角色提取;④事件类型提取。

1.1 字向量表示

使用WORD2VEC中的SKIP-GRAM模型预训练词向量。每个句子以字符为单位由向量 $T = [C_1, C_2, C_3, \dots, C_{N_t}]$ 构成, N_t 是句子最大长度,利用预训练的字向量将句子中的字 C_i 映射到低维稠密的字

向量 W_i 文档级数据中包含多个句子,多个句子可合并到一个高维的列表,得到文档的向量表示 $S = [W_1, W_2, W_3, \dots, W_{N_s}]$, N_s 为文档中最大字符个数。

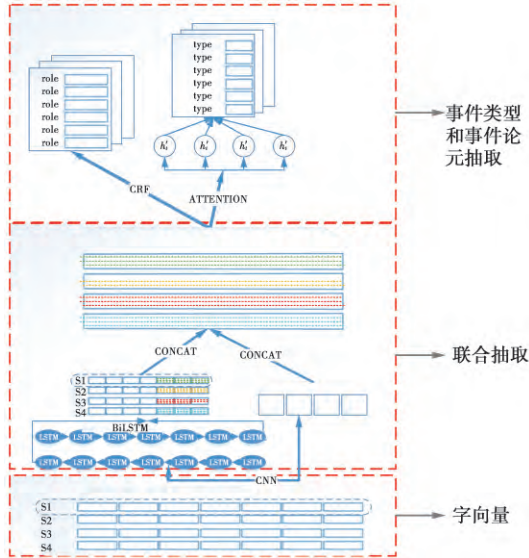


图 1 文档级事件抽取模型整体框架

Fig.1 Overall framework of document level event extraction model

1.2 联合抽取

文献[16]证明 CNN 能从词的字符表示中提取出形态学信息,如词的前缀和后缀。对字符向量进行卷积操作可获得单词的字符表示,而改变 CNN 的卷积核的大小并合并多层 CNN,可提取不同长度的词向量信息。合并 CNN 提取的词向量特征与 WORD2VEC 的字符向量特征进行残差连接,并作为 LSTM 层输入,LSTM 层计算当前层输入与前一层的输出,可学习长期依赖。将 LSTM 作为参数共享层,下游任务事件抽取与事件论元共享该层网络参数,实现联合抽取。

LSTM 只能获取前端序列的信息,使用 BILSTM 可同时获取前后向信息。BILSTM 编码层分为前向 LSTM 层、后向 LSTM 层和他们的拼接层,其中 LSTM 可获取序列上文信息,用于特征提取。相对于普通的循环神经网络 (recurrent neural network, RNN), LSTM 加入了 3 个门机制可更好地获取长范围的依赖关系,并使用特殊的记忆细胞取代 RNN 隐藏层的更新方式。

第 1 个门机制是遗忘门,遗忘门决定上一时刻的单元状态 C_{t-1} 有多少能保留到当前状态 C_t ,值越大的位置执行点乘操作后的值也越大,即当前位置保留的历史信息越多,通过激活函数 SIGMOID 将值

约束在 $[0, 1]$, 1 表示完全保留,0 则表示完全舍弃。遗忘门公式为

$$f_t = \sigma(W_f[h_{t-1}, X_t] + b_f) \quad (1)$$

(1) 式中: σ 是逻辑 SIGMOID 函数; h_{t-1} 是上一层的隐藏层; X_t 是当前层输入的词向量; W_f, b_f 分别为权重和偏置矩阵; f_t 是遗忘门矩阵。

第 2 个门机制是输入门,决定当前状态有多少信息得到更新,通过 2 步实现: ①SIGMOID 层决定需要被更新的信息; ②TANH 层生成备选的更新内容,并更新细胞状态,计算当前时刻的单元状态 C_t 。输入门公式为

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, X_t] + b_c) \quad (3)$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (4)$$

(2) — (4) 式中: i_t 是输入门矩阵; \tilde{C}_t 是当前输入层的状态矩阵; C_{t-1}, C_t 分别为前一层与当前层的状态矩阵; W_i, W_c, b_i, b_c 分别为权重和偏置。

第 3 个门机制是输出门,确定输出值。首先,通过 SIGMOID 计算输出门矩阵;接着,将当前状态值经过 TANH 处理,将值约束在 $[-1, 1]$,再和输出门矩阵相乘,作为最终的输出。其公式为

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o) \quad (5)$$

$$h_t = O_t \tanh(C_t) \quad (6)$$

(5) — (6) 式中: O_t, h_t, W_o, b_o 分别为输出门矩阵、隐藏层输出值、输入门权重和偏置。

前向 LSTM 层和后向 LSTM 层各自产生一个隐藏层 h_t , 包含前后端序列信息, BILSTM 则将正向和反向的隐层拼接作为最终的隐层向量 $h_t = [\vec{h}_t, \overleftarrow{h}_t]$ 。在 BILSTM 层共享参数,提取的特征被子任务共享,2 个子任务依赖相同的特征,在反向传播时,2 个子任务在 LSTM 层合并损失值,通过优化 LSTM 层,同时优化 2 个子任务,然后再进入各自的网络层,提取对应任务的独立特征,完成子任务。

1.3 事件角色抽取

事件角色抽取任务采用 BIEO 标签^[17]模式 (Begin: 单词开始, Inside: 单词中间, End: 单词结束, Outside: 非事件元素单词), 标签中还包含单词的相对位置信息。BILSTM 虽然能够捕获到长距离的字级别依赖关系,但是序列标注任务还需要考虑相邻标签之间的依赖性。CRF 可计算给定一组输入序列条件下另一组输出序列的条件概率分布,将

LSTM 层输出作为 CRF 的发射概率,CRF 计算转移概率,便可计算标签之间的依赖关系。CRF 层转移概率矩阵 A_{ij} 表示第 i 个标签到第 j 个标签的转移得分,该矩阵可以利用已经标注过的标签。当已知序列 $S = [W_1, W_2, \dots, W_{N_w}]$ 对应的输出标签结果为 $y = [y_1, y_2, \dots, y_{N_w}]$ 时,可定义当前序列得分为

$$\text{score}(x, y) = \sum_{i=1}^n P_{i, y_i} + \sum_{i=1}^{n+1} A_{y_{i-1} y_i} \quad (7)$$

(7) 式中: y_i 是第 i 个位置的标签值; P_{i, y_i} 是第 i 个位置 SOFTMAX 输出为 y_i 的概率; $A_{y_{i-1} y_i}$ 为 y_{i-1} 到 y_i 的转移概率。每个位置的打分之和作为整个序列的打分值,而每个位置的打分由 2 部分组成,一部分由上层网络的输出决定,另一部分由 CRF 的转移矩阵决定。最后利用 SOFTMAX 计算归一化后的概率,公式为

$$P(y|x) = \frac{\exp(\text{score}(x, y))}{\sum \exp(\text{score}(x, y'))} \quad (8)$$

采用最大化对数似然函数优化目标函数,训练样本 (x, y) 的对数似然为

$$\log P(y|x) =$$

$$\text{score}(x, y) - \ln(\sum \exp(\text{score}(x, y'))) \quad (9)$$

在预测(解码)时,使用动态规划的 Viterbi 算法求解最优路径,得到序列标注任务中每个字对应的标签概率,最大概率对应的标签即为正确标签,概率公式为

$$y^* = \underset{y}{\operatorname{argmax}} \text{score}(x, y') \quad (10)$$

1.4 事件类型抽取

事件类型抽取和以往的分类任务不同,采用序列标注方式,使用自注意力机制学习序列内部的词依赖关系,捕获序列的内部结构。合并识别的事件角色的编码信息和参数共享层的编码信息,然后送入到自注意力层,该层关联单个序列的不同位置的注意力,计算同一个句子内间隔较远的单词之间的联系,ATTENTION 计算公式为

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (11)$$

(11) 式中: $Q \in R^{n \times d}$, $K \in R^{n \times d}$ 和 $V \in R^{n \times d}$ 都是需要计算的矩阵; d 是网络中隐藏层神经元的数量; n 是输入句子长度。自注意力机制将句子中任意 2 个词通过计算产生直接联系,更容易捕获相互依赖的特征。

模型采用改进的多头注意力结构,将输入向量矩阵通过不同的线性变换映射成多个矩阵,分别执

行点乘操作,拼接起来的 h 次点乘结果作为编码向量,注意力机制公式为

$$M_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (12)$$

$$M = \text{Concat}(M_1, M_2, \dots, M_h) \quad (13)$$

$$Y = MW \quad (14)$$

(12) — (14) 式中: h 为平行头的数量; Concat 指将多个矩阵拼接操作; W_i^Q, W_i^K, W_i^V, W 是对应线性变换的权重矩阵。

2 仿真实验

为了验证本文方法的性能,在司法数据集上做了 3 组实验,分别为基准对比实验、网络模型对比实验、联合抽取与管道方式的对比实验,并在金融数据集上与基准实验对比,验证文中方法的可移植性。

2.1 数据来源

实验使用的司法数据集来自法律网站裁判文书网(<http://www.wenshu.com/>),网站数据包含法律案件和争议焦点以及关键字。争议焦点是司法案件中需要解决的主要问题,以事件抽取方式抽取答辩状中的争议焦点,并将争议焦点和争议焦点中的参与者分别作为事件类型与事件论元。

司法案件分为民事、刑事和行政 3 类,其中民事案件更符合实验要求,实验仅研究民事案件中的婚姻纠纷案例。采用 BRAT 工具对司法数据人工打标,打标婚姻数据总共 1 540 篇,按照 6:2:2 划分为训练集、测试集和验证集,事件类型如表 1。

表 1 婚姻数据集事件类型分布

Tab.1 Event type distribution of marriage data set

事件类型	训练集	验证集	测试集	总共
感情破裂	838	118	126	1 082
抚养问题	623	91	86	800
财产问题	561	117	96	774
出轨	757	97	82	936
债权债务	210	120	108	438
总共	1 132	204	204	1 540

2.2 实验步骤

首先,在司法网站爬取答辩状数据;接着,使用模型抽取婚姻纠纷类型,争议焦点的重要参与者以及原被告对于焦点的情感词(如赞同、否定和听法官判决等);然后,先根据事件类型判断是否属于同一事件,再判断事件中的主语及拥有者是谁(如股权拥有者和股权转让者)和判断情感词是否一致;

最后,得出该事件是否产生争议焦点。实验具体流程如图 2。

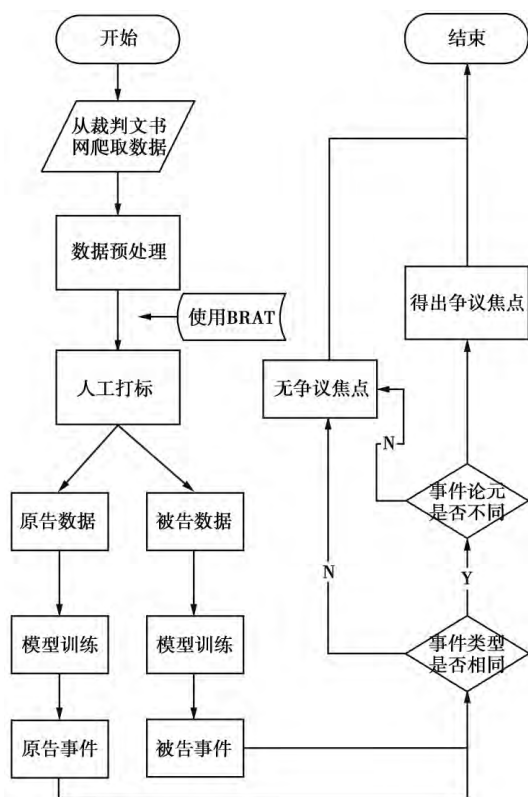


图 2 争议焦点事件抽取与分析流程图

Fig.2 Flow chart of legal dispute focus extraction

2.3 参数说明

模型具体参数如下: 字向量 EMBEDDING 维度为 200, 使用 WORD2VEC 预训练完成; 文本最大长度为 500 维; LSTM 隐层维度为 128; BATCH 设定为 32; 学习率大小为 0.000 1; DROUPOUT 为 0.5; 优化器采用 ADAM。

2.4 实验环境和评价指标

具体实验环境如下: 操作系统为 windows10 64 位; 处理器为 Intel (R) Core (TM) i5-4590 @ 3.3 GHz; 内存大小为 16 GByte。使用的主要编程环境为: python 3.6.5, tensorflow 1.14.0, keras2.2.4。

实验的评估度量方式为准准确率 (precision, P)、召回率 (recall, R)、 F_1 (F_1 -measure) 值, 其中, 序列标注是基于字级别的多分类任务, 而评估时论元是基于词级别的, 所以词语的 BIO 标签预测都计入评估。数据集集中的每个事件类型通过微平均进行评估为

$$R = \frac{\text{事件类别预测正确的标签数目}}{\text{数据集中该类别实际的标签数目}}$$

$$F_1 = \frac{2PR}{P + R}$$

2.5 实验结果及分析

为了验证基于序列标注的事件抽取模型 (DJEE) 在司法领域的性能, 将它和 3 个基准系统做比较, 这 3 个系统分别为: ①面向文档级金融数据的事件抽取框架 DCFEE^[11]; ②利用基于注意力机制和多层感知机的篇章级事件抽取模型 ILP^[18]; ③基于 LSTM 结合自注意力和门控卷积同时抽取实体和事件的模型 HNN-EE^[19]。所有模型都采用预训练好的 WORD2VEC 向量, 实验结果如表 2。

表 2 事件类型对比实验 F_1

Tab.2 Event type comparison experiment F_1 %

模型	感情	抚养	财产	出轨	债务	AVG
ILP	52.8	56.9	68.1	64.4	71.1	62.6
HNN-EE	69.6	86.9	69.4	68.4	85.7	76.0
DCFEE	60.6	79.7	62.2	65.3	83.1	70.1
DJEE	72.4	85.2	77.9	77.4	86.0	79.7

分析表 2 实验结果可知, 基于联合抽取方法的 HNN-EE 和 DJEE 可以避免误差传递, 并捕捉触发词和事件论元之间的相互依赖关系, 提取效果更好。基于联合抽取的方式中, 相比于 HNN-EE 模型将事件类型识别作为分类任务, 然后再采取相应的算法去填充事件论元角色方式, DJEE 模型采用序列标注方式, 基于字级别实现分类任务, 直接匹配事件类型与事件角色。

为了进一步分析网络结构对模型性能的影响, 还设计了 3 组对比试验。①NO_CNN: 原网络未使用卷积神经网络 CNN, 分析 CNN 提取局部特征对模型的影响; ②NO_LSTM: 原模型未使用双向 LSTM, 分析 LSTM 提取全局特征对模型的影响; ③Class: 事件论元角色识别仍使用序列标注, 但事件类型识别作为一个分类任务, 分析序列标注对模型的影响。3 组实验都是联合抽取, 实验结果如表 3。

表 3 网络模型对比试验

Tab.3 Network model comparison experiment %

模型	P	R	F_1
NO_CNN	69.5	83.1	74.8
NO_LSTM	50.1	18.1	23.1
Class	70.5	66.3	68.2
DJEE	72.4	85.2	79.7

由表3可知,LSTM提取词级别特征,可大幅度提升模型的召回率;Class模型和DJEE模型相比于NO_CNN模型,NO_LSTM模型准确率相对提高,表明CNN与LSTM提取全局与局部特征,并在LSTM层共享,提取了更多语义特征,可提高模型准确率。

为了验证联合抽取的有效性,设计了1组对比实验,Pipeline为管道方式抽取,分步抽取事件类型和事件角色,评价方式和联合抽取一样。实验结果如表4。

表4 联合抽取性能对比

Tab.4 Performance comparison of combined extraction %

模型	P	R	F ₁
Pipeline	75.5	68.0	71.2
DJEE	72.4	85.2	79.7

由表4可知,加入了参数共享的联合抽取,在反向传播时同时更新2个子任务,不仅在任务间传导不确定信息,还共享有价值的信息,2个子任务间信息相互增强,在准确率保持不变的同时,召回率得到提高。

为了进一步验证模型抽取文档事件的性能,在金融数据上进行了3组对比实验,实验数据集来自<https://github.com/dolphin-zs/Doc2EDAG>,实验效果如表5。

表5 金融数据对比实验 F1

Tab.5 Financial data comparison experiment F1 %

模型	EF	ER	EU	EO	EP	AVG
ILP	52.0	56.1	55.1	54.0	65.5	56.5
HNN-EE	67.9	86.5	67.7	67.7	86.5	75.2
DCFEE	65.4	86.1	62.7	68.4	84.2	73.3
DJEE	71.2	86.9	64.5	69.5	84.6	75.3

从表5可知,DJEE在金融数据集上仍取得了较好的性能,证明该模型可有效缓解抽取文档数据时跨事件句与论元分散带来的影响。

3 结束语

本文针对事件抽取任务存在的论元分散和跨事件句问题,提出基于序列标注的事件抽取方法。模型通过LSTM层共享参数实现联合抽取,以序列标注形式抽取并匹配事件类型和事件论元。基于序列标注的事件抽取,可简单快速实现事件类型与事件论元的匹配,无需额外的特征与步骤。通过在婚姻纠纷数据集上的多组仿真实验表明,文中提出的方

法能有效提取文档的事件信息。

参考文献:

- [1] KIM J T, MOLDOVAN D I. Acquisition of linguistic patterns for knowledge-based information extraction [J]. IEEE Transactions on Knowledge and Data Engineering, 1995, 7(5): 713-724.
- [2] GUPTA S, PATEL D. NE2: named event extraction engine [J]. Knowledge and Information Systems, 2019, 59(2): 311-335.
- [3] YUBO C, LIHENG X, KANG L, et al. Event extraction via dynamic multi-pooling convolutional neural networks [C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China: ACL, 2015: 167-176.
- [4] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [J/OL]. (2013-09-07) [2020-05-12]. Computation and Language. <https://arxiv.org/abs/1301.3781>.
- [5] 施喆尔, 陈锦秀. 基于语言模型及循环卷积神经网络的事件检测 [J]. 厦门大学学报(自然科学版), 2019, 58(3): 442-448.
- SHI J R, CHEN J X. Event detection via recurrent and convolutional networks based on language model [J]. Journal of Xiamen University (natural science edition), 2019, 58(3): 442-448.
- [6] ZENG Y, YANG H, FENG Y, et al. A convolution BiLSTM neural network model for Chinese event extraction [C]//Natural Language Understanding and Intelligent Applications. Berlin, Germany: Springer, Cham, 2016: 275-287.
- [7] 马晨曦, 陈兴蜀, 王文贤. 基于递归神经网络的中文事件检测 [J]. 信息安全学报, 2018(5): 75-81.
- MA C X, CHEN X S, WANG W X. Chinese Event Detection Based on Recurrent Neural Network [J]. Netinfo Security, 2018(5): 75-81.
- [8] HONG Y, ZHOU W, ZHANG J, et al. Self-regulation: Employing a generative adversarial network to improve event detection [C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia: ACL, 2018: 515-526.
- [9] LI Y, LI C, XU W, et al. Prior Knowledge Integrated with Self-attention for Event Detection [C]//China Conference on Information Retrieval. Guilin, China: Springer, 2018: 263-273.
- [10] NGUYEN T H, CHO K, GRISHMAN R. Joint event extraction via recurrent neural networks [C]//Proceedings

- of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego, California: NAACL, 2016: 300-309.
- [11] 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 Association for Computational Linguistics 2018, System Demonstrations. Melbourne, Australia: ACL, 2018: 50-55.
- [12] MINTZ M, BILLS S, SNOW R, et al. Distant supervision for relation extraction without labeled data [C]//Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP. Suntec, Singapore: ACL, 2009: 1003-1011.
- [13] YU W, YI M, HUANG X, et al. Make It Directly: Event Extraction Based on Tree-LSTM and Bi-GRU [J]. IEEE Access, 2020(8): 14344-14354.
- [14] 贺瑞芳, 段绍杨. 基于多任务学习的中文事件抽取联合模型 [J]. 软件学报, 2019, 30(4): 1015-1030.
DUAN R F, DUAN S Y. Joint Chinese Event Extraction Based Multi-task Learning [J]. Journal of Software, 2019 30(4): 1015-1030.
- [15] LIU X, LUO Z, HUANG H. Jointly multiple events extraction via attention-based graph information aggregation [J/OL]. (2018-10-23) [2020-05-12]. Empirical Methods on Natural Language Processing. <https://arxiv.org/abs/1809.09078v1>.
- [16] SANTOS C D, ZADROZNV B. Learning character-level representations for part-of-speech tagging [C]//Proceedings of the 31th International Conference on Machine Learning. Beijing, China: ICML, 2014: 1818-1826.
- [17] NGUYEN T H, GRISHMAN R. Event detection and domain adaptation with convolutional neural networks [C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China: ACL, 2015: 365-371.
- [18] 仲伟峰, 杨航, 陈玉博. 基于联合标注和全局推理的篇章级事件抽取 [J]. 中文信息学报, 2019, 33(9): 88-95.
ZHONG W F, YANG H, CHEN Y B. Document-level Event Extraction Based on Joint Labeling and Global Reasoning [J]. Journal of Chinese Information Processing, 2019, 33(9): 88-95.
- [19] 吴文涛, 李培峰, 朱巧明. 基于混合神经网络的实体和事件联合抽取方法 [J]. 中文信息学报, 2019, 33(8): 77-83.
WU W T, LI P F, ZHU Q M. Joint Extraction of Entities and Events by a Hybrid Neural Network [J]. Journal of Chinese Information Processing, 2019, 33(8): 77-83.

作者简介:



王晓浪(1995-) 男, 四川成都人, 硕士研究生, 主要研究方向为自然语言处理。E-mail: 2249127476@qq.com。



邓 蔚(1979-) 男, 四川自贡人, 博士, 硕士研究生导师, 分别在 2002 年和 2005 年毕业于重庆邮电大学, 获学士学位和硕士学位, 2018 年 10 月开始在重庆邮电大学做博士后研究。主要研究方向为知识图谱、机器学习、计算社会科学和算法伦理。E-mail: dengwei@swufe.edu.cn。



胡 峰(1978-) 男, 湖北武汉人, 教授, 2003 年至今在重庆邮电大学任教。主要研究方向为粗糙集、机器学习、数据挖掘。E-mail: hufeng@cqupt.edu.cn。

(编辑: 陈文星)