

基于自注意力机制和平均池化下图卷积网络的 商品新闻事件抽取

罗茜雅, 李红军, 王子怡, 甘晨灼, 胡正浩

成都理工大学 计算机与网络安全学院(示范性软件学院), 成都 610059

[摘要] 商品新闻事件抽取是对新闻非结构化语句进行归纳和表达,以提取出该语句所包含的事件以及相关信
息,能够为供货需求预测、价格预测、问答系统等提供基础。现有研究工作普遍存在候选触发词与实体向
量关联性利用不强以及参数角色提取准确率不够的问题,因此本文在已有研究工作的基础上,提出了一
种基于自注意力机制和平均池化图卷积网络及依赖解析树的提取模型(SAT-GCN-DPT)。该模型主要分
为3个模块,ComBERT预训练模块、self-attention机制下的触发分类模块、利用平均池化图卷积和依
赖解析树的参数角色分类模块。模型利用self-attention机制对输入数据进行操作增强候选触发词与实
体向量之间关联性,同时对图卷积结果使用平均池化函数进行信息聚合来更大程度地还原事件之间关
联性和提高分类准确率。实验结果表明,在CON数据集上,本文提出的模型在触发分类以及参数角色分
类的准确度以及F1值均有了提高。

[关键词] 商品新闻事件抽取;自注意力机制;平均池化函数;图卷积网络;依赖解析树

[分类号] TP391.1

[文献标志码] A

Extracting commodity news events based on self-attention mechanism and average pooling-based graph convolutional network

LUO Xiya, LI Hongjun, WANG Ziyi, GAN Chenzhuo, HU Zhenghao

College of Computer Science and Cyber Security(Pilot Software College), Chengdu University of Technology,
Chengdu 610059, China

Abstract: Commodity News event extraction involves analyzing unstructured sentences in news items to extract the information contained in them. Extracting information from news events on commodities can provide the basis for forecasting supply and demand, predicting prices, and developing question-answering systems. The existing researches generally have the problems that the correction between

[收稿日期] 2023-07-06。

[基金项目] 国家自然科学基金(42050104);自然资源部深时地理环境重建与应用重点实验室开放基金项目(DGERA20221102)。

[第一作者] 罗茜雅(1999—),女,硕士研究生,研究方向:自然语言处理、知识图谱,E-mail:1752591095@qq.com。

[通信作者] 李红军(1977—),男,副教授,研究方向:数据挖掘、人工智能,E-mail:lihongjun2012@cdut.edu.cn。

[引用格式] 罗茜雅,李红军,王子怡,等,2024. 基于自注意力机制和平均池化下图卷积网络的商品新闻事件抽取[J]. 成都理工大学学报(自然科学版),51(2):500-512.

Luo X Y, Li H J, Wang Z Y, et al., 2024. Extracting commodity news events based on self-attention mechanism and average pooling-based graph convolutional network[J]. Journal of Chengdu University of Technology (Science & Technology Edition), 51(2): 500-512.

candidate trigger words and entity vector is not strong and the accuracy of parameter role extraction is not enough. In this study, we propose a model to extract commodity news events (SAT-GCN-DPT) based on a self-attention mechanism, the average pooling-based graph convolutional network, and a dependency parse tree. The model is mainly divided into three modules: a ComBERT pre-training module, a module to classify trigger words based on the self-attention mechanism, and a module to classify parameter roles by using the average pooling-based graph convolution network and dependency parsing tree. The model uses the self-attention mechanism to manipulate the input data and enhance the association between the trigger words, while the results of graph convolution are aggregated by using the average pooling function to restore the association between events and improve the accuracy of classification. The results of experiments on the CON dataset showed that that the proposed model achieved high values of accuracy and the *F1* score on tasks of classifying the trigger words and parameter roles.

Key words: commodity news event extraction; self-attention mechanism; average pooling function; graph convolutional network; dependency parse tree

事件抽取是指从非结构化信息中抽取出事件,并以结构化的形式展现,抽取方式通常分为两种:第一种是基于流水线的抽取方式,即 pipeline 管道模式,其流程为先对事件触发词进行抽取,判断整体事件类型,再根据所给出的事件类型对参数进行识别,最后进行参数角色分类,但此方法容易产生因事件触发不准确而造成参数识别不准确的联级误差;第二种是基于联合模型的事件抽取,通常利用神经网络进行特征学习和联合学习,将事件触发词和参数进行联合抽取,即事件类别触发词辅助论元角色分类,同时论元信息辅助识别触发词,很好地解决了联级误差。

在事件抽取中,抽取分为两个大的子任务:(1)通过触发词对事件进行识别,该过程主要通过触发词进行识别从而对事件类型进行分类。(2)通过已确定的事件类型对论元进行识别和角色分类。目前大部分事件抽取模型所使用的数据集均为经典的 ACE05 数据集,该数据集注释了 8 个事件类别以及细分的 33 种子类型,如“life”“movement”等事件类别,属于通用事件数据集,在该数据集上进行的事件抽取任务则称为通用事件抽取;而金融领域的商品新闻语句中的事件类型与 ACE05 数据集中通用事件类型有着很大不同(Lee et al., 2021b),商品新闻句中通常存在着大量数字,包括价格及日期等,这些数字对金融事件的表达至关重要,因此商品新闻语句中事件的准确提取对商品价格预测、供货需求预测等下游任务有着重要作用(Yang et al., 2018);除此之

外,商品新闻事件句中还存在着参数同质性的特征,即同一类型的很多参数在金融事件中扮演着不同的参数角色,例如在句子(图 1)“The IEA boosted its estimate of output from U. S. shale drillers to 1.8 million bpd in July 4 holiday weekend, from 1.6 million bpd previously”中,数字“1.8”“1.6”和“4”均为特定数值但它们在句中却代表着不同的参数角色,其中“1.8”代表 IEA 组织对美国页岩钻探公司预估产出量要求增加后的产量数,“1.6”代表要求增加前产量数,“4”则表示该事件发生的具体日期。

针对以上提到的商品新闻事件句中的特定特征,通过相关实验发现现存的针对通用事件抽取的模型(Li et al., 2022)和部分针对商品新闻事件抽取的模型(Ding et al., 2014)无法很好地对商品新闻事件句中的特征进行处理,出现了对语句中候选触发词与实体向量关联性利用不足导致触发词识别不准确以及参数角色提取准确率不够的问题,即通用事件抽取模型不适合对有大量数值数据的金融领域文本进行事件抽取。为此,本文提出了一种新的针对商品新闻事件语句中几大特征的 SAT-GCN-DPT 联合提取模型,该模型大体包含 3 个部分:(1)通过 ComBERT 预训练模型对商品新闻数据集 CON 进行预处理,更好地对商品新闻数据集进行自适应训练,把握其特征;(2)触发分类和参数角色联合提取,其中触发分类模块采用 self-attention 机制对预训练后得到的数据进行操作来获得事件触发器列表,参数角色

分类模块利用剪枝解析树和平均池化图卷积对信息进行聚合;图卷积从事务句中的实体提及池中识别参与该事件的论元,同时根据已经识别出的论元对实体提及池中剩余实体的论元角色进行标记分类;(3)最后通过平均池化函数计算进行信息汇总,提高参数角色提取准确率。

The IEA **boosted** its estimate of output from U.S.shale drillers to **1.8 million bpd** in **July 4** holiday weekend, from **1.6 million bpd** previously Due to the increase in user demand for its products.

图 1 一条商品新闻的例句

Fig. 1 A example sentences from a commodity news article

注:展示了两个事件(1)IEA 组织对美国页岩钻探公司预估产出量增加的要求;(2)供不应求,其中触发词用粗体表示,具有相同实体类型的事件参数为 1.8 million bpd 和 1.6 million bpd,用彩色标注

本文在以往通用事件抽取模型和针对金融领域的商品新闻事件抽取模型存在的对候选触发词与实体向量关联性利用不足,而导致触发词识别不准确以及参数角色提取率不够的问题上做出了改进,主要贡献为:提出了一种新的针对商品新闻事件抽取的模型,能更好地对商品新闻事件中候选触发词与实体进行关联以及信息聚合,验证引入自注意力机制以及平均池化图卷积能够提升模型的相关性能。该模型在商品新闻数据集 CON 上取得了比以往基线模型更好的效果,验证了模型有效性。

1 事件抽取模型介绍

目前,事件抽取工作在通用事件抽取以及金融领域事件抽取中均提出了很多模型,例如在通用事件抽取中,Liu et al. (2018b)提出基于注意力机制和图卷积网络的模型,同时引入句法捷径弧增强信息流对文章中多个事件进行联合提取,提出了 JMEE 框架,针对普遍出现的多事件具有相关性给出了解决办法,提高了事件提取准确性同时也依赖了最短路径。受前者启发,Xu et al. (2020)提出了 JMCEE 框架,一种新的多中文事件联合抽取模型,该框架基于与训练语言模型的共享特征表示,将论元分类任务转化为关系抽取。GCN-FT 模型由 Zhang et al. (2018)提出,该模型使用完整的依赖解析树以及图卷积进行通用事件关系抽取。王士浩等(2021)则利用门控图卷积和动态依存池化对参数进行抽取,在 BERT 提取

句子语义特征后根据句子依存句法树构造邻接图进行抽取,除此之外在通用事件抽取中还有着大量其他利用深度学习模型的抽取方法,在通用事件抽取任务上取得了较好的效果。

在金融领域事件抽取中,陈星月等(2021)针对中文金融事件提出了基于 ELECTRA 模型与词性特征的抽取模型,将数据通过两个 ELECTRA 模型预处理后进行融合增强语义,同时利用 CRF 克服标签偏差来进行抽取。为了专门提取商品新闻中数字或数量单位短语的信息提取元组,Saha and Pal(2017)针对金融事件领域引入了的首个开放数字关系提取器;Lee et al. (2021a)针对商品新闻语句的特征提出 GCN-CST 模型。该模型使用 ComBERT 模型进行数据预处理,其次利用上下文子树联合 GCN 通过句法依赖关系进行触发词和参数联合抽取将句子表现为有向树,在相关单词间通过头修饰依赖弧对最相关单词进行卷积,最后通过最大池化 max-pooling 函数对将信息汇总完成对商品新闻事件的抽取。以上模型虽然在金融领域的商品新闻事件抽取上有了较好的实验结果,但在候选触发词与实体向量关联性利用率以及参数角色提取准确率上还存在着不足。

在以上方法中,通用事件抽取模型目前存在着无法准确利用商品新闻句中的大量数值特征的问题,金融领域事件抽取部分模型存在着对商品新闻事件中数值特征利用不足、触发词与实体向量关联性利用不够以及参数角色提取准确率不够等问题。综合上述相关工作,本文在商品新闻事件抽取任务中针对现有模型存在的问题,提出用基于自注意力机制和平均池化图卷积及依赖解析树(SAT-GCN-DPT)的抽取模型解决,该模型能更准确地识别事件类型、更好地对事件触发词与实体进行有效关联,在提高参数角色识别准确率的同时还提升了模型对商品新闻的抽取性能以及有效性。

2 方法

本文提出的 SAT-GCN-DPT 模型构架如(图 2)所示。该模型整体结构分为 3 个板块:ComBERT 预训练模块,Self-attention 机制下增强候选触发词与实体向量相关性的触发分类模块和利用平均池化图卷积与依赖解析树的参数角色分类模块。其中,ComBERT 预训练模块包含数

据输入,含有嵌入向量特征表示、词性标注、划分实体类别等操作;触发分类模块使用自注意力机制增强商品新闻事件句中候选触发词与实体向量的关联性;参数角色分类模块利用修剪过的唯一上下文子树而不使用完整依赖解析树,来避免完

整依存关系树中包含除关键信息外的其他噪音信息对参数角色分类造成的影响,同时在图卷积后利用平均池化函数将子树信息进行向量值的计算,以获得子树、触发以及实体的 3 个变量并将其连接最后得到参数角色分类。

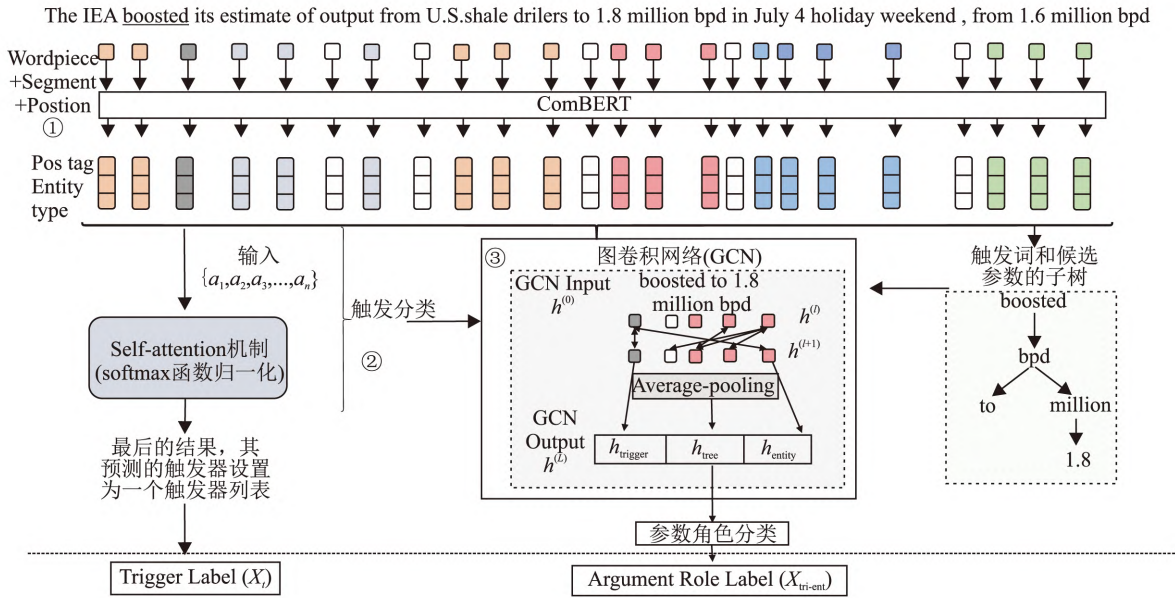


图 2 模型整体结构图
Fig. 2 Overall structure of the model

2.1 数据预处理

2.1.1 输入

针对金融领域中商品新闻语料库事件句的特征(Brandt and Gao, 2019),本文在数据输入过程前使用 CoreNLP 工具包对数据集中的商品新闻例句进行相应的句法分析,其中包括对句子进行拆分、分词、词性标注、黄金实体提及和实体类型识别标注,以及依存句法分析数据的输入以生成依存关系分析树等操作。同时,根据语料库中的商品新闻文章从 Bart Annotation standoff 格式转化为文本文件,在该过程中将新闻句中所有单词进行原形化,其中涉及动词时态以及人称的还原,名词复数变单数以及过去式还原等操作,更有利于对事件句中词语进行主干分析,同时产生基本依存关系并且进行数据增强。

对于模型的输入本文采用多通道模式(如图 2 中“①”所示)将以下几个组件进行连接。假设 $S = \{s_1, s_2, s_3, \dots, s_n\}$ 是输入长度为 n 的例句,其中 s_i 表示第 i 个节点。

(1) s_i 的词嵌入向量:属于预训练模型

ComBERT 单词嵌入的特征表示,该嵌入向量由 WordPiece 标记化组成,其中[CLS]和[SEP]分别位于句子的开头和结尾。

(2) s_i 的词性标注(POS 标注)标签嵌入向量:通过查找 POS 标记标签嵌入生成。

(3) s_i 的实体类型标签嵌入向量:与词性标注标签嵌入向量类似,使用 BIO 标注模式进行,即对句子的开头、中间以及不属于任何类型的部分进行标注,并将标注后的实体类型标签转化为实体向量进行计算。

2.1.2 自适应预训练模型:ComBERT

经典预训练模型 BERT(张震等, 2021)作为 Transformer 的编码部分,在通用事件抽取模型常使用的 ACE05 数据集上进行数据预处理时取得了极好的效果(Liu et al., 2021)。但本文没有使用 BERT 模型进行预训练,而是在商品新闻语料库上进一步预训练 BERT 模型来得到适应金融领域的 ComBERT 模型,利用 ComBERT 进行预训练。这是因为商品新闻数据集 CON 中数据存在着有特定多义词、数值量大以及一词多义的

特征,例如“corporate charter”一词,在句中既可表达为营业执照,也可表达为公司章程;“bond”一词既有票据、债券之意,同时也有“投标保证金”的意思(Sheng et al., 2021)。通常情况下,通用事件抽取方法所使用的 BERT 模型在预训练时对于这类词语的提取不太敏感,无法准确地对句中相关元素进行提取,ComBERT 模型则针对上述提到的商品新闻句中多数值、参数同质性、一词多义性等特征,在商品新闻语料库 CON 上做自适应训练,更完整准确地表达出有着特定特征的词语,对输入事件句中拥有该类特征的词语进行更好的理解,以确保事件触发识别以及参数角色分类的准确性。

同时,对 ComBERT 预训练模型的设置按照 Devlin et al. (2018)对 BERT 模型的变压器和超参进行相同设置,具体设置为:将模型层数,即 Transformer 变压器块表示为 L ,将 Hidden size 隐藏大小表示为 H ,同时设置 self-attention 自我注意头部数量为 A ,模型参数大小表示为 ComBERT($L=12, H=768, A=12$, Total Parameters = 110 M),且用 bert-base-cased 初始化 ComBERT 预训练模型,对事件中的词汇大小写进行区分,因为在金融领域事件句中常常出现人名地名以及国家等类似名词,区分大小写有利于在提取过程中进行更精准地识别,同时有利于理解相关上下文信息的表达。

2.2 基于 Self-attention 的触发分类模块

使用 ComBERT 模型对数据进行预处理后进入模型的联合抽取模块,该模块细分为触发提取分类和参数角色分类。模型在触发分类模块中利用 self-attention 机制(Liu et al., 2018a),该机制相比令牌分类会增加句中词性的关联,即候选触发词与实体向量之间关联。若是通过 MLP 进行简单的邻居信息汇聚可能会造成事件触发的不准确性,而 self-attention 机制不仅可以得到源端与目标端词与词之间的依赖关系,同时还可以有效获取源端或目标端自身词与词之间的依赖关系且更容易捕获句子中长距离的相互依赖的特征。MLP 则需要依次序进行序列计算,对于远距离的相互依赖的特征则需要经过若干时间步骤的信息累积才能将两者联系起来。但自注意力机制在计算过程中会直接将句子中任意两个单词的联系通过一个计算步骤直接串联起来,因此远距离依赖

特征之间的距离被极大缩短,有利于更有效地利用这些特征(Qiao et al., 2022)。除此之外,自注意力机制对于增加计算的并行性也有直接帮助作用。

将 ComBERT 预训练后的最后一层输出实体向量 $A = \{a_1, a_2, \dots, a_n\}$ 作为触发分类模块自注意力机制的输入,在自注意力机制中会计算新矩阵 Q_i, K_i, V_i (矩阵值会在反向传播中一直更新)分别是输入的实体向量 a_i 与一个随机初始化矩阵相乘的结果。因为要满足该矩阵的第二个维度与向量 a_i 的维度相同,本文中单词嵌入的维度数为 768,则令该矩阵维度为 (64, 768)。将 A 输入自注意力机制,得到 Q, K, V 后经过 softmax 函数进行归一化再乘以矩阵 V 得到最终的输出结果 $B = \{b_1, b_2, \dots, b_n\}$,详细计算过程如下:

$$Q_i = W_{q(i)} \cdot a_i \quad (1)$$

$$V_i = W_{v(i)} \cdot a_i \quad (2)$$

$$K_i = W_{k(i)} \cdot a_i \quad (3)$$

$$b_i = \text{softmax}\left(\frac{Q_i \cdot K_i^T}{\sqrt{d_k}}\right) \cdot V_i \quad (4)$$

$$B = \{b_1, b_2, \dots, b_n\} \quad (5)$$

式中: $W_q \in R^{d_m \times d_q}, W_v \in R^{d_m \times d_v}, W_k \in R^{d_m \times d_k}$ 均属于随机初始化矩阵;(4)式得到自注意力机制下捕获的嵌入字符向量间潜在的各个语义信息矩阵,通过其关联性强弱把在此过程中预测的所有触发器添加到一个列表 X_i 中,其中 $X_i = \{X_1, X_2, X_3, \dots, X_i\}$ 。

2.3 参数识别角色分类模块

2.3.1 上下文子树

在利用图卷积进行参数角色分类之前,本文将图 1 例句进行依存关系分析得到其依存关系树,但在该部分并不使用其完整的依赖解析树,而是对依赖树进行剪枝得到其子树(详见图 2 右侧子树),该子树以触发器候选和参数提及候选之间的最小公共祖先为根。同时由于解析树中的偏离节点路径会包含重要的隐藏信息以及上下文信息使参数角色分类能够得到更准确的结果,因此进行剪枝的子树会包含这些偏离节点路径(图 3 左右两侧分别显示了带有偏离路径的 LCA 子树和不带偏离路径的 LCA 子树)来更大程度地利用隐藏信息以及上下文信息,其中偏离路径信息由与依赖路径距离为 DIST 的标记节点组成。

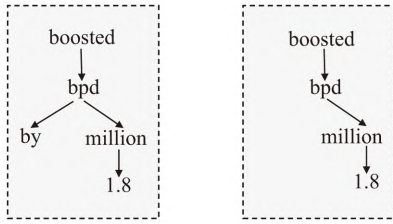


图 3 带有偏离路径的 LCA 子树和不带偏离路径的 LCA 子树

Fig. 3 LCA subtree with deviating paths and LCA subtree without deviating paths

在以往的剪枝依赖解析树的工作中,Zhang et al. (2018)给出了子树构建的相关具体步骤,其中 $DIST=1$ 在最大程度保留了相关上下文信息以及在去掉部分不相关信息上实现了最有效的平衡。该算法的目的是得到包含偏离路径的子树结构并且将头索引转化为树对象。在算法整体思想中,若偏离路径信息与依赖路径信息的 $DIST$ 值小于零,代表两路径的距离为负,即不存在定义的偏离路径,那么将会构建整棵子树;而如果两路径距离 $DIST$ 大于零,则首先查找触发器和参数的所有祖先节点,在找到两者最小公共祖先时生成两者之间的公共节点,最后根据距离远近再插入更多的标记节点来远离前一步得到的公共节点。

上述提到的(该)算法通过对完整依赖解析树进行剪枝,去掉了冗长噪音信息的同时保留了影响较大的上下文相关信息,因此该算法在对候选参数在已有触发候选对应的角色扮演判断上起到重要的作用。本文模型利用该算法,同时使用上下文子树会针对参数在商品新闻事件中扮演的角色进行分类,将候选参数所扮演的角色划分为 19 个参数角色之一,同时产生的带有偏离路径的子树会更加准确地对商品新闻事件句中参数所对应的参数角色进行提取识别,在剪枝操作完成之后会将得到的子树句转化为有向图输入到图卷积网络中进行下一步操作。

2.3.2 平均池化 GCN 下的参数角色分类

将前一模块产生的句子句法分析树当作一个有向图 T ,假设 $T=\{V,E\}$ 是句子 S 的依存树,其中 V 代表 T 的节点, E 代表树的边,那么 V 即为句子 S 中 n 个节点的集合 $\{s_1, s_2, s_3, \dots, s_n\}$,每条边 $(V_i, V_j) \in E$ 则表示从中心词 s_i 指向具有依赖关系的依赖词 s_j ;给定一个句子,由前文可知其句

法依存树就拥有对应的 n 个节点,接着将得到的句法分析树转化成其对应的 $n \times n$ 邻接矩阵 A ;同时在图卷积 GCN 层堆叠 L 次(L 为超参数),在每次卷积过程中都会收集汇总每个图节点的信息(Nguyen and Grishman, 2018)。设 $h^{(0)}$ 为卷积过程中的初始输入向量, $h^{(L)}$ 为输出向量,GCN 卷积过程中的具体运算如下:

$$h_i^{(L)} = \alpha \left(\sum_{j=1}^n \frac{A_{ij} S^{(L)} h_j^{(L-1)}}{d_i} + b^{(L)} \right) \quad (6)$$

式中: $h_j^{(L-1)}$ 为初始输入字向量; $h_i^{(L)}$ 为集合隐藏表示; $S^{(L)}$ 为权重矩阵; α 是 sigmoid 函数; $b^{(L)}$ 是偏差项; $d_i = \sum_{j=1}^n A_{ij}$ 是结果图中边的数量。

在该模块中定义触发器为 t ,实体为 e ,给定一个触发器—实体对 $t_x e_y$,根据参数角色分类模块中第一部分上下文子树构造里提到的剪枝算法对完整依赖解析树进行剪枝,然后得到触发器 t_x 和实体 e_y 之间的上下文子树,子树中的节点(句子 S 的节点)会形成进行图卷积的初始输入向量 $h^{(0)}$,在经过 L 次卷积之后子树的表示结果为:

$$h_{tree} = g(h^{(L)}) = g(GCN(h^{(0)})) \quad (7)$$

式中: $h^{(L)}$ 为经过 L 层图卷积之后的输出字向量; g 是将输入向量映射到子树的平均池化函数。

一般来说,在利用池化函数进行特征提取时存在的误差主要来自两个方面,一是领域大小受限造成的估计值方差增大,二是卷积层参数误差造成的估计值偏移(Chen et al., 2015),在常见的两种池化函数:最大池化和平均池化中,平均池化能减少第一种误差,更多地保留事件句中的语义背景信息,很大程度上契合本文针对商品新闻事件句抽取时对语句中事件所属语境及背景要求准确的特征。例如“cash”一词,在不同的金融语境下既可理解为“现货”,又可理解为“现金”,句子“the commodity’s cash market price”的意思是“商品的现货市场价”,“cash”在该语境下意为“physical market”,即“现货市场”;而在句子“a liquid cash market”中 cash 则表示现金,整句意为“流动现金市场”。利用平均池化函数能很好地根据该问题进行特征提取来准确识别金融事件句中的事件语境以及背景,同时句中候选参数的角色识别分类也需要一个准确的结果,因此本文在该部分使用平均池化图卷积网络对商品新闻事件

进行特征提取。

在该模块中,除了对子树进行了表示之外还对句中的触发器和实体进行了表示:

$$h_{\text{trigger}} = g(h_t^{(L)}) \quad (8)$$

$$h_{\text{entity}} = g(h_e^{(L)}) \quad (9)$$

这里依旧采用平均池化函数进行计算,最后将表示出的 3 个向量 $h_{\text{tree}}, h_{\text{trigger}}, h_{\text{entity}}$ 连接成一个向量,然后通过全连接层进行传播实现对参数角色进行分类的目的,具体运算如下:

$$X_{\text{tri-ent}} = F(S_a[h_{\text{tree}}; h_{\text{trigger}}; h_{\text{entity}}]) + b_a \quad (10)$$

式中: F 是识别参数角色的 *softmax* 函数; $X_{\text{tri-ent}}$ 是实体即参数 *entity* 在触发器 *trigger* 上角色表达的最终输出,最后通过最小化交叉熵损失函数对该联合提取进行训练:

$$L_{\text{joint}} = L_{\text{trigger}} + \beta(L_{\text{arg}}) \quad (11)$$

式中: β 是参数角色提取任务损失的权重,在实验中,我们使用其值为 2,训练模型时使用双倍的参数提取权重。

3 实验设置

3.1 数据集

本文使用的 CON 数据集是由带有注释的多个商品新闻文章的集合组成的(Lee et al., 2022),其注释基于 ACE2005 规范程序引入的标准。与通用事件抽取所常用的数据集 ACE2005 不同,此数据集包含 21 个参数类型,18 个事件类型和 19 个参数角色类型,其中 18 个事件类型主要可分为 4 个大类:(1)供需关系;(2)宏观经济;(3)地缘政治;(4)商品价格变化。并且每个事件类别都有属于自己的参数角色集,CON 数据集总共包含 8 850 个参数提及和 3 949 个金融事件,表 1 显示了其事件类型。

3.2 实验参数设置

深度学习模型的搭建采用 PyTorch1.2,模型参数如表 2 所示,模型使用 ComBERT 作为编码器,数据集中训练集和测试集的划分比例为 7 : 3,单词嵌入维度为 768,词性标注嵌入和实体类型嵌入维度均为 50,在反向传播和训练阶段时使用 *batch_size* 为 4, *dropout* 为 0.35,隐藏层维度为 200 的双层 GCN 进行实验,采用 Adam 优化器,学习率设置为 $2e-05$,训练轮次为 40,并且在实验最后使用最小交叉熵损失函数对模型进行训练以达到最优效果。

表 1 数据集事件类别以及句子数

Table 1 Categories of events and number of sentences in the dataset

事件类型	类型比例/%	句子数/个
Cause-movement-down-loss	13.35	524
Cause-movement-up-gain	2.23	88
Civil-unrest	2.53	100
Crisis	0.76	30
Embargo	3.75	148
Geopolitical-tension	1.70	67
Grow-strong	6.03	238
Movement-down-loss	22.69	896
Movement-flat	1.52	60
Movement-up-gain	22.13	874
Negative-sentiment	4.79	189
Oversupply	2.63	104
Position-high	3.82	151
Position-low	3.11	123
Prohibiting	1.06	42
Shortage	1.04	41
Slow-weak	5.47	216
Trade-tensions	1.39	55

表 2 模型参数设置

Table 2 Model parameters

参数	值
Hidden layer dimension	200
Optimizer	Adam
Training batch size	4
Dropout	0.35
POS tagging embedding	50
Epoch	40
Entity type embedding	50
Learning rate	$2e-05$
Word embedding dimension	768

3.3 评价指标

本文采用 P 准确率、 R 召回率、 $F1$ 值作为评价指标对提出的针对金融领域商品新闻事件抽取模型 SAT-GCN-DPT 的准确性进行评估, $F1$ 的计算公式选择最常见的方法,具体如下:

$$F1 = \frac{2 \times P \times R}{P + R} \times 100\%$$

式中: P 表示该模型正确预测的样本数量占全部预测为正确样本数量的百分比; R 表示模型正确

预测的样本数量占实际预测正确样本数量的百分比;F1 为 P 和 R 线性计算得出。

4 实验结果及分析

4.1 预训练模型比较及结果

本文所使用的预训练模型是基于 BERT 模型在金融领域语料库进行自适应训练后得到的更适用于商品新闻事件抽取的 ComBERT 模型,下面将在本文所使用的 CON 数据集上分别用其他两种预训练模型做对比实验,包括 BERT, RoBERTa 模型,实验结果如图 4.5 所示。

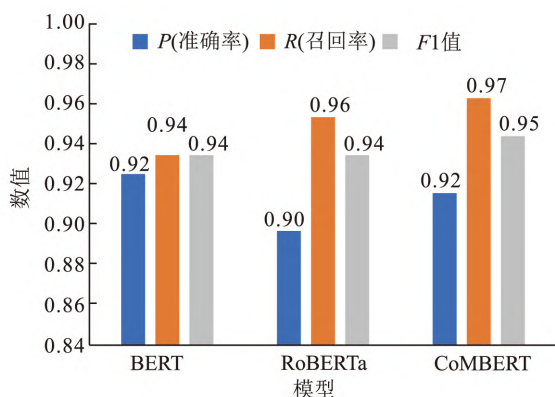


图 4 各预训练模型在触发分类部分的 P, R, F1 的值
Fig. 4 Values of P, R, and F1 of each pre-trained model in terms of the classification of trigger words

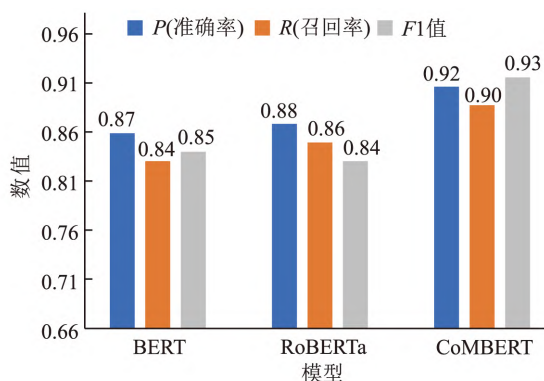


图 5 各预训练模型在参数角色分类的 P, R, F1 的值
Fig. 5 Values of P, R, and F1 of each pre-trained model in terms of classifying parametric roles

图 4 可以看出在触发分类部分, ComBERT 预训练模型在 P 值上与 BERT 相比虽然无变化, 但比 RoBERTa 增加了 2%; 其 R 召回率的值相比 BERT 高出了 3%, 相比 RoBERTa 高出了 1%; F1 值也均比 BERT, RoBERTa 模型高出 1%。

由图 5 可以看出在参数角色分类部分, 预训练模型 ComBERT 的 P 值相比 BERT 和 RoBERTa 分别高出 5% 和 4%, R 值相比于前两种模型分别高出 6% 和 4%; 而其 F1 值较前两个模型至少有 8% 的提升, 与 RoBERTa 相比更是高出了 9%。由此可见基于 BERT 在金融领域语料库上进行自适应训练后的 ComBERT 预训练模型能够更好地对商品新闻事件句中参数角色以及分类进行准确识别, 同时在触发分类部分也具有更高的准确率。

4.2 基线模型设置

为验证本文提出模型的更优性, 将其与以下基线模型在 P, R 以及 F1 值进行对比。

(1) 模型 A: JMEE (Xu et al., 2020) 模型进行的是对多事件的联合抽取, 其中运用到了图卷积网络并且引入了句法快捷弧来增强信息流。

(2) 模型 B: GCN-FT 模型使用的是未修剪过的完整依赖树对参数识别及分类部分进行卷积运算。

(3) 模型 C: GCN-LST 模型依靠触发候选和对应实体候选的最短依赖路径进行事件抽取。

(4) 模型 D: GCN-CST (Lee et al., 2021a) 模型利用 GCN 和经过修剪的上下文子树对金融领域事件进行触发词和参数的联合抽取。

(5) 模型 E: SAT-GCN-DPT 本文模型, 在触发分类模块利用 self-attention 机制增强候选触发词关联性, 同时利用平均池化图卷积和剪枝依赖解析树进行参数角色识别分类, 联合对金融领域事件进行抽取。

4.3 实验结果以及分析

表 3 展示了各基线模型以及本文提出模型在触发分类以及参数角色分类模块的 P, R, F1 值。由表 3 结果对比可知, 对于触发词分类模块, 各个模型在准确率、召回率以及 F1 上都有着较高的值, 但本文所提出的 SAT-GCN-DPT 模型因在触发分类模块使用了 self-attention 机制使输入向量对其自身词与词之间的依赖关系有了更充分的利用, 且更容易捕获句子中长距离的相互依赖的特征, 使得触发提取的 3 个评估结果相比于前 4 个模型有了更好地提升, 同时可以看出不管是依赖完全树还是经过修剪的唯一上下文子树都对触发提取有着很好的辅助效果。在参数角色分类模块, 因为模型 A 未使用解析树, 其 3 个评估指标

结果显然没有 B, C, D, E 模型好, 而针对模型 B 跟 C 来说, 其 $P, R, F1$ 值虽较模型 A 有了提升, 但其结果却没有模型 D 更优, 原因在于模型 D 即 Lee et al. (2021a) 提出的 GCN-CST 模型利用的

是修剪过的唯一上下文子树, 而不是完整的依存关系树, 完整依存关系树中会包含除关键信息外的其他噪音信息对论元角色分类的工作不仅没有帮助反而会对其造成一定的影响。

表 3 各类模型触发分类和参数角色分类比较结果 (P 准确度、 R 召回率、 $F1$ 值)

Table 3 Comparative results of classification of trigger words and parametric roles by the models (P , accuracy; R , recall; $F1$, $F1$ score)

方法	触发分类			参数角色分类		
	P	R	$F1$	P	R	$F1$
模型 A(JMEE)	0.89	0.91	0.90	0.72	0.76	0.75
模型 B(GCN-FT)	0.92	0.94	0.92	0.74	0.75	0.73
模型 C(GCN-LST)	0.93	0.95	0.94	0.87	0.86	0.85
模型 D(GCN-CST)	0.91	0.96	0.93	0.90	0.89	0.90
模型 E(SAT-GCN-DPT)	0.91	0.97	0.94	0.93	0.93	0.93

总体来说, 在触发分类模块中, 本文提出的 SAT-GCN-DPT 模型相比于经典基线模型 B 在 R 值上有了 3.0% 的提升, 在 $F1$ 值上有了 2.0% 的提升。相比于结果较好的模型 D, 本文提出的模型在 $F1$ 值和 R 值上也均有 1.0% 的提升; 而在参数角色分类模块, 本文提出的模型较前几个基线模型 A, B, C 在 $P, R, F1$ 值上分别至少有 6.0%, 7.0%, 8.0% 的提升, 较结果较好的模型 D 在 P 值和 $F1$ 值上也均有 3.0% 的提升, 且在 R

值上有了 4.0% 的提升, 这些数据足以说明该模型的有效性。

表 4 则列举出了本文提出的 SAT-GCN-DPT 模型以及各基线模型在 CON 数据集上进行金融事件抽取的具体结果, 结果以抽取对应金融事件类型的句子数进行表示, 表中给出了部分金融事件类型各自在 CON 数据集中所占句子总数, 同时展示了各模型抽取出的对应金融事件类型句子数, 表 5 则对各模型抽取的 $F1$ 值进行了展示。

表 4 各类模型从 CON 数据集中提取部分金融事件所对应的句子数

Table 4 The number of sentences corresponding to the extraction of partial financial events from the CON dataset by each model

方法	金融事件类型名称和相应句子数量			
	Crisis(30)	Movement-down-loss(896)	Oversupply(104)	Position-low(123)
模型 A(JMEE)	25	816	94	110
模型 B(GCN-FT)	26	833	96	112
模型 C(GCN-LST)	28	842	97	114
模型 D(GCN-CST)	27	851	98	116
模型 E(SAT-GCN-DPT)	28	869	99	117

表 5 从 CON 数据集中提取部分金融事件各类模型 $F1$ 值对比

Table 5 Comparison of $F1$ values of each model for the extracted financial events from the CON dataset

方法	抽取金融事件类型名对应的 $F1$ 值			
	Crisis	Movement-down-loss	Oversupply	Position-low
模型 A(JMEE)	0.85	0.91	0.90	0.89
模型 B(GCN-FT)	0.88	0.93	0.92	0.91
模型 C(GCN-LST)	0.95	0.94	0.93	0.92
模型 D(GCN-CST)	0.94	0.95	0.94	0.93
模型 E(SAT-GCN-DPT)	0.95	0.97	0.95	0.96

由表 4 可知, 针对 Crisis 事件类型, 因其在数据集中所占句子总数较少, 因此各个模型所触发

识别出为该事件类型的句子数差别不是很大; 而对于 Movement-down-loss 事件类型的抽取, 由

表 4 可知模型 A 因未使用解析树导致其抽取结果相比于后面 4 个模型来说较差,模型 B,C 因使用完全解析树进行辅助识别使得其抽取结果为 Movement-down-loss 事件类型的句子数相较于模型 A 来说有了很大提升,其句子数分别对应增加了 17 和 26,但相比于使用修剪过的唯一上下文子树的模型 D 来说其抽取效果还是存在一定差距。而本文提出的模型 E 因在触发分类模块加入了 self-attention 机制,进一步增强了触发候选与实体向量之间的关联性,使得最终抽取结果相较于前 4 个模型有了大幅提升,具体表现在相较于结果较好的模型 D 来说其抽取为 Movement-down-loss 事件的句子数也有 18 的数值增加,为 5 个模型中最好的结果;对于 Position-low 事件类型的抽取结果,本文提出的模型 E 在抽取结果上相比于模型 A,B,C 也均有了不同提升,进一步验证了本文提出模型 E 的有效性。

总体来说,由表 5 可知,在金融事件类型抽取的具体结果中,本文所提出的 SAT-GCN-DPT 模

型的抽取性能领先于各基线模型,所展示的抽取结果 F1 值也均优于前几个基线模型;具体表现为:本文所提出的模型 E 相较于模型 A,B 在所选 4 个金融事件类型抽取结果的 F1 值上分别至少有 7.0%,4.0%,3.0%,5.0% 的提升,相较于模型 C 在 Movement-down-loss, Oversupply 以及 Position-low 事件类型的抽取结果的 F1 值分别有 3.0%,2.0%,4.0% 的提升,相比于结果较好的模型 D 来说在 Crisis, Movement-down-loss, Oversupply 以及 Position-low 事件类型的抽取结果上也分别有着 1.0%,2.0%,1.0%,3.0% 的提升。这些数据足以说明本文提出的 SAT-GCN-DPT 模型在触发识别部分相较于基线模型的更优性。

表 6 展示了各个模型针对实体参数同质性对其参数角色进行细分工作的 F1 值,表中的 ♠、❖、♣ 符号表示按照实体类型对参数角色进行分类。图 6~8 则对表 6 中数据进行了可视化,绘制出其折线图。

表 6 各类模型针对实体的参数同质性对其参数角色细分的 F1 值结果

Table 6 F1 values of each model for the parametric homogeneity of entities with regard to their role in segmentation

参数角色	实体类型	参数角色细分 F1 值				
		模型 A	模型 B	模型 C	模型 D	模型 E
Final value ♠	Money/Production unit/Price unit/Percentage/Quantity	0.38	0.70	0.74	0.78	0.85
Initial value ♠	Money/Production unit/Price unit/Percentage/Quantity	0.57	0.73	0.68	0.76	0.84
Difference ♠	Money/Production unit/Price unit/Percentage/Quantity	0.68	0.83	0.88	0.89	0.92
Reference point ❖	Date	0.69	0.79	0.70	0.80	0.83
Initial point ❖	Date	0.62	0.63	0.62	0.66	0.71
Contract date ❖	Date	0.53	0.70	0.67	0.80	0.86
Supply consumer ♣	Country/State/Province	0.71	0.72	0.73	0.79	0.83
Impacted countries ♣	Country	0.68	0.72	0.71	0.76	0.83
Participating countries	Country	0.74	0.78	0.81	0.82	0.87

由表 6 可知,CON 数据集中的数据表现出了商品新闻事件中特有的参数同质性的特征,即相同实体类型的参数在事件中扮演着不同角色。对比各模型的 F1 值可知,模型 A 在该工作中的效果不佳,因其未使用依赖解析树对参数角色分类进行辅助,对于有着参数同质性的实体进行角色区分的结果不尽人意;模型 B,C 虽使用了依赖解析树,但因其解析树是完整的,所包含的除关键信息外的噪声信息会极大的影响有着参数同质性实体的参数角色划分,故两者只在极少数实体类型的参数角色细分工作中有着较好的结果,例如在

图 6 中可知模型 B,C 只在实体类型为 Money/Production unit/Price unit/Percentage/Quantity 的实体细分其参数角色为“Difference”时才有较高 F1 的值,在图 8 中模型 C 在对实体类型为 Country 的实体细分其参数角色为“Participating countries”时才有较高的 F1 值。

而模型 D,E 因使用了进行剪枝的依赖解析树对参数角色进行辅助分类,去掉了噪音信息对分类造成的不良影响,故两者在该项工作中除了在极个别实体类型的实体对应的参数角色细分时有着不太高的 F1 值外,对其他实体类型进行角

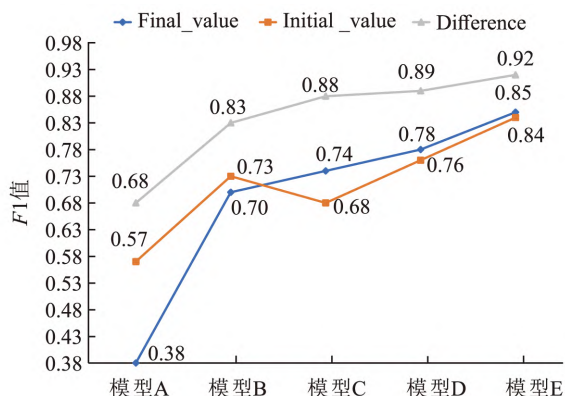


图 6 各模型对实体类型为 Money/Price 实体细分其参数角色 F1 值折线图

Fig. 6 Each model subdivides the parameter roles. The line graph of F1 values for entities of type Money/Price

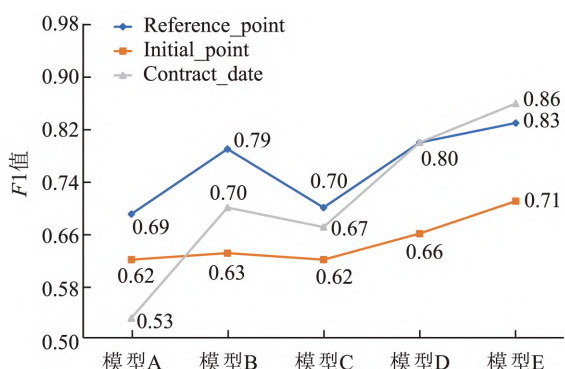


图 7 各模型对实体类型为 Date 实体细分其参数角色 F1 值折线图

Fig. 7 Each model subdivides the parameter roles. Line graph of F1 values for entities of type Date

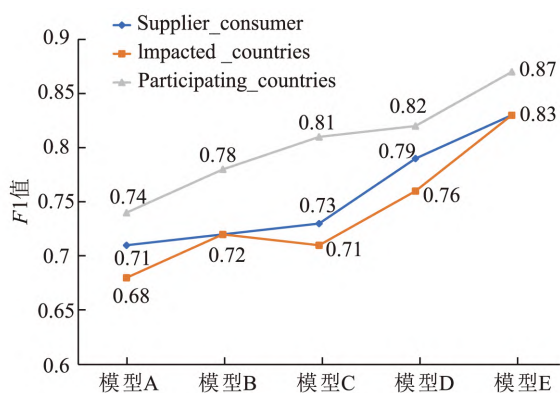


图 8 各模型对实体类型为 Country 实体细分其参数角色 F1 值折线图

Fig. 8 Each model subdivides the parameter roles. Line graph of F1 values for entities of type Country

色细分的结果均优于前 3 个模型。

特别地,相比于模型 D,因模型 E 在参数角色

分类模块的 GCN 中使用了平均池化函数对信息进行聚合,针对商品新闻事件中对背景信息要求高的条件进行了更好地满足,因此在对相同实体类型的参数进行参数角色细分时有着比模型 D 更好的 F1 值。例如在图 6 中可知模型 E 在对实体类型为 Money/Production unit/Price unit/Percentage/Quantity 的实体细分其参数角色为“Difference”时,其 F1 值高达 0.92,相比于模型 D 有着 3%的提升。同样地,在图 7 可知,模型 E 在对实体类型为 Date 的实体细分其参数角色为“Contract date”时,其 F1 值高达 0.86,相比于模型 D 的 F1 值有着 6%的提升。

综上所述,表 6 的结果清楚地展现出使用进行剪枝的依赖解析树能更好地对有着相同实体类型的参数进行角色分类,同时展现出本文提出的在 GCN 中利用平均池化函数进行信息聚合的模型有利于针对商品新闻事件中参数同质性的特征对参数角色进行更好的细分,同时得到更高的 F1 值。

5 模型实际意义

本文提出的 SAT-GCN-DPT 模型通过在触发分类模块设置 self-attention 机制增强候选触发与实体向量的关联性以提高触发分类的准确性,在参数角色分类模块利用平均池化函数进行图卷积以及利用上下文子树来提高参数角色识别准确率,提取出的准确信息可以用于商品供求关系以及价格预测等下游任务,例如某新闻报道某国家的核能发电厂由于技术问题需要进行紧急停机维护,则可通过提出的抽取模型进行准确抽取后得到该事件的供求关系和价格预测:①供求关系:由新闻中触发词“shutdown maintenance”可将其事件类型划分为“shortage”即“短缺”,得到该国家电力供应可能会暂时短缺的关键信息,判断会出现因电力短缺而导致电力价格上涨的情况,供求关系会发生变化,同时电力供应不足也可能导致其他行业的生产能力下降,进一步影响到其相关商品的供求关系;②价格预测:针对触发结果对参数进行角色分类,将文本中出现的“family”“enterprise”识别为“victim”角色,即家庭和企业会因电力短缺面临更高用电成本,家庭可能会减少对电力设备(如空调、电热水器)的购买和使用,而企业也会考虑以提高产品价格来应对成本

上涨,由此可根据出现的家电销售量下降或者制造商上调价格的行为来对相关商品价格的趋势走向进行进一步预测。

6 总结

在金融领域的商品新闻事件抽取任务中,文本触发词与实体向量关联性的有效利用以及参数角色的准确提取对该任务结果有着至关重要的影响,针对现存抽取方法的不足,本文提出一种基于自注意力机制和平均池化图卷积网络及依赖解析树的提取模型(SAT-GCN-DPT),实验结果表明,本文提出的模型:

a. 对该领域事件数值数据量大、触发词与实体向量间关联性利用率低以及同一类型的参数扮演不同角色的问题即参数同质性问题进行了解决,利用自适应金融领域的ComBERT预训练模型进行数据预处理,通过self-attention机制增加触发词与实体向量之间关联以及在图卷积部分利用平均池化函数进行信息汇聚提高角色细分准确性。

b. 在CON数据集上无论是触发分类板块还是参数角色分类板块都获得了比基线模型更好的效果,其 $P, R, F1$ 值均有了一定程度的提高,同时对商品信息提取的准确性也有了提高。

该模型的提出对未来的工作,例如在获得关键上下文后对商品供求关系以及价格进行预测的工作上也提供了坚实的基础。

[参 考 文 献]

陈星月,倪丽萍,倪志伟,2021. 基于ELECTRA模型与词性特征的金融事件抽取方法研究[J]. 数据分析与知识发现,5(7):36-47.

Chen X Y, Ni L P, Ni Z W, 2021. Extracting financial events with ELECTRA and part-of-speech[J]. Data Analysis and Knowledge Discovery, 5(7):36-47. (in Chinese)

王士浩,王中卿,李寿山,等,2021. 基于门控图卷积与动态依存池化的事件论元抽取[J]. 计算机科学,48(增刊2):52-56.

Wang S H, Wang Z Q, Li S S, *et al.*, 2021. Event argument extraction using gated convolution and dependency pooling[J]. Computer Science, 48(S2):52-56. (in Chinese)

张震,谷雨,耿小航,等,2021. 基于预训练模型和特征融合的事件触发词抽取[J]. 杭州电子科技大学学报(自然科学版),41(4):40-47.

Zhang Z, Gu Y, Geng X H, *et al.*, 2021. Event trigger word extraction based on pre-trained model and feature fusion[J]. Journal of Hangzhou Dianzi University (Natural Sciences), 41(4):40-47. (in Chinese)

Brandt M W, Gao L, 2019. Macro fundamentals or geopolitical events? A textual analysis of news events for crude oil [J]. Journal of Empirical Finance, 51:64-94.

Chen Y B, Xu L H, Liu K, *et al.*, 2015. 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: Association for Computational Linguistics:167-176.

Devlin J, Chang M W, Lee K, *et al.*, 2018. Bert: Pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the North-American-Chapter of the Association-for-Computational-Linguistics-Human Language Technologies. Minneapolis: Association for Computational Linguistics:4171-4186.

Ding X, Zhang Y, Liu T, *et al.*, 2014. Using structured events to predict stock price movement: An empirical investigation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics:1415-1425.

Lee M, Soon L K, Siew E G, 2021a. Effective use of graph convolution network and contextual sub-tree for commodity news event extraction [C]//Proceedings of the 3rd Workshop on Economics and Natural Language Processing. Punta Cana: Association for Computational Linguistics:69-81.

Lee M, Soon L K, Siew E G, *et al.*, 2021b. An annotated commodity news corpus for event extraction[J]. <https://doi.org/10.48550/arXiv.2105.08214>.

Lee M, Soon L K, Siew E G, *et al.*, 2022. CrudeOilNews: An annotated crude oil news corpus for event extraction[C]//Proceedings of the 13th International Conference on Language Resources and Evaluation. Marseille: European Language Resources Association:465-479.

Li Q, Peng H, Li J X, *et al.*, 2022. Reinforcement learning-based dialogue guided event extraction to exploit argument relations[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 30:520-533.

Liu J, Chen Y B, Liu K, *et al.*, 2018a. Event detection via gated multilingual attention mechanism [C]//Proceedings of the 30th Innovative Applications of Artificial Intelligence Conference. New Orleans: AAAI

- Press;4865—4872.
- Liu X, Luo Z C, Huang H Y, 2018b. Jointly multiple events extraction via attention-based graph information aggregation[C]//Proceedings of the Empirical Methods in Natural Language Processing, Brussels; Association for Computational Linguistics;1247—1256.
- Liu Z, Lin W, Shi Y, *et al.*, 2021. A robustly optimized BERT pre-training approach with post-training[C]//Proceedings of the 20th China National Conference on Computational Linguistics. Hohhot; Springer Science and Business Media Deutschland GmbH;471—484.
- Nguyen T H, Grishman R, 2018. Graph convolutional networks with argument-aware pooling for event detection[C]//Proceedings of the 32nd AAAI Conference on Artificial Intelligence, New Orleans; AAAI Press; 5900—5907.
- Qiao X M, Tang Y, Liu Y H, *et al.*, 2022. A event extraction method of document-level based on the self-attention mechanism[C]//Proceedings of the 4th International Conference on Machine Learning for Cyber Security. Guangzhou; Springer Science and Business Media Deutschland GmbH;609—619.
- Saha S, Pal H, 2017. Bootstrapping for numerical open IE [C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver; Association for Computational Linguistics;317—323.
- Sheng J W, Li Q, Hei Y M, *et al.*, 2021. A joint learning framework for the CCKS-2020 financial event extraction task[J]. *Data Intelligence*, 3(3):444—459.
- Xu N, Xie H H, Zhao D Y, 2020. A novel joint framework for multiple Chinese events extraction[C]//Proceedings of the 19th China National Conference on Computational Linguistics, Haikou; Chinese National Conference on Computational Linguistic;174—183.
- Yang H, Chen Y B, Liu K, *et al.*, 2018. 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, Melbourne; Association for Computational Linguistics;50—55.
- Zhang Y H, Qi P, Manning C D, 2018. Graph convolution over pruned dependency trees improves relation extraction[C]//Proceedings of the Empirical Methods in Natural Language Processing, Belgium; Association for Computational Linguistics;2205—2215.

(责任编辑:甘娟娟)



(上接第 499 页)

- 武仁杰,李海波,于崇,等,2018. 基于统计分级判别的爆破块度预测模型[J]. *岩石力学与工程学报*, 37(1):141—147.
- Wu R J, Li H B, Yu C, *et al.*, 2018. Model for blasting fragmentation prediction based on statistical classification [J]. *Chinese Journal of Rock Mechanics and Engineering*, 37(1):141—147. (in Chinese)
- 谢宝香,高文华,易梅辉,等,2022. 红砂岩粗粒料颗粒破碎的分形维数指标及影响因素分析[J]. *应用力学学报*, 39(6):1117—1124.
- Xie B X, Gao W H, Yi M H, *et al.*, 2022. Analysis of fractal dimension index and influencing factors of red sandstone coarse granular material breaking[J]. *Chinese Journal of applied mechanics*, 39(6):1117—1124. (in Chinese)
- 徐日庆,畅帅,李雪刚,等,2013. 基于裂离参量的粒状土颗粒破碎量化方法[J]. *岩土工程学报*, 35(12):2179—2185.
- Xu R Q, Chang S, Li X G, *et al.*, 2013. Quantitative method for crushing of granular soils based on discarding parameters[J]. *Chinese Journal of Geotechnical Engineering*, 35(12):2179—2185. (in Chinese)
- 张晨曦,王洋,魏玉峰,等,2022. 无胶结粗粒土初始结构强度研究[J]. *水文地质工程地质*, 49(2):54—63.
- Zhang C X, Wang Y, Wei Y F, *et al.*, 2022. Research on the critical strength of the initial structure of the uncemented coarse-grained soil [J]. *Hydrogeology & Engineering Geology*, 49(2):54—63. (in Chinese)
- Einav I, 2007. Breakage mechanics-part I: Theory[J]. *Journal of the Mechanics and Physics of Solids*, 55(6):1274—1297.
- Hardin B O, 1985. Crushing of soil particles[J]. *Journal of Geotechnical Engineering*, 111(10):1177—1192.
- Marsal R J, 1967. Large-scale testing of rockfill materials [J]. *Journal of the Soil Mechanics and Foundations Division*, 93(2):27—43.
- Rahman M M, Lo S R, Baki M A L, 2011. Equivalent granular state parameter and undrained behaviour of sand-fines mixtures[J]. *Acta Geotechnica*, 6:183—194.
- Zhao X L, Zhu J G, Wu Y L, *et al.*, 2022. A numerical study on the influence of coordination number on the crushing of rockfill materials[J]. *Rock Mechanics and Rock Engineering*, 55(10):6279—6300.

(责任编辑:魏 崑)