

# 文本立场检测综述

李 洋 孙宇晴 景维鹏

(东北林业大学信息与计算机工程学院 哈尔滨 150040)

(yli@nefu.edu.cn)

## Survey of Text Stance Detection

Li Yang, Sun Yuqing, and Jing Weipeng

(College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040)

**Abstract** Text stance detection is a basic study of text opinion mining, which aims to analyze the stance expressed in the text towards a specific target. Due to the rapid development of the Internet, the discussions of users for public events and consumer products are growing exponentially. The research of text stance detection is of great importance for product marketing and public opinion decision-making. This paper reviews the research of text stance detection from three angles: target type, text granularity and research method. First, from the perspective of target type, this paper focuses on three aspects: single-target stance detection, multi-target stance detection and cross-target stance detection; from the perspective of text granularity, the paper compares different application scenarios and methods of sentence level stance detection, document level stance detection and debate text stance detection; from the perspective of research methods, the paper introduces the traditional machine learning, topic model, deep learning and “two-stage” methods, and points out the advantages and disadvantages of various methods. Then, the evaluation tasks of text stance detection and the open data resources are summarized. Finally, based on the current research, the paper summarizes the application fields and looks forward to the future development trends and challenges of text stance detection.

**Key words** text; opinion mining; stance detection; target; stance

**摘要** 文本立场检测是文本意见挖掘领域的基础性研究,旨在分析文本中对特定目标所表现的立场倾向。随着互联网的飞速发展,用户对于公共事件、消费产品等的讨论文本呈指数级增长,文本立场检测研究对产品营销、舆情决策等具有重要意义。从目标类型、文本粒度以及研究方法3个角度对文本立场检测研究工作展开综述。首先,从目标类型角度,围绕单目标、多目标以及跨目标立场检测3个方面梳理了文本立场检测的不同研究任务;从文本粒度角度,对比了句子级、篇章级以及辩论文本立场检测的不同研究场景和方法;从研究方法角度,介绍了基于传统机器学习、主题模型、深度学习以及“2阶段”的方法,并指出各种方法的可取与不足之处。接着,对文本立场检测评测任务以及公开数据资源进行了归纳。最后,立足当前研究形势,总结了文本立场检测研究的应用领域,展望了未来发展趋势以及面临的挑战。

收稿日期:2020-06-29;修回日期:2021-01-22

基金项目:国家自然科学基金青年科学基金项目(61806049);黑龙江省自然科学基金项目(F2018001);黑龙江省应用技术研究与开发计划重大项目(GA18B301)

This work was supported by the National Natural Science Foundation of China for Young Scientists (61806049), the Natural Science Foundation of Heilongjiang Province (F2018001), and the Major Projects of Applied Technology Research and Development Plan in Heilongjiang Province (GA18B301).

通信作者:景维鹏(jwp@nefu.edu.cn)

关键词 文本;意见挖掘;立场检测;目标;立场

中图法分类号 TP391

随着社交媒体的快速发展,每天有大量的以产品评论、博客、论坛和微博等形式出现的用户生成内容(user-generated content, UGC).此类文本数据中往往蕴含着对产品、服务、政策和新闻等的宝贵反馈,能够为决策者提供来自于公众的、与其关注目标相关的意见数据,为自动化文本意见挖掘带来了前所未有的研究契机.

文本立场是文本作者表露出的对于特定目标的认识或处理问题时所持有的态度,目标主要包括人物、组织、政策、活动或者产品等.例如,为确定消费者对新产品的反应和提升用户体验,营销者可以通过网络发布与其目标产品相关的帖子,收集消费者的立场反馈.为确定民众对政策、法规等的反应和维护社会安全稳定,决策者可以通过收集民众对其的立场反馈来了解民情、引导舆论.

文本立场检测是指从用户发表的文本中自动判断其对于预先给定目标的立场.文本立场检测与文本情感分析是文本意见挖掘领域的重要研究方向,与文本情感分析不同的是,文本立场检测需要判别表达方式更为复杂的“支持、反对或中立”的立场,而不是对指定对象的积极或消极的情感极性.情感通常作为立场表达的载体,同一立场可以通过不同的情感来表达.

文本立场检测主要包括支持、反对或中立3个立场.

1) 支持.作者对目标持支持态度.例如,直接或间接地支持某人/某事,反对或批评与目标相反的某人/某事,或附和他人对目标支持的立场.

2) 反对.作者对目标持反对态度.例如,直接或间接反对或批评某人/某事物,支持某人/某事物反对目标,或附和他人对目标反对的立场.

3) 中立.作者对目标未表现任何立场.

如图1所示,对于用户A,用户B,C,D分别持有直接支持、直接反对以及中立的立场,由于用户E对用户B持有支持立场,用户F对用户C持有支持立场,因此,用户E和用户F对用户A分别持有间接支持立场和间接反对立场,由此可见立场关系具有一定的传递性.

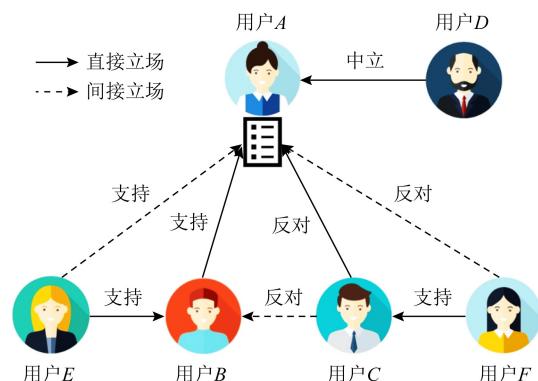


Fig. 1 The transitivity of text stance

图1 文本立场的传递性

文本立场检测早期的研究工作大多集中于英文在线论坛以及政治辩论.2006年Thomas等人<sup>[1]</sup>收集了2005年美国国会发言辩论记录,并进行支持与反对立场的标注,随后,不同角度的文本立场检测研究工作相继出现.近年来,立场检测的研究得到了学术界和工业界的广泛关注,在商业智能、舆情分析、政治选举、医疗决策和心理健康等许多不同领域均有应用.同时,对于诸多研究均可起到辅助作用,其研究成果可以推广到对话系统、个性化推荐、谣言验证、社会感知等研究领域,具有重要的学术研究与应用价值.

本文从3个角度对文本立场检测研究展开综述,

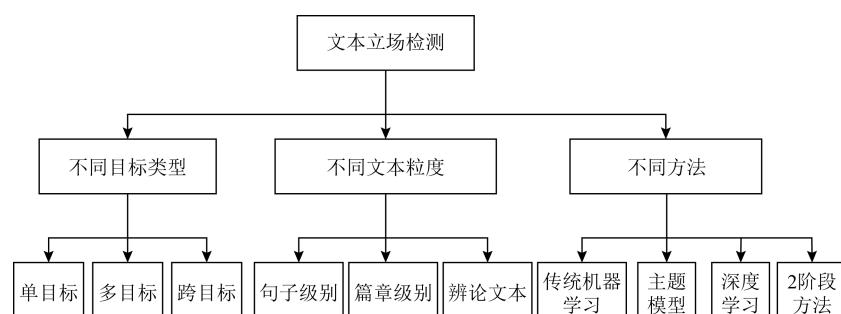


Fig. 2 Classification of text stance detection

图2 文本立场检测分类

组织架构如图 2 所示。文本立场检测的目标类型包括单目标、多目标以及跨目标。文本立场检测的文本粒度涉及到句子级别、篇章级别以及辩论文本。文本立场检测的研究方法涵盖了传统机器学习、主题模型、深度学习以及“2 阶段”方法。通过上述 3 个角度，梳理了不同目标类型、文本粒度立场检测的特点，对比了不同方法的可取与不足之处。接着，对文本立场检测评测任务以及公开数据资源进行了归纳。最后，立足当前研究形势，总结文本立场检测研究的应用，

并展望了文本立场检测研究的未来发展趋势以及面临的挑战。

## 1 文本立场检测的不同目标类型

根据文本立场检测任务中目标类型的不同，可分为单目标立场检测<sup>[2-3]</sup>、多目标立场检测<sup>[4]</sup>以及跨目标立场检测<sup>[5]</sup>，如表 1 所示，接下来将分别对这 3 个方向展开介绍。

Table 1 Stance Detection with Different Target Granularity

表 1 不同目标粒度的立场检测

目标类型	目标	文本	立场
单目标	Open two-child policy	We have a two-child. Let's think of a name for him/her.	支持
多目标	Bernie Sanders & Hillary Clinton	The scum on the left will choose to kill 100 white Americans.	反对和反对
跨目标	Legalization of abortion→feminism	All people, men and women alike, should enjoy equal political, economic and social rights.	支持

### 1.1 单目标立场检测

单目标立场检测是指给定单一的文本（推特、微博、新闻文章、辩论文本等）以及目标，需要确定文本对给定的目标的态度是支持、反对或者中立，即寻找可确定文本对于目标所持立场的映射函数，其定义为：

**定义 1.** 给出 1 组文本  $D$  及 1 个目标  $T$ ，确定映射的值  $S_T : D \rightarrow \{\text{支持}, \text{反对}, \text{中立}\}$ ,  $\forall d \in D$ 。

单目标文本立场检测的研究是文本立场检测中最为广泛的，现有的研究工作主要分为针对人物的立场检测、针对政策法规的立场检测、针对公共事件的立场检测以及针对产品的立场检测。

#### 1.1.1 针对人物目标的立场检测

针对人物的立场检测大多是将政治人物作为目标，而有关政治人物的文本中常常没有直接针对该政治人物的信息。例如，用户针对某候选人未发表任何提及其实内容的文本，但可以通过用户对其支持的某政治事件的反对态度，推断该用户对该候选人的反对立场。因此通过分析用户在推特、微博等社交媒体上对某政治人物的政党、支持者、反对者以及对相关事件的支持或批评性言论，可以间接推断用户对该政治人物的立场。Dias 等人<sup>[6]</sup> 主要关注该角度，将用户使用的与目标或立场有关的常用词语的  $n$ -gram 作为特征训练支持向量机（support vector machine, SVM）分类器进行立场检测。

#### 1.1.2 针对政策法规目标的立场检测

用户在表达对政策法规这类目标的立场时，通常伴随着具体支持或反对的理由。例如：针对“深圳

禁摩限电”这一政策法规，用户发表微博“广州的也给全部禁了吧，特别是摩托车，容易出事！”，从“全部禁了”“出事”这类词语可知用户对该目标的支持立场。尽管用户对目标的表达方式千差万别，但对于同一个目标，支持者（或反对者）所使用的关键词是相近的，因此，如果能利用目标文本捕捉到用户文本中支持（或反对）的关键词，就能够大幅度提升立场检测的效果。Du 等人<sup>[7]</sup> 建立了基于目标主题增强的注意力模型，用注意力机制定位到文本中与目标显著相关的关键部分，在中文立场检测数据集上“开放二胎政策”和“是否允许春节放鞭炮的法规”2 个目标的实验证明了对目标文本和用户文本构建注意力机制的有效性。

#### 1.1.3 针对公共事件目标的立场检测

当一个事件成为舆论热点时，就称为公共事件，如“对全球变暖的担忧”“朝鲜核试验”“俄罗斯在叙利亚的反恐行动”等。对于“俄罗斯在叙利亚的反恐行动”这一目标，某用户发表微博“9月30日开始至今，俄空爆叙利亚，共死亡1331人，其中403人是一般的民众……其中的三分之一是无辜平民陪葬。”文本中表述了对无辜平民死亡的看法，没有谈及对于该事件的态度。针对公共事件目标的立场检测往往需要额外的背景知识来丰富文本的词汇和语义信息。莫雨洁等人<sup>[8]</sup> 针对“俄罗斯在叙利亚反恐行动”目标进行了立场检测，使用维基百科中文语料训练了一个 400 维的 word2vec 模型，得到了语料的字向量、词向量表示，同时包含了词语的语义信息，对目标

相关的词汇与语义信息的扩展,能够更加高效地判定用户的立场。

#### 1.1.4 针对产品目标的立场检测

产品目标的立场检测,大都是基于产品评论文本进行的。用户在评论文本中描述较多的是产品属性,因此针对产品的立场检测的关键是对文本中产品属性关键词和描述词的抽取,此时针对产品评论的立场检测与方面级情感分析任务极为相似。Wang 等人<sup>[9]</sup>提出从产品评论中自动挖掘用户对于产品的立场,帮助客户和商家做出明智的购买决策和有效的营销策略。一般来说,方面级情感分析的目的是提取五元组  $\langle e, a, s, h, t \rangle$ ,其中  $e$  是实体或目标,  $a$  是实体  $e$  的方面(aspect),  $h$  是意见持有人,  $t$  是意见持有人对实体  $e$  表达意见的时间,  $s$  是  $h$  在  $t$  持有实体  $e$  的方面  $a$  的意见。例如,用户发表文本“我今天买了一个新的 iPhoneX,屏幕是很好,但是语音质量很差。”输出的 2 个五元组分别是  $\langle \text{iPhoneX}, \text{screen}, \text{great}, \text{I}, \text{today} \rangle$  和  $\langle \text{iPhoneX}, \text{voice quality}, \text{poor}, \text{I}, \text{today} \rangle$ 。此时,进行文本立场检测,仅需要提取五元组中的针对特定目标  $e$  的意见  $s$ ,用户在评论文本中描述较多的是产品属性,因此捕获  $s$  的关键是对文本中产品属性关键词(如屏幕、电池)和描述词(如可移植性、可用性)的抽取。早期大都采用基于规则的方法,利用词性标注模板、句法依赖模板和关联规则挖掘来提取关键词和描述词。由于基于规则的方法很难从一个领域移植到另一个领域,且往往忽略了属性之间的语义相似性,主题模型、聚类等无监督方法被提出以缓解该问题。Titov 等人<sup>[10]</sup>发现标准主

题模型 LDA (latent Dirichlet allocation), PLSA (probabilistic latent semantic analysis) 更倾向于生成全局属性(如产品类型或品牌),他们提出了多粒度主题模型 LDA(multi-grain topic model, MG-LDA) 区分了全局主题和局部主题。局部主题用于捕获属性描述词,全局主题用于捕获产品的属性词。有效地提取属性词以及属性描述词,利于提升文本立场检测性能。

#### 1.2 多目标立场检测

多目标立场检测与单目标立场检测相似,给定文本和  $K$  个相关的目标,需要确定文本对给定  $K$  个目标的立场,区别在于其中一个目标的预测对其他目标的预测可能有潜在影响,其定义为:

**定义 2.** 给定 1 组文本  $D$  与  $K$  个目标  $T_1, T_2, \dots, T_K$ , 目的是  $\forall d \in D$ , 确定映射  $S: T_1 \times T_2 \times \dots \times T_K, D \rightarrow \{\text{支持, 反对, 中立}\}$ .

Sobhani 等人<sup>[4]</sup>首先注意到了多个目标之间的依赖性,并创建了一个用于多目标立场检测的数据集。该数据集收集了与 2016 年美国大选相关的 4 455 条推文,并对多个目标对的立场进行了人工标注。在文献[11]中,Sobhani 等人提出一个基于注意力机制的多目标立场检测方法(multi-target stance detection, MTSD),使用编码器-解码器框架,如图 3 所示。将推特文本通过双向循环神经网络(bidirectional recurrent neural network, BRNN)映射为编码器端的一个向量,在解码器端使用循环神经网络(recurrent neural network, RNN)生成指向给定目标对的立场标签,该模型利用注意力机制根据目标动态地聚焦于输入

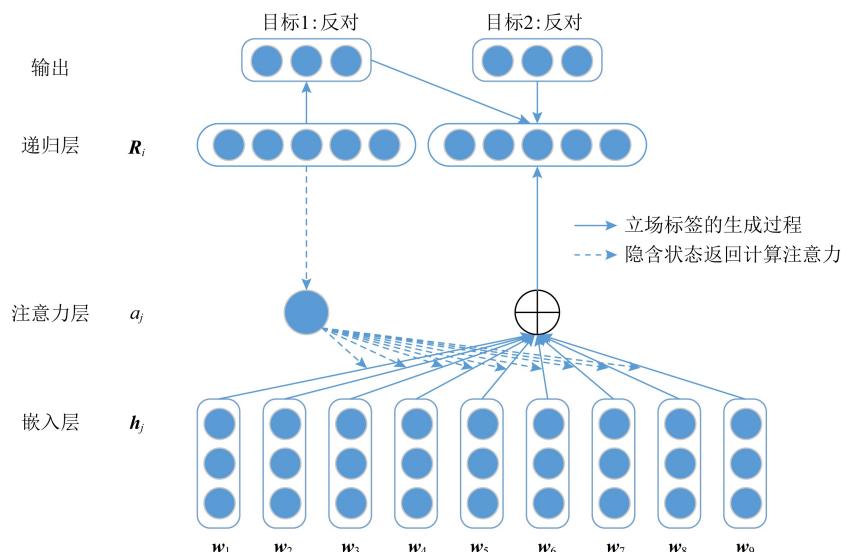


Fig. 3 Multi-target stance detection based on attention mechanism

图 3 基于注意力机制的多目标立场检测

文本中的不同单词,从而针对每个目标生成立场标签.在训练时,分别针对不同的目标顺序,为每个目标对训练 2 个模型(如:针对候选人 A - 候选人 B 目标对进行立场检测,首先训练一个模型预测对候选人 B 的立场,然后预测对候选人 A 的立场,另一个模型则按照与之相反的顺序预测立场标签),此外,在为每个目标生成立场标签时,同时利用前一个目标的立场标签作为输入,因此可以捕获不同目标标签之间的依赖.通过多次实验表明:基于注意力机制的多目标编码器-解码器框架比独立使用机器学习方法或其他联合学习方法(如级联分类和多任务学习)更有效地对多个目标的立场进行联合建模.

为了捕获所有目标的相关信息,一些学者尝试利用记忆网络.Wei 等人<sup>[12]</sup>提出了一种动态记忆增强网络模型(dynamic memory-augmented network, DMAN),该模型首先使用 2 个双向长短时记忆网络模型(bidirectional long short term memory, Bi-LSTM)分别编码推特文本和目标,接着利用注意力机制学习针对不同目标的文本表示.然后,利用共享的动态记忆单元自适应地提取多个目标的相关立场指示信息,该动态记忆网络在训练过程中能够捕获和保存有关全部目标的隐式状态信息,最后通过 softmax 分类分别对  $K$  个目标进行立场检测.

Siddiqua 等人<sup>[13]</sup>提出了一种神经网络集成模

型(proposed neural ensemble model, PNEM),将基于注意力机制的密集连接的双向长短期记忆网络(densely connected Bi-LSTM, DC-Bi-LSTM)<sup>[14]</sup>、嵌套的长短期记忆网络(nested LSTM, NLSTM)<sup>[15]</sup>与多核卷积结合在一个统一的结构中.该模型针对目标对中的每个目标进行立场检测,首先将目标与文本向量拼接起来,通过多核卷积滤波器生成不同的特征向量,再分别输入到带有注意力机制的 DC-Bi-LSTM 和 NLSTM 中,最后将 2 个部分的表示拼接起来以确定最终的立场标签,在实验结果上超越了前面 2 个方法.

### 1.3 跨目标立场检测

跨目标立场检测任务的基本思想是给定领域相关的 2 个目标,从源目标学习一组特定领域的知识,将它们应用于给定目标立场的预测,其定义为:

**定义 3.** 给定一组文本  $D$  与源目标  $P$ 、目的目标  $T$ ,目的是确定映射  $S_T: S_P \cup D \rightarrow \{\text{支持}, \text{反对}, \text{中立}\}$ ,  $\forall d \in D$ .

Xu 等人<sup>[5]</sup>针对该问题提出交叉神经网络模型,该模型由 4 层神经网络组成,将来自源目标的带有立场的句子  $P$  和给定目标  $T$  作为输入,并采用了自注意机制,最后输出预测的立场,模型结构如图 4 所示.该模型能从源目标中发现有用的领域信息,并很好地应用在对给定目标的立场检测上.

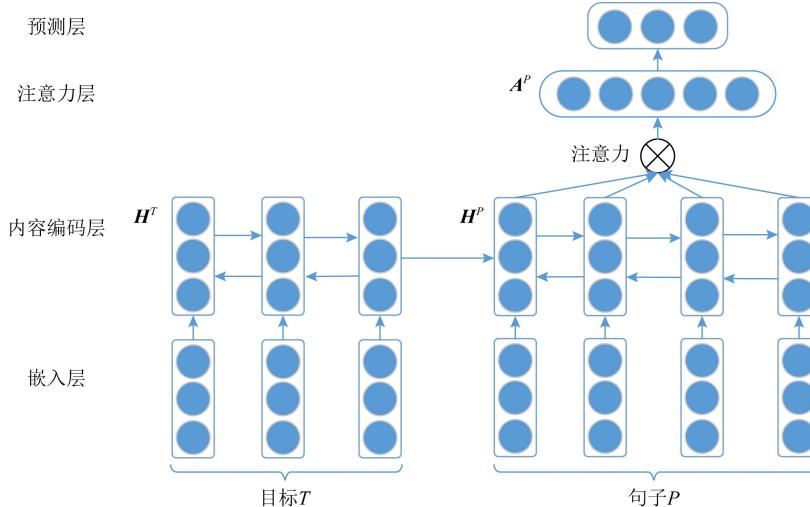


Fig. 4 Cross-target stance detection based on attention mechanism

图 4 基于注意力机制的跨目标立场检测

### 1.4 小结

本节介绍了不同目标类型的文本立场检测工作,总结 3 方面:

#### 1) 单目标立场检测

单目标立场检测中目标与文本是一一对应的关

系,在进行立场检测时通常需要针对目标和文本的特点,因此模型和方法简单且针对性强.

#### 2) 多目标立场检测

多目标立场检测中文本对于每一个目标的立场取决于其他目标对它的影响,因为存在语言模型可

能无法兼容不同目标的问题,通常需考虑联合建模.

### 3) 跨目标立场检测

给定领域相关的 2 个目标,从源目标学习一组特定领域的知识,然后利用学到的知识判断相同领域内其他给定目标的立场.跨目标立场检测多利用迁移学习思想,捕捉相同领域不同目标的相似性.

## 2 基于不同文本粒度的立场检测

根据待检测文本粒度的不同,文本立场检测分为句子级别立场检测、篇章级别立场检测以及辩论文本立场检测.

### 2.1 句子级别立场检测

当立场检测中给定的用户文本是较短的单个句子形式,即为句子级别立场检测.句子级别立场检测的待检测文本常见于短文本,如给定推特或微博文本,判断其对给定目标的立场是支持、反对或中立.

Zarrella 等人<sup>[16]</sup>认为句子级别立场检测的挑战在于用户表达立场观点经常伴有讽刺、反讽、类比和

隐喻等修辞性的语言表达,由于立场文本较短,缺乏背景知识,算法很难准确识别出文本立场.为了解决这些问题,他们采用迁移学习的思想,通过在大规模无标签推特数据集上远程监督训练了单词和短语的初始嵌入向量,将其作为 4 层循环神经网络模型的输入,这种背景知识信息的引入,在句子级别立场检测任务中取得了较好的结果.Wei 等人<sup>[17]</sup>提出了“2 步骤”无监督的方法进行句子级别立场检测,同样使用远程监督通过英文推特立场检测领域语料库(domain corpus)建立了 2 个类别的训练数据集(支持/反对),然后利用改进的 CNN(convolutional neural network)模型<sup>[18]</sup>提出的模型将推特文本分成支持、反对与中立 3 类,最后设计了一个投票方案(vote scheme)进行预测.

### 2.2 篇章级别立场检测

篇章级别立场检测通常是指给定文章标题作为目标,判断文章与该标题的一致性,分为支持、反对、需要讨论或与标题无关 4 种类别,其目标-文本对举例如表 2 所示:

Table 2 Document Level Stance Detection

表 2 篇章级别立场检测

目标(标题)	文本(正文)	立场
Apple watch's battery life is "much better" than expected due to its power saving features	According to a report, the battery life of Apple watch should be much longer than that suggested by some early customers, and the device can fundamentally change the way users interact with their iPhones. TechCrunch cited reports from several people who spent "a long time" on products, saying that Apple watch has been trying to improve its service life. Therefore, after one day's use, the watch should still have about 25% electricity. At the same time, it is noted that if you use the Apple watch, you can immediately perform many tasks and receive notifications without using an iPhone.	支持

篇章级别的立场检测是虚假新闻检测中的一项基础研究工作,Pomerleau 等人<sup>[1]</sup>在 2017 年提出虚假新闻检测任务,并将立场检测作为该任务的第一阶段.Riedel 等人<sup>[19]</sup>提出一个端到端的立场检测模型,将输入文本和目标分别用 2 个简单的词袋(bag of words, BOW)模型来表示,然后使用具有隐藏单元的多层次感知机网络对其进行一致性的分类.Bhatt 等人<sup>[20]</sup>着重从外部特征获取信息,首先利用启发式的特征工程方法获取目标与文本相似词语数、相似 n-gram 数等外部特征,然后采用循环神经网络将所有特征组合起来进行立场检测,取得了比前者词袋特征更好的效果.还有一些学者借助神经网络模型在特征提取方面的优越表现,Conforti 等人<sup>[21]</sup>提出

了 2 种神经网络模型:双向条件编码模型与联合匹配注意力模型(co-matching attention),将文章按照句子的层次顺序进行编码,判断新闻文章对标题的立场,该模型对支持和反对 2 个立场的检测效果非常突出.

### 2.3 辩论文本立场检测

Hasan 等人<sup>[22]</sup>指出辩论文本的立场检测包含的主要文本类型为:国会辩论、公司内部讨论以及在线论坛辩论 3 种情境.辩论文本立场检测任务通常是将给定辩论主题作为目标,判断用户的辩论文本对该目标的立场是支持或反对,如表 3 所示.本文将对辩论文本的立场检测分为基于政治辩论文本的立场检测与基于在线社交网络辩论文本的立场检测.

① <http://www.fakenewschallenge.org/>

Table 3 Debate Text Stance Detection

表 3 辩论文本立场检测

目标	文本	立场
Self defense; is self-defense a good reason to hold a gun?	1) As a means of self-defense, citizens have the “right” to hold guns. 2) Citizens who are not allowed to carry guns are also at a disadvantage in the struggle against lawbreakers and armed elements.	支持
	1) Although people have the right to own their own property, it should not be a justification for wielding lethal guns. 2) Robert F. Drinan, former congressman of the Democratic Party of the United States, <gun control:good is greater than evil>.	反对

### 2.3.1 针对政治辩论文本

Thomas 等人<sup>[1]</sup>首次提出政治辩论文本的立场检测任务,他们收集了 2005 年美国国会发言辩论记录,并进行“支持”与“反对”的标注.Gottipati 等人<sup>[23]</sup>从辩论百科 Debatepedia<sup>①</sup>中获取政治辩论数据,利用用户发表的辩论文本以及所表达的情感信息等构建了一个主题模型,以推断出用户的辩论文本对于各个目标的立场.Somasundaran 等人<sup>[24]</sup>对 2 种类型的特征,即情感特征和辩论(argue)特征进行了实验研究,结果表明辩论特征在立场分类中的结果始终优于情感特征。

刘兵<sup>[25]</sup>提到将针对辩论文本的立场检测看作传统情感分析的一种扩展,他们认为用户表达立场的文本也蕴含着情感信息,例如表达支持的立场可以认为具有褒义的情感信息,反对的立场则具有贬义的情感信息.因此,称支持与反对的文本表达为“支持-反对”(agreement disagreement, AD)情感表达,该概念的提出对于辩论文本立场检测与情感分析具有推动意义.

### 2.3.2 针对在线社交网络辩论文本

随着社交网络的普遍使用,对在线社交网络辩论文本进行立场检测逐步受到人们的关注.将社交网络辩论的主题视作目标,识别用户发布的文本对于该目标的立场,目前此任务大致分为 2 种途径:直接立场检测与间接立场检测.

#### 2.3.2.1 直接立场检测

相对于间接立场检测而言,直接立场检测是直接识别用户的辩论文本对于目标的立场.Hasan 等人<sup>[26]</sup>使用 2 种语言特征:上下文信息和人工标注获得的知识训练辩论文本立场分类器,接着学习用户的立场,以作为辩论文本立场分类器的补充,最后使用整数线性规划(integer linear programming, ILP)方法联合推断立场.Yuan 等人<sup>[27]</sup>对问答形式

的辩论文本展开了研究,首次从百度知道、搜狗问问、明医等网站获取并标注 13 591 条问题-回答对数据,主要涉及怀孕、食品、安全、疾病等话题.采用一种循环条件注意力模型(recurrent conditional attention, RCA)判断用户的辩论文本对目标问题的立场倾向性.首先使用基于门控循环单元(gated recurrent unit, GRU)结构的循环神经网络对问题句和回答句进行建模;接着利用注意力机制来筛选问题句中对于立场比较重要的词,得到与立场相关的问题句的表示向量,回答句中也采用该方法获取向量表示;最终将回答句中所有词的隐含向量加权平均融入到立场状态中,完成 RCA 模型对问答对的一次阅读过程.在循环阅读问题对的过程中,RCA 模型交错提炼问题句、回答句语义表示,挖掘问题句-回答句之间的相互依赖关系,逐步推理获得回答句对于目标问题真实的立场.

#### 2.3.2.2 间接立场检测

大多数意见的表达并不是直接针对目标主题,而是针对相似目标或者其他用户的辩论文本,这时,直接立场检测很难确定用户对某个目标主题的立场,因此需要间接立场检测.Murakami 等人<sup>[28]</sup>在调研前人的工作时,发现在基于链接的方法中使用已标记的文本内容会提高立场检测的效果,因此使用基于规则的分类器将用户相对于相邻辩论文本的立场分为支持、反对或中立,从而确定观点网络中相应链接的权重,根据该权重计算表示 2 个用户之间的分歧程度的反应参数,最后运用最大割(max-cut)算法将用户相对于目标主题的立场分为支持或反对.Durmus 等人<sup>[29]</sup>创建了一个包含 741 个辩论文本主题的数据集,针对每一个目标主题都构建一个论点树,其中根节点代表主题,每个子节点代表论点,可能支持或反对其父节点,如图 5 所示,每个子节点对其父节点的立场都标识在他们之间的边上.针对主

① [http://dbp.idebate.org/en/index.php/Debate:\\_Gun\\_control](http://dbp.idebate.org/en/index.php/Debate:_Gun_control)

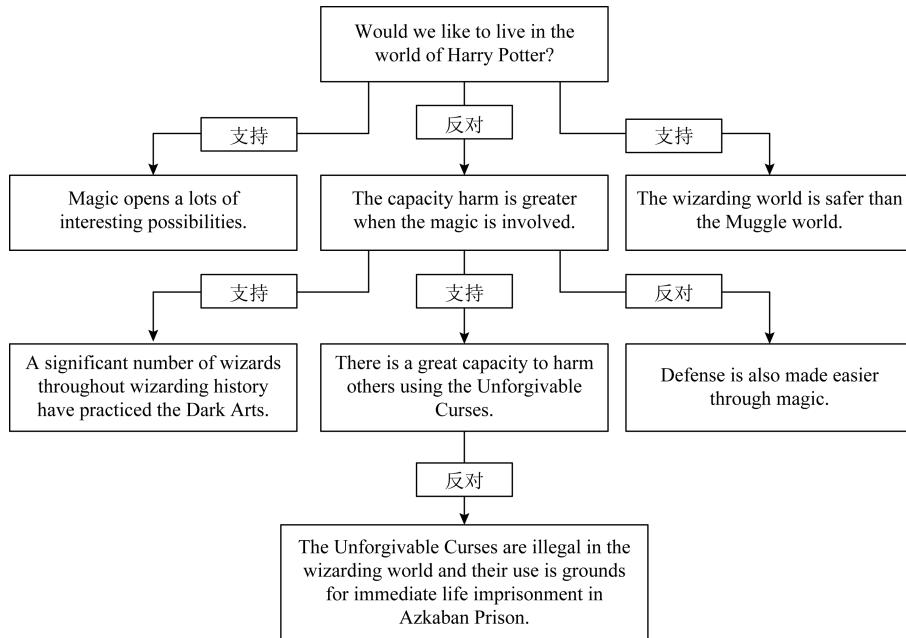


Fig. 5 Debate tree: “would we like to live in the world of harry potter?”

图5 辩论主题树：“我们愿意生活在哈利·波特的世界里吗？”

题“我们想住在哈利波特的世界吗？”，它的子节点“魔术开启了许许多多有趣的可能性”是支持立场，而另一个子节点“当魔法介入时，伤害的能力会更强”则持有反对立场。采用 BERT (bidirectional encoder representation from transformers) 模型对每一个论点相对于主题或者同一棵论点树上其他论点的立场进行检测。还对比了 3 种微调 BERT 模型：朴素的 BERT 模型、带有路径信息的 BERT 模型、带有层次结构和路径信息的 BERT 模型。结果表明，考虑层次结构和路径信息有助于确定 2 个论点之间的相对立场，通过相邻观点间接检测用户对主题的立场与直接立场检测相比能取得更准确的结果，模型结构也相对复杂。

### 3 文本立场检测方法分类

文本立场检测的方法大致可分为基于传统机器学习的方法、基于主题模型的方法、基于深度学习的方法以及基于“2 阶段”的立场检测方法。

#### 3.1 基于传统机器学习的立场检测

在传统机器学习方法中逻辑回归、朴素贝叶斯、决策树、SVM 等是最常用的机器学习方法，也是常常用作与其他方法进行比较的基线方法，其重点是如何选取合适的特征表示。特征类型可分为简单文本特征、情感特征和混合特征。

#### 3.1.1 简单文本特征

简单文本特征主要包括词汇特征、词频-逆文档频率(term frequency-inverse document frequency, TF-IDF)、单词向量(word2vec, para2vec)、主题模型相关特征(LDA, LSA)、命名实体、依赖关系、句法规则等。Xu 等人<sup>[30]</sup>采用 para2vec, LDA, LSA 等语义特征来表示推特文本中的语义信息，对比了随机森林(random-forest, RF)、基于线性核函数的支持向量机(support vector machine-linear, SVM-Linear)以及基于 RBF 核函数的支持向量机(support vector machine-RBF, SVM-RBF)等机器学习算法在使用不同语义特征时立场检测的效果。类似地，Sun 等人<sup>[31]</sup>分析了语义特征、词法特征、形态特征和句法特征多种语言特征在中文微博立场检测任务中的作用，实验表明上述特征均有助于立场检测性能提升。

#### 3.1.2 情感特征

为了证明情感信息与给定目标之间存在相互依赖关系，Mohammad 等人<sup>[32]</sup>首先发布了一个融合情感信息和立场标签的标注数据集，用于研究情感和立场之间的相互影响关系。基于该数据集，Ebrahimi 等人<sup>[33]</sup>将情感极性融入到目标和立场中，通过一个图模型对目标、立场标签和文本中情感词的交互进行建模，建立了对数线性的情感-目标-立场联合模型(joints sentiment-target-stance modeling,

STS), 证明了联合情感信息对目标进行立场建模的有效性。此外, 情感词典也是将情感特征引入立场检测的重要手段<sup>[34]</sup>。

### 3.1.3 混合特征

除了简单文本特征、情感特征之外,一些学者提出了基于混合特征的方法。莫雨洁等人<sup>[8]</sup>利用融入同义词词典的词袋模型和字向量、词向量等特征, 分别使用 SVM、RF 和决策树等进行立场检测, 最后对所有的分类器进行后期融合, 取得了中文微博立场检测任务(NLPCC2016 Task4)最好的效果。

## 3.2 基于主题模型的立场检测

Ahmed 等人<sup>[35]</sup>提出一个多视角的潜在的狄利克雷分配模型(multi-view latent Dirichlet allocation, mview-LDA), 将文本中的每一个单词看成是立场倾向和主题的相互结合。使用因子主题模型在主题级别解决了建模立场的问题, 并通过折叠的吉布斯采样进行后验推断:首先进行学习, 找到模型参数的点估计, 再根据文本和已获得的点估计对进行推断。对模型进行了评估和说明, 明显优于基线模型。Gottipati 等人<sup>[23]</sup>提出一种隐含主题模型, 将信息提取和情感分析整合到该模型中, 并且使用辩论百科数据集中论点的层次结构来构建主题, 以推断出社会政治辩论领域中的主题和立场, 最后通过定性和定量评估展示了该模型的优点和缺点。

## 3.3 基于深度学习的立场检测

近年来, 深度学习在自然语言处理方向取得了卓越的成果, 相对于传统机器学习方法, 深度学习有效地减少了手工构造特征的工作量, 并且拥有优秀的表示学习能力。

### 3.3.1 卷积神经网络

卷积神经网络是一类基于空间上卷积操作的神经网络模型,CNN 在文本情感分类、虚假新闻检测等许多文本分类任务中都取得了很好的效果, 因此在立场检测工作中被广泛使用。Vijayaraghavan 等人<sup>[36]</sup>为英文推特立场检测数据集中的 5 个目标分别训练了字符级别和词级别卷积神经网络模型, 并选择表现最佳的模型或是两者的组合进行预测。结果表明在数据量足够大的立场检测中, 字符级别卷积神经网络的表现远远好于词级别卷积神经网络模型。

此外, Taulé 等人<sup>[37]</sup>借助了 CNN 在图像处理等方面的优势, 在针对加泰罗尼亚公投的立场检测中对推特文本、链接的上下文信息以及图片多模态信息进行建模, 认为使用加泰罗尼亚国旗代表着支持加

泰罗尼亞的独立, 进一步提高了立场检测的效果。

### 3.3.2 递归神经网络

一些学者认为, 普通的卷积神经网络不能有效地捕捉语法和句法信息, Zarrella 等人<sup>[16]</sup>通过在 2 个较大的无标签文本数据集上进行弱监督学习获得词向量特征, 然后采用递归神经网络(recursive neural network, RNN)融合语法和句法信息学习输入文本的向量表示, 对推特文本进行立场检测。递归神经网络的优势在于其可以充分利用语法和句法结构的信息, 但其检测效果对语法句法分析的结果也具有较强的依赖。

### 3.3.3 长短时记忆网络

长短时记忆网络(long short-term memory, LSTM)是由门(gate)控制的循环神经网络, 可以捕获较长距离的依赖关系, 缓解 RNN 中梯度消失的问题。Augenstein 等人<sup>[38]</sup>建立了基于弱监督学习的双向条件 LSTM 编码模型, 分别使用 2 个 LSTM 先后建模, 将立场目标与推特文本结合起来: 第 1 个 LSTM 使用额外的训练数据对预定目标进行编码, 第 2 个 LSTM 使用预定目标的编码作为初始状态对推特进行编码, 完成对目标和文本的联合建模。Sun 等人<sup>[39]</sup>提出使用一个共享的 LSTM 层来学习立场和情感信息之间的深层共享表示, 通过参数共享实现相互的语义表示; 然后将共享的表示信息进行叠加, 将情感检测的隐含层输出作为立场检测的附加输入进行联合学习, 利用情感信息来促进立场检测的效果。

注意力(attention)机制最早是在视觉图像领域提出来的, 源于人类视觉的注意力机制。近年来, 注意力机制被广泛应用到基于深度学习的自然语言处理各个任务中, 常与长短时记忆网络结合使用。Du 等人<sup>[2]</sup>为了捕捉目标信息对文本表示的影响, 引入了注意力的机制。他们提出了一种基于目标主题增强的注意力模型, 定位到文本中与目标显著相关的关键部分。基于 Du 的模型, 岳天驰等人<sup>[40]</sup>提出基于“2 阶段”注意力机制的模型, 完成“中国政府在新疆反恐行动”目标的立场检测, 构建的“2 阶段”注意力机制模型 ATA (attention-target-attention) 结构如图 6 所示。图 6 中第 1 阶段将目标与文本向量匹配得到注意力权重  $\beta$ , 进而得到目标向量表示; 第 2 阶段将目标向量与文本词向量匹配得到注意力权重  $\alpha$ , 将该权重与双向 LSTM 文本表示进行交互得到融入话题的文本表示, 最后将该文本表示用于立场检测, 由此提高立场检测结果。

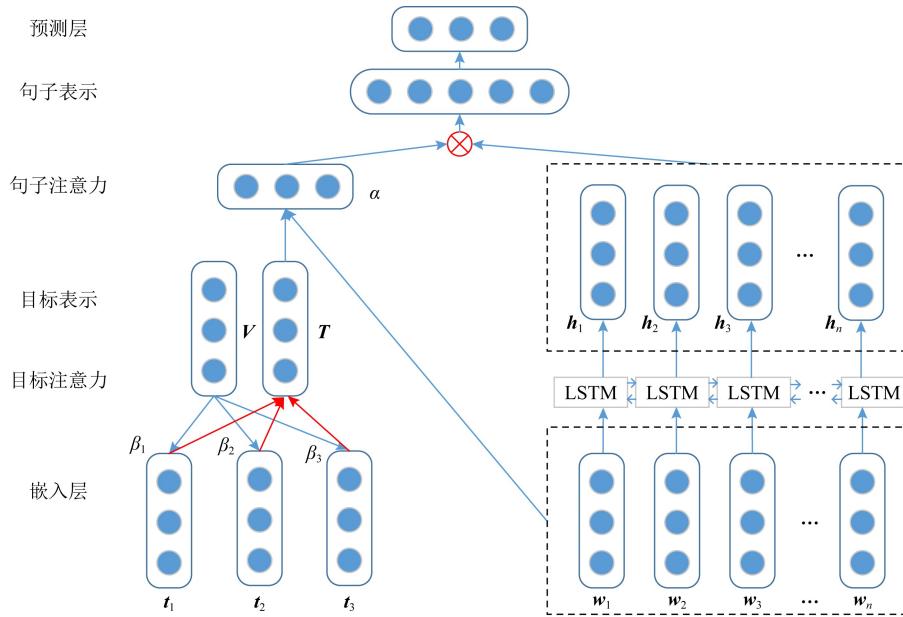


Fig. 6 Stance detection model based on the combination of attention mechanism and LSTM

图 6 注意力机制与长短时记忆网络结合的立场检测模型

颜瑶<sup>[41]</sup>以条件编码的方式结合主题目标信息与文本信息,条件编码在“编码”阶段引入主题目标信息,而注意力机制在“解码”阶段引入主题目标信息,对目标编码时采用单向LSTM,而对文本编码时使用双向LSTM模型.在英文推特立场检测数据集和中文微博立场检测数据集上的实验结果证明了结合注意力机制与条件编码的方法在文本立场检测任务中的有效性.Sun等人<sup>[42]</sup>指出应权衡不同语言信息的重要性,提出一种多层次注意力网络(hierarchical attention network, HAN)模型,模型包括2个部分:语言注意力部分和超注意力(hyper attention)部分,前者用于学习具有不同语言特征(包括情感词、依赖性以及目标)的表示,后者调整不同特征集的权重,经过实验证明,该模型在英文推特立场检测数据集取得了很好的效果.

### 3.3.4 卷积神经网络与长短时记忆网络结合

Yu等人<sup>[43]</sup>使用了卷积神经网络与LSTM类神经网络结合的方法,首先利用CNN层来提取局部n-gram特征,然后利用双向LSTM来提取隐含的全局语义特征,最后将表示这2个特征的固定维度向量结合起来进行结果预测.受Yu等人<sup>[43]</sup>工作的启发,白静等人<sup>[44]</sup>将词向量分别输入到CNN和BiLSTM中,获得卷积特征和隐含语义表示,然后将基于注意力的池化策略应用到模型中,通过BiLSTM模型获得的隐含文本表示指导卷积特征注意力加权,最后将CNN和BiLSTM得到的文本表示进行

拼接作为最终文本表示,基于此预测立场标签.

### 3.3.5 记忆网络

记忆网络(memory network)是为了弥补循环神经网络等无法存储长时间记忆而提出的.它具有独立存储器,能够按需读写“知识库”来增强模型的记忆能力<sup>[45]</sup>.Mohtarami等人<sup>[46]</sup>使用一种端到端的记忆网络模型,包含卷积神经网络和循环神经网络以及相似矩阵,确定文本相对于给定目标的立场,并提取相关证据片段以支持预测结果.特别注意的是,与以往研究工作不同,这里提出的端对端的记忆网络是一个六元组{M,I,F,G,O,R},并引入一个相似矩阵,用来改进证据片段的提取效果,证明了模型能够从给定的文本文档中提取有意义的片段从而提高预测效果.

### 3.3.6 预训练语言模型

近两年ELMo<sup>[47]</sup>,OpenAI GPT<sup>[48]</sup>,BERT<sup>[49]</sup>等预训练语言模型在多项自然语言处理任务中获得了提升,很多学者在进行文本立场检测工作时也采用了预训练语言模型的方法.首先在一个较大的数据集上使用语言模型进行预训练,然后针对特定的任务对它们进行微调.Rao<sup>[50]</sup>讨论并比较了2个方法:ULMFiT(universal language model fine-tuning)<sup>[51]</sup>和GPT(generative pre-training),并展示如何对它们进行微调,以便完成对英文推特的立场检测.ULMFiT方法首先引入通用领域语料库来训练语言模型(父模型),再利用父模型优化立场分类模型

(子模型),最后使用目标任务数据微调分类器并训练立场分类器,得到比前两者更好的预测效果.GPT方法分2个步骤进行:首先是无监督的预训练,语言模型以无监督的方式对来自谷歌图书库的几千本书进行了训练;其次是有监督的微调,为有监督的立场检测任务调整参数,输入通过预先训练的模型最终获得预测的结果,优于ULMFiT方法.

### 3.4 基于“2阶段”的立场检测

3.1~3.3 节介绍的方法都是直接进行立场检测

的,还有另外一类“2阶段”立场检测方法,即以流水线的形式分2个阶段完成立场检测任务。现有的方法中,有2种分段方式:1)首先判断文本是否具有立场,随后判断立场是支持或反对;2)将问题建模为相关性检测与立场检测,最终目标都是判断用户文本对于所给目标的立场。基于“2阶段”的立场检测,将三分类问题转化为2个二分类问题,易于建立模型,实现简单且判断精度较高,基于“2阶段”的立场检测分段方式如图7所示:

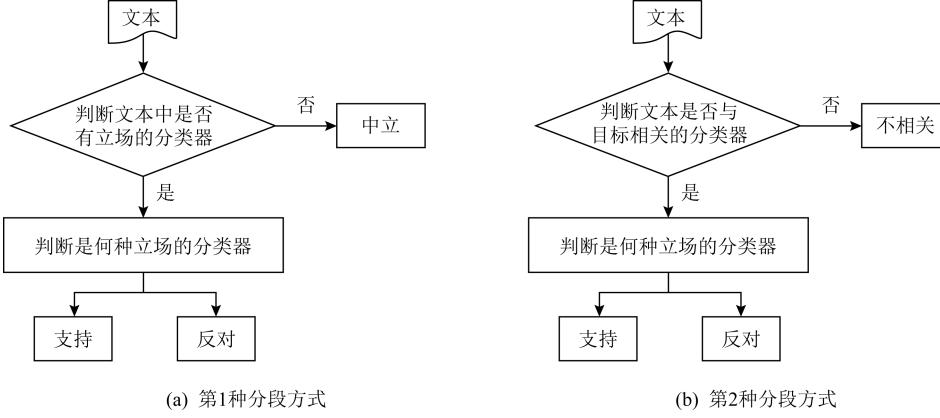


Fig. 7 Flow chart of two-stage stance detection model

图 7 “2 阶段”立场检测模型流程图

第 1 种“2 阶段”方法将常见的文本立场检测中“支持、反对、不相关”的三分类问题转化成 2 个二分类问题：在第 1 个阶段，将数据集中的文本数据分为中立（不带有立场）和非中立（带有立场）这 2 个类别，再对第 1 个阶段中非中立（带有立场）类别的数据进行分类，分为支持和反对 2 种立场，如图 7(a) 所示。Wojatzki 等人<sup>[52]</sup> 分别在有监督和弱监督的情况下对推特文本立场进行预测。在有监督学习时，使用了  $n$ -gram 特征、句法特征、立场词典特征等多种丰富的特征，2 个阶段使用了大致相同的特征集，但是由于 2 个阶段目标不同，针对每个阶段的目标训练单独的 SVM 分类器。在弱监督学习时，采用了迁移学习的思想，首先从有监督任务中选择与推特文本最相似的目标。如果目标被选定，使用其被训练的分类器来分类该推特文本，实现的效果很好。后续的消融实验证明对于所有目标来说，最有效的特征是立场词典特征。基于前人的工作，Dey 等人<sup>[53]</sup> 提出在 2 阶段立场检测方法中融合新的特征。在第 1 阶段，受到主观性检测工作的启发，采用加权 MPQA (Multiple-Perspective QA) 主观-极性分类和基于词网 (WordNet) 的潜在形容词 2 个特征，将整个数据集的文本分为中立的和非中立的 2 个类别；第 2 阶

段结合基于情感词典(SentiWordNet)和MPQA的情感分类特征、框架语义特征、目标检测特征、单词 $n$ -grams和字符 $n$ -grams等多个特征,对第1阶段分类预测出的非中立文本进行分类,分为支持和反对2个类别。在中文微博立场检测中,刘勘等人<sup>[54]</sup>建立了一个结合情感加权算法和朴素贝叶斯算法的组合分类模型(serial sentiment weighted and naïve Bayes model, SWNB)。在第1阶段,为每一个目标扩充了情感词库并建立了关联实体集,提出了能够同时对复杂句式、目标相关实体进行处理的情感加权规则,计算加权分数后将微博文本分为带有立场和不带有立场2个类别。第2阶段,利用朴素贝叶斯分类器,从各个目标的关联实体库提取出每条微博文本中的情感词、关联实体、否定词、程度副词,并将它们作为特征计算这些特征项和各类别的联合概率,从而预测给定微博文本的具体立场即支持或反对。

第2种“2阶段”方法首先实现文本和目标之间的相关性检测，判断文本与目标之间是否相关；接着再完成立场检测，即判断用户文本对目标的立场是支持或者反对，如图7(b)所示。Zhang等人<sup>[55]</sup>在有监督的立场检测中，根据5个预定的目标将训练数据分为5个子集，为每个子集训练2个分类器，对所

使用的包括:语言特征、主题特征、相似性特征、情感词特征、推特文本特征(标签、表情符等)以及词向量。对这 6 种特征进行了对比分析,其中相似性特征可以有效地判断推文是否与给定目标相关,为立场检测的方法提供了新思路。

### 3.5 小结

综合 3.1~3.3 节所述,文本立场检测从方法上可分为基于传统机器学习的立场检测、基于主题模型的立场检测、基于深度学习的立场检测、基于“2 阶段”的立场检测,现在进行总结与对比。

#### 1) 基于传统机器学习的立场检测

基于机器学习的立场分析方法需要结合语言学规则人工构建特征组合,对语言学知识具有较高的要求,也会造成大量的人力消耗。通过构造特征训练机器学习模型来完成文本立场检测任务能取得较好的效果,最常用的分类模型是支持向量机和朴素贝叶斯分类器。相比深度学习的分类方法,该方法模型构成比较简单,定义明确,但存在着特征稀疏、维度灾难等缺陷。并且,特征的提取与选择对分类效果影响较大,因此,探索更复杂、更有意义的分类特征是这类方法努力的方向。

#### 2) 基于主题模型的立场检测

主题模型因其在挖掘文本背后隐含信息方面具有优势,已成功地应用于特征识别等意见挖掘任务,同时这种方法具有一定的理论基础。该模型可以很好地将目标与文本中重要的关键词联系起来,能够解决诸如文本蕴含等其他方法难以判断的问题。但主题模型中文本特征也是稀疏和离散的,需要较多的

训练语料,且目标的数量与文本的对齐处理不够灵活,例如,不能很好地解决对多目标立场检测的工作。

#### 3) 基于深度学习的立场检测

基于深度学习的立场检测方法随机初始化或使用谷歌新闻、维基百科等语料预训练词嵌入,或者利用预训练语言模型学习在大规模无监督语料所蕴含的背景知识。基于深度学习的立场检测方法最常见的就是 LSTM 和记忆网络的方法,得益于其分析长期内容、过渡信息和检测上下文的能力。与基于传统机器学习的立场检测以及基于主题模型的立场检测不同,深度学习有效地减少了手工构造特征的工作量,能够自动抽取最优的特征表示,且连续化的向量表示可以克服特征稀疏的问题。但构造复杂的模型来进行更深度的表示学习,模型体量大,训练时间长,利用知识蒸馏等手段在保证模型泛化能力的前提下,对模型进行简化是这类方法值得深入研究的问题。与此同时,应用到立场检测任务中,探索所持立场的原因,增强模型的可解释性也是进一步值得思考的问题。

#### 4) 基于“2 阶段”的立场检测

与前 3 种立场检测方法不同,基于“2 阶段”的立场检测是根据文本立场检测的不同阶段对其进行分类,将三分类问题转化为 2 个二分类问题,多增加的一个阶段使分类任务更加细化,但是也容易产生错误级联。

最后,对文本立场检测研究中应用的方法进行总结,如表 4 所示。按照使用的机器学习方法不同,对检测对象粒度、特征选择以及是否双语 3 个角度进行了总结。

Table 4 Comparison of Existing Methods of Stance Detection

表 4 立场检测现有方法的比较

立场检测方法	具体方法	文本粒度	特征	是否双语
机器学习	SVM	句子级别	Unigram, TFIDF, Synonym, Word Vector, Character Vector <sup>[8]</sup>	否
			Para2vec, LDA, LSA <sup>[30]</sup>	否
		TFIDF, Sentiment Dictionary <sup>[34]</sup>		否
	RF	Weighted MPQA Subjectivity-Polarity Classification, WordNet Based Potential Adjective Recognition SentiWordNet and MPQA Based Sentiment Classification <sup>[53]</sup>		否
		Unigram, Bigrams, LIWC <sup>[24]</sup>		否
		Context <sup>[26]</sup>		否
	GBDT	句子级别	Unigram, TFIDF, Synonym, Word Vector, Character Vector <sup>[8]</sup>	否
			Para2vec, LDA, LSA <sup>[30]</sup>	否
		TFIDF, Sentiment Dictionary <sup>[24]</sup>		否
	MLP	篇章级别	Unigram, TFIDF, Synonym, Word Vector, Character Vector <sup>[8]</sup>	否
			BOW <sup>[19]</sup>	否

续表 4

立场检测方法	具体方法	文本粒度	特征	是否双语
主题模型	Factorization Model	辩论文本	Stance Dictionary <sup>[35]</sup>	否
	RNN	句子级别	Dependency Feature <sup>[11]</sup> Sentiment Dictionary, Dependency Feature, Target Feature <sup>[39]</sup>	否
	CNN	篇章级别	Hierarchical Order <sup>[21]</sup>	否
深度学习	RNN&CNN	句子级别	National-Flag Pictures <sup>[37]</sup> Word Vector <sup>[44]</sup>	是
	NBM	句子级别	Local $n$ -gram, Global Syntactic Features <sup>[43]</sup> Sentiment Dictionary, Associated Entity Set <sup>[54]</sup>	否
“2 阶段”方法	LR	句子级别	Traditional Linguistic, Topic Feature, Similarity Feature, Sentiment Lexicon, Tweet Specific Feature, Word Vector <sup>[55]</sup>	否

## 4 文本立场检测的评测会议与数据资源

### 4.1 立场检测的评测会议

2016 年以来,国内外的很多研究机构组织了文本立场检测公共评测任务,其中包括国际语义评测大会(International Workshop on Semantic Evaluation,

SemEval)和自然语言处理与中文计算大会(International Conference on Natural Language Processing and Chinese Computing, NLPCC)等,推动了立场检测研究的发展。

表 5 对文本立场检测 4 个评测任务按照立场检测对象粒度、语料种类以及效果最佳的团队进行总结,以供参考。

Table 5 Summary of Open Evaluations for Stance Detection

表 5 立场检测公开评测会议总结

评测任务	文本粒度	语言	得分最高的队伍
SemEval 2016 Task 6	句子级别	英	A: MTTRE <sup>[16]</sup> ; B: Pkudblab <sup>[17]</sup>
NLPCC-ICCPOL 2016 Task 4	句子级别	汉	A: RUC_MMC <sup>[8]</sup> ; B: March
FNC-I	篇章级别	英	Cisco Talos <sup>①</sup>
IberEval-2017/IberEval-2018	句子级别	西班牙语/加泰罗尼亚语	Spanish: uc3m <sup>[37]</sup>

#### 4.1.1 SemEval 2016 Task 6

该评测使用由 Mohammad 等人<sup>[2]</sup>基于英文推特构建的英文立场检测数据集,Task 6 任务包含 2 项子任务:子任务 A 是一个有监督分类任务,使用带有标注的训练数据检测用户对 5 个给定的目标的立场;子任务 B 是一个弱监督分类任务,使用大量未标注的训练数据检测推特文本对于“唐纳德·特朗普”目标的立场。

#### 4.1.2 NLPCC-ICCPOL 2016 Task 4

2016 年自然语言处理与中文计算大会提出了基于中文微博文本(来自新浪微博应用)的评测任务——NLPCC-ICCPOL (International Conference on Computer Processing of Oriental Languages)

2016 Task 4.该任务也是分成 2 项子任务:子任务 A 是有监督分类任务,子任务 B 是弱监督的分类任务。

#### 4.1.3 FNC-I

事实上,2017 年 Dean 等人<sup>[56]</sup>提出将文本立场检测作为虚假新闻检测任务的第一个子任务(fake-news-challenge I, FNC-I).该任务使用的数据是从文献<sup>[57]</sup>发布的数据集中提取的,给定文章标题,判断文章对该标题的立场,分为支持、反对、需要讨论或与标题无关 4 种类别,属于篇章级别立场检测。

#### 4.1.4 IberEval-2017/IberEval-2018

IberEval-2017 (evaluation of human language technologies for Iberian languages)<sup>[58]</sup> 评测提供针对“加泰罗尼亚独立”的推特文本检测,包含 4 319

① <https://blog.talosintelligence.com/2017/06/talos-fake-news-challenge.html>

条西班牙语推文和 4 319 条加泰罗尼亚语推文,参与者判断用户对于“加泰罗尼亚的独立”这一目标的立场是支持、反对或中立。IberEval-2018 提出了 MultiStanceCat(multimodal stance detection in tweets on Catalan #1Oct referendum)任务<sup>①</sup>,是对 IberEval-2017“加泰罗尼亚独立”的立场检测任务的延伸。该任务提出了在双语立场检测任务中结合多模态信息,如推文文本中链接信息、图像和上下文信息等,

旨在从多模态角度检测推特用户对加泰罗尼亚独立公投事件的立场。

## 4.2 文本立场检测的公开数据资源

目前,文本立场检测公开数据资源除了 4.1 节评测公布的数据集之外,还有一些是由学者们自行构建并公开的,本节将对文本立场检测的公开数据资源进行梳理与综述,数据集详细信息及获取方式如表 6 所示。希望对有兴趣的研究者提供启发。

Table 6 Details of Public Data Resources

表 6 公开数据集详细说明

数据集	目标类型	目标数量	文本粒度	数据数量	数据集链接
英文推特立场 检测数据集	单目标	A:5	句子级别	训练集:2914	
				测试集:1249	<a href="http://altqcri.org/semeval2016/task6/">http://altqcri.org/ semeval2016/task6/</a>
		B:1		测试集:707	
英文推特多目标 立场检测数据集	多目标	3 对	句子级别	训练集:3119	<a href="https://engineering.uottawa.ca/school-EECS">https://engineering.uottawa.ca/ school-EECS</a>
				测试集:890	
中文微博立场 检测数据集	单目标	A:5	句子级别	训练集:3 000	<a href="http://tcc.ccf.org.cn/conference/2016/pages/page05_evadata.html">http://tcc.ccf.org.cn/ conference/2016/ pages/page05_evadata.html</a>
				测试集:1 000	
		B:2		测试集:400	
MultiStanceCat 数据集	单目标	1	句子级别	加泰罗尼亚语:5 853	<a href="https://www.autoritas.net/MultiStanceCat-IberEval2018/corpus/">https://www.autoritas.net/ MultiStanceCat-IberEval2018/ corpus/</a>
				西班牙语:5 545	
FNC-1 数据集	单目标	300	篇章级别	训练集:49 972	<a href="https://github.com/FakeNewsChallenge/fnc-1">https://github.com/ FakeNewsChallenge/fnc-1</a>
				测试集:25 147	
土耳其语推特 立场检测数据集	单目标	2	句子级别	700	<a href="http://ceng.metu.edu.tr/">http://ceng.metu.edu.tr/</a>

在 4.1 节所述的 4 个评测任务中,Mohammad 等人<sup>[2]</sup>使用了英文推特单目标立场检测数据集,Xu 等人<sup>[3]</sup>创建了中文微博单目标立场检测数据集,Dean 等人<sup>[56]</sup>构建了虚假新闻立场检测数据集,Taulé 等人<sup>[58]</sup>则是分别注释加泰罗尼亚语和西班牙语推特的数据集。

除了以上 4 个国际/国内评测所提供的语料,也有其他学者提供了用于特殊情景下的语料,Sobhani 等人<sup>[4]</sup>认为立场检测中 2 个或多个目标之间的依赖性不应被忽略,创建了英文推特多目标立场检测数据集,Vasiliki 等人<sup>[59]</sup>收集了一个关于英国脱欧事件的 2016 年英国公投博客文章数据集,用于描述用户文本表达的立场的可能性。

此外,还有学者进行了小语种的立场检测数据集构建工作,如土耳其语<sup>[60]</sup>、阿拉伯语<sup>[61]</sup>、捷克

语<sup>[62]</sup>、英语-印地语<sup>[63]</sup>、意大利语<sup>[64]</sup>、日语<sup>[28]</sup>、俄语<sup>[65]</sup>,这些数据集为语料资源稀少的语言开展立场检测工作提供了很好的施展平台。

目前,SemEval 2016 英文推特立场数据集是当前针对特定目标立场检测使用较为广泛的基准数据集,该数据集构建时间较早。从数据集包含目标的类别来看,大多数现有数据集中在社会、政治和宗教问题上。因此,构建涵盖其他领域的新的数据集对于立场检测任务具有推动性作用。

## 5 文本立场检测的应用

作为意见挖掘的重要研究内容之一,立场检测的相关研究近年来得到学术界和工业界的广泛关注,成为自然语言处理领域的一个研究热点。文本立场检测研究的兴起与发展,不仅能够在商业智能领域

① <https://www.autoritas.net/MultiStanceCat-IberEval2018/>

帮助用户、消费者、生产者了解更多信息,起到建议指导的作用;也是政府了解民生、体察民情的重要手段和渠道<sup>[66]</sup>.同时,文本立场检测对于其他研究领域也具有较高的学术和应用价值,如推荐系统、对话系统、谣言验证以及社会感知等.

### 5.1 推荐系统

推荐系统<sup>[67]</sup>是目前解决互联网用户信息过载问题的重要手段.“推荐”是系统为用户提供的一种主动的信息推送方式,需要进行用户兴趣的推断.持有相同立场的用户倾向于具有相似的兴趣,基于这一发现,Qiu 等人<sup>[68]</sup>针对 CreateDebate<sup>①</sup>平台上辩论话题推荐问题,提出了隐含因子模型,对用户评论的立场、属性以及交互行为同时建模.

其次,对用户评论进行立场检测,可以分析用户对某产品或品牌的态度,预测产品或品牌的销量,辅助营销方进行决策.Amiri 等人<sup>[69-70]</sup>收集了 3 个电信品牌 Verizon, T-Mobile, AT&T 相关的推特,通过立场检测发现品牌客户流失(churn).

### 5.2 对话系统

自然语言对话是人工智能领域最具挑战性的任务之一,其中包含语言理解、推理和对现有知识的利用,近年来已取得很大进展.从理论角度,一些研究者试图在对话系统中探索立场表达的基本理论及其语言学特点,在对话回复中融入立场信息,建立立场倾向和情感极性驱动的自动回复生成,完善回复生成领域的理论体系<sup>[71]</sup>.

从应用角度,将立场倾向融入对话系统使得回复文本的语言方式拟人化,覆盖更加细分的人群和更深入的场景.目前,一些典型对话系统已经初步具备情感陪护、语音助手、商业客服等功能,如 Siri (speech interpretation & recognition interface)、小度、微软小冰等.

### 5.3 谣言验证

互联网具有开放性、虚拟性、隐蔽性、发散性、渗透性和随意性等特点,逐渐成为舆情话题产生和谣言传播的主要场所,谣言的验证与清理成为亟需解决的问题.通过不同来源(例如新闻网站、社交媒体等)检索相关的文本,确定每个文本相对于目标的立场,最后汇总立场的强度,可以对谣言的真实性进行验证.因此,谣言验证是立场检测的一个重要的应用方向.

一方面,立场检测被列为谣言验证的关键步骤.

Zubiaga 等人<sup>[72]</sup>对谣言引起的交流对话进行研究,发现谣言中的对话形式是树形结构,由对谣言持不同立场的回复组成,挖掘用户之间辩论立场特征进行谣言验证,有助于提升谣言验证的准确率.此外,Kumar 等人<sup>[73]</sup>利用立场信息对推特对话中的谣言真实性进行验证,也得到了类似结论.

另一方面,随着从传统新闻媒体向互联网、社交媒体的转变,“回音室效应”使得用户只能看到自己感兴趣和认同的内容,难以获得不同的意见,导致最终的决策基于片面或者存在偏见的信息,甚至可能成为假新闻和虚假信息的牺牲品.Orbach 等人<sup>[74]</sup>提出了反对立场发言的检测任务,给定输入文本和语料库,从该语料库中检索一个包含与输入文本中提出的论点立场相驳斥的反文本.

此外,谣言验证希望通过提供文本中的相关证据来解释事实的真假.但是,现有的语料库并不能很好地统一检索文本、立场检测、谣言验证的工作,Baly 等人<sup>[75]</sup>提出一个阿拉伯语语料库来支持这些工作之间的相互依赖关系,通过实验进一步证明,统一的标注对于立场检测与谣言验证的效果均有提升.

### 5.4 社会感知

立场检测可以作为一种社会感知技术来测量与社会、政治和公共健康等人话题相关的公众态度.

1) 在社会方面.Rebollo 等人<sup>[76]</sup>对使用不同语言的推特用户的讨论文本进行立场检测,从而分析他们针对难民问题的态度;Bartlett 等人<sup>[77]</sup>对英国移民问题的公众意见进行了研究,发现大约 5% 的用户反对移民,23% 的用户支持移民;立场检测还被用来分析民众对于破坏性事件相关的态度,Darwish 等人<sup>[78]</sup>, Demszky 等人<sup>[79]</sup>的工作研究了 2015 年巴黎恐怖袭击后公众对穆斯林的态度.

2) 在政治方面.Lai 等人<sup>[80]</sup>、Grčar 等人<sup>[81]</sup>、Simaki 等人<sup>[59,82]</sup>利用推特文本数据分析公众对英国脱欧的反应;2016 年的美国大选中,研究者使用推特数据分析民众对 2 名候选人的立场,从而进行选举结果的预测<sup>[83-84]</sup>;Xi 等人<sup>[85]</sup>通过分析用户在脸书(Facebook)上发布的文字和图片分析其政治倾向.

3) 在公共健康方面,一些科学医学有关研究的重要结果经常在社会媒体广泛传播,进而引发公众讨论,对公共健康进行监测,了解用户立场,对于指导制定、完善和评价公共健康干预措施与策略具有积极作用.Zhang 等人<sup>[86]</sup>在线健康论坛上发布了有关补充

<sup>①</sup> <https://www.createdebate.com/>

和替代医学(complementary and alternative medicine, CAM)的具有不同立场的争议性辩论,首先确定现有的在线乳腺癌论坛中有争议的CAM治疗方法,以及检测患者对这些方法有效性的立场,以探索哪种特定的CAM疗法比其他疗法更容易引发辩论。

## 6 总结与展望

本文在充分调研和深入分析的基础上,对文本立场检测的研究进展进行了综述,其中重点介绍了立场检测研究中的关键问题,包括基于不同目标粒度的立场检测、文本立场检测方法、文本立场检测的评测会议与公开数据资源。文本立场检测是一个有价值的研究课题,其研究结果不仅可以推广到情感分析、意图挖掘,同时它与推荐系统、对话系统、谣言验证以及社会感知许多相关研究密切相关。前人的研究工作为文本立场检测研究奠定了坚实的基础,但文本立场检测研究仍然有许多值得挖掘的问题。在本文的最后,基于大量的调研,基于大数据、深度神经网络以及相关研究3个方面带来的促进作用,提出了文本立场检测研究领域一些值得进一步探索的研究点,希望对本领域的其他研究者有所启发。

### 6.1 大数据促进文本立场检测研究

#### 6.1.1 伪标注数据

现实世界中与立场相关的语料分布较广、数量巨大,但难以广泛采集准确标注的数据。大规模、合理的有标注语料有助于获取数据的特征,促进文本立场检测的发展。一种解决策略是运用众包方式采集,通过众包平台发布立场标注任务,召集参与者通过回答与立场相关的问题将立场标注为支持、反对与中立,从而构建立场检测数据集。然而通过此方法标注的数据有规模的限制,只能在一定程度上验证算法的有效性。事实上,互联网上存在着大规模的伪标注数据,这些数据是带标签的训练数据,但标签并不是通过人工标注而获得的,而是自然产生或自动构造的。因此,在数据不充分的情况下,将大规模伪标注数据与人工标注的小数据结合在一起,以提升文本立场检测的性能值得进一步探索。此外,预训练语言模型在多项自然语言处理任务中获得了提升,其思想是在大规模无标注数据集上进行模型预训练,然后在小规模任务数据集上进行模型的微调,也是值得借鉴的方法。

#### 6.1.2 多模态数据

目前来看,立场检测的大部分工作都集中在对

文本类型的内容进行分类,然而,除了文本以外,互联网上充满了图片、语音以及视频等多模态数据信息。在现有的评测任务MultiStanceCat中,Casacufans团队创造性地使用了图片信息,取得了较好的效果。然而,这方面的研究工作还比较少且并不深入。多模态信息通常包括文本、语音、图像、视频等信息,如何在立场检测任务中有效融合多模态信息,使得不同模态的信息相互补充和验证,是一个亟待解决的问题。

#### 6.1.3 外部知识

人们发表的立场文本中通常蕴含一些常识性的知识,例如,在美国大选数据中,某用户在文本中表达支持某个党派,则可根据候选人所属党派推断其对于该候选人的支持立场。引入常识知识能够简化模型构建,一定程度上减少对标注数据的依赖,从而使得模型的泛化能力更强。如何在立场检测中引入这些显然的、易于获取的知识,从而促进和提高立场检测的性能是值得深入研究的问题。

### 6.2 深度神经网络促进文本立场检测研究

本文中介绍的方法大多对文本进行顺序化处理,这样导致每个单词的学习仅依赖于其附近的单词,无法捕获中长距离的依赖关系,并且重复的单词会学习多次,增加了学习的难度。近年来,深度学习方法在图上的扩展吸引了越来越多学者的注意,进而定义和设计了图神经网络(graph neural network, GNN)。将图神经网络应用在立场检测任务中,不再以完全顺序化的方式学习文本内容,而是将文本构建为复杂关联的图结构(同质图或者异质图)进行学习,借助图神经网络的强大的全局学习能力和推理能力,将文本分类问题转化为节点的分类问题,是一个值得积极探索的任务。

### 6.3 相关研究促进文本立场检测研究

#### 6.3.1 文本蕴含

在文本蕴含任务中,一个文本作为前提(premise),另一个文本作为假设(hypothesis),如果根据目标 $P$ 可以从给定的源文本中推断出一个文本陈述(假设) $H$ ,那么就称 $P$ 蕴含 $H$ 。立场检测描述的是文本对于目标的立场,也可以看作是2个文本之间的关系,与文本蕴含关系(textual entailment)十分类似。文本立场检测可以视为一种特殊的文本蕴含,其假设是固定的,立场是支持目标或者反对目标。因此,将文本蕴含的研究方法运用在立场检测任务中,以深入理解文本意义,快速获取有效信息,能够一定程度地推动立场检测研究的发展。

### 6.3.2 社会计算

在社会网络中,用户之间体现了相似即相聚的同质性特点,即具有相同立场的用户倾向于使用相同的词汇或者相同的行为来表达他们的观点.因此,用户之间的相似性可作为推断其立场的核心属性.借助社会计算中的数据挖掘相关算法对用户网络(交互网络、偏好网络、连接网络或者标签网络等)进行建模,挖掘用户关系并在此基础上进行立场检测,也是未来立场检测的研究中值得重点讨论的方向.

## 参 考 文 献

- [1] Thomas M, Pang Bo, Lee L. Get out the vote: Determining support or opposition from congressional floor-debate transcripts [C] //Proc of the 2006 Conf on Empirical Methods in Natural Language Processing. New York: ACM, 2006: 327–335
- [2] Mohammad S, Kiritchenko S, Sobhani P, et al. SemEval-2016 task 6: Detecting stance in tweets [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 31–41
- [3] Xu Ruiheng, Zhou Yu, Wu Dongyi, et al. Overview of NLPCC Shared Task 4: Stance detection in Chinese microblogs [C] //Proc of the ICCPOL 2016, Co-Located with NLPCC 2016. Berlin: Springer, 2016: 907–916
- [4] Sobhani P, Inkpen D, Zhu Xiaodan. A dataset for multi-target stance detection [C] //Proc of the 15th Conf of the European Chapter of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2017: 551–557
- [5] Xu Chang, Paris C, Nepal S, et al. Cross-target stance classification with self-attention networks [C] //Proc of the 56th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2018: 778–783
- [6] Dias M, Becker K. INF-UFRGS-OPION-MINING at SemEval-2016 Task 6: Automatic generation of a training corpus for unsupervised identification of stance in tweets [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 378–383
- [7] Du Jiachen, Xu Ruiheng, He Yulan, et al. Stance classification with target-specific neural attention networks [C] //Proc of the 26th Int Joint Conf on Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 2017: 3988–3994
- [8] Dian Yujie, Jin Qin, Wu Huimin. Stance detection in Chinese microblogs via fusing multiple text features [J]. Computer Engineering and Applications, 2017, 53(21): 77–84 (in Chinese)  
奠雨洁, 金琴, 吴慧敏. 基于多文本特征融合的中文微博的立场检测[J]. 计算机工程与应用, 2017, 53(21): 77–84
- [9] Wang Rui, ZhouDeyu, Jiang Mingmin, et al. A survey on opinion mining: From stance to product aspect [J]. IEEE Access, 2019, 7: 41101–41124
- [10] Titov I, McDonald R. Modeling online reviews with multigrain topic models [C] //Proc of Int Conf of World Wide Web. New York: ACM, 2008: 111–120
- [11] Sobhani P, Inkpen D, Zhu Xiaodan. Exploring deep neural networks for multitarget stance detection [J]. Computational Intelligence, 2019, 35(1): 82–97
- [12] Wei Penghui, Lin Junjie, Mao Wenji. Multi-target stance detection via a dynamic memory-augmented network [C] //Proc of the 41st Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2018: 1229–1232
- [13] Siddiqua U A, Chy A N, Aono M. Tweet stance detection using an attention based neural ensemble model [C] //Proc of the 2019 NAACL-HLT. Stroudsburg, PA: ACL, 2019: 1868–1873
- [14] Ding Zixiang, Xia Rui, Yu Jianfei, et al. Densely connected bidirectional LSTM with applications to sentence classification [J]. arXiv preprint, arXiv: 1802.00889, 2018
- [15] Moniz J R A, KruegeR D. Nested LSTMs [J]. arXiv preprint, arXiv: 1801.10308, 2018
- [16] Zarrella G, Marsh A. MITRE at SemEval-2016 Task 6: Transfer learning for stance detection [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 458–463
- [17] Wei Wan, Zhang Xiao, LiuXuqin, et al. Pkudblab at SemEval-2016 task 6: A specific convolutional neural network system for effective stance detection [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 384–388
- [18] Kim Y. Convolutional neural networks for sentence classification [J]. arXiv preprint, arXiv: 1408.5882, 2014
- [19] Riedel B, Augenstein I, Spithourakis G P, et al. A simple but tough-to-beat baseline for the fake news challenge stance detection task [J]. arXiv preprint, arXiv: 1707.03264, 2017
- [20] Bhatt G, Sharma A, Sharma S, et al. On the benefit of combining neural, statistical and external features for fake news identification [J]. arXiv preprint, arXiv: 1712.03935, 2017
- [21] Conforti C, Pilehvar M T, Collier N. Towards automatic fake news detection: Cross-level stance detection in news articles [C] //Proc of the 1st Workshop on Fact Extraction and Verification. Stroudsburg, PA: ACL, 2018: 40–49
- [22] Hasan K S, Ng V. Stance classification of ideological debates: Data, models, features, and constraints [C] //Proc of the 6th Int Joint Conf on Natural Language Processing. Stroudsburg, PA: ACL, 2013: 1348–1356
- [23] Gottipati S, Qiu Minghui, Sim Y, et al. Learning topics and positions from Debatepedia [C] //Proc of the 2013 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2013: 1858–1868
- [24] Somasundaran S, Wiebe J. Recognizing stances in ideological online debates [C] //Proc of the 2010 NAACL-HLT. Stroudsburg, PA: ACL, 2010: 116–124

- [25] Liu Bing. Sentiment Analysis: Mining Opinions, Sentiments, and Emotions [M]. Translated by Liu Kang. 1st ed. Beijing: China Machine Press, 2017  
([美]刘兵. 情感分析: 挖掘观点、情感和情绪[M]. 刘康, 译. 第1版. 北京: 机械工业出版社, 2017)
- [26] Hasan K S, Ng V. Predicting stance in ideological debate with rich linguistic knowledge [C] //Proc of Int Conf on Computational Linguistics. New York: ACM, 2012: 451–460
- [27] Yuan Jianhua, ZhaoYanyan, Xu Jingfang, et al. Exploring answer stance detection with recurrent conditional attention [C] //Proc of Association for the Advancement of Artificial Intelligence. Palo Alto, CA: AAAI, 2019: 7426–7433
- [28] Murakami A, Raymond R. Support or oppose? Classifying positions in online debates from reply activities and opinion expressions [C] //Proc of the 23rd Int Conf on Computational Linguistics. Stroudsburg, PA: ACL, 2010: 869–875
- [29] Durmus E, Ladhak F, Cardie C. Determining relative argument specificity and stance for complex argumentative structures [J]. arXiv preprint, arXiv: 1906.11313, 2019
- [30] Xu Jiaming, Zheng Suncong, Shi Jing, et al. Ensemble of feature sets and classification methods for stance detection [C] //Proc of the ICCPOL 2016, Co-Located with NLPCC 2016. Berlin: Springer, 2016: 679–688
- [31] Sun Qingying, Wang Zhongqing, Zhu Qiaoming, et al. Exploring various linguistic features for stance detection [C] //Proc of the ICCPOL 2016, Co-Located with NLPCC 2016. Berlin: Springer, 2016: 840–847
- [32] Mohammad S M, Sobhani P, Kiritchenko S. Stance and sentiment in tweets [J]. ACM Transactions on Internet Technology, 2017, 17(3): 1–26
- [33] Ebrahimi J, Dou Dejing, Lowd D. A joint sentiment-target-stance model for stance classification in tweets [C] //Proc of Int Conf on Computational Linguistics. New York: ACM, 2016: 2656–2665
- [34] Liu Liran, Feng Shi, Wang Daling, et al. An empirical study on Chinese microblog stance detection using supervised and semi-supervised machine learning methods [C] //Proc of the ICCPOL 2016, Co-Located with NLPCC 2016. Berlin: Springer, 2016: 753–765
- [35] Ahmed A, Xing E P. Staying informed: Supervised and semi-supervised multi-view topical analysis of ideological perspective [C] //Proc of the 2010 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2010: 1140–1150
- [36] Vijayaraghavan P, Syssoev I, Vosoughi S, et al. Deep stance at SemEval-2016 Task 6: Detecting stance in tweets using character and word-level CNNs [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 413–419
- [37] Taulé M, Pardo F M R, Martí M A, et al. Overview of the task on multimodal stance detection in tweets on Catalan #1Oct referendum [C/OL]. 2018 [2019-10-23]. <http://ceur-ws.org/Vol-2150/>
- [38] Augenstein I, Rocktäschel T, Vlachos A, et al. Stance detection with bidirectional conditional encoding [C] //Proc of the 2016 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2016, 2016: 876–885
- [39] Sun Qingying, Wang Zhongqing, Li Shoushan, et al. Stance detection via sentiment information and neural network model [J]. Frontiers of Computer Science, 2019, 13(1): 127–138
- [40] Yue Tianchi, Zhang Shaowu, Yang Liang, et al. Stance detection method based on two-stage attention mechanism [J]. Journal of Guangxi Normal University: Natural Science Edition, 2019, 37(1): 46–53 (in Chinese)  
(岳天驰, 张邵武, 杨亮, 等. 基于两阶段注意力机制的立场检测方法[J]. 广西师范大学学报: 自然科学版, 2019, 37(1): 46–53)
- [41] Yan Yao. An analysis of the text stance of social media based on the subject target information [D]. Harbin: Harbin Institute of Technology, 2018 (in Chinese)  
(颜瑶. 结合主题目标信息的社交媒体文本立场分析[D]. 哈尔滨: 哈尔滨工业大学, 2018)
- [42] Sun Qingying, Wang Zhongqing, Zhu Qiaoming, et al. Stance detection with hierarchical attention network [C] //Proc of Int Conf on Computational Linguistics. New York: ACM, 2018: 2399–2409
- [43] Yu Nan, Pan Da, Zhang Meishan, et al. Stance detection in Chinese microblogs with neural networks [C] //Proc of the ICCPOL 2016, Co-Located with NLPCC 2016. Berlin: Springer, 2016: 893–900
- [44] Bai Jing, Li Fei, Ji Donghong. Attention based BiLSTM-CNN Chinese microblog stance detection model [J]. Computer Applications and Software, 2018, 35(3): 266–274 (in Chinese)  
(白静, 李霏, 姬东鸿. 基于注意力的BiLSTM-CNN中文微博立场检测模型[J]. 计算机应用与软件, 2018, 35(3): 266–274)
- [45] Sukhbaatar S, Szlam A, Weston J, et al. End-to-end memory networks [J]. arXiv preprint, arXiv: 1503.08895, 2015
- [46] Mohtarami M, Baly R, Glass J, et al. Automatic stance detection using end-to-end memory networks [C] //Proc of the 2018 Conf of the 2018 NAACL-HLT. Stroudsburg, PA: ACL, 2018: 767–776
- [47] Peters M E, Neumann M, Iyyer M, et al. Deep contextualized word representations [C] //Proc of the 2018 Conf of the 2018 NAACL-HLT. Stroudsburg, PA: ACL, 2018: 2227–2237
- [48] Radford A, Narasimhan K, Salimans T, et al. Improving language understanding by generative pre-training [OL]. 2018 [2020-09-20]. <https://s3-us-west-2.amazonaws.com/openaiassets/research-covers/languageunsupervised/language-understanding-paper.pdf>

- [49] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [J]. arXiv preprint, arXiv:1810.04805, 2018
- [50] Rao P. Transfer learning in NLP for Tweet stance classification [OL]. 2019 [2019-12-27]. <https://towardsdatascience.com/transfer-learning-in-nlp-for-tweet-stance-classification-8ab014da8dde>
- [51] Howard J, Ruder S. Universal language model fine-tuning for text classification [J]. arXiv preprint, arXiv: 1801.06146, 2018
- [52] Wojatzki M, Zesch T. ltl.uni-due at semeval-2016 task 6: Stance detection in social media using stacked classifiers [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 428–433
- [53] Dey K, Shrivastava R, Kaushik S. Twitter stance detection—A subjectivity and sentiment polarity inspired two-phase approach [C] //Proc of the 2017 IEEE Int Conf on Data Mining. Piscataway, NJ: IEEE, 2017: 365–372
- [54] Liu Zhan, Tian Ningmeng, Wang Hongyu, et al. Research on the stance discrimination of Chinese microblog [J]. Knowledge Management Forum, 2017, 2(3): 175–185 (in Chinese)  
(刘勘, 田宁梦, 王宏宇, 等. 中文微博的立场判别研究[J]. 知识管理论坛, 2017, 2(3): 175–185)
- [55] Zhang Zhihua, Lan Man. ECNU at SemEval 2016 task 6: Relevant or not? Supportive or not? A two-step learning system for automatic detecting stance in tweets [C] //Proc of SemEval-2016. Stroudsburg, PA: ACL, 2016: 451–457
- [56] Dean P, Delip R. Fake news challenge stage 1 (FNC-1): Stance detection [OL]. [ 2020-07-16 ]. <http://www.fakenewschallenge.org>
- [57] Ferreira W, Vlachos A. Emergent: A novel dataset for stance classification [C] //Proc of the 2018 Conf of the 2018 NAACL-HLT. Stroudsburg, PA: ACL, 2016: 1163–1168
- [58] Taulé M, Martí M A, Rangel F, et al. Overview of the task on stance and gender detection in tweets on Catalan independence at IberEval 2017. [C/OL]. 2017 [2019-10-23]. <http://ceur-ws.org/Vol-1881/>
- [59] Simaki V, Carita P, Maria S, et al. Annotating speaker stance in discourse: The Brexit blog corpus [J/OL]. Corpus Linguistics & Linguistic Theory, 2017 [2020-07-06]. <http://lup.lub.lu.se/record/18e17da2-9c3f-4c27-8980-7a04d12aacda>
- [60] Küçük D. Stance detection in Turkish tweets [J]. arXiv preprint, arXiv:1706.06894, 2017
- [61] Darwish K, Magdy W, Zanouda T. Improved stance prediction in a user similarity feature space [C] //Proc of the 2017 IEEE/ACM Int Conf on Advances in Social Networks Analysis and Mining. New York: ACM, 2017: 145–148
- [62] Hercig T, Krejzl P, Hourová B, et al. Detecting stance in Czech news commentaries [C/OL]. 2017 [2020-07-20]. <http://ceur-ws.org/Vol-1885/>
- [63] Swami S, Khandelwal K, Singh V, et al. An English-Hindi code-mixed corpus: Stance annotation and baseline system [J]. arXiv preprint, arXiv:1805.11868, 2018
- [64] Lai M, Patti V, Ruffo G, et al. Stance evolution and Twitter interactions in an Italian political debate [C] //Proc of the 23rd Int Conf on Natural Language and Information Systems. Berlin: Springer, 2018: 15–27
- [65] Lozhnikov N, Derczynski L, Mazzara M. Stance prediction for Russian: Data and analysis [J]. arXiv preprint, arXiv: 1809.01574, 2018
- [66] Li Ran, Lin Zheng, Lin Hailun, et al. Text emotion analysis: A survey [J]. Journal of Computer Research and Development, 2018, 55(1): 30–52 (in Chinese)  
(李然, 林政, 林海伦, 等. 文本情绪分析综述[J]. 计算机研究与发展, 2018, 55(1): 30–52)
- [67] Zhang Shuai, Yao Lina, Sun Aixin, et al. Deep learning based recommender system: A survey and new perspectives [J]. ACM Computing Surveys, 2017, 52(1): 1–38
- [68] Qiu Minghui, Sim Y, Smith N A. Modeling user arguments, interactions, and attributes for stance prediction in online debate forums [J]. Journal of Computational & Applied Mathematics, 2015, 237(1): 62–82
- [69] Amiri H, Daumé H III. Target-dependent churn classification in microblogs [C] //Proc of Association for the Advancement of Artificial Intelligence. Palo Alto, CA: AAAI, 2015: 2361–2367
- [70] Amiri H, Daumé H III. Short text representation for detecting churn in microblogs [C] //Proc of Association for the Advancement of Artificial Intelligence. Palo Alto, CA: AAAI, 2016: 2566–2572
- [71] Pan Da. Research and application of stance detection method for Chinese short text dialogue [D]. Harbin: Heilongjiang University, 2017 (in Chinese)  
(潘达. 面向中文短文本对话的立场检测方法研究及其应用[D]. 哈尔滨: 黑龙江大学, 2017)
- [72] Zubaga A, Kochkina E, Liakata M, et al. Stance classification in rumors as a sequential task exploiting the tree structure of social media conversations [C] //Proc of Int Conf on Computational Linguistics. New York: ACM, 2016: 2438–2448
- [73] Kumar S, Carley K M. Tree LSTMs with convolution units to predict stance and rumor veracity in social media conversations [C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 5047–5058
- [74] Orbach M, Bilu Y, Toledo A, et al. Out of the echo chamber: Detecting countering debate speeches [C] //Proc of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 7073–7086
- [75] Baly R, Mohtarami M, Glass J, et al. Integrating stance detection and fact checking in a unified corpus [C] //Proc of the 2020 NAACL-HLT. Stroudsburg, PA: ACL, 2020: 21–27

- [76] Rebollo C, Gualda E. The refugee crisis on Twitter: A diversity of discourses at a European crossroads [J]. Journal of Spatial and Organizational Dynamics, 2016, 4(3): 199–212
- [77] Bartlett J, Norrie R. Immigration on Twitter: Understanding public attitudes online [OL]. 2015 [2019-12-27]. <https://apo.org.au/node/54788>
- [78] Darwish K, Magdy W, Rahimi A, et al. Predicting onlineislamophobic behavior after # parisattacks [J]. The Journal of Web Science, 2018, 4(2): 34–52
- [79] Demszky D, Garg N, Voigt R, et al. Analyzing polarization in social media: Method and application to tweets on 21 mass shootings [C] //Proc of the 2019 NAACL-HLT. Stroudsburg, PA: ACL, 2019: 2970–3005
- [80] Lai M, Cignarella A T, Hernandez-Farias D I, et al. Multilingual stance detection in social media political debates [J]. Computer Speech & Language, 2020, 63: 1–27
- [81] Grčar M, Cherepnalkoski D, Mozetič I, et al. Stance and influence of Twitter users regarding the Brexit referendum [J/OL]. 2017 [2020-09-20]. <https://link.springer.com/article/10.1186/s40649-017-0042-6>
- [82] Simaki V, Paradis C, Kerren A. Stance classification in texts from blogs on the 2016 British referendum [C/OL] //Proc of Speech and Computer. Berlin: Springer, 2017 [2020-09-20]. [https://link.springer.com/chapter/10.1007%2F978-3-319-66429-3\\_70](https://link.springer.com/chapter/10.1007%2F978-3-319-66429-3_70)
- [83] Lai M, Farias D I H, Patti V, et al. Friends and enemies of Clinton and Trump: Using context for detecting stance in political tweets [C] //Proc of the Mexican Int Conf on Artificial Intelligence. Berlin: Springer, 2016: 155–168
- [84] Darwish K, Magdy W, Zanouda T. Trump vs Hillary: What went viral during the 2016 US Presidential Election [C] //Proc of the Int Conf on Social Informatics. Berlin: Springer, 2017: 143–161
- [85] Xi Nan, Ma Di, Liou M, et al. Understanding the political ideology of legislators from social media images [C] //Proc of Association for the Advancement of Artificial Intelligence. Palo Alto, CA: AAAI, 2020: 726–737
- [86] Zhang Shaodian, Qiu Lin, Chen F, et al. We make choices we think are going to save us: Debate and stance identification for online breast cancer CAM discussions [C] //Proc of Int Conf of World Wide Web. New York: ACM, 2017: 1073–1081



**Li Yang**, born in 1987. PhD. Associate professor. Member of CCF. Her main research interests include natural language processing and information recommendation.  
李 洋,1987 年生.博士,副教授,CCF 会员.主要研究方向为自然语言处理和信息推荐.



**Sun Yuqing**, born in 1996. Master candidate. Student member of CCF. Her main research interests include natural language processing and stance detection.  
孙宇晴,1996 年生.硕士研究生,CCF 学生会员.主要研究方向为自然语言处理、立场检测.



**Jing Weipeng**, born in 1979. PhD. Professor. Member of ACM and IEEE, and senior member of CCF. His main research interests include resource management in cloud computing, AI/machine learning, big data analytics.  
景维鹏,1979 年生.博士,教授,ACM 会员,IEEE 会员,CCF 高级会员.主要研究方向为云计算中的资源管理、人工智能/机器学习、大数据分析.