

低资源方面级情感分析研究综述

陈 壮 钱铁云 李万理 张 婷 周 燊 钟 鸣 祝园园 刘梦赤

(武汉大学计算机学院 武汉 430072)

摘 要 方面级情感分析 (Aspect-Based Sentiment Analysis) 旨在识别文本中用户对于特定对象所表达的观点信息, 是情感计算领域的前沿课题, 近年来受到学术界和工业界的广泛关注. 方面级情感分析涉及到方面词、观点词、方面类别和情感极性等多种要素, 并相应地构成了多种目标任务. 随着自然语言处理技术的高速发展, 研究者们针对方面级情感分析提出了众多解决方案, 在性能表现上取得了一定进展. 然而, 受限于方面级情感分析高昂的标注代价, 现有研究大多局限于特定的语言和领域, 训练样本不足导致的低资源场景阻碍了相关方法的进一步拓展与应用. 本文首先从方面级情感分析相关定义出发, 对其所涉及的各种要素及任务进行了阐述, 并介绍了常用数据集及评价指标; 其次, 针对方面级情感分析包含的基本任务和扩展任务, 详细深入地综述和分析了相关模型的发展历程; 再次, 对于低资源场景下的方面级情感分析任务, 从模型和数据两个层面对现有改进策略进行了归纳和对比; 最后, 对低资源方面级情感分析的未来研究趋势进行了展望和总结.

关键词 低资源场景; 方面级情感分析; 数据增强; 领域适应; 信息抽取

中图法分类号 TP391 **DOI号** 10.11897/SP.J.1016.2023.01445

Low-Resource Aspect-Based Sentiment Analysis: A Survey

CHEN Zhuang QIAN Tie-Yun LI Wan-Li ZHANG Ting
ZHOU Shen ZHONG Ming ZHU Yuan-Yuan LIU Meng-Chi

(School of Computer Science, Wuhan University, Wuhan 430072)

Abstract Aspect-based sentiment analysis (ABSA) aims to identify the opinions expressed by users about specific targets in subjective texts. In recent years, ABSA has been a frontier topic in the field of affective computing, and has received extensive attention from both academia and industry. ABSA involves various elements such as aspect terms, opinion terms, aspect categories, and sentiment polarity, and accordingly constitutes various research tasks. With the rapid development of natural language processing, researchers have proposed a large amount of methods for ABSA, and have made certain progress in performance. However, limited by the high annotation cost, most existing researches are limited to specific languages and fields, and the low-resource scenarios caused by insufficient training samples hinder the further expansion and application of ABSA methods. In this paper, we start from the relevant definitions of ABSA and introduces the elements, tasks, commonly-used datasets, and evaluation metrics. Then, for the basic tasks and extended tasks included in ABSA, we thoroughly review and analyze the development of related ABSA methods. Further, for the low-resource problem in ABSA, we summarize and compare the

收稿日期: 2022-09-09; 在线发布日期: 2023-04-18. 本课题得到国家自然科学基金项目(62276193, 61572376)资助. 陈 壮, 博士, 博士后研究员, 主要研究领域为信息抽取、情感分析. E-mail: zhchen18@whu.edu.cn. 钱铁云 (通信作者), 博士, 教授, 主要研究领域为自然语言处理、推荐系统. E-mail: qty@whu.edu.cn. 李万理, 博士研究生, 主要研究方向为信息抽取. 张 婷, 硕士研究生, 主要研究方向为情感分析、对话系统. 周 燊, 硕士研究生, 主要研究方向为情感分析. 钟 鸣, 博士, 副教授, 主要研究方向为大数据、社会网络分析. 祝园园, 博士, 副教授, 主要研究方向为图数据库. 刘梦赤, 博士, 教授, 主要研究领域为数据库、大数据管理和分析、智能信息系统.

existing improving strategies for low-resource ABSA, which can be categorized into the model level and the data level. Finally, we prospect the future research trends of low-resource ABSA.

Keywords low-resource scenario; aspect-based sentiment analysis; data augmentation; domain adaptation; information extraction

1 引 言

互联网时代的来临,使得用户在网络上的活动方式发生了重大变化,由被动接收信息逐渐转变为主动获取并产生内容.同时,移动终端和社交媒体的高速发展也进一步加剧了网络对人们日常生活的渗透.根据数据显示,国内社交平台微博和知乎的月活跃用户数量分别超过 5.73 亿和 1.01 亿,而购物平台淘宝和拼多多的月活跃用户数则分别超过 8.74 亿和 8.24 亿.用户现在已经逐渐适应通过在线撰写发帖和评论的方式,表达自己对特定对象的观点和看法.这种趋势导致互联网上累积了大量包含主观态度的文本数据.

情感分析(Sentiment Analysis)是一项旨在通过分析自然语言文本,了解人们对某一对象表达的情感、态度、观点、评价或情绪的任务.它是自然语言处理领域中的一个关键问题,已有大量的研究投入.由于对情感和观点的理解是分析人类行为的必要途径,情感分析在各种场景下都有着重要应用.首先,对于管理机构,舆情监控、热点追踪和快速定位等均是情感分析的重要应用.运用自动化的情感分析工具,相比以往的人工采访和调查问卷等手段更有效,可以提升工作效率,实现决策科学化、治理精准化和服务高效化.其次,对于企业商户,通过收集大量真实且有针对性的情感数据,企业主体可以有效预测行业的发展趋势,制定良好的战略并采取积极的应对措施,进而提高产品质量、改进营销策略并增强个性化推荐.再次,对于个体用户,人们现在已经习惯于向不同平台的各种用户群体寻求帮助,而不仅仅局限于现实中的家人和朋友.情感分析技术能为个人用户提供详细的参考信息,辅助其决策过程并建立相应社交网络,极大地提高了可用信息量.

情感分析可根据目标粒度的范围分为文档级、语句级和方面级三类任务.在该领域的早期研究中,文档级和语句级是主要任务,旨在判断每个文档或句子的整体情感倾向.作为常规的文本分类任务,上述任务相对简单,相关方法已经取得了优异的性

能^[1].然而,文档级和语句级任务因缺乏对目标对象的建模,不能对文本进行深入挖掘,导致分析结果的信息量受限.在 20 世纪初,Hu 等学者^[2]提出了同时提取商品特征和相关情感信息的想法,这为更细粒度的情感分析任务开创了新的方向.随后,多项后续研究^[3-5]陆续提出方面级情感分析(Asspect-Based Sentiment Analysis)概念,在领域内引发关注.在 2014 至 2016 年期间,国际语义评测大赛 SemEval^[6-8]亦连续三年设置了方面级情感分析赛道,有力地推动了该领域研究的发展.

目前的大部分研究采用有监督学习的方式来完成方面级情感分析任务.一般而言,有监督学习需要大量标注语料来训练模型,以获得较好的性能表现^[9].然而,由于数据标注的人工代价高昂,在真实应用中往往只能获得有限的标注样本,造成相关模型的性能表现受限.以上这种标注样本数量不足的学习场景被称为低资源场景(Low-Resource Scenario)^[10].

低资源场景包括语言、领域和任务等多种不同的资源条件.从语言角度看,受到经济和政治等多重因素的影响,全球现存的 7000 种语言只有大约十种拥有充足的标注资源,因此小语种自然语言处理普遍存在低资源场景.从领域角度看,标注样本的数量在热门和冷门领域之间存在巨大差异.关注度低、标注样本少的冷门领域在学术研究和工业落地中都面临低资源场景.从任务角度看,复杂任务因标注难度高,相较于简单任务而言标注样本数量少.例如,粗粒度的文本分类任务拥有数十万计的训练数据,而细粒度的方面级情感分析等任务则通常仅有数千条训练数据,这使得多数复杂任务也面临低资源场景.因此,低资源问题在多个方面阻碍了自然语言处理模型的泛化与应用,Ruder 等学者将其称为领域内现存的四大问题之一^①.关注细粒度的词级(token-level)特征是方面级情感分析任务的关键之处,但是词级标注过程需要花费大量的人力成本,因此常用的方面级情感分析数据集如 SemEval-2014 Task 4 等,仅包含约三千条乃至更少的训练样本.在面对多

① <https://ruder.io/4-biggest-open-problems-in-nlp>

种不同要素和交互关系时,标注样本数量的不足使得模型难以获得充分的训练,使得相关方法在低资源学习场景下的性能表现不佳.因此,开展面向低资源场景的方面级情感分析研究,提高模型在资源受限条件下的性能表现,促进相关方法在更多语言、领域和任务中的推广,具有重要的理论价值和实际意义.

2 问题定义、数据集和评价指标

2.1 问题定义

方面级情感分析的目的是对文本中用户表达的主观信息进行总结,以深入理解其中的情感语义.该任务主要涉及文本中的四种要素,具体如图 1 中的样例所示.

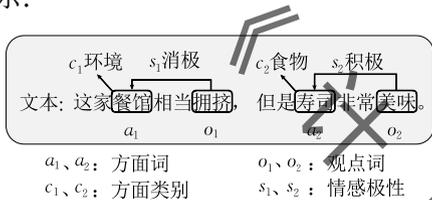


图 1 对主观文本进行方面级情感分析的示例

(1) 方面词 (Aspect Term). 一般来说,用户在文本中表达情感时,会指明具体对象.这些对象可以是单个词,也可以是由多个词组成的连续片段.在上述例子中,餐馆(a_1)和寿司(a_2)就是两个方面词.

(2) 观点词 (Opinion Term). 用户会在文本中以主观描述的形式表达对目标对象的看法.在上述例子中,拥挤(o_1)和美味(o_2)就是两个观点词.

(3) 方面类别 (Aspect Category). 用户对所评价对象进行的语义类别划分,通常归属于领域相关的预定义有限概念集合.例如,在上述例子中,方面词餐馆(a_1)和方面类别环境(c_1)相对应,而寿司(a_2)则与食物(c_2)相对应.

(4) 情感极性 (Sentiment Polarity). 用户对目标对象的情感信息可以被归类为预定义的且与领域无关的有限极性集合,包括积极、中立、消极和矛盾.在上面的例子中,方面词餐馆(a_1)和寿司(a_2)分别与观点词拥挤(o_1)和美味(o_2)相关联,呈现出消极(s_1)和积极(s_2)的情感极性.

方面级情感分析中的两个首要独立任务^①是方面词抽取和方面情感分类.它们由 Hu 等学者^[2]最先在细粒度情感分析领域中提出,且是早期相关研究中唯二的子任务^[3].这两个任务的目的是挖掘文本中用户描述的对象和用户对该对象的看法.一旦这两个问题得到解决,文本所包含的最重要的主观

信息就已被呈现出来.它们还可配对形成一个组合任务,即方面-情感对抽取^②,旨在通过一步式的方法直接从文本中抽取方面词并判断其情感极性.以上三个任务可被认为是方面级情感分析的基本任务.除此之外,由于方面级情感分析还涉及观点词和方面类别这两种要素,还形成了其它的独立任务,如观点词抽取^[11]、方面观点抽取^[12]、类别识别^[13]和类别情感分类^[14]等,并可相应配对成方面-观点对抽取^[15]和类别-情感对抽取^[16]等组合任务.由于这些额外的任务扩展了领域研究范围,因此可被认为是方面级情感分析的扩展任务.随着研究的不断深入,更多的扩展任务逐渐被提出,例如三元组抽取^[17-23]、四元组抽取^[24-25]等.此处以图 1 中文本为例,对方面级情感分析所涉及的任务进行详细定义,并在图 2 中展示了各任务的具体构成和相互关联.

(1) 方面词抽取 (Aspect Term Extraction). 该任务以文本为输入,旨在提取涉及的方面词(如餐馆、寿司).方面词抽取的目标显式出现在文本中,其属于抽取任务.方面级情感分析以方面词为基本对象,因此方面词抽取是核心上游任务.

(2) 方面情感分类 (Aspect-Level Sentiment Classification). 该任务通过输入文本和某一给定方面词(如餐馆),判断方面词所对应的情感极性(如消极).方面情感分类的结果通常属于特定的预定义集合,包括积极、中立、消极和矛盾,其属于分类任务.因为方面词需要被提前给定,所以方面情感分类是方面词抽取的下游任务.

(3) 方面-情感对抽取 (Aspect-Sentiment Pair Extraction). 方面词抽取和方面情感分类的组合任务.该任务以文本为输入,目标是同时抽取方面词并判断每个方面词的情感极性(餐馆-消极和寿司-积极),包含了抽取和分类两类任务形式.这是一个端到端的组合任务,一步式地归纳了文本对所有方面词的情感极性,是方面级情感分析的最简可行形式 (Minimum Viable Format).

(4) 观点词抽取 (Opinion Term Extraction). 该任务以文本为输入,目标是抽取所提及的所有观点词(如拥挤、美味).观点词抽取的结果是文本中所出现的词,其属于抽取任务.

① 独立任务定义为输出只包含一种要素的任务,如方面词抽取的输出只包含方面词要素.组合任务定义为输出包含多种要素的任务,如方面-情感对抽取任务的输出同时包含方面词和情感极性两种要素.

② 该任务在文献中也常被称为端到端方面级情感分析 (End-to-End Aspect-Based Sentiment Analysis, E2E-ABSA) 或统一方面级情感分析 (Unified Aspect-Based Sentiment Analysis).

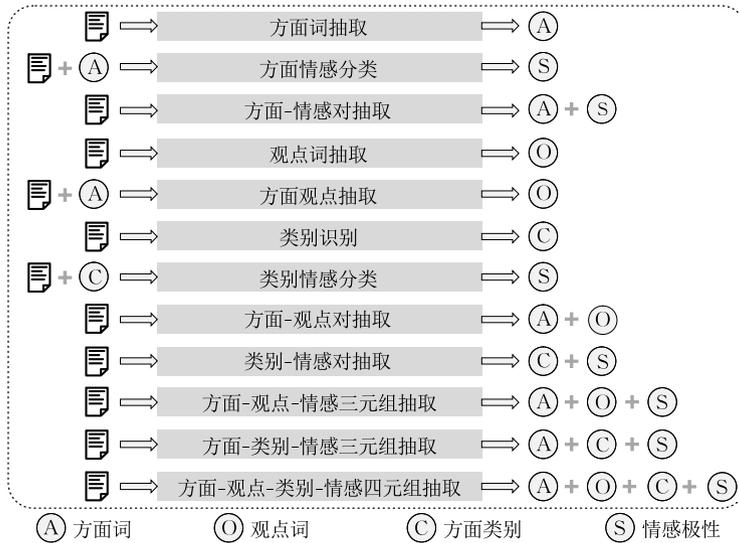


图 2 方面级情感分析中的任务一览

(5) 方面观点抽取 (Aspect-Oriented Opinion Extraction). 该任务以文本和某一给定方面词为输入, 目标是提取方面词所对应的观点词 (如拥挤). 观点词显式出现在输入文本中, 对其的提取属于抽取任务. 由于该任务需要给定方面词, 因此属于方面词抽取的下游任务.

(6) 类别识别 (Aspect Category Detection). 该任务以文本为输入, 目标是识别所有方面类别 (如环境、食物). 方面类别的分类结果一般属于特定的预定义集合, 其属于分类任务.

(7) 类别情感分类 (Aspect Category Sentiment Classification). 该任务以文本某一给定的方面类别为输入, 目标是判断方面类别的情感极性. 方面情感分类的结果通常也属于特定的预定义集合, 即积极、中立、消极和矛盾, 其属于分类任务. 由于此处的方面类别需要事先给定, 类别情感分类是类别识别的下游任务.

(8) 方面-观点对抽取 (Aspect-Opinion Pair Extraction). 该任务以文本为输入, 目标是识别出输入文本中所有的方面-观点对 (如餐馆-拥挤). 由于方面词与观点词都是输入文本中所出现的词, 其属于抽取任务. 方面-观点对抽取, 一次性抽取出了所有的方面词以及对应的观点词, 为后续研究更细致的任务 (如三元组抽取、四元组抽取) 奠定了基础.

(9) 类别-情感对抽取 (Category-Sentiment Pair Extraction). 该任务以文本为输入, 目标是识别出输入文本中所有的方面类别以及对应的情感极性 (如环境-消极). 由于方面类别与情感极性都归属于预定义的集合, 该任务本质上仍可划分为分类任务,

但需要额外对分类结果进行配对抽取.

(10) 方面-观点-情感三元组抽取 (Aspect-Opinion-Sentiment Triplet Extraction). 该任务以文本为输入, 目标是识别出所有的 (方面词, 观点词, 情感极性) 三元组 (如餐馆-拥挤-消极). 该任务不仅要求抽取出方面-观点对, 还要求判断其情感极性, 包含了抽取和分类两种任务形式.

(11) 方面-类别-情感三元组抽取 (Aspect-Category-Sentiment Triplet Extraction). 该任务以文本为输入, 目标是识别出所有的 (方面词-方面类别-情感极性) 三元组 (如餐馆-环境-消极). 该任务同样包含了抽取与分类两种任务形式.

(12) 方面-观点-类别-情感四元组抽取 (Aspect Sentiment Quad Extraction). 该任务以文本为输入, 目标是输出方面级情感分析中所有的情感要素 (如餐馆-拥挤-环境-消极). 该任务是当前方面级情感分析中最为完备的任务, 需要关注所有要素, 包含了抽取与分类两种任务形式.

2.2 数据集

现有的方面级情感分析数据集涉及多种领域和语言. 除了一般的基于评论文本的数据集, 还有基于对话、问答等形式的数据集. 近年来, 用于多模态场景和用于鲁棒性测试场景的数据集也陆续出现. 表 1 中汇总了常用数据集的语言、领域、样本数量和链接, 并按照标签信息不同划分为六大类, 即“A, S”、“A, C, S”、“C, S”、“A, O”、“A, O, S”和“A, O, C, S”, 以便参考查阅. 其中, A、O、C、S 分别代表方面级情感分析中的方面词、观点词、方面类别和情感极性. 由标签信息可见, 在方面级情感分析中, 方面词和情

表 1 方面级情感分析常用数据集

数据集	语言	领域	样本数量	标签	链接
Device ^[2]	English	Digital Device	3.8 k	A,S	https://github.com/hsqmlzno1/Transferable-E2E-ABSA
Targeted Sentiment system ^[26]	Multilingual	Twitter	3.2 k 6.6 k	A,S	http://www.m-mitchell.com/code/MitchellEtAl-13-Open-Sentiment.tgz
Twitter 2014 ^[27]	English	Twitter	6.9 k	A,S	https://docs.google.com/forms/d/e/1FAIpQLSd6Hh-Eitpzv-vpdrB63Jbx-fYWohnOZRCUb0ibKjpO21Q_tIQ/viewform
BabyCare ^[28]	Chinese	Baby Care	36.7 k	A,S	http://www.marcpoint.com/junyang.html
ASC-QA ^[29]	Chinese	Electronics Beauty Bags	2.4 k	A,S	https://github.com/jjwangnlp/ASC-QA
MultiModal-ABSA ^[30]	English	Twitter	11.3 k	A,S	https://github.com/jefferyYu/TomBERT
ARTS ^[31]	English	Laptops Restaurants	1.8 k 3.5 k	A,S	https://github.com/zhijing-jin/ARTS_TestSet
Automobiles ^[32]	Multilingual	Automobiles	58.5 k	A,S	https://github.com/dmhyun/alsadata
SemEval-2014 ^[6]	English	Laptops Restaurants	3.8 k 3.8 k	A,S A,C,S	https://alt.qcri.org/semEval2014/task4/
SemEval-2015 ^[7]	English	Laptops Restaurants	2.5 k 2.0 k	C,S A,C,S	https://alt.qcri.org/semEval2015/task12/
SemEval-2016 ^[8]	Multilingual	Laptops Restaurants	3.3 k 2.6 k	C,S A,C,S	https://alt.qcri.org/semEval2016/task5/
TdParse ^[33]	English	Twitter	12.5 k	A,C,S	https://github.com/bwang482/tdparse
MAMS ^[34]	English	Twitter Laptops Restaurants	6.9 k 3.0 k 4.8 k	A,S A,S A,C,S	https://github.com/siat-nlp/MAMS-for-ABSA
SemEval-2017 ^[35]	Multilingual	Twitter	50 k	C,S	https://alt.qcri.org/semEval2017/task4/
ASAP ^[36]	Chinese	Restaurants	46.7 k	C,S	https://github.com/Meituan-Dianping/asap
TOWE ^[12]	English	Laptops Restaurants	1.4 k 4.6 k	A,O	https://github.com/NJUNLP/TOWE
ASTE-Data-V1/2 ^[22,37]	English	Laptops Restaurants	1.4 k 4.5 k	A,O,S	https://github.com/xuuluuu/Position-Aware-Tagging-for-ASTE
MPQA ^[38]	English	News	9.4 k	A,O,S	https://mpqa.cs.pitt.edu/
JDPA ^[39]	English	Automobiles Digital Cameras	0.5 k	A,O,S	https://verbs.colorado.edu/jdpacorpus/
Service ^[40]	English	Online Service	2.2 k	A,O,S	https://tudatalib.ulb.tu-darmstadt.de/handle/tudatalib/2448
CASA ^[41]	Chinese	Daily Dialog News	3.0 k 0.2 k	A,O,S	https://github.com/freesunshine0316/lab-conv-asa
ACOS ^[24]	English	Laptops Restaurants	4.0 k 2.2 k	A,O,C,S	https://github.com/NUSTM/ACOS
ABSA-QUAD ^[25]	English	Restaurants	3.7 k	A,O,C,S	https://github.com/IsakZhang/ABSA-QUAD

感极性所受的关注程度最高,而观点词和方面类别常作为辅助要素出现,以进一步扩展任务范围.这一特点在随后的文献综述章节也将体现.数据集内容详述如下:

(1) Digital Device^[2]为数码设备领域数据集.该数据集收集自 Amazon 网站,包括对于两种数码相机、一种 DVD 播放器、一种 MP3 播放器以及一种手机等五种电子产品的评论文本.该数据集对评论文本中出现的方面词及其极性进行了标注.

(2) Targeted Sentiment System^[26]收集自 Twitter,其中包括大约 3288 条英语推文及 6658 条西班牙语推文.该数据集以推文中的人物、组织、地点、时间等

10 种类型的实体为目标,对实体本身、实体类别及其情感极性进行了标注.

(3) Twitter 2014^[27]收集自 Twitter 社交网站,其中包含了用户对于名人、产品、公司等评论.该数据集对评论文本中的方面词以及其情感极性进行了标注.

(4) BabyCare^[28]为婴儿护理领域中文数据集.该数据集收集自 babytree 网站,其中包括关于尿布、奶粉等话题的经验分享帖子.在该数据集中,实体包括品类名(尿布、奶粉等)、品牌名以及产品名,方面包括防漏能力、价格等. BabyCare 数据集对用户关于实体各个方面的情感极性进行了标注.

(5) ASC-QA^[29] 是基于问答的方面级情感分析数据集. 该数据集从淘宝平台的“问大家”功能中收集得到, 所涉及的领域为电子设备等. 该数据集对问题-回答对中的方面词、方面类别以及细粒度情感极性进行了标注.

(6) MultiModal-ABSA^[30] 由收集自 Twitter 社交网站的多模态数据组成. 在该数据集中, 每条数据由推文和其对应的图像所构成. MultiModal-ABSA 对推文中的方面词以及方面词对应的情感极性进行了标注.

(7) ARTS^[31] 以 SemEval-2014 为基础, 利用 SemEval-2014 的测试集, 对目标方面词所在句子进行了目标方面词情感极性反转、非目标方面词情感极性反转、添加更多不同极性的方面词等扰动操作, 并将所得到的新样本用于测试方面级情感分析模型的鲁棒性.

(8) Automobiles^[32] 为汽车领域数据集. 该数据集收集自美国和韩国的汽车社区网站, 以 12 家最大的汽车厂商所出品的畅销车型作为方面词, 对相应的情感极性进行了标注. Automobiles 中包含英语和韩语 2 个数据集, 相比于其它方面级情感分析数据集, 它们的规模较大, 训练集的样本都超过了 20 000 条.

(9) SemEval-2014^[6] 包括餐馆领域的 Restaurant 数据集、电脑领域的 Laptop 数据集. Restaurant 数据集对评论中的方面词、方面情感极性以及方面类别进行了标注. Laptop 数据集与 Restaurant 数据集类似, 但不含方面类别的标注.

(10) SemEval-2015^[7] 是 SemEval-2014 的延续. 相比 SemEval-2014, SemEval-2015 中的每条数据均为完整的评论, 可能由多个句子组成, 有着更加完整的上下文. 同时, SemEval 提供了一个新的领域, 即 Hotel 数据集作为测试集, 因此可被用于进行跨领域的 ABSA 研究. SemEval-2015 对方面词、方面类别、情感极性进行了标注.

(11) SemEval-2016^[8] 是 SemEval-2015 的延续. 该数据集中提供了 39 个子数据集, 其中 19 个用于训练, 20 个用于测试. 涉及电子、旅馆、餐馆等 7 种领域, 以及英语、阿拉伯语、汉语等 8 种语言, 可用于跨领域、跨语言研究. 标注内容同 SemEval-2015.

(12) TdParse^[33] 收集自 Twitter 社交网站, 由关于英国大选的推文组成. 该数据集中平均每条推文包括 3.09 个标注实体, 且有 49.7% 的推文中含 2~3 种不同的情感极性. TdParse 对句子中的实体(方面词)及其情感极性进行了标注.

(13) MAMS^[34] 为餐馆领域数据集, 分为 MAMS-ATSA 以及 MAMS-ACSA 两个版本, 分别用于方面情感分类和类别情感分类任务. 该数据集具有鲜明特点, 其中每条评论文本中至少包含两个具有不同情感极性的方面词或方面类别, 因此可认为是一个具有挑战性的困难样本集合.

(14) SemEval-2017^[35] 收集自 Twitter 社交网站, 包括英语和阿拉伯语两种语言的数据样本. 该数据集对推文中的事物、事件、产品等进行了情感极性标注.

(15) ASAP^[36] 为餐馆领域的中文数据集. 该数据集针对餐馆评论标注了用户的整体评分(0~5), 同时包含对于餐馆的地理位置、服务、价格等 18 个预定义方面类别的打分. ASAP 数据集的规模较大, 训练集数据达到了 36 850 条.

(16) TOWE^[12] 以 SemEval-2014、SemEval-2015、SemEval-2016 为基础, 标注了这些数据集中的方面词、观点词以及它们的配对关系.

(17) ASTE-Data-V1^[22] 数据集以 SemEval-2014、SemEval-2015、SemEval-2016 数据集为基础, 标注了方面词、观点词、情感极性三元组. Xu 等学者^[37]发现, ASTE-Data-V1 中对同一观点词对应多个方面词的情况标注不完全, 因此补充了相应三元组的标注, 形成了 ASTE-Data-V2 版本.

(18) MPQA^[38] 为新闻领域数据集, 数据来自各种新闻文章. MPQA 数据集对于新闻中的文本片段进行了标注, 并提供了新闻对于该文本片段的情感极性、情感强度, 以及情感相关的对象(如作者、新闻中所提到的人名等).

(19) JDPA^[39] 由汽车领域的博客组成. 该数据集以汽车领域的具体对象(如车座、车门等)和抽象对象(如操作、动力等)为实体, 标注了这些实体对应的情感表达和情感极性. 本数据集还对观点持有者, 以及实体不同提及之间的共指关系进行了标注.

(20) Online Service^[40] 为在线评论领域数据集. 该数据集由收集自 RateItAll 和 eOpinion 网站的评论文本所组成, 其中包括对于 PayPal、eGroups、eTrade、eCircles 等在线服务的评价. Online Service 数据集对于评论中出现的方面词、观点词、情感持有者、情感极性以及情感强度等进行了详细标注.

(21) CASA^[41] 是一个中文数据集, 针对对话场景下的方面级情感分析任务. 该数据集包括来自日常闲聊数据集 DuConv 和新闻讨论数据集 NewsDialogue 中的数据, 涵盖的话题包括电影、电视剧、名人、新闻等. 其中, NewsDialogue 的数据仅用于测试模型在

跨领域情景下的鲁棒性。该数据集对方面词、观点词以及情感极性进行了标注,相比一般的数据集,本数据集中的方面词和观点词可能不在同一句话中,而是分散在对话的各个部分。

(22) ACOS^[24] 包含餐馆和电脑两个领域的数据集。餐馆领域数据集以 SemEval-2016 为基础进行了拓展,电脑领域数据集由作者自行从 Amazon 网站收集所得。ACOS 数据集对于评论中的方面词、方面类别、观点词、情感极性四元组进行了标注,其中,方面词和观点词可能缺省。

(23) ABSA-QUAD^[25] 也是四元组数据集,标注内容同 ACOS。该数据集以 SemEval-2015、SemEval-2016 为基础进行了四元组的标注。

2.3 评价指标

2.3.1 分类评价指标

对于方面情感分类、类别识别等分类型任务,常用的分类评价指标为准确率 (Accuracy) 和 F1-Score。准确率 (Accuracy) 主要用于衡量模型的分类型总体性能,其值等于预测正确样本数占所有样本数量的比值:

$$\text{准确率 (Accuracy)} = \frac{\text{预测正确样本数}}{\text{样本总数}} \quad (1)$$

F1-Score 则关注模型对于不同类别样本的分类性能。为了计算某一类别的 F1-Score 值,先需要统计模型分类结果中的 TP(真正例)、TN(真负例)、FP(假正例)、FN(假负例)数量,再计算精确率 (Precision) 和召回率 (Recall):

$$\text{精确率 (Precision)} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{召回率 (Recall)} = \frac{TP}{TP + FN} \quad (3)$$

据此, F1-Score 可计算为

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

此外,分类任务常涉及到多个不同类别,因而还涉及 Micro-F1 和 Macro-F1 的指标评价。计算 Micro-F1 时,不区分类别地统计 TP、TN、FP、FN,合并计算精确率和召回率后,计算得到 F1-Score。此时,样本数越多的类别,在 Micro-F1 计算过程中的重要性就越高。而计算 Macro-F1 时,则是预先计算出每一类别的 F1-Score,再求所有类别 F1-Score 的算数平均值。此时,所有类别无论样本数多少,对于计算 Macro-F1 是同等重要的。在实际应用中,常采用 Macro-F1 进行评价,以免忽略样本数较少的“中立”情感极性。

2.3.2 抽取评价指标

对于方面词抽取和观点词抽取等抽取型任务,常用的评价指标也为精确率 (Precision)、召回率 (Recall) 和 F1-Score,但计算方式稍有不同。具体地,观察抽取结果时,仅有抽取出的片段与标签片段完全匹配时,才认为是抽取正确。此时,精确率和召回率计算方式如下:

$$\text{精确率 (Precision)} = \frac{\text{正确预测的实体数}}{\text{预测的实体总数}} \quad (5)$$

$$\text{召回率 (Recall)} = \frac{\text{正确预测的实体数}}{\text{标注的实体总数}} \quad (6)$$

此处实体指代方面词、观点词等任务目标词, F1-Score 计算方式则与分类任务相同。由于抽取时通常不考虑方面词、观点词归属的高层类别,因而此处计算的 F1-Score 可认为是 Micro-F1 指标。

需要注意的是,对于方面-情感对抽取、三元组抽取等组合任务,通常也采用 Micro-F1 进行评价。此时,仅需将抽取正确的标准做细微调整即可,即当且仅当抽取出的多元组中的每一个要素都与目标多元组完全匹配,才可视为抽取正确。

3 方面级情感分析方法

本节介绍针对方面级情感分析中基本任务和扩展任务的常用方法,并进行深入对比分析。

3.1 [基本任务] 方面词抽取

方面词抽取的核心目标是通过单词特征判断该单词是否是用户评价的对象。基于不同的学习方式,可将现有研究分为三种类型,即基于无监督学习的方法,基于有监督浅层模型的方法,和基于有监督深层模型的方法。

3.1.1 基于无监督学习的方法

如图 3 所示,传统的方面词和观点词的抽取工作通常依赖于无监督的手工规则和统计信息。基于规则的方法其主要思路是设计方面词和观点词之间

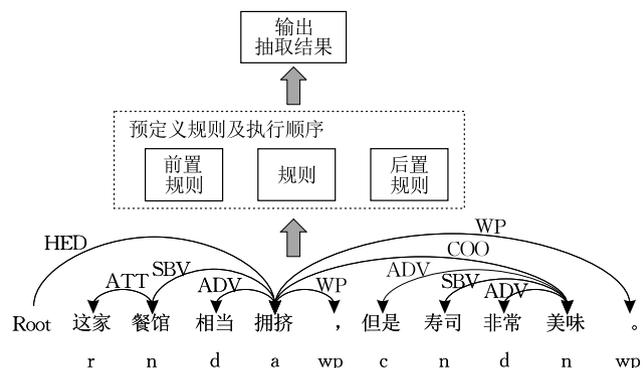


图 3 基于规则的方面词抽取

的依赖关系规则,并将这些规则与文本进行匹配以完成抽取。

Hu 等学者^[2]使用关联挖掘算法从名词或名词组中提取方面词,对于罕见的方面词,会同时借助观点词的邻居来进行补充。Qiu 等学者^[42]提出双向传播(Double Propagation, DP)算法,借助情感词典和依赖路径,制定出一系列与方面词和观点词相关的手工规则,并以自举的方式扩充方面词和观点词集合。与此不同,赵妍妍等学者^[43]自动挖掘大规模语料中潜在的方面词和观点词依赖路径规则,并通过计算测试样本和规则样本的编辑距离进行匹配抽取。Liu 等学者^[44]应用了词级翻译模型,实现了方面词和观点词的对齐,并在共现关系的帮助下实现抽取性能的提高。Gindl 等学者^[45]使用词性模板匹配方面词,并利用跨句情感传播以增大方面词和观点词的相互作用范围。在对潜在的方面词和观点词进行聚类之后,Zhao 等学者^[46]还进行了不同高层次类别之间的关联信息挖掘。Liu 等学者^[47]为了在不同方面词之间建立更准确的关联,提出了一种基于语义相似性和方面词关联挖掘的方法,进一步增强了双向传播算法。Gunes 等学者^[48]提出通过词性标注和互信息对方面词进行聚类、分段和层次模型归纳。江腾蛟等学者^[49]为了进行金融领域的方面词和观点词抽取,以样本的词性标注、依存句法分析和语义角色标注结果为基础,完成了一系列手工规则的制定。Shafie 等学者^[50]通过实验,对比了在抽取方面词和观点词的过程中,不同类型的依赖关系和词性信息规则的作用。

如图 4 所示,基于统计信息的方法为了从文本中挖掘潜在的方面类别,并进一步挖掘与每个类别高度相关的潜在方面词,通常会设计和改进各类主题模型。现有工作主要沿着两种研究思路改进主题模型,以便使其更好地应用于方面词抽取任务。第一种思路是将与方面词相关的先验知识补充到主题模型中。例如,Mukherjee 等学者^[51]通过使用少量人工挑选的种子方面词,来提升传统主题模型在方面词抽取上的性能。为了补充领域相关的知识,Chen 等学者^[52-53]将来自多个领域的大规模数据作为先验知识源,从中自动学习先验知识,并构建了领域相关的 must-link 集合和 cannot-link 集合。彭云等学者^[54]从句法分析、词义理解和语境相关等角度,进一步补充了上述 must-link 和 cannot-link 集合,实现了弱监督的 LDA。第二种思路旨在提升主题模型的学习能力。He 等学者^[55]提出 ATAE(Attention-Based Aspect Extraction)模型,借助词向量和注意

力机制来实现基于神经网络的主题模型,效果显著超越了传统 LDA 模型。Liao 等学者^[56]提出使用注意力模型以及变分自编码器(Variational Auto-Encoder, VAE)模型,分别建模局部上下文和全局上下文,进一步改进了基于神经网络的主题建模。

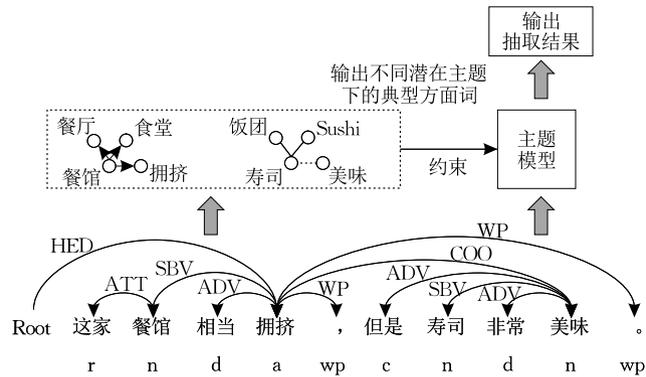


图 4 基于统计的方面词抽取

从以上描述中可以看出,基于手工规则的方法主要依赖词性信息和依存关系,通过构建方面词和上下文词之间的关联规则来帮助实现抽取任务;而基于统计信息的方法则一般借助主题模型挖掘不同潜在主题下的典型方面词。虽然基于手工规则和统计信息的方法不需要人工标注,但规则集合难以制定完备,统计规则难以适应所有场景,因而性能受到限制。近年来领域内的研究重心逐渐转向设计更有效的有监督学习方法。

3.1.2 基于有监督浅层模型的方法

早期针对方面词抽取的有监督方法通常使用手工构建的特征表示单词,再以此为基础训练浅层的有监督学习模型,如图 5 所示。

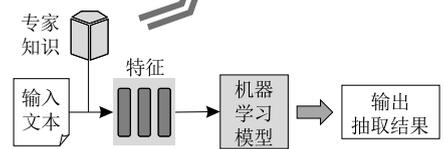


图 5 基于有监督浅层模型的方面词抽取

Jin 等学者^[57]使用了三种特征,分别是词性(Part-of-Speech, POS)、单词从属类别和单词上下文来表示单词。接着,他们采用自举(Bootstrapping)增强方法训练隐马尔可夫模型(Hidden Markov Model, HMM)用于标注单词序列。Mitchell 等学者^[26]同时使用表层特征和语义特征来表示单词。其中,表层特征包括单词位置、是否大写、单词长度、评论长度等,语义特征包括是否为俚语、是否为常用动词词尾、是否为普通名词词尾等。获取特征后,训练条件随机场(Conditional Random Field, CRF)进行标注。NILC_USP^[58]、NRC-Canada^[59]、DLIREC^[60]

和 ECNU^[61] 这些在 SemEval-2014 竞赛榜单上排名前列的模型, 均使用 CRF 进行序列标注, 许多手工构建的特征, 如单词、词性、词组、实体类别、语义角色、知识库内类别、依赖路径、词形态等等都可用于表示单词. Liu 等学者^[62] 提出基于训练数据选择高质量的依赖路径规则的方法, 从而实现了自适应的方面词和观点词抽取. Shu 等学者^[63] 提出将依赖路径抽象为依赖模式, 并将其作为单词特征, 在此基础上利用终身学习增强的 CRF 来提升模型的标注性能.

在词表示学习算法 Word2Vec^[64] 的启发下, 一些工作借助无监督学习算法将以往的具体特征抽象为向量表示, 进一步提升了方面词抽取的性能. Yin 等学者^[65] 首先使用无监督算法学习词向量和依赖路径向量, 再利用条件随机场进行标注. 实验结果表明, 基于这两类向量特征的效果可与传统的基于复杂手工设计特征的方法效果相当.

早期研究广泛采用基于特征工程的方法, 利用词性信息、单词长度、单词位置等多种特征对单词进行表示, 再结合机器学习模型进行方面词抽取. 由于通常需要专家知识进行特征的选取和构建, 该方法引入了额外的人力资源代价.

3.1.3 基于有监督深层模型的方法

随着深度学习的发展, 将词向量和神经网络相结合, 实现一种深层有监督学习模型用于序列标注, 逐渐成为方面词抽取的主流方法. 深度学习采用端到端的学习方式, 可以自动地学习和构建单词特征, 大幅降低了模型设计和应用的开销, 如图 6 所示. 具体而言, 用于方面词抽取的神经网络模型可分为四种: 循环神经网络、递归神经网络、卷积神经网络和 Transformer 结构.

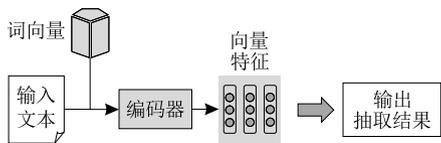


图 6 基于有监督深层模型的方面词抽取

由于循环神经网络可以建模单词上下文依赖, 因此在方面词抽取领域应用最为广泛. Liu 等学者^[66] 提出了第一个将循环神经网络与词向量结合用于方面词抽取的工作, 并尝试了 Elman-type、Jordan-type 和 LSTM-type 三种不同的网络架构, 在正反双向的循环建模机制的基础上, 融合了词性、组块信息等语言学特征. 与以往基于 CRF 模型相比, LSTM-RNN 结构的标注效果显著更优. 后续针对循环神经网络又有一系列的改进. Wang 等学者^[67]

结合了循环神经网络与转移机制, 通过转移动作和状态来进行方面词的预测. Ma 等学者^[68] 利用循环神经网络构建了序列到序列模型, 能够直接生成标注序列. Wei 等学者^[69] 使用了 BiLSTM-CRF 的组合模型, 并利用指针网络修正序列标注过程中的微小边界误差. Yang 等学者^[70] 为 LSTM 增加了成分晶格编码机制, 显式考虑了词组信息, 这种方法能够更好地处理包含多个单词的方面词的情况.

递归神经网络具有建模层次信息的能力, 与依存句法分析的结果相结合可以实现依赖路径上的特征传播. Wang 等学者^[71] 通过 Stanford 依赖解析工具解析评论, 并根据单词间的依赖关系, 在依赖路径上进行词之间的特征传播. Luo 等学者^[72] 进一步将上述单向传播的过程改进为双向传播, 更好地建模了单词特征.

卷积神经网络更适合提取局部特征, 因此广泛用于建模 N-gram 特征, 利用单词在固定大小内的上下文信息对单词进行判断. Xu 等学者^[73] 结合通用词向量和领域词向量, 提出了一种用于方面词抽取的双重词嵌入卷积神经网络. 领域词向量的加入进一步提升了模型效果, 以并行方式运行的卷积网络同时有着高效的优势. Shu 等学者^[74] 在卷积网络中加入了额外的控制模块, 以异步学习的方式加入噪声, 以提升卷积网络的鲁棒性, 避免过拟合有限的训练数据, 从而达到更优的抽取效果.

随着预训练-微调模式的流行, 越来越多基于 Transformer 结构实现大规模预训练模型的工作开始出现. Xu 等学者^[75] 首先从在线评论平台 Yelp 和 Amazon 收集了大量的领域内数据, 再以 BERT 模型为基础, 在这些数据上进行后训练, 将领域知识蕴含于 BERT 中. 最后, 基于 BERT 输出的高层语义特征进行序列标注, 在方面词抽取任务上达到了目前最优的效果.

3.2 [基本任务]方面情感分类

方面情感分类的目的是识别与方面词相关的上下文, 以判断其情感极性. 现有的方面情感分类方法可以分为三类, 分别是基于注意力机制的方法、基于依存句法分析的方法和基于预训练模型的方法, 这些方法根据识别关键上下文的策略不同而区分开来.

3.2.1 基于注意力机制的方法

注意力机制是目前最为普遍的提取关键上下文的方法, 已被广泛用于方面情感分类. 其核心原理是为上下文中的每个单词赋予不同的权重, 从而增强相关词的影响, 并减弱无关词的干扰, 如图 7 所示. 注意力权重是数据驱动的, 由神经网络自主计算确定.

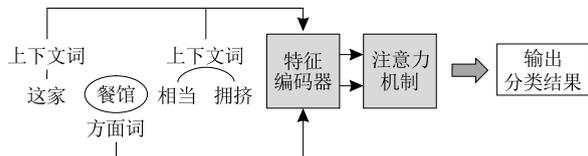


图 7 基于注意力机制的方面情感分类

当前有多种与注意力机制结合的神经网络架构。Tang 等学者^[76]提出 MemNet (Memory Network) 模型, 结合了注意力机制与记忆网络, 在堆叠计算注意力的同时, 会对查询记忆单元也进行不断更新。Wang 等学者^[77]提出 ATAE-LSTM (Attention-based LSTM with Aspect Embedding) 模型, 结合了注意力机制与循环神经网络, 以达到保留与方面词相关的上下文特征的目的。Fan 等学者^[78]结合了注意力机制与卷积神经网络, 并计算了 N-gram 特征的注意力权重。Li 等学者^[79]提出 TNet (Transformation Network) 模型, 将注意力机制与残差网络相结合, 进而实现了针对深层语义的特征变换。Zhao 等学者^[80]提出 SDGCN (Sentiment Dependencies with Graph Convolutional Networks) 模型, 将注意力机制与图神经网络结合, 通过图神经网络显式建模多个方面词情感极性之间的依赖关系。

注意力机制中的权重分数可以使用多种不同的计算方式。Tay 等学者^[81]提出 DyMemNN (Dyadic Memory Network) 模型, 通过张量模块计算注意力权重。Chen 等学者^[82]提出 RAM (Recurrent Attention on Memory) 模型, 先计算多层的注意力权重, 再使用 GRU 建模多层注意力之间的关系。Tay 等学者^[83]提出全息模块以计算注意力权重, 该模块基于快速傅里叶变换实现。Wang 等学者^[84]引入了非线性打分函数到多层注意力计算的过程中。Gu 等学者^[85]显式建模方面词的位置信息, 提出在注意力计算过程中结合绝对位置向量。Hu 等学者^[86]为了使得每个方面词关注的上下文尽量稀疏, 提出为注意力机制增加正则化约束, 使得不同方面词关注的上下文趋于正交。杜成玉等学者^[87]提出在计算注意力权重的过程中, 螺旋式地交替更新方面词和上下文的表示。

注意力机制可以融合不同粒度的信息。Ma 等学者^[88]提出 IAN (Interactive Attention Network) 模型, 在针对上下文使用注意力机制的基础上, 同时对方面词也应用注意力机制。梁斌等学者^[89]考虑到卷积神经网络缺乏对全局信息建模, 进而提出多重

注意力机制以改善这一问题。Fan 等学者^[90]提出 MGAN (Multi-Grained Attention Network) 模型即多粒度注意力网络, 结合粗粒度和细粒度注意力来捕捉方面词和上下文的交互信息。Majumder 等学者^[91]在多层注意力机制的基础上考虑单条评论文本中不同方面词之间的关系。为了帮助建模方面词与上下文之间的交互关系, Zhu 等学者^[92]引入了隐式的方面类别单元。Huang 等学者^[93]提出 AOA (Attention-over-Attention) 模型, 旨在基于细粒度注意力同时建模方面词与上下文的双向交互关系。Wang 等学者^[94]提出一个多层次注意力网络以建模单词级和句子级的注意力权重。曾锋等学者^[95]提出在引入多种额外语义特征的同时, 将单词级注意力和句子级注意力上下堆叠。Zhang 等学者^[96]提出 KG CapsAN (Knowledge Guided Capsule Attention Network) 模型, 在胶囊注意力网络的计算中引入句法和 N-gram 特征作为指导。

除此之外, 门控机制也可以看作是一种未经过归一化处理的注意力机制的实现方式, 也可以用于自适应地缩放上下文特征。Zhang 等学者^[97]在循环神经网络 RNN 的基础上加入门控机制, 以汇聚方面词的左右上下文特征。Xue 等学者^[98]提出 GCAE (Gated Convolutional network with Aspect Embedding) 模型, 该模型基于卷积神经网络与门控机制实现, 可以高效理解给定方面词的情感信息。Huang 等学者^[99]利用参数化的过滤器与门控来将方面词信息融入到卷积神经网络中。冯超等学者^[100]为了让方面词和上下文之间的交互建模更加深入, 同时利用了门控机制与层次注意力机制。宋威等学者^[101]则基于门控机制实现两步式的蒸馏, 旨在过滤出与方面词相关联的上下文。

3.2.2 基于依存句法分析的方法

在神经网络模型中融合先验的文本语法结构, 可为方面情感分类任务提供方面词与上下文之间的关联信息, 有助于提升性能。在实现上, 先行使用依存句法分析工具 (如 Stanford CoreNLP 和 Spacy 工具包) 获取依存句法信息, 再将其注入模型的计算过程, 可以帮助模型更加高效准确地判别与方面词相关的关键上下文, 如图 8 所示。

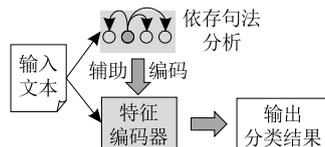


图 8 基于依存句法的方面情感分类

Dong 等学者^[27]采用自适应的递归神经网络 RNN,将上下文中关联方面词的情感特征和句法特征传播到方面词的表示中以帮助情感分类。He 等学者^[102]通过注意力机制,参考句法依赖路径距离来融合上下文单词与方面词。Sun 等学者^[103]提出 CDT(Convolution over Dependency Tree)模型,基于图卷积神经网络在依赖树上建模句子的语法特征。Yang 等学者^[104]利用依赖解析的结果与预定义规则以计算注意力权重。Phan 等学者^[105]将依赖解析之后的词性信息与依赖关系信息融入到了单词向量中,并在注意力机制中融合了句法成分之间的关联。Zheng 等学者^[106]在依赖路径上引入了随机游走算法,强化了上下文中的关联单词特征。Wang 等学者^[107]提出 RGAT(Relational Graph Attention Network)模型,将传统的依赖分析树转化为了一个方面词主导的依赖树结构,并使用关系图注意力网络来编码新的依赖树结构。王光第等学者^[108]深度融合了基于依存句法树的图卷积网络和基于上下文的记忆网络。

3.2.3 基于预训练语言模型的方法

伴随着预训练-微调模式的流行,利用大规模预训练语言模型中隐含的语义知识进行方面情感分类已经成为主流,如图 9 所示。在预训练语言模型中,注意力机制被广泛应用。然而,预训练过程中所学习的外部知识是实现优良性能的基础和前提,因此这里需要对其进行单独讨论。

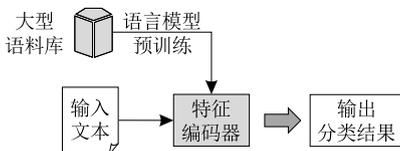


图 9 基于预训练语言模型的方面情感分类

Xu 等学者^[75]提出 BERT-PT 模型,利用掩码还原任务,在大规模的领域内语料上对 BERT 进行后训练。然后,使用经过后训练的 BERT 模型在方面级数据集上进行微调。相较于基于词向量的模型,该方法的效果提升显著。Rietzeler 等学者^[109]提出 BERT-ADA 模型,该模型结合了自监督的语言模型微调和有监督的任务微调。Xu 等学者^[110]在前人的基础上进行了一系列实验,探究了 BERT-PT 模型在进行方面情感分类时内部注意力权重的分布情况。为了更有效地捕捉单词间的关系,Dai 等学者^[111]利用预训练语言模型诱导出的树结构代替使用传统工具所解析的依存句法树。为了让模型更加深入地分辨隐式情感信息,Li 等学者^[112]提出一种

结合对比学习的有监督预训练方法,以对齐隐式情感与显式情感。Seoh 等学者^[113]将方面情感分类形式化为自然语言推理任务,以结合 Prompt 来诱导预训练语言模型蕴含的内部知识。Yang 等学者^[114]提出 LSA(Local Sentiment Aggregating)模型,在结合局部情感聚合机制的同时,使用模式驱动的情感依赖代替传统的句法依赖,在方面情感分类上取得了较好的性能。预训练-微调方式的广泛应用,显著提升了方面情感分类的性能,目前已成为通用训练方法。

上述提到的方法主要是基于注意力机制、依存句法分析和预训练语言模型三类。此外,方面情感分类任务还可以基于分段建模^[115-118]、概率图模型^[119]、胶囊网络^[120-121]和逻辑规则^[122]等方法实现。本文在此不再详细介绍,具体信息可参考相关论文。

3.3 [基本任务]方面-情感对抽取

方面-情感对抽取任务旨在同时抽取出评论文本中所有的方面词及其相对应的情感极性,是方面词抽取任务与方面情感分类任务的组合。

一个最简单的想法就是将前面两节所介绍的方面词抽取方法与方面情感分类方法进行组合,即先进行方面词抽取,再对所预测的方面词进行方面情感分类,如图 10 所示,这便是流水线模型^[123]。但是,这类模型存在两个严峻的问题,一是无法建模方面词抽取与方面情感分类这两个任务之间的交互关系,而该关系可以有效提升方面-情感对抽取的整体性能,二是存在级联错误传播问题,即方面词抽取任务的误差会直接导致下游的方面情感分类任务的性能下降。因此,在一般情况下,流水线模型对于方面-情感对的抽取能力相对受限。

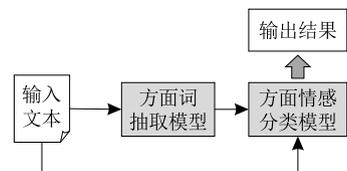


图 10 基于流水线模型的方面-情感对抽取

为了应对这一挑战,当前领域内通常使用端到端的学习方式来同时建模方面词抽取和方面情感分类两个任务。针对不同的任务结合方式,现有的方面-情感对抽取方法可以分为三类:基于统一标签的方法、基于联合学习的方法和基于文本生成的方法。

3.3.1 基于统一标签的方法

现有的方法通常采用序列标注的方式来解决方面词抽取问题,常见的标注方案有 BIO 标注。同时,方面情感分类的标签大多只分为四类,分别为积极、

中立、消极和矛盾. 因此, 一种最直接的方式就是融合两类标签, 即在 BIO 标注的基础上加入对情感极性的标注. 如图 11 所示, “B+ 积极”代表该单词为方面词的第一个单词, 并且其情感极性为积极.

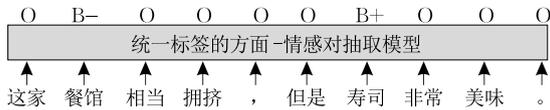


图 11 基于统一标签的方面-情感对抽取

Zhang 等学者^[124]首次在方面级情感分析任务中引入统一标签, 并设计了一个前馈神经网络来预测统一标签. Li 等学者^[125]使用了层次化的预测方案, 第一层用于预测方面词的边界, 即 BIO 边界, 第二层用于预测最后的统一标签, 其使用的模型基于 LSTM 实现. 在此基础上, Peng 等学者^[22]额外引入了方面词与观点词的配对. Chen 等学者^[126]关注到图卷积过程中左侧、自身、右侧的不同影响, 因此利用不同的转换矩阵先进行特征传播, 最后预测统一标签.

使用统一标签来同时处理方面词抽取和方面情感分类两个子任务, 是最直接的方面-情感对抽取方法. 但该方法会导致两个子任务共享大部分可学习特征, 缺乏对于私有特征的建模, 这可能会造成潜在的干扰, 并阻碍模型建模子任务之间的交互关系.

3.3.2 基于联合学习的方法

联合学习是一种实用的建模方法, 其思路是在一个模型内部分别建模方面词抽取和方面情感分类任务, 再将两个任务的结果融合后作为输出, 从而实现方面-情感对抽取的目标, 其过程如图 12 所示.

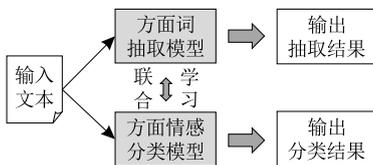


图 12 基于联合学习的方面-情感对抽取

Wang 等学者^[127]提出了一种使用注意力机制学习方面词和情感极性之间关系表示的方法. 该方法联合训练两个子任务, 实现了方面词和情感极性之间的关联性建模. Luo 等学者^[128]先利用残差门控单元来建模方面和情感, 再使用交叉共享单元在两个任务之间传递信息, 实现了任务之间的信息共享. 为了进行领域知识迁移, He 等学者^[129]提出 IMN (Interactive Multi-task learning Network) 模型, 其核心在于利用文档级领域分类任务来辅助方面词抽

取, 并利用文档级情感分类任务辅助方面情感分类. Chen 等学者^[130]考虑到现有方法没有充分建模方面词抽取、观点词抽取与方面情感分类之间的关系, 提出了 RACL (Relation-Aware Collaborative Learning) 模型, 其使用多层堆叠结构完成对关系的深层传播. Wang 等学者^[131]提出 LCM (Label Correction Model) 模型, 通过建模方面情感分类任务和方面词抽取之间的交互关系以完成标签校正. Liu 等学者^[132]提出一种动态异质图模型来完成方面词抽取任务与方面情感分类任务之间的交互, 其核心做法在于将评论中单词和情感标签同时放入图内, 在训练中完成特征传播与更新. Mao 等学者^[133]提出 Dual-MRC (Dual Machine Reading Comprehension) 模型, 以多轮问答的形式来建模方面情感对抽取, 分别询问所有出现的方面词以及其所对应的情感极性.

相较于统一标签的学习方式, 联合学习同时建模两个子任务, 更加符合该任务的需求. 在该方案中, 子任务可以建模私有特征, 这为建模任务之间的交互关联打下了基础, 因而取得了较为优异的性能.

3.3.3 基于文本生成的方法

伴随生成式预训练模型 (如 T5^[134] 和 BART^[135]) 的不断进步, 将抽取任务转化为文本生成已成为一种切实可行的方式, 其可以充分利用预训练模型的知识. 研究人员也开始将方面-情感对抽取任务看作文本生成任务来解决, 图 13 是一个基本的解决流程.

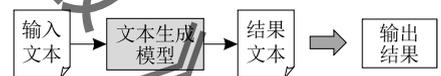


图 13 基于文本生成的方面-情感对抽取

Yan 等学者^[136]将指针网络的思想结合进生成式模型 BART, 利用指针生成的方式统一了方面级情感分析的不同子任务. 在此基础上, Li 等学者^[137]提出了 Senti-Prompt, 其核心在于额外加入了一个自监督的关系判别任务以提示方面词与观点词. 为了考虑标签词的语义知识而非将其转化为特定的标签索引, Zhang 等学者^[138]提出了两种不同的文本生成模版, 一是在原有的输入文本上插入标签词序列 (GAS-Annotation), 二是构建仅包含标签词的序列 (GAS-Extraction), 通过这两种预定义模版可将抽取任务转化为文本生成任务, 并以此训练生成式模型 T5. 进一步, Zhang 等学者^[25]提出 Paraphrase 模型, 设计了一个更加贴合自然语言的文本生成模版, 将方面情感分析任务转化为语义复述任务, 同样采

用 T5 模型进行训练。

虽然基于文本生成的方法在严格意义上仍属于联合学习的范畴,但其与传统联合学习方法所使用的序列标注形式在任务定义上存在明显区别。该方法使用自然语言的形式作为预训练模型的学习目标,其中目标文本要以特定的格式囊括方面词和情感极性要素,从而利用这些要素的语义信息辅助模型理解。不同于传统方法,文本生成消除了预训练和微调阶段的任务形式差异,从而更加有效地挖掘了预训练模型所蕴含的知识,统一了方面级情感分析的众多子任务,成为了近期热门的研究方向。

3.4 [扩展任务]观点词抽取

观点词抽取任务的目标是抽取出输入文本中所有的观点词,是一种较为简单的扩展任务。在方面级情感分析中,观点词一般依赖于方面词,是用户对于方面词的主观描述。大多数的研究不局限于完成单一的观点词抽取任务,而是将对方面词和观点词两者的抽取合并处理。注意,这类研究与方面-观点抽取任务不同,不需要抽取出方面词与观点词之间的配对关系。

观点词抽取任务大多采用序列标注模型,即采用一套标签集合来标注文本中的每一个单词^[66]。序列标注模型中最常使用的是 BIO 标注方案,其中 B 表示目标的第一个单词, I 表示目标中除第一个单词以外的其他单词, O 表示非目标单词。近年来,针对该任务的模型主要从学习方面词与观点词之间的交互关系出发,以同时提升对两者的抽取效果。Wang 等学者^[71]利用基于句法依存树的递归神经网络来学习单词的高维特征表示,再利用条件随机场进行标签分类。Wang 等学者^[15]提出耦合的多层注意力机制用于学习方面词与观点词之间的交互关系。Yu 等学者^[139]提出一种多任务学习框架来捕捉方面词与观点词之间的交互关系,使其适用于数据稀疏的场景。

3.5 [扩展任务]方面观点抽取

方面观点抽取任务的目标是在给定输入文本与某一个方面词的情况下,抽取出与方面词所对应的所有观点词。

Fan 等学者^[12]首次提出了方面观点抽取任务,称作 Target-oriented Opinion Words Extraction,并设计了一个序列标注模型来解决该任务。受限于方面观点抽取任务的数据稀缺,Wu 等学者^[140]提出了一种隐式观点知识迁移网络,将观点知识从丰富的评论情感分类数据集中迁移到低资源的方面观点抽

取任务。Veysseh 等学者^[141]提出利用单词之间的依存关系来优化单词的表示,以提高模型在方面观点抽取任务上的性能。Mensah 等学者^[142]测试了若干个常用的文本编码器,并考虑了位置信息与词性信息。结果显示,BiLSTM 可以有效地利用位置信息,在此基础上再加入图卷积网络融合语法信息只能获得较小的收益。Li 等学者^[143]发现已有方面观点抽取模型所用的数据集中,每一个方面词都至少对应一个观点词,而在真实的场景中,方面词不一定有所对应的观点词。因此作者相应扩充了数据集,并对现有方法进行了重新训练,发现模型的性能得到了进一步增强。

3.6 [扩展任务]类别识别

类别识别的目标是给定文本作为输入,输出所有的方面类别,其中方面类别来自于预定义的集合。

Zhou 等学者^[144]提出了一种半监督的词向量学习算法来捕获单词之间的语义关系,并且在训练分类器之前,利用两种不同的神经网络生成了更深层更多元的特征以增强类别识别。Tulkens 等学者^[13]提出一种无监督学习的方法用于类别识别,只需要领域内词向量及词性标注即可完成。Shi 等学者^[145]提出一种自监督的对比学习框架用于无监督的类别识别任务。Liu 等学者^[16]将类别识别任务转换为文本生成任务,例如将方面类别 *food* 填入模板“*The food category is discussed*”作为备选文本计算得分,再选择得分高的类别作为输出。

3.7 [扩展任务]类别情感分类

类别情感分类任务是给定文本和某一方面类别作为输入,对方面类别的情感极性进行分类。

Ruder 等学者^[146]通过一个多层次的双向长短期记忆网络来建模评论文间的内部依赖关系。Sun 等学者^[147]对于每一个方面类别构造了一个辅助句子,将该任务转换成了句子对配对问题。Zhu 等学者^[14]考虑方面类别与评论文本之间的交互关系,并设计一种双路记忆网络将方面感知学习和情感分析框架结合起来。Liu 等学者^[16]将该任务转换为文本生成任务,例如,将方面类别 *food* 加入模板“*The sentiment polarity of food is positive*”之后,交由模型计算可能性得分。

3.8 [扩展任务]方面-观点对抽取

方面-观点对抽取任务的目标是抽取出输入文本中所有的方面-观点对。相较于独立的方面词抽取和观点词抽取,该任务要求在抽取方面词与观点词的同时,预测它们之间的配对关系。

Chen 等学者^[148]提出了一个同步双通道循环神经网络模型,利用双通道结构同时抽取目标实体(方面词与观点词)及其相互关系,再利用一个同步单元让两个通道内的信息产生交互. Zhao 等学者^[149]提出了一种基于片段的多任务学习框架,其主要思想是利用一个片段枚举器来枚举文本中可能为方面词或者观点词的片段,每一个片段的表示会参与计算实体分数与关系分数以用于预测. Gao 等学者^[150]提出了一个问题驱动的片段标注框架,首先利用片段标注模型抽取所有的方面词,再对于每一个方面词生成一个问题,形如“*What do you think about the <aspect>?*”,以抽取对应观点词. Wu 等学者^[151]引入了丰富句法知识来增加方面-观点对的抽取效果. 除此以外,生成模型也可以应用于该任务^[136-138].

3.9 [扩展任务] 类别-情感对抽取

类别-情感对抽取任务的目标是识别输入文本中所包含的方面类别并判断其情感极性.

针对该任务最为直观的做法就是流水线模型,将其拆解为类别识别与类别情感分类两个任务,但是这样容易导致错误的累积传播. Schmitt 等学者^[152]在一个统一的框架中联合训练两个任务. Cai 等学者^[153]发现先前的模型难以学习方面类别之间以及方面类别与情感极性之间的关系,因此提出一种多层次图神经网络,同时识别出文本中的所有方面类别以及相对应的情感极性.

3.10 [扩展任务] 方面-观点-情感三元组抽取

方面-观点-情感三元组抽取任务的目标在于一步抽取所有的(方面词,观点词,情感极性)三元组,相较于方面-情感对抽取,该任务在其基础上进一步抽取出了观点词,可以清晰地表明为什么方面词会表现出其对应的情感极性.

Peng 等学者^[22]首次提出了方面-观点-情感三元组抽取任务,并利用一个两阶段的序列标注模型来流水线式地抽取三元组. 第一阶段首先抽取所有的方面-情感对与观点词,第二阶段配对方面词与观点词. 为了避免流水线模型产生的错误传播, Xu 等学者^[37]提出了位置感知的标注框架,可以一次性预测完整的三元组. Wu 等学者^[20]提出了基于表格标注的三元组抽取模型. Chen 等学者^[154-155]进一步考虑单词之间的语言学特征,利用结合了语法知识的图卷积模块捕捉词语之间的联系,以进行方面-观点-情感三元组抽取. Zhang 等学者^[23]提出了一个多任务学习框架,对方面词抽取与观点词抽取进行联合学习. Mao 等学者^[133]将三元组抽取任务转换为

阅读理解任务,首先在输入文本前面拼接上询问方面的问题,以抽取所有的方面词,再根据每一个方面词来设计问题以抽取其对应的所有观点词,并判断方面词所属的情感极性. Chen 等学者^[17]提出了双向阅读理解模型,对方面词和观点词进行交互抽取. Xu 等学者^[21]考虑了方面词与观点词整个片段的语义信息,提出了基于片段交互的三元组抽取模型.

随着生成模型在信息抽取领域的快速应用, Mukherjee 等学者^[19]提出一个基于指针网络的序列生成模型来解码三元组. Yan 等学者^[136]将三元组抽取转换为序列索引生成任务,即生成每个要素的起止位置. Li 等学者^[137]在此基础上,对于预训练语言模型的编码器加入了提示信息,以更好地诱导模型生成. Zhang 等学者^[138]同样将三元组抽取转换为文本生成任务,并提出了两种生成模板,注释样式与抽取样式. Zhang 等学者^[25]设计了一个包含语义的自然语言模板来帮助模型更好地理解不同要素之间的关系. Fei 等学者^[156]将三元组抽取任务视为无序的三元组集合预测任务,提出一种非自回归的生成模型来生成所有的三元组. Jian 等学者^[18]使用强化学习来进行三元组抽取.

3.11 [扩展任务] 方面-类别-情感三元组抽取

方面-类别-情感三元组抽取任务旨在抽取所有的(方面词,方面类别,情感极性)三元组.

Wu 等学者^[15]利用图卷积网络捕捉三种要素之间的内部依赖关系. Zhang 等学者^[41]将该任务转换为文本生成任务,设计了两套模板来将三元组映射为文本输出. Zhang 等学者^[25]将三元组集合映射为语义通顺的自然语言表达后再进行生成.

3.12 [扩展任务] 方面-观点-类别-情感四元组抽取

方面-观点-类别-情感四元组抽取任务的目标在于一步抽取方面级情感分析中所有的要素,包括方面词、观点词、方面类别和情感极性,是目前最为完备但也最困难的方面级情感分析任务. 由于尚处于起步阶段,相关工作仍然较少.

Cai 等学者^[24]首次提出了方面-观点-类别-情感四元组抽取任务,发布了两个四元组抽取数据集,并设计了流水线模型完成该任务. Zhang 等学者^[25]将四元组抽取任务转换为句子复述任务,以便于模型更好地理解不同要素之间的语义信息.

表 2 汇总整理了本节所涉及的所有相关文献,以便参考查阅.

表 2 方面级情感分析中基本任务和扩展任务的相关工作

定位	序号	名称	输入	输出	类型
基本任务	1	方面词抽取 [2,42-56] [26,57-63,65] [66-75]	文本	a_1, a_2	抽取
	2	方面情感分类 [76-101] [27,102-108] [75,109-114]	文本 + a_1	s_1	分类
			文本 + a_2	s_2	
	3	方面-情感对抽取 [22,124-126] [127-133] [25,136-138]	文本	$(a_1, s_1),$ (a_2, s_2)	抽取 + 分类
	4	观点词抽取 [15,66,71,139]	文本	o_1, o_2	抽取
	5	方面观点抽取 [12,140-143]	文本 + a_1	o_1	抽取
文本 + a_2			o_2		
扩展任务	6	类别识别 [13,16,144,145]	文本	c_1, c_2	分类
	7	类别情感分类 [14,16,146,147]	文本 + c_1	s_1	分类
			文本 + c_2	s_2	
	8	方面-观点对抽取 [136-138,148,151]	文本	$(a_1, o_1),$ (a_2, o_2)	抽取
	9	类别-情感对抽取 [152,153]	文本	$(c_1, s_1),$ (c_2, s_2)	分类 + 配对
	10	方面-观点-情感三元组抽取 [17-23,25,37,133,136-138,154,156]	文本	$(a_1, o_1, s_1),$ (a_2, o_2, s_2)	抽取 + 分类
	11	方面-类别-情感三元组抽取 [25,141,157]	文本	$(a_1, c_1, s_1),$ (a_2, c_2, s_2)	抽取 + 分类
	12	方面-观点-类别-情感四元组抽取 [24,25]	文本	$(a_1, o_1, c_1, s_1),$	抽取 + 分类
(a_2, o_2, c_2, s_2)					

4 面向低资源场景的改进策略

本节首先分析方面级情感分析在低资源场景下面临的挑战,随后介绍面向低资源场景方面级情感分析的两类改进策略,最后进行小结。

4.1 低资源场景下的问题与挑战

在自然语言处理领域,文本数据对整体类别和情感的粗粒度标注多,对方面、观点、目标情感的细粒度标注少,导致方面级情感分析数据集的语种、领域和规模均受限。本文着重分析低资源场景下的样本匮乏对于方面级情感分析模型带来的问题与挑战。

在模型层面,低资源限制了方面级情感分析的特征质量和学习过程。一方面,对于方面词抽取、观点词抽取等抽取型子任务而言,模型需要对目标词提取高质量的特征以完成判断。在低资源下,样本不足直接导致大部分目标词均为长尾词,即在训练集中出现的频次极低。受此影响,即使模型的特征抽取模块设计得再为精细,也难以获得长尾词的有效特

征。最终,模型能成功抽取的目标词仅占词表的极小部分,难以应用至现实场景。例如,在餐馆领域中,现有模型对有限的高频方面词(如“寿司”)识别成功率很高,而对大量的低频方面词(如“贝尔普里”)识别成功率则很低。另一方面,对于方面情感分类、类别情感分类等分类型子任务而言,模型需要仔细分析上下文与目标词之间的依赖关系,以理清用户的情感倾向。在低资源下,样本不足直接限制了模型所能学习到的情感知识总量,使其无法理解复杂语境中的语义信息。例如,电脑领域评论中常见的观点词“很快”,在修饰方面词“硬盘”时体现积极情感,表达运行高效;而在修饰方面词“电池”时则可能体现消极情感,表达电量续航较差。在样本不足时,模型则难以理解此类依赖关系,容易做出错误判断。除了上述对于特定子任务的负面影响,低资源场景还进一步削弱了模型对于任务的理解程度,降低了其从样本中学习子任务间关联的能力。因此,在低资源场景下,如何提升模型的学习能力,是现存挑战之一。

在数据层面而言,低资源导致的样本匮乏问题

促使研究者们思考如何获得更多的方面级标注数据。对于粗粒度的文档级/语句级任务而言,已有众多工作基于数据增强和领域适应等手段扩充训练数据。然而,对于方面级情感分析而言,标注信息往往与文本中的特定词绑定,使得随机删除、随机替换、领域对抗等经典方法难以发挥作用,所产生的合成样本质量无法达到预期。因此,在低资源场景下,如何生成或迁移高质量的方面级标注样本,是现存挑战之一。

4.2 模型层面的改进策略

从模型的角度看,监督学习可以在有适当标注数据的训练下,实现比无监督学习更好的性能。但是,由于方面级情感分析数据的标注成本高,导致了低资源的数据环境,使得监督学习模型的能力受到极大限制。为了克服该问题,现有研究试图从不同的角度改善模型在有限样本下的学习过程,已有方法主要包括基于交互建模的方法和基于知识注入的方法两个类别。

4.2.1 基于交互建模的方法

方面级情感分析任务涉及多种要素之间的交互关系,例如观点词和方面词之间的直接修饰关系。多项现有研究利用这些子任务之间的交互关系,使得某些子任务的特征成为其它子任务的辅助判据,以提升方面级情感分析模型在低资源场景下的性能表现。这些方法旨在通过利用不同子任务之间的关联性,提高模型对于数据的利用效率,从而提高模型的泛化性和准确性。

在先前的研究中,Li等学者^[158-159]、程梦等学者^[160]和Wang等学者^[15]使用远程监督或人工标注的观点词,在模型训练过程中构建观点词与方面词之间的关联,以缓解有限样本所导致的方面词特征学习困难的问题。He等学者^[161]和Luo等学者^[128]则在多任务学习框架下,显式地考察了方面情感分类和方面词抽取任务之间的关联,并采用额外的观点词辅助情感极性的判别。这些研究使用基于交互建模的方法来实现子任务之间的相互辅助,能够有效提高各子任务和整体任务的性能。

4.2.2 基于知识注入的方法

在低资源场景下,由于只有少量标注样本可用于训练,因此方面级情感分析模型的学习能力受到了限制。为了克服这一问题,现有研究提出了多种基于知识注入的改进方法,旨在为方面级情感分析模型提供更多的情感知识和语义模式,以弥补训练数据的不足。

第一种方法是利用大规模的领域相关粗粒度文档级语料库中的相关知识,以帮助细粒度的方面级任务。例如,为了增强方面词的特征表示,Xu等学者^[73]在模型中引入了在领域内语料上预训练的词向量。为了促进模型对情感语义的判断,He等学者^[129]采用文档级数据预训练LSTM模型,然后将文档级和方面级数据一起输入,进行多任务联合学习。为了将预训练的文档情感分类模型应用于方面情感分类,Angelidis等学者^[162]提出了多示例学习方法,并利用手工方面词种子实现了弱监督的方面词抽取。为了提高对上下文的理解,Chen等学者^[120]利用文档级数据中的情感语义知识,通过方面路由和动态路由的设计,构建出一种迁移胶囊网络。此外,Chen等学者^[163]还使用领域相关的语言模型生成软模板信息,以增强模型对于长尾词的理解。为了提取关键上下文信息,Zhao等学者^[164]采用文档级数据训练注意力模型参数,然后使用固定的文档级注意力结果来指导方面级模型。

第二种方法是利用专门构建或者已有的情感知识图谱,为模型增加情感维度的表示。例如,Ma等学者^[165]将SentiNet中的情感知识与LSTM模型相结合,以此来提高模型在低资源场景下理解情感上下文的能力。Zhao等学者^[166]通过人工构建的领域内情感知识图谱,有效加强了BERT对情感信息的理解能力。

第三种方法是以预训练语言模型作为基础模型,充分利用其在自监督学习过程中所捕获的丰富语义信息。例如,Mao等学者^[133]和Chen等学者^[17]将方面级情感分析任务转换为问答(Question Answering,QA)形式,这有助于利用预训练语言模型BERT所学习的语义知识在低资源场景下进行任务处理。为了利用强大的序列到序列(Sequence-to-Sequence,Seq2Seq)语言模型(如BART和T5)中的丰富语义信息,Yan等学者^[136]和Zhang等学者^[138]将方面级情感分析任务转化为文本生成(Text Generation)任务。为了在低资源场景下更加有效地引导预训练语言模型进行生成,Li等学者^[137]和Zhang等学者^[25]分别在预训练语言模型的编码端和解码端加入适当的提示(Prompt)信息。

总体而言,外部知识注入是一种提升低资源场景下方面级情感分析性能的有效手段。

4.3 数据层面的改进策略

从数据的角度看,生成或迁移标注信息,是缓解低资源场景下标注样本不足问题的一种可行解决方

案,现有的相关工作主要可分为数据增强和领域适应两个类别.

4.3.1 基于数据增强的方法

数据增强能够有效提升数据总量及多样性, Li 等学者^[167]总结了自然语言处理任务领域常用的数据增强方法. 与文本分类等粗粒度的任务不同, 细粒度的方面级情感分析需要进行细粒度的数据增强, 不能通过排序/增加/删除单词等简单的方法来实现. 为了解决这个问题, 针对细粒度任务的数据增强方法主要通过以下三种方式来实现.

第一种方法是增加额外约束来改进粗粒度的数据增强方法, 以达到生成符合要求的细粒度伪样本的目的. Dai 等学者^[168]提出了多种词级扰动方法, 如标签单词替换、同义词替换、提及替换和段内乱序等, 以生成用于序列标注的细粒度伪样本. Chen 等学者^[169]则使用依赖路径和词性标签来约束样本生成过程, 以提高伪样本的质量.

第二种方法是利用模型为未标注数据添加伪标签, 以直接扩充数据量. 例如 Wang 等学者^[170]结合课程学习和自监督学习方法, 为大规模未标注样本添加伪标签后, 作为额外样本进行训练.

第三种方法是利用预训练语言模型来引导生成符合语义和标注要求的伪样本. 例如 Ding 等学者^[171]将已有的样本标签插入到原始的单词序列中, 形成组合序列对语言模型进行训练. 接着, 固定语言模型并对其进行采样, 进一步生成新的组合序列, 并最终将其还原为单词序列及其所对应的标签序列. Li 等学者^[172]使用 MASS 这一预训练模型进行受控的序列数据增强, 在保持方面词不变的条件下替换上下文, 以生成伪样本. Hsu 等学者^[173]首先通过选择性扰动掩码来对文本进行处理, 然后用预训练语言模型进行恢复, 以生成语义相似且方面词和情感极性不变的伪样本.

上述基于额外约束和伪标签的方法成本较低, 但是生成的伪样本质量较差. 基于语言模型生成的伪样本质量较高, 但会引入较高的计算代价. 因此, 针对方面级情感分析的数据增强方法还需要进一步探索效率和性能的平衡.

4.3.2 基于领域适应的方法

不同领域内的方面级情感分析任务的标注数据量不同, 但是直接跨领域应用标注样本难以达到可用的性能, 其原因在于源领域和目标领域存在知识差异. Ramponi 等学者^[174]总结了自然语言处理领域中常用的领域适应方法. 因为方面级情感分析任

务的关注点是方面词和对应的情感信息, 而大部分方面词是领域特定的, 所以常用的只迁移领域通用枢纽词的领域适应方法难以适用于方面级情感分析任务. 为了解决这个问题, 现有针对细粒度任务的领域适应方法主要采用三种方式实现.

第一种方法是利用领域通用的特征来迁移领域特有的方面词及其情感信息. 例如, Jakob 等学者^[175]使用词性和依赖路径等不受领域影响的特征, 训练了跨领域的条件随机场来标注方面词. Li 等学者^[176]预先确定一些高置信度的主题词和情感, 并使用关系自适应的自举方法在目标域内探测观点词和方面词. Wang 等学者^[177-178]充分利用了依存句法在不同领域之间的相似性, 并在依赖路径上采用递归神经网络传播特征, 训练具备跨领域能力的序列标注器. Wang 等学者^[179]通过学习领域固定的摘要向量, 在不同领域的的数据上探测观点词和方面词. Pereg 等学者^[180]将基于句法的跨领域注意力加入到预训练语言模型中, 有效提升 BERT 跨领域特征建模的能力. 赵光耀等学者^[181]将样本的特征空间划分为领域特有和领域共有两个子空间, 并协同训练两个子空间分类器用于跨领域方面情感分类. Yang 等学者^[182]在情感向量中添加显式的领域信息, 促使方面情感分类器有效适应多个不同领域. Chen 等学者^[183]基于语法角色实现了能够辅助方面词迁移的语义桥和语法桥. Chen 等学者^[184]利用多种领域通用信息检索可迁移原型, 为领域特定词补充可迁移特征, 以在不同领域间有效迁移其标注信息.

第二种方法是通过为源域词设置不同的自适应权重, 以动态地提高或降低它们的影响. 例如, Li 等学者^[185]使用对抗学习来自适应地降低领域特有词的特征权重, 并提高领域通用词的特征权重, 以减小目标域和源域之间的差异. Gong 等学者^[186]提出了一种特征级适应的方法, 在学习过程中倾向于提取领域不变的特征, 并且还同时采用了样本级适应的方法, 以提高领域不变样本(单词)的权重. Chen 等学者^[187]设计了一种互信息最大化组件, 能够有效提升现有领域适应方法的性能.

第三种方法是为目标域的无标注样本打上伪标签, 再将其加入对模型的训练过程中. 例如, Ding 等学者^[188]使用规则在目标域和源域同时创建伪样本, 然后使用这些伪样本共同训练 LSTM 序列标注器, 从而减少领域差异. Zhou 等学者^[189]则使用基于滑动窗口的平均教师方法, 为目标域无标注

数据打上高置信度的标签,以辅助训练并降低领域差异.

4.4 低资源下改进策略小结

为了解决低资源场景对方面级情感分析带来的问题与挑战,研究者们从模型和数据两个层面,提出了众多的改进策略.在模型层面,现有方法主要基于交互建模和知识注入,改进方面级情感分析的特征质量和学习过程,其目标在于减轻低资源下样本不

足对于模型的负面影响.在数据层面,现有方法主要基于数据增强和领域适应,扩充有限的方面级标注样本,其目标在于增广更多的高质量训练数据.总体而言,两类改进策略的动机互补,可相互促进,实现在更充足的标注样本上训练更可靠鲁棒的机器学习模型,从而共同提高方面级情感分析在低资源场景下的性能表现.表 3 中汇总整理了本节所涉及的相关文献,以便参考查阅.

表 3 低资源场景方面级情感分析的改进策略

定位	方法	思路	文献
模型 层面	交互建模	利用子任务间存在的交互关系,使某一子任务的特征成为其他子任务的辅助判据,从而提升方面级情感分析模型在低资源场景下的整体性能表现.	[15,128,158-161]
	知识注入	在低资源场景下为方面级情感分析模型补充更多的情感知识和语义模式,缓解训练样本不足对于模型学习能力的限制.	[17,25,73,120,129,133,136-138,162-166]
数据 层面	数据增强	基于细粒度的数据增强,生成额外的方面级标注样本,补充至训练集中,缓解样本匮乏问题.	[168-173]
	领域适应	基于细粒度的领域适应,从标注丰富的源域迁移方面级标注信息至标注匮乏的目标域,降低对于目标域训练数据的需求.	[175-189]

5 未来展望

本文综述了方面级情感分析的常用方法,以及现有面向低资源场景的改进策略.实际上,细粒度情感分析仍存在众多待深入研究的任务和场景,值得进一步探索.以下从多个角度对未来工作进行展望.

第一,如何令方面级情感分析适用于复杂语义结构.当前的方面级情感分析方法主要针对较短的陈述型文本,其语义结构相对简单.然而在实际应用中,存在各种语义结构复杂的场景.例如在社交媒体和购物平台中,用户会基于问答对话相互交流,情感信息在对话者之间流动^[41,190-192].又如新闻网站和评测文章中,对特定对象的描述通常是篇章级文本,其中包含对各个方面的详细评价^[193-201].针对复杂语义结构,方面级情感分析模型理应动态地捕捉情感的发展和转变,并在较大范围的语境中有效提取关键信息.因此,一种潜在的方向是研究复杂场景下的方面级情感分析方法,实现更准确全面的信息获取.

第二,如何使方面级情感分析与大规模语言模型深度结合.现有研究从模型和数据层面提出了一系列改进策略,但模型在低资源场景下的性能仍难以接近充分监督学习的水平.随着大规模预训练模型的飞速发展,众多研究已证明大模型中蕴含着大量的世界常识和语义知识,在 Few-Shot 乃至 Zero-Shot 场景下均能发挥优良性能^[202-204].将方面级情

感分析与大规模预训练模型深度结合,一方面可以借助大模型的强大能力,减少对于标注数据的过度依赖;另一方面可以引导大模型的知识外放,快速经济地合成更多高质量训练样本^[205-206].通过对模型和数据的双重增强,有望大幅提升低资源场景下的方面级情感分析性能.因此,一种潜在的方向是研究提示学习等技术,令大模型充分赋能方面级情感分析任务.

第三,如何从多种模态中获取多源信息辅助方面级情感分析.在众多应用平台中,用户发言不仅包含文本,还对应着图像、视频、语音等多种模态.近期有研究针对多模态下的方面级情感分析进行了探索^[207],发现同时建模文本和图像能得到显著的性能提升.显然,不同模态之间可能存在一定的交互关系,例如文本中的情感信息可能与图像中的表情等视觉特征相呼应.通过对模态内和模态间信息进行建模,可以更好地挖掘不同模态之间的关联,从而降低模型对于标注样本数量的要求,提升低资源方面级情感分析的性能.因此,一种潜在的方向是考虑多种模态的方面级数据,并设计有效的跨模态特征交互方法.

第四,如何在方面级情感分析过程中引入因果推断.现有的方面级情感分析方法绝大多数基于机器学习或深度学习实现,其本质是试图寻找输入特征和输出结果之间的相关性,但相关性并不是可靠的判断依据.因此,现有方法大多面临可解释性差、鲁棒性弱等致命缺陷,尚不足以支撑其在医疗、司

法、金融等高决策风险领域的可靠应用. 反事实思维是人类思考的主要方式和因果推理的基石. 正如图灵奖得主 Judea Pearl 所指出, 机器学习理论存在局限, 其突破口在于因果革命, 这同样也适用于方面级情感分析. 因此, 一种潜在的方向是从因果分析的角度出发, 通过反事实样本增强和决策路径构建, 提升方面级情感分析模型的鲁棒性和可解释性.

6 总 结

方面级情感分析是一种细粒度的自然语言处理任务, 旨在总结文本中用户对于特定方面的情感极性. 本文从方面级情感分析的概念定义着手, 介绍了所涉及的四种基本要素和十二种子任务, 并收集了领域内常用数据集及其详细信息. 在此基础上, 针对不同子任务, 调研综述了大量相关文献, 并进行了比较分析. 随后, 阐述了低资源场景对方面级情感分析带来的问题和挑战, 并归纳了模型和数据两个层面的低资源改进策略. 最后, 对方面级情感分析做出了四点展望, 指出了未来的研究方向.

参 考 文 献

- [1] Kim Y. Convolutional neural networks for sentence classification//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha, Qatar, 2014: 1746-1751
- [2] Hu M, Liu B. Mining and summarizing customer reviews//Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, USA, 2004: 168-177
- [3] Thet T T, Na J-C, Khoo C S G. Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science*, 2010, 36(6): 823-848
- [4] Fang J, Price B, Price L. Pruning non-informative text through non-expert annotations to improve aspect-level sentiment classification//Proceedings of the 2nd Workshop on the People's Web Meets NLP: Collaboratively Constructed Semantic Resources. Beijing, China, 2010: 37-45
- [5] Brody S, Elhadad N. An unsupervised aspect-sentiment model for online reviews//Proceedings of the 2010 Conference of Human Language Technologies: The North American Chapter of the Association for Computational Linguistics. Los Angeles, USA, 2010: 804-812
- [6] Pontiki M, Galanis D, Pavlopoulos J, et al. SemEval-2014 Task 4: Aspect based sentiment analysis//Proceedings of the 8th International Workshop on Semantic Evaluation. Dublin, Ireland, 2014: 27-35
- [7] Pontiki M, Galanis D, Papageorgiou H, et al. SemEval-2015 Task 12: Aspect based sentiment analysis//Proceedings of the 9th International Workshop on Semantic Evaluation. Denver, USA, 2015: 486-495
- [8] Pontiki M, Galanis D, Papageorgiou H, et al. SemEval-2016 Task 5: Aspect based sentiment analysis//Proceedings of the 10th International Workshop on Semantic Evaluation. San Diego, USA, 2016: 19-30
- [9] Zhang W, Li X, Deng Y, et al. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. arXiv preprint, abs/2203.01054, 2022
- [10] Hedderich M A, Lange L, Adel H, et al. A survey on recent approaches for natural language processing in low-resource scenarios//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Online, 2021: 2545-2568
- [11] Broß J, Ehrig H. Automatic construction of domain and aspect specific sentiment lexicons for customer review mining//Proceedings of the 22nd ACM International Conference on Information and Knowledge Management. San Francisco, USA, 2013: 1077-1086
- [12] Fan Z, Wu Z, Dai X-Y, et al. Target-oriented opinion words extraction with target-fused neural sequence labeling//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, USA, 2019: 2509-2518
- [13] Tulkens S, Van Cranenburgh A. Embarrassingly simple unsupervised aspect extraction//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 3182-3187
- [14] Zhu P, Chen Z, Zheng H, Qian T. Aspect aware learning for aspect category sentiment analysis. *ACM Transactions on Knowledge Discovery from Data*, 2019, 13(6): 55:1-55:21
- [15] Wang W, Pan S J, Dahlmeier D, Xiao X. Coupled multi-layer attentions for co-extraction of aspect and opinion terms//Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA, 2017: 3316-3322
- [16] Liu J, Teng Z, Cui L, et al. Solving aspect category sentiment analysis as a text generation task. arXiv preprint, abs/2110.07310, 2021
- [17] Chen S, Wang Y, Liu J, Wang Y. Bidirectional machine reading comprehension for aspect sentiment triplet extraction//Proceedings of the 35th AAAI Conference on Artificial Intelligence. 2021: 12666-12674
- [18] Jian S Y B, Nayak T, Majumder N, Poria S. Aspect sentiment triplet extraction using reinforcement learning//Proceedings of the 30th ACM International Conference on Information and Knowledge Management. Queensland, Australia, 2021: 3603-3607

- [19] Mukherjee R, Nayak T, Butala Y, et al. PASTE: A tagging-free decoding framework using pointer networks for aspect sentiment triplet extraction//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 9279-9291
- [20] Wu Z, Ying C, Zhao F, et al. Grid tagging scheme for end-to-end fine-grained opinion extraction//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing. Online, 2020: 2576-2585
- [21] Xu L, Chia YK, Bing L. Learning span-level interactions for aspect sentiment triplet extraction//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. 2021: 4755-4766
- [22] Peng H, Xu L, Bing L, et al. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA, 2020: 8600-8607
- [23] Zhang C, Li Q, Song D, Wang B. A multi-task learning framework for opinion triplet extraction//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020: 819-828
- [24] Cai H, Xia R, Yu J. Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. 2021: 340-350
- [25] Zhang W, Deng Y, Li X, et al. Aspect sentiment quad prediction as paraphrase generation//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 9209-9219
- [26] Mitchell M, Aguilar J, Wilson T, Durme B V. Open domain targeted sentiment//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Seattle, USA, 2013: 1643-1654
- [27] Dong L, Wei F, Tan C, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, USA, 2014: 49-54
- [28] Yang J, Yang R, Wang C, Xie J. Multi-entity aspect-based sentiment analysis with context, entity and aspect memory//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018: 6029-6036
- [29] Wu H, Liu M, Wang J, et al. Question-answering aspect classification with multi-attention representation//Proceedings of the Information Retrieval-24th China Conference. Guilin, China, 2018: 78-89
- [30] Yu J, Jiang J. Adapting bert for target-oriented multimodal sentiment classification//Proceedings of the 28th International Joint Conference on Artificial Intelligence. Macao, China, 2019: 5408-5414
- [31] Xing X, Jin Z, Jin D, et al. Tasty burgers, soggy fries: Probing aspect robustness in aspect-based sentiment analysis//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020: 3594-3605
- [32] Hyun D, Cho J, Yu H. Building large-scale English and Korean datasets for aspect-level sentiment analysis in automotive domain//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 961-966
- [33] Wang B, Liakata M, Zubiaga A, Procter R. TDParse: Multi-target-specific sentiment recognition on Twitter//Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. Valencia, Spain, 2017: 483-493
- [34] Jiang Q, Chen L, Xu R, et al. A challenge dataset and effective models for aspect-based sentiment analysis//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong, China, 2019: 6279-6284
- [35] Rosenthal S, Farra N, Nakov P. SemEval-2017 Task 4: Sentiment analysis in Twitter//Proceedings of the 11th International Workshop on Semantic Evaluation. Vancouver, Canada, 2017: 502-518
- [36] Bu J, Ren L, Zheng S, et al. ASAP: A Chinese review dataset towards aspect category sentiment analysis and rating prediction//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2021: 2069-2079
- [37] Xu L, Li H, Lu W, Bing L. Position-aware tagging for aspect sentiment triplet extraction//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020: 2339-2349
- [38] Wiebe J, Wilson T, Cardie C. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 2005, 39(2-3): 165-210
- [39] Kessler J S, Eckert M, Clark L, Nicolov N. The 2010 ICWSM JDDA sentiment corpus for the automotive domain//Proceedings of the 4th International AAAI Conference on Weblogs and Social Media Data Workshop Challenge. Washington, USA, 2010
- [40] Toprak C, Jakob N, Gurevych I. Sentence and expression level annotation of opinions in user-generated discourse//Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Uppsala, Sweden, 2010: 575-584
- [41] Song L, Xin C, Lai S, et al. CASA: Conversational aspect sentiment analysis for dialogue understanding. *Journal of Artificial Intelligence Research*, 2022, 73: 511-533
- [42] Qiu G, Liu B, Bu J, Chen C. Opinion word expansion and target extraction through double propagation. *Computational Linguistics*, 2011, 37(1): 9-27

- [43] Zhao Yan-Yan, Qin Bing, Che Wan-Xiang, Liu Ting. Appraisal expression recognition based on syntactic path. *Journal of Software*, 2011, 22(5): 887-898(in Chinese)
(赵妍妍, 秦兵, 车万翔, 刘挺. 基于句法路径的情感评价单元识别. *软件学报*, 2011, 22(5): 887-898)
- [44] Liu K, Xu L, Zhao J. Opinion target extraction using word-based translation model//*Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Jeju Island, Korea, 2012: 1346-1356
- [45] Gindl S, Weichselbraun A, Scharl A. Rule-based opinion target and aspect extraction to acquire affective knowledge//*Proceedings of the 22nd International World Wide Web Conference*. Rio De Janeiro, Brazil, 2013: 557-564
- [46] Zhao L, Huang M, Sun J, et al. Sentiment extraction by leveraging aspect-opinion association structure//*Proceedings of the 24th ACM International Conference on Information and Knowledge Management*. Melbourne, Australia, 2015: 343-352
- [47] Liu Q, Liu B, Zhang Y, et al. Improving opinion aspect extraction using semantic similarity and aspect associations//*Proceedings of the 30th AAAI Conference on Artificial Intelligence*. Phoenix, USA, 2016: 2986-2992
- [48] Gunes O, Furche T, Orsi G. Structured aspect extraction//*Proceedings of the 26th International Conference on Computational Linguistics*. Osaka, Japan, 2016: 2321-2332
- [49] Jiang Teng-Jiao, Wan Chang-Xuan, Liu De-Xi, et al. Extracting target-opinion pairs based on semantic analysis. *Chinese Journal of Computers*, 2017, 40(3): 617-633(in Chinese)
(江腾蛟, 万常选, 刘德喜等. 基于语义分析的评价对象-情感词对抽取. *计算机学报*, 2017, 40(3): 617-633)
- [50] Shafie A S, Sharef N M, Murad M A A, Azman A. Aspect extraction performance with pos tag pattern of dependency relation in aspect-based sentiment analysis//*Proceedings of the 2018 4th International Conference on Information Retrieval and Knowledge Management*. Le Méridien Kota Kinabalu, Malaysia, 2018: 1-6
- [51] Mukherjee A, Liu B. Aspect extraction through semi-supervised modeling//*Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*. Jeju Island, Korea, 2012: 339-348
- [52] Chen Z, Mukherjee A, Liu B. Aspect extraction with automated prior knowledge learning//*Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Baltimore, USA, 2014: 347-358
- [53] Chen Z, Mukherjee A, Liu B, et al. Exploiting domain knowledge in aspect extraction//*Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Seattle, USA, 2013: 1655-1667
- [54] Peng Yun, Wan Chang-Xuan, Jiang Teng-Jiao, et al. Extracting product aspects and user opinions based on semantic constrained LDA model. *Journal of Software*, 2017, 28(3): 676-693(in Chinese)
(彭云, 万常选, 江腾蛟等. 基于语义约束 LDA 的商品特征和情感词提取. *软件学报*, 2017, 28(3): 676-693)
- [55] He R, Lee W S, Ng H T, Dahlmeier D. An unsupervised neural attention model for aspect extraction//*Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. Vancouver, Canada, 2017: 388-397
- [56] Liao M, Li J, Zhang H, et al. Coupling global and local context for unsupervised aspect extraction//*Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*. Hong Kong, China, 2019: 4578-4588
- [57] Jin W, Ho H H, Srihari R K. A novel lexicalized hmm-based learning framework for web opinion mining//*Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal, Canada, 2009: 465-472
- [58] Filho P B, Pardo T. NILC_USP: Aspect extraction using semantic labels//*Proceedings of the 8th International Workshop on Semantic Evaluation*. Dublin, Ireland, 2014: 433-436
- [59] Kiritchenko S, Zhu X, Cherry C, Mohammad S M. NRC-Canada-2014: Detecting aspects and sentiment in customer reviews//*Proceedings of the 8th International Workshop on Semantic Evaluation*. Dublin, Ireland, 2014: 437-442
- [60] Toh Z, Wang W. DLIREC: Aspect term extraction and term polarity classification system//*Proceedings of the 8th International Workshop on Semantic Evaluation*. Dublin, Ireland, 2014: 235-240
- [61] Zhang F, Zhang Z, Lan M. ECNU: A combination method and multiple features for aspect extraction and sentiment polarity classification//*Proceedings of the 8th International Workshop on Semantic Evaluation*. Dublin, Ireland, 2014: 252-258
- [62] Liu Q, Gao Z, Liu B, Zhang Y. Automated rule selection for aspect extraction in opinion mining//*Proceedings of the 24th International Joint Conference on Artificial Intelligence*. Buenos Aires, Argentina, 2015: 1291-1297
- [63] Shu L, Xu H, Liu B. Lifelong learning CRF for supervised aspect extraction//*Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. Vancouver, Canada, 2017: 148-154
- [64] Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space//*Proceedings of the 1st International Conference on Learning Representations*. Scottsdale, USA, 2013
- [65] Yin Y, Wei F, Dong L, et al. Unsupervised word and dependency path embeddings for aspect term extraction//*Proceedings of the 25th International Joint Conference on Artificial Intelligence*. New York, USA, 2016: 2979-2985
- [66] Liu P, Joty S R, Meng H M. Fine-grained opinion mining with recurrent neural networks and word embeddings//*Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon, Portugal, 2015: 1433-1443

- [67] Wang W, Pan S J. Transition-based adversarial network for cross-lingual aspect extraction//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden, 2018: 4475-4481
- [68] Ma D, Li S, Wu F, et al. Exploring sequence-to-sequence learning in aspect term extraction//Proceedings of the 57th Conference of the Association for Computational Linguistics. Florence, Italy, 2019: 3538-3547
- [69] Wei Z, Hong Y, Zou B, et al. Don't eclipse your arts due to small discrepancies: Boundary repositioning with a pointer network for aspect extraction//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 3678-3684
- [70] Yang Y, Li K, Quan X, et al. Constituency lattice encoding for aspect term extraction//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 844-855
- [71] Wang W, Pan S J, Dahlmeier D, Xiao X. Recursive neural conditional random fields for aspect-based sentiment analysis//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016: 616-626
- [72] Luo H, Li T, Liu B, et al. Improving aspect term extraction with bidirectional dependency tree representation. *IEEE ACM Transactions on Audio, Speech and Language Processing*, 2019, 27(7): 1201-1212
- [73] Xu H, Liu B, Shu L, Yu P S. Double embeddings and CNN-based sequence labeling for aspect extraction//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 592-598
- [74] Shu L, Xu H, Liu B. Controlled CNN-based sequence labeling for aspect extraction. *arXiv preprint, abs/1905.06407*, 2019
- [75] Xu H, Liu B, Shu L, Yu P S. BERT post-training for review reading comprehension and aspect-based sentiment analysis//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, USA, 2019: 2324-2335
- [76] Tang D, Qin B, Liu T. Aspect level sentiment classification with deep memory network//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016: 214-224
- [77] Wang Y, Huang M, Zhu X, Zhao L. Attention-based LSTM for aspect-level sentiment classification//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016: 606-615
- [78] Fan C, Gao Q, Du J, et al. Convolution-based memory network for aspect-based sentiment analysis//Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. Ann Arbor, USA, 2018: 1161-1164
- [79] Li X, Bing L, Lam W, Shi B. Transformation networks for target-oriented sentiment classification//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 946-956
- [80] Zhao P, Hou L, Wu O. Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification. *Knowledge-Based Systems*, 2020, 193: 105443
- [81] Tay Y, Tuan L A, Hui S C. Dyadic memory networks for aspect-based sentiment analysis//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. Singapore, 2017: 107-116
- [82] Chen P, Sun Z, Bing L, Yang W. Recurrent attention network on memory for aspect sentiment analysis//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, 2017: 452-461
- [83] Tay Y, Tuan L A, Hui S C. Learning to attend via word-aspect associative fusion for aspect-based sentiment analysis//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018: 5956-5963
- [84] Wang S, Mazumder S, Liu B, et al. Target-sensitive memory networks for aspect sentiment classification//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 957-967
- [85] Gu S, Zhang L, Hou Y, Song Y. A position-aware bidirectional attention network for aspect-level sentiment analysis//Proceedings of the 27th International Conference on Computational Linguistics. Santa Fe, USA, 2018: 774-784
- [86] Hu M, Zhao S, Zhang L, et al. CAN: Constrained attention networks for multi-aspect sentiment analysis//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong, China, 2019: 4600-4609
- [87] Du Cheng-Yu, Liu Peng-Yuan. Helical attention networks for aspect-level sentiment classification. *Journal of Chinese Information Processing*, 2020, 34(9): 70-77 (in Chinese)
(杜成玉, 刘鹏远. 基于螺旋注意力网络的方面级别情感分析模型. *中文信息学报*, 2020, 34(9): 70-77)
- [88] Ma D, Li S, Zhang X, Wang H. Interactive attention networks for aspect-level sentiment classification//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017: 4068-4074
- [89] Liang Bin, Liu Quan, Xu Jin, et al. Aspect-based sentiment analysis based on multi-attention CNN. *Journal of Computer Research and Development*, 2017, 54(8): 1724-1735 (in Chinese)
(梁斌, 刘全, 徐进等. 基于多注意力卷积神经网络的特定目标情感分析. *计算机研究与发展*, 2017, 54(8): 1724-1735)
- [90] Fan F, Feng Y, Zhao D. Multi-grained attention network for aspect-level sentiment classification//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium, 2018: 3433-3442

- [91] Majumder N, Poria S, Gelbukh A F, et al. IARM: Inter-
aspect relation modeling with memory networks in aspect-
based sentiment analysis//Proceedings of the 2018 Conference
on Empirical Methods in Natural Language Processing.
Brussels, Belgium, 2018: 3402-3411
- [92] Zhu P, Qian T. Enhanced aspect level sentiment classification
with auxiliary memory//Proceedings of the 27th International
Conference on Computational Linguistics. Santa Fe,
USA, 2018: 1077-1087
- [93] Huang B, Ou Y, Carley K M. Aspect level sentiment clas-
sification with attention-over-attention neural networks//
Proceedings of the Social, Cultural, and Behavioral Modeling-
11th International Conference. Washington, USA, 2018,
10899: 197-206
- [94] Wang J, Li J, Li S, et al. Aspect sentiment classification
with both word-level and clause-level attention networks//
Proceedings of the 27th International Joint Conference on
Artificial Intelligence. Stockholm, Sweden, 2018: 4439-
4445
- [95] Zeng Feng, Zeng Bi-Qing, Han Xu-Li, et al. Double attention
neural network for aspect-based sentiment analysis. *Journal
of Chinese Information Processing*, 2019, 33(6): 108-115
(in Chinese)
(曾锋, 曾碧卿, 韩旭丽等. 基于双层注意力循环神经网络的
方面级情感分析. *中文信息学报*, 2019, 33(6): 108-
115)
- [96] Zhang B, Li X, Xu X, et al. Knowledge guided capsule
attention network for aspect-based sentiment analysis.
*IEEE/ACM Transactions on Audio, Speech and Language
Processing*, 2020, 28: 2538-2551
- [97] Zhang M, Zhang Y, Vo D-T. Gated neural networks for
targeted sentiment analysis//Proceedings of the 30th AAAI
Conference on Artificial Intelligence. Phoenix, USA, 2016:
3087-3093
- [98] Xue W, Li T. Aspect based sentiment analysis with gated
convolutional networks//Proceedings of the 56th Annual
Meeting of the Association for Computational Linguistics.
Amelbourne, Australia, 2018: 2514-2523
- [99] Huang B, Carley K M. Parameterized convolutional neural
networks for aspect level sentiment classification//Proceedings
of the 2018 Conference on Empirical Methods in Natural
Language Processing. Brussels, Belgium, 2018: 1091-1096
- [100] Feng Chao, Li Hai-Hui, Zhao Hong-Ya, et al. Aspect-level
sentiment analysis based on hierarchical attention and gate
networks. *Journal of Chinese Information Processing*,
2020, 35(10): 128-136(in Chinese)
(冯超, 黎海辉, 赵洪雅等. 基于层次注意力机制和门机制
的属性级别情感分析. *中文信息学报*, 2020, 35(10): 128-
136)
- [101] Song Wei, Wen Zi-Jian. Feature dual distillation network
for aspect-based sentiment analysis. *Journal of Chinese
Information Processing*, 2021, 35(7): 126-133(in Chinese)
(宋威, 温子健. 基于特征双重蒸馏网络的方面级情感分
析. *中文信息学报*, 2021, 35(7): 126-133)
- [102] He R, Lee W S, Ng H T, Dahlmeier D. Effective attention
modeling for aspect-level sentiment classification//Proceedings
of the 27th International Conference on Computational
Linguistics. Santa Fe, USA, 2018: 1121-1131
- [103] Sun K, Zhang R, Mensah S, et al. Aspect-level sentiment
analysis via convolution over dependency tree//Proceedings
of the 2019 Conference on Empirical Methods in Natural
Language Processing and the 9th International Joint Conference
on Natural Language Processing. Hong Kong, China, 2019:
5678-5687
- [104] Yang Y, Qian T, Chen Z. Aspect-level sentiment classifica-
tion with dependency rules and dual attention//Proceedings
of the 26th International Conference on Neural Information
Processing. Sydney, Australia, 2019: 643-655
- [105] Phan M-H, Ogunbona P O. Modelling context and syntactical
features for aspect-based sentiment analysis//Proceedings of
the 58th Annual Meeting of the Association for Computa-
tional Linguistics, 2020: 3211-3220
- [106] Zheng Y, Zhang R, Mensah S, Mao Y. Replicate, walk,
and stop on syntax: An effective neural network model for
aspect-level sentiment classification//Proceedings of the 34th
AAAI Conference on Artificial Intelligence. New York,
USA, 2020: 9685-9692
- [107] Wang K, Shen W, Yang Y, et al. Relational graph attention
network for aspect-based sentiment analysis//Proceedings of
the 58th Annual Meeting of the Association for Computa-
tional Linguistics, 2020: 3229-3238
- [108] Wang Guang, Li Hong-Yu, Qiu Yun-Fei, et al. Aspect-
based sentiment classification via memory graph convolu-
tional network. *Journal of Chinese Information Processing*,
2021, 35(8): 98-106(in Chinese)
(王光, 李鸿宇, 邱云飞等. 基于图卷积记忆网络的方面级
情感分类. *中文信息学报*, 2021, 35(8): 98-106)
- [109] Rietzler A, Stabinger S, Opitz P, Engl S. Adapt or get left
behind: Domain adaptation through BERT language model
finetuning for aspect-target sentiment classification//Pro-
ceedings of the 12th Language Resources and Evaluation
Conference. Marseille, France, 2020: 4933-4941
- [110] Xu H, Shu L, Yu P S, Liu B. Understanding pre-trained
BERT for aspect-based sentiment analysis//Proceedings
of the 28th International Conference on Computational
Linguistics. Barcelona, Spain, 2020: 244-250
- [111] Dai J, Yan H, Sun T, et al. Does syntax matter? A strong
baseline for aspect-based sentiment analysis with RoBERTa//
Proceedings of the 2021 Conference of the North American
Chapter of the Association for Computational Linguistics;
Human Language Technologies, 2021: 1816-1829
- [112] Li Z, Zou Y, Zhang C, et al. Learning implicit sentiment in
aspect-based sentiment analysis with supervised contrastive
pre-training//Proceedings of the 2021 Conference on Empirical

- Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 246-256
- [113] Seoh R, Birlle I, Tak M, et al. Open aspect target sentiment classification with natural language prompts//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 6311-6322
- [114] Yang H, Zeng B, Xu M, Wang T. Back to reality: Leveraging pattern-driven modeling to enable affordable sentiment dependency learning. arXiv preprint, abs/2110.08604, 2021
- [115] Tang D, Qin B, Feng X, Liu T. Effective LSTMs for target-dependent sentiment classification//Proceedings of the 26th International Conference on Computational Linguistics. Osaka, Japan, 2016: 3298-3307
- [116] Liu Quan, Liang Bin, Xu Jin, Zhou Qian. A deep hierarchical neural network model for aspect-based sentiment analysis. Chinese Journal of Computers, 2018, 41(12): 2637-2652 (in Chinese)
(刘全, 梁斌, 徐进, 周倩. 一种用于基于方面情感分析的深度分层网络模型. 计算机学报, 2018, 41(12): 2637-2652)
- [117] Wang B, Lu W. Learning latent opinions for aspect-level sentiment classification//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018: 5537-5544
- [118] Liu Q, Zhang H, Zeng Y, et al. Content attention model for aspect based sentiment analysis//Proceedings of the 2018 World Wide Web Conference. Lyon, France, 2018: 1023-1032
- [119] Wang Y, Chen Q, Liu X, et al. SenHint: A joint framework for aspect-level sentiment analysis by deep neural networks and linguistic hints//Proceedings of the Companion of the Web Conference 2018. Lyon, France, 2018: 207-210
- [120] Chen Z, Qian T. Transfer capsule network for aspect level sentiment classification//Proceedings of the 57th Conference of the Association for Computational Linguistics. Florence, Italy, 2019: 547-556
- [121] Wang Y, Sun A, Huang M, Zhu X. Aspect-level sentiment analysis using as-capsules//Proceedings of the 2019 World Wide Web Conference. San Francisco, USA, 2019: 2033-2044
- [122] Yadav R K, Jiao L, Granmo O-C, Goodwin M. Human-level interpretable learning for aspect-based sentiment analysis//Proceedings of the 35th AAAI Conference on Artificial Intelligence. 2021: 14203-14212
- [123] Hu M, Peng Y, Huang Z, et al. Open-domain targeted sentiment analysis via span-based extraction and classification //Proceedings of the 57th Conference of the Association for Computational Linguistics. Florence, Italy, 2019: 537-546
- [124] Zhang M, Zhang Y, Vo D-T. Neural networks for open domain targeted sentiment//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal, 2015: 612-621
- [125] Li X, Bing L, Li P, Lam W. A unified model for opinion target extraction and target sentiment prediction//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu, USA, 2019: 6714-6721
- [126] Chen G, Tian Y, Song Y. Joint aspect extraction and sentiment analysis with directional graph convolutional networks //Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 272-279
- [127] Wang F, Lan M, Wang W. Towards a one-stop solution to both aspect extraction and sentiment analysis tasks with neural multi-task learning//Proceedings of the 2018 International Joint Conference on Neural Networks. Rio De Janeiro, Brazil, 2018: 1-8
- [128] Luo H, Li T, Liu B, Zhang J. DOER: Dual cross-shared RNN for aspect term-polarity co-extraction//Proceedings of the 57th Conference of the Association for Computational Linguistics. Florence, Italy, 2019: 591-601
- [129] He R, Lee W S, Ng H T, Dahlmeier D. Exploiting document knowledge for aspect-level sentiment classification//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 579-585
- [130] Chen Z, Qian T. Relation-aware collaborative learning for unified aspect-based sentiment analysis//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online, 2020: 3685-3694
- [131] Wang Q, Ren J. Label correction model for aspect-based sentiment analysis//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 827-832
- [132] Liu S, Li W, Wu Y, et al. Jointly modeling aspect and sentiment with dynamic heterogeneous graph neural networks. arXiv preprint, abs/2004.06427, 2020
- [133] Mao Y, Shen Y, Yu C, Cai L. A joint training dual-MRC framework for aspect based sentiment analysis//Proceedings of the 35th AAAI Conference on Artificial Intelligence. 2021: 13543-13551
- [134] Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 2020, 21: 140:1-140:67
- [135] Lewis M, Liu Y, Goyal N, et al. BART: Denoising sequence- to-sequence pre-training for natural language generation, translation, and comprehension//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 7871-7880
- [136] Yan H, Dai J, Ji T, et al. A unified generative framework for aspect-based sentiment analysis//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. 2021: 2416-2429

- [137] Li C, Gao F, Bu J, et al. SentiPrompt: Sentiment knowledge enhanced prompt-tuning for aspect-based sentiment analysis. arXiv preprint, abs/2109.08306, 2021
- [138] Zhang W, Li X, Deng Y, Bing L, Lam W. Towards generative aspect-based sentiment analysis//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. 2021; 504-510
- [139] Yu J, Jiang J, Xia R. Global inference for aspect and opinion terms co-extraction based on multi-task neural networks. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2019, 27(1): 168-177
- [140] Wu Z, Zhao F, Dai X-Y, et al. Latent opinions transfer network for target-oriented opinion words extraction//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA, 2020; 9298-9305
- [141] Veysel A, Nouri N, Dernoncourt F, et al. Introducing syntactic structures into target opinion word extraction with deep learning//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020; 8947-8956
- [142] Mensah S, Sun K, Aletras N. An empirical study on leveraging position embeddings for target-oriented opinion words extraction//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021; 9174-9179
- [143] Li Y, Wang F, Zhong S-H. Training entire-space models for target-oriented opinion words extraction//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022; 1875-1879
- [144] Zhou X, Wan X, Xiao J. Representation learning for aspect category detection in online reviews//Proceedings of the 29th AAAI Conference on Artificial Intelligence. Austin, USA, 2015; 417-424
- [145] Shi T, Li L, Wang P, Reddy C K. A simple and effective self-supervised contrastive learning framework for aspect detection//Proceedings of the 35th AAAI Conference on Artificial Intelligence. 2021; 13815-13824
- [146] Ruder S, Ghaffari P, Breslin J G. A hierarchical model of reviews for aspect-based sentiment analysis//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016; 999-1005
- [147] Sun C, Huang L, Qiu X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies. Minneapolis, USA, 2019; 380-385
- [148] Chen S, Liu J, Wang Y, et al. Synchronous double-channel recurrent network for aspect-opinion pair extraction//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020; 6515-6524
- [149] Zhao H, Huang L, Zhang R, et al. SpanMlt: A span-based multi-task learning framework for pair-wise aspect and opinion terms extraction//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020; 3239-3248
- [150] Gao L, Wang Y, Liu T, et al. Question-driven span labeling model for aspect-opinion pair extraction//Proceedings of the 35th AAAI Conference on Artificial Intelligence. 2021; 12875-12883
- [151] Wu S, Fei H, Ren Y, et al. Learn from syntax: Improving pair-wise aspect and opinion terms extraction with rich syntactic knowledge//Proceedings of the 30th International Joint Conference on Artificial Intelligence. Montreal, Canada, 2021; 3957-3963
- [152] Schmitt M, Steinheber S, Schreiber K, Roth B. Joint aspect and polarity classification for aspect-based sentiment analysis with end-to-end neural networks//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium, 2018; 1109-1114
- [153] Cai H, Tu Y, Zhou X, et al. Aspect-category based sentiment analysis with hierarchical graph convolutional network//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020; 833-843
- [154] Chen Z, Huang H, Liu B, et al. Semantic and syntactic enhanced aspect sentiment triplet extraction//Findings of the Association for Computational Linguistics, 2021; 1474-1483
- [155] Chen H, Zhai Z, Feng F, et al. Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Dublin, Ireland, 2022; 2977-2985
- [156] Fei H, Ren Y, Zhang Y, Ji D. Nonautoregressive encoder-decoder neural framework for end-to-end aspect-based sentiment triplet extraction. IEEE Transactions on Neural Networks and Learning Systems, 2021; 1-13
- [157] Wu C, Xiong Q, Yi H, et al. Multiple-element joint detection for aspect-based sentiment analysis. Knowledge-Based Systems, 2021, 223: 107073
- [158] Li X, Lam W. Deep multi-task learning for aspect term extraction with memory interaction//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, 2017; 2886-2892
- [159] Li X, Bing L, Li P, et al. Aspect term extraction with history attention and selective transformation//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden, 2018; 4194-4200
- [160] Cheng Meng, Hong Yu, Wei Zhen-Kai, Yao Jian-Min. Aspect extraction via interactive attention with sentiment words. Journal of Chinese Information Processing, 2021, 35(10): 90-100(in Chinese)

- (程梦, 洪宇, 尉桢楷, 姚建民. 融合情感词交互注意力机制的属性抽取研究. 中文信息学报, 2021, 35(10): 90-100)
- [161] He R, Lee W S, Ng H T, Dahlmeier D. An interactive multi-task learning network for end-to-end aspect-based sentiment analysis//Proceedings of the 57th Conference of the Association for Computational Linguistics. Florence, Italy, 2019: 504-515
- [162] Angelidis S, Lapata M. Summarizing opinions: Aspect extraction meets sentiment prediction and they are both weakly supervised//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium, 2018: 3675-3686
- [163] Chen Z, Qian T. Enhancing aspect term extraction with soft prototypes// Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020: 2107-2117
- [164] Zhao F, Wu Z, Dai X. Attention transfer network for aspect-level sentiment classification//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 811-821
- [165] Ma Y, Peng H, Cambria E. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018: 5876-5883
- [166] Zhao A, Yu Y. Knowledge-enabled BERT for aspect-based sentiment analysis. Knowledge-Based Systems, 2021, 227: 107220
- [167] Li B, Hou Y, Che W. Data augmentation approaches in natural language processing: A survey. arXiv preprint, abs/2110.01852, 2021
- [168] Dai X, Adel H. An analysis of simple data augmentation for named entity recognition//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 3861-3867
- [169] Chen Z, Qian T. Description and demonstration guided data augmentation for sequence tagging. World Wide Web, 2022, 25(1): 175-194
- [170] Wang Q, Wen Z, Zhao Q, et al. Progressive self-training with discriminator for aspect term extraction//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 257-268
- [171] Ding B, Liu L, Bing L, et al. DAGA: Data augmentation with a generation approach for low-resource tagging tasks//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020: 6045-6057
- [172] Li K, Chen C, Quan X, et al. Conditional augmentation for aspect term extraction via masked sequence-to-sequence generation//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 7056-7066
- [173] Hsu T-W, Chen C-C, Huang H-H, Chen H-H. Semantics-preserved data augmentation for aspect-based sentiment analysis//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 4417-4422
- [174] Ramponi A, Plank B. Neural unsupervised domain adaptation in NLP — A survey//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 6838-6855
- [175] Jakob N, Gurevych I. Extracting opinion targets in a single and cross-domain setting with conditional random fields//Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Boston, USA, 2010: 1035-1045
- [176] Li F, Pan S J, Jin O, et al. Cross-domain co-extraction of sentiment and topic lexicons//Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics. Jeju Island, Korea, 2012: 410-419
- [177] Wang W, Pan S J. Recursive neural structural correspondence network for cross-domain aspect and opinion co-extraction//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia, 2018: 2171-2181
- [178] Wang W, Pan S J. Syntactically meaningful and transferable recursive neural networks for aspect and opinion extraction. Computational Linguistics, 2019, 45(4): 705-736
- [179] Wang W, Pan S J. Transferable interactive memory network for domain adaptation in fine-grained opinion extraction//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu, USA, 2019: 7192-7199
- [180] Pereg O, Korat D, Wasserblat M. Syntactically aware cross-domain aspect and opinion terms extraction//Proceedings of the 28th International Conference on Computational Linguistics. Barcelona, Spain, 2020: 1772-1777
- [181] Zhao Guang-Yao, Lv Cheng-Guo, Fu Guo-Hong, et al. Domain specific sentiment words based attention model for cross-domain attribute-oriented sentiment analysis. Journal of Chinese Information Processing, 2021, 35(6): 93-102(in Chinese)
- (赵光耀, 吕成国, 付国宏等. 基于领域特有情感词注意力模型的跨领域属性情感分析. 中文信息学报, 2021, 35(6): 93-102)
- [182] Yang M, Yin W, Qu Q, et al. Neural attentive network for cross-domain aspect-level sentiment classification. IEEE Transactions on Affective Computing, 2021, 12(3): 761-775
- [183] Chen Z, Qian T. Bridge-based active domain adaptation for aspect term extraction//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics

- and the 11th International Joint Conference on Natural Language Processing. 2021: 317-327
- [184] Chen Z, Qian T. Retrieve-and-edit domain adaptation for end2end aspect based sentiment analysis. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 2022, 30: 659-672
- [185] Li Z, Li X, Wei Y, et al. Transferable end-to-end aspect-based sentiment analysis with selective adversarial learning // *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*. Hong Kong, China, 2019: 4589-4599
- [186] Gong C, Yu J, Xia R. Unified feature and instance based domain adaptation for aspect-based sentiment analysis // *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020: 7035-7045
- [187] Chen X, Wan X. A simple information-based approach to unsupervised domain-adaptive aspect-based sentiment analysis. *arXiv preprint, abs/2201.12549*, 2022
- [188] Ding Y, Yu J, Jiang J. Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction // *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. San Francisco, USA, 2017: 3436-3442
- [189] Zhou Y, Zhu F, Song P, et al. An adaptive hybrid framework for cross-domain aspect-based sentiment analysis // *Proceedings of the 35th AAAI Conference on Artificial Intelligence*. 2021: 14630-14637
- [190] Wang J, Sun C, Li S, et al. Aspect sentiment classification towards question-answering with reinforced bidirectional attention network // *Proceedings of the 57th Conference of the Association for Computational Linguistics*. Florence, Italy, 2019: 3548-3557
- [191] Zeng J, Liu T, Jia W, Zhou J. Fine-grained question-answer sentiment classification with hierarchical graph attention network. *Neurocomputing*, 2021, 457: 214-224
- [192] Zhang W, Deng Y, Li X, et al. Aspect-based sentiment analysis in question answering forums // *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Punta Cana, Dominican Republic, 2021: 4582-4591
- [193] Ou G, Chen W, Liu P, et al. Aspect-specific polarity-aware summarization of online reviews // *Proceedings of the Web-Age Information Management-14th International Conference (Waim 2013)*. Beidaihe, China, 2013: 289-300
- [194] Yang M, Qu Q, Shen Y, et al. Aspect and sentiment aware abstractive review summarization // *Proceedings of the 27th International Conference on Computational Linguistics*. Santa Fe, USA, 2018: 1110-1120
- [195] Pecar S. Towards opinion summarization of customer reviews // *Proceedings of ACL 2018 Student Research Workshop*. Melbourne, Australia, 2018: 1-8
- [196] Kunneman F, Wubben S, Van den Bosch A, Kraemer E. Aspect-based summarization of pros and cons in unstructured product reviews // *Proceedings of the 27th International Conference on Computational Linguistics*. Santa Fe, USA, 2018: 2219-2229
- [197] Li H, Yuan P, Xu S, et al. Aspect-aware multimodal summarization for chinese e-commerce products // *Proceedings of the 34th AAAI Conference on Artificial Intelligence*. New York, USA, 2020: 8188-8195
- [198] Liang J, Bao J, Wang Y, et al. CUSTOM: Aspect-oriented product summarization for e-commerce // *Proceedings of the Natural Language Processing and Chinese Computing-10th CCF International Conference*. Qingdao, China, 2021: 124-136
- [199] Amplayo RK, Angelidis S, Lapata M. Aspect-controllable opinion summarization // *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Punta Cana, Dominican Republic, 2021: 6578-6593
- [200] Wang K, Wan X. TransSum: Translating aspect and sentiment embeddings for self-supervised opinion summarization // *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 2021: 729-742
- [201] Ke W, Gao J, Shen H, Cheng X. ConsistSum: Unsupervised opinion summarization with the consistency of aspect, sentiment and semantic // *Proceedings of the 15th ACM International Conference on Web Search and Data Mining*. Tempe, USA, 2022: 467-475
- [202] Li X L, Liang P. Prefix-tuning: Optimizing continuous prompts for generation // *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 2021: 4582-4597
- [203] Liu X, Ji K, Fu Y, et al. P-Tuning v2: Prompt tuning can be comparable to fine-tuning universally across scales and tasks. *arXiv preprint, abs/2110.07602*, 2021
- [204] Liu X, Zheng Y, Du Z, et al. GPT understands, too. *arXiv preprint, abs/2103.10385*, 2021
- [205] Yoo K M, Park D, Kang J, et al. GPT3Mix: Leveraging large-scale language models for text augmentation // *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Punta Cana, Dominican Republic, 2021: 2225-2239
- [206] Mehri S, Altun Y, Eskenazi M. LAD: Language models as data for zero-shot dialog // *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Edinburgh, UK, 2022: 595-604
- [207] Ju X, Zhang D, Xiao R, et al. Joint multi-modal aspect-sentiment analysis with auxiliary cross-modal relation detection // *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Punta Cana, Dominican Republic, 2021: 4395-4405



CHEN Zhuang, Ph. D. , postdoctoral researcher. His research interests include information extraction, and sentiment analysis.

QIAN Tie-Yun, Ph. D. , professor. Her research interests include natural language processing, and recommendation systems.

LI Wan-Li, Ph. D. candidate. His research interest is information extraction.

ZHANG Ting, M. S. candidate. Her research interests include sentiment analysis, and dialogue systems.

ZHOU Shen, M. S. candidate. His research interest is sentiment analysis.

ZHONG Ming, Ph. D. , associate professor. His research interests include big data, social network analysis.

ZHU Yuan-Yuan, Ph. D. , associate professor. Her research interest is graph database.

LIU Meng-Chi, Ph. D. , professor. His research interests include databases, big data management and analytics, intelligent information systems.

Background

This paper investigates the research progress of aspect-based sentiment analysis (ABSA) in the field of natural language processing. Traditionally, ABSA includes three basic sub-tasks, including aspect term extraction, aspect-level sentiment classification, and end-to-end aspect-based sentiment analysis. Recently, more extended sub-tasks have been introduced into ABSA, such as aspect opinion pair extraction, aspect-opinion-sentiment triplet extraction, and aspect sentiment quad extraction. While the basic sub-tasks have been extensively studied, the newly extended sub-tasks are less discussed and still face great challenges. More importantly, existing methods for ABSA tasks are mostly supervised methods, which require sufficient labeled data to train a good classifier. However, in real-world applications, it is hard and expensive to obtain a large number of labeled samples. In view of this, we present this survey to show the detail of both classic and new sub-tasks in ABSA, as well as the techniques for tackling ABSA problems in the low-resource scenario.

The work described in this paper is supported by the grants from the National Natural Science Foundation of China (NSFC) projects (Nos. 62276193 and 61572376). Our group has published more than ten papers along this direction, including top conferences like ACL, EMNLP, and COLING, and premier journals like TASLP, TKDD. Our studies cover both the basic and extended sub-tasks, and most of them are toward low-resource ABSA. Below is the list of some selected papers.

(1) Chen Z, Qian T. Retrieve-and-edit domain adaptation for end2end aspect based sentiment analysis. *IEEE/ACM Transactions on Audio, Speech and Language Processing*,

2022, 30: 659-672.

(2) Chen Z, Qian T. Bridge-based active domain adaptation for aspect term extraction//*Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. 2021: 317-327.

(3) Chen Z, Qian T. Relation-aware collaborative learning for unified aspect-based sentiment analysis//*Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020: 3685-3694.

(4) Chen Z, Qian T. Enhancing aspect term extraction with soft prototypes//*Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. 2020: 2107-2117.

(5) Zhang M, Qian T. Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis //*Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. 2020: 3540-3549.

(6) Chen Z, Qian T. Transfer capsule network for aspect level sentiment classification// *Proceedings of the 57th Conference of the Association for Computational Linguistics*, Florence, Italy, 2019: 547-556.

(7) Zhu P, Chen Z, Zheng H, Qian T. Aspect aware learning for aspect category sentiment analysis. *ACM Transactions on Knowledge Discovery from Data*, 2019, 13(6): 55:1-55:21.

(8) Zhu P, Qian T. Enhanced aspect level sentiment classification with auxiliary memory//*Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, USA, 2018: 1077-1087.