

网络评论方面级观点挖掘方法研究综述*

韩忠明^{1,2}, 李梦琪¹, 刘雯¹, 张梦玫¹, 段大高¹, 于重重^{1,2}



¹(北京工商大学 计算机与信息工程学院, 北京 100048)

²(食品安全大数据技术北京市重点实验室(北京工商大学), 北京 100048)

通讯作者: 韩忠明, E-mail: hanzhongming@btbu.edu.cn

摘要: 网络评论的观点挖掘任务是文本分析的关键问题之一。随着网络评论的快速增长,用户在浏览评论时更加关注细粒度的信息,因此,对评论进行方面级观点挖掘能够帮助消费者更好地做出决策。过去的10多年间,研究人员在大量网络评论语料库上进行观点挖掘等相关研究,并取得了丰硕的研究成果和广泛的应用价值,更不乏优秀学者对观点挖掘方法现状进行综述总结。然而,有针对性地对观点挖掘中的方面提取与观点提取进行综述总结的成果较少。综述了近年来网络评论方面级观点挖掘的研究现状:首先,介绍了方面级观点挖掘的相关问题描述;然后,重点分类介绍方面提取方法及观点内容提取的主要方法;随后,总结了方面级观点挖掘的常见评价指标以及在社会中的广泛应用价值;最后,根据对现有方法提出具有挑战性的方向并进行系统总结。对方面级观点挖掘进行综述有助于比较不同方法的差异,从而发现有价值的研究方向。

关键词: 观点挖掘;方面提取;观点表达提取;序列标注;网络评论

中图法分类号: TP311

中文引用格式: 韩忠明,李梦琪,刘雯,张梦玫,段大高,于重重.网络评论方面级观点挖掘方法研究综述.软件学报,2018,29(2): 417-441. <http://www.jos.org.cn/1000-9825/5386.htm>

英文引用格式: Han ZM, Li MQ, Liu W, Zhang MM, Duan DG, Yu CC. Survey of studies on aspect-based opinion mining of Internet. Ruan Jian Xue Bao/Journal of Software, 2018,29(2):417-441 (in Chinese). <http://www.jos.org.cn/1000-9825/5386.htm>

Survey of Studies on Aspect-Based Opinion Mining of Internet

HAN Zhong-Ming^{1,2}, LI Meng-Qi¹, LIU Wen¹, ZHANG Meng-Mei¹, DUAN Da-Gao¹, YU Chong-Chong^{1,2}

¹(School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

²(Beijing Key Laboratory of Big Data Technology for Food Safety (Beijing Technology and Business University), Beijing 100048, China)

Abstract: Opinion mining (OM) of Internet reviews is one of the key issues in text analysis. As the rapid growth of the Internet reviews, users pay more attention to all this fine-grained information when browsing comments. Therefore, aspect-level OM can help consumers make better decisions. In last decade, researchers conducted opinion extraction and analysis on a large number of Internet reviews corpus, and have achieved fruitful research results and broaden the scope of application. There were also some scholars conducted summaries on the present situation of OM methods. To rectify the lack of specific summaries on aspect extraction and opinion expression extraction, this paper analyzes and summarizes the recent research status of aspect-level OM on Internet reviews. The paper describes the aspect-level OM, introduces the different methods of aspect extraction and opinion expression extraction, and summarizes the evaluation measures of aspect-level OM and application values. In the end, it provides an overview of the future challenges along with a synopsis on the existing techniques. This specific survey on aspect-level OM helps to evaluate the different methods and find valuable research direction.

Key words: opinion mining; aspect extraction; opinion expression extraction; sequence labeling; Internet review

* 基金项目: 国家自然科学基金(61170112, 61532006); 北京市自然科学基金(4172016)

Foundation item: Foundation item: National Natural Science Foundation of China (61170112, 61532006); Beijing Natural Science Foundation (4172016)

收稿时间: 2017-01-10; 修改时间: 2017-06-09, 2017-08-25; 采用时间: 2017-09-09; jos 在线出版时间: 2017-10-09

CNKI 网络优先出版: 2017-10-09 16:20:55, <http://kns.cnki.net/kcms/detail/11.2560.TP.20171009.1620.006.html>

随着网络技术的迅速发展以及计算机与手机等互联设备的普及,互联网已经成为日常生活中不可或缺的一部分,人们也愈加倾向于借助互联网来发表自己对电子商品、餐饮服务及热门时事等事物的看法与观点,这样就产生了大量的网络评论.网络评论通常针对某一实体描述,且涵盖较多包含主观倾向的内容信息.网络评论文本的观点挖掘任务可看作是从评论文本中获取对评论实体有价值的观点评论信息,从而有助于进一步地对文本信息进行如情感倾向性分析等有应用价值的处理.

用户在针对某实体发表观点时,除了在评论中给出总体评价外,通常也会针对该实体的多个方面发表观点评论.事实上,这些细粒度的信息同样会受到广大用户阅读时的重点关注,同时也更加值得研究者们进一步去挖掘、分析与总结.举例分析,图 1 给出了两条小米 5s 手机评论,两条手机总体给出的评分均为 5 分,没有区分度.当用户阅读该评论并试图寻找产品特征时,则会更关注这两条评论中细粒度的信息.如,第 1 条评论反映了快递速度快,急切买到的用户可能因此而下单;第 2 条评论提到了手机运行流畅但电池容量小,因此对电池要求过高的用户可能不会下单或继续观望.因此,面对互联网中海量的评论内容,我们需要一种能够自动对大量评论数据进行方面级观点挖掘的方法来挖掘和分析评论中细粒度方面的观点,这样才能使用户更快速、更准确地发现评论中有价值的观点,卖家也能因此为用户提供更有效的推荐决策等服务.



Fig.1 MI 5s' phone comments

图 1 小米 5s 手机评论

观点挖掘研究技术最早在 1997 年由 Hatzivassiloglou 等人^[1]提出,随着互联网的发展,观点挖掘逐渐成为数据挖掘中最重要的研究领域之一.在观点挖掘研究不断发展与壮大的同时,不乏众多优秀学者对该领域问题进行分析总结.国内学者中,姚天昉等人^[2]在 2008 年对观点挖掘的研究现状进行了总结,主要内容包括:介绍观点挖掘的定义及研究目的;分别从方面提取、观点发出者识别、观点陈述提取和情感分析这几方面进行分析,同时总结了一些应用于生活中的应用系统;详细叙述中文文本观点挖掘的研究现状.王辉等人^[3]在 2009 年针对观点挖掘的各种应用形式进行了较为详尽的阐释,其内容主要以文本观点挖掘、观点搜索与观点作弊这 3 个方面为主.陈旻等人^[4]在 2014 年以观点挖掘的任务流程为引导,分别从观点提取、观点倾向性分析、观点总结等方面进行了较为全面、详细的分析.国外学者中,Pang 等人^[5]在 2008 年分别从观点挖掘的技术方法、观点挖掘的应用及未来面临的挑战这 3 个方面进行综述.Liu^[6]在 2012 年对观点挖掘领域涉及到的各项方法技术进行了较为全面的综述.Ghag 等人^[7]列出了 38 种观点挖掘所适用的方法,但缺少对问题有针对性的分类.2015 年,Ravi 等人^[8]、Bhatia 等人^[9]与 Guellil 等人^[10]分别对观点挖掘领域进行了较为详尽的综述,其中,Guellil 等人总结了整个社交媒体领域的相关问题与方法研究,包括用户数据挖掘、用户关系挖掘及用户所发表的文本数据挖掘等.Rana 等人^[11]在 2016 年分别从显式方面提取及隐式方面提取两个角度对方面挖掘技术做了较全面的综述.Sun 等人^[12]于 2017 年分别对自然语言处理(natural language processing,简称 NLP)开源工具、不同层次的观点挖掘方法及现有资源进行了综述,此外,还特别对“比较级观点挖掘”这一新兴领域进行了初步综述.李建华等人^[13]针对不同层次的情感分析技术进行了综述,并对词库资源、语料库资源、评测方法及重要会议等进行了总结分

析.Chen 等人^[14]在近期,具体从评价对象、观点持有者与情感词的抽取、主观性文本识别和文本情感分类等角度进行综述,并对它们的相关研究和技术方法进行了详细介绍.通过分析以上相关综述,有 5 个原因促使我们对网络评论观点挖掘的研究文献进行重新梳理和总结.

- 1) 现有的综述主要将观点挖掘与情感分析(观点挖掘与情感分析并不完全相同,下一节会对其进行阐述)合并在一起进行总结,缺少单独针对观点挖掘领域现状的更有深度及广度的分析与总结.
- 2) 现有的观点挖掘综述大多针对文档级与句子级观点挖掘进行总结,涉及到方面级观点挖掘的信息较为笼统、分散,缺少专门针对方面级观点挖掘方法较为深入的综述总结.
- 3) 现有的综述中,针对观点挖掘过程中观点内容的提取方法总结较少、较浅,一些先进方法(如深度学习方法)并未全面总结到位.
- 4) 现有的观点挖掘综述缺乏对方面提取及观点内容提取评估指标的系统总结与对比.
- 5) 对方面级观点挖掘进行综述,能够帮助初入此领域的研究者更快、更详细地了解整个领域的研究情况,同时对该领域的研究人员在选择相关的方法与技术上有较好的指导意义.

本文对网络评论方面级观点挖掘的定义、方法、评估以及应用等相关研究进行深入细致的综述,为短文本观点挖掘的进一步研究奠定基础,提供思路.本文第 1 节介绍相关问题描述及定义.第 2 节分别从显式方面提取及隐式方面提取两个角度介绍方面提取过程的现有方法及相应优缺点.第 3 节分门别类介绍基于方面的观点内容提取方法及相应优缺点.第 4 节介绍方面级观点挖掘的评价指标及应用领域.第 5 节提出挑战性问题与值得进一步研究的方向.最后在第 6 节对全文进行总结.

1 问题描述

1.1 网络评论

概括地讲,评论包括人对某种事物持有的观点,以及对所持有观点的阐释.按照评论对象划分,最主要的几类网络评论有:新闻评论、图书评论、影视评论及产品与服务评论^[15].

网络评论整体特点为:数据海量、复杂多样、可用价值高与非结构化.不同类型与领域的评论各有其特点,对相应特征表现的总结见表 1.由此可见,对不同领域的评论信息进行观点挖掘的方法也会有所不同.

Table 1 Categories and characteristics of Internet reviews

表 1 网络评论类别及特点

评论类别	评论特性
新闻评论	形式上较为规范,观点倾向明显
图书评论	形式松散,专业性不强,情感倾向不明显,评论对象相对容易判断
影视评论	类似图书评论,往往由感而发,随意性更强
产品和服务评论	内容简短,评论对象和情感倾向明显

1.2 方面级观点挖掘

根据 Kim 等人的定义^[16],观点由 4 个元素组成:方面(aspect)、持有者(holder)、观点内容(expression)及情感(sentiment).这四者之间所存在的联系为:观点的持有者针对某一方面发表了具有情感的观点内容.

观点挖掘表示一种对实体及实体方面的观点及态度的挖掘研究.观点挖掘与情感分析不尽相同,二者是两个经常容易混淆的研究领域.情感分析是指通过自动分析网络评论的文本内容,挖掘评论用户对这方面的褒贬态度倾向^[17,18],大部分网络评论情感分析集中于评论的情感极性分析^[19-21].观点挖掘往往伴随情感分析,但并不绝对.常用的观点挖掘流程为:方面挖掘、基于方面的观点内容挖掘、观点情感分析、观点总结.本文将分析总结的着重点置于流程的第 1 步、第 2 步,即方面与观点内容挖掘.

观点内容挖掘是指在评论词序列中挖掘出针对不同方面的相应观点子序列,即观点内容.需要指出的是,观点内容不一定是情感词,而情感词则包含于观点内容之中.以电影《神奇动物在哪里》的豆瓣影评“这比垃圾奇

异博士好看”为例,句中情感词为“垃圾”“好看”,观点内容则可以是“比奇异博士好看”.由此可见,只提取情感词并不能充分反映句子观点.

文本观点挖掘可从 3 个目标方向开展研究:文档级观点挖掘、句子级观点挖掘与方面级观点挖掘.文档级观点挖掘由 Pang 等人^[17]提出,目标是将所有文档数据所表达出来的情感分为正、负、平这 3 类^[22].然而该类方法太过宽泛,大多数情况下,正、负情感均出现在同篇文档中,所以基于文档级的观点挖掘显然对于理解一篇文档来说并不充分;句子级观点挖掘首先判断句子主观性,而后对主观性句子进行观点挖掘或判断情感极性^[23,24],虽然句子级的观点挖掘在众多场合都有较大的用处,但仍有较大的空间值得探索;方面级观点挖掘是将包含评价实体及实体方面的文本筛选出来,然后针对实体的不同方面对其进行相应的观点内容挖掘^[25],而不是提取评论文本的总体性评价.本文将综述重点主要集中于面向方面的观点挖掘上.实现对评论中不同方面进行观点挖掘这一过程主要需两步工作:方面提取以及基于方面的观点内容提取,得到观点二元组(方面,观点).

对于图 1 中的第 1 条评论,方面级观点挖掘的目标即是从该评论中提取出二元组(快递,速度快);对于第 2 条评论,则可以提取出(电池,容量小)与(速度,很流畅)等.

2 方面提取

方面是指在一条互联网评论语句中粒度最细的评论对象,即观点词所指向的最小对象.目前的研究工作中,方面提取任务大多数应用于在线消费者评论语料库.方面通常被分为两类:显式方面与隐式方面.显式方面是指在评论中正向指明观点所对应的方面,如电子产品评论“电池不禁用”中的方面“电池”;隐式方面通常不直接指明,而是暗含在观点评论中,如电子产品评论“一只手握不住手机”暗含隐式方面“手机尺寸”.

2.1 显式方面提取

总结相关研究,显式方面提取的方法按照学习方式不同可以分为带监督学习的方法、无监督学习的方法与半监督学习的方法.

2.1.1 带监督学习方法

方面的提取过程可看作是一个文本序列标注问题,因此可利用带监督的统计模型对序列标注过程进行训练,从而提取方面.适用此问题的典型带监督学习的方法有隐马尔可夫模型(hidden Markov model,简称 HMM)、条件随机场模型(conditional random field,简称 CRF)等.Jin 等人^[26,27]采用一种编入词汇的 HMM 模型来提取显式方面.该模型能够识别方面及其对应观点内容,同时能够判断观点极性.该方法首先建立由不同词汇及其对应词性组成的词集,以此来对评论文本的方面及观点内容进行人工标注;然后,将标注完毕的数据送入 HMM 进行训练.对于 CRF 模型而言,其基本思想是,通过依赖于少数变量的局部函数的乘积来表示一个依赖大量随机变量的概率分布.由于方面的抽取过程可视作一个序列标注问题,因此可采用 CRF 中的线性链状结构来解决此问题,图 2 即为 CRF 中的线性链状结构.

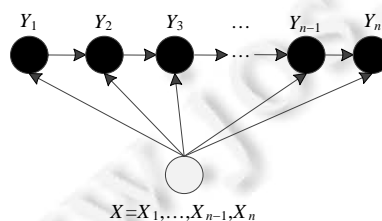


Fig.2 Chain of undirected graph model

图 2 链式的无向图模型

序列标注任务中,随机变量是观察序列.随机变量表示观察序列对应的标记序列.Zhang 等人^[28]采用带监督学习的条件随机场模型,在该模型中引入词性等特征进行模型训练,并构建相应领域词典,利用该词典识别显性产品方面,最终取得较好的实验效果.Xu 等人^[29]同样基于带监督的条件随机场模型来提取显性方面,并在模型

中引入了位置、词性和上下文信息等浅层语法特征,并最终采用精确重叠与比例重叠两种模型评估方式来评价模型.实验结果显示:将上述特征添加到条件随机场模型后,模型的各项指标提高明显,模型性能明显优于参照模型——最大熵模型的性能.

主题模型中,以 pLSA(probabilistic latent semantic analysis)^[30]和 LDA(latent Dirichlet allocation)^[31]模型为主的方法也被许多研究者应用于方面提取的任务中^[32,33].其主要思想是:在文档与单词间建立“主题”这一桥梁,通常的主题模型为无监督模型.Mcauliffe 等人^[34]在 LDA 方法中为每个样本添加用户反馈,实现通过带监督的学习过程来提取名词短语.Kessler 等人^[35]创建了一个由汽车与数码相机产品评论组成的庞大英文语料库.该语料库标记了情感类型、指代关系、从属关系及实体方面,能够很好地被用于产品方面的识别,同时利用 RankSVM(rank support vector machine)模型来提取方面.Jiang 等人^[36]基于对标准句法树的改进,提出了一种基于广义对象情感树(generalized aspect-sentiment tree,简称 GFST)的方法,通过事先定义的 4 种树内核空间来识别评论中的显式方面.此种基于内核的方法实际上是一个分类任务,利用机器学习方法对所有可能的〈显示方面,观点〉二元组进行匹配正误分类.模型的特征输入为树与树间的相似度,因此不必从每棵树中提取个体特征,模型召回率可达 89.83%.在近期的工作中,深度学习中卷积神经网络(convolutional neural network,简称 CNN)^[37,38]网络被应用于方面提取任务,并展现出卓越的效果.Poria 等人^[39]在 2016 年提出一种基于 7 层深度卷积网络的模型来对句子进行标记训练,从而提取方面.同时,他们定义了一种基于语言学的规则模式,并将其融合进该模型中.实验结果表明,该方法相对于目前其他方法有更佳的表现.

2.1.2 无监督学习方法

无监督的方面提取方法中,Hu 等人^[40]最早认为:在多数情况下,名词及名词短语即为显式方面.该方法首先将电子产品评论数据进行词性标注(part-of-speech tagging,简称 POS),并使用以 Apriori 算法为基础的关联规则挖掘方法找出频繁出现的名词及名词短语作为候选方面;然后,将错误的词语通过剪枝算法将其进行过滤,最终形成方面集合,方法流程图如图 3 所示,效仿此做法的还有 Chinsha 等人^[41].

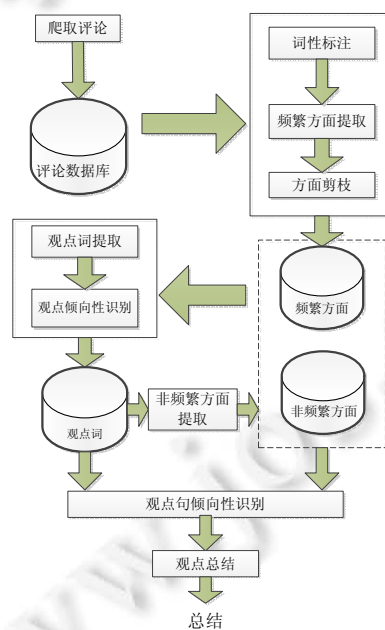


Fig.3 Flow chart of opinion mining of Hu, et al. ^[40]

图 3 Hu 等人的观点挖掘流程图^[40]

基于文献[40]的方法,Popescu 等人^[42]在剪枝策略中加入 PMI(pointwise mutual information)计算步骤来确定方面词语.针对某一特定领域的产品评论,他们首先定义了一系列产品相关的部分-整体间的关系判别短语,

如扫描仪类中的“扫描仪的”“扫描仪有”“扫描仪自带的”等,通过计算评论中频繁的候选方面短语与上述判别短语间的 PMI 值来确定该候选短语是否表示该类产品某个方面.与文献[40]的方法相比,该方法在精确度上提高了 22%(同时,召回率下降 3%).刘鸿宇等人^[43]使用了基于句法分析的方法,对评论文本首先进行词性标注,继而利用句法依存模板进行分析,根据既定的规则抽取出频繁候选方面,最终使用剪枝算法去掉错误的方面词.由于上述方法是基于频繁挖掘规则来提取显示方面,因此仅能较好地识别出频繁方面,非频繁的方面并不能被有效地识别.

Yi 等人^[44]依据方面名词短语的语法结构特点,采用相似性测试(likelihood test)的方法来识别显式方面,但这种方法不能高效地解决方面术语的覆盖性问题.若研究目标领域较小,可以构建方面词典,用于评论文本中方面的匹配抽取.Samha 等人^[45]采用 WordNet 词典(wordnet.princeton.edu)来进行显式方面提取,通过查询词典中相关领域方面名称与同义词信息来识别显式方面.这种基于词典的无监督的方法较为朴素、常见,也简单、快捷,但由于词典无法全面地涵盖各个领域下词语间的关系,当目标文本涉及多领域或结构较复杂时,此方法并不适用. Zhu 等人^[46]在 2011 年提出另一种无监督式 MAB(multi-aspect bootstrapping)模型,对中文餐厅评论进行显式方面挖掘.Bagheri 等人^[47]利用 Bootstrapping 模型与 POS 机制来对方面进行提取:首先,将文本进行词性标注;然后,利用启发式规则筛选出符合标准的方面组成模型所需要的种子词汇集;最后,采用 bootstrapping 方法从数据中自助提取方面.Quan 等人^[48]利用点互信息和词频-逆文档频率(TF-IDF)来发现显式方面与相关实体类间的联系,由此判断方面所属的实体.如方面“照片质量”相对于实体类“MP3”更接近于实体类“数码相机”.Zhou 等人^[49]在 2016 年开发了一个实现方面提取到观点总结的 CMiner 系统,首次将其应用于微博话题评论数据中.他们采用基于动态规划的算法来进行中文微博中的命名实体标签分割,基于“同一话题内的微博可能关注相同或相似的方面”这一假设实现了无监督的标签传播算法,从而生成方面候选集.

2.1.3 半监督学习方法

Lu 等人^[50]在 2008 年提出一个半监督模型,在标准 pLSA 的基础上,将一些名词性显式方面词语作为先验知识,利用这些先验知识去训练半监督 pLSA 的模型参数,因此模型性能较为依赖先验词汇的选择.随后, Andrzejewski 等人^[51]提出了基于 LDA 的半监督模型:DF-LDA 模型.模型为任意两词设置“must-link”或“cannot-link”关系,前者表示两词倾向于来自同一个主题,后者则表示两词倾向于来自不同主题.随后,利用词语间的这种约束关系建立狄利克雷森林,并将此作为模型先验知识.与标准 LDA 相比,DF-LDA 能够识别更多显示方面. Mukherjee 等人^[52]提出了两个新型半监督模型:SAS(seeded aspect and sentiment model)和 ME-SAS(maximum entropy-SAS model).前者先提取方面后提取观点内容,后者将方面与内容联合提取.同时,模型中很好地融合了用户的需求——加入相应种子词汇.实验结果显示,该模型比 DF-LDA 模型有更出色的表现.Wang 等人^[53]同样提出了两个半监督模型,从电商产品描述中提取种子词汇,并依据该种子词汇将评论预先分类,此举能够引导模型学习更多相关关键词.2016 年,刘倩^[54]提出了一种基于回答集编程语言的方面抽取框架,同时提出了两种自动选择规则的方法和一种基于语义相似性和相关性的评价对象推荐方法,并引入种子词汇.这一体系的方法能够从质量参差不齐的抽取规则中自主选择质量较高的规则,并从大量无标注的数据中快速抽取观点方面.

多种显式方面提取方法各有利弊,方法各项属性详细对比可见表 2.同时,将各种方法实验结果按使用数据集划分并进行对比,具体见表 3.从表 3 中可看到:大部分学者在研究显示方面提取方法时会采用现有数据集作为方法测评数据集,如 Hu 等人的电子产品评论数据集^[40]和 Qiu 的酒店评论数据集^[55]等;同时,也有不少学者选择爬取在线文本数据并对其进行人工标注.结合表 2 能够观察到:带监督方法中,主要模型的精度较高,缺陷在于领域的局限性较强,因此领域间的可移植性较差;同时,训练过程中需大量人工标注语料,相对无监督与半监督方法更为费时.大多数无监督方法相对来说可操作性强,无需标注大量数据,适用于目标领域较小的数据集,难点在于需提前建立规则模板,并且要对海量的评论数据进行语义分析等预处理工作,人工干预相对较多,不适用结构较为复杂的数据.半监督方法能够利用局部标注来完成显式方面的全局挖掘,但对数据领域有着较强的依赖.总之,能够自主挖掘显式方面,大量减少人工干涉同时提高结果的精准度仍是显式方面提取方法的目标.

Table 2 Comparison of explicit aspect extraction methods
表 2 显式方面提取方法对比

方法类	文献	模型	使用算法	应用语种	应用领域	隐式方面提取	备注
带监督学习方法	Jin 等人 ^[26,27]	HMM	lexicalized HMM+ Bootsrapping	英文	相机评论	无	不依赖 NLP 技术及统计信息;融入语言学特征;能够自主学习新词
	Zhang 等人 ^[28]	CRF	CRF+领域词典+POS	中文	产品评论	无	融入词特征与领域词典知识,精度高且有效;需要人工定义并选择特征,耗时、耗力
	Xu 等人 ^[29]	CRF	CRF+评价词词典+语言学特征	中文	产品\汽车评论	无	引用语言学特征和启发式位置特征;可跨领域使用
	Kessler 等人 ^[35]	RankSVM	SVM	英文	汽车\相机评论	无	能够发现观点指向的多个方面
	Jiang 等人 ^[36]	GFST	基于 tree kernels	英文	产品评论	无	降低了特征选择复杂性,考虑情感信息特征;模型依赖树内核空间的定义,定义树内核空间较困难
	Poria 等人 ^[39]	Deep CNN	CNN+词向量+POS	中文/英文	电脑\酒店评论	无	精度较高,适用于复杂数据;网络结构对参数较为敏感
无监督学习方法	Hu 等人 ^[40]	FBS	基于词频	英文	产品评论	无	简单易行,有剪枝策略;频繁方面挖掘效果不如非频繁方面
	Chinsha 等人 ^[41]	AOMR	基于规则+依存关系+评价词词典	英文	餐饮评论	无	无需任何训练数据;领域依赖性强,复杂句解析错误率较高
	Popescu 等人 ^[42]	OPINE	基于 PMI	英文	产品评论	有	引入剪枝策略,查准率高,可跨领域使用
	刘鸿宇等人 ^[43]	基于语义关系	基于句法依存关系+ PMI	中文	产品评论	无	仅能提取高频方面.短语边界识别不佳
	Yi 等人 ^[44]	SA	基于短语结构+似然比检验	英文	相机\音乐评论	无	仅识别名词性短语,在不同领域查准率差异较大
	Samha 等人 ^[45]	基于词典	基于名词词典+自定义方面词典+ POS	英文	产品评论	无	对方面词典质量与领域依赖性强
	Zhu 等人 ^[46]	MAB	Bootstrapping	中文	餐饮评论	无	英语评论适用,对领域无依赖性
	Bagheri 等人 ^[47]	BST1*	Bootstrapping+启发式规则	英文	产品评论	有	可识别方面短语,可自由移植至其他语言与领域
	Quan 等人 ^[48]	基于 PMI	PMI+TFIDF	英文	产品评论	无	引入方面与实体类关系
	Zhou 等人 ^[49]	CMiner	基于规则+hashtag+百科词典	中文	博客	有	方面短语边界识别不佳,模型性能对 NLP 工具依赖性强
半监督学习方法	Lu 等人 ^[50]	PLSA	PLSA+最大后验概率估计量	英文	手机\政治评论	无	任意领域的评论数据适用
	Andrzejews 等人 ^[51]	DF-LDA	狄里克雷森林 LDA+ Gibbs&MCMC 采样	英文	虚构+真实文本	无	允许用户设置约束,增加先验条件
	Mukherjee 等人 ^[52]	SAS/ME-SAS	LDA+最大熵先验+种子词	英文	酒店评论	无	迎合用户需求引入种子词汇;限定领域,可移植性差
	Wang 等人 ^[53]	FL-LDA/UFL-LDA	LDA+种子词	英文	相机\酒店评论	无	考虑产品描述信息,学习能力强,人为干涉少;过度依赖于种子词集
	刘倩 ^[54]	ASP-DP	ASP 语言+贪心算法+模拟退火算法+词频过滤	英文	产品评论	无	能够实现向新领域推荐高质量方面词汇

Table 3 Comparison of explicit aspect extraction methods' evaluation results**表 3** 显式方面提取方法实验结果对比

数据集	文献	实验结果		
		P (%)	R (%)	F
Hu 等人的数据集 ^[40]	Hu 等人 ^[40]	72.0	80.0	0.76
	Popescu 等人 ^[42]	88.0	77.0	0.82
	Samha 等人 ^[45]	99.0	64.0	0.77
	Quan 等人 ^[48]	84.0	91.0	0.87
	刘倩 ^[54]	81.9	84.6	0.83
	Jiang 等人 ^[36]	89.1	90.5	0.89
Qiu 数据集 ^[55]	Mukherjee 等人 ^[52]	88.0	79.2	0.83
	Wang 等人 ^[53]	79.0	80.7	0.80
	Poria ^[39]	87.4	85.0	0.86
	Chinsha 等人 ^[41]	83.0	89.3	0.86
	Zhu 等人 ^[46]	81.3	78.2	0.80
COAE 2008 数据集 ^[56]	Zhang 等人 ^[28]	41.1	45.1	0.43
	Xu 等人 ^[29]	75.0	62.9	0.68
	刘鸿宇等人 ^[43]	72.0	57.9	0.52
亚马逊数码相机评论(人工标注)	Jin 等人 ^[26,27]	62.0	88.6	0.73
汽车/数码相机评论(人工标注)	Kessler 等人 ^[35]	74.8	65.4	0.70
数码相机/音乐评论(人工标注)	Yi 等人 ^[44]	97.0	-	-
电子产品评论. http://www.cs.uic.edu/~liub/FBS/sentimentanalysis.html#datasets	Bagheri 等人 ^[47]	84.1	66.2	0.74
CSMAE 2012 数据集. http://tcci.ccf.org.cn/conference/2012/pages/page04_eva.html	Zhou 等人 ^[49]	62.1	55.7	0.58
iPhone/名人评论(人工标注)	Lu 等人 ^[50]	-	-	-
合成语料库(人工标注)	Andrzejewski 等人 ^[51]	-	-	-

2.2 隐式方面提取

针对隐式方面中指示词指代特定方面的情况, Yang 等人^[57]利用模板来获取指示代词和方面名词之间存在的语义关系.他们首先利用维基百科来获取大量语义模板,并对这些模板进行评分;然后,利用此模板获取指示代词与方面词之间的语义特征.同样致力于此问题的还有郎君等人^[58],他们从维基百科与 WordNet 上获取背景知识,并将这些先行语义知识与句子本身含有的特征(如词汇间的语义关系、句子上下文特征等)共同作为指代消解模型的输入特征.实验结果表明:当特征集合加入背景语义知识后,模型的召回率提高了 3 个百分点, F 值提高了 5 个百分点,同时证明了模型获取背景知识后能够识别更多的隐式方面.

Hai 等人^[59]提出一种基于 ABOOT 模型(association-based bootstrapping method)的方法来提取隐式方面:模型先对已提取出的显式方面进行聚类构成方面集合;然后,利用该方面集合与相应观点间的关系建立强健的语义关系对应集;最后,将指代不清的新观点通过语义关系集定位到相应方面集合.Zeng 等人^[60]提出一种基于规则与文本分类的方法来提取隐式方面:首先,将显式方面与相应观点词聚类组成相应集合;然后,将隐式方面识别过程看作文本分类(隐式方面句分类)问题,利用文本分类计数来提取隐式方面.Sun 等人^[61]同样利用显式方面与观点词间已有关系来确定隐式方面,并在此基础上添加隐式方面候选集及打分机制来进一步精确选择.使用该映射关系的还有马京苗等人^[62],他们利用关联规则挖掘办法来提取隐式方面.Zhou 等人^[49]在对中文微博进行观点挖掘时发现:同一话题内,意思相近的句子拥有相同的方面.因此,基于该结论来提取隐式方面.

Xu 等人^[63]采用 LDA 模型和 SVM 分类器来提取隐式方面:首先,用 LDA 来提取显式方面;而 LDA 主题模型由两个约束性和关联性先验参数来引导建立,这些参数构成 SVM 分类器的训练特征;最后,由 SVM 分类器来分类识别隐式与显式方面.该方法在京东商城 14 218 条中文评论上实验的 $F1$ -Score 可达到 77.78%.采用 LDA 模型来提取隐式方面的还有 Lau 等人^[64],他们在模型中还引入了吉布斯抽样方法.

Zhang 等人^[65]提出一种新型的基于共现关系来识别隐式方面的模型:首先计算语料库中所有词汇的共现频率,通过修正方面与观点间的关系来确定词与词关系矩阵,识别所有观点词,并通过关系矩阵找出所有可能的隐式方面.Bagheri 等人^[66]提出一种无监督的领域独立模型来提取隐式方面,观点词作为节点映射到相应的一组方面节点上,方面与观点词间有联系则联通边的权重较高,然后,采用基于图的评分法来计算方面与观点节点间

的距离从而识别隐式方面.Fei 等人^[67]尝试用词典中形容词与名词的对应关系来提取隐式方面.Xia 等人^[68]提出一种利用主题图(topic maps)来提取隐式方面的方法,主题图中定义了大量概念间的多种关系,其中不仅包含方面间的层级关系,同时也包含了方面与对应情感词间的关系,当评论中情感词无明显指向的方面时,利用主题图中的上述关系则能够提取隐式方面.

多种隐式方面方法各有利弊,详细对比及优缺点可见表 4.为了增强不同方法的可比性,将各种方法的实验结果依照评估数据集划分,对比详情可见表 5.

Table 4 Comparison of implicit aspect extraction methods

表 4 隐式方面提取方法对比

文献	模型	使用算法	应用语种	应用领域	备注
Yang 等人 ^[57]	基于语义特征	基于语义特征+规则+百科词典	英文	新闻数据	考虑了语义特征,能够自主发现有效特征,无需人工设计
郎君等人 ^[58]	基于语义特征	浅层语义知识+模板特征+百科词典+WordNet	英文	新闻数据	考虑了语义特征及背景知识,充分利用现有词典/语料库资源
Hai 等人 ^[59]	ABOOT	LRTs+LSA+Bootstrapping	中文	手机\酒店评论	利用显式方面与观点词间语义关系实现隐式方面提取
Zeng 等人 ^[60]	基于分类	基于规则+分类器	中文	手机\相机评论	利用显式方面与观点词间语义关系实现隐式方面提取
Sun 等人 ^[61]	基于上下文关系	词共现矩阵+规则	中文	手机\电脑评论	利用显式方面与观点词间语义关系实现隐式方面提取,融入隐式方面候选集及隐式方面打分机制
马京苗等人 ^[62]	基于关联规则	关键规则+模拟退火算法	中文	手机评论	未考虑上下文信息与低频词
Zhou ^[49]	CMiner	基于规则+hashtag+百科词典	中文	博客	方面短语边界识别不佳,模型性能对 NLP 工具依赖性强
Xu 等人 ^[63]	LDA	LDA+SVM	中文	产品评论	充分利用评论主题信息,融合显式方面特征,性能较高
Lau 等人 ^[64]	LDA	LDA+本体挖掘算法	英文	酒店\电影评论	融合半监督模糊产品本体挖掘算法,评价指标待完善
Zhang 等人 ^[57]	基于词共现关系	词共现关系矩阵+修正矩阵	中文	手机\服装评论	无需先验知识与人工标注,模型同样适用于其他语言
Bagheri 等人 ^[66]	基于启发式规则	启发式规则+POS+Bootstrapping	英文	产品评论	领域无关,无需人工标注,适用于多语言任务,可识别多词语方面
Fei 等人 ^[67]	基于词典	词典+分类+动态关系特征	英文	在线词典	多词典共同作用,减少语料库大小的限制
Xia 等人 ^[68]	基于图	主题图	中文	洗衣机评论	主题地图充分利用词语间语义关系,多语言多领域适用.廉价、省力

Table 5 Comparison of implicit aspect extraction methods' evaluation results

表 5 隐式方面提取方法实验结果对比

数据集	文献	实验结果		
		P (%)	R (%)	F
京东电子产品评论	Sun 等人 ^[61]	70.0	62.0	0.66
	Xu 等人 ^[63]	87.4	70.1	0.78
	Xia 等人 ^[68]	81.2	79.4	0.80
Qiu 数据集 ^[55]	Hai 等人 ^[59]	72.9	69.3	0.61
	Lau 等人 ^[64]	71.4	70.0	0.71
ACE-2 Version 1.0 数据集. https://catalog ldc.upenn.edu/LDC2003T11	Yang 等人 ^[57]	80.8	57.4	0.67
	郎君等人 ^[58]	50.4	70.4	0.62
Hu 等人的数据集 ^[40]	Bagheri 等人 ^[66]	87.5	65.0	0.75
	Fei 等人 ^[67]	76.3	77.3	0.77
亚马逊手机/数码相机评论(人工标注)	Zeng 等人 ^[60]	84.9	59.9	0.66
淘宝手机评论(人工标注)	马京苗等人 ^[62]	75.3	68.4	0.72
CSMAE 2012 数据集. http://tcci.ccf.org.cn/conference/2012/pages/page04_eva.html	Zhou ^[49]	-	-	0.58
	Zhang 等人 ^[65]	81.3	79.5	0.80

概括来讲,基于语义特征、观点内容与采用 LDA 的方法需要人工建立规则库,且需事先选择定义特征,相对耗时、耗力,且模型性能对研究人员的操作能力与资源质量有较强的依赖性,但可操作性强,是用于结构简单、主题单一的评论语句;基于词共现关系与基于图的方法无需人工标注,较为省时、省力,但隐式方面提取不全面。因此,较高的查准率与查全率依然是隐式方面提取的目标。

提取观点方面的方法多种多样,但仍存在一些亟待解决的问题。

- 1) 复杂方面短语的挖掘问题。互联网评论形式繁多,有时评论中的方面是由多个词语组成的复杂短语,如“手机屏幕的大小”“手机按键的灵敏度”等,而不仅限于单个词语,此时需要能够完整提取复杂方面短语的模型;同时,获取方面短语中细粒度方面的层次关系也是未来的研究方向之一。
- 2) 方面挖掘的自主性与普适性问题。目前,主流的方面挖掘方法大多依赖于人工标注、语料库等外部资源,且领域依赖性较强。因此,能够自主挖掘方面,大量减少人工干涉,同时提高方法及模型的领域普适性,是在方面提取过程中仍需达到的目标。

3 基于方面的观点内容提取

目前,根据方法使用模型种类,观点内容提取方法主要可分为 3 类:基于规则提取、基于统计模型提取和基于深度模型提取。

3.1 基于规则提取

研究者们最初利用语料库来构建规则,其基本思想是对大量评论语料进行预处理并进行词频统计,然后对统计结果挖掘出一定的频繁特征与规则,利用该规则对大量评论文本进行观点词语提取。随后,基于词典的方法也相继被开发使用,以 WordNet 与 HowNet(www.keenage.com)为代表的观点词典已被众多研究者们广泛使用。早在 2005 年,Smrz^[69]提出了一种使用 WorldNet 词典对文本进行主观内容挖掘。朱嫣岚等人^[70]基于 HowNet 词典提出了基于语义相似度、基于语义相关场的两种计算词语情感倾向性的方法。实验结果表明:该方法在中文常用词汇上有相对较好的效果,词频加权后的判别正确率达到了 80%以上。李钝等人^[71]结合了评论语句中上下文信息,并同样基于 HowNet 词典建立了由情感词组成的种子集,通过计算句中词语与种子词间的相似度来获取这些词的褒贬倾向。该方法主要侧重于词汇的情感倾向性判别。杜伟夫等人^[72]同样基于词语的情感倾向计算与图划分法提出计算词汇情感倾向性的模型。该模型可扩展性较强;同时,将词语的情感极性计算问题转化为一种优化问题。Su 等人^[73]同样将语义关系转换为图表示,用最小切割算法来对词语进行主观性判别。为了解决词语多义性问题,Esuli 等人^[74]提出了一种半监督模型,利用词典中词语的注释信息来进行观点词语的提取与倾向性判别。Turney 等人^[75]利用正、负向两类情感词语库,对近 4 000 个词语(形容词、副词、名词和动词)进行关联统计。二人在关联统计计算过程中提出了两种语义关联计算方法:基于 PMI 计算与基于潜在语义分析计算。部分研究人员采用人工标注的方式获得较多数量的观点词语集,并将此词集中的词语作为种子观点词,利用现有的观点词典对观点词汇集合进行扩充,最终计算该集合与文本中词语的相似度来获取评论语句中的观点词语^[76,77]。由此可推断出:对现有词典进行扩充的方法相对简单易行,但种子观点词的数量与质量直接决定了方法性能的好坏。因此,这些人工构建语料资源虽然可行性较高,但耗时、耗力,这也成为这种方法得以广泛应用的瓶颈。

为了免去大量人工标注工作,不少学者利用句法依存关系来对观点内容进行提取,基于句法依存关系的观点提取基本流程如图 4 所示:首先,从网络获取评论数据,对其进行分句、分词及词性标注等预处理工作;然后进行依存分析,利用人工定义的规则库对分析结果进行(方面,观点词)二元组提取;最后,利用情感词典、程度副词词典及否定词规则对二元组中的观点词进行情感倾向性判别。

Liu 等人^[78]在句法依存分析法的基础上提出了一种基于观点内容中形容词与副词从属关系的假设,并利用 n -gram 窗口围绕方面进行观点提取,在口碑网(beijing.koubei.com)下载 54 208 条评论并取得 60.23% 的召回率,指标高于 Hu 等人^[26]提出的方法。陆浩等人^[79]在 2014 年提出了一种利用句法依存关系分析来对中文微博进行观点倾向分析的方法。该方法首先针对大量微博建立特定领域的方面词汇集合,然后对微博进行分词等预处理

工作,随后进行句法依存关系分析,最后依据设定的规则提取出观点二元组(方面,观点词),进而判断出微博评论的观点倾向性.该模型能够对常见复杂句式、否定词与程度副词进行处理,最终情感分类的准确率可达 86%.在近期工作中,Chinsha 等人^[41]利用 Stanford 句法依存分析器对文本进行观点提取,他们认为,观点内容主要由形容词、动词、副词形容词组合及副词动词组合这 4 类构成,因此,其工作对这 4 种词汇进行提取.最终采用餐厅评论数据用于模型性能的评估,实验结果表示,平均准确率为 78.04%.

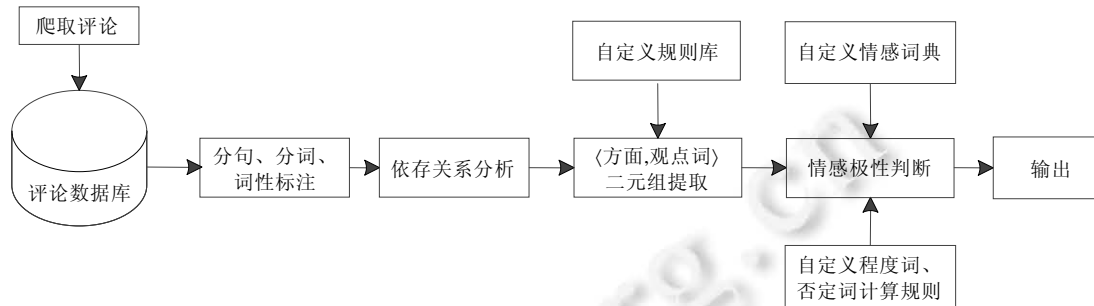


Fig.4 Sentiment analysis flow chart based on syntactic parsing

图 4 基于依存句法分析的观点倾向分析流程图

3.2 基于统计模型提取

Matsuo 等人^[80]在 2002 年提出了基于词共现网络的关键词提取方法.该模型较为简单,同时也忽略了对非频繁词语的提取.Lin 等人^[81]提出了一种联合情感主题模型(joint sentiment-topic,简称 JST)和逆向联合情感主题模型(reverse-JST,简称 RJST),两模型均基于利用 4 层贝叶斯模型改进的 LDA 模型建立,并均能从文本中提取情感分别为正/负向的主题.JST 与 RJST 在 Pang 等人的数据集^[82]及 Blitzer 的数据集^[83]上分别取得了较好的效果,精确率均可达到 76.6%.Duric 等人^[84]提出 3 种不同机制来从文本中提取观点词:隐马尔可夫-潜在狄利克雷分布模型(HMM-LDA)、基于产品方面的语法语义类模型和基于产品特征的最大分值模型.将 3 种方法均应用于 Pang 和 Li 数据集^[17]上,后续分类准确率最高可达 87.5%.考虑到 LDA 模型抽取的主题粒度较粗,无法较为准确地识别观点方面,Zheng^[85]提出了细粒度的主题情感混合模型(fine-grained topic sentiment unification model,简称 FG-TSU),将主题分为局部主题和全局主题,同时引入滑动窗口将词共现信息降到句子级.这种方法不仅能够判别整个文档的情感极性,还能针对细粒度的观点方面获取情感极性.由于标准 LDA 通常选用同类别文档来构建语料库,因此对于混合多文档语料库进行观点挖掘时模型效果并不好.Li 等人^[86]针对这一问题引入文本分类技术,将标准 LDA 模型扩展为 4 层结构:类别-文档-主题-单词,利用 LDA 对混合多类型文档进行观点挖掘.

在大多数统计模型框架中,基于方面的观点内容提取被认作是序列标记问题^[87]:传统的序列标记方法为 B-I-O 标记,B 代表目标片段的开端,I 代表目标片段中剩下的部分,O 则表示原句中不在目标片段中的词.Du 等人^[88]提出了一种用来提取关键片段的新型标记方法:将传统的标记分为直接表达 DES 与间接表达 ESE 两类,即总共 5 类标签:B_DSE,I_DSE,B_ESE,I_ESE 以及 O.CRF 模型^[89]针对情感分析领域的各种序列标注问题(包括观点片段提取问题^[90]、观点评论人鉴别^[91])都取得了不小的成果.Zhang^[92]提出利用 CRF 模型来训练序列模型,从而提取关键片段.Laddha 等人^[93]提出了囊括生成模型与判别模型的混合模型——PSM 模型(phrase sentiment model),无监督生成模型基于大量语料库中的术语搭配对观点内容进行生成建模,带监督判别模型基于 CRF 对评论进行序列建模从而提取观点词组.目前,基于统计模型的关键片段提取技术比较成功的模型依然是基于 CRF 的模型^[94]或加入马尔可夫假设^[95]的条件随机场变体.然而,条件随机场对特征集人工筛选的细致性有较强的依赖性,而这需要许多研究员在工程上的努力.

3.3 基于深度模型提取

近几年来,深度学习对自然语言处理领域的技术发展无疑起到了强有力的推进作用.深度学习模型的关键

优势在于:其能够对大量原始数据中的潜在特征进行自动学习,同时用分布式向量加以表示,而不是人工手动设计特征^[96],这在自然语言处理的众多任务中都远强过基于 CRF 的模型.在众多深度学习方法中,递归神经网络(recursive neural network,简称 RNN)与卷积神经网络(convolutional neural network,简称 CNN)^[97]脱颖而出.庞亮等人^[98]对目前的基于深度学习模型来处理文本序列问题做了综述研究.Socher 等人^[99]提出了递归神经张量网络(recursive neural tensor network,简称 RNTN),首次将短语用分布式向量表示.Yao 等人^[100]将 RNN 应用到命名实体识别任务中,该模型展现了在任务中前所未有的卓越性能.Kalchbrenner 等人^[101]研究出一种动态卷积神经网络(dynamic convolutional neural network,简称 DCNN),用于句子级语义建模.方面级观点挖掘在深度框架中通常被看作是序列标注问题.Huang 等人^[102]利用双向长短时记忆网络(bidirectional-LSTM,简称 BLSTM)与 CRF 模型来进行序列标注工作.该实验主要应用于词性标注、组块分析及命名实体识别任务.Ma 等人^[103]在此基础上添加一层 CNN 网络,从而使模型能够更好地学习英文单词中前缀、后缀等拼写特征.基于 LSTM 框架,Du 等人^[88]探索出一种基于注意力机制的 RNN 模型来进行观点片段提取任务.模型由两层 BLSTM 网络组合而成.该模型能够分析文本中的关键片段,同时能够用标签“DSE”“BSE”标注过的文本进行句中词语关联性评估.在数据集 MPQA^[104]上的实验结果表明:这种模型能够在复杂句上实现关键片段提取任务,召回率可达 74.89%,远高于对比模型——条件随机场模型,网络结构如图 5 所示.

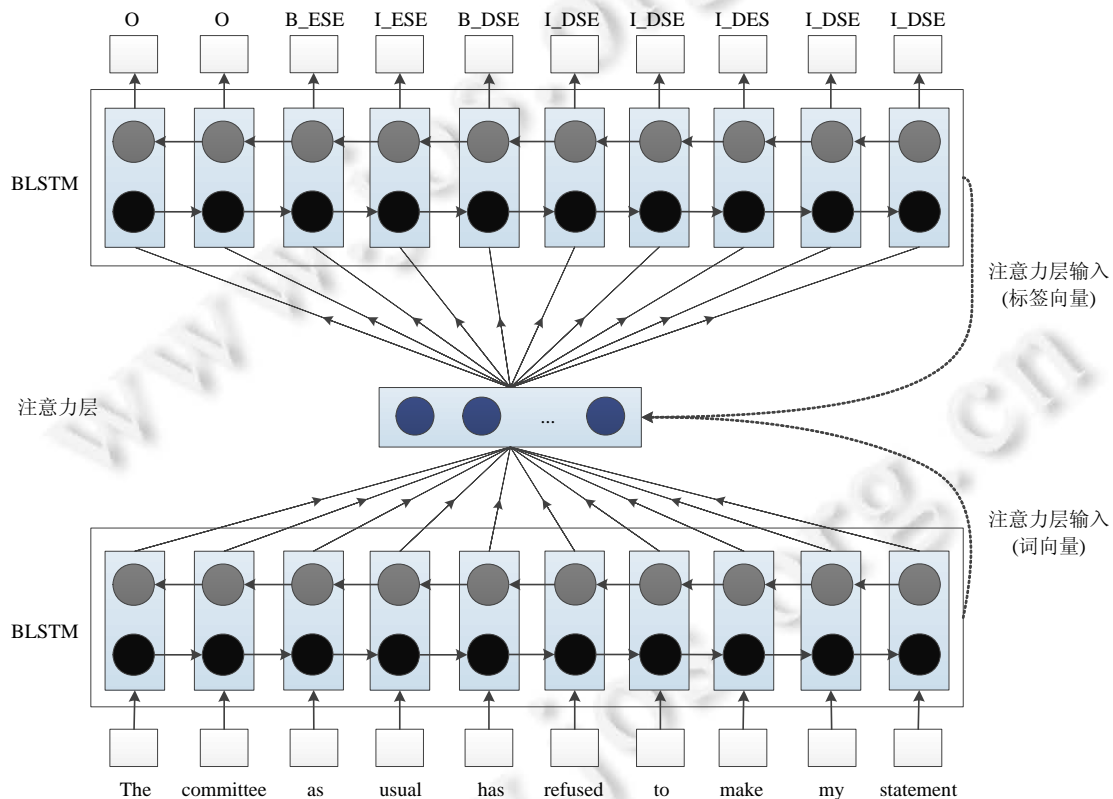


Fig.5 Structure of stacked bidirectional-LSTM based on attention mechanism^[88]

图 5 基于注意力机制的两层双向 LSTM 网络结构^[88]

从上述分析中发现,3 类方法各有利弊,详细说明可见表 6.为了增强不同方法间性能的可比性,将上述各方法的实验结果按评估数据集划分,详情可见表 7.

总体而言,基于规则的提取方法相对简单易行,但大多数基于规则的提取方法结果仅为情感词,而不是观点内容,这样不利于对相对复杂的句子进行观点挖掘;基于统计模型的观点内容提取方法最终精确度较高,但需人

工筛选特征集,且模型对人工选择的细致性有较强的依赖性;由于深度模型的端到端架构,基于深度网络提取的方法体现出自动学习特征的特性,减少大量的人工干预,且精度较高,但美中不足的是,深度模型的性能对网络参数较为敏感,如何快速、有效地调节参数仍是一个令研究者头疼的问题。

Table 6 Comparison of opinion expression extraction methods

表 6 观点内容提取方法对比

方法类	文献	模型	使用算法	观点内容	情感极性	应用语种	应用领域	备注
基于规则提取	朱嫣岚等人 ^[70]	基于语义信息	HowNet+语义相似度/相关场+词频加权	无	含	中文	褒贬词集	简单直观,模型易操作,对基准词依赖性强
	李钝等人 ^[71]	基于语义信息	HowNet+上下文信息+语义相似度	无	含	中文/英文	话题评论	将种子词集分级,种子词应有较强的覆盖面及褒贬倾向
	杜伟夫等人 ^[72]	基于图	图划分法+语义相似度+模拟退火算法	无	含	中文/英文	情感博客, 电脑\电影评论	充分利用词语间全局信息,性能较语义相似度方法提升较大,扩展性高,相似度度量方法不够准确
	Esuli 等人 ^[74]	基于词典	词典+种子词集	含	含	英文	多领域评论	将词汇进行向量表示,模型效果依赖于词典质量
	Turney 等人 ^[75]	SO-A	PMI+LSA+词典	无	含	英文	网页文档	模型易实现,可扩展性强;模型性能受语料库规模的限制
	Kim 等人 ^[76,77]	基于语义关系	基于句法依存关系+词典	含	含	英文	新闻, 话题评论	人工标注少量数据,实现词典的自动扩充,考虑词语间同义关系
	Liu 等人 ^[78]	PORE	基于副词特性+POS+规则	含	含	中文	餐饮评论	有效解决数据稀疏问题,可挖掘潜在观点
	陆浩等人 ^[79]	基于句法与主题扩展	句法分析+HowNet+PMI	无	含	中文	影视\名人\产品评论	引入微博表情符号权值.比较依赖分词与句法分析的结果
	Chinsha 等人 ^[41]	AOMR	基于句法依存关系+POS+评价词词典	含	含	英文	餐饮评论	不适用于包含多个方面词的句子的观点提取
基于统计模型提取	Matsuo 等人 ^[80]	基于词共现	词共现网络+频繁挖掘	含	无	英文	技术类文档	简单,无需语料库,忽略了非频繁观点词
	Lin 等人 ^[81]	JST/Reverse-JST	LDA+贝叶斯	含	含	英文	电影\产品评论	引入领域无关词典,弱监督,模型可移植性强模型中参数的增量式学习能力较弱
	Duric 等人 ^[84]	FS	句法语义+LDA+HMM+最大熵分类	无	含	英文	电影评论	不依赖于特定词性提取,语法语义两类特征均考虑文档级观点挖掘,未针对特定方面提取
	Zheng ^[85]	FG-TSU	LDA+情感词词典	含	含	中文	酒店\手机评论	将主题分层,在 LDA 中添加情感层;依赖情感词典质量
	Li 等人 ^[86]	DC-LDA	LDA	含	无	中文	多领域文档	可用于多领域语料库
	Zhang ^[92]	CRF	CRF	含	无	中文	金融学术文档	能够发现文档内特征,适用于序列标注问题;精度较低
	Laddha 等人 ^[93]	PSM	CRF+GPU 采样	含	含	英文	电子产品评论	领域可移植性强,集成无监督生成模型与带监督判别模型
基于深度模型提取	Kalchbrenner 等人 ^[101]	DCNN	动态卷积神经网络+k-max pooling	无	含	英文	电影微博评论	不依赖句法信息,适用于多语言,适应变长句子输入,无需引入其他特征,可理解语义关系
	Du 等人 ^[88]	BLSTM	BLSTM+注意力机制	含	无	英文	新闻评论	有效评估词语间语义关系,强调句中重点成分,适应复杂句输入,模型评估指标有待完善

Table 7 Comparison of opinion expression extraction methods' evaluation results

表 7 观点内容提取方法实验结果对比

数据集	文献	实验结果		
		P (%)	R (%)	F
Qiu 的数据集 ^[55]	Hu 等人 ^[40]	56.1	41.5	0.48
	Chinsha 等人 ^[41]	83.0	89.2	0.86
	Liu 等人 ^[78]	56.8	60.2	0.59
MPQA 数据集 ^[104]	Kim 等人 ^[76,77]	61.0	85.0	0.71
	Du 等人 ^[88]	67.8	74.9	0.71
Pang 等人的数据集 ^[82]	Lin 等人 ^[81]	75.6	80.1	0.78
	Duric ^[84]	87.5	76.9	0.82
Turney 等人的数据集 ^[75]	Turney 等人 ^[75]	85.8	81.3	0.83
	Esuli 等人 ^[74]	87.0	82.1	0.84
HowNet 褒贬词集. http://www.keenage.com/	朱嫣岚等人 ^[70]	78.6	76.1	0.77
新浪微博“超级女声”评论(人工标注)	李钝等人 ^[71]	75.7	72.0	0.74
电影/笔记本评论,情感博客(人工标注)	杜伟夫等人 ^[72]	-	-	-
新浪微博影视/名人/产品评论(人工标注)	陆浩等人 ^[79]	83.1	85.9	0.86
学术文章(人工标注)	Matsuo 等人 ^[80]	51.0	-	-
酒店/手机评论(人工标注)	Zheng ^[85]	81.2	72.7	0.78
SougouC.Mini 中文语料库. http://download.csdn.net/download/qianxun168168/1835507	Li 等人 ^[86]	83.1	72.6	0.78
中国人民大学社科信息中心数据库. http://art.zlzx.org/	Zhang ^[92]	66.4	42.0	0.51
亚马逊电子产品评论(人工标注)	Laddha 等人 ^[93]	80.3	72.8	0.76
Stanford Sentiment Treebank 电影评论 ^[99]	Kalchbrenner 等人 ^[101]	-	-	-

观点内容的提取研究愈发受到领域学者们的关注,但仍存在一些问题值得研究者们去进一步改进与提高.

- 1) 观点内容提取模型的领域普适性.目前,主流的观点内容提取方法大多是领域相关的,对于新领域的可移植性较差,模型的可扩展性与多领域可移植性仍需进一步提高.
- 2) 观点内容边界识别问题.在众多观点内容提取方法中,研究者们通常会头疼于观点内容边界的无法精准识别与结果的无法准确评估问题.更为合理、适用于相关领域的识别观点内容边界的策略需要研究者们进一步去探索挖掘.
- 3) 复杂句的观点内容提取问题.互联网评论形式多种多样,如何完美地应对结构不整齐的复杂句(如包含网络用语、习语或隐喻观点等复杂句),仍是领域中值得分析与研究的重要问题.

4 方面级观点挖掘的评价与应用

4.1 评价方法

方面级观点挖掘中,两个重要的步骤为观点句中的方面提取与基于方面的观点内容提取,不同步骤中需分别利用不同的评价方法来评估性能.下面分别对方面提取及观点内容提取常用的评估方法与指标进行总结.

4.1.1 方面提取评价方法

观点句中的方面通常为单个词的形式,因此,通常选用直接与标注对比计数的方式;同时,选取精确率(precision)、召回率(recall)和 $F1$ 度量值来评价模型整体性能.为了方便后续观点内容提取评价指标的描述,这里对 Precision, Recall 和 $F1$ 采用通用的描述形式.令 C 与 P 分别表示人工标注的真正短语集合与模型输出短语集合, c 与 p 分别表示每个标注短语与模型相应输出短语.传统方法中 Precision, Recall 和 $F1$ 的计算方法可依次表示为(在此处的方面提取中,由于目标短语与输出短语通常为单个词,因此, C 与 P 可分别看作方面单词的集合):

$$Precision = \frac{|\{p | p \in P \wedge \exists c \in C \text{ s.t. } \alpha(c, p)\}|}{|P|},$$

$$Recall = \frac{|\{c | c \in C \wedge \exists p \in P \text{ s.t. } \alpha(c, p)\}|}{|C|},$$

$$F1 = \frac{2Precision \cdot Recall}{Precision + Recall},$$

其中,s.t. $\alpha(c,p)$ 表示约束条件:标签短语 c 与输出短语 p 完全匹配(短语间的完全匹配意为两个短语中的单词逐个对应相同,此处可看作是单词相同),“| \cdot ”表示计算符合要求的短语数量.

当评论数据中方面存在层次关系,即待提取的方面为复杂短语(如“空调遥控器的按键”“空调的制冷效果”等)时,我们可将方面提取看作与观点提取相同步骤处理,详见第 4.1.2 节.根据应用场合的不同,有时待提取的方面被赋予了重要等级,从而在结果中更加关注模型提取出的那些频次较高的方面.此时,在评估模型阶段可以采用信息检索中常用的 Precision@ N 等评价指标.这里考虑进了位置因素,Precision@ N 表示检测前 N 条数据的准确率,通过计算前 N 条数据结果的算数平均值来评判该模型的整体性能.

4.1.2 观点内容提取评价方法

在传统度量方法中,每个预测短语与真正标注短语的一次完全匹配被看作一次预测正确,也称为精确评价方式.然而在提取观点内容时,提取结果中文本表达正确的界限往往难以定义,人工识别也有一定的困难^[104],如模型往往会挖掘出一些与真正标签相近却不完全相同的表达,例如标签“严厉地批评”,模型可能只提取出“批评”.这样既不是完全错误,也不算绝对正确.由此可见,采用传统的度量方法来评估模型性能并不科学.

因此,研究人员采用两种基于传统度量指标改进后的软性评估方法——二元重叠方法(binary overlap)和比例重叠方法(proportional overlap)来评估观点挖掘任务.Binary Overlap 方法首次由 Breck 等人^[94]提出.该方法认为:每个预测值与真正值间若存在局部匹配,则作为一次预测正确,也成为覆盖评价;Proportional Overlap^[105]则是按比例匹配计数.对于这两种方法的详细描述如下.

Binary Overlap 在传统 Precision,Recall 及 $F1$ 的基础上加入了弹性制度,精确率、召回率的计算方法可重新依次表示为($F1$ 的计算方法同第 4.1.1 节)

$$Precision^b = \frac{|\{p | p \in P \wedge \exists c \in C \text{ s.t. } \alpha'(c,p)\}|}{|P|},$$

$$Recall^b = \frac{|\{c | c \in C \wedge \exists p \in P \text{ s.t. } \alpha'(c,p)\}|}{|C|},$$

其中,s.t. $\alpha'(c,p)$ 表示约束条件:标签短语 c 与输出短语 p 存在重叠.“| \cdot ”表示计算符合要求的短语数量.显然,此方法易对模型评估过高.Proportional Overlap 则相对严格,定义片段 c 与片段 p 间的重叠比例 $r(c,p)$:

$$r(c,p) = \frac{|c \cap p|}{|p|},$$

其中,“| \cdot ”表示符合要求的字符数量.定义测试集上的重叠比例为

$$R(C,P) = \sum_{c \in C} \sum_{p \in P} r(c,p).$$

因此,Proportional Overlap 方法中的精确率、召回率的计算方法可重新依次表示为($F1$ 的计算方法同第 4.1.1 节)

$$Precision^p = \frac{R(C,P)}{|P|},$$

$$Recall^p = \frac{R(P,C)}{|C|}.$$

表 8 对比了不同的评价指标.整体来看,传统指标要求预测与真实标签完全匹配,最为严格,容易对模型评估过低,但方面提取中由于标签与预测值通常为单个词,因此,该方法较为适用;而 Binary Overlap 认为预测值与标签存在局部匹配则看作一次预测正确,指标较为宽松,易对模型评估过高;Proportional Overlap 依据预测值与标签间重合比例来计算各项指标,在提取观点内容任务中,该方法相对合理与严谨,也更为通用.在数据量较小的情况下,许多工作采用人工校对的方式来评估模型,选取多位领域专家来为模型结果进行评判打分,采用该方法的评估结果最为可靠、实际.但过于耗时、耗力,数据量过大的情况下往往并不易实现.

Table 8 Comparison of evaluation metrics

表 8 评价指标对比

评价类型	评价思想	方面提取	观点提取	优点	缺点
传统指标	每个预测短语与标签短语的完全匹配被认作一次预测正确	适用	不适用	计算简单,适用于方面提取	指标严格,对模型评估过低
二元重叠法	每个预测短语与标签短语存在字符重叠被认作一次预测正确	适用	不适用	计算简单	对模型评估过高
比例重叠法	依据预测短语与标签短语间字符重合比例来计算指标	适用	适用	合理、严谨,更为通用	计算相对繁琐
人工校对	多位领域专家人工校对,算术平均求整体指标	偶尔适用	偶尔适用	评估结果最可靠	耗时、耗力

4.2 应用场景

利用对互联网评论信息的方面级观点挖掘结果,叶强^[106]等人发现:对于商品评论,消费者可以了解其他用户的态度倾向,做出更好的购买决策;销售商和生产商可以获得用户对其商品和服务的反馈,了解用户对自己和对竞争对手的评价,从而改进产品、改善服务,获得竞争优势.因此,对网络评论进行方面级观点挖掘,在实际生活中具有很高的实用价值.目前为止,基于观点挖掘技术已开发出不少成型的系统可供使用,并已应用于现实生活中.如:Dave 等人研究并开发的 Review Seer 是世界上首个情感分析工具,同时也是首个通过分析给定产品评论能够判别其褒贬性的系统^[107];Gamon 等人研究开发的 Pulse 系统能够自动从汽车评论中挖掘相应的褒贬信息以及强弱程度^[108];Liu 等人开发的 Opinion Observer 系统可以挖掘在线的消费者产品评价^[109],同时对涉及产品的各种方面的优缺点进行统计,并采用可视化的方式对若干产品方面的综合质量进行比对.此外,更有面向多类型数据、多领域数据的开放性观点挖掘研究平台——Web Fountain.Yi 等人在此基础上开发了应用于 WebFountain 系统中的观点挖掘器^[110].Wilso 等人研究并开发了一个能够自主判别句子主观性同时提取句子中与主观性相关成分(如观点持有者、主观观点内容及观点情感等)的系统——Opinion Finder^[111].

5 方面级观点挖掘可用数据资源

5.1 语料库

可用于方面级观点挖掘的数据资源尤为重要.现将常用的可应用于该领域的语料库资源进行综述总结,同时附上资源链接,见表 9.

Table 9 Summary and comparison of corpora resources for opinion mining

表 9 观点挖掘语料库总结与对比

语料库名称	语言	方面标注	观点标注	情感极性分类	情感程度	备注
MPQA 语料库 ^[104]	英文	含	含	正、负两类	含	每句标注了观点持有者、观点对象、主观表达式及其情感极性与强度;适合于新闻评论领域任务的研究. http://mpqa.cs.pitt.edu/corpora/mpqa_corpus/
Amazon 产品评论消极语料 ^[112]	英文	含	含	仅负类	无	包含网站中耳机、键盘、GPS、鼠标、路由器等产品的 1~3 星评论. http://arjun3.rcc.uh.edu/dumps/issue_sent_labeled_data.Zip
细粒度汽车评论标注语料	中文	含	含	正、负两类	无	详细标注了用户评论中的评价对象和评价特征.可用于情感计算、评论抽取、口碑分析等用途.以 xml 的形式提供.数据量有 160M 左右. http://www.datatang.com/detail/86
SE-ABSA15 语料库 ^[113]	英文	含	无	正、负两类	无	含方面名词短语标注,含笔记本、饭店、酒店等 3 类产品语料. http://metashare.ilsp.gr:8080/
康奈尔大学 (Cornell) 影评数据集 ^[17]	英文	无	无	正、负两类	无	由影评组成,其中,持肯定和否定态度的各 1 000 篇;标注了褒贬极性的句子各 5 331 句,标注主客观标签的句子各 5 000 句. http://www.cs.cornell.edu/people/pabo/movie-review-data/

Table 9 Summary and comparison of corpora resources for opinion mining (Continued)

表 9 观点挖掘语料库总结与对比(续)

语料库名称	语言	方面标注	观点标注	情感极性分类	情感程度	备注
Hu 等人的产品领域评论语料 ^[40]	英文	含	无	正、负两类	含	主要包括从亚马逊和 Cnet 下载的 5 种电子产品的网络评论.这些语料按句子为单元详细标注了评价对象,情感句的极性 & 强度等信息. http://www.cs.uic.edu/~liub/FBS/sentiment-analys
多角度餐馆评论语料 ^[113]	英文	含	无	多类	含	麻省理工学院(MIT)构建的多角度餐馆评论语料 ^[114] .共 4 488 篇,每篇语料按 5 个角度(饭菜、环境、服务、价钱、整体体验). http://pages.cs.wisc.edu/~bsnyder/data/naac107
Amazon 多领域情感分析语料库 ^[83]	英文	无	无	正、负两类	含	囊括多领域海量商品评论;包含用户评分;大小为 1.53G. https://www.cs.jhu.edu/~mdredze/datasets/sentime
酒店评论语料	中文	无	无	正、负、平 3 类	无	谭松波整理的大规模酒店评论语料.规模 10 000 篇.携程网自动采集.含正、负、平这 3 类情感语料. http://www.datatang.com/data/11936
豆瓣网影评情感测试语料	中文	无	无	多类	含	来自豆瓣网对电影《ICE AGE3》的评论.含打分标注(5 星).共计 11 323 条评论. http://www.datatang.com/data/13539
酒店、电脑与书籍评论语料	中文	无	无	正、负两类	无	含酒店、电脑、书籍这 3 类评论语料,未去重. http://www.datatang.com/data/11937

目前来看,同时包含方面及观点内容的语料资源相对稀缺^[115],且在以下 3 个著名的语料库中:MPQA 语料库^[104]倾向于句子整体性观点表达与观点对象的标注,而不是方面级观点表达标注;Arjun 提供的人工标注语料^[112]形式比较吻合方面级观点挖掘的任务需求,但仅包含消极情感语料;细粒度汽车评论标注语料相对更为符合任务的需求,但语料领域相对单一,不太适合跨领域观点挖掘任务.因此,方面级观点挖掘语料的建设还需完善.

5.2 词 典

词典资源在众多观点挖掘的任务中都起着十分重要的作用.表 10 总结与对比了 NLP 中应用广泛的一些情感词典,并给出了相应的资源链接.

Table 10 Summary and comparison of lexicon resources for opinion mining

表 10 观点挖掘词典总结与对比

词典名称	语言	情感极性分类	情感程度	备注
Hu 等人的情感词典 ^[40]	英文	正、负两类	无	含 4 783 个消极词汇与 2 006 个积极词汇. http://www.cs.uic.edu/~liub/FBS/sentiment-analys
MPQA 主观性词典 ^[104]	英文	正、负两类	含	含 8 222 词语及其主观性(强\弱)、词性和情感极性. http://mpqa.cs.pitt.edu/lexicons/subj_lexicon
SentiWordNet ^[116]	英文	正、负、平 3 类	含	每个词语有正、负、平这 3 项情感打分.分值总和相加为 1. http://sentiwordnet.isti.cnr.it/
GI(General inquirer) 评价词典 ^[117]	英语	正、负两类	含	收集了 1 914 个褒义词和 2 293 个贬义词,并为每个词语按照极性、强度、词性等进行标注. http://www.wjh.harvard.edu/~inquirer/
HowNet 词典 ^[118]	中文+英文	正、负两类	无	包含 9 193 个中文评价词语/短语、9 142 个英文评价词语/短语,并被分为褒贬两类. http://www.keenage.com/html/e_index.html
NTUSD 评价词典 ^[119]	繁+简体中文	正、负两类	无	由台湾大学收集,含有 2 812 个褒义词与 8 276 个贬义词. http://academiasinicanlplab.github.io/
LIWC ^[120]	英文	正、负两类	无	提供了带情感极性标注的情感表达常用词. http://liwc.wpengine.com
台湾大学情感极性词典	中文	正、负两类	无	包括 2 810 个正极性词语和 8 276 个负极性词语. http://www.datatang.com/data/11837
Xsimilarity 情感词典	中文+英文	正、负两类	含	包含情感极性、情感程度、主观性标注. http://www.datatang.com/data/11851
OpinionFinder 情感词典 ^[111]	英文	正、负两类	含	主观词语来自 OpinionFinder 系统,该词典含有 8 221 个主观词,并为每个词语标注了词性、情感极性及其情感程度. http://more.datatang.com/data/11838

此外,Wang 等人^[121]对历年来的情感词典自动构建方法进行了较为全面的综述,主要罗列了目前已有的情感词典资源,对比分析了若干情感词典构建方法的优缺点,并指出了目前构建情感词典方法中依然存在的问题与未来发展的方向.评价词典大多与领域无关,由于有时相同或相近的评价词在不同领域中所表达的情感极性并不一致,可能导致在跨领域观点判别任务中出错.因此,在目前所拥有的词典基础上建立独特领域的情感词典也是十分必要的.

6 面临的挑战与研究方向

6.1 面临的挑战

近年来,方面级观点挖掘任务获得了越来越多的研究者的关注,新的研究成果也在不断涌现.同时我们也应该认识到,该领域目前还处于学术研究阶段,方面级观点挖掘仍然面临着许多挑战,主要体现在以下几点.

- 1) 语料库不够丰富和规范.现有的用于方面级观点挖掘任务的语料资源不够丰富和全面,文献中的数据与方法多用于英文数据,中文数据支持度不够^[122];且现有语料库针对方面级观点挖掘任务的标注也有待规范,领域不够广泛,这直接降低了互联网评论观点挖掘的识别精度^[123].
- 2) 数据降噪、指代消解问题.网络评论由大量用户生成,评论质量与风格差异较大,错别字和垃圾信息比比皆是,评论中也经常出现连续分句对一个方面进行评价的情况,并用指示代词代替该方面,这给观点挖掘任务带来了一定的困难.文本预处理是观点挖掘中十分重要的一环,目前的方法还并不能完美应对各种冗杂数据,因此需要进一步完善数据降噪、指代消解等文本预处理技术,从而提高模型的整体效果.
- 3) 复杂句的方面级观点挖掘与理解问题.互联网用户多种多样,评论风格也大相径庭.有些互联网评论句中含有观点词,但整体并未表达出一定的情感;有些句不包含观点词,但却也表达出一定的情感.这些针对句意较难捕捉的复杂句的方面级观点挖掘工作对于领域中众多学者来说仍是不小的挑战;此外,对含反语、疑问句及双重否定句等句式的处理,同样值得研究者们进一步探究与挖掘.

6.2 未来的研究方向

根据对国内外研究现状的了解与分析,我们认为,未来值得关注的研究方向主要有以下几点.

- 1) 方面层次关系的自动获取.观点句中的方面即为观点内容所评价的细粒度对象,但不少句中观点内容指向具有层次关系或者从属关系的两个对象,如电脑包含屏幕,而屏幕又包含大小、清晰度等更细粒度的属性,因此,需要能够自动识别对象间层次关系的方法来帮助提高整个模型的有效性.
- 2) 复杂句观点内容挖掘的准确性.互联网评论用户人多且杂,表达观点的文字风格及所用句式也大不相同.在句子结构上,如反语、疑问句、复杂从句或成分嵌套等句式,均需要有效而准确的观点内容提取模式来进行识别.在句子内容上,如何准确地获取包含歇后语、具有时效性的网络用语等新型、偏门词汇的观点内容,同样是研究者们下一步需探究的目标之一.
- 3) 方面级观点挖掘模型的领域普适性.现有的方面级观点挖掘模型大多聚集在互联网产品评论区,且观点挖掘模型往往依赖于评论领域,而博客、短消息等信息源也包含大量的观点信息,因此同样值得挖掘分析.因此,模型和方法是否具有普适性、能否构建有效的通用模型,都是该领域未来值得研究的问题.

7 总结

随着社交网络、电商网站等平台的飞速发展,互联网所承载的观点评论日益增长,观点挖掘也愈发成为 NLP 领域关注的问题.对网络评论进行方面级观点挖掘,不仅能够帮助人们更充分地了解到其所关心实体的各项特征,进而做出选择或购买决策,也能帮助政府部门及商家依据实时的舆论导向或商品评价做出相应的调整.因此,对网络评论进行方面级观点挖掘研究具有十分重要的理论价值及应用价值.目前,研究者已在此方面做了

大量的工作,本文对近年来该领域国内外的主要研究成果进行了回顾与总结,综述了观点挖掘中方面提取、观点提取等主要方法的研究现状.同时,对比分析了观点挖掘中常用的评价指标,并罗列和总结了方面级观点挖掘常见的应用场景及相关可用资源.本篇综述能够帮助该领域初学者较为快速、全面地认识与了解方面级观点挖掘的整体情况,同时能够为该领域今后的研究提供一定的借鉴与指导意义.

References:

- [1] Hatzivassiloglou V, Mckeown KR. Predicting the semantic orientation of adjectives. In: Proc. of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conf. of the European Chapter of the Association for Computational Linguistics (ACL/EACL'97). Stroudsburg: ACL, 1997. 174–181. [doi: 10.3115/976909.979640]
- [2] Yao TF, Cheng XY, Xu FY, Han W, Wang R. A survey of opining mining for texts. *Journal of Chinese Information Processing*, 2008,22(3):71–80 (in Chinese with English abstract).
- [3] Wang H, Wang HY, Zuo WL. Survey on opinion mining. *Application Research of Computers*, 2009,26(1):25–29 (in Chinese with English abstract).
- [4] Chen M, Zhu FW, Wu MH, Ying J. Survey of opining mining. *Journal of Zhejiang University (Engineering Science)*, 2014, 48(8):1461–1472 (in Chinese with English abstract).
- [5] Pang B, Lee L. Opinion mining and sentiment analysis. *Foundations & Trends in Information Retrieval*, 2008,2(1-2):1–135. [doi: 10.1561/1500000011]
- [6] Liu B. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 2012,5(1):1–167. [doi: 10.2200/S00416ED1V01Y201204HLT016]
- [7] Ghag K, Shah K. Comparative analysis of the techniques for sentiment analysis. In: Proc. of the 2013 Int'l Conf. on Advances in Technology and Engineering (ICATE 2013). Washington: IEEE Computer Society, 2013. 1–7. [doi: 10.1109/ICAdTE.2013.6524752]
- [8] Ravi K, Ravi V. A survey on opinion mining and sentiment analysis. *Knowledge-Based Systems*, 2015,89(C):14–46. [doi: 10.1016/j.knosys.2015.06.015]
- [9] Bhatia S, Sharma M, Bhatia KK. Strategies for mining opinions: A survey. In: Proc. of the 2nd Int'l Conf. on Computing for Sustainable Global Development (INDIACom 2015). Washington: IEEE Computer Society, 2015. 262–266.
- [10] Guellil I, Boukhalfa K. Social big data mining: A survey focused on opinion mining and sentiments analysis. In: Proc. of the 12th Int'l Symp. on Programming and Systems (ISPS). Washington: IEEE Computer Society, 2015. 1–10. [doi: 10.1109/ISPS.2015.7244976]
- [11] Rana TA, Cheah YN. Aspect extraction in sentiment analysis: Comparative analysis and survey. *Artificial Intelligence Review*, 2016,46(4):459–483. [doi: 10.1007/s10462-016-9472-z]
- [12] Sun S, Luo C, Chen J. A review of natural language processing techniques for opinion mining systems. *Information Fusion*, 2017, 36:10–25. [doi: 10.1016/j.inffus.2016.10.004]
- [13] Li JH, Liu GS, Lin X. Survey on sentiment orientation analysis and its applications. *Journal of Cyber Security*, 2017,2(2):48–62 (in Chinese with English abstract).
- [14] Chen QH, Sun CH, Jia YB. Overview of text data mining techniques. *Industrial Control Computer*, 2017,30(2):94–95,102 (in Chinese with English abstract).
- [15] Chen XM, Wang FG, Wu HW, Sun ZQ. Social network comments based on opinion mining and application. *Information Science*, 2013,31(11):119–124 (in Chinese with English abstract).
- [16] Kim SM, Hovy E. Determining the sentiment of opinions. In: Proc. of the 20th Int'l Conf. on Computational Linguistics (ICCL 2004). Stroudsburg: ACL, 2004. 1367. [doi: 10.3115/1220355.1220555]
- [17] Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques. In: Proc. of the ACL-02 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2002). Stroudsburg: ACL, 2002. 79–86. [doi: 10.3115/1118693.1118704]
- [18] Turney PD. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: Proc. of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002). Stroudsburg: ACL, 2002. 417–424. [doi: 10.3115/1073083.1073153]

- [19] Barbosa L, Feng J. Robust sentiment detection on Twitter from biased and noisy data. In: Proc. of the 23rd Int'l Conf. on Computational Linguistics (COLING 2010). Stroudsburg: ACL, 2010. 36–44.
- [20] Jiang L, Yu M, Zhou M, Liu X, Zhao T. Target-Dependent Twitter sentiment classification. In: Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011). Stroudsburg: ACL, 2011. 151–160.
- [21] Speriosu M, Sudan N, Upadhyay S, Baldrige J. Twitter polarity classification with label propagation over lexical links and the follower graph. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP 2011). Stroudsburg: ACL, 2011. 53–63.
- [22] Valakunde ND, Patwardhan MS. Multi-Aspect and multi-class based document sentiment analysis of educational data catering accreditation process. In: Proc. of the Int'l Conf. on Cloud & Ubiquitous Computing & Emerging Technologies (CUBE 2013). Washington: IEEE Computer Society, 2013. 188–192. [doi: 10.1109/CUBE.2013.42]
- [23] Karamibekr M, Ghorbani AA. Sentence subjectivity analysis in social domains. In: Proc. of the 2013 IEEE/WIC/ACM Int'l Joint Conf. on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). Washington: IEEE Computer Society, 2013. 268–275. [doi: 10.1109/WI-IAT.2013.39]
- [24] Wiebe J, Wilson T, Bruce R, Bell M, Martin M. Learning subjective language. Computational Linguistics, 2004,30(3): 277–308. [doi: 10.1162/0891201041850885]
- [25] Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis. Int'l Journal of Computer Applications, 2005,7(5):347–354. [doi: 10.3115/1220575.1220619]
- [26] Jin W, Ho HH. A novel lexicalized HMM-based learning framework for Web opinion mining. In: Proc. of the 26th Annual Int'l Conf. on Machine Learning (ICML 2009). New York: ACM Press, 2009. 465–472. [doi: 10.1145/1553374.1553435]
- [27] Jin W, Ho HH, Srihari RK. Opinion miner: A novel machine learning system for Web opinion mining and extraction. In: Proc. of the 15th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2009). New York: ACM Press, 2009. 1195–1204.
- [28] Zhang S, Jia WJ, Xia YJ, Meng Y, Yu H. Opinion analysis of product reviews. In: Proc. of the 6th Int'l Conf. on Fuzzy Systems and Knowledge Discovery (FSKD 2009), Vol.2. Washington: IEEE Computer Society, 2009. 591–595. [doi: 10.1109/FSKD.2009.200]
- [29] Xu B, Zhao TJ, Wang SY, Zheng DQ. Extraction of opinion targets based on shallow parsing features. Acta Automatica Sinica, 2011,37(10):1241–1247 (in Chinese with English abstract).
- [30] Hofmann T. Probabilistic latent semantic indexing. In: Proc. of the 22nd Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR'99). New York: ACM Press, 1999. 50–57. [doi: 10.1145/312624.312649]
- [31] Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. Journal of Machine Learning Research, 2003,3:993–1022.
- [32] Hu Y, Boyd-Graber J, Satinoff B, Smith A. Interactive topic modeling. Machine Learning, 2014,95(3):423–469. [doi: 10.1007/s10994-013-5413-0]
- [33] Chen Z, Liu B. Mining topics in documents: Standing on the shoulders of big data. In: Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2014). New York: ACM Press, 2014. 1116–1125. [doi: 10.1145/2623330.2623622]
- [34] Mcauliffe JD, Blei DM. Supervised topic models. In: Proc. of the Advances in Neural Information Processing Systems. 2008. 121–128.
- [35] Kessler J, Nicolov N. Targeting sentiment expressions through supervised ranking of linguistic configurations. In: Proc. of the 3rd Int'l AAI Conf. on Weblogs and Social Media (ICWSM 2009). Menlo Park: AAI Press, 2009. 90–97.
- [36] Jiang P, Zhang C, Fu H, Niu Z, Yang Q. An approach based on tree kernels for opinion mining of online product reviews. In: Proc. of the 10th IEEE Int'l Conf. on Data Mining (ICDM 2010). Washington: IEEE Computer Society, 2010. 256–265. [doi: 10.1109/ICDM.2010.104]
- [37] Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 2011,12(1):2493–2537.
- [38] Poria S, Cambria E, Gelbukh A. Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis. In: Proc. of the 2015 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2015). Stroudsburg: ACL, 2015. 2539–2544. [doi: 10.18653/v1/D15-1303]

- [39] Poria S, Cambria E, Gelbukh A. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 2016,108(C):42–49. [doi: 10.1016/j.knosys.2016.06.009]
- [40] Hu M, Liu B. Mining and summarizing customer reviews. In: *Proc. of the 10th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2004)*. New York: ACM Press, 2004. 168–177. [doi: 10.1145/1014052.1014073]
- [41] Chinsha TC, Joseph S. A syntactic approach for aspect based opinion mining. In: *Proc. of the 9th IEEE Int'l Conf. on Semantic Computing (ICSC 2015)*. Washington: IEEE Computer Society, 2015. 24–31. [doi: 10.1109/ICOSC.2015.7050774]
- [42] Popescu AM, Etzioni O. Extracting product features and opinions from reviews. In: *Proc. of the Conf. on Human Language Technology and Empirical Methods in Natural Language Processing (HLT-EMNLP 2007)*. Stroudsburg: ACL, 2007. 9–28. [doi: 10.3115/1220575.1220618]
- [43] Liu HY, Zhao YY, Qin B, Liu T. Comment target extraction and sentiment classification. *Journal of Chinese Information Processing*, 2010,24(1):84–88 (in Chinese with English abstract).
- [44] Yi J, Nasukawa T, Bunescu R, Niblack W. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In: *Proc. of the 3rd IEEE Int'l Conf. on Data Mining (ICDM 2003)*. Washington: IEEE Computer Society, 2003. 427–434. [doi: 10.1109/ICDM.2003.1250949]
- [45] Samha AK, Li Y, Zhang J. Aspect-Based opinion extraction from customer reviews. *Int'l Journal of Computer Science & Information Technolo*, 2014,4(4):149–160.
- [46] Zhu J, Wang H, Zhu M, Tsou BK, Ma M. Aspect-Based opinion polling from customer reviews. *IEEE Trans. on Affective Computing*, 2011,2(1):37–49. [doi: 10.1109/T-AFFC.2011.2]
- [47] Bagheri A, Saraee M, Jong FD. An unsupervised aspect detection model for sentiment analysis of reviews. *Lecture Notes in Computer Science*, 2013,7934:140–151. [doi: 10.1007/978-3-642-38824-8_12]
- [48] Quan C, Ren F. Unsupervised product feature extraction for feature-oriented opinion determination. *Information Sciences*, 2014, 272(C):16–28. [doi: 10.1016/j.ins.2014.02.063]
- [49] Zhou X, Wan X, Xiao J. CMiner: Opinion extraction and summarization for Chinese microblogs. *IEEE Trans. on Knowledge and Data Engineering*, 2016,28(7):1650–1663. [doi: 10.1109/TKDE.2016.2541148]
- [50] Lu Y, Zhai C. Opinion integration through semi-supervised topic modeling. In: *Proc. of the 17th Int'l Conf. on World Wide Web (WWW 2008)*. New York: ACM Press, 2008. 121–130. [doi: 10.1145/1367497.1367514]
- [51] Andrzejewski D, Zhu X, Craven M. Incorporating domain knowledge into topic modeling via Dirichlet forest priors. In: *Proc. of the 26th Annual Int'l Conf. on Machine Learning (ICML 2009)*. New York: ACM Press, 2009. 25–32. [doi: 10.1145/1553374.1553378]
- [52] Mukherjee A, Liu B. Aspect extraction through semi-supervised modeling. In: *Proc. of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1 (ACL 2012)*. Stroudsburg: ACL, 2012. 339–348.
- [53] Wang T, Cai Y, Leung H, Lau RYK, Li Q, Min H. Product aspect extraction supervised with online domain knowledge. *Knowledge-Based Systems*, 2014,71(1):86–100. [doi: 10.1016/j.knosys.2014.05.018]
- [54] Liu Q. Research on approaches to opinion target extraction in opinion mining [Ph.D. Thesis]. Nanjing: Southeast University, 2016 (in Chinese with English abstract).
- [55] Qiu G, Liu B, Bu J, Chen C. Opinion word expansion and target extraction through double propagation. In: *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*. 2011. 9–27. [doi: 10.1162/coli_a_00034]
- [56] Zhao J, Xu H, Huang X, Tan S, Liu K, Zhang Q. Overview of Chinese opinion analysis evaluation 2008. In: *Proc. of the 1st Chinese Opinion Analysis Evaluation (COAE 2008)*. Beijing: The Professional Committee of Information Retrieval, 2008. 1–20.
- [57] Yang X, Su J. Coreference resolution using semantic relatedness information from automatically discovered patterns. In: *Proc. of the 45th Annual Meeting of the Association for Computational Linguistics (ACL 2007)*. Stroudsburg: ACL, 2007. 528–535.
- [58] Lang J, Xin Z, Qin B, Liu T, Li S. Coreference resolution with integrated multiple background semantic knowledge. *Journal of Chinese Information Processing*, 2009,23(3):3–9 (in Chinese with English abstract).
- [59] Hai Z, Chang K, Cong G, Yang CC. An association-based unified framework for mining features and opinion words. *ACM Trans. on Intelligent Systems and Technology (TIST)*, 2015,6(2):1–21. [doi: 10.1145/2663359]

- [60] Zeng L, Li F. A classification-based approach for implicit feature identification. In: Song SM, Min Z, eds. Proc. of the 12th China National Conf. on Computational Linguistics (CCL 2013). Berlin, Heidelberg: Springer-Verlag, 2013. 190–202. [doi: 10.1007/978-3-642-41491-6_18]
- [61] Sun L, Li S, Li JY, Lv JT. A novel context-based implicit feature extracting method. In: Proc. of the 2014 Int'l Conf. on Data Science and Advanced Analytics (DSAA 2014). Washington: IEEE Computer Society, 2014. 420–424. [doi: 10.1109/DSAA.2014.7058106]
- [62] Ma JM. Research on extraction method of implicit evaluation objects in online user reviews [MS. Thesis]. Beijing: Beijing Jiaotong University, 2017 (in Chinese with English abstract).
- [63] Xu H, Zhang F, Wang W. Implicit feature identification in Chinese reviews using explicit topic mining model. *Knowledge-Based Systems*, 2015,76(1):166–175. [doi: 10.1016/j.knosys.2014.12.012]
- [64] Lau RYK, Li C, Liao SSY. Social analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis. *Decision Support Systems*, 2014,65:80–94. [doi: 10.1016/j.dss.2014.05.005]
- [65] Zhang Y, Zhu W. Extracting implicit features in online customer reviews for opinion mining. In: Proc. of the 22nd Int'l Conf. on World Wide Web Companion (WWW 2013). New York: ACM Press, 2013. 103–104. [doi: 10.1145/2487788.2487835]
- [66] Bagheri A, Saraee M, De JF. Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews. *Knowledge-Based Systems*, 2013,52:201–213. [doi: 10.1016/j.knosys.2013.08.011]
- [67] Fei G, Liu B, Hsu M, Castellanos M, Ghosh R. A dictionary-based approach to identifying aspects implied by adjectives for opinion mining. In: Proc. of the 24th Int'l Conf. on Computational Linguistics (COLING 2012). Stroudsburg: ACL, 2012. 309.
- [68] Xia L, Wang Z, Chen C, Zhai S. Research on feature-based opinion mining using topic maps. *The Electronic Library*, 2016,34(3): 435–456. [doi: 10.1108/EL-11-2014-0197]
- [69] Smrř P. Using WordNet for opinion mining. In: Sojka P, Choi KS, eds. Proc. of the 3rd Int'l WordNet Conf. (GWC 2006). Brno: Masaryk University, 2006. 333–335.
- [70] Zhu YL, Min J, Zhou YQ, Huang XJ, Wu LD. Semantic orientation computing based on HowNet. *Journal of Chinese Information Processing*, 2006,20(1):16–22 (in Chinese with English abstract).
- [71] Li D, Qiao BJ, Cao YD, Wan YL. Word orientation recognition based on semantic analysis. *Pattern Recognition and Artificial Intelligence*, 2008,21(4):482–487 (in Chinese with English abstract).
- [72] Du WF, Tan SB, Yun XC, Cheng XQ. A new method to compute semantic orientation. *Journal of Computer Research and Development*, 2009,46(10):1713–1720 (in Chinese with English abstract).
- [73] Su F, Markert K. Subjectivity recognition on word senses via semi-supervised mincuts. In: Proc. of the Human Language Technologies: The 2009 Annual Conf. of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2009). Stroudsburg: ACL, 2009. 1–9.
- [74] Esuli A, Sebastiani F. Determining the semantic orientation of terms through gloss classification. In: Herzog O, Schek HJ, eds. Proc. of the 14th ACM Int'l Conf. on Information and Knowledge Management (CIKM 2005). New York: ACM Press, 2005. 617–624. [doi: 10.1145/1099554.1099713]
- [75] Turney PD, Littman ML. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Trans. on Information Systems (TOIS)*, 2003,21(4):315–346. [doi: 10.1145/944012.944013]
- [76] Kim SM, Hovy E. Automatic detection of opinion bearing words and sentences. In: Dale R, Wong KF, eds. Proc. of the 2nd Int'l Joint Conf. on Natural Language Processing (IJCNLP 2005). Berlin: Springer-Verlag, 2005. 61–66.
- [77] Kim SM, Hovy E. Identifying and analyzing judgment opinions. In: Robert C, Jeff A, Jennifer C, Mark S, eds. Proc. of the Main Conf. on Human Language Technology Conf. of the North American Chapter of the Association of Computational Linguistics (HLT-NAACL 2006). Stroudsburg: ACL, 2006. 200–207.
- [78] Liu H, He J, Wang T, Song W, Du X. Combining user preferences and user opinions for accurate recommendation. *Electronic Commerce Research and Applications*, 2013,12(1):14–23. [doi: 10.1016/j.elerap.2012.05.002]
- [79] Lu H, Niu ZD, Zhang N, Sun XK, Liu WL. A model for sentiment classification of Chinese microblog based on parsing and theme extension. *Trans. of Beijing Institute of Technology*, 2014,34(8):824–830 (in Chinese with English abstract).
- [80] Matsuo Y, Ishizuka M. Keyword extraction from a single document using word co-occurrence statistical information. *Int'l Journal on Artificial Intelligence Tools*, 2004,13(1):157–169. [doi: 10.1142/S0218213004001466]

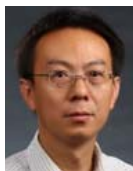
- [81] Lin C, He Y, Everson R, Ruger S. Weakly supervised joint sentiment-topic detection from text. *IEEE Trans. on Knowledge and Data Engineering*, 2012,24(6):1134–1145. [doi: 10.1109/TKDE.2011.48]
- [82] Pang B, Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: *Proc. of the 42nd Annual Meeting on Association for Computational Linguistics (ACL 2004)*. Stroudsburg: ACL, 2004. 271–278. [doi: 10.3115/1218955.1218990]
- [83] Blitzer J, Dredze M, Pereira F. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In: *Proc. of the 45th Annual Meeting on Association for Computational Linguistics (ACL 2004)*, Vol.7. Stroudsburg: ACL, 2007. 440–447.
- [84] Duric A, Song F. Feature selection for sentiment analysis based on content and syntax models. *Decision Support Systems*, 2012, 53(4):704–711. [doi: 10.1016/j.dss.2012.05.023]
- [85] Zheng YG. Research on extended topic model for fine-grained opinion mining of online reviews [MS. Thesis]. Hangzhou: Zhejiang Gongshang University, 2017 (in Chinese with English abstract).
- [86] Li CX, Xie LD. Text classification and opinion mining based on LDA. *Electronic Technology & Software Engineering*, 2017,4: 209–210 (in Chinese with English abstract).
- [87] Sterckx L, Caragea C, Demeester T, Develder C. Supervised keyphrase extraction as positive unlabeled learning. In: *Proc. of the 2016 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2016)*. Stroudsburg: ACL, 2016. 1924–1929.
- [88] Du J, Gui L, Xu R. Extracting opinion expression with neural attention. In: Li YM, Xiang GX, eds. *Proc. of the 5th of Chinese National Conf. on Social Media Processing (SMP 2016)*. Singapore: Springer-Verlag, 2016. 151–161. [doi: 10.1007/978-981-10-2993-6_13]
- [89] Lafferty J, McCallum A, Pereira F. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Brodley C, Danyluk AP, eds. *Proc. of the 8th Int'l Conf. on Machine Learning (ICML 2001)*, Vol.1. New York: ACM Press, 2001. 282–289.
- [90] Pontiki M, Galanis D, Papageorgiou H, Manandhar S, Androutsopoulos I. Semeval-2015 task 12: Aspect based sentiment analysis. In: Nakov P, Zesch T, eds. *Proc. of the 9th Int'l Workshop on Semantic Evaluation (SemEval 2015)*. Stroudsburg: ACL, 2015. 486–495.
- [91] Lu B. Identifying opinion holders and targets with dependency parser in Chinese news texts. In: *Proc. of the Human Language Technologies: The 11th Annual Conf. of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2010)*. Stroudsburg: ACL, 2010. 46–51.
- [92] Zhang C. Automatic keyword extraction from documents using conditional random fields. *Journal of Computational Information Systems*, 2008,4(3):1169–1180.
- [93] Laddha A, Mukherjee A. Extracting aspect specific opinion expressions. In: *Proc. of the 2016 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2016)*. Stroudsburg: ACL, 2016. 627–637.
- [94] Breck E, Choi Y, Cardie C. Identifying expressions of opinion in context. In: Veloso M, ed. *Proc. of the 20th Int'l Joint Conf. on Artificial Intelligence (IJCAI 2007)*. Menlo Park: AAAI Press, 2007. 2683–2688.
- [95] Yang B, Cardie C. Extracting opinion expressions with semi-Markov conditional random fields. In: *Proc. of the 2012 Joint Conf. on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012)*. Stroudsburg: ACL, 2012. 1335–1345.
- [96] Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015,521(7553):436–444. [doi: 10.1038/nature14539]
- [97] Li Q, Jin Z, Wang C, Zeng DD. Mining opinion summarizations using convolutional neural networks in Chinese microblogging systems. *Knowledge-Based Systems*, 2016,107(C):289–300. [doi: 10.1016/j.knosys.2016.06.017]
- [98] Pang L, Lan YY, Xu J, Guo JF, Wan SX, Cheng XQ. A survey on deep text matching. *Chinese Journal Computers*, 2016,40(4): 1–19 (in Chinese with English abstract).
- [99] Socher R, Perelygin A, Wu JY, Chuang J, Manning CD, Ng AY, Potts C. Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proc. of the 2013 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2013)*. Stroudsburg: ACL, 2013. 1631–1642.
- [100] Yao K, Zweig G, Hwang MY, Shi Y, Yu D. Recurrent neural networks for language understanding. In: *Proc. of the 14th Annual Conf. of the Int'l Speech Communication Association (INTERSPEECH 2013)*. Portland: INTERSPEECH, 2013. 2524–2528.

- [101] Kalchbrenner N, Grefenstette E, Blunsom P. A convolutional neural network for modelling sentences. In: Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014). Stroudsburg: ACL, 2014. 655–665.
- [102] Huang Z, Xu W, Yu K. Bidirectional LSTM-CRF models for sequence tagging. *Computer Science*, 2015,20(2):508–517.
- [103] Ma X, Hovy E. End-to-End sequence labeling via bi-directional LSTM-CNNs-CRF. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume (ACL 2016). Stroudsburg: ACL, 2016. [doi: 10.13140/RG.2.1.2182.5685]
- [104] Wiebe J, Wilson T, Cardie C. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 2005,39(2-3):165–210. [doi: 10.1007/s10579-005-7880-9]
- [105] Johansson R, Moschitti A. Syntactic and semantic structure for opinion expression detection. In: Proc. of the 14th Conf. on Computational Natural Language Learning (CoNLL 2010). Stroudsburg: ACL, 2010. 67–76.
- [106] Ye Q, Zhang ZQ, Luo ZX. Automatically measuring subjectivity of Chinese sentences for sentiment analysis to reviews on the Internet. *China Journal of Information Systems*, 2007,1(1):79–91 (in Chinese with English abstract).
- [107] Dave K, Lawrence S, Pennock DM. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In: Proc. of the 12th Int'l Conf. on World Wide Web (WWW 2003). New York: ACM Press, 2003. 519–528. [doi: 10.1145/775152.775226]
- [108] Gamon M, Aue A, Corston-Oliver S, Ringger E. Pulse: Mining customer opinions from free text. In: Famili A, Joost N, eds. Proc. of the 6th Int'l Conf. on Intelligent Data Analysis (IDA 2005). Berlin, Heidelberg: Springer-Verlag, 2005. 121–132. [doi: 10.1007/11552253_12]
- [109] Liu B, Hu M, Cheng J. Opinion observer: Analyzing and comparing opinions on the Web. In: Proc. of the 14th Int'l Conf. on World Wide Web (WWW 2005). New York: ACM Press, 2005. 342–351. [doi: 10.1145/1060745.1060797]
- [110] Yi J, Niblack W. Sentiment mining in WebFountain. In: Proc. of the 21st IEEE Int'l Conf. on Data Engineering (ICDE 2005). Washington: IEEE Computer Society, 2005. 1073–1083. [doi: 10.1109/ICDE.2005.132]
- [111] Wilson T, Hoffmann P, Undaran SS, Kessler K, Wiebe J, Choi Y, Cardie C, Riloff E, Patwardhan S. OpinionFinder: A system for subjectivity analysis. In: Proc. of the Human Language Technology Conf. on Empirical Methods in Natural Language Processing (HLT-EMNLP 2005). Stroudsburg: ACL, 2005. 34–35.
- [112] Mukherjee A. Extracting aspect specific sentiment expressions implying negative opinions. In: Proc. of the 17th Int'l Conf. on Intelligent Text Processing and Computational Linguistics (CICLING 2016). 2016. 1–20.
- [113] Pontiki M, Galanis D, Papageorgiou J, Manandhar S, Androutsopoulos I. Semeval-2015 task 12: Aspect based sentiment analysis. In: Proc. of the 9th Int'l Workshop on Semantic Evaluation (SemEval 2015). Stroudsburg: Association for Computational Linguistics, 2015. 486–495.
- [114] Snyder B, Barzilay R. Multiple aspect ranking using the good grief algorithm. In: Proc. of the Human Language Technologies: The 2007 Annual Conf. of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2007). Stroudsburg: ACL, 2007. 300–307.
- [115] Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis. In: Proc. of the Conf. on Human Language Technology and Empirical Methods in Natural Language Processing. 2005. 347–354.
- [116] Baccianella S, Esuli A, Sebastiani F. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Proc. of the 7th Int'l Conf. on Language Resources and Evaluation. 2010. 2200–2204.
- [117] Stone PJ, Dunphy DC, Smith MS, Ogilvie DM. *The General Inquirer: A Computer Approach to Content Analysis*. Oxford: The M.I.T. Press, 1968. 375–376.
- [118] Chen WT, Lin SC, Huang SL, Chung YS, Chen KJ. E-HowNet and automatic construction of a lexical ontology. In: Proc. of the Int'l Conf. on Computational Linguistics: Demonstrations (ICCL 2010). Stroudsburg: ACL, 2010. 45–48.
- [119] Ku LW, Chen HH. Mining opinions from the Web: Beyond relevance retrieval. *Journal of the Association for Information Science and Technology*, 2007,58(12):1838–1850. [doi: 10.1002/asi.20630]
- [120] Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ. The development and psychometric properties of LIWC 2007. *LIWC 2007 Manual*, 2015,29(11):1020–1025.
- [121] Wang K, Xia R. A survey on automatic construction methods of sentiment lexicons. *Acta Automatica Sinica*, 2016,42(4): 495–511 (in Chinese with English abstract).

- [122] Chen H, Zimbra D. AI and opinion mining. *IEEE Intelligent Systems*, 2010,25(3):74–80. [doi: 10.1109/MIS.2010.75]
- [123] Huang XJ, Zhang Q, Wu YB. A survey on sentiment analysis. *Journal of Chinese Information Processing*, 2011,25(6):118–126 (in Chinese with English abstract).

附中文参考文献:

- [2] 姚天昉,程希文,徐飞玉,汉思·乌思克尔特,王睿.文本意见挖掘综述. *中文信息学报*,2008,22(3):71–80.
- [3] 王辉,王晖昱,左万利.观点挖掘综述. *计算机应用研究*,2009,26(1):25–29.
- [4] 陈旻,朱凡微,吴明晖,应晶.观点挖掘综述. *浙江大学学报:工学版*,2014,48(8):1461–1472.
- [13] 李建华,刘功申,林祥.情感倾向性分析及应用研究综述. *信息安全学报*,2017,2(2):48–62.
- [14] 陈巧红,孙超红,贾宇波.文本数据观点挖掘技术综述. *工业控制计算机*,2017,30(2):94–95,102.
- [15] 陈晓美,王付国,吴宏伟,孙中秋.社会化网络评论观点挖掘的研究热点与应用进展. *情报科学*,2013,31(11):119–124.
- [29] 徐冰,赵铁军,王山雨,郑德权.基于浅层句法特征的评价对象抽取研究. *自动化学报*,2011,37(10):1241–1247.
- [43] 刘鸿宇,赵妍妍,秦兵,刘挺.评价对象抽取及其倾向性分析. *中文信息学报*,2010,24(1):84–88.
- [54] 刘倩.观点挖掘中评价对象抽取方法的研究[博士学位论文].南京:东南大学,2016.
- [58] 郎君,忻舟,秦兵,刘挺,李生.集成多种背景语义知识的共指消解. *中文信息学报*,2009,23(3):3–9.
- [62] 马京苗.网购用户评论中隐式评价对象的提取方法研究[硕士学位论文].北京:北京交通大学,2017.
- [70] 朱嫣岚,闵锦,周雅倩,黄萱菁,吴立德.基于 HowNet 的词汇语义倾向计算. *中文信息学报*,2006,20(1):16–22.
- [71] 李钝,乔保军,曹元大,万月亮.基于语义分析的词汇倾向识别研究. *模式识别与人工智能*,2008,21(4):482–487.
- [72] 杜伟夫,谭松波,云晓春,程学旗.一种新的情感词汇语义倾向计算方法. *计算机研究与发展*,2009,46(10):1713–1720.
- [79] 陆浩,牛振东,张楠,孙星恺,刘文礼.基于句法与主题扩展的中文微博情感倾向性分析模型. *北京理工大学学报*,2014,34(8):824–830.
- [85] 郑玉桂.面向电商评论细粒度观点挖掘的拓展主题模型研究[硕士学位论文].杭州:浙江工商大学,2017.
- [86] 李晨曦,谢罗迪.基于 LDA 模型的文本分类与观点挖掘. *电子技术与软件工程*,2017,4:209–210.
- [98] 庞亮,兰艳艳,徐君,郭嘉丰,万圣贤,程学旗.深度文本匹配综述. *计算机学报*,2016,40(4):1–19.
- [106] 叶强,张紫琼,罗振雄.面向互联网评论情感分析的中文主观性自动判别方法研究. *信息系统学报*,2007,1(1):79–91.
- [121] 王科,夏睿.情感词典自动构建方法综述. *自动化学报*,2016,42(4):495–511.
- [123] 黄萱菁,张奇,吴苑斌.文本情感倾向分析. *中文信息学报*,2011,25(6):118–126.
- [123] 黄萱菁,张奇,吴苑斌.文本情感倾向分析. *中文信息学报*,2011,25(6):118–126.



韩忠明(1972—),男,山西文水人,博士,副教授,CCF 专业会员,主要研究领域为社会网络,数据挖掘,大数据处理.



李梦琪(1993—),女,硕士生,主要研究领域为自然语言处理.



刘莹(1992—),男,硕士生,主要研究领域为社交网络挖掘.



张梦玟(1995—),女,硕士生,主要研究领域为自然语言处理.



段大高(1976—),男,博士,副教授,CCF 专业会员,主要研究领域为数据挖掘,人工智能.



于重重(1971—),女,博士,教授,CCF 高级会员,主要研究领域为模式识别,机器学习.