句子情感分析及其关键问题*

李　纲　程洋洋　寇广增
武汉大学信息管理学院　武汉 430072

〔摘要〕情感分析关注具有情感倾向的评价性信息，具有广泛的应用。情感分析按照粒度的不同分为三种：词汇情感分析、句子情感分析和文档情感分析。文中对句子情感分析及其关键问题进行介绍，首先简要描述句子情感分析的任务，然后介绍句子情感分析中主客观句分类方法及两种主观句情感分类方法——基于情感词的方法和机器学习方法，最后对情感分析中的三个关键问题——词汇上下文极性判定、评价主题识别、意见持有者识别进行总结。

〔关键词〕句子情感分析　词汇上下文极性　评价主题　意见持有者
〔分类号〕TP391

Key Problems of Sentence Level Sentiment Analysis

Li Gang　Cheng Yangyang　Kou Guangzeng
Information Management School of WhuHan University, WuHan 430072

〔**Abstract**〕Sentiment analysis studies on evaluative information, it has a wide range of application. Divided by granularity, there are three kinds of sentiment analysis, including word, sentence and document level sentiment analysis. This paper mainly focuses on the sentence level sentiment analysis. First describes its tasks briefly, then introduces the methods of sentence S/O classification and two ways of subjective sentence sentiment classification, finally summarizes three key problems of sentence level sentiment analysis: word contextual polarity disambiguation, topic identification and opinion holder identification.

〔**Keywords**〕sentence level sentiment analysis　word contextual polarity　evaluation topic　opinion holder

1　引　言

随着网络的发展与普及，由普通用户发表的包含个人情感倾向的评价性信息越来越多。评价性信息包含四个部分[1]：评价主题、意见持有者、评价和情感倾向，情感分析通过分析和挖掘评价性信息，识别其情感倾向。情感分析根据粒度的不同分为三种：词汇情感分析、句子情感分析和文档情感分析，三者相比较，句子情感分析能够得到评价主题及各个方面特征的情感关系，具有更广泛的应用范围。

本文将句子情感分析作为研究重点，首先简要描述句子情感分析的主要任务，然后介绍了主客观句分类方法以及两种主观句情感分类方法——基于情感词的方法和机器学习方法，最后对句子情感分析中的关键问题进行总结。

2　句子情感分析的任务

句子情感分析的任务是按照句子所表达的情感倾向对其进行识别，包含以下两个子任务[2]：①主观句识别，提取文本中包含的主观句；②主观句的情感分类，识别主观句的情感倾向，通常是褒/贬二元分类。

2.1　主观句识别

主观句识别是对文本进行分析，过滤掉其中的客观句，得到更能反映文本情感倾向的主观句集合。根据词性的不同，Hu Minqing 和 Liu Bing[3]将形容词作为主客观句的分界线，当句子中同时包含形容词和评价主题时，即认为该句为主观句。句子之间的关系同样可以作为判定标准，Pang 和 Lee[4]采用最小图割的方法获取文档中的句子与已知主观句的关系；Yu[5]将事实性文档看作客观句集合，评价性文档看作主观句集合，通过判断句子与这两种文档之间的相互关系识别

*本文系国家自然科学基金资助项目“文本集特征提取方法及应用研究”（项目编号：70673070）研究成果之一。
收稿日期：2010－01－13　修回日期：2010－05－06　本文起止页码：104－107，127　本文责任编辑：高　丹

句子的主客观性。从句子中筛选出具有情感倾向的情感词和短语作为特征[6]，选择不同的分类算法如贝叶斯、K 邻近等，采用机器学习的方法进行主客观分类，同样取得不错的效果。基于规则的方法可以从文档中提取精度高、观点清晰的主观句[7]，但需要人工编写语言规则且覆盖面较窄。

从以上研究可以看出，识别文档中的主观句关键是提取句子中包含的情感词或者直接判断，或者结合其它信息作为特征项送入标准分类器中判断。句子的主客观分类能够有效提高文本情感分析的准确度，在以上方法中，客观句的识别一般在 80% 左右，而主观句的识别比较低，只有 60% 左右。

2.2 主观句的情感分类

主观句的情感分类是对主观句所表达的情感倾向进行褒贬识别，主要包括两种分类方法：基于情感词的方法和机器学习方法。

2.2.1 *基于情感词的方法* 基本思路是通过判定句子中包含情感词的语义倾向，加上句法结构等信息，间接得到句子的情感倾向。其流程如图 1 所示：

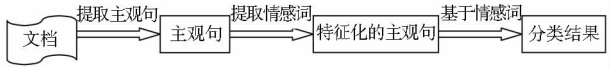

图 1 基于情感词的方法

通过情感词判断句子情感倾向时，Yu[5]、Kim 和 Hovy[1]、Hu Minqing 和 Liu Bing[3]首先构建一个情感词集，并为集合中的每个情感词标记正/负整数值作为情感值。

在得到情感词集后，Hu Minqing 和 Liu Bing 通过统计句子中褒义词和贬义词的数量判断句子的情感倾向；Yu 将句子中所有情感词的情感平均值作为句子的情感值；Kim 和 Hovy 则将否定词纳入到句子情感倾向的判定之中，采用乘积方法来判断句子的极性，该方法能够处理否定、双重否定对句子情感倾向的反向影响。

采用基于情感词的方法判定句子情感倾向时，能否得到情感倾向准确、包含全面的情感词集是关键，同时也要考虑一些特殊的句法结构对结果的影响，如否定句、比较句等。

2.2.2 *机器学习方法* 机器学习方法的基本思想是根据已知训练样本求取对系统输入输出之间依赖关系的估计，使它能够对未知输出作出尽可能准确的预测。使用机器学习方法进行情感分类时，分类算法的选择和特征项的选取是最重要的两个方面。运用机器学习方法进行情感分类的过程（见图 2）。

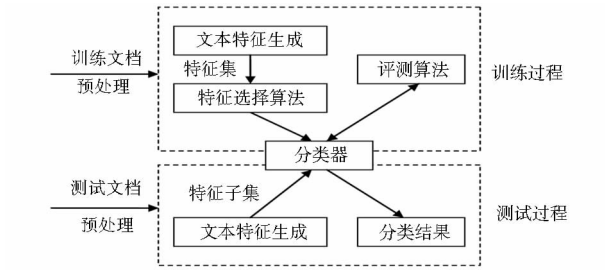

图 2 运用机器学习方法进行情感分类

PangBo[8]最早将机器学习方法应用于情感分类领域，他分别利用朴素贝叶斯、最大熵、SVM 算法对电影评论进行分类，当以 unigram 作为特征项时，SVM 表现最好，准确率为 82.9%，最大熵和朴素贝叶斯的效果相当。

与 PangBo 不同，Dava[9]在对几种产品的评论进行情感分类时，采用 bigram 作为特征项训练分类器的效果最好，这表明分类器效果的好坏与所选取的特征项息息相关。

在特征项的选择上，崔彩霞和王素格[10]提出一个特征项选择函数，用来替代传统的文档频率和互信息选择方法。除此之外，王素格等[11]还研究了停用词对中文文本情感分类的影响，它构造了五种停用词表作为特征项选择的依据，实验表明停用词表的选择对文本情感分类的影响很大。

在采用机器学习方法分类时，同时选取形容词、副词、名词作为特征项比选取单一词性的效果要好，对否定词进行处理能明显提高分类的准确性[12]。

在上述机器学习方法中，选取的特征项是相互独立的，然而句子中词汇之间的语义关系对判断文本的情感倾向也很重要。Matsumoto 等[13]从句子提取出频繁子序列和频繁子树，与 unigram、bigram 共同作为特征项，采用 SVM 方法分类时准确率达到了 92% 以上。Whitelaw[14]将评价组作为文本情感倾向识别的最小单位，同样采用 SVM 方法分类，准确率在 78% 左右，当其它特征项增大文本的覆盖范围时，准确率上升到 90% 以上。

3 句子情感分类的关键问题

在许多应用中，不但需要对句子的主客观性和整体情感倾向进行识别，还需要深入句子内部分析评价主题和各个特征的情感倾向以及与意见持有者的从属关系。本文从实际应用的角度出发总结出句子情感分

析的三个关键问题,下面分别介绍。

3.1 词汇上下文极性的判定

词汇含有两种极性,原极性和上下文极性。原极性指词汇本身的极性;上下文极性指词汇在文本中的极性。在上下文中,由于受到周围词汇影响,词汇的情感强度可能发生变化,甚至与原极性相反。正确识别词汇的上下文极性能够有效提高情感分类的准确率。

娄德成等[15]和徐琳宏等[16]研究了否定词和强度词对词汇极性的影响。前者构建否定词字典和强度词字典,对文本进行词性标注后,根据词性找到词汇间的依存关系,计算词汇的上下文倾向。后者采用否定规则匹配文档中的否定句,同时处理强度词附近具有明显语义倾向的词汇,得到经过否定处理和强度处理的特征项,分类效果比处理前提高了5%左右。

Wilson等[17]首先判断句子中短语的主客观性,从主观性短语中选取词汇特征和极性特征,对短语的上下文极性采用机器学习方法判断。同样可以采用人工编写规则的方法[7]来判断词汇的上下文极性,该方法可以达到非常高的准确率,然而查全率很低,并且只能对部分情感表达进行判断。

词汇上下文极性的识别是句子情感分析的关键,然而由于自然语言的差异以及句法结构的复杂性,使得词汇的上下文极性很难判断;另外,人为因素如书写不规范、人造词语等也增加了这方面的困难。要准确地判断词汇上下文极性,还需要吸收一些语言学方面的研究成果。

3.2 评价主题的识别

评价主题包括显性评价主题和隐性评价主题,前者可以直接从句子中得到,而隐性评价主题只能根据句子中词汇之间的关系来判断。

3.2.1 显性评价主题的识别 Hu Minqing和Liu Bing[3]认为,虽然在一篇文档中会涉及到对评价主题多个方面的评价,但他们所用的词汇具有收敛关系,可以通过关联挖掘方法从文本中得到经常出现和较少出现的评价主题。

娄德成等[15]提出SBV算法及其补充算法,利用词汇间的语义关系从汉语主观句中识别评价主题。由于汉语语义关系的复杂性和网络中用户评论结构的不规范性,该方法在实验中的准确率只有40%。

天网知名度系统将名人作为评价主题,能够从网络中自动抽取名人的姓名以及相应的评价,但它只能提供名人的总体评价,缺乏对某一方面的具体评价。

苏祺[18]通过对主观句进行词性标注,将名词和名词短语作为候选主题,在对候选主题过滤后进行聚类。该方法不但能识别显性评价主题,还可以得到具有内在联系的评价主题集合。

3.2.2 隐性评价主题的识别 很多情况下句子中并没有将评价主题显示出来,而是通过一些词语表达。例如"这辆车很灵活",就隐性地对汽车的操控性进行评价。隐性评价主题的识别依赖于上下文语义分析,目前只有少数的研究涉及。

在隐性评价主题的识别上,可以将评价词汇与评价主题映射[18],如图3所示:

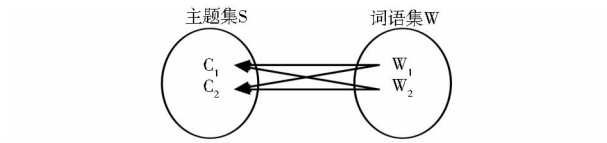

图3 评价主题与评价词语的映射关系

它在PMI-IR的基础上提出了用于计算评价性词语与评价主题之间相互关系的FB-PMI-IR方法,可以根据值的大小确定评价性词语与主题的关系。例如"漂亮"与"动力性"FB-PMI-IR值为-12.01,与"外观"的值为-4.79,从而可以得到"漂亮"更有可能修饰"外观"。

评价主题的识别是句子情感分析应用的重要方面。在领域中,评价主题及对应的情感词是一个有限集合,两者之间存在着多对多的关系。在通过机器学习识别这些关系时,需要大量的评价文本做支撑,如何获得高质量的标注评价文本,是提高评价主题识别准确率的关键。

3.3 意见持有者的识别

意见持有者是对评价主题进行评价的主体,包括个人、机构等实体,对意见持有者进行识别能够得到某人对某事的具体态度。

一般来说,可以通过命名实体识别将人或机构名作为意见持有者[7],但该方法的语言覆盖率较差且领域独立性较弱。Kim和Hovy[19]、Xu和Wong[20]将个人、机构、国家和群体四种实体作为候选意见持有者,前者通过最大熵评测算法从中选取可能性最大的实体作为意见持有者;而后者通过在句子中找到意见算子进而确定意见持有者,同时考虑到修饰实体的词汇及其附近实体对意见持有者识别的影响,如短语"美国总统布什"作为一个整体被看做意见持有者。

Choi等[21]把意见持有者的识别看做是一个信息抽取任务,考虑到句子表达的情感强度,将基于规则的

信息抽取和机器学习方法相结合来识别意见持有者。

总体来说,相对于词汇上下文极性的判定和评价主题的识别,句子中意见持有者识别的难度更高,这主要表现在:① 一个句子中可能会包含多个评价,需要为每个评价确定对应的意见持有者;② 一个句子可能包含多个意见持有者,需要判断它们之间的关系;③ 需要考虑句子之间评价与意见持有者的关系。

4 结 语

由于在商业方面巨大的应用价值,情感分析受到许多研究机构的重视。在技术上,产生了多种情感特征抽取方法和分类策略;在应用上,基于情感分析的应用系统层出不穷,如意见挖掘系统、舆情分析系统等。但是由于自然语言情感表达方式的多样性,情感分析仍然面临着许多困难,笔者认为,未来情感分析研究的热点主要集中在:

- 文本情感强度判断。情感分析不仅得出好/坏、正面/负面这样的二元分类,还应该对句子的情感强度进行分析。例如可以通过加权,为各个情感词设置不同权重的方法对句子及其中包含的主观性短语进行强度分析。
- 更有效的特征抽取方法。通过改进现有的或设计新的特征抽取方法,从文本中提取出更能表达文本情感的主题、情感词特征以及影响文本情感倾向的句法信息、特殊词汇等,提高文本情感分类的效率。
- 情感语料库的建设。对非频繁特征来说,数据稀疏一直是基于机器学习方法的瓶颈,作为情感分析的知识来源,需要建设大规模情感语料库,在语料的采集、标注规范的制定及语料库应用等方面提供规则。
- 中文特殊句法结构处理。汉语表达方式的多样性、句法结构的复杂性使中文文本情感分析更加复杂,其中否定句、比较句是最常见的两种类型,对中文句子的句法分析需要借助于语言学领域的研究成果。
- 更广泛的应用。情感分析需要与其它领域相结合,形成更有价值的应用。如可以将情感词、主题等作为查询条件的情感检索;对产品评论进行情感分析后得到规范的情感摘要;识别不同网站相互转载的重复信息、竞争对手发布的恶意信息等垃圾信息识别等。
- 跨领域研究。在目前的情感分析中,情感词本身所表示的情感极性与主题所属的领域密切相关,在大部分情况下,这种领域依赖是不同主题领域中常用词汇变化的结果。当同一个情感词与不同的主题、不同的特征相关联时,可能表达了截然相反的情感极性,这就需要对情感分析的跨领域问题进行研究。

参考文献:

[1] Kim S M, Hovy E. Determining the sentiment of opinions//Proceedings of the 20th International Conference on CL. Morristown: ACL, 2004: 1367 - 1373.

[2] 王根, 赵军. 基于多重冗余标记 CRFS 的句子情感分析研究. 中文信息学报, 2007, 21(5): 51 - 55.

[3] Hu Minqing, Liu Bing. Mining and summarizing customer reviews// Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. New York: ACM, 2004: 168 - 177.

[4] Pang B, Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts// Proceedings of the ACL. Morristown: ACL, 2004: 271 - 278.

[5] Yu H, Hatzivassiloglou V. Towards answering opinion questions: Separating facts from opinions and Identifying the polarity of opinion sentences//Proceedings of the Conference on Empirical Methods in NLP. Morristown: ACL, 2003: 129 - 136.

[6] Hatzivassiloglou V, Wiebe J. Effects of adjective orientation and gradability on sentence subjectivity//DFKI. Proceedings of 18th International Conference on CL. Morristown: ACL, 2000: 299 - 305.

[7] Nasukawa T, Yi J. Sentiment analysis: Capturing favorability using natural language processing//Proceedings of the International Conference on Knowledge Capture. New York: ACM, 2003: 70 - 77.

[8] Pang Bo, Lee L. Thumbs up? Sentiment classification using machine learning techniques//Proceedings of the Conference on Empirical Methods in NLP. Morristown: ACL, 2002: 79 - 86.

[9] Dave K, Lawrence S, Pennock D. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews// ACM. Proceedings of the 12th International World Wide Web Conference. New York: ACM, 2003: 519 - 528.

[10] 崔彩霞, 王素格. 基于内类频率的文本分类特征选择方法. 计算机工程与设计, 2007, 28(17): 4249 - 4251.

[11] 王素格, 魏英杰. 停用词表对中文文本情感分类的影响. 情报学报, 2008,27(2): 175 - 179.

[12] 徐军, 丁宇新, 王晓龙. 使用机器学习方法进行新闻的情感自动分类. 中文信息学报, 2007, 21(6): 95 - 100.

[13] Matsumoto S, Takamura H, Okumura M. Sentiment classification using word sub-sequences and dependency sub-trees// Proceedings of the 9th Pacific-Asia Conference on Knowledge Discovery and Data Mining. Berlin: Springer Verlag, 2005: 301 - 310.

[14] Whitelaw C, Garg N, Argamon S. Using appraisal groups for sentiment analysis//Proceedings of the 14th ACM International Conference on Information and Knowledge Management. New York: ACM, 2005: 625 - 631.

[15] 娄德成, 姚天昉. 汉语语句主题语义倾向分析方法的研究. 中文信息学报, 2007, 21(5): 73 - 79.

(下转第127页)

交流有关外,不得要领,根本揭示不了它们的本质,因为图书馆内部活动不直接参与交流。这似乎已为后来的知识组织理论所补充,但怎样将两种理论有机地捏合成一个整体,仍然需要进一步探讨。

12 思索和评价

• 我国图书馆学的"知识流派"虽都以知识为基点展开,研究知识的本质属性、知识的流程循环、知识的发现、揭示和组织以及知识管理制度等一系列方面,但这些研究的群体性和关联性不太强;即使是关于一些基本观点的论述也体现了不同的视角。这也直接造成了图书馆学的学科知识体系更加难以捉摸,图书馆学研究的价值取向更加模糊。

• 关于知识各个层面的研究还不够系统,表现在各个层面的研究中虽然都有一个或几个学术代表人物,但这些研究体系尚未成熟,处于观点体系的不断探讨、补充和继续探索阶段,缺乏较大规模的后继研究及群体性研究,学术共同体一时难以组建成规模。

• 关于知识理论与技术的研究,两者难以统一协调。纵观已有文献的研究,各观点的代表学者大多只关注于自身钻研的有限领域,这就出现了各学说理论体系中缺乏技术、方法论等实践模块的充实,而理论体系本身的研究也缺乏如语言学、经济学、管理学、法学等多学科知识联合的、交叉的、深入的借鉴和指导,论证层面显得有些肤浅。

• 令人遗憾的是,虽然意识到学科科际整合的重要性,但不同学术团体以及专业人员之间仍未达成有效的沟通和交流。这里尤其值得注意的是,各理论体系所提出的基本原理与基本观点缺乏对图书馆一线工作的阐释与指导,缺乏图书馆事业这个图书馆学载体的实践支撑,因而也就无法迎合图书馆职业精神的弘扬这个时代课题。

• 我国图书馆学研究对象学术思想"知识流派"的学科共同体还缺乏一个统一的研究规范,缺乏一个学科共同体应有的加强的韧性与相对稳定性。

参考文献:

[1] [2004 - 08 - 21]. http://biz.163.com/05/0411/20/1H38H51E00021ED1.html.
[2] 彭修义. 关于开展知识学研究的建议. 图书馆学通讯,1981(3):85 - 88.
[3] 蒋永福. 客观知识·人·图书馆——兼论图书馆学的研究对象. 中国图书馆学报,2003(5):11 - 15.
[4] 王子舟. 知识集合初论——对图书馆学研究对象的探索. 中国图书馆学报,2000(4):7 - 12.
[5] 柯平. 知识资源论——关于知识管理与图书馆学的研究对象. 图书馆论坛,2004(12):58 - 63.
[6] 龚蛟腾. 公共知识管理学——关于图书馆学本质的思考. 大学图书馆学报,2003(6):2 - 6.
[7] 马恒通. 知识传播论——图书馆学研究对象新探. 图书馆,2007(1):15 - 21.
[8] 倪延年. 知识传播功能论. 中国图书馆学报,2002(5):25 - 27.
[9] 王睿,张开凤. 关于图书馆学研究对象的再思考. 情报杂志,2003(5):103 - 104.
[10] 梁灿兴. 可获得性论的文献及机关概念. 图书馆,2002(1):9 - 15.
[11] 刘洪波. 知识组织论——图书馆内部活动的一种说明. 图书馆,1991(2):13 - 16.

〔作者简介〕葛园园,女,1980 年生,馆员,硕士,发表论文 10 余篇。

(上接第 107 页)

[16] 徐琳宏, 林鸿飞, 杨志豪. 基于语义理解的文本倾向性研究机制. 中文信息学报, 2007, 21(1): 96 - 100.
[17] Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis// Proceedings of Human Language Technology Conference on Empirical Methods in NLP. Morristown: ACL, 2005: 347 - 354,
[18] 苏祺. 面向问答系统的情感倾向分析研究.[学位论文]. 北京:北京大学, 2006.
[19] Kim S M, Hovy E. Identifying opinion holders for question answering in opinion texts//Proceedings of AAAI Workshop on Question Answering in Restricted Domains. California: AAAI Press, 2005: 254 - 261.
[20] Xu R, Wong K F, Xia Y. Opinmine-opinion analysis system by CUHK for NTCIR-6 pilot task// Proceedings of NTCIR-6 Workshop Meeting. Tokyo: Temberlake Consultants Press, 2007:350 - 357.
[21] Choi Y, Cardie C, Riloff E, et al. Identifying sources of opinions with conditional random fields and extraction patterns// Proceedings of Human Language Technologies Conference on Empirical Methods in NLP. Morristown: ACL, 2005: 355 - 362.

〔作者简介〕李 纲,男,1966 年生,教授,博士生导师,发表论文 50 余篇。
程洋洋,男,1987 年生,硕士研究生,发表论文 1 篇。
寇广增,男,1983 年生,博士研究生,发表论文 10 余篇。