文章编号:1007-130X(2019)04-0750-08

文本情感分析方法研究综述

洪 巍1,2,李 敏2

(1. 江南大学江苏省食品安全研究基地,江苏 无锡 214122; 2. 江南大学商学院,江苏 无锡 214122)

摘 要:文本情感是信息挖掘的一个新兴领域,近年受到管理学等相关领域的广泛关注。目前,文本情感分析使用的方法主要有情感词典方法和机器学习方法。由于文本情感分析的结果对优化政府、企业以及消费者决策具有重大意义,以文本情感分析的方法为视角,对情感词典的方法、有监督的机器学习方法和弱监督的深度学习方法以及其他方法的相关文献进行了梳理并做出评述。此外,指出虽然文本情感分析领域的学者基于情感词典和有监督的机器学习方法已提出许多情感分析模型,但准确率和效率普遍不高,进一步的研究重点应在于使用深度学习的方法处理文本情感,并提出未来的研究方向。

关键词:文本情感分析;情感词典;机器学习;深度学习

中图分类号:TP393

文献标志码:A

doi:10.3969/j.issn.1007-130X.2019.04.024

A review: Text sentiment analysis methods

HONG Wei^{1,2}, LI Min²

(1. Food Safety Research Base of Jiangsu Province, Jiangnan University, Wuxi 214122; 2. School of Business, Jiangnan University, Wuxi 214122, China)

Abstract: Text sentiment analysis is an emerging field of information mining, and it has attracted wide attention of management fields and other related areas in recent years. Currently, the methods used to perform text sentiment analysis tasks are mainly sentiment dictionary and t machine learning. The results of the text sentiment analysis are of great importance for the optimization of the decisions made by governments, enterprises and consumers. From the perspective of text sentiment analysis methods, we sort out and comment on methods such as sentiment dictionary, supervised machine learning, weak supervised deep learning, and some other methods. In addition, we find out that although a number of scholars in the field of text sentiment analysis have proposed many sentiment analysis models based on sentiment dictionary and supervised machine learning, the accuracy and efficiency of these models are generally not high. We suggest that further research should focus on using deep learning to deal with the text sentiment analysis, and also make proposal for future research directions.

Key words: text sentiment analysis; sentiment dictionary; machine learning; deep learning

1 引言

文本情感分析,又称意见挖掘(Opinion Mining),是指对带有情感色彩的主观性文本进行采集、处理、分析、归纳和推理的过程,涉及到人工智

能、机器学习、数据挖掘、自然语言处理等多个研究 领域。互联网的快速发展推动了电子商务以及各 大网络平台的普及,互联网用户由信息接受者向信 息发布者转变,不仅关注其他人对某件事或某件商 品的评论和看法,也越来越乐于分享自己的意见看 法或表明态度。因此,有必要对各大平台上的带有

基金项目:国家自然科学基金(71303094);国家社会科学基金重大项目(14ZDA069);中央高校基本科研业务费专项资金(UISPP51641A)

(JUSRP51641A)

^{*} 收稿日期:2018-04-24;修回日期:2018-09-14

情感色彩或情感倾向的言论进行整理分析,这可能带来多方面的效益,例如分析购物网站上的商品评论便于消费者详细了解商品信息,从而优化消费决策,也便于企业根据消费者的反馈信息掌握自己的优劣势,从而优化战略决策;针对某一舆情,通过对相关微博的评论进行整理分析,政府可以将其作为颁布相关政策的有力依据。所以,对含有情感色彩的文本进行情感极性判断具有巨大的商业价值和社会价值。由于文本情感分析是目前比较热门的研究领域,众多学者对此展开了研究,但研究方法比较分散,难以进行对比分析,本文根据文本情感分析的方法对已有的研究进行整理归纳。

2 情感词典方法

基于情感知识构建情感词典并将其作为工具 是判断主观性文本情感极性的传统方法,大部分的 情感词典是人工构建的,基本原理是根据经验将广 泛使用的情感词进行归纳整理,当文本输入后就与 词典内容进行匹配,寻找文本中与情感词典中重合 的情感词,从而判断文本的情感极性。情感词典最 早可以追溯到 1998 年, Whissell [1] 要求 148 名受 试者用5个附加单词来描述数学、物理学、电视、报 纸、生物学和技术等术语,然后与情感词典中广泛 使用的情感词进行匹配。在之后的几十年中,众多 学者在研究中对情感词典不断进行拓展完善, Whissell^[2]对原先构建的情感词典进行了修订,以 提高词典对自然语言的适用性。情感词典较早出 现在国外,英文词典资源丰富且具有优势,于是李 寿山等[3]利用英文种子词典,借助机器翻译系统, 构建了中文情感词典。上述的情感词典是最基础 的情感词典,即只对广泛使用的情感词如"美""棒" "差"等词进行整理形成情感词典,然后将需要处理 的文本与情感词典中的词进行匹配,统计正面和负 面情感词出现的次数,以进行文本情感极性判断。 但是,基础情感词典中的情感词覆盖率较低,无法 结合语境、语义,也无法识别同义词、近义词等,并 且主要依赖人工统计文本中情感词出现的次数,所 以在后来的研究中学者们对情感词典进行了进一 步完善。考虑到语境迁移的影响,现有的情感词典 应用到旧词新用的语料中分类效果较差,阳爱民 等[4] 用若干个情感种子词计算基础情感词的情感 倾向值,利用搜索引擎返回的共现数构建了情感词 典。Rao等[5]提出了一种有效算法和三种删减策 略自动构建词语级的和话题级的社交情感检测词 典。王志涛等[6]利用 40 万条微博数据构建新词词 典,对已有情感资源进行拓展,并对不同语言层次 定义不同的规则,还以表情符号作为附加信息提供 辅助作用。周杰[7]通过比较分析基于情感词典的 方法和基于机器学习的方法,针对两者的不足提出 一种新的基于情感词典和句型分类的中文情感分 析方法,并利用拉普拉斯平滑的情感倾向点互信息 SO-PMI (Semantic Orientation Pointwise Mutual Information)算法对微博情感词典进行扩展,深入 分析不同句型对句子情感倾向的影响。张克亮 等[8]融合了情感词典资源和概念层次网络语境框 架的优势,将文本的情感分析分为两个阶段,特征 词、句子和句群判定阶段;基于 HNC 语境框架的 句与句群情感分析阶段。周哲等[9] 将动态情感词 典与决策粗糙集中的三支决策思想相结合,能有效 提高分类器效率。上述学者改进的情感词典主要 是用于中文文本的情感分析,但随着英文的国际 化,网友们在网络平台上发布的言论中除了中文还 夹杂着少许英文甚至混合着多种语言,可见上述研 究存在一定的局限性。栗雨晴等[10]考虑到目前文 本情感分析工作多针对单一语种的状况,提出了一 种双语的情感词典,然后利用半监督高斯混合模型 分类算法和基于对称相对熵的 K 近邻算法对微博 文本进行情感分类,双语词典的提出和成功应用是 一种包含时尚性的进步。Vilares 等[11] 为解决 Twitter上多语言极性分类问题引入了带有情感 标签的代码转换 Twitter 语料库。

以上学者在研究文本情感方面虽然已经意识到了传统的情感词典的局限性并做出了改进,但是仍有一定的局限性,他们并没有突破情感词典的限制。情感词典无论怎样拓展完善都存在"词典"这一边界,它无法涵盖所有情感表达形式且随着时代发展出现的新词无法及时涵盖进去,这使得文本情感判断准确率较低。

3 机器学习方法

学习是人类具有的一种持续性智能行为,目前 计算机也已经初步具备了这种能力,即机器学习。 机器学习的核心是学习,如何让机器像人类一样学 习是机器学习领域的研究重点。基于机器学习对 文本进行情感分析的原理是人工提取文本特征后 由计算机根据某种特定的算法对文本进行处理然 后输出情感分类。相较于完全依赖人工构建情感 词典的方法,机器学习具有明显的优势,一方面能 有效地缓解劳动力的负担且减少非理性判断,另一方面能构建庞大的数据库且能根据时代发展及时对词库进行更新。本文根据机器学习的发展阶段将机器学习分为有监督的机器学习和弱监督的深度学习。

3.1 有监督的机器学习方法

在机器学习方法中朴素贝叶斯 NB(Naive Bayes)和支持向量机 SVM(Support Vector Machine)是常用的监督学习算法,但是有研究指出, NB和 SVM 单独使用时分别会面临独立条件假设 和核函数选择方面的问题,所以 Sharma 等[12] 通过 使用 Boosting 技术整合"弱"支持向量机分类器, 利用了 Boosting 的分类性能,同时使用 SVM 作为 基础分类器,研究结果表明集成分类器在准确率上 明显优于单纯的 SVM 分类器。唐晓波等[13] 提出 了一种基于旋进原则和 AdaBoost 集成技术的回 归 SVM 情感分类模型,实现了文本情感强度阈值 的可视化。Manek 等[14]提出了基于基尼指数的支 持向量机分类器的特征选择方法。考虑到目前以 支持向量机为代表的文本情感倾向分析方法中,几 乎没有考虑文本中包含的语义信息,对文本情感分 类过度依赖于情感词的统计, Abbasi 等[15] 提出一 个基于规则的多元文本特征选择方法,不仅考虑了 语义信息还应用了语法特性之间的关系,可以有效 剔除文本中的杂质、不相关信息和冗余特征。此 外,韩开旭[16] 在考虑语义的基础上提出了基于概 率潜在语义分析的 Fisher 核函数,让带有概率特 征的潜在语义信息作为分类特征,并且根据潜在语 义分析方法 LSA (Latent Semantic Analysis)改进 了概率潜在语义分析 PLSA (Probabilistic Latent Senmantic Indexing)的参数初始化方法。Shi 等[17] 在分析了相关的基本知识和语义特征的基础 上介绍了在随机条件基础上的情感信息联合识别 模型,并给出了词语情感强度计算公式。

计算机在执行文本情感分析任务时通常先是 对文本的每个词单独进行分析,挖掘可能带有情感 色彩的词,然后整合句子中的情感词判断句子的情感,层层叠加,进而判断整个文本的情感极性。然 而,文本中的每个词对文本情感分析不是同等重要 的,而计算机无法自动判断词语的重要性,于是 Deng 等[18] 根据词语在整个文本中的重要性和表 达情感的重要性这两个因素提出监督的词语权重 赋予方案。评论类的文本缺乏逻辑性,文本多呈无 序性,一般的监督学习算法处理无序文本时准确率 较低,Perikos等[19]设计了一种集成分类器,它是基于3个分类器:第1个和第2个是统计学(朴素贝叶斯和最大熵),第3个是基于知识的工具,对自然语言句子进行深入分析。类似地,Tripathy等[20]将文本分别以1个词、2个词、3个词以及其组合的方式划分,然后分别用朴素贝叶斯、最大熵、随机梯度下降和支持向量机的方法进行评论情感分析。Tripathy等不仅对不同方法进行对比,还与组合的形式进行对比。

有监督的机器学习方法相较于构建情感词典的方法虽然有了一定的进步,但是局限性也比较明显。首先,有监督的机器学习方法主要是依赖分类器,还是需要人工对文本特征进行标记。其次,有监督的机器学习方法是计算机根据已有程序对文本进行重复机械操作,并没有"学习"的过程,在进行文本情感分析时不可避免地会产生无效作业。效率不高的有监督学习模型无法适应大数据时代的要求。

3.2 弱监督的深度学习方法

传统的文本情感分析方法主要有人工构建情感词典的方法或基于监督的机器学习模型,但是这2种方法不仅耗费大量的人力,而且在大数据时代任务完成效率和任务完成质量较低。深度学习可以通过构建网络模型模拟人脑神经系统对文本进行逐步分析、特征抽取且自动学习优化模型输出,以提高文本分类的正确性。

神经网络模型的使用不可避免地要涉及词向 量嵌入技术,即将人类语言转换成机器语言,例如 Word2Vec, Giatsoglou 等[21] 将 Word2Vec 提供的 上下文敏感编码与词典提供的情感信息相结合。 虽然词向量嵌入技术考虑了单词的上下文,但是忽 略了整体文本的情感,Tang等[22]提出通过在情感 嵌入中将文本的情感信息连同词的情境一起编码 来解决这个问题,并开发了一个具有裁剪损失功能 的神经网络自动收集情感信号。Fernández-Gavilanes等[23]提出新的无监督情感分析算法,该算法 使用了依存句法来判断情感的极性。在进行文本 情感分析任务时,经常出现同一句子中有情感极性 不一致的多个情感词,梁斌等[24]认为注意力机制 (Attention Mechanism)能有效解决上述问题,于 是将词向量注意力机制、词性注意力机制和位置注 意力机制相结合构造了多注意力卷积神经网络 MATT-CNN (Multi-ATTention Convolution Neural Networks).

RNN是深度学习中经常用到的网络模型之

一,已经在自然语言处理中广泛应用。一般提到 RNN 网络模型是指循环神经网络(Recurrent Neural Network),是一种基于时间序列的模型,但 RNN 也可指代递归神经网络(Recursive Neural Network),是一种注重结构层次的网络模型。刘 金硕等[25]将预训练好的词向量作为下层递归神经 网络的输入,然后将递归神经网络输出的句子向量 以时序逻辑作为上层循环神经网络的输入,有效地 联合2种网络,能解决分类器正确率低下的问题。 谢铁等[26] 通讨利用深度递归神经网络算法来捕获 句子语义信息,并引入中文"情感训练树图资料库" 作为训练数据来发现词语情感信息。在进行文本 情感分析任务时,多数使用的数据库是评论类的短 文本数据库,由于文本较短、特征比较稀疏,传统特 征抽取方法准确率较低。Sun 等[27] 将深度信念网 络和特征抽取方法结合得到拓展特征抽取方法,以 解决短文本抽取特征稀疏的问题。曹宇慧[28] 着力 于解决循环神经网络 RNN 无法学习长距离依赖 信息的问题和卷积神经网络 CNN 的文本情感分 析模型中全连接分类层不能够有效地对非线性分 布的数据进行情感分类的问题。考虑到传统方法 无法获取文本语义信息的弊端,朱少杰[29]引入了 基于深度学习的半监督 RAE(Recursive AutoEncoders)方法,该方法在特征维度较低时能取得较 高的准确率。长期短期记忆网络 LSTM (Long Short Term Memory Network) 是一种特殊的 RNN, RNN 无法解决长距离依赖问题, 而 LSTM 可以捕获文本之间的依赖关系,并能够长时间存储 信息。周瑛等[30]提出基于多注意力机制的 LSTM 模型,应用于华为 P10 闪存门的微博评论中用以 分析网民情绪。Hu等[31]在 LSTM 模型的基础上 建立了一个关键词词库,能帮助挖掘文本中的潜在 语言,可以进一步提高文本极性判断的正确性。 Ma等[32]提出在LSTM中添加一个由目标级和句 子级的注意模型组成的叠加注意机制,称为感知 LSTM,特别关注于利用深层神经序列模型中的常 识知识。LSTM 是一种有效的链式循环神经网 络,链式结构不能有效表征语言的结构层次信息, 梁军等[33]将 LSTM 扩展到基于树结构的递归神 经网络,用于捕获文本深层次的信息。对语句级和 篇章级的文本情感分析任务难度是不一样的,所以 一般模型不具有普适性。田竹[34]针对语句级的文 本分析任务提出了结合空间递归网络的具有记忆 能力的门控制单元 GRU(Gated Recurrent Unit) 的结构递归门控制单元 RSGRU(RecurSive Gated

Recurrent Unit)网络模型,针对篇章级的任务提出 了结合卷积神经网络和双向 GRU 网络的卷积神 经-门控制单元网络模型 CNN-BGRN (Convolutional Neural Networks-Bidirectional Gated Recurrent Unit)。该学者虑到了文本分析任务的差 异性,但是没有进一步提出解决任务中情感强度辨 识问题的模型。在使用深度学习方法进行文本情 感分析时卷积神经网络模型使用较为广泛,但是卷 积神经网络不考虑文本的潜在话题。Zhou 等[35] 提出了一种基于 CNN 的多样化受限波尔兹曼机 RBM(Restricted Boltzmann Machine) 方法,即使 用 CNN 将句子映射到序列级别的特征空间,然后 利用多样化的 RBM 来模拟潜在的文本主题。类 似地,考虑到卷积神经网络通常忽视文本的结构信 息,并且容易在训练过程中出现过拟合情况,杜昌 顺等[36]采用分段池化策略,分段提取句子的主要 特征,并引入 Dropout 算法有效解决了模型过拟合 问题并提升了模型的泛化能力。卷积网络不仅能 有效处理纯文本,并由于卷积神经网络的多个卷积 核能分别处理部分信息的抽取特征工作,相较于其 他神经网络的全连接方式效率更高,所以卷积网络 也能处理图片形式的文本,特别是对边缘模糊的文 本图像,CNN 也能以较高的准确率完成任务。Xu 等[37]提出了一种基于深度卷积神经网络的模糊图 像边缘提取算法,减少卷积核的负荷。张志华[38] 除了构建情感模块,将情感词向量应用于整句情感 分析任务中,还将情感词向量应用在词或短语的情 感强度预测任务中。

情感分析是自然语言处理领域的重要研究问题,但是现有的分析方法往往难以克服样本偏置与领域依赖问题,严重制约了情感分析的发展和应用。为此,吴冬茵等[39]提出先利用深度神经网络获得文本的分布式表示,然后基于高斯过程迁移与测试集数据分布相符的高质量样例扩充训练数据集,以提高文本情感分类系统的性能。

从上述研究中可以看出众多学者已经意识到 了深度学习的优越性,并着力于将深度学习应用于 文本情感分析任务,从而达到优化分类的目的。但 是,目前深度学习在文本情感分析这一领域还处于 初步探索阶段。

4 其他方法

情感词典和机器学习是众多学者在研究文本情感时普遍采用的方法,事实上除了这2种方法还

有很多其他的方法,部分学者已经将这些方法应用 到了文本情感分析任务中。高歌等[40]利用概念层 次网络理论 HNC(Hierarchical Network of Concepts)中的概念层次和对偶理论,并根据符号生成 规则将一些新词用符号表示,便于计算情感值。针 对带有主观性质的文本如产品评论, Pandey 等[41] 提出了一种基于 K 值和杜宇搜索的新型启发式算 法。杨立公等[42]提出一种基于马尔科夫逻辑网的 句子情感分析方法,并与深度学习相结合实现了不 同领域知识的迁移。以往的情感分析研究多注重 词语的情感识别,但是同一词语在不同语境中不一 定表现出同样的情感极性,而且也不一定与整个文 本情感倾向一致,因此赵妍妍等[43]提出了一种自 动采集大量句法路径的方法,继而基于句法路径精 确匹配算法自动获取情感句中的情感评价单元。 进一步地,她还提出了一种基于编辑距离的句法路 径匹配改进策略来提高系统的性能。同样认为局 部情感无法代表整个文本情感,樊娜等[44]先通过 两层随机场模型获取局部文本情感,然后采用加权 K 近邻算法获得文本的全局情感。通常对文本讲 行情感分类时,假设发布文本的个体是独立的,但 Grandi 等[45]认为这样处理一组数据可能导致结果 不准确或错误,于是提出一种将个人情感和偏好信 息相结合的博尔达算法,以便对集体情感进行更准 确的分析。Eliacik等[46]也认同个体不是独立的, 一个个体对团体的信任程度会影响情感分类结果, 所以他们将社交网络考虑在内并重点关注有影响 力的用户。在进行文本情感分析时,不可避免地会 出现杂质导致某些信息不明确,针对这一问题, Dragoni 等[47] 提出通过训练集的学习用模糊逻辑 表示极性。为了减少对特征选择技术的依赖, Yousefpour 等[48] 对特征子集选择进行了研究并致 力于获得高质量的最小特征子集,提出了一种基于 序列和频次整合的选择方法。事实上,不少研究者 仅仅重点关注文本情感分析方法的优化研究,倾向 于忽视结果的检测, Mukhtar 等[49] 使用了3种标 准评估方法: McNemar 检验法、kappa 统计和均方 根误差对文本情感分析结果的可靠性进行检验。 Appel 等[50] 用诱导有序加权平均量子对多个情感 分析系统的结果进行整合,整合之后的结果才是实 际意义上的"大多数人的观点",能更客观地判断文 本的情感极性。事实上,多数文本情感研究倾向于 只将情感大致分为积极、中立和消极3种状态,而 这不能准确地反映情感强度和情感状态。针对这 一问题唐晓波等[51] 改进了 TF-IDF 算法计算出了

情感词与特征的权重,能更准确地得出情感强度。

上述的其他情感分析方法虽然不是主流的方法,应用范围不广,但这些方法颇具创造性。这些方法是文本情感分析发展过程中不可缺少的研究成果,它们有望成为文本情感研究领域新的突破口。

5 结束语

信息技术的快速发展带来了电子商务的发展 热潮和网络平台的急剧增加,对这些平台上的大量 言论信息进行情感分析所得到的结果可以用于网 络平台评论分类、产品分析推荐、消费预测等方面, 具有极高的商业价值。传统的文本情感分析方法 耗费大量人力资源,然而人工提取的特征覆盖面有 限且人工的非理性行为会影响结果的正确性,因此 传统方法不具有普适性。随着深度学习能自动提 取特征、学习修正输出、可以处理非线性复杂数据 等优势凸显,深度学习的方法在自然语言处理方面 正受到众多学者的关注,可以预测深度学习的方法 将成为文本情感分析研究的重要趋势。

本文综述了文本情感分析的主要方法,主要阐述了相关方法的研究发展状况并进行了比较研究。 国内外相关学者都已经意识到传统文本情感极性分析方法的局限性,也投入到了深度学习的应用研究中,并通过对基于深度学习的文本情感分析的研究得出了一些研究成果,但是还有待进一步发展。本文针对文本情感分析现状,提出如下未来研究中值得关注的几个方向。

(1)深度学习在文本情感分析领域的应用。弱监督的深度学习在进行文本情感分析时不仅能自动提取文本特征还能在程序执行过程中自动学习并修正模型,与情感词典和有监督的机器学习模型相比,它的优越性显而易见。在今后的研究中应重点关注深度学习的方法在具体文本情感分析问题中的应用。

(2)情感强度的划分。目前的文本情感分类普遍分为3类标签:积极的、中立的和消极的,但无法对某一类型中的情感强度进行区分。例如,对"毒奶粉"这一事件的微博评论进行情感分析,"极其愤怒"比"消极"更能说明公众的心理,更能显示事件的严重性。在今后的研究中实现情感强度的划分是非常必要的。

(3)关注其他研究方法。学者在研究文本情感 分类时不能将注意力全部集中在深度学习的应用 上,对其他方法的关注也是必不可少的,其他方法也可能是未来研究的突破口。

参考文献:

- [1] Whissell C. Objective analysis of text; II. Using an emotional compass to describe the emotional tone of situation comedies [J]. Psychological Reports, 1998, 82(2):643-646.
- [2] Whissell C. Using the revised dictionary of affect in language to quantity the emotional undertones of samples of natural language [J]. Psychological Reports, 2009, 105(2):509-521.
- [3] Li Shou-shan, Li Yi-wei, Huang Ju-ren, et al. Construction of Chinese sentiment lexicon using bilingual information and label propagation algorithm [J]. Journal of Chinese Information Processing, 2013, 27(6):75-81. (in Chinese)
- [4] Yang Ai-min, Lin Jiang-hao, Zhou Yong-mei. Method on building Chinese text sentiment lexicon [J]. Journal of Frontiers of Computer Science and Technology, 2013, 7(11):1033-1039. (in Chinese)
- [5] Rao Yang-hui, Lei Jing-sheng, Liu Wen-yin, et al. Building e-motional dictionary for sentiment analysis of online news [J].
 World Wide Web-internet & Web Information Systems,
 2014,17(4):723-742.
- [6] Wang Zhi-tao, Yu Zhi-wen, Guo Bin, et al. Sentiment analysis of Chinese micro blog based on lexicon and rule set [J], Computer Engineering and Applications, 2015, 51(8):218-225. (in Chinese)
- [7] Zhou Jie. Research of sentiment analysis for the Chinese microblog based on emotional dictionary and sentence classification [D]. Yinchuan, Ningxia University, 2016. (in Chinese)
- [8] Zhang Ke-liang, Huang Jin-zhu, Cao Rong, et al. Text sentimental orientation analysis based on HNC contextual framework and sentimental dictionaries [J]. Journal of Shandong University (Natural Science), 2016, 51 (7): 51-58. (in Chinese)
- [9] Zhou Zhe, Shang Lin. A sentiment analysis method based on dynamic lexicon and three-way decision [J]. Journal of Shandong University (Engineering Science), 2015, 45(1): 19-23. (in Chinese)
- [10] Li Yu-qing, Li Xin, Han Xu, et al. A bilingual lexicon-based multi-class semantic orientation analysis for microblogs [J]. Acta Electronica Sinica, 2016, 44 (9): 2068-2073. (in Chinese)
- [11] Vilares D, Alonso M A, Gmez-Rodrguez C. Supervised sentiment analysis in multilingual environments [J]. Information Processing and Management, 2017, 53(3):595-607.
- [12] Sharma A, Dey S. A boosted SVM based ensemble classifier for sentiment analysis of online reviews [J]. ACM SIGAPP Applied Computing Review, 2013, 13(4): 43-52.
- [13] Tang Xiao-bo, Yan Cheng-xi. A sentimental classification model based on SPIPRO principle and support vector ma-

- chine [J]. Information Studies: Theory & Application, 2013,36(1):98-103, (in Chinese)
- [14] Manek A S, Shenoy P D, Mohan M C, et al. Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier [J]. World Wide Web-internet & Web Information Systems, 2017, 20(2):135-154.
- [15] Abbasi A, France S, Zhu Zhang, et al. Selecting Attributes for sentiment classification using feature relation networks [J]. IEEE Transactions on Knowledge & Data Engineering, 2011, 23(3):447-462.
- [16] Han Kai-xu. Text sentiment analysis based on SVM [D].
 Daging; Northeast Petroleum University, 2014. (in Chinese)
- [17] Shi Han-xiao, Zhan Wen-ping, Li Xiao-jun. A supervised fine-grained sentiment analysis system for online reviews [J]. Intelligent Automation & Soft Computing, 2015, 21 (4):589-605.
- [18] Deng Zhi-hong, Luo Kun-hu, Yu Hong-liang. A study of supervised term weighting scheme for sentiment analysis [J]. Expert Systems with Applications, 2014, 41(7): 3506-3513.
- [19] Perikos I, Hatzilygeroudis I. Recognizing emotions in text using ensemble of classifiers [J]. Engineering Applications of Artificial Intelligence, 2016, 51, 191-201.
- [20] Tripathy A, Agrawal A, Rath S K. Classification of sentiment reviews using n-gram machine learning approach [J]. Expert Systems with Applications, 2016, 57:117-126.
- [21] Giatsoglou M, Vozalis M G, Diamantaras K I, et al. Sentiment analysis leveraging emotions and word embeddings
 [J]. Expert Systems with Applications, 2017, 69:214-224.
- [22] Tang Du-yu, Wei Fu-ru, Qin Bing, et al. Sentiment embeddings with applications to sentiment analysis [J]. IEEE Transactions on Knowledge & Data Engineering, 2016, 28 (2):496-509.
- [23] Fernández-Gavilanes M, álvarez-López T, Juncal-Martínez J, et al. Unsupervised method for sentiment analysis in online texts [J]. Expert Systems with Applications, 2016, 58: 57-75.
- [24] Liang Bin, Liu Quan, Xu Jin, et al. Aspect-based sentiment analysis based on multi-attention CNN [J]. Journal of Computer Research and Development, 2017, 54(8): 1724-1735.

 (in Chinese)
- [25] Liu Jin-shuo, Zhang Zhi. Sentiment analysis on food safety news using joint deep neural network model [J]. Computer Science, 2016, 43(12):277-280. (in Chinese)
- [26] Xie Tie, Zheng Xiao, Zhang Lei, et al. Sentiment classification of Chinese short text based on parallelized recursive neural network [J]. Computer Applications and Software, 2017,34(3):205-211. (in Chinese)
- [27] Sun Xiao, Li Cheng-cheng, Ren Fu-ji. Sentiment analysis for Chinese microblog based on deep neural networks with convolutional extension features [J]. Neurocomputing, 2016,

- 210(3):227-236.
- [28] Cao Yu-hui. The research on text sentiment analysis based on deep learning [D]. Harbin: Harbin Institute of Technology, 2016. (in Chinese)
- [29] Zhu Shao-jie. Research on text sentiment classification based on deep learning [D]. Harbin; Harbin Institute of Technology, 2014. (in Chinese)
- [30] Zhou Ying, Liu Yue, Cai Jun. Sentiment analysis of microblogs based on attention mechanism [J], Information Studies; Theory & Application, 2018, 41(3):89-94. (in Chinese)
- [31] Hu Fei, Li Li, Zhang Zi-li, et al. Emphasizing essential words for sentiment classification based on recurrent neural networks [J]. Journal of Computer Science and Technology, 2017, 32(4), 785-795.
- [32] Ma Yu-kun, Peng Hai-yun, Khan T, et al. Sentic LSTM: A hybrid network for targeted aspect-based sentiment analysis [J]. Cognitive Computation, 2018, 10(4): 639-650.
- [33] Liang Jun, Chai Yu-mei, Yuan Hui-bin, et al. Polarity shifting and LSTM based recursive networks for sentiment analysis [J]. Journal of Chinese Information Processing, 2015, 29(5):152-159. (in Chinese)
- [34] Tian Zhu. Research on sentiment analysis based on deep feature representation [D]. Weihai: Shandong University, 2017. (in Chinese)
- [35] Zhou Yu, Xu Rui-feng, Gui Lin. A sequence level latent topic modeling method for sentiment analysis via CNN based diversified restrict Boltzmann machine [C] // Proc of International Conference on Machine Learning and Cybernetics, 2016:356-361.
- [36] Du Chang-shun, Huang Lei. Sentiment analysis with piecewise convolution neural network [J]. Computer Engineering & Science, 2017, 39(1):173-179. (in Chinese)
- [37] Xu Xiang-yu, Pan Jin-shan, Zhang Yu-jin, et al. Motion blur kernel estimation via deep learning [J]. IEEE Transactions on Image Processing, 2018, 27(1):194-205.
- [38] Zhang Zhi-hua, Deep learning based sentiment word vector towards sentiment analysis[D]. Weihai; East China Normal University, 2016. (in Chinese)
- [39] Wu Dong-yin, Gui Lin, Chen Zhao, et al. Sentiment analysis based on deep representation learning and gaussian processes transfer learning [J]. Journal of Chinese Information Processing, 2017, 31(1):169-176. (in Chinese)
- [40] Gao Ge, Luo Jun-mei, Wang Yu. Analyzing textual sentiment based on HNC theory [J]. Data Analysis and Knowledge Discovery, 2017, 1(8):85-91. (in Chinese)
- [41] Pandey A C, Rajpoot D S, Saraswat M, Twitter sentiment analysis using hybrid cuckoo search method [J], Information Processing and Management, 2017, 53(4), 764-779.
- [42] Yang Li-gong, Tang Shi-ping, Zhu Jian. A markov logic network based sentence sentimental analysis method [J].

 Transactions of Beijing Institute of Technology, 2013, 33

- (6):600-604. (in Chinese)
- [43] Zhao Yan-yan, Qin Bin, Che Wan-xiang, et al. Appraisal expression recognition based on syntactic path [J], Journal of Software, 2011, 22(5):887-898. (in Chinese)
- [44] Fan Na, An Yi-sheng, Li Hui-xian. Research on analyzing sentiment of texts based on k-nearest neighbor algorithm
 [J]. Computer Engineering and Design, 2012, 33(3):11601164. (in Chinese)
- [45] Grandi U, Loreggia A, Rossi F, et al. A Borda count for collective sentiment analysis [J]. Annals of Mathematics & Artificial Intelligence, 2016, 77(3-4): 281-302.
- [46] Eliacik A B, Erdogan N. Influential user weighted sentiment analysis on topic based microblogging community [J]. Expert Systems with Applications, 2018, 92:403-418.
- [47] Dragoni M, Petrucci G, A fuzzy-based strategy for multi-domain sentiment analysis [J]. International Journal of Approximate Reasoning, 2018, 93(10); 59-73.
- [48] Yousefpour A, Ibrahim R, Hamed H N A. Ordinal-based and frequency-based integration of feature selection methods for sentiment analysis [J]. Expert Systems with Applications, 2017,75:80-93.
- [49] Mukhtar N, Khan M A, Chiragh N. Effective use of evaluation measures for the validation of best classifier in Urdu sentiment analysis [J]. Cognitive Computation, 2017, 9(4): 446-456.
- [50] Appel O, Chiclana F, Carter J, et al. A consensus approach to the sentiment analysis problem driven by support-based IOWA majority [J]. International Journal of Intelligent Systems, 2017, 32(9):947-965.
- [51] Tang Xiao-bo, Lan Yu-ting. Sentiment analysis of microblog product reviews based on feature ontology [J], Library and Information Service, 2016, 60(16), 121-127. (in Chinese)

附中文参考文献:

- [3] 李寿山,李逸薇,黄居仁,等. 基于双语信息和标签传播算法的中文情感词典构建方法[J]. 中文信息学报,2013,27(6):75-81.
- [4] 阳爱民,林江豪,周咏梅.中文文本情感词典构建方法 [J]. 计 算机科学与探索,2013,7(11):1033-1039.
- [6] 王志涛,於志文,郭斌,等.基于词典和规则集的中文微博情感分析[J].计算机工程与应用,2015,51(8):218-225.
- [7] 周杰.基于情感词典与句型分类的中文微博情感分析研究 [D].银川:宁夏大学,2016.
- [8] 张克亮,黄金柱,曹蓉,等.基于 HNC 语境框架和情感词典的 文本情感倾向分析 [J]. 山东大学学报(理学版),2016,51 (7):51-58.
- [9] 周哲,商琳. 一种基于动态词典和三支决策的情感分析方法 [J]. 山东大学学报(工学版),2015,45(1):19-23.
- [10] 栗雨晴,礼欣,韩煦,等. 基于双语词典的微博多类情感分析方法[J]. 电子学报,2016,44(9):2068-2073.
- [13] 唐晓波,严承希.基于旋进原则和支持向量机的文本情感分析研究[J].情报理论与实践,2013,36(1):98-103.

- [16] 韩开旭. 基于支持向量机的文本情感分析研究 [D]. 大庆: 东北石油大学,2014.
- [24] 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 [J]. 计算机研究与发展,2017,54(8):1724-1735.
- [25] 刘金硕,张智.一种基于联合深度神经网络的食品安全信息情感分类模型[J].计算机科学,2016,43(12):277-280.
- [26] 谢铁,郑啸,张雷,等.基于并行化递归神经网络的中文短文本情感分类[J].计算机应用与软件,2017,34(3);205-211.
- [28] 曹字慧. 基于深度学习的文本情感分析研究 [D]. 哈尔滨: 哈尔滨工业大学硕士论文,2016.
- [29] 朱少杰. 基于深度学习的文本情感分类研究 [D]. 哈尔滨: 哈尔滨工业大学硕士论文,2014,
- [30] 周瑛,刘越,蔡俊. 基于注意力机制的微博情感分析 [J]. 情报理论与实践,2018,41(3);89-94.
- [33] 梁军,柴玉梅,原慧斌,等. 基于极性转移和 LSTM 递归网络的情感分析 [J]. 中文信息学报,2015,29(5):152-159.
- [34] 田竹. 基于深度特征提取的文本情感极性分类研究 [D]. 威海:山东大学,2017.
- [36] 杜昌顺,黄磊.分段卷积神经网络在文本情感分析中的应用 [J]. 计算机工程与科学,2017,39(1):173-179.
- [38] 张志华. 基于深度学习的情感词向量及文本情感分析的研究 [D]. 上海: 华东师范大学, 2016.
- [39] 吴冬茜,桂林,陈钊,等.基于深度表示学习和高斯过程迁移 学习的情感分析方法[J].中文信息学报,2017,31(1):169-176.
- [40] 高歌,罗珺玫,王宇. 基于 HNC 理论的文本情感倾向性分析

- 「J]. 数据分析与知识发现,2017,1(8):85-91.
- [42] 杨立公,汤世平,朱俭.基于马尔科夫逻辑网的句子情感分析方法[J],北京理工大学学报,2013,33(6):600-604.
- [43] 赵妍妍,秦兵,车万翔,等. 基于句法路径的情感评价单元识别[J]. 软件学报,2011,22(5):887-898.
- [44] 樊娜,安毅生,李慧贤. 基于 K-近邻算法的文本情感分析方法研究[J]. 计算机工程与设计,2012,33(3);1160-1164.
- [51] 唐晓波, 兰玉婷. 基于特征本体的微博产品评论情感分析 [J]. 图书情报工作, 2016, 60(16): 121-127.

作者简介:



洪巍(1983-),男,安徽滁州人,博士, 副教授,研究方向为食品安全与网络舆情。 E-mail; hongwei@jiangnan. edu. cn

HONG Wei, born in 1983, PhD, associate professor, his research interests include

food safety, and internet public opinion.



李敏(1995-),女,江苏扬州人,硕士生,研究方向为基于深度学习对食品與情进行文本情感分析。E-mail:17851308760@163,com

LI Min, born in 1995, MS candidate,

her research interest includes text sentiment analysis of food public opinion based on deep learning.