

文章编号: 1003-0077 (2011) 00-0000-00

基于情感特征向量的 Twitter 情感分类方法研究*

易顺明¹, 易昊², 周国栋³

(1. 沙洲职业工学院 电子信息工程系, 江苏 苏州 215600;

(2. 美国 Rutgers 大学 数学系, 美国 新泽西;

3. 苏州大学 计算机科学与技术学院, 江苏 苏州 215006)

摘要: 公众情感在包括电影评论、消费者信心、政治选举、股票走势预测等众多领域发挥着越来越大的影响力。面向公共媒体内容开展情感分析是分析公众情感的一项基础工作。经典的基于词频特征向量的特征提取方法, 主要利用词频作为文本分类的依据, 而词频与情感信息之间的关系并不紧密。针对 Twitter 推文的特点, 本文提出了一种采用基于情感特征向量的 Twitter 推文情感分类方法。该方法首先通过对推文进行数据清洗、词形还原、词性标注和词汇向量化; 其次, 将单词匹配到情感词典中; 最后, 利用每个单词的正向情感、负向情感取值生成情感特征向量, 通过 MNB、SVM 等机器学习方法训练模型, 对推文的情感进行分类。实验结果表明采用情感特征向量的 Twitter 推文情感分类方法能够获得更佳的性能。

关键词: 推文; 情感分类; 情感词典

中图分类号: TP391

文献标识码: A

Twitter Sentiment Classification with Sentimental Feature Vector

YI Shunming¹, YI Hao², ZHOU Guodong³

(1. Department of Electronic Information Engineering, Shazhou Institute of Technology, Suzhou, Jiangsu 215600, China;

2. Department of Mathematics, Rutgers University, New Brunswick, NJ, USA;

3. School of Computer Sciences and Technology, Soochow University, Suzhou, Jiangsu 215006, China;)

Abstract: Public sentiment plays an increasingly important influence in a variety of areas, including movie reviews, consumer confidence, political elections and stock trend predictions. Sentiment classification on the content of the public media is a fundamental task for analyzing public sentiment. A classic model for sentiment classification is based on word frequency model which takes advantages of word frequency in text classification. However, the relationship between word frequency and sentiment information is not actually close as we expected. Based on the characteristics of Twitter tweets, this paper presents a distinct approach of sentiment classification using sentimental feature vector instead of word-frequency feature vector. First, preprocessing is done to clean, lemmatize, and POS tag each word in a single tweet; Second, with the sentiment dictionary, each word is attached with a score corresponding to positive or negative sentiment respectively so as to get the sentimental feature vector for each tweet; Third, sentiment of tweets are classified by training models of different algorithms such as Multinomial Naïve Bayes (MNB) and Support Vector Machine (SVM). Empirical studies show that our sentimental feature vector is beneficial for Twitter sentiment classification.

Key words: Tweets; Sentiment Classification; Sentiment Lexicon

* 收稿日期:

定稿日期:

基金项目: 本文承国家自然科学基金项目 (61003155, 61273320) 资助

1 引言

Twitter 是一个被公众广泛接受的社交媒体平台, 上面有 6.45 亿的活跃用户, 每天有 135000 个全新用户, 信息呈现海量增长。Twitter 有许多特点, 首先, 与 Facebook 不同的是, Twitter 是基于文本的, 可以通过 Twitter 接口注册下载, 便于作为自然语言处理所需的语料库。其次, Twitter 规定了每一个推文不超过 140 个字, 实际推文中的文本长短不一、长度一般较短, 有些只有一个句子甚至一个短语, 这对其开展情感分类标注带来许多困难。再者, 推文常常是随性所作, 内容中包含情感的元素较多, 口语化内容居多, 缩写随处都在, 并且使用了许多网络用语, 情绪符号、新词和俚语随处可见。因此, 与正式文本非常不同。如果采用那些适合处理正式文本的情感分类方法来对 Twitter 推文进行情感分类, 效果将不尽人意。

情感分类研究旨在对文本所表达的情感倾向进行分类, 将其标注为正向情感 (positive)、负向情感 (negative)、中性情感 (neutral) 等。目前, 主流的情感分类的研究方法, 仍然是机器学习方法中的监督学习方法, 这种方法的关键在于特征提取和分类器选择。在特征提取阶段, 最典型的莫过于词汇向量化方法, 目前通用的词汇向量化方法是将文本中所使用的全部单词作为特征向量, 以每个单词在文本中出现的词频作为数据, 辅之于期限频率逆文档频率 (TF-IDF) 作为权重指标, 也可采用 n -gram 语言模型进行处理, 这种基于情感特征向量的词袋 (bag-of-words) 方法在许多文本分类中有很好的效果。然而, 这种情感分类方法并不可靠, 原因在于这种方法只将词频作为特征数据, 而词频与情感之间却并无逻辑上比较严密的关系。

本文提出了一种针对 Twitter 特点的基于情感特征向量的监督学习方法。该方法通过匹配情感词典, 给推文文本中的每个单词赋予情感值。由于推文的情感分类, 取决于推文文本中那些有主观色彩的单词的情感倾向。因此, 用这些情感值作为特征数据来训练监督学习模型, 理论上可以取得较好的效果。实验结果表明本文提出的方法情感分类性能优于基于词频特征向量的监督学习方法。

本文其他部分安排如下: 第 2 节介绍了情感分类的相关工作; 第 3 节针对 Twitter 推文文本的特点进行了语料的预处理, 并介绍了基于词频特征向量模型的词汇向量化过程和机器学习的模型选择和参数设置; 第 4 节提出了基于情感特征向量的监督学习方法; 第 5 节给出两种方法的实验结果和比对分析, 并详细讨论了情感特征向量模型的一些问题; 第 6 节给出相关结论。

2 相关工作

Pang 等^[1]最早用机器学习的方法来研究情感分类问题, 他们将朴素贝叶斯、最大熵、支持向量机等多种分类方法引入情感分类任务中并取得了较好的分类效果。许多研究集中在特征向量的抽取上, Hu 和 Liu^[2]提出了将形容词作为情感分类特征, 苏祺^[3]在中文文本情感分析中则将名词和名词短语作为主题词, Wiebe 等^[4]尝试将代名词、形容词、基数词、情态动词、副词作为分类特征, 用多种词性的单词作为共同作为特征项比选取单一词性的效果要好。同时, 在选取特征向量时, 也有考虑 n -gram 语言模型的, Pang 等^[1]采用的是 Unigram, Dave 等^[5]采用 Bigram 作为特征项对产品评论进行情感分类, 得到分类器训练的效果更好。Matsumoto 等^[6]将 unigram、bigram 一起作为特征项, 分类准确率较高。此外, 也有学者对否定词对文本的情感分类的影响开展了研究, 经研究发现利用否定词词典和否定句的规则匹配, 分类效果有所提高^[7-8]。随着机器学习主题模型 (Topic Model) 的兴起, 也有研究者将此应用到情感分类研究中来, 像 Mei 等^[9]就将文档的分类看作是主题模型和情感模型的混合。

情感分类最直接的方法就是情感词的选择和使用情感词典。Turney^[10]在早期的研究中就采用了基于词典的无监督学习方法, 但这种方法的效果较差。Hatzivassiloglou 等^[11]通过筛选获得具有情感倾向的情感词和短语作为特征, Kim 和 Hovy^[12]介绍了构建情感词集为每个

情感词标注正向、负向情感值的方法,并探讨了否定词、双重否定对句子情感倾向的反向影响。Yu 等^[13]计算句子中所有情感词的平均情感值直接作为句子的情感值,Nasukawa 等^[14]通过主客观分类用规则方法提取主观词句。也有学者^[15-18]使用词典中词语的注释信息来完成评价词语的识别与极性判断。此外,由于很多词存在一词多义现象,采用的情感词典往往有较多的歧义词,Wiebe 等^[19]分析了词义和情感极性的不一致现象。由于在情感分类中标注样本比较难以获取,半监督情感分类也是一个获得了很多研究的方向。多视图学习(Multi-View Learning)是半监督学习中常用的方法,有研究将文本区分为个人(Personal)和非个人视图(Impersonal)^[23],肯定句和否定句视图^[24],新闻和评论视图^[25],或者直接将文本区分为多个随机子空间^[26]等,以上方法都在半监督情感分类中获得了比较好的效果。另外,情绪分类也是情感分析中一个比较重要的研究方向。除了简单的将文本内容作为特征进行情绪分类之外,有研究发现将情绪分类和情感分类结合起来^[27],或者利用情绪标签和上下文的联系^[28]都能够比较好的提升情绪分类的效果。

随着 Twitter 的发展,基于 Twitter 的情感分析也成为研究的一个方向。Pak 等^[20]对 Twitter 语料进行收集与自动标注,建立包含正向、负向情感的主观文本语料库,开展情感分析与意见挖掘的研究,同时也有相关研究将 Twitter 上的笑脸等表情符号作为标签进行分类^[21-22]。

3 语料分析与处理

3.1 语料选择和预处理

本文选取手工标记情感的奥巴马-麦凯恩选举推文语料库(OMD)和 Sentiment140 手动注释数据集用于训练模型,其中 OMD 包含了 2392 条推文,Sentiment140 手动注释数据集包含了 498 条推文,共计 2890 条。

Twitter 上的每个推文不超过 140 个字,相当于一段句子,或者一个句子、一个短语。由于网络上书写的随意性,推文的内容比较杂乱,第一步要通过使用小写和删除字符“RT”、URL 链接、用户名、标点符号、控制字符和数字来清理文本,并删除所有的标点符号。

对语料库的预处理,分成四步:分词、符号转换、去停用词和词形还原。在 Twitter 推文中,由于存在着大量多个单词连在一起的情形。因此,需要进行文本分词。符号转化具体包括情绪符号处理、俚语和缩写处理。情绪符号处理是情感标注任务中一个重要的工作,对推文中的情绪符号的一般操作都是直接删除。但是,在情感标注中,由于这些情绪符号表示了比较强烈的情感色彩,直接去除会丢失重要的信息。在处理中,我们采用了比较简单的方法,即将那些重要的情绪符号直接转换为单词。例如,将小写化后表示正向情感的情绪符号用单词“good”替代,将表示负向情感的表情符号等用单词“sad”或“bad”替代。同样也对推文中大量俚语和缩写进行了处理。在本文中,我们采用了自己定制的停用词表,将那些跟情感相关的词汇从标准停用词表中剔除,同时我们采用了 NLTK 语料库的 WordNet 中的 Morphy 查找词元方法来进行词形还原。

3.2 机器学习模型训练

我们的实验用传统的基于词频特征向量的方法作为比较,采用 Python 机器学习工具包 *scikit-learn* 将预处理过的 Twitter 推文进行向量化,还允许加入 N-gram 语言模型的处理。经过词汇向量化之后,就可以用机器学习的方法来训练模型,进行结果分析了。分别选用 MNB (Multinomial NB)、kNeighbors Classifier、Logistic Regression、SVC 和 kMeans 聚类进行模型拟合训练,根据性能分析选择最佳分类器。

4 基于情感特征向量的情感分类方法

传统的词袋模型使用词频作特征向量作为样本表示方法,并取得了一定的效果。然而,作为推文情感自动标注的输入值,单单考量句子中单词的频次来做情感分析信息量显得不足。这主要是因为推文非常短,词的词频表示很难捕捉足够的情感信息用于情感分类。从推文的情感标注角度看,人们判断一段推文是积极还是消极、是赞成还是反对、是喜欢还是厌恶、

是满意还是不满意，都是从字里行间获得的。也就是说，推文之所以有情感值，可以表达情感，是因为推文是由一个个具有一定情感值的单词和符号组成的。如果能够获得每一个单词和符号的情感倾向或情感取值，机器就可能更准确地分类出推文的情感取向，完成推文的情感分类。

4.1 情感词典的结构分析

本文的方法首先需要利用情感词典来得到每个单词的情感值。目前常用的情感分析方面的词典有 GI(General Inquirer)¹、LIWC(Linguistic Inquiry and Word Count)²、Janyce Wiebe 等人所开发的 MPQA³、美国伊利诺伊芝加哥分校 Bing Liu 开发的 Opinion Lexicon^[23]和 SentiWordNet⁴等。本文采用的是 SentiWordNet。

SentiWordNet 是意大利信息科学研究所研制的英文语义情感极性词典。我们采用 2013 年版的 SentiWordNet 3.0.0，包含 117659 条记录。SentiWordNet 通过对 WordNet 中的词条进行情感分类，并标注出每个词条属于正向情感和 n 负向情感的分值。在 SentiWordNet 中，只对 4 类词性赋予了情感分值，分别为名词、形容词、动词、和副词。

4.2 基于情感特征向量的分类方法

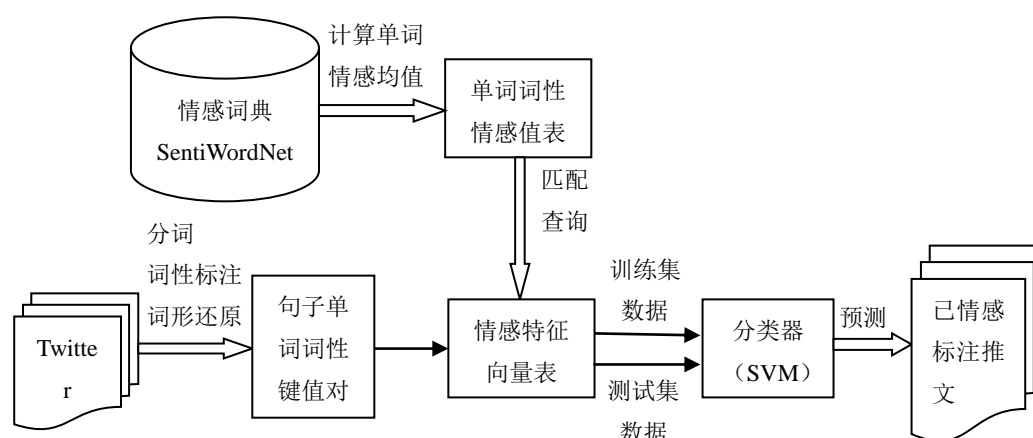


图 1 基于情感特征向量的推文情感分类流程

图 1 给出了基于情感特征向量的分类方法具体实现示意图。首先，需要对经预处理后的推文进行词性标注（Part Of Speech tagging, POS 标注），以确定每个单词的词性。本来采用 NLTK 提供的一款较成熟的 POS 标注器。然后，对单词进行词形还原和词性标签的转换操作。本文利用 WordNet 的 Morphy 查找词元方法来进行词形还原，以保证与情感词典匹配的一致性。最后，经过以上处理所得到的“单词原形#词性”，可以与英文情感词典 SentiWordNet 中的具有相同单词和词性的词项进行匹配，找到相应的 PosScore（正向情感值）和 NegScore（负向情感值），进而计算出每个推文句子的不同情感（正向、负向及中性情感）的平均值，并与 4 种词性的单词词频在句中的占比一起，形成每个推文在情感向量空间上的特征向量。具体格式如“中性情感平均值、正向情感平均值、负向情感平均值、名词词频占比、形容词词频占比、动词词频占比、副词词频占比”。然后将抽取出的情感特征向量作为输入，选用 MNB、SVM 等机器学习模型进行训练，得到拟合后的预测模型。

5 实验

¹ <http://www.wjh.harvard.edu/~inquirer/>

² <http://www.liwc.net>

³ <http://mpqa.cs.pitt.edu>

⁴ <http://sentiwordnet.isti.cnr.it>

在本节中，我们将详细介绍我们的实验方法，以及对于实验结果进行深入的分析。

5.1 实验方法

实验语料库来自 Twitter，在 2890 条推文中，手工标注为正向情感（positive）的推文 723 个、负向情感（negative）的推文 1558 个、中性或无关情感（neutral）的推文 609 个。推文经去除停用词和词形还原之后，Unigram 特征共计 5849 个特征，Unigram+Bigram、Unigram+Bigram+Trigram 特征共计 40536。

本文重点比较基于词频特征向量和基于情感特征向量这两种情感分类方法。我们选取了 80% 的语料作为训练集，20% 的语料作为测试集。评价标准为精确度（Precision），召回率（Recall）和 F1 值（f1-score），并用 P/R AUC 曲线图来比较不同情形下分类器的效果。

5.2 实验结果分析

表 1 给出了基于词频特征向量的情感分类方法在不同词汇向量化处理模式下的分类性能比较。其中 CountVectorizer 表示是基于词个数权重的词频特征向量方法，TfidfVectorizer 是基于 Tf-Idf 权重的词频特征向量，TfidfVectorizer ngram_range=(1,3)表示利用一元、二元、三元词特征的 Tfidf 词频特征向量方法。

表 1 基于词频特征向量的词袋分类模型分类结果结果

词向量化方法	评价指标	MNB	kNeighbors Classifier	Logistic Regression (L2,C=0.48)	SVC (Linear, C=0.08)	SVC (RBF, C=10, gamma=0.10)	kMeans (k=5)
Count Vectorizer	P.	0.651	0.589	0.682	0.645	0.670	0.429
	R.	0.631	0.514	0.689	0.661	0.682	0.258
	F1.	0.637	0.526	0.684	0.639	0.670	0.249
Tfidf Vectorizer	P.	0.636	0.619	0.695	0.316	0.684	0.455
	R.	0.635	0.628	0.637	0.562	0.690	0.190
	F1.	0.633	0.621	0.562	0.405	0.686	0.184
Tfidf Vectorizer ngram_range=(1,3)	P.	0.633	0.619	0.695	0.316	0.684	0.558
	R.	0.642	0.628	0.637	0.562	0.690	0.206
	F1.	0.634	0.621	0.695	0.405	0.686	0.299

从表 1 的结果可以看出以下几点：

- 1) 模型方面：SVC(RBF, C=10, gamma=0.10) 的结果最佳，Logistic Regression(L2,C=0.48)和 MNB 次之，kMeans 的性能最差。综合而言，支持向量机分类算法在情感分类中是一种较佳的分类器。由于 SVC(RBF)存在明显的过拟合，实际使用中选择 SVC(Linear)。
- 2) 词汇向量化的三种方法对模型性能的影响比较：TfidfVectorizer 与 CountVectorizer 相比，总体上性能并无明显提升。例如，在 SVC(RBF)中基本持平，相反在许多训练模型中不升反降，如 MNB、Logistic Regression 中，而在 SVC(Linear)中性能下降了很多。其次，同样是 TfidfVectorizer，采用 Unigram+Bigram+Trigram 混合语言模型也并不比采用 Unigram 语言模型性能变好。

表 2 给出了不同分类目标下基于词频特征向量和基于情感特征向量两种模型的 MNB、SVM 比较结果。具体分类目标包括：区分推文中有无情感倾向的分类（正向/负向情感 vs.

其它情感)、区分正向情感和其它情感的分类, 区分负向情感和其他情感的分类, 以及区分正向情感和负向情感分类。

从表 2 的结果可以看出:

- 1) 基于词频特征向量方法的分类器在上述四种不同分类目标下, MNB 与 SVM (线性核) 的平均得分分别为 0.750、0.692, P/R AUC 均值最高为 0.921, 而本文提出的基于情感特征向量的分类器在相同的分类目标和分类模型下, MNB 与 SVM (线性核) 的平均得分分别为 0.784、0.698, 分别提高了 4.5% 和 0.87%, 其 P/R AUC 均值最高为 0.935, 性能提高了 1.52%。
- 2) 从四种分类目标的性能来看, 当分类器分类对推文作有无情感之分 (pos+neg vs rest) 时, 无论是 SVM 还是 MNB, 其得分均值最高, 分类效果最好。因此, 在实际对推文进行情感分类时, 可以采用的策略是, 首先用分类器来确定推文中是否包含情感, 然后再使用分类器来确定是“正向情感”还是“负向情感”。

表 2 不同分类目标下基于词频和情感特征向量分类方法比较结果

分类目标	分类模型		得分均值	标准差	P/R AUC 均值	标准差
正向/负向情感 vs. 其它情感	词频特征向量模型	MNB	0.783	0.031	0.900	0.022
		SVM	0.784	0.030	0.921	0.025
	情感特征向量模型	MNB	0.822	0.014	0.935	0.008
		SVM	0.783	0.011	0.930	0.011
正向情感 vs. 其它情感	词频特征向量模型	MNB	0.705	0.023	0.773	0.035
		SVM	0.551	0.026	0.805	0.040
	情感特征向量模型	MNB	0.804	0.011	0.615	0.025
		SVM	0.752	0.013	0.613	0.020
负向情感 vs. 其它情感	词频特征向量模型	MNB	0.763	0.019	0.481	0.047
		SVM	0.750	0.026	0.593	0.051
	情感特征向量模型	MNB	0.733	0.018	0.815	0.021
		SVM	0.567	0.013	0.801	0.020
正向情感 vs. 负向情感	词频特征向量模型	MNB	0.749	0.028	0.627	0.047
		SVM	0.681	0.043	0.725	0.045
	情感特征向量模型	MNB	0.775	0.011	0.692	0.040
		SVM	0.688	0.017	0.710	0.034

5.3 问题分析

虽然基于情感特征空间的情感分类方法优于传统基于词频特征空间的情感分类方法, 但是通过分析我们发现我们提出的方法还存在下面一下问题。

- 1) 基于情感特征空间的情感分类方法, 比较依赖于词典所设定的单词的情感值。例如, 在 SentiWordNet 词典中有大量的单词没有正向情感值和负向情感值。总共 117659 条同义词项中, 只有 17016 个同义词项含有正向情感值, 占比 14.46%; 只有 18029 个同义词项含有负向情感值, 占比 15.32%; 多达 88564 个同义词项没有任何情感值, 占比 75.27%。SentiWordNet 和其它的情感词典一样, 并不是手工设定情感值的, 而是通过半监督的 bootstrapping 方法训练模型估计出的, 其自身的准确率也不高。
- 2) 从实验过程看, 基于 WordNet, SentiWordNet 词典中存在着大量相同单词的多项语

义，每个语义项下的情感值可能各不相同。然而，要明确文中使用的是哪些同义词是一个比较有挑战的问题。我们在实际处理中，只是采用了比较简单的方法，即简单地对所有同义词的情感值求平均值。例如，下面例子是一个被手动标注为负向情感的推文句子，经推文小写化后如下：

“coburn every appropriation will be outlined if u sold your vote the voters will know you were bribed tcof”

推文中的一个单词“sold”经词性标注为动词，经词形还原成“sell”，在 SentiWordNet 中作动词时，“sell”有 8 个同义词项，“sell#1”至“sell#8”的正向情感值、负向情感值设置分别为 (0, 0)、(0, 0)、(0, 0)、(0, 0)、(0, 0)、(0.125, 0)、(0.25, 0)、(0, 0.5)。最后我们计算得到的“sell”作为动词的正向情感平均值为 0.046875，负向情感平均值为 0.0625。实际上，从推文语义理解，这里的“sell your vote”，“sell”应作“出卖”解，属于贬义词，情感词至少取 (0, 0.5)，与 0.0625 相比，误差确实太大。

- 3) SentiWordNet 中只有名词、形容词、动词和副词四种词性类别，其它词性的单词没有情感值，一方面导致推文中情感的单词比例太小；另一方面，像感叹词（如“haha”）、疑问词等可能有较强情感倾向的单词被忽略。此外，在推文中存在大量的短句和新词、无语义词，一旦被多次过滤或赋予 0 值，用于情感分析的有效词汇变得很少，导致与手动标注的结果毫无关联。

6 结语

本文旨在研究 Twitter 中推文的情感自动分类问题。针对该问题，本文提出了一种基于情感特征向量的情感分类方法。该方法的核心思想是借助情感词典将推文中的单词转换为情感特征向量的方法，结合分词、数据清洗、去停用词、词形还原、词性标注等语料库处理方法，并将处理结果作为机器学习模型训练的输入数据，选择最佳分类器。实验结果表明，本文提出的基于情感特征向量的分类方法在 Twitter 推文情感分类任务中明显优于传统的基于词频特征向量的分类方法。

在下一步的工作中，我们将围绕本方法的不足，例如词语覆盖面不够、一词多义等问题，改进本文的方法，进一步提升 Twitter 推文情感分类性能。同时，我们将考虑在别的类型的情感分类任务中应用本文提出的基于情感特征向量的分类方法，考虑该方法在其他类型的情感分类任务中的表现。

参考文献

- [1] Pang B, Lee L and Vaithyanathan S. Thump up? Sentiment Classification using Machine Learning Technoques[C] //Proceedings of EMNLP-2002:79-86.
- [2] Hu M., Liu B. Mining and Summarizing Customer Reviews[C]//Proceedings of KDD-2004: 168-177.
- [3] 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.
- [4] Wiebe J, Bruce R, O'Hara T. Development and Use of a Gold Standard Dataset for Subjectively Classification[C]//Proceedings of ACL-1999:246-253.
- [5] Dave K, Lawrence S, Pennock D. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews[C]//Proceedings of WWW-2003:519-528.
- [6] Matsumoto S, Takamura H, Okumura M. Sentiment Classification using Word Sub-sequences and Dependency Sub-trees[C]//Proceedings of PAKDD-2005:301-111.
- [7] 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[J]. 中文信息学报, 2007, 21(1):96-100.
- [8] 娄德成, 姚天叻. 汉语语句主题语义倾向分析方法的研究[J]. 中文信息学报, 2007, 21(5):73-79.
- [9] Mei Q, Ling X, Wondra M, Su H, and Zhai C. Topic Sentiment Mixture Modeling Facets and Opinions in Weblogs[C]//Proceedings of WWW-2007:171-180.
- [10]Turney P. Thumps up or Thumbs down ? Semantic Orientation Applied to Unsupervised Classification of Reviews [C]//Proceedings of ACL-2002:417-424.

- [11] Hatzivassiloglou V, Wiebe J. Effects of Adjective Orientation and Gradability on Sentence Subjectivity[C]//Proceedings of COLING-2000:299-305.
- [12] Kim S M, Hovy E. Determining the Sentiment of Opinions [C]//Proceedings of COLING-2004.
- [13] Yu H, Hatzivassiloglou V. Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences[C]//Proceedings of EMNLP-2003:129-136.
- [14] Nasukawa T, Yi J. Sentiment Analysis: Capturing Favorability using Natural Language Processing[C]//Proceedings of the International Conference on Knowledge Capture, 2003: 70-77.
- [15] Esuli A, Sebastiani F. Determining the Semantic Orientation of Terms through Gloss Analysis[C]// Proceedings of CIKM-2005:617-624.
- [16] Esuli A, Sebastiani F. Determining Term Subjectivity and Term Orientation for Opinion Mining[C]//Proceedings of EACL-2006.
- [17] Andreevskaia A, Bergler S. Mining WordNet for a Fuzzy Sentiment: Sentiment Tag Extraction from WordNet Glosses[C]//Proceedings of EACL-2006:209-216.
- [18] Su F., Markert K. Subjectivity Recognition on Word Senses via Semi-supervised Mincuts[C]//Proceedings of the NAACL-2009:1-9..
- [19] Wiebe J, Mihalcea R. Word Sense and Subjectivity[C]// Proceedings of ACL-2006:1065-1072.
- [20] Pak A, Paroubek P. Twitter as a Corpus for Sentiment Analysis and Opinion Mining[C]//Proceedings of LREC-2010:1320-1326.
- [21] Go A, Bhayani R, Huang R, Huang L. Twitter Sentiment Classification using Distant Supervision[R]. CS224N Project Report. 2009.
- [22] Davidov D, Tsur O, Rappoport A. Enhanced Sentiment Learning using Twitter Hashtags and Smileys[C]//Proceedings of COLING-2010:241-249.
- [23] Ding X., B. Liu and P. Yu. A Holistic Lexicon-Based Approach to Opinion Mining[C] //Proceedings of WSDM-2008:231-240.
- [23] Li S., C. Huang, G. Zhou and S. Lee. 2010. Employing Personal and Impersonal Views in Supervised and Semi-supervised Sentiment Classification[C]. //Proceedings of ACL-2010:414-423.
- [24] Li S, S. Lee, Y. Chen, C. Huang and G. Zhou. 2010. Sentiment Classification and Polarity Shifting[C]. // Proceeding of COLING-2010:635-643.
- [25] Liu H., S. Li, G. Zhou, C. Huang, and P. Li. 2013. Joint Modeling of News Reader's and Comment Writer's Emotions[C]. //Proceedings of ACL-2013:511-515.
- [26] Li S., Z. Wang, G. Zhou and S. Lee. 2011. Semi-supervised Learning for Imbalanced Sentiment Classification[C]. //Proceedings of IJCAI-2011:1826-1831.
- [27] Gao W., S. Li, S. Lee, G. Zhou, C. Huang. 2013. Joint learning on sentiment and emotion classification[C]. //Proceedings of CIKM-2013:1505-1508.
- [28] Li S., L. Huang, R. Wang, and G. Zhou. 2013. Sentence-level Emotion Classification with Label and Context Dependence[C]. //Proceedings of ACL-2015:1045-1053.

作者简介:



易顺明(1966—),男,副教授,主要研究方向为自然语言处理。邮箱: ysm2501@qq.com。



周国栋（1967－），男，教授，博士生导师，主要研究方向为自然语言处理。邮箱：
gdzhou@suda.edu.cn



易昊（1991－），男，硕士生，主要研究方向为金融数据挖掘。邮箱：yihao_zg@hotmail.com