

# 微博情感分析的情感词典构造及分析方法研究

杨立月<sup>1</sup> 王移芝<sup>2</sup>

(1-2 北京交通大学 计算机与信息技术学院, 北京市 100044)

**摘 要:** 为了提高微博情感分析的准确性, 对微博情感分析中的语义规则和情感词典进行了研究。在传统基于情感词典的微博情感分析的基础上对情感词典中的微博情感词典的构造方法做了改进。首先构造情感词典, 主要包括开源情感词典、具有时代特征的网络情感词典、根据情感词的位置特点构造的微博情感词典, 具有明显情感倾向的语气情感词典。在词典构造完成的基础上结合中文语法规则, 主要包括句间关系规则和句型关系规则, 根据句间和句型关系算法计算微博句子的情感倾向性, 将微博文本分为正向、负向和中性三个方面。为了提高微博分类的准确率, 提出构建语气词词典, 并且在语气词权重计算的方法上做出创新, 同时对微博情感词典的构造方法做出了改进。实验证明了此方法能够提高微博情感分析的正确率。

**关键词:** 情感词典 微博情感词典 语气词词典 语义规则 情感分析

**中图分类号:** TP393 **文献标识码** A

## Research of the Construction of Emotion Dictionary and Emotion Analysis in the Emotion Analysis of micro-blog

Yang Li-yue<sup>1</sup>, Wang Yi-zhi<sup>2</sup>

(1-2 College of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

**Abstract** In order to improve accuracy of the sentiment analysis of micro-blog sentences, the semantic rules and sentiment dictionaries in the micro-blog sentiment analysis are studied. the constructing method of micro-blog sentiment dictionary in the sentiment dictionary is improved on the basis of the traditional sentiment analysis of the sentiment dictionary. First of all, constructs the sentiment dictionary including the open-source sentiment dictionary, the net sentiment dictionary with the characteristics of the times, and micro-blog sentiment dictionary based on position of emotional words on the basis of constructed dictionaries, Emotional sentiment dictionary with obvious emotion tendency. Based on the completion of the construction of the lexicon, combined with the Chinese grammar rules, including the rule of relationship between sentences and rules of sentence patterns, the sentiment tendencies of Weibo sentences are calculated according to algorithm of relationship between sentences and sentence patterns, The microblogging texts are divided into three categories: positive, negative and neutral section. In order to improve the classification accuracy, this paper proposes to construct a modal dictionaries and make innovations in the modal weight calculation of the modal particles, and at the same time, to improve the construction method of the micro-blog sentiment dictionary. Experiments show that this method can improve the accuracy of the emotional analysis of micro-blog.

**Keywords** Emotional Dictionary Micro-blog Sentiment Dictionary Modal Word Dictionary Semantic Rules Emotional Analysis

**基金项目:** 课题来源: 国家自然科学基金“面上”(K13A300050)。杨立月(1992-), 女, 硕士, 研究方向 web 开发, 机器学习, 情感分析。王移芝(1953-), 女, 教授, 研究方向为计算机网络与数据库技术, 计算机远程教育。

### 0 引言

近年来随着微博的广泛应用,微博用户越来越多,微博成为人们关注世界、社会、记录生活等的重要平台之一。随着网络的发展,微博情感分析也越来越重要。通过微博情感分析,商家可以及时得到消费者的反馈,并对下一步产销计划作出调整;政府机关根据微博用户对某一热点事件的评论得到大众对热点事件的观点,从而及时控制网络舆情的来保证公共秩序安全,对于消费者来说可以对微博中某一商品评论的情感分析统计,得出这一商品的优缺点,从而根据自身需求决定是否购买此商品。此研究正是基于这样的研究目的,在前人研究的基础上改进了微博情感词典和创建了语气词词典,利用中文语法规则判断微博文本的情感倾向性。

## 1 相关工作

近些年来,微博的广泛使用使得微博情感分析成为热点方向,国外的微博研究较早并且相对较成熟,英文情感词典相对较丰富,如:SentiWordNet 和 Inquirei。而国内对于微博的研究是近几年才开始的,情感词典还不太成熟,目前主要有知网 HowNet 词典和同义词词林、台湾大学和大连理工大学等高校提供的情感词汇库,此研究采用分类效果较好的 HowNet 词典作为基准情感词典。目前对于微博情感分析的研究主要分为两大类:基于情感词典的微博情感分析和机器学习。基于情感词典的微博情感分析就是利用现有的情感词典根据语义规则通过算法求出文本的情感值,机器学习是通过选取文本的一些特征标注训练集和测试集,使用朴素贝叶斯(Naive Bayes)、支持向量机(support vector machine,SVM)和最大熵(maximum entropy)等分类器进行文本分类。文献[1-2]表明,常用的分类方法和特征选取方法有各自适应情况。但都无法带来性能上实质性的变化,还有一些基于模型及研究模型提高分类效果的研究,如王伟<sup>[3]</sup>等人提出一种基于 LDA 主题模型的评论文本情感分类方法,黄发良<sup>[4]</sup>等人提出一种基于多特征融合的微博主题情感挖掘模型 TSMME,证明了该模型有更好的微博主题情感检测功能。

国内研究主要针对于新浪微博,腾讯微博等平台的微博信息进行研究。M Taboada<sup>[5]</sup>等人提出了一种基于词典的文本情感提取方法。谢丽星<sup>[6]</sup>等提出基于 SVM 的层次结构多策略方法进行分类,且引入主题相关特征进行分类。周剑锋<sup>[7]</sup>提出在分析海量微博数据过程中,自动构建情感词典的方法。自动从语料中筛选情感词汇,获取情感新词。肖江<sup>[8]</sup>等人构建了具有

自动识别和扩展功能的领域情感词典具有可行性和准确率。王志涛<sup>[9]</sup>等人提出基于词典和规则集的中文微博情感分词方法,根据规定的语义规则结合词典进行情感分析。王振宇<sup>[10]</sup>等人提出一种 HowNet 和 PMI(Point Mutual Information)相融合的词性极性计算方法。实验结果中微平均和宏平均性能提高了 5%。

分析以上理论,通过改进词典对微博进行分析和研究是微博分析的重点,本文与传统的方法不同的是改进了微博情感词典的构建方法,在传统新词根据点互信息计算进行筛选的基础上结合新词根据积极训练语料和消极训练语料的文档频数差筛选出新词并加入词典,同时研究提出构建语气词词典。实验结果证明该方法具有可行性和准确率。

## 2 中文情感词典的构建

### 2.1 分词

利用 NLPIR 分词系统进行分词,此研究先构建好开源情感词典,表情情感词典和网络情感词典,在分词系统中将构建好的词典加入系统中,从而提高分词效果,例如,未加入词典之前,网络新词“坑爹”得到的分词效果是“坑/v 爹/n”,加入词典后的分词效果是“坑爹/a”。

### 2.2 开源情感词典

目前中文情感词典中没有成熟完整的情感词典可供研究。本文采用情感区分较明显的 HowNet 中文情感词典和台湾大学简体中文情感极性词典。知网的情感词典分为正面情感词,正面评价词,负面情感词,负面评价词,程度副词分为 6 个等级分别为极其、很、较、稍、欠、超。知网情感词典中正面情感词,正面情感评价词均赋权值为 1,而负面情感词和负面情感评价词均赋权值为-1。台湾大学简体中文情感极性词典中分为两类词,积极词汇和消极词汇,我们将积极词汇赋权值为 1,而消极词汇赋权值为-1,表 1 是开源词典情感词情况。

表 1 情感词示例表

情感分类	实例	个数
正面情感词	关心、欢愉、开心	7376
负面情感词	哀痛、抱憾、抱歉	12646

### 2.3 否定副词词典

否定副词是整理的 33 个常用于否定后面词语的否定副词,将否定副词的权值全部设置为-1,否定词如“不”、“没”、“没有”、“无”、“非”等。

## 2.4 表情符号词典

本文在分析文本情感词的基础上融入分析表情符号,随着表情符号的流行人们越来越趋向于用表情符号表达自己的观点,所以本文首先加入了新浪微博中常用的默认的表情符号并赋予相应的权值。提取的微博语料中的表情词以“[]”形式出现,表情词列表见表2。

表2 表情词典示例

表情符号倾向	权值	表情符号示例	个数
正向	1	[偷乐]、[呵呵]	77
负向	-1	[汗]、[可怜]	36

## 2.5 程度副词情感词典

采用知网提供的程度副词,示例表如表3:

表3 程度副词示例表

量级	权重	程度副词实例	个数
极其	5	百分之百	69
超	3	超、过、过于	30
很	2	不过、不胜	42
较	1.75	大不了、更	37
稍	1.5	稍微、略微	29
欠	12	不甚、不怎么	12

## 2.6 网络情感词典

网络新词:即多在网络上流行的非正式语言。多为谐音、错别字改成,也有象形字词。

基于微博文本非正式、口语化的表达特点,网络新词在微博评论中越来越被广泛应用。

由于网络新词的流行使得微博语言更多样化,如“我”不叫我,叫“偶”,“这样子”不叫这样子,叫“酱紫”,“喜欢”不叫喜欢叫“稀饭”,“新手”不叫新手叫“菜鸟”等。

本文选用百度引擎和搜狐引擎里的网络新词,集合了现在广为流行的网络用语,有较好的网络用词覆盖率,网络词示例见表4。

表4 网络情感词示例

网络情感词典	权值	举例	个数
正面	1	白富美、暖男	36
负面	-1	坑爹、逗比	128

## 2.7 微博情感词典构建

由于HowNet情感词典并不完整,在进行微博情感分析的过程中仍然有许多情感词汇无法判别情感倾向,刘培玉<sup>[1]</sup>等人在传统词典的基础上构建领域微博

情感词典,该方法在微博倾向性分析中得到了很好的效果。

### 2.7.1 基准词选取方法

根据微博数据分词结果对各个词语按词频统计并排序,对排好序的词语从上到下进行筛选,筛选出含有明显情感倾向且词频高的词语作为基准词,所以分词的好坏直接影响到排序结果,本文中用的是NLPIR分词算法并且加入了知网情感词典,台湾情感词典,网络新词词典,表情词典中的词语,从而使得分词效果更好,得到更准确的基准词词语,朱嫣岚<sup>[12]</sup>等人证明种子词越多,词的倾向性判断准确性越高,故本文在词频阈值范围内选取了27对种子词,同时考虑到表情符号元素,选取了2对高频表情符号做基准词。

### 2.7.2 点互信息

点互信息主要用于计算语义间的语义相似度,基本思想是统计两个词语在文本中同时出现的概率,概率越大相关性越紧密,关联度越高,两个词语word1和word2的PMI计算公式如下:

$$PMI(word1, word2) = \log_2 \left( \frac{P(word1 \& word2)}{P(word1)P(word2)} \right) \quad (1)$$

$PMI(word1 \& word2)$ 表示两个词语同时出现的概率,即word1和word2共同出现的文档数, $PMI(word1), PMI(word2)$ 分别表示word1, word2单独出现的概率。

褒义和贬义基准词分别用Pw和Nw来表示这两组词语,基准词的选取是具有明显情感倾向性的且极具领域代表性的词语。则候选集中候选词word1的情感倾向计算公式如下:

$$SO\_PMI(word1) = \sum_{i=1}^n (PMI(word1, Pw_i) - PMI(word1, Nw_i)) \quad (2)$$

若 $SO\_PMI(word1) > 0$ : word1具有正面情感倾向,为正向情感词。

$SO\_PMI(word1) < 0$ : word1具有负面情感倾向,为负向情感词。

$SO\_PMI(word1) = 0$ : word1没有情感倾向,为中性情感词。

中文情感词情感倾向值计算方法首先选取基准词,通过计算基准词与候选词间的共现概率来初步确定该情感词的情感倾向值。计算情感词点互信息的计算过程

如下：首先根据词频和情感词强度筛选得到基准词，其次根据程度副词提取新词并按词频阈值筛选，最后将新词和基准词进行 PMI 计算。

## 2.7.3 文档频数

$$w_i = F_{pos}(w_d) - F_{neg}(w_d) \quad (3)$$

式中  $w_i$  表示词  $w_d$  在语料中的频数差，其中  $F_{pos}(w_d)$  表示词  $w_d$  在正向语料中出现的次数， $F_{neg}(w_d)$  表示表示词  $w_d$  在负向语料中出现的次数。当  $w_i$  为 0 时，表示  $w_d$  在正向语料和在负向中出现的次数相等，直接删除该词。

## 2.7.4 微博新词抽取方法

陈建美<sup>[13]</sup>等人分析了不同语法规律对情感词汇自动获取的作用大小，并且证明了情感词汇自动获取方法是有效的，其中否定词与词性搭配规律和程度副词修饰规律最为通用。汉语中程度副词经常修饰形容词及部分心理动词，而这些词多为情感词，如：非常开心，很郁闷等。所以在进行情感词汇抽取时选择程度副词和否定词后 2/3/4 个词范围内的词抽取出来。然后按照这些新词出现的频率进行排序，设定阈值，删除掉频率小于这个阈值的词汇，将筛选出来的新词和基准词做点互信息计算，若计算得出的点互信息结果为 0，则不可以将该词加入情感词典，故删除该词。否则对新词进行文档频数统计，结合点互信息值进行筛选，若新词计算的点互信息值和计算得出的文档频数同为正数或同为负数，则将该新词加入词典，将新词的 PMI 值经处理后(如若 PMI 值在 0 到 10 范围内赋权值为 1，在 -10 到 0 范围内赋值权值为 -1，以此类推)作为对应的权值加入词典，具体步骤如图 1 所示。

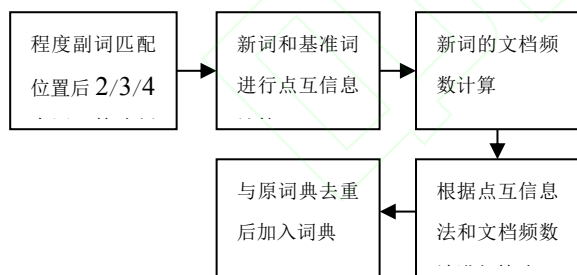


图 1 点互信息计算过程

## 2.8 语气词词典

姜杰<sup>[14]</sup>等人将语气词数量作为语义特征提取，证明了语气词在情感表达中有重要作用。语气词在中文表达中经常流露出感情倾向，如，“哎”，“啊”，“哇”，“耶”等。

本文从百度百科中查找的语气词大全，如，“罢了”“哈”，“呵”“呸”等得到 76 个语气词，并扩展了没有提到的语气词，如“哎”、“哎呀”。共 22 个。语气词在词典中的权值设置和新词加入新词词典的权值计

算方法一样，同样根据点互信息法和文档频数法筛选，最后将语气词和其权值加入语气词词典中。

## 3 语句情感分析

### 3.1 词语情感值的计算

语义是语句进行情感分类的重要特征<sup>[15]</sup>，文档分类判断应按照词汇、句子、微博短文的步骤进行判断<sup>[16]</sup>。情感倾向情感词前经常有程度副词修饰。当情感词前有程度副词修饰时，则会使情感词的情感倾向加强或减弱。如：开心是正向情感词，其权值为 1，则很开心使得情感词情感倾向加强。而情感词前有否定词修饰时会使情感词的情感倾向反转，如“伤心”表示负向情感词，而“不伤心”表示正向情感倾向。但是存在一个问题，形如：“否定词+程度副词+情感词”此形式对情感词的强度有减弱作用。“程度副词+否定词+情感词”对情感词的情感强度具有加强作用。如很不开心和不是很开心的情感强度截然不同。故需要区分两种形式的区别。情感词语的情感值计算公式如下：

否定词+程度副词+情感词：

$$w = d \times a \times t \times 0.5 \quad (4)$$

程度副词+否定词+情感词：

$$w = d \times a \times t \times 2 \quad (5)$$

其中  $w$  表示计算得到的情感词语的情感强度值， $t$  表示情感词的权值， $a$  表示该情感词  $t$  前的程度副词的权值， $d$  表示否定词的权值。

### 3.2 句间关系原则

赵天奇<sup>[17]</sup>等人在语义规则描述部分，基本涵盖了汉语中最常用的几种句型规则和句间关系规则，从而使得对复杂句的情感分析更加准确。由于中文表达的传统习惯，一个句子可能包含多个关联词，而这些关联词对句子的情感表达有很大影响。如：虽然  $s_1$  但是  $s_2$ ，其中， $s_2 = w_3 w_6 w_7$ ，其中  $w_i$  是一些情感词语。显然句子  $s_1$  表达的情感强度没有  $s_2$  的情感强度大。总结句间规则如下。

#### 3.2.1 转折关系原则

当句子中出现转折时，如：虽然。。。但是。。。由于出现转折，则强调转折后的句子的情感倾向而减弱转折前的句子的情感倾向。邸鹏<sup>[18]</sup>等人提出依据已有资源，并根据转折句式否定词、转折词、情感词的组合规律进行情感分析的启发式规则，证明该方法能更好的对转折句式进行情感倾向性分析。转折语句情感分析规则如下，因为转折关系连接词并非都是成对出现，故我们分为以下三种情况。



我们将转折关系连接词分为转折前接词和转折后接词，其中转折前接词如：虽然、虽说、尽管等。转折后接词如：但是、可是、不过等。根据中文表达习惯，转折前接词会削弱用户要表达的意思，而转折后接词会加强用户要表达的意思。

(1)当分句只含有转折前接词的时候，设转折前接词出现在分句 $s_k$ ，则各句句间关系系数为：

$$s_k, s_{k+1}, \dots = 0.5$$

(2)当分句只含有转折后接词的时候，设转折后接词出现在 $s_k$ ，则句间关系系数 $s_k, s_{k+1}, \dots = 1.5$

(3)当分句既含有转折前接词又含有转折后接词时，句间系数为默认值 1 不变。

### 3.2.2 假设关系原则

当句子出现假设关系连接词时，句子强调的是条件，例如：如果 $s_1$ ，那么 $s_2$ ，则句子强调的是“如果”后面的内容即 $s_j$ 。设假设关系后接词如上例中的“那么”出现在分句 $s_i$ ，则令 $s_1 s_2 \dots s_{i-1} = 1, s_i s_{i+1} \dots s_n = 0.5$ 。

### 3.2.3 因果关系原则

当句子中出现因果关系连接词时，表示强调的是因果关系前接词后的句子。例如：因为 $s_j$ ，所以 $s_2$ ，强调的是子句 $s_j$ 的情感倾向。设因果关系前接词出现在分句 $s_i$ 上，因果关系后接词出现在分句 $s_j$ 上，则 $s_i, s_{i+1}, \dots s_{j-1} = 1.5, s_j, s_{j+1} \dots s_n = 1$ 。

### 3.2.4 分句情感倾向计算公式

$$E(s_j) = \sum E(w_i) * r_j \quad (6)$$

其中 $E(s_j)$ 表示第 $j$ 个句子的情感倾向值，

$E(w_i)$ 表示第 $j$ 个句子中的情感词的权值。 $r_j$ 表示第 $j$ 个句子的句间关系系数。

### 3.3 句型关系

刘楠<sup>[9]</sup>将标点符号作为特征进行提取。当一个句子以“!”结尾时，该句的情感倾向明显加强。例如：“海燕儿，你可长点心吧!”表示的是加强句子的感叹作用，当句子以“!”结尾时，该句的情感倾向值变成原来的二倍。

当句子以“?”结尾并且该句子存在反向疑问词时，该句子表示的是质疑。例如：“难道你就不能开开心心独立自主吗?”，该句子中出现的情感词“开开心心、独立自主”均为正向情感词，该句子以“?”号结尾，并且出现反向疑问词“难道”，故该句的情感倾向将反向加倍。复句情感倾向值计算公式如下：

$$E(T_i) = \sum E(s_j) * M_i \quad (7)$$

式中 $E(T_i)$ 表示第 $i$ 个复句的情感倾向值， $E(s_j)$ 表示第 $i$ 个复句中各个子句的情感倾向值， $M_i$ 表示句型关系系数，若该复句 $T_i$ 以“!”结尾则 $M_i=2$ ，若 $T_i$ 以“?”结尾，且 $T_i$ 中含有反向疑问词，则 $M_i=-2$ 。

### 3.4 首句和尾句

中文表达中有开门见山的表达习惯，故首句在短文本情感判断中具有很重要的作用。如：中文表达中尾句的情感倾向在微博情感表达中占得权重很大，故设置首句的的加权关系系数为 2，微博语句的尾句具有很重要的情感倾向判断价值，如“今年是交大 120 周年校庆，我很荣幸作为母校的学子。今天好开心呀，哈哈。但是没有和曾经的老同学们团聚真的太可惜了”。这段微博包含 3 句话，其中第 1, 2 句表达了正向情感，而最后一句表达了负向情感。这段微博整体表达的是负向情感，以尾句的情感倾向为准，所以在微博表达中尾句的情感倾向具有非常重要的作用故设尾句的加权关系系数为 2，如计算首尾句的情感倾向值为 $E(T_i)$ ，则 $E(T_i)=E(T_i)*2$ 。

## 4 实验与分析

实验人工选取 NLPCC 官方网站中 Task A Annotated Training Data 中的数据，主题包括 IphoneSE、春节放鞭炮、俄罗斯在叙利亚的反恐行动、开放二胎、深圳禁摩限电，同时本实验还加入了 coae2014 评测数据的 task3&4 任务中对蒙牛的评论。正面数据 2019 条，负面数据 2139 条。判断标准如下：当微博情感值 $>1$ 时，判断该微博为正面，当微博情感值 $<-1$ ，判断该条微博为负面，当微博情感值为 0 时，判断该微博为中性。实验结果如图下所示：

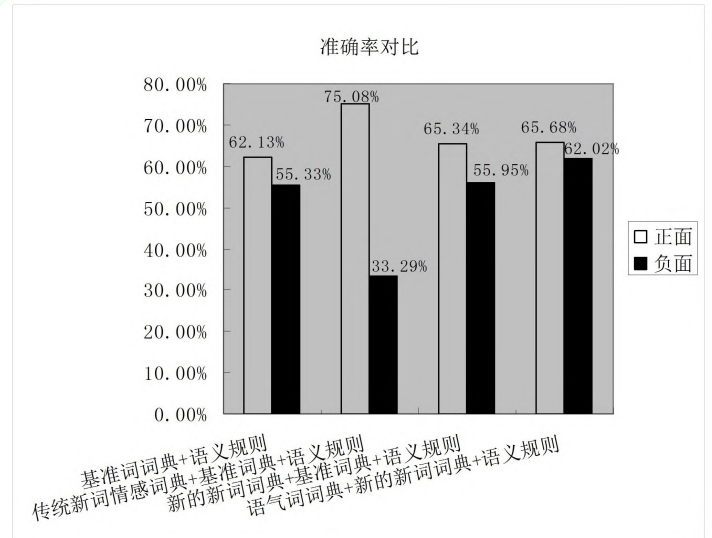


图 2 准确率对比

各方法准确率对比如图 2，对比基准词典+语义规则的结果和传统新词情感词典+基准词典+语义规则的结果可以发现，加入传统的新词词典判断的结果有个严重的弊端，负面微博的分类准确率骤降，而正面微博的准确率骤增，这是因为微博数据正面情感词比

例远远大于负面情感词的比例。这对于分类判断来说是无效的。通过新的新词词典+基准词典+语义规则的结果可以观察到,加入新的新词词典后,正负面的微博分类的准确率都有不同程度的提高,这是因为当我们加入了文档频数筛选方法后可以筛掉那些在负面微博中出现的大量正面情感词。

对比新的新词词典+基准词典+语义规则的结果和语气词词典+新的新词词典+语义规则的结果可以发现,加入语气词词典后,正面微博的分类准确率没有明显变化,而负面微博的分类准确率有明显的提高。

## 5 结论

本文在分析已有情感词典的基础上,建立了更全面完整的微博情感词典,在现有网络资源的基础上对微博语料进行统计建立网络情感词典和表情情感词典,通过词典的构建形成了更加完备的情感词典,用于情感分析。建好情感词典后根据中文的语义规则包括句间规则和句型规则,进行句子级的微博情感分析。本文的特点是加入了首句和尾句情感分析倾向的权重系数,从而提高了微博情感分析的正确率。本文的创新之处首先在传统的微博情感词典的构造上做出了改进,对新词通过点互信息和文档频数法进行筛选再加入词典。其次创建了语气词词典,有了语气词词典后负面微博分类的准确性有了明显的提高。

## 6 结束语

微博情感分析的关键是词典构建,词典的囊括范围及准确性对分类的准确率具有关键作用,本研究构建了开源情感词典,表情情感词典,网络情感词典。并且改进了微博情感词典的构造方法,创建了语气词词典。利用改进后的词典,微博分类的准确率有了明显的提高,通过实验证明了此方法的准确性。但是微博分类还有有待研究的方向。

1. 微博文本中经常出现反讽现象,具有反讽现象的微博文本的情感倾向和用算法算出来的情感分析值相反。
2. 微博用户的情感表现不仅包括文本,还有微博的点赞信息,转载量等,这些都是微博分析需要进一步研究的内容。
3. 由于微博用户发表的微博多接近于口语化,故微博文本经常出现拼写错误或符号省略等现象,如何识别这些错误并在情感分析中考虑这些因素带来的影响需要进一步的研究。
4. 本文采用的是基于词典和语义规则的方法进行情感分类,将其作为特征,结合深度学习方式并与社会背景结合<sup>[20]</sup>是将来需要研究的方向。

## 7 参考文献

- [1] Turney P D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews[C]// Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2002:417-424.
- [2] Cui H, Mittal V, Datar M. Comparative experiments on sentiment classification for online product reviews[C]// National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, Usa. DBLP, 2006:61-80.
- [3] 王伟,周咏梅,阳爱民,等.一种基于LDA主题模型的评论文本情感分类方法[J].数据采集与处理,2017,32(3):629-635.
- [4] 黄发良,冯时,王大玲,等.基于多特征融合的微博主题情感挖掘[J].计算机学报,2017,40(4):872-888.
- [5] Taboada M, Brooke J, Tofiloski M, et al. Lexicon-based methods for sentiment analysis[J]. Computational Linguistics, 2011, 37(2):267-307.
- [6] 谢丽星,周明,孙茂松.基于层次结构的多策略中文微博情感分析和特征抽取[J].文信息学报,2012,26(1):73-83.
- [7] 周剑峰,阳爱民,周咏梅.基于中文微博的情感词典构建及分类方法[J].计算机与数字工程,2014,26(10):1773-1776.
- [8] 肖江,丁星,何荣杰.基于领域情感词典的中文微博情感分析[J].电子设计工程,2015,23(12):18-21.
- [9] 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.
- [10] 王振宇,吴泽衡,胡方涛.基于HowNet和PMI的词语情感极性计算[J].计算机工程,2012,38(15):187-189.
- [11] 刘培玉,张艳辉,朱振方,等.融合表情符号的微博文本倾向性分析[J].山东大学学报(理学版),2014,49(11):8-13.
- [12] 朱嫣岚,闵锦,周雅倩,等.基于HowNet的词汇语义倾向计算[J].中文信息学报,2006,20(1):14-20.
- [13] 陈建美,林鸿飞,杨志豪.基于语法的情感词汇自动获取[J].智能系统学报,2009,4(2):100-106.
- [14] 姜杰,夏睿.机器学习与语义规则融合的微博情感分类方法[J].北京大学学报(自然科学版),2017,53(2):247-254.
- [15] Whitelaw C, Garg N, Argamon S. Using appraisal groups for sentiment analysis[C]//Proceedings of the 14th

ACM international conference on Information and knowledge management. ACM, 2005: 625-631.

[16] Bing Liu. Sentiment Analysis and Opinion Mining[J]. Synthesis Lectures on Human Language Technologies, 2016, 30(1):152-153.

[17] 赵天奇, 姚海鹏, 方超, 等. 语义规则与表情加权融合的微博情感分析方法[J]重庆邮电大学学报(自然科学版), 2016, 28(4):503-510.

[18] 邸鹏, 李爱萍, 段利国. 基于转折句式的文本情感倾向性分析[J]. 计算机工程与设计, 2014(12):4289-4295.

[19] 刘楠. 面向微博短文本的情感分析研究[D]. 武汉大学, 2013.

[20] Zou X, Yang J, Zhang J, et al. Microblog Sentiment Analysis with Weak Dependency Connections[J]. Knowledge-Based Systems, 2017, 142:170-180.

