

文章编号: 1003-0077(2010)03-0037-07

面向话题的新闻评论的情感特征选取

陶富民, 高 军, 王腾蛟, 周 凯

(北京大学 信息科学技术学院计算机系网络与信息技术研究院, 北京 100871)

摘 要: 情感特征的提取是进行文本情感分析的一个非常重要的步骤,也是影响其结果好坏的主要因素。在该文中,作者提出一种新的特征提取方法来解决新闻评论的情感分析问题。在该方法中,首先根据评论和新闻的对比分析获得候选情感特征,然后经过相关的扩充和验证操作得到通用的情感特征,并将其用于新闻评论的情感分析。对新闻进行话题划分后进行更细粒度的情感分析:根据新闻话题信息,设计相应的话题相关的特征对比和验证过程,选取出面向话题的情感特征,最后用面向话题的情感特征对相应话题进行情感分析。实验证明,这种情感特征提取方法,对于新闻评论这种语句短、评论对象相对分散的评论,情感分析效果有较大的改进。

关键词: 计算机应用; 中文信息处理; 情感分析; 特征选取; 特征扩展

中图分类号: TP391 **文献标识码:** A

Topic Oriented Sentimental Feature Selection Method for News Comments

TAO Fumin, GAO Jun, WANG Tengjiao, ZHOU Kai

(Institute of Network and Information System, Department of Computer Science,
School of EECS, Peking University, Beijing 100871, China)

Abstract: Feature extraction is essential to the quality of text based sentiment analysis. This paper proposes a novel approach to feature extracttion for the sentiment analysis of the news comments. Firstly, the candidate sentimental features are extracted according to the comparison between contents of the news comments and the corresponding news. Then, the general sentimental features for sentiment analysis on various news comments are selected by several extension and validation processes.

The proposed method is featured by capable of providing finer-grained sentimental analysis for specific news topic. Specifically, based on the topic information of news comments, it can be adapted for corresponding feature comparison and validation policies to extract topical sentiment features. The experiments show a high performance for the sentiment analysis in sparse data sets, such as comments from news.

Key words: computer application; Chinese information processing; sentiment analysis; feature selection; feature extension

1 引言

随着网络的普及特别是 Web2.0 的兴起和发展,网络上用户生成内容(UGC)越来越多,比如博客、评论、论坛帖子等。这些用户生成内容在网络中占据越来越重要的地位。它们包含大量的主观性内

容,这些主观性内容含有很多潜在的有用信息,比如针对商品的评价会直接影响到用户的购买行为;政府机构会关注发布的政策法规在网络中的反响;关注网上的舆情信息,特别是热点事件的舆情信息。

传统的网络信息处理主要针对基于事实性的文本,比如基于关键字的检索,文本的分类,聚类等,这些处理忽略了其中的情感信息。情感分析主要针

收稿日期: 2009-09-18 定稿日期: 2010-03-01

基金项目: 国家自然科学基金资助项目(60873062)

作者简介: 陶富民(1984—),男,硕士生,主要研究方向为数据库与信息系统;高军(1975—),男,副教授,主要研究方向为数据库与信息系统;王腾蛟(1973—),男,副教授,主要研究方向为数据库与信息系统。

对用户生成内容来进行情感信息的挖掘,其最重要的方面是情感倾向性分析。情感的特征选取不仅是影响情感倾向性分析结果好坏的重要因素,也是其主要难点。主要体现在 1) 仅仅通过词频或其他简单统计量很难或提取不出有效的情感特征;2) 一些特征词在不同的领域具有不同的情感倾向,比如“布什是中国人民的老朋友”和“老布什这个老不死的”这两句话中,“老”这个词在其出现的三个地方中有三种不同的情感特征。由于情感特征提取的特殊性,传统文本分析方法不能满足情感分析的要求。本文的目标就是找出一个有效的情感特征提取方法。

在本文中,我们主要关注新闻领域中与政治相关评论的情感倾向。我们抓取搜狐一年的国内、国际、社会相关的 67 190 条新闻和对应的 6 156 840 条评论并进行了简要的分析:1) 对新闻进行聚类,共聚集出 15 190 个类别。从图 1 中可以看出聚类后的类别的新闻数目大致符合齐普夫定律(zipf’s law^[1]),表明热点新闻是用户的主要关注点。2) 新闻的评论数目分布也大致符合齐普夫定律(见图 2),进一步说明用户的关注度集中在热点新闻中。从上面两个结论说明改善热点新闻评论的情感分析效果具有重要的意义,可以通过提高它们的情感分析效果来提高整体情感分析效果。

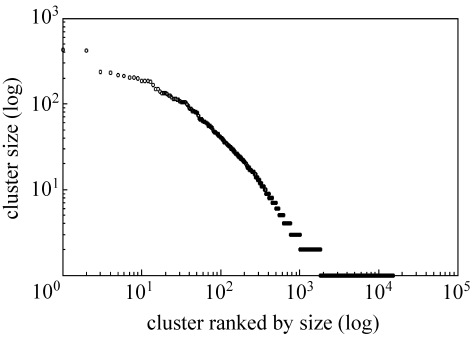


图 1 聚类的新闻数目

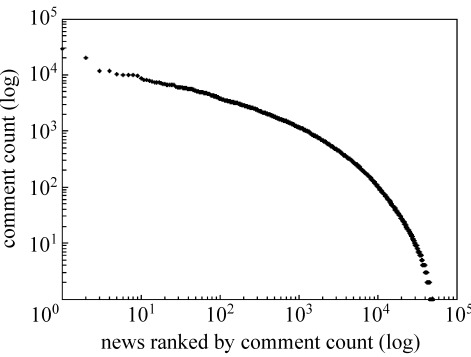


图 2 新闻的评论数目

目前的情感特征提取的方法主要集中在电影、书籍、商品等评论上,针对新闻评论的情感特征提取则很少,而且相对更加困难。其困难之处在于新闻领域中的评论很难像文献[2-3]中描述的那样,可以比较容易地获取高质量评论或与所评论的内容相关性较高的评论来进行情感分析。此外新闻评论的情感特征提取的主要难点还在于:评论的内容普遍较短(见图 3,图 4);评论的主题十分发散,评论容易偏离其所针对的主题;普通用户评论用词口语化、错字、别字、简写、俚语较多;另外还有很多评论有很强的背景(比如“农夫与蛇”这样与典故相关的评论)。新闻评论中的这些特性使得我们需要在评论本身与评论所针对的对象之间进行更多的分析处理。

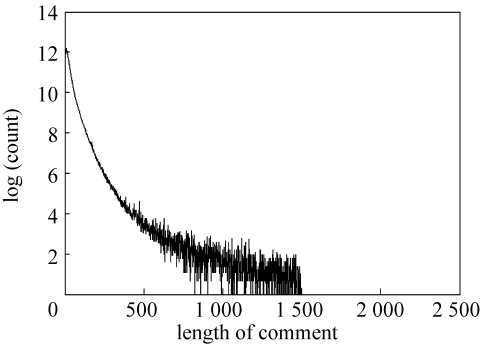


图 3 评论长度与数量的关系

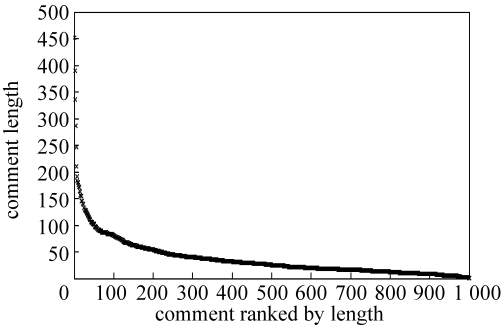


图 4 随机选取 1 000 条评论长度

2 相关术语定义和工作流程

- 1. 对比新闻内容和评论内容得到的通用候选情感特征词表,简称通用候选特征表 GC;
- 2. 对通用候选特征进行人工筛选和倾向性标注得到新闻评论的情感种子特征词表,称之为基础情感特征表 BF;
- 3. 在基础特征词表的基础上进行扩充和验证得到扩充的情感特征词表,称之为通用情感特征表 GF;

4. 根据新闻内容对新闻进行聚类,选取数目大于预定义数值的新闻类簇,这些新闻类簇称为话题 $\{T\}$;

5. 对比话题的评论和话题的新闻得到与话题相关的候选情感特征词表,简称话题候选特征表 $\{TC\}$;

6. 在基础情感特征词表 BF 的基础上,对话题候选特征词表 $\{TC\}$ 进行验证得到与话题相关的情感词表,称之为话题情感特征表 $\{TF\}$ 。

在得到上述的特征表和话题信息之后,我们用下列方式对新闻的评论进行情感分析:

1. 确定新闻 N 所在的话题 t :首先查看新闻 N 是否属于某个话题 t 中,如果不属于任何一个话题,则根据新闻 N 与各个话题的相似关系判定新闻 N 是否能够划分到某个话题 t 中,如果没有被划入任何一个话题中,则令 t 为空话题 nil 。

2. 根据新闻 N 所在的话题 t 确定对应的情感特征表 f :如果 t 为 nil ,则令情感特征表 f 为通用情感词表 GF ,否则根据 t 从 $\{TF\}$ 中寻找对应的情感特征表 f 。

3. 对新闻 N 的每一个评论 c ,通过其情感特征表 f 确定其情感倾向 p 。

3 基于话题的情感特征提取

3.1 获取评论的通用候选特征和基础情感特征

一般而言,评论除了包含情感特征词之外,还包含许多与评论内容相关的词和高频词。直接采用词频等统计信息很难提取出有效的情感特征;同时 Pang 等人在文献[3]中发现用统计的方法选取出的评论特征词比人工直接选取的效果要好。这表明对评论进行情感分析时,评论的情感特征应该从所分析的评论中取出。我们采取的方法是分步骤逐渐优化提取情感特征。

首先我们采用公式:

$$\text{Score1}(\text{Term}) = (\text{R_DF}(\text{Term}) / (\text{N_DF}(\text{Term}) + 1)) \quad (1)$$

$$\text{Score2}(\text{Term}) = (\text{R_TF}(\text{Term}) / (\text{N_TF}(\text{Term}) + 1)) \quad (2)$$

对评论中出现的词打分后进行排序,选取得分较高的词作为通用候选特征表 GC 。其中 $\text{R_DF}(\text{Term})$ 、 $\text{N_DF}(\text{Term})$ 、 $\text{R_TF}(\text{Term})$ 和 $\text{N_TF}(\text{Term})$ 分别表示 Term 出现在所有不同评论中的次数、出现在所有不同新闻中的次数、在所有评论中

出现的总次数和在所有新闻中出现的总次数。排序规则是先按 Score1 的得分排序,如果 Score1 的得分一样,再按 Score2 的得分进行排序。

利用上述公式可以使高频词和与新闻背景相关的词(比如在讲述中美关系的新闻中,“中国”,“美国”这样的词)得分降低。这样一来,这些词就不容易出现在通用候选特征表 GC 中。

在通用候选特征表 GC 中存在着大量的噪音,比如不规范的用词,错别字;评论中的广告或其他垃圾信息,它们在评论中出现频率比较高,而在新闻中却很少出现,所以这些噪音的得分也会比较高,之外我们还需要明确那些与情感相关的候选特征词的倾向性。因此需要进行必要的人工筛选和倾向性标注,通过筛选和标注后得到基础情感特征表 BF 。

3.2 构建词的同位关系对基础情感特征词表进行扩充和验证

虽然基础情感特征词表 BF 里面的情感词都是经过人工干预的,但考虑人工筛选和标注的代价以及在获取候选特征词表中我们不能保证所有与情感相关的特征都能得到较高的分成为候选特征,需要对基础情感特征词表 BF 进行自动扩充操作以增加召回率。

3.2.1 同位关系的构建

定义 1:同位关系 如果两个词在意思表达可以替换或表示同一类事物时,则称这两个词存在同位关系。例如“丰田”和“宝马”是同位关系。但“丰田”和“汽车”就不能看成同位关系。同理,“高兴”和“悲伤”是同位关系,但“高兴”和“情绪”就不是同位关系。

我们在大规模语料中利用设定的模板模拟词与词之间的同位关系,如“($@x$ [、 $@y$] * [和|与|以及] $@z$)”。 $@x$ 和 $@z$ 分别表示一个词, $\{ @y \}$ 表示包含 0 到多个词的集合。该模板定义词 $@x$, 词 $@z$ 和词集 $\{ @y \}$ 中词之间存在同位关系。我们利用自设的一些模板提取了近 1 240 万个模拟词同位关系对。用匹配次数 (MatchCnt) 以及同位关系的置信度 $\text{confidence}(\text{term1}, \text{term2}) = p(\text{term1}, \text{term2}) / p(\text{term1})$ 来衡量模拟词的同位关系对的维度。

3.2.2 扩充和验证

根据词的同位关系,我们对基础情感特征表 BF 中的所有特征词按公式 $\log(\text{MatchCnt}) \times \text{confidence}$ 的结果大小顺序找出 k 个同位词(不足则全取)来进行扩充操作。把正向情感特征和相关扩充特征称为正向扩充词典,把负向情感特征和相关扩

充特征称为负向扩充词典。由于我们所得到的同位关系只是模拟的同位关系,另外还存在处于同位关系的两个特征词具有不同的情感倾向性,所以需要对这些扩充词进行验证来过滤引入的错误的情感特征以及对扩充词进行情感标注。验证方法如下:

1) 首先对每一个在扩充阶段得到的词,找出其同位词集,根据它们在正负扩充词典的个数和比率进行过滤。

个数过滤方法:如果正向扩充词在基础情感特征词表 BF 中的同位词标注为正向倾向的特征词小于设定的最小个数 min_cnt (设定为 3) 则将该词过滤,负向扩充词同理。

比率过滤方法:通过随机选取基础情感词表 BF 中的一些词,并对它们的扩充词进行人工标注后,把过滤这些扩充词的最优的参数值设定为阈值 a 。如果正负向同位词的比率小于设定的阈值 a ,则将该词过滤。

2) 经过上一步过滤后,还存在一些正负扩充词典中共同包含的词。采用比较其同位词在正负向扩充词典的个数的方法来确定其情感倾向,即如果词 T 同时出现在正负向扩充词典中,并且出现在正向扩充词典中的个数多,那么就认为词 T 的情感倾向为正向,将其从负向扩充词典中删除。如果数目一样,则同时删除,否则将其从正向扩充词典删除。

按上述方式进行多次迭代过滤后就可以得到两个过滤后的情感词典。把这两个词典合并成一个带情感倾向标注的词典就是我们定义的通用情感特征表 GF 。

3.3 话题构建和话题情感特征的构建

根据我们对搜狐的新闻对应的评论数目,新闻聚类中对应的新闻数目的分析,可以得出用户评论主要集中在热点新闻中的结论。因此可以把改善情感分析效果的重心放在热点新闻类别中。我们通过聚类将新闻划分成多个话题,对每一个话题提取出与该话题相关的情感特征。通过这些话题相关的情感特征来提高话题中的新闻的情感分析效果,以达到提高整体的情感分析效果的目的。

3.3.1 构建话题

我们利用 CMU 与 UMASS 的 Lemur 系统提供的索引和聚类工具对新闻进行聚类。采用中心点 (centroid) 的聚类方法,用向量之间的余弦值 (cosine similarity) 计算新闻相似度,通过设置较高的相似度阈值来提高同一类簇中的新闻的内聚性。把新

闻数目大于等于 10 的类簇定义为话题。所有的话题的集合定义为话题 $\{T\}$ 。

3.3.2 获取话题的候选特征

与获取评论的通用候选特征类似,采用公式:

$$Score1(t, term) = (R_DF(t, term) / (N_DF(t, term) + 1)) \quad (3)$$

$$Score2(t, term) = (R_TF(t, term) / (N_TF(t, term) + 1)) \quad (4)$$

对评论中出现的词进行打分排序。 $R_DF(t, term)$, $N_DF(t, term)$, $R_TF(t, term)$ 和 $N_TF(t, term)$ 分别表示 $term$ 在话题 t 中出现在不同评论中的次数、出现在不同新闻中的次数、出现在评论中的总次数和在新闻中的出现总次数。排序规则也是先按 $Score1(t, term)$ 的得分排序,在该分值一致的情况下再按 $Score2(t, term)$ 的得分进行排序。然后取得分较高的前 k 个特征词作为话题候选特征。所有话题的候选特征的集合为话题候选特征表 $\{TC\}$ 。

3.3.3 对话题的候选特征进行验证

我们将基础情感特征表 BF 和话题对应的话题候选特征表 tc 一起作为话题的扩充词典,但只用模拟的同位关系去验证和过滤话题的候选特征词,不做进一步的迭代以达到在线处理的要求。

处理方式:设置最小个数 min_cnt 和差别率 a 两个阈值进行过滤,令 P 为候选特征词 $term$ 的同位词中正向情感特征数, N 为负向情感特征数,如果同时满足 $P > N$, $P \geq min_cnt$, $P / (N + 1) \geq 1 + a$ 三个条件,就保留 $term$ 作为话题的正向情感特征。同理,如果同时满足 $N > P$, $N \geq min_cnt$, $N / (P + 1) \geq 1 + a$ 三个条件,就把 $term$ 保留下来作为话题的负向情感特征。这样就可以得到话题的情感特征表,所有话题的情感特征表构成的集合为话题情感特征表 $\{TF\}$ 。

4 实验结果和分析

4.1 实验说明

我们从 NTCIR-6 的中文语料中抽取出带有标注信息的情感特征词,仅仅使用这些情感特征词进行情感倾向性分析的召回率很低。为此我们设计了一个提供标注评论的情感特征词,情感倾向和标注评论的情感倾向等功能的反馈系统。在本实验中,我们把从 NTCIR-6 中抽取的情感特征词和用户反

馈的情感特征词组成的情感词典作为我们的基准情感特征表 NF。

我们一共标注了 419 篇新闻的 5 414 条评论，其中 3 528 条负向评论，1 886 条正向评论。从新闻的聚类结果中选取了 3 个新闻和评论数最高的话题 (A, B, C) 进行了标注，从话题 A 中选取了 146 条新闻，对其中的 1 052 条评论进行了标注，其中正负向评论分别为 380 条，672 条；从话题 B 中选取了 119 篇新闻，对其中的 671 条评论进行了标注，其中正负向评论分别为 179 条，492 条；从话题 C 选取了 75 篇新闻，对其中的 1 106 条评论进行了标注，其中正负向的评论分别为 471 条，635 条。

表 1 通用候选特征表 GC 和基准情感特征表 NF 在机器学习方法上的对比结果

	分类方法	考虑特征数目		考虑特征出现与否	
		召回数	准确度	召回数	准确度
基准情感特征表 NF	Naïve Bayes	2 434	81.635%	2 434	81.841%
基准情感特征表 NF	Naïve Bayes Multinomial	2 434	82.827%	2 434	83.032%
通用候选特征表 GC	Naïve Bayes	2 834	77.735%	2 834	77.770%
通用候选特征表 GC	Naïve Bayes Multinomial	2 834	81.757%	2 834	81.863%

以发现，尽管与评论相关的通用候选特征表 GC 的特征数目 (1 000 个) 比采用基准情感特征表 NF 的特征数目 (1 914 个) 要少近一半，但其召回数目要高出 16.4%，而准确度 (Accuracy) 却相差不大，特别是在较高的准确的 Naïve Bayes Multinomial 方法下只相差了不到 1% 的准确度。

4.3 基准情感特征表 NF 和通用情感特征表 GF 的实验对比

我们从通用候选特征表 GC 中标注了 1 195 个情感特征词，利用这些特征词进行扩充和验证得到通用情感特征表 GF。采用的方法是利用词的情感倾向特征结合句子的否定和转折处理对评论进行情感倾向性判定。从表 2 中可以看出，在构造通用情

表 2 通用情感特征表 GF 和基准情感特征表 NF 的情感分类的对比结果

	考虑特征数目		考虑特征出现与否	
	召回数	准确度	召回数	准确度
基准情感特征表 NF	3 429	80.723%	3 390	81.180%
通用情感特征表 GF (20)	3 426	79.072%	3 401	79.359%
通用情感特征表 GF (100)	3 820	78.325%	3 788	78.696%

4.2 基准情感特征表 NF 和通用候选特征表 GC 的实验对比

在 3.1 节中选取的通用候选特征表 GC 时，选取候选特征词的前 1 000 个特征词和基准情感特征表 NF 进行对比实验：采用 Weka 提供的 Naïve Bayes 和 Naïve Bayes Multinomial 方法，用 10 折交叉验证方式的对结果进行对比 (见表 1)。从实验结果中可以看出考虑特征词是否出现和考虑其出现的次数之间差别很小。采用 Naïve Bayes Multinomial 方法的效果要好于 Naïve Bayes 的方法。另外还可

感特征表 GF 时对每个情感特征词选取 20 个以内的同位词 (不足 20 个选取所有) 进行扩充和验证时，就能达到和基准情感特征表 NF 相近的召回数和准确度。当对每个情感特征词选取 100 个以内的同位词进行扩充和验证时，以降低约 2% 的准确度为代价提高了 11.4%~11.7% 的召回数，相当于在总体上提高了 9% 以上的准确率。与采用机器学习的方法类似，考虑特征的出现与否的准确率稍高于考虑特征的数目 (约 1%)。

对比表 1 和表 2 的结果发现，采用词的情感倾向特征结合句子的否定和转折处理的情感分析方法，虽然在准确度上有所下降 (1%~2%)。但带来了很大的召回数目的提升 (34.8%~40.9%)。

4.4 面向话题的情感分析实验对比

首先我们利用机器学习的方法比较使用通用候选特征表 GC 与采用话题候选特征表 {TC} 的效果差异。从表 3 中可以看出，尽管采用话题候选特征表 {TC} 在准确度上有所下降 (在效果较好的 Naïve Bayes Multinomial 方法分别为 0.7%，4.7% 和 2.6%)，但召回率却分别高出 91%，105% 和 98%。该结果说明利用话题候选特征表 {TC} 能极大地改善情感分析的效果。

表 3 通用候选特征表 GC 和话题候选特征表{TC}的机器学习方法上的对比结果

	类别	分类方法	考虑特征数目		考虑特征出现与否	
			召回数	准确度	召回数	准确度
通用候选特征表 GC	A	Naïve Bayes	515	78.641%	515	79.029%
话题 A 候选特征表	A	Naïve Bayes	983	74.771%	983	77.009%
通用候选特征表 GC	A	Naïve Bayes Multinomial	515	83.107%	515	83.495%
话题 A 候选特征表	A	Naïve Bayes Multinomial	983	81.689%	983	82.808%
通用候选特征表 GC	B	Naïve Bayes	311	84.012%	311	85.465%
话题 B 候选特征表	B	Naïve Bayes	638	74.295%	638	79.937%
通用候选特征表 GC	B	Naïve Bayes Multinomial	311	85.465%	311	85.756%
话题 B 候选特征表	B	Naïve Bayes Multinomial	638	81.035%	638	81.035%
通用候选特征表 GC	C	Naïve Bayes	532	79.323%	532	79.699%
话题 C 候选特征表	C	Naïve Bayes	1 053	76.828%	1 053	77.398%
通用候选特征表 GC	C	Naïve Bayes Multinomial	532	80.263%	532	80.075%
话题 C 候选特征表	C	Naïve Bayes Multinomial	1 053	76.828%	1 053	77.398%

下面的实验是根据话题情感特征词表{TF},利用词的情感倾向特征结合句子的否定和转折处理进行情感倾向性判定。从实验结果中(见表 4,5)可以看出使用话题情感特征词表{TF}与通用情感特征表 GF(对每个特征词选取 100 个以内的同位词扩展)相比,在召回率相近(70%左右)的情况下,其准

确度高出 2%~5%。与使用基准情感特征表 NF 相比,召回率高 5%~10%。

5 相关工作

文本情感分析最常用的方法是情感分类。虽然在普通的文本分类里有许多经典的特征提取方法^[4-5],如 DF,卡方检验,互信息,信息增益等。但 these 方法不能直接应用在情感分类中,比如“中美关系”是政治新闻中的一个重要特征,但不能作为情感分类的特征。

在早期的研究中^[6],Pang 等人用人工挑选情感特征词和基于词频统计加人工检验的方法比较时发现,基于统计的方法在准确度上有近 5%到 10%的提高。在文献[7]中,Turney 等利用两个词之间的逐点互信息量 PMI(Pointwise Mutual Information)定义词的情感倾向 SO(Semantic Orientation),在汽车、银行等领域的情感分类能达到 80%以上的准确率,但使用搜索引擎计算 SO 值的开销比较大,不适用于大规模的快速分析。

还有一些研究^[8]从一个核心的情感种子词典出发,利用 WordNet 定义的关系(如同义,近义)进行倾向性判定来扩充情感特征词。Popescu^[9]在其 OPINE 系统中,使用“Relaxation Labeling”的方法,通过 Support Function 和迭代操作,同时标识出产品的特征和用户的情感观点。与文献[7]相比,以降

表 4 基于话题情感特征词表(TF)的情感分析结果

	考虑特征数目		考虑特征出现与否	
	召回数	准确度	召回数	准确度
话题 A 情感特征表	730	84.932%	723	85.201%
话题 B 情感特征表	483	80.331%	475	80.842%
话题 C 情感特征表	781	81.178%	769	81.795%

表 5 召回率对比

	类别	召回率 (考虑特征数目)	召回率 (考虑特征出现与否)
话题 A 情感特征表	A	69.40%	68.73%
基准情感特征表 NF	A	64.92%	64.92%
话题 B 情感特征表	B	71.98%	70.79%
基准情感特征表 NF	B	59.46%	58.42%
话题 C 情感特征表	C	70.61%	69.53%
基准情感特征表 NF	C	63.56%	62.29%
通用情感特征表 GF (100)	无	70.56%	69.97%

低 3% 的召回率的代价提高了 22% 的精度。Kobayashi^[10]通过层次的方法获取情感特征改善情感分析效果,但需要在每一个层次进行人工干预。

6 结论

本文的主要贡献在下列几个方面:

1) 通过构建词的同位关系,对人工标注的情感特征进行扩充,并对噪声比较大的候选情感特征和扩充情感特征进行了验证。

2) 构建了一个面向话题的新闻评论的情感特征提取框架,通过对那些热门话题构造对应的情感特征表来达到改善情感分析的效果。由于用户的关注主要集中在热门话题上,故该工作有重要的意义。

同时在我们的工作中还发现目前的情感分析方法对新闻评论中反语效果很差,如果能够找到一种针对反语的处理方法,就可以有效的改善情感分析的效果。另外找出评论对象之间的关系以及和评论者之间的关系也是一个值得深入研究的课题。

参考文献

[1] Wentian Li. Random Texts Exhibit Zipf’s-Law-Like Word Frequency Distribution[J]. IEEE Transactions on Information Theory 38 1992, 6: 1842-1845.

[2] J. Liu, Y. Cao, C. Y. Lin and et al. Low-quality product review detection in opinion summarization [C]//Proc. of EMNLP-CoNLL, 2007: 334-342.

[3] S. M. Kim, P. Pantel, T. Chklovski and M. Pennacchiotti. Automatically assessing review helpfulness [C]//Proc. of EMNLP, 2006: 423-430.

[4] Yiming Yang and Jan O. Pedersen. A Comparative Study on Feature Selection in Text Categorization [C]//Proc. of ICML, 1997: 412-420.

[5] G. Forman. An extensive empirical study of feature selection metrics for text classification[J]. Journal of Machine Learning Research, 2003, 3: 1289-1305.

[6] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques[C]//Proc. of EMNLP, 2002: 79-86.

[7] P. Turney. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews[C]// Proc. of ACL, 2002: 417-424.

[8] M. Hu and B. Liu. 2004. Mining and summarizing customer reviews [C]//Proc. of ACM SIGKDD, 2004: 168-177.

[9] Ana-M. Popescu and O. Etzioni. Extracting product features and opinions from reviews [C]//Proc. of HLT/EMNLP, 2005: 339-346.

[10] N. Kobayashi, K. Inui, Y. Matsumoto and et al. Collecting evaluative expressions for opinion extraction[C]//Proc. of IJCNLP, 2004: 584-589.



COLING 2010 即将在北京召开

经国务院批准,中国中文信息学会将于 2010 年 8 月 23 日—27 日举办“第 23 届国际计算语言学大会”。会前会后有 Tutorials 和 Workshops。经大会筹备组一年多的不懈努力,现已步入正轨。

希望国内专家踊跃参会,登陆 <http://www.coling-2010.org>,即可注册。想了解详情可致电 13671171606 或 13701393267。

中国中文信息学会
2010 年 5 月 5 日