文章编号：

基于 BootStrapping 的集成分类器的中文观点句识别方法*

吕云云[1]，李旸[1]，王素格[1,2]

（1. 山西大学计算机与信息技术学院，太原，030006；

2. 山西大学计算智能与中文信息处理教育部重点实验室，太原，030006）

摘要：领域相关的大规模和高质量的标注训练数据是分类器性能的重要保证，而标注训练语料是一件费时费力的工作。本文提出了一种采用小规模标注语料识别中文观点句的方法。首先采用 Bootstrapping 方法扩展训练语料，分别训练贝叶斯、支持向量机和最大熵分类器。最后，通过给三个训练好的分类器赋权获得一个集成分类器。实验结果表明，集成后的分类器性能优于单分类器，并且该方法在使用部分标注训练数据的情况下也能取得与采用全部标注训练数据相近的实验结果。

关键词：观点句识别；BootStrapping；集成分类器

中图分类号：TP391　　**文献标识码：**A

A Method for Chinese Opinion Sentence Identification Based on the Ensemble Classifier with BootStrapping

Lv Yun[1], Li Yang[1], Wang Suge[1,2]

(1. School of Computer & Information Technology, Shanxi University, Taiyuan, 030006,China ;

2. Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006,China)

Abstract: The large scale and high quality domain training data is an important guarantee for constructing a high performance classifier. However, it is an expensive work to label a large scale corpus in a domain. In this paper, we propose a method for identifying Chinese opinion sentences using a small-scale labeled corpus. At first, the method uses BootStrapping to expand the small-scale labeled corpus. Using the expanded labeled corpus we then train three classifiers that are based on naive Bayes, support vector machine and maximum entropy respectively. At last, an ensemble classifier is obtained by assigning a set of probability weights to the three trained classifiers. Experimental results indicate that the ensemble classifier is superior to the three single classifiers. And the proposed method can achieve the analogous experimental results by using partially labeled training data or using totally labeled training data.

Key words: opinion sentence identifying; BootStrapping; ensemble classifier

1 引言

目前，互联网已经成为思想文化的集散地和社会舆论的放大器。博客、微博、论坛、贴吧等平台使得人们可以自由发表观点和意见，这些信息蕴藏着巨大的潜在价值。例如美国竞选总统期间，奥巴马的竞选团队根据选民的微博，实时分析选民对总统候选人的喜好，为奥巴马竞选成功提供了技术支撑；产品生产商可依据购物网站上顾客的评论，了解用户对产品的满意程度，以此改进产品性能或售后服务；政府及相关部门通过人们对法律、法规、时事政治的看法和观点，了解民情民意。观点句识别技术可以帮助人们尽快获得这些带有观点的句子，为智能导购、市场调查、舆情分析等诸多领域提供数据基础，因此，观点句识别已成

* **收稿日期：**　　**定稿日期：**

基金项目：国家自然科学基金资助项目(No.61175067, 61272095,60970014)；山西省自然科学基金资助项目(No.2010011021-1)；山西省科技攻关项目(No.20110321027-02).

作者简介：吕云云(1988-),女,硕士. 主要研究方向为智能检索；李旸(1988-)，女，硕士生，主要研究方向为智能检索；王素格(1964—),女,教授，博士. 主要研究方向为自然语言处理、智能检索。

为自然语言处理以及文本挖掘领域的热点研究之一。与观点句识别相关的主要技术有基于无监督的识别方法、半监督的识别方法和有监督的识别方法。

基于无监督的观点句识别方法大都以情感词汇为主要识别依据。对于自然语言来言，一个句子的意义是由句子内部各个词汇的意义组合而成。情感词[1-3]往往是句子情感倾向识别的主导因素。Wiebe[4]提出了一个无指导主客观分类方法，该方法主要依据句子中是否包含主观表达来判断句子是否为主观句。Kim 和 Hovy[5]采用基于词典和语料库的方法获得观点词集和非观点词集，并对观点词的极性强度进行度量，最后依据句子中所有词语的极性强度大小或者句子中是否有极性较强的词语出现来判定句子的主客观性。在 COAE2011 评测中，王中卿等[6]通过扩展特征重构观点词表和情感词表，再对两个词表中出现的词赋予不同的权重，采用加权求和的方法判断句子的极性，从而得到观点句。李岩等[7]采用基于句法和 CRFs 的方法首先获得每个词在句子中的句法结构及其是否为实体等特征，利用人工标注了正负倾向的包含形容词的句子训练 CRFs 模型，对测试数据进行标注。利用情感词词典中的词语作为观点句识别的主要手段，这样观点句识别的结果直接依赖情感词典或情感词识别的质量。

基于半监督的方法是标注少量的目标领域或其他领域的标注数据，用于对目标数据进行文本和句子的情感倾向类别预测。Qiu 等[8]提出了 SELC 模型，分两阶段对评论进行分类。Wiebe 等[9]用分布相似度对低频词、搭配、形容词和动词进行主观性聚类，然后利用已知主观词汇作为初始种子，用抽取的模板和概率分类器抽取主观性句子。Pang 等[10]用一种基于寻找文档的最小图割的方法来寻找文档中的主观性部分，以此建立与观点句的对应关系，从而达到对主客观句分类的目的。Riloff 等[11]提出了 bootstrapping 方法，该方法使用 HP-Subj 和 HP-Obj 两个高精确率分类器，用于抽取主观表示的模式，以此建立主观句分类器。

基于有监督的机器学习方法，是以标注类别信息的语料库为基础，训练分类器并用于句子或文本类别识别。Pang 等[12]首先将朴素贝叶斯（Naïve Bayes）、支持向量机（SVM）、最大熵模型（ME）等应用于电影评论的分类研究中，结果表明使用词袋作为特征能够取得很好的结果。徐睿峰等[13]采用多分类器表决的方法进行观点句抽取。该方法在本次评测中取得了较好的成绩，不难看出，具有与领域相关的大规模和高质量的标注训练数据为提高分类器的性能提供了重要的保证。而其他缺乏高质量训练数据的方法则效果明显受到影响，基于集成学习的观点句识别是一种融合多种分类方法，赵立东等[14]基于机器学习和基于规则方法的结果进行了集成，从而得到观点句集；韩先培等[15]采用集成学习的策略，构建基于情感词典的分类器和自学习的领域特定分类器，并使用这些分类器的分类结果作为句子的特征表示，再使用 SVM 算法构建了元分类器将三个领域的所有句子划分三个类别上，将褒义句集和贬义句集作为观点句抽取结果；董喜双等[16]采用最大熵分类的方法判断情感句类别，他们采用词典筛选出观点句，再将观点句切分成短句并用最大熵模型预测短句极性最后通过短句预测长句极性。结果表明这种相结合能够取得相对较好的结果。

本文根据已有的研究，将半监督 BootStrapping 学习方法和集成学习方法融合，提出一个基于 BootStrapping 的集成分类器的中文观点句识别方法，通过实验，验证了方法的有效性。

2 词汇特征选择

本文将每个句子作为一个“词袋”，“词袋”中每个词之间都相互独立。为了有效地表示句子，词袋中的词需要依据主客观区分能力和主观语义进行选择。

（1）基于主客观区分能力的词汇选择[17]

Fisher 线性识别准则的分类思想是寻找空间中的一条直线，使两类样本点在该直线上的投影之间距离最大，而两类样本内部之间的距离即方差最小。本文借用 Fisher 识别准则的分类思想，使用 Fisher 准则函数作为识别观点句子具有区分能力的词选择依据。

$$F(word_k)=\frac{(E(word_k|S)-E(word_k|O))^2}{D(word_k|S)+D(word_k|O)} \quad (1)$$

其中，$word_k$ 代表某一词，S 和 O 分别代表主观和客观，$E(word_k|C)$ 为词 $word_k$ 在类别 C 条件下的期望，$D(word_k|C)$ 为特征项 $word_k$ 在类别 C 条件下的方差，$C\in\{S,O\}$。

（2）基于主观语义的词汇选择

主张词和程度词分别表示主观表达和强度的词汇，本文采用 Hownet 中的具有主观语义信息的主张词和程度副词，共 255 个。

3 分类器选择

现有的机器学习中，朴素贝叶斯分类器、SVM分类器和最大熵模型分类器在文本分类和句子主客观分类取得了较好的成绩[12]。

（1）Naïve Bayes分类器

假定句子sent是由 n 个词$word_1,word_2,\ldots,word_n$组成，则对句子sent ($word_1,word_2,\ldots,word_n$)的识别问题可以转化成求解在该句子出现的条件下主客观类别C_1,C_2出现的概率 $p(C_i|sen)$(i=1,2)如下：

$$p_{Bayes}(C_i|sent)=\frac{p(sent|C_i)p(C_i)}{p(sent)} \quad (2)$$

其中，句子sent在类别 C_i 条件下出现的概率 $p(sent|C_i)$、类别 C_i 出现的概率 $p(C_i)$ 以及句子sent出现的概率 $p(sent)$ 均可从训练语料库中估计。

（2）SVM分类器

对于观点句识别，可以看成二分类问题，而SVM旨在寻找一个最优的超平面 $\mathbf{w}\cdot sent+b=0$，使得该平面将输入空间分成两部分，一部分为观点句，另一部分为非观点句。为了获得高可信度的未标注的类标签，利用文献[18,19]，判断句子为观点句的概率输出函数见公式（3）。

$$p_{SVM}(C|sent)=\frac{1}{1+e^{(Af(sent)+B)}} \quad (3)$$

这里C为观点句类别，参数A和B可通过最小化训练数据的负对数似然值来确定.

（3）ME分类器

ME分类器以最大熵理论为基础，它将已知事实作为约束条件，求使得熵最大化的概率分布作为正确的概率分布。这一理论要求句子sent ($word_1,word_2,\ldots,word_n$)属于某一类别 C_i 的概率 $p(C_i|sent)$ 要使得熵 $H(p)$ 最大化，其中

$$H(p)=-\sum p(C_i|sent)\log p(C_i|sent) \quad (4)$$

采用“特征-类别”对作为特征函数 $f(word,C)$，若当前句子sent包含特征word且类别为 C 时值为1，否则为0。

$$f(word,C)=\begin{cases}1 & C_i=C \text{ and } word\in sent\\ 0 & other\end{cases} \quad (5)$$

解出在该特征约束下的最优概率分布，满足指数函数

$$p_{ME}(C_i \mid sent) = \frac{1}{Z_\lambda(sent)} \exp(\sum_i \lambda_i f_i(word, C)) \quad (6)$$

其中，$Z_\lambda(x)$ 是归一化因子，λ 为特征函数的权重。

4 基于 Bootstrapping 的集成分类器的观点句识别算法

（1）Bootstrapping方法

BootStrapping[20]是一种被广泛应用的自学习策略机器学习方法，本文采用BootStrapping方法对观点句识别的训练集进行扩展。算法思想：采用少量种子集训练初始分类器，并对未标注数据集进行观点分类，把分类结果中具有高置信度的样本加入种子集中，重新训练分类器，直到没有新数据加入种子集为止。最后用训练好的分类模型对测试集进行测试，输出识别的最终观点句。

（2）集成分类器

自学习策略的一个缺点是不能保证所学知识的正确性，当在某一次迭代过程中被错分的样本加入训练集中重新训练分类器，用错误的分类模型进行观点分类，将会导致错误的级联传递，对最终的结果造成不利的影响。因此，应选取较可靠的高置信度分类样本。本文采用性能优良的 naïve Bayes 分类器、SVM 分类器和 ME 分类器进行集成，在此基础上，对各分类器线性加权组成一个更高分类性能的集成分类器，该分类器的分类概率函数如下：

$$p(C \mid sent) = \sum_{i=1}^{3} w^i M^i(C \mid sent) \quad (7)$$

其中，$M^i(C \mid sent)$ 分别表示 $p_{bayes}(C \mid set)$、$p_{svm}(C \mid sent)$ 和 $p_{ME}(C \mid sent)$，C 代表观点句类别，$w^i(i=1,2,3)$ 为赋予各分类器的权重，$w^i \in [0,1](i=1,2,3)$，且 $\sum_{i=1}^{3} w^i = 1$。

（3）BootStrapping的集成分类器的观点句识别算法

利用（1）和（2），BootStrapping的集成分类器的观点句识别算法流程，如图1所示。

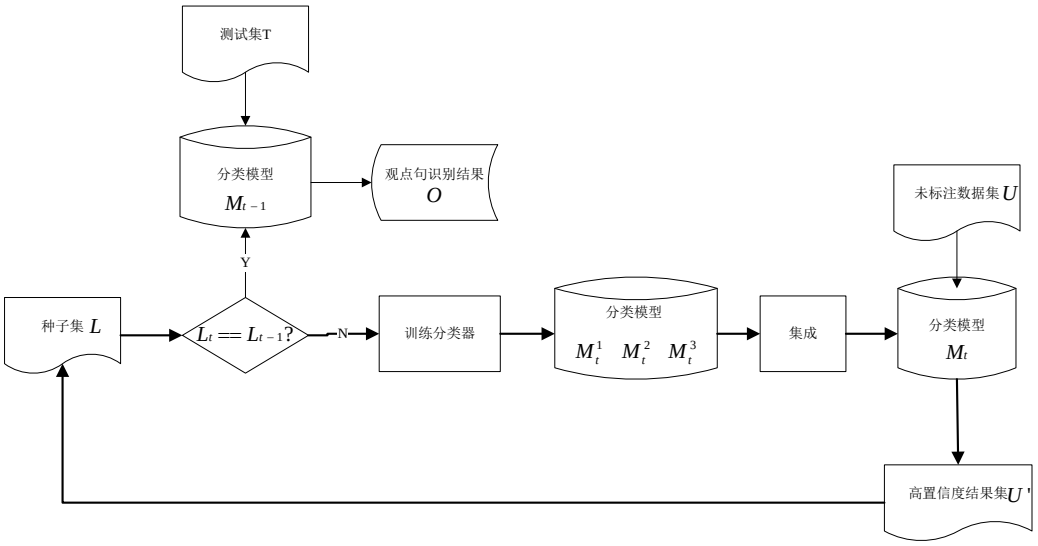

图1 BootStrapping的集成分类器的观点句识别算法流程图

根据图1算法流程，BootStrapping的集成分类器的观点句识别算法（**BFSSIA**）描述如下。

输入：少量已标记的训练句子集L，大量未标注的训练句子集U，测试句子集T

输出：测试集T中的观点句集O

Step 1 令$t=0$，初始化种子集L_t，未标注句子集U；

Step 2 $U'=\varnothing$；

Step 3 用L_t训练分类器$\mathrm{F}_t^i(i=1,2,3$，得到分类模型$M_t^i(i=1,2,3$，集成分类器$p_t(C\mid sent)$；

Step 4 对未标注集U中的每个句子sent利用$p_t(C\mid ser$进行识别。若$p_t(C\mid sent)\geq sub_threshold$，则该句子sent为高置信度观点句且$U'=U'\cup\{sent\}$；否则，若$p_t(C\mid sent)\leq obj_threshold$，则该句子sent为高置信度非观点句且$U'=U'\cup\{sent\}$。

Step 5 令$t=t+1$，$L_t=L_{t-1}\cup U'$，$U=U-U'$；

Step 6 若$L_t==L_{t-1}$，则转Step7，否则，重复Step2-Step6；

Step 7 用$p_{t-1}(C\mid sent)$对测试集T进行观点识别，输出观点句识别结果集O；

Step 8 算法结束

5 实验设计

5.1 实验数据

本文采用第三届中文倾向性分析评测(COAE2011)所发布的语料集《COAE2011_Corpus_Sample_Sentence》中的电子产品语料，该语料共包含2000篇电子产品领域的文档，并且每篇文档均已进行了断句处理。

对该语料进行预处理：首先对语料中的噪音数据进行了清理，去除广告性的句子，如“摩托罗拉600行货，参考价格2399元，公司名称天禧通行货手机网，订购热线01068319570”等类句子，处理后整个语料集共有14928条句子。其中，观点句有5662句，非观点句9266句，比例将近1:2。为了保证数据的平衡性，本文采用随机裁剪的方法，将非观点句集中随机裁剪一些句子，使得到的非观点句与观点句数量相等，此时得到了平衡的观点句与非观点句的语料库。然后再采用中科院的分词与词性标注软件[①]，对所有的句子进行分词与词性标注处理。

5.2 实验方案

利用第3节的三种分类器和集成分类器，采用Bootstrapping方法设计观点句识别的两种方案。

方案1：利用Bootstrapping方法，Naïve Bayes、SVM分、ME三种分类器分别进行观点句的识别；

方案2：利用Bootstrapping方法，将Naïve Bayes、SVM分、ME三种分类器进行集成，得到集成分类器的观点句识别。

为了验证方法的有效性，我们对每种方案中的识别方法分别进行了5折交叉验证实验。观点句识别的评价指标采用常用的三种方法，精确率p，召回率r和F值。

5.3 词汇选取与置信度阈值选择

（1）词汇选取

采用第2节基于主客观区分能力的词汇选择选取1500维特征，再将Hownet中的主张词和程度词(共255个)并入该特征集作为最终的特征词集。实验工具借助了开源的机器学习工具

① http://ictclas.org

包weka[②]和张乐博士的最大熵工具包maxent[③]。

（2）高置信度阈值选择

基于Bootstrapping的观点识别方法需要每次选取具有高置信度的分类样本加入训练集中，以不断提升分类器的性能。通过试验，本文选取每一次对未标注样本集进行分类中的概率值的观点句的阈值 $sub\_threshold = 0.5$ 和非观点句的阈值 $obj\_threshold = 0.3$。

6 实验结果及分析

根据第5.1节的实验设置，不同的标注率标注的数据分别作为种子集，采用第4节的BFSSIA算法，对方案1和方案2进行观点句识别。

方案1：Naïve Bayes、SVM分、ME分类器的实验结果如图2-图4所示。

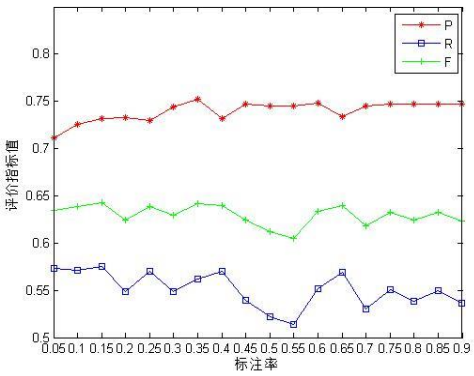

图2 Bayes分类器观点句识别结果

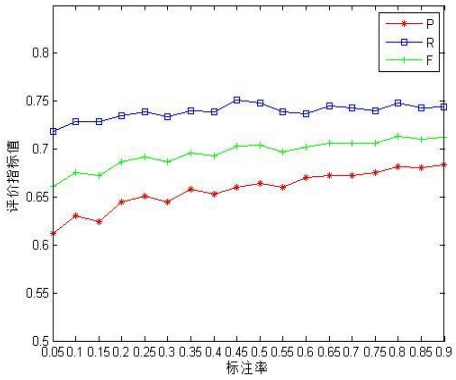

图3 SVM分类器观点句识别结果

由图2-图4可以看出：

（1）采用Bootstrapping方法进行观点句识别，各单分类器当标注率在0.45附近时的识别结果均与较高标注率时识别结果相当，充分证明了采用Bootstrapping方法对训练集进行扩展是可行的。

（2）当使用标注率0.45下的数据作为种子集时，Naive Bayes分类器的分类精度有明显提升，SVM分类器在分类精度损失较小的情况下召回率有所提升，而最大熵分类器的各项性能指标均有明显提升。

（3）三种标注率相同的情况下，Naive Bayes分类器的观点句识别的精确率比较好，而SVM观点句识别的召回率比较高，最大熵模型的观点句识别的精确率、召回率比较接近。

方案2：根据方案1的实验结果，Naïve Bayes、SVM分、ME各分类器的优点也各不相同，因此，设置两种权重，由各分类器线性加权构成集成分类器。

（1）等值权重：集成分类器的权重设置 $w_i = \frac{1}{3}(i = 1,2,3)$。

（2）非等值权重：根据图2-图4，Naïve Bayes的效果较差，因此，设 $w_1 = \frac{1}{5}$，$w_i = \frac{2}{5}(i = 2,3)$。

上述两种情况的集成分类器的观点句识别实验结果如图5-图7所示。

② http://www.cs.waikato.ac.nz/ml/weka

③ http://homepages.inf.ed.ac.uk/lzhang10/maxent_toolkit.html

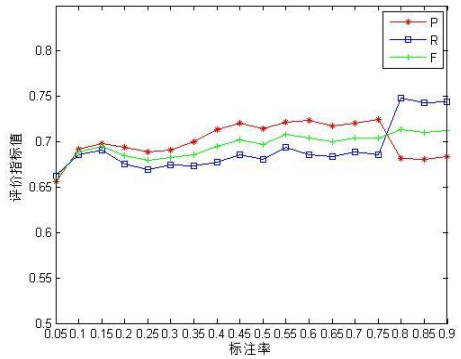

图4 ME分类器观点句识别结果

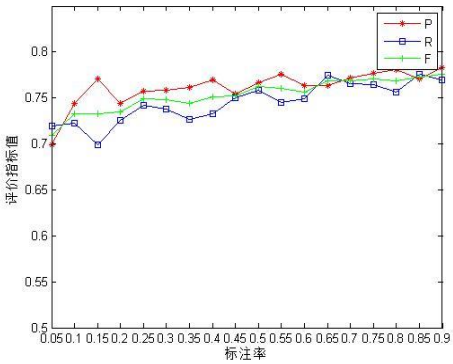

图5 等权集成分类器的观点句识别结果

由图5-图6可以看出，两种权值的集成分类器均取得较单分类器的观点句识别的结果稳定且效果优。

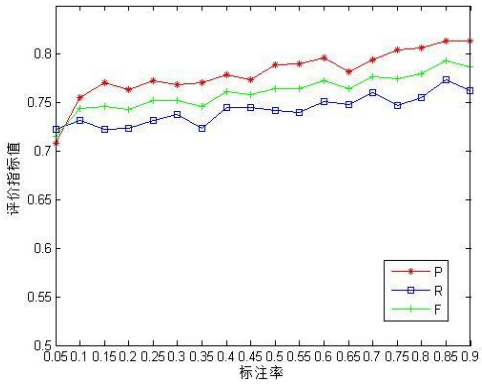

图6 在不等权集成分类器观点句识别结果

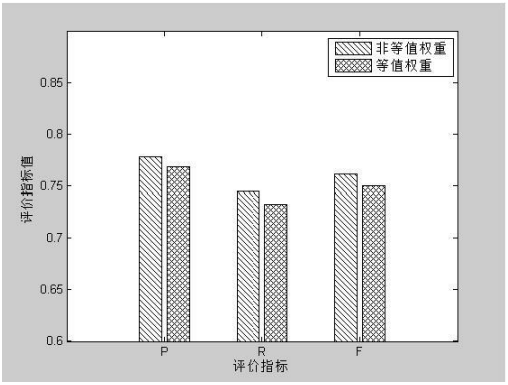

图7 在标注率为0.4时两种集成分类器观点句识别结果

对于图7，在标注率为0.4时，非等权重的集成分类器的观点句识别的各项指标均优于等权重的集成分类器，说明SVM和ME分类器在集成分类器中起着重要的作用。

为了说明各单分类器和集成分类器在训练集的标注率为少数时可以取得比较好的结果，表 1 列出了三种情况的观点句识别结果。

（1）在 Bootstrapping 方法下，各单分类器和集成分类器在训练集的标注率为 0.4 的观点句识别结果。

（2）各单分类器和集成分类器在训练集的标注率为 1，即有监督学习方法观点句的识别结果。

（3）标注数据为 0 时，即无监督方法。该方法将能识别观点句的主要要素评价对象、评价搭配、主张词和程度副词作为特征，评价对象和主张词分别用布尔值表示、评价搭配和程度副词分别用频次表示，通过累加，选取最前面的分值作为观点句识别的依据。

由表1可知：

（1）无论利用Bootstrapping的观点句识别方法在训练集的标注率为0.4和1时，集成分类器优于单分类器的观点句识别结果。

（2）当训练集的标注率为0.4时，各分类器的观点句识别结果可达到训练集为全部带标注类别的水平，即标注率为1。

（3）无监督分类方法中选取分值排名在前60%的观点句识别的结果整体劣于各分类器利用Bootstrapping的观点句识别方法，说明未标注数据对观点句识别的有一定的支持作用。

表 1　各分类器在 Bootstrapping 和有监督方式与无监督的观点句识别实验结果比较

分类器	标注率	Precision	Recall	F
Naive Bayes	1	0.6281	0.7919	0.7001
	0.4	0.7318	0.5700	0.6392
SVM	1	0.6885	0.7352	0.7110
	0.4	0.6525	0.7389	0.6930
ME	1	0.6548	0.6148	0.6342
	0.4	0.7135	0.6776	0.6951
集成分类器	1	0.7724	0.7703	0.7713
	0.4	0.7786	0.7456	0.7617
无监督观点句分值排名前60%	0	0.5514	0.6616	0.6015

7 结束语

本文在三种经典的机器学习分类算法的基础上研究了基于Bootstrapping的观点句识别方法，为了提高分类结果的置信度对三种分类器进行加权集成构建了比单个分类器性能更好的集成分类器，该集成分类器采用少量标注数据做种子集，在Bootstrapping识别思想下取得了比单个分类器进行Bootstrapping识别更好的结果，甚至达到了与训练集全标注时观点句的识别结果。因此，该方法适应于在标注少量数据的情况下的识别任务，尤其适合网络环境下的大数据相关问题的处理。

由于观点句中句法结构复杂，只考虑词汇方面的特征，可能会降低观点句识别的结果，今后可尝试加入句法特征对其进行深入的研究。另外，如何更加科学的选取各种参数也是今后值得研究的问题。

参考文献

[1] 宋乐, 何婷婷, 王倩. 极性相似度计算在词汇倾向性识别中的应用[J]. 中文信息学报. 2010,24(4): 63-67

[2] 王素格,李德玉,魏英杰等. 基于同义词的词汇情感倾向识别方法[J]. 中文信息学报. 2009,23(5): 68-74

[3] Peter D. Turney and Michael L. Littman. Measuring Praise and Criticism: Inference of Semantic Orientation from Association[J]. ACM Transactions on Information Systems. 2003,21(4):315-346.

[4] Janyce Wiebe. Learning subjective adjectives from corpora[A]. Proceeding of National Conference on Artificial Intellifence[C]. 2000.

http://www.cs.columbia.edu/~vh/courses/LexicalSemantics/Orientation/wiebe-aaai2000.pdf

[5] Kim Soo Min and Hovy Eduard. Determining the Sentiment of Opinions[A]. Proceedings of the COLING Conference[C]. Geneva. 2004: 1367-1373

[6] 王忠卿，王荣洋，庞磊等. Suda_SAM_OMS 情感倾向性分析技术报告[A]. 第三届中文倾向性分析评测论文集[C].2011:25-32

[7] 李岩，张佳玥，林宇航等. FRIS_COAE COAE2011评测报告[A]. 第三届中文倾向性分析评测论文集[C]. 2011:42-51

[8] L. Qiu, W. Zhang, C. Hu et al. Selc: A self-supervised model for sentiment classification[A]. Proceeding of the 18th ACM conference on information and knowledge management[C]. 2009. 929–936

[9] Janyce Wiebe, Theresa Wilson et al. Learning Subjective Language[J]. Computational Linguistics. 2004, 30(3): 277-308

[10] Pang Bo and Lillian Lee. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts [A].Proceedings of the ACL[A]. 2004: 271-278

[11] Riloff Ellen and Janyce Wiebe. Learning extraction patterns for subjective expressions[A]. Proceedings of Conference on Empirical Methods in Natural Language Processing[C] (EMNLP-2003). 2003.

[12] Pang Bo, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up sentiment classification using machine learning techniques[A]. Proceedings of Conference on Empirical Methods in Natural Language Processing[C]. 2002

[13] 徐睿峰，王亚伟，徐军等. 基于多知识源融合和多分类器表决的中文观点分析[A]. 第三届中文倾向性分析评测论文集[C]. 2011:77-87.

[14] 赵立东，王素格，王瑞波等. 基于多策略的中文文本倾向分析技术[A].第三届中文倾向性分析评测论文集[C]. 2011:88-96.

[15] 韩先培，孙乐，江雪. 第三届中文文本倾向分析评测ISCAS-Opinion系统报告[A]. 第三届中文倾向性分析评测论文集[C]. 2011:120-125.

[16] 董喜双，邹启波，关毅等. 基于最大熵模型和最小割模型的中文词与句褒贬极性分析[A].第三届中文倾向性分析评测论文集[C]. 2011:97-105.

[17] Suge Wang, Deyu Li Xiaolei Song et al. A feature selection method based on improved fishe's discriminant ratio for text sentiment classification[J]. Expert Systems with Applications. 2011, 38(2011): 8696-8702

[18] J.Platt. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods[A]. Advances in large margin classifiers[C]. In A. Smola, P. Bartlett, B. Schölkopf, & D. Schuurmans (Eds.), Cambridge: MIT Press. 2000

[19] Kai-Bo Duan1 and S. Sathiya Keerthi. Which Is the Best Multiclass SVM Method? An Empirical Study [A] // MCS 2005, LNCS 3541[C]. Springer-Verlag Berlin Heidelberg. 2005: 278–285.

[20] Steven Abney. Bootstrapping[A]. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics[C] . 2002

作者联系方式：姓名：王素格 地址：山西大学计算机与信息技术学院 邮编：030006
电话：13934649855 电子邮箱：wsg@sxu.edu.cn