Journal of Computer Applications
计算机应用,2016,36(6):1613-1618
ISSN 1001-9081
CODEN JYIIDU
2016-06-10
http://www.joca.cn

文章编号:1001-9081(2016)06-1613-06
DOI:10.11772/j.issn.1001-9081.2016.06.1613

# 基于朴素贝叶斯与潜在狄利克雷分布相结合的情感分析

苏 莹[1],张 勇[2*],胡 珀[2],涂新辉[2]
(1.武昌首义学院 信息科学与工程学院,武汉 430064; 2.华中师范大学 计算机学院,武汉 430079)
(* 通信作者电子邮箱 ychang@mail.ccnu.edu.cn)

**摘 要:** 针对情感分析需要大量人工标注语料的难点,提出了一种面向无指导情感分析的层次性生成模型。该模型将朴素贝叶斯(NB)模型和潜在狄利克雷分布(LDA)相结合,仅仅需要合适的情感词典,不需要篇章级别和句子级别的标注信息即可同时对网络评论的篇章级别和句子级别的情感倾向进行分析。该模型假设每个句子而不是每个单词拥有一个潜在的情感变量;然后,该情感变量再以朴素贝叶斯的方式生成一系列独立的特征。在该模型中,朴素贝叶斯假设的引入使得该模型可以结合自然语言处理(NLP)相关的技术,例如依存分析、句法分析等,用以提高无指导情感分析的性能。在两个情感语料数据集上的实验结果显示,该模型能够自动推导出篇章级别和句子级别的情感极性,该模型的正确率显著优于其他无指导的方法,甚至接近部分半指导或有指导的研究方法。

**关键词:** 情感分析;主题模型;潜在狄利克雷分布;朴素贝叶斯;意见挖掘
**中图分类号:** TP391.1 **文献标志码:** A

## Sentiment analysis research based on combination of naive Bayes and latent Dirichlet allocation

SU Ying[1], ZHANG Yong[2*], HU Po[2], TU Xinhui[2]
(1. *College of Information Science and Engineering, Wuchang Shouyi University, Wuhan Hubei* 430064, *China*;
2. *School of Computer, Central China Normal University, Wuhan Hubei* 430079, *China*)

**Abstract:** Generally the manually labeled corpus is a critical resource for sentiment analysis. To circumvent laborious annotation efforts, an unsupervised hierarchical generation model for sentiment analysis was presented, which was based on the combination of Naive Bayes (NB) and Latent Dirichlet Allocation (LDA), named Naive Bayes and Latent Dirichlet Allocation (NB-LDA). Just needing the right emotional dictionary, the emotional tendencies of network comments were analyzed at sentence level and document level simultaneously without sentence level and document level markup information. In particular, the proposed model assumed that each sentence instead of each word had a latent sentiment label, and then the sentiment label generated a series of features for the sentence independently by the NB manner. The proposed model could combine the advanced Natural Language Processing (NLP) correlation technologies such as dependency parsing and syntactic parsing by the introduction of NB assumption and could be used to improve the performance for unsupervised sentiment analysis. The experimental results conducted on two sentiment corpus datasets show that the proposed NB-LDA can automatically derive the emotional polarities of sentence level and document level, and significantly improve the accuracy of sentiment analysis compared to the other unsupervised methods. Moreover, as an unsupervised model, the NB-LDA can achieve comparable performance to some supervised or semi-supervised methods.

**Key words:** sentiment analysis; topic model; Latent Dirichlet Allocation (LDA); Naive Bayes (NB); opinion mining

## 0 引言

近年来,情感分析,或意见挖掘,已经成为国内外的研究热点,其主要目标在于从评论性的文本中揭示作者的观点、意见、态度、喜好等潜在情感状态,例如,从用户的网络评论中,自动识别出用户的喜好、态度,分析商品的质量缺陷等。因此,情感分析在商业上具有非常广阔的应用前景,也广泛应用在股票经济预测、政治人物支持率分析等多个方面。

目前,情感分析的主要方法通常严重依赖人工标注的标注语料,常见的方法以文本分类为主,即以人工标注语料为训练数据来训练情感分类器。因此,对情感分析而言,人工标注数据成为了非常宝贵的资源。为了节省人工标注成本,本文尝试提出一种无指导的情感分析模型。

从分类粒度而言,情感分析可以分为篇章级别的情感分析、句子级别的情感分析,以及短语级别的情感分析等。以往的研究[1-2]表明,不同粒度之间情感分析能够相互强化,互相促进。目前为止,已有一些研究者提出来了将篇章级别和句子级别的情感分析相结合的情感分析方法[1,3-4],但是这些方

收稿日期:2015-11-30;修回日期:2016-02-23。 基金项目:国家社会科学基金重大项目(12&2D223);国家自然科学基金资助项目(61402191,61300144,61572223);国家语委科研项目(WT125-44);华中师范大学自主科研项目(CCNU14A05014,CCNU14A05015)。

作者简介:苏莹(1982—),女,河南信阳人,讲师,硕士,主要研究方向:文本挖掘、机器学习; 张勇(1978—),男,湖北仙桃人,副教授,博士,CCF会员,主要研究方向:文本挖掘、自然语言处理; 胡珀(1980—),男,湖北武汉人,副教授,博士,CCF会员,主要研究方向:自动文摘、机器学习; 涂新辉(1979—),男,湖北应城人,副教授,博士,CCF会员,主要研究方向:信息检索、自然语言处理、机器学习。

法仍然属于有指导的,或半指导的方法。因此,本文则尝试提出一种无指导的篇章级别和句子级别相结合的情感分析模型,利用篇章级别和句子级别的互强化来提高整体的情感分析性能。

相对于篇章级别的情感分析而言,句子级别的情感分析更为复杂,其主要表现在于句子级别的情感分析需要考虑更多的语法、句法结构。例如,否定的表达就是句子级别情感分析的一个主要问题,特别是长距离依存的否定表达(如,没有人认为这是一个好现象)。此外,句子级别的情感分析还受到时态、词义、句法成分、修饰等影响(如:“我一直以为这是一个值得看一看的景点”)。要解决句子级别情感分析的这些问题,就需要引入句法分析、依存分析等自然语言处理技术。因此,本文在主题模型(Topic Model)的基础上,引入了朴素贝叶斯模型,用以结合自然语言处理技术与统计模型,以提高情感分析的性能。

综上所述,本文提出的模型有三个主要的优点:首先,该模型是一种无指导的情感分析模型,不需要人工标注语料进行训练;其次,该模型能对篇章级别和句子级别的情感同时进行分析,结合篇章级别和句子级别情感之间关系,通过两者之间的互强化来提高情感分析性能;最后,在传统的主题模型基础上,该模型能够结合自然语言处理的相关技术,引入句子级别的语言学特征,以解决句子级别情感分析的主要难点。该模型的实现策略则是在句子级别引入了朴素贝叶斯假设,使其能够对句子级别的特征进行建模。

## 1 相关研究工作

已有的情感分析研究包括很多方面,主要有情感极性分类[2,5]、意见抽取[6]和意见成分识别[7]等,不同的方法和系统处理的粒度也有较大区别。通常而言,细粒度的情感分析(如句子级别,或短语级别的情感分析)一般缺乏足够的细粒度标注语料,难度也较粗粒度的情感分析(如篇章级别的情感分析)困难。为了解决细粒度情感分析的困难,有研究者编撰了相关的情感词典[8],并在句子依存分析的基础上提出了一系列的分析特征,通过机器分类的方法来对细粒度的短语进行情感分类。但是,通常来说,这类方法需要大量细粒度人工标注语料作为训练数据。因此,本文试图研究一种无指导的情感分析方法,以避免或减少对人工标注数据的依赖。

同时,在无指导的情感分析方面,主题模型显示出了其特有的优势。主题模型广泛应用在多个研究领域[9-11],也有相关研究在主题模型的基础上提出了一系列的情感分析模型[12-14]。这些方法主要试图同时识别篇章中的主题和情感,以提高篇章级别的情感分析性能,因为相关研究表明,情感分析是依赖于主题分析的。但是,这些方法也继承了主题模型中“词袋”(Bag of Words)假设,即每个单词都存在一个潜在的情感变量,这显然是与人类的语言思维是不符的。近年来,相关研究者也尝试突破“词袋模型”的假设,如文献[11]提出了两个情感分析模型:TME(Topic and Multi-Expression)和ME-TME(Maximum-Entropy Topic and Multi-Expression)模型,同时从在线评论中抽取主题词语和评价表达,但是他们的模型限定每个情感表达最多只包括连续的4个词语。而本文假设情感表达的基本单位是句子,而不是单词,克服了传统的主题模型在情感应用上的不足。

另一方面,相关的研究显示,将篇章级别和句子级别的情感分析相结合,能够有效地提高情感分析的性能。文献[2]为了减少客观性句子对情感分析的影响,采用了对句子关系图的最小划分策略来排除客观性句子,以篇章级别的情感分析性能。但是,该研究并没有对句子进行情感分析。

近年来,同时结合篇章级别和句子级别情感分析的研究也取得了一些不错的进展[1,3-4,15]。文献[1]提出了一种结构化的图模型,同时进行粗粒度和细粒度的情感分析,该方法主要采用一种条件随机场(Conditional Random Field, CRF)的序列分类方法进行建模,但是该方法同时需要粗粒度和细粒度的人工标注数据作为训练语料。为了避免繁重的细粒度人工标注,文献[3]则提出了一种基于因子图(Factor Graph)的结构化模型,该模型只需要粗粒度的标注数据(篇章级别的情感标注)作为训练语料,可同时进行篇章级别和句子级别的情感分析;但是,由于该模型在句子级别情感分析的性能有限。文献[4]则提出了两种结构化条件模型,采用大量粗粒度的标注数据(比较容易从互联网上获取)和少量细粒度的人工标注数据,以提高细粒度的分析性能。但是这些方法都依赖于标注数据,属于有指导或半指导的情感分析方法。

与本文提出的方法最相近的模型是主题和情感统一分析模型(Aspect and Sentiment Unification Model, ASUM)[16]。但是,本文的模型与其主要不同点在于两个方面:首先,ASUM主要用于抽取情感和主题的表达,而不是情感分类;其次,ASUM无法对句子的特征进行建模,无法引入自然语言处理的相关技术来提高情感分析的性能。另外一个与本文模型比较接近的是应用于股票预测的主题和情感的潜在狄利克雷分布(Topic Sentiment Latent Dirichlet Allocation, TSLDA)模型[17],TSLDA假设每个句子具有一个潜在的主题和情感变量,同时该模型还需要标注每个词语为主题词,或者情感词,或者其他词语,这种细粒度的标注工作需要耗费大量的人力,且不容易推广应用。而且,TSLDA模型也没有引入朴素贝叶斯的概念,无法对句子内部的特征进行建模。

在以上相关研究的基础上,本文结合朴素贝叶斯模型和主题模型,提出了一种层次性的生成模型——朴素贝叶斯与潜在狄利克雷分布(Naive Bayes and Latent Dirichlet Allocation, NB-LDA)联合模型。该模型能对句子内部的语言学特征进行建模,同时在篇章级别和句子级别进行情感分类。当然,也有一些相关的研究综合了朴素贝叶斯和主题的优点,如文献[18]的基于朴素贝叶斯的潜在狄利克雷(Latent Dirichlet Conditional Naive-Bayes, LD-CNB)模型,以及基于贝叶斯定律的朴素贝叶斯(Bayesian Naive Bayes)模型,但这些模型无论在应用领域、应用目的,还是模型的设计上,都与本文的模型有显著不同。

## 2 NB-LDA模型

### 2.1 模型提出的基础

本文研究的基础在于篇章级别的情感极性与句子级别的情感极性具有很强的一致性。文献[3]的研究表明,篇章的

情感与其中句子的情感具有一致性,如表 1 所示,在情感极性为正面的篇章中,有 53% 的句子是正面的,只有 8% 的句子是负面的,另外 39% 的句子是中性的;而在负面的篇章中,则有 62% 的句子是负面的,只有 5% 的句子是正面的,剩下的 33% 的句子是中性的;在情感极性为中性的篇章中,也有 51% 的句子是中性的。显然,篇章级别和句子级别的情感极性的分布一致性能够被利用,以提高情感分析的性能。因此,本文采用主题模型的基本原理,即利用篇章中的句子情感的聚集性,自动识别出具有相同情感极性的句子。同时,本文的 NB-LDA 模型还引入了朴素贝叶斯模型,用以对句子内部特征进行建模,使得该模型能够利用句子内部的语言学特征来识别句子的情感极性。

表 1　不同情感极性的篇章中的句子极性分布

| 篇章情感极性 | 句子极性所占比例 | | |
|---|---|---|---|
| | 正面 | 负面 | 中性 |
| 正面 | 0.53 | 0.08 | 0.39 |
| 负面 | 0.05 | 0.62 | 0.33 |
| 中性 | 0.14 | 0.35 | 0.51 |

相对于其他相关的情感分析模型,本文提出的 NB-LDA 模型有三个主要的优点。首先,NB-LDA 能对句子内部的语言学特征进行建模,充分利用句子内部的句法和依存关系,来提高细粒度的情感分类性能。其次,NB-LDA 是一个统一的生成模型,其主要原因在于朴素贝叶斯和 LDA 模型都是生成模型,因此,两者的结合就具有很强的合理性。最后,NB-LDA 模型避免了传统的主题模型的"词袋"假设。一般来说,情感表达是以句子为单位,而不是以词语为单位。在 NB-LDA 模型中,句子的情感是通过句子内部的语言学特征来识别的,同时利用篇章中句子情感的聚集性进行互强化,以提高整体的分析性能。

### 2.2　NB-LDA 模型描述

本文的 NB-LDA 模型属于生成模型,其中,每个文档包含一定数据量的句子,每个句子表达了一个潜在的情感,用 $z$ 表示。句子中所有词语都可以看作是潜在情感变量 $z$ 生成的特征。在本文中,朴素贝叶斯模型通过句子内部的词语,以及一系列的语言学特征来识别句子的情感极性(正面、负面或者中性),而不是假设每个词语都具有潜在情感极性[3-4]。对于每篇文档而言,本文假设文档是句子的组合,每篇文章有一个在 $T$ 种情感极性上的潜在混合分布变量 $\boldsymbol{\theta}$,该变量服从参数为 $\alpha$ 的 Dirichlet 分布。

为了更好地描述该模型的生成过程,本文假设存在 $T$ 种情感极性(一般是三种,正面,负面和中性),$T$ 类似于 LDA 模型中主题数目。$D$ 表示文档的数目,$F$ 表示句子中所有特征的数目,$S_d$ 表示文档 $d$ 中的句子数目,$S_{d,i}$ 表示文档 $d$ 中的第 $i$ 个句子,$F_{d,i}$ 表示文档 $d$ 中的第 $i$ 个句子中的特征数目。所有文档在情感极性上的分布,用 $\boldsymbol{\Theta}$ 表示,该变量是一个 $D * T$ 的矩阵,其中每一行用 $\boldsymbol{\theta}_d$ 表示,代表文档 $d$ 的情感分布。类似的,$T* F$ 的矩阵 $\boldsymbol{\Phi}$ 表示情感极性在特征上的分布,其中每一行,$\varphi_t$,表示该情感极性在特征上的分布。其中,$\boldsymbol{\Theta}$ 和 $\boldsymbol{\Phi}$ 分别服从参数为 $\alpha$ 和 $\beta$ 的 Dirichlet 分布。

因此,整个生成过程可以描述如下:

1) 对每篇文档 $d$,生成 $\boldsymbol{\theta}_d \sim Dir(\alpha)$;

2) 对每个情感极性 $t$,生成 $\varphi_t \sim Dir(\beta)$;

3) 对文档 $d$ 中的每个句子 $i$:

3.1) 生成情感变量 $z_{d,i} \sim Multi(\boldsymbol{\theta}_d)$;

3.2) 对句子中的每个特征 $j$;

3.2.1) 生成特征值 $f_{d,i,j} \sim Multi(\varphi_{z_{d,i}})$。

在以上生成过程中,与传统的 LDA 模型最大的区别在于,NB-LDA 模型假设文档 $d$ 中的每个句子存在一个潜在情感变量 $z$,而不是每个单词;然后,潜在情感变量 $z$ 以朴素贝叶斯的方式生成句子中的单词以及其他语言学特征。NB-LDA 模型的图形表示如图 1 所示。

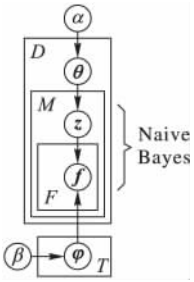


图 1　NB-LDA 模型的图形化描述

### 2.3　NB-LDA 模型推导

目前有几种不同主题模型参数推导方法[9,19]。本文采用了 Gibbs Sampling 的方法[19],该方法具有推导简单、直观的特点,同时能够综合多个局部最大值以获取更优的参数估计。

根据 2.2 节的生成假设,潜在情感变量 $z$ 和特征 $f$ 的联合分布可以分解成式(1):

$$p(\boldsymbol{z},\boldsymbol{f}) = \prod_{d=1}^{D}\prod_{i=1}^{S_d} p(z_{d,i}) \prod_{j=1}^{F_{d,i}} p(f \mid z_{d,i}) \tag{1}$$

根据 Gibbs Sampling 的推导原理,本文首先要构建一个马尔可夫链,该马尔可夫链将收敛到在 $z$ 上的后验分布,通过不断的对单个 $z$ 的抽样来评估整体的后验分布,其抽样公式如式(2)所示:

$$p(z_{d,i} = k \mid s_{d,i}, \boldsymbol{z}_{d,-i}, \boldsymbol{s}_{d,-i}) \propto \frac{C^{DT}_{d,k} + \alpha}{\sum_{t=1}^{T} C^{DT}_{d,t} + T\alpha} \cdot \prod_{j \in F_{d,i}} \frac{C^{FT}_{j,k} + \beta}{\sum_{f=1}^{F} C^{FT}_{f,k} + F\beta} \tag{2}$$

其中:$C^{DT}_{d,k}$ 表示在文档 $d$ 中,情感极性为 $k$ 的句子数目,不包括当前的句子;$C^{FT}_{j,k}$ 表示情感极性为 $k$ 的句子中包括有特征 $j$ 的次数,不包括当前的句子。

经过一定数目的迭代抽取后,潜在变量 $z$ 将趋近真实的后验分布,模型的主要参数则可以通过式(3)和(4)进行评估:

$$\theta_{d,k} \propto (C^{DT}_{d,k} + \alpha) \Big/ \Big(\sum_{t=1}^{T} C^{DT}_{d,t} + T\alpha\Big) \tag{3}$$

$$\varphi_{k,j} \propto (C^{FT}_{j,k} + \beta) \Big/ \Big(\sum_{f=1}^{F} C^{FT}_{f,k} + F\beta\Big) \tag{4}$$

其中:$\varphi_{k,j}$ 表示在情感极性 $k$ 中,特征 $j$ 出现的概率;$\theta_{d,k}$ 表示文档 $d$ 中,情感极性 $k$ 所占的比重。

## 3　实验设置

本章将主要介绍本文实验所使用的数据集、情感词典、特征空间,以及 Baseline 方法。在本文中,先验参数 $\alpha$ 和 $\beta$ 分别设置为 0.3 和 0.01,与文献[18]类似。

### 3.1 数据集和情感词典

本文使用了两个英文数据集来评估NB-LDA模型：第一个数据集是电影评论数据集(http://www.cs.cornell.edu/People/pabo/movie-review-data/)；第二个数据集是细粒度情感标注数据集[3]。在预处理阶段，本文采用了斯坦福的自然语言处理工具(http://nlp.stanford.edu/software/corenlp.shtml)，对文档进行了句子切分、句法分析和依存分析，同时去除了部分停用词。

实验中，本文采用了一个公开的英文情感词典[3]。在预处理阶段，本文将文档中的词语，与情感词典中词语的词形、词性等进行比较，如果匹配成功，则将该词标注为词典中标注的情感极性，整个句子的情感极性则选择句子内部占大多数的情感极性。需要注意的是，情感词典仅仅在预处理阶段使用，作为先验知识来引导模型的推导。

### 3.2 特征空间

本文所采用的特征空间主要参考了文献[3]的研究，如图2所示。其中，本文采用的特征与文献[3]主要的不同之处有两点：首先，本文采用词语的lemma作为特征，而不是词语本身的词形；其次，本文增加了句子位置特征，包括句子是否在文档开头(first)或者文章结尾(last)，还是在文章中间(mid)。增加该特征的主要原因在于文档开头或结尾的句子往往与文档本身的情感极性高度相关。所有特征的识别是在依存分析的结果上，根据文献[3]的相关描述，在预处理阶段进行分析识别。

### 3.3 Baseline方法

本文采用了三种Baseline方法用来对比实验，以评估本文模型的有效性。

lemma－prior：该方法仅仅使用词语的lemma作为特征，而没有使用情感词典，以及其他任何特征。该方法主要用来对比分析情感词典和句子内部特征对情感分析性能的影响。

lemma＋prior：该方法在lemma－prior的基础上增加了情感词典的使用，主要用来对比分析情感词典的作用。

ASUM＋prior：该方法采用了ASUM[14]，并同时使用情感词典作为先验知识，以对比分析本文NB-LDA模型与ASUM的差异。

| 词语特征 |
| --- |
| word lemma |
| word prior polarity: positive, negative, both, neutral |
| reliability class: strongsubj or weak subj |
| **极性特征** |
| negated: binary |
| negated subject: binary |
| modifies polarity: positive, negative, neutral, both, notmod |
| modified bypolarity: positive, negative, neutral, both, notmod |
| conj polarity: positive, negative, neutral, both, notmod |
| **句子特征** |
| cardinal number in sentence:binary |
| pronoun in sentence: binary |
| modal in sentence(other than will): binary |
| **结构化特征** |
| Sentenceposition: first, mid, last |

图2 NB-LDA模型采用的特征

## 4 实验结果

### 4.1 电影评论上的实验结果

在电影评论数据集上，所有的文档被标注为正面，或者负面，因此实验中只需要考虑两种情感状态：正面或负面，因而$T$被设置为2。在模型推导结束后，根据式(3)可以评估每篇文章的情感分布$\theta_d$，如果$\theta_{d,pos}$大于$\theta_{d,neg}$，则文档$d$被认为是正面的，否则被认为是负面的。该数据集上的情感分类结果如表2所示，其中正确率为10次实验的平均结果。

表2 电影评论语料上的情感分类正确率 %

| 算法策略 | 正面正确率 | 负面正确率 | 总计正确率 |
| --- | --- | --- | --- |
| lemma－prior | 55.0 | 59.5 | 57.25 |
| lemma＋prior | 72.7 | 63.2 | 67.95 |
| ASUM＋prior | 72.1 | 63.3 | 67.70 |
| NB-LDA | 75.3 | 68.4 | 71.85 |
| 文献[12]算法 | 74.1 | 66.7 | 70.40 |
| 文献[20]算法 | | | 70.90 |
| 文献[21]算法(使用10%的标注数据) | | | 60.00 |
| 文献[21]算法(使用35%标注数据) | | | 69.00 |

从表2可以看出，方法lemma－prior的性能很一般，正确率只有57.25%，但是在增加了情感词典后，方法lemma＋prior的正确率提高了10.7个百分点，两个结果的对比表明，情感词典对情感分析具有非常重要的作用。

相对于lemma＋prior方法和ASUM＋prior方法，NB-LDA模型加入了句子内部的特征，实验中取得了71.85%的正确率，在lemma＋prior方法和ASUM＋prior的基础上又至少提高了3.9个百分点，该结果表明，句子内部的语言学特征起到了非常积极的作用。

同时，ASUM＋prior方法的性能与lemma＋prior的结果类似，两者的主要区别在于ASUM同时对主题和情感进行了建模，而lemma＋prior只对情感进行了建模，从情感评估的角度而言，两种方法高度相似。

表2同时显示了最近的相关研究文献的实验结果。与文献[20]的谱聚类算法相比较，NB-LDA模型的正确率提高了近1个百分点。但是需要注意的是，文献[20]算法需要用户根据每个维度的特征来指定每个维度的情感极性，然后再进行聚类以获得最后的输出结果，而本文的方法则不需要任何的人工辅助。文献[21]采用了一种非负矩阵分解的方法来进行半指导的情感分类，当使用10%的篇章级标注数据训练时，其也只取得了60%左右的正确率，甚至在使用35%的标注数据时，其正确率仍然比本文的NB-LDA方法略差，而本文的NB-LDA方法没有使用任何的标注数据。

从表2的实验结果可以发现另外一个有趣的现象：情感词典对正面文档和负面文档的影响效果是不同的。在没有使用情感词典时，正面文档的分类正确率要低于负面文章；但是，在使用情感词典后，正面文档的情感分类正确率大幅度提高了，显著高于负面文档的正确率。研究人员进一步的人工检查发现，在整个电影评论语料中，根据情感词典匹配后，正面词语的数目远远大于负面词语的数目，从而导致正面先验知识和负面先验知识的不平衡性。同时，在细粒度情感标注数据集中，实验结果也显示出了类似的现象，文献[3－4,12]的实验结果也具有类似的现象。而文献[12]的研究则针对电影评论语料的特点，修改了情感词典，从而获得较好的正确率，但依然略低于本文模型的性能。

### 4.2 细粒度情感标注数据集上的实验结果

在细粒度情感标注数据集中，共有5类标注：POS(正面)、NEG(负面)、NEUT(中性)、MIX(混合情感)和NR(与情

感无关)。与文献[3]类似的处理策略,本文将 MIX 和 NR 两类标注归为 NEUT 类别。因此,在该实验中,$T$ 被设置为 3,所有句子的情感极性以抽样迭代后 $z$ 变量的值为最终结果。

表 3 显示了本实验的结果,包括篇章级别的正确率和句子级别的正确率,以及 F1 值。同时,表 3 的下半部分列举了文献了[3-4]在该数据集上的实验结果。从 lemma - prior、lemma + prior、ASUM + prior 和 NB-LDA 的实验结果来看,lemma + prior 在句子级别要高于 lemma - prior 约 6 个百分点,在篇章级别要高约 8 个百分点,而 NB-LDA 的性能仍然是四个方法中最好的。该结果表明,合适的先验情感词典和语言学特征对提高情感分析的性能非常重要。

表 3　细粒度情感标注数据集上的分类结果　%

| 算法策略 | 句子 | | | | 篇章 |
|---|---|---|---|---|---|
| | 总正确率 | 正面 F1 值 | 负面 F1 值 | 中性 F1 值 | 正确率 |
| lemma - prior | 36.0 | 25.7 | 38.6 | 38.8 | 40.5 |
| lemma + prior | 41.9 | 46.4 | 49.5 | 31.6 | 48.6 |
| ASUM + prior | 41.4 | 45.9 | 49.0 | 31.2 | 47.3 |
| NB-LDA | 46.8 | 53.4 | 49.6 | 38.0 | 54.4 |
| VoteFlip[3] | 41.5 | 45.7 | 48.9 | 28.0 | — |
| SaD[3] | 47.6 | 52.9 | 48.4 | 42.8 | — |
| DaS[3] | 47.5 | 52.1 | 54.3 | 36.0 | 66.6 |
| HCRF(soft)[3] | 53.9 | 57.3 | 58.5 | 47.8 | 65.6 |
| HCRF(hard)[3] | 54.4 | 57.8 | 58.8 | 48.5 | 64.6 |
| Interpolated[4] | 59.1 | — | — | — | — |

注:"—"表示内容为空。

与已有的研究相比较,NB-LDA 模型的性能显著高于无指导的 VoteFlip[3] 方法,与有指导的 SaD(Sentence as Document)[3] 和 DaS(Document as Sentence)[3] 的结果相当,该实验结果说明,无指导的 NB-LDA 模型在该数据集上取得相当不错的表现。

但是,相对于隐藏条件随机场(Hidden Conditional Random Field, HCRF)[3] 等方法,本文的 NB-LDA 方法落后 7 到 12 个百分点,但是需要注意的是,HCRF[3] 等方法均使用了大量的粗粒度标注数据,而文本的 NB-LDA 则没有使用任何的标注数据。

### 4.3　情感词语的检测

虽然 NB-LDA 模型主要用于情感分类,但与传统的主题模型类似的是,NB-LDA 模型也能够通过式(4)来评估参数 $\varphi_{k,j}$,以判断每个词语的情感极性。细粒度情感标注数据集上获取到的部分正面词语有:good、great、more、want、best、love、better、well、only、original;负面词语有:bad、little、too、hard、just、so、out、also、even、more。

大部分词语的情感分类都比较准确。同时也发现,这些词语都属于比较常见的词语,其主要原因在于主题模型对高频词语非常敏感,即高频的常见词语一般都会排在每个类别的前列,而一些与领域相关的情感词语由于出现的频次不高,则难以排到每个情感类别的前 50 位。因此,如何检测特定领域的情感词语也将是本文将来的研究工作之一。

## 5　结语

本文结合朴素贝叶斯方法和主题模型,提出了一种层次性生成模型,用于篇章级别和句子级别的无指导情感分析。该模型假设文档中每个句子具有一个潜在的情感变量,而该情感变量以朴素贝叶斯的方式生成句子内部的语言学特征或词语。该假设使得 NB-LDA 模型既具有主题模型的聚类特点,同时避免了主题模型的"词袋"假设。朴素贝叶斯的引入使得 NB-LDA 模型具有对句子内部语言学特征建模的能力,从而能够引入自然语言处理的相关技术来提高情感分析的性能。在两个情感数据集上的实验结果显示,相对于其他无指导的方法,NB-LDA 模型在篇章级别和句子级别均能取得相当不错的实验结果。当然,相对于有指导或半指导的方法而言,NB-LDA 的性能仍然有一定的差距。

针对 NB-LDA 模型的特点,未来工作将主要集中在以下方面:

1)既然无指导的情感分析方法难以企及有指导的策略,在 NB-LDA 模型的基础上,引入粗粒度的情感标注信息将能显著提高情感分析的性能,而带有篇章级别的情感标注信息可以从网络上大量自动获取到(如淘宝评论)。在利用篇章级别的评分信息进行情感分析方面,相关研究已经取得一定的成果,如文献[22]提出了一种将协同过滤和情感分析相结合的主题模型分析方法,而文献[23]则提出了一种神经网络方法来学习用户评分的情感倾向,以构造一个情感评分的预测模型;当然,这些文献主要研究用户对情感评分的影响,而本文侧重挖掘句子内部的结构化特征。

2)目前本文采用的语言学特征相对比较简单,如何针对情感分析的特点,挖掘更加有效的句子、词汇特征将成为研究重点,例如,在词汇的表达层面引入深度学习的相关研究成果,特别是 Word Embedding 的词汇表示方法,如文献[24]利用卷积神经网络进行图文情感分析取得了较好的效果。

3)在 NB-LDA 模型中,特征选择主要使用的是句子内部的依存关系,如何引入更加高级的特征,如文献[25]所研究的评价对象、情感词语、否定词等与情感分析密切相关的成分特征,将是进一步需要研究的问题。

参考文献:


[1] MCDONALD R, HANNAN K, NEYLON T, et al. Structured models for fine-to-coarse sentiment analysis [C] // ACL 2007: Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2007: 432 - 439.

[2] PANG B, LEE L. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts [C] // ACL 2004: Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2004: 271 - 278.

[3] TÄCKSTRÖM O, MCDONALD R. Discovering fine-grained sentiment with latent variable structured prediction models [C] // ECIR 2011: Proceedings of the 33rd European Conference on Information Retrieval. Berlin: Springer, 2011: 368 - 374.

[4] TÄCKSTRÖM O, MCDONALD R. Semi-supervised latent variable models for sentence-level sentiment analysis [C] // ACL 2011: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2011: 569 - 574.

[5] ZHOU G, ZHAO J, ZENG D. Sentiment classification with graph co-regularization [C] // COLING 2014: Proceedings of the 25th International Conference on Computational Linguistics. Stroudsburg: ACL, 2014: 1331 - 1340.

[6] DING X, LIU B, ZHANG L. Entity discovery and assignment for opinion mining applications [C] // KDD 2009: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2009: 1125 - 1134.

[7] ZHAO W X, JIANG J, YAN H, et al. Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid [C] // EMNLP 2010: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2010: 56 - 65.

[8] WILSON T, WIEBE J, HOFFMANN P. Recognizing contextual polarity in phrase-level sentiment analysis [C] // HLT'2005: Proceedings of the 2005 Conference on Human Language Technology and Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2005: 347 - 354.

[9] BLEI D, NG A, JORDAN M. Latent Dirichlet allocation [J]. Journal of Machine Learning Research, 2003, 3(5): 993 - 1022.

[10] GRUBER A, WEISS Y, ROSEN-ZVI M. Hidden topic Markov models [C] // AISTATS 2007: Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics. [S. l.]: JMLR, 2007, 2:163 - 170.

[11] MUKHERJEE A, LIU B. Modeling review comments [C] // ACL 2012: Proceedings of the 23rd International Conference on Computational Linguistics. Stroudsburg: ACL, 2012: 320 - 329.

[12] LIN C, HE Y. Joint sentiment/topic model for sentiment analysis [C] // CIKM 2009: Proceedings of the 18th ACM Conference on Information and Knowledge Management. New York: ACM, 2009: 375 - 384.

[13] MEI Q, LING X, WONDRA M, et al. Topic sentiment mixture: modeling facets and opinions in weblogs [C] // WWW 2007: Proceedings of the 16th International Conference on World Wide Web. New York: ACM, 2007: 171 - 180.

[14] ZHANG Y, JI D, SU Y, et al. Sentiment analysis for online reviews using an author-review-object model [C] // AIRS 2011: Proceedings of the Seventh Asia Information Retrieval Societies Conference. Berlin: Springer, 2011: 362 - 371.

[15] BAGHERI A, SARAEE M, JONG F D. ADM-LDA: An aspect detection model based on topic modelling using the structure of review sentences [J]. Journal of Information Science, 2014, 40(5): 621 - 636.

[16] JO Y, OH A H. Aspect and sentiment unification model for online review analysis [C] // WSDM 2011: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. New York: ACM, 2011: 815 - 824.

[17] NGUYEN T H, SHIRAI K. Topic modeling based sentiment analysis on social media for stock market prediction [C] // ACL-IJCNLP 2015: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg: ACL, 2015: 1354 - 1364

[18] BANERJEE A, SHAN H. Latent Dirichlet conditional Naive-Bayes models [C] // ICDM2007: Proceedings of the 2007 IEEE International Conference on Data Mining. Washington, DC: IEEE Computer Society, 2007: 421 - 426.

[19] GRIFFITHD T L, STEYVERS M. Finding scientific topics [J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101(Supp1): 5228 - 5235.

[20] DASGUPTA S, NG V. Topic-wise, sentiment-wise, or otherwise? Identifying the hidden dimension for unsupervised text classification [C] // EMNLP 2009: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2009, 2: 580 - 589

[21] LI T, ZHANG Y, SINDHWANI V. A non-negative matrix tri-factorization approach to sentiment classification with lexical prior knowledge [C] // ACL-IJCNLP 2009: Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP. Stroudsburg: ACL, 2009: 244 - 252.

[22] WU Y, ESTER M. FLAME: a probabilistic model combining aspect based opinion mining and collaborative filtering [C] // WSDM 2015: Proceedings of the 8th ACM International Conference on Web Search and Data Mining. New York: ACM, 2015: 199 - 208.

[23] TANG D, QIN B, LIU T, et al. User modeling with neural network for review rating prediction [C] // IJCAI'15: Proceedings of the 24th International Conference on Artificial Intelligence. Menlo Park: AAAI Press, 2015: 1340 - 1346.

[24] 蔡国永,夏彬彬. 基于卷积神经网络的图文融合媒体情感预测[J]. 计算机应用,2016,36(2):428 - 431. (CAI G Y, XIA B B. Multimedia sentiment analysis based on convolutional neural network [J]. Journal of Computer Applications, 2016, 36(2): 428 - 431.)

[25] 徐学可,谭松波,刘悦,等. 面向在线顾客点评的属性依赖情感知识学习[J]. 中文信息学报,2015,29(3):121 - 129. (XU X K, TAN S B, LIU Y, et al. Learning aspect-dependent sentiment knowledge for online customer reviews [J]. Journal of Chinese Information Processing, 2015, 29(3): 121 - 129.)

**Background**

This work is partially supported by the Major Projects of National Social Science Foundation of China (12&2D223), the National Natural Science Foundation of China (61402191, 61300144, 61572223), the Project of State Language Commission (WT125-44), the Self-Determined Research Funds of Central China Normal University (CCNU14A05014, CCNU14A05015).

**SU Ying**, born in 1982, M. S., lecturer. Her research interests include text mining, machine learning.

**ZHANG Yong**, born in 1978, Ph. D., associate professor. His research interests include text mining, natural language processing.

**HU Po**, born in 1980, Ph. D., associate professor. His research interests include text summarization, machine learning.

**TU Xinhui**, born in 1979, Ph. D., associate professor. His research interests include information retrieval, natural language processing, machine learning.