

# 基于主题与三支决策的文本情感分析

王磊 黄河笑 吴兵 郑任儿

(上海开放大学信息与工程学院 上海 200433)

(上海开放大学上海开放远程教育工程技术研究中心 上海 200433)

**摘要** 近年来,情感计算已经成为自然语言处理与人工智能领域的一个研究热点,而文本情感分析是情感计算的一个重要组成部分。提出了一个基于主题特征与三支决策理论相融合的多标记情感分类方法。首先采用基于主题的情感识别模型判断句子的多标记情感类别,在此基础上结合三支决策理论,最终实现对文本篇章的多标记情感分类。实验结果表明,该方法在文本篇章的多标记情感类别识别上取得了令人满意的结果。

**关键词** 三支决策,主题特征,多标签分类,情感计算

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.6.021

## Emotion Analysis of Text Based on Topics and Three-way Decisions

WANG Lei HUANG He-xiao WU Bing ZHENG Ren-er

(School of Information and Engineering, Shanghai Open University, Shanghai 200433, China)

(Shanghai Engineering Research Center of Open Distance Education, Shanghai Open University, Shanghai 200433, China)

**Abstract** Affective computing has received much attention and has been a hot research in the field of natural language processing and artificial intelligence in recent years. Emotion analysis of text is one of important parts in affective computing. A novel method was proposed to analyze the multi-label emotion classification of the text based on topics features and three-way decisions. The multi-label emotions of sentence are judged by using the topic emotion model, and then the multi-label emotions of text are recognized, combining the theory of three-way decision. Experiment results show that the method is reasonable and effective in recognizing the classifications of text emotion.

**Keywords** Three-way decisions, Topic feature, Multi-label classification, Affective computing

## 1 引言

互联网技术日新月异,正迅速改变人们日常交流与沟通的方式。个人博客、微博、产品点评及时事评论等在线信息的不断涌现,使得互联网上产生了许许多多带有个人主观情感的互联网在线信息。这些在线信息大都以文本形式出现,包含用户个人的观点、态度、立场和情绪,反映了人们的喜、怒、哀、乐等情感特征及情感变化。因此,文本不再仅仅局限于描述客观事实,而更侧重于表达个人观点与情感,促进文本情感分析技术的不断发展,使之成为人工智能与自然语言处理领域的一个研究热点。

文本情感分析是通过挖掘与分析文本中的观点、看法、情绪及好恶等主观信息,对文本中词、句和篇章的情感色彩做出判断。文本情感分析是情感计算领域的一个重要组成部分,并对自然语言处理提出了新的挑战。针对文本情感分析对象的不同,文本情感分析可以分为词与短语的情感分析、句的情感分析和篇章的情感分析3个从低到高的不同研究层次<sup>[1]</sup>。在这个层次结构中,词的情感分析是基础,为句与篇章的情感

分析提供依据。本文就是在分析与研究词的情感信息的基础上,推断句的情感类别,并最终判断篇章整体的情感类别。

篇章的情感分析是情感计算中较高级别的表现形式,主要研究方法有两类:监督学习与非监督学习。非监督学习方法是利用文本中词或短语的情感信息来判断整个文本篇章的情感类别。Turkey<sup>[2]</sup>介绍了基于语义倾向的非监督学习方法,即根据词语褒贬含义的倾向信息对评论性文章进行情感分类。监督学习方法采用机器学习方法,采用文本分类的思想将篇章归入不同的情感类别中。Pang<sup>[3]</sup>等人首次将机器学习的方法应用于文本篇章情感分类任务中,并对比了NB、ME和SVM3种分类模型。国内,徐琳宏<sup>[4]</sup>等人提出一种结合语义特征和机器学习的汉语文本极性自动识别机制。徐军<sup>[5]</sup>等人利用朴素贝叶斯和最大熵方法来研究新闻与评论语料的情感分类,通过一系列的实验总结各种方法的优劣。王素格<sup>[6]</sup>则应用粗糙集理论中的数据表示模型,提出了带情感倾向强度的文本向量表示模型,构造了赋权粗糙隶属函数,并将其用于文本的情感类别分类。

本文的研究重点是文本篇章的情感多标记分类问题,通

到稿日期:2014-04-27 返修日期:2014-06-03 本文受上海市科学技术委员会科研计划项目课题“上海开放远程教育工程技术研究中心”(13DZ2252200),国家开放大学2012年科研课题(Q1301F-Y)资助。

王磊(1976-),男,博士生,讲师,CCF学生会员,主要研究方向为粗糙集理论、数据挖掘和机器学习等,E-mail:dragon\_wlei@126.com;黄河笑(1963-),男,博士,教授,主要研究方向为人工智能、多智能体研究、教育技术等;吴兵(1976-),男,博士,副教授,主要研究方向为数据挖掘、推荐系统、Agent系统等;郑任儿(1981-),女,硕士,讲师,主要研究方向为语义网、多媒体技术等。

通过对词的情感特征与主题特征的分析,获得文本中句的情感类别,再进一步融合三支决策理论,从而最终获得文本篇章的多标记情感类别。实验结果表明,基于主题特征与三支决策理论的文本情感多标记分类方法取得了令人满意的效果。

本文第 2 节简要介绍了相关工作,包括 LDA 模型、三支决策理论、Ren\_CECps 中文情感语料库;第 3 节详细阐述了本文提出的基于主题特征与三支决策理论的文本情感多标记分类方法;第 4 节描述了实验过程并对结果进行分析;最后对全文的工作进行了总结。

## 2 相关工作

### 2.1 LDA 模型

2003 年,Blei 等人提出 LDA 模型<sup>[7]</sup>。LDA 模型是一个“词-主题-文本”的层次贝叶斯式模型,在文档与词之间引入主题特征,模型中的参数被看作随机变量,而且参数数量不会随文档数量的增加而增加,因此被广泛应用于大规模语料处理中。

在 LDA 模型中,语料库中的每一篇文档都可表示为若干主题变量所构成的一个概率分布,同时每个主题变量又是若干词或短语所构成的一个概率分布。最初的 LDA 模型针对文档的主题概率分布引入一个超参数并使其服从 Dirichlet 分布,随后 Griffiths 等人针对主题的词的概率分布又引入一个超参数,使其服从 Dirichlet 分布,从而得到一个完整的产生式模型,如图 1 所示。

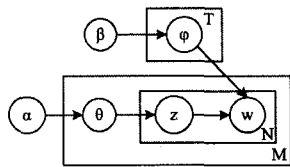


图 1 LDA 图模型

### 2.2 三支决策理论

本文采用三支决策粗糙集模型<sup>[8]</sup>来识别文本篇章的多标记情感类别。与其他已存在的各类分类模型相比,三支决策粗糙集模型不只是简单地增加一个判断类别,将两类分类问题转化为三类分类问题,而是在经典粗糙集理论的基础上,引入一对阈值,将一个集合划分为 3 个两两互不相交的区域,从而引入了第三个选择项,即延迟决策,避免了直接决策带来的风险。

在三支决策粗糙集中,依据最小代价做出不同的决策,通过计算对象分类类别的概率与阈值,将对象划分到该类别相应的正区域、负区域和边界域中,分别对应于决策中的接受决策、拒绝决策和延迟决策。延迟决策对象是位于边界域中的对象,需要进一步地收集信息加以分析,然后根据新的分析结果与实际情况再做出相应决策,即接受还是拒绝。三支决策理论可以有效地处理和分类数据,减少错误决策,从而提高分类的准确率。同时,三支决策理论可以很好地模拟人类解决实际决策问题时的思维模式,可以广泛应用于现实生活的各个领域。

### 2.3 中文情感语料库

本文实验对象采用 Ren\_CECps 中文情感语料库<sup>[9]</sup>,该语料库以中文博客网站爬取的中文博客文章作为初始文本语料,经过处理和标注后,共包含了 1487 篇中文博客文章,共计

11255 个段落,35096 个句子,878164 个词语。

在 Ren\_CECps 中文情感语料库中,所有文本中与情感表达相关联的语言信息均为人工标记,整个文本标注共分为 3 个层次:篇章级、句子级和词语级。词语级的情感标注是整个中文情感语料标注的基础,其标注对象有:词与短语的情感类别及情感强度、词性标注等。句子级的情感标注对象有:句的情感类别及强度、情感主体及情感对象、修辞手法等。篇章级的标注对象有:篇章主题词、篇章中心段落、篇章情感类别及其强度等。

目前,各国学者对于情感类别的划分有着不同认识,还没有形成统一的标准,或者只是简单地将情感划分为褒义与贬义,同时中文与其他语言的情感标注存在较大差别。在 Ren\_CECps 中文情感语料库中,标注者将所有情感分为 8 类最基本的情感类别,分别是:惊讶(surprise)、悲伤(sorrow)、喜爱(love)、高兴(joy)、憎恨(hate)、期待(expect)、焦虑(anxiety)、生气(anger)。篇章、句子与词语的情感类别及强度都被表示为一个 8 维情感向量,表示形式如下:

$$\vec{e} = (e^1, e^2, e^3, e^4, e^5, e^6, e^7, e^8) \quad (1)$$

其中, $e^i$  的取值范围为 0.1 到 1.0,表示 8 类情感类别中一个基本情感类别的情感强度。本文中,如果  $e^i$  的值大于 0,则认为该篇文档具有第  $i$  个基本情感。

## 3 基于主题特征与三支决策理论的文本情感分析方法

作者认为文本篇章的情感由文本中所含句子的情感决定,而每个句子的情感可以由句子中所含词或短语的情感决定。因此,本文的核心思想是基于词的情感信息与主题特征来判断句子的情感类别,在获取句子情感类别的基础上,利用三支决策分类器进一步识别文本篇章的多标记情感类别。

### 3.1 融合主题特征的文本句情感分析

#### 3.1.1 多标记情感主题模型框架

通过对 Ren\_CECps 中文情感语料库的深入研究,发现文本中句子的情感类别与词的主题特征之间有着密不可分的联系。因此利用这种联系,提出一个多标记情感主题模型(MLETM)<sup>[10]</sup>来识别句子的情感类别,该模型如图 2 所示。

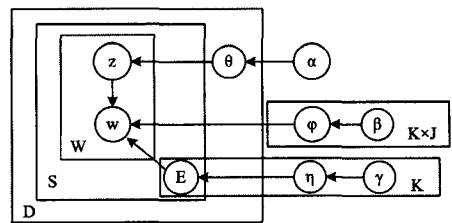


图 2 多标记情感主题模型(MLETM)

在图 2 中,结点代表随机变量,如词结点  $w$ ;有向边描述结点之间的条件依赖关系,如有向边  $z \rightarrow w$ 。整个图模型中包含 3 类变量:分类变量、比例变量和观察变量。 $E, z$  和  $w$  代表结点,属于分类变量。为了识别文本中句子的  $K$  类情感类别,作者定义了  $K$  个二元随机变量  $E_{dsk}$  来表征文本  $d$  中句子  $s$  是否具有第  $k$  类情感。 $w_{dsi}$  作为文本  $d$  中句子  $s$  的第  $i$  个词,它服从一个  $\varphi \rightarrow w$  的随机分布,同时也受到主题  $z$  和情感  $E$  的影响。 $\theta, \eta$  和  $\varphi$  是比例变量,分别表示为  $E, z$  和  $w$  的先验概率。 $\theta_d$  是一个  $J$  维向量,每个  $\theta_{dj}$  表示文本  $d$  中第  $j$  个主题的先验概率。 $\eta$  是一个  $K$  维向量,描述不同情感类别的先

验概率。 $\varphi$ 是一个 $K \times J \times N$ 维向量,描述基于主题特征与情感特征的词的先验概率。 $\alpha, \beta$ 和 $\gamma$ 是3个从训练集中获取的观察变量。

### 3.1.2 多标记情感主题模型概率假设

根据多标记情感主题模型的定义,模型中的有向边描述了随机变量之间的条件依赖关系,依据这些条件依赖关系,作者给出了概率假设。

对于文本中的每个句子,作者定义有 $K$ 个情感分类器 $E_{dsk}$ ,并假设每一个 $E_{dsk}$ 相互独立,它与主题 $z_{di}$ 一起影响词 $w$ 的概率分布。假设词 $w$ 服从随机变量 $\varphi$ 的分类分布,同时条件依赖于情感 $E$ 和主题 $z$ ,公式表示如下:

$$w_{dsi} | E_{dsk}^1, z \sim \text{Categorical}(\varphi_{E_{dsk}^1}^{z_{di}}) \quad (2)$$

$$w_{dsi} | E_{dsk}^0, z \sim \text{Categorical}(\varphi_{E_{dsk}^0}^{z_{di}}) \quad (3)$$

假设主题 $z$ 是条件依赖变量 $\theta$ 的一个分类分布,公式如下:

$$z_{di} \sim \text{Categorical}(\theta_d) \quad (4)$$

由于 $K$ 个情感分类器 $E_{dsk}$ 是相互独立的,因此假设 $E_{dsk}$ 是服从参数 $\eta$ 的一个Bernoulli分布,公式如下:

$$E_{dsk} \sim \text{Bernoulli}(\eta_k) \quad (5)$$

针对随机变量 $\varphi$ ,定义 $\varphi_{kjt}^1$ 和 $\varphi_{kjt}^0$ ,分别描述词是否具有某个情感,它们都是服从参数 $\beta$ 的Dirichlet分布,公式如下:

$$\varphi_{kjt}^1 \sim \text{Dirichlet}(\beta_{kjt}^1) \quad (6)$$

$$\varphi_{kjt}^0 \sim \text{Dirichlet}(\beta_{kjt}^0) \quad (7)$$

$K$ 维随机变量 $\eta$ 是二类情感分类器 $E_{dsk}$ 的先验概率,假设 $\eta$ 服从参数 $\gamma$ 的Bernoulli分布的共轭分布——Beta分布,公式如下:

$$\eta_k \sim \text{Beta}(\gamma_k^1, \gamma_k^0) \quad (8)$$

### 3.1.3 多标记情感主题模型推导

依据图模型理论中许多潜在随机变量用来描述需要预测的潜在特征,根据概率分布假设观察变量,可以推导出这些潜在变量的值。在多标记情感主题模型中,需要预测每个句子的 $K$ 个 $E_{dsk}$ 值,它描述文本 $d$ 中句子 $s$ 具有情感 $k$ 的概率,同时条件依赖于其他变量,推导公式如下:

$$p(E_{dsk} | w, z, E_{-dsk}; \alpha, \beta, \gamma) \propto \frac{n_k^1 + \gamma_k^1}{n_k^0 + n_k^1 + \gamma_k^0 + \gamma_k^1} \times \exp\left(\sum_{i \in W_{ds}} \log \frac{n_{kz_{di}^1}^1 w_{dsi} + \beta_{kz_{di}^1}^1}{\sum_t n_{kz_{di}^1}^1 t + \beta_{kz_{di}^1}^1}\right), E_{dsk} = 1 \quad (9)$$

$$p(E_{dsk} | w, z, E_{-dsk}; \alpha, \beta, \gamma) \propto \frac{n_k^0 + \gamma_k^0}{n_k^0 + n_k^1 + \gamma_k^0 + \gamma_k^1} \times \exp\left(\sum_{i \in W_{ds}} \log \frac{n_{kz_{di}^0}^0 w_{dsi} + \beta_{kz_{di}^0}^0}{\sum_t n_{kz_{di}^0}^0 t + \beta_{kz_{di}^0}^0}\right), E_{dsk} = 0 \quad (10)$$

主题 $z_{di}$ 描述文本 $d$ 中第 $i$ 个词的主题概率分布,公式如下:

$$p(z_{di} | w, z_{-di}, E, \alpha, \beta, \gamma) \propto \frac{n_{dz_{di}} + \alpha_{z_{di}}}{W_d + \alpha^*} \times \prod_{k \in K_d} \frac{n_{kz_{di}}^1 w_{dsi} + \beta_{kz_{di}^1}^1}{\sum_t n_{kz_{di}^1}^1 t + \beta_{kz_{di}^1}^1} \times \prod_{k \in K_d} \frac{n_{kz_{di}}^0 w_{dsi} + \beta_{kz_{di}^0}^0}{\sum_t n_{kz_{di}^0}^0 t + \beta_{kz_{di}^0}^0} \quad (11)$$

### 3.2 基于三支决策理论的文本篇章情感分类

对文本中所有句子的情感进行多标记分类后,文本中的每一个句子都被赋予了一定的情感标记。在每个文本中,每个句子的情感都表示为一个8维情感向量,从而可以统计文本中每类情感出现的次数,这个值作为文本篇章情感三支

策分类器的输入数据。每个文本都可以用一个8维向量表示,向量的长度就是情感类别,每一维的值为对应情感在文本中的标记。

根据三支决策理论的决策规则<sup>[8]</sup>和事先设定的阈值 $\alpha$ 和 $\beta$ ,判断给定的测试文本 $x$ 是否拥有情感 $k$ 的决策过程如下:

(1)如果 $P(k|[x]) \geq \alpha$ ,则文本 $x$ 拥有情感 $k$ ;

(2)如果 $P(k|[x]) \leq \beta$ ,则文本 $x$ 不拥有情感 $k$ ;

(3)如果 $\beta < P(k|[x]) < \alpha$ ,则表示该文本 $x$ 可能拥有情感 $k$ ,也可能不拥有情感 $k$ ,需要进一步处理。

对于文本无法确定是否拥有情感 $k$ ,设定一个阈值 $\theta$ ,并进行如下处理:

(1)如果文本 $x$ 的情感等价类的个数多于等于 $\theta$ ,则判断文本 $x$ 拥有情感 $k$ ;

(2)如果文本 $x$ 的情感等价类的个数少于 $\theta$ ,则判断文本 $x$ 不拥有情感 $k$ 。

### 3.3 基于主题与三支决策相融合的文本情感分析框架

本文中,文本篇章的多标记情感分析框架如图3所示,左侧是训练过程,右侧是测试过程,并将主题特征和三支决策理论应用于篇章的多标记情感分类中。文本篇章的情感类别判断流程共分5步,具体描述如下:

步骤1 从Ren\_CECps中文情感语料库中抽取1000篇文档作为实验数据,从而构成训练语料与测试语料;

步骤2 分别对训练语料与测试语料进行预处理,去除少量不含有任何情感类别的句子,并根据停用词表去除所有句子中的停用词;

步骤3 针对训练数据集,通过训练分别获取多标记情感主题模型(MLETM)和三支决策分类器所需要的所有参数;

步骤4 针对测试数据集,利用多标记情感主题模型生成文本中句子的情感类别;

步骤5 再利用三支决策分类器对文本篇章的多标记情感类别进行判别,并评价实验分类结果。

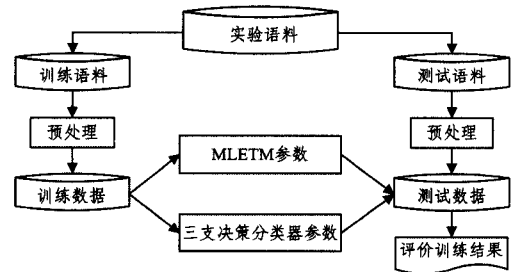


图3 文本篇章的多标记情感分析框架

## 4 实验与分析

### 4.1 实验数据

在实验中,从Ren\_CECps中文情感语料库随机选取了1000篇博客文章,共21225个句子。对数据集进行预处理:1)去除文本中少量的不含有任何情感类别的句子;2)去除所有句子中的停用词。将预处理后的1000篇中文博客文章平均分为10份,采用10折交叉验证。

### 4.2 实验评价标准

本文实验的目的是实现对文本篇章的情感类别进行多标签分类,采用基于标签的评价方法来评价实验结果<sup>[11,12]</sup>。对

于某个单个标签  $k$ , 采用公式  $M(tp_k, tn_k, fp_k, fn_k)$  来评价的分类效果,  $tp_k$  表示正确识别拥有情感标签  $k$  的文本篇数,  $tn_k$  表示正确识别不拥有情感标签  $k$  的文本篇数,  $fp_k$  表示错误识别拥有情感标签  $k$  的文本篇数,  $fn_k$  表示错误识别不拥有情感标签  $k$  的文本篇数。多标签分类的宏平均和微平均公式如下:

$$M_{macro} = \frac{1}{|K|} \sum_{k=1}^{|K|} M(tp_k, fp_k, tn_k, fn_k) \quad (12)$$

$$M_{micro} = M(\sum_{k=1}^{|K|} tp_k, \sum_{k=1}^{|K|} fp_k, \sum_{k=1}^{|K|} tn_k, \sum_{k=1}^{|K|} fn_k) \quad (13)$$

### 4.3 实验结果与分析

本文实验的情感类别为(惊讶、悲伤、喜爱、高兴、憎恨、期待、焦虑、生气)8类, 不考虑中性情感的句子。首先利用多标记情感主题模型(MLETM)来识别文本中句子的情感, 再进一步利用三支决策分类器对文本篇章的情感进行分类。三支决策分类器中参数设置如下:  $\alpha=0.6, \beta=0.3, \theta=12$ , 以上参数均为从训练数据集中获得的实验经验值。

对比朴素贝叶斯方法、单一主题方法和主题与三支决策融合方法, 在测试数据上进行文本篇章的情感类别多标签分类实验。表1给出3类方法的宏平均值与微平均值。

表1 3类不同方法的多标签分类结果的宏平均值与微平均值

	贝叶斯方法	单一主题方法	主题与三支决策相融合方法
宏平均精确率	0.655	0.719	0.722
宏平均正确率	0.521	0.745	0.754
微平均精确率	0.655	0.718	0.732
微平均正确率	0.554	0.776	0.796

表1的实验结果充分显示出基于主题与三支决策相融合方法在文本篇章的多标记情感识别中的优势。

针对基于主题与三支决策相融合的方法, 8类基本情感类别单标签情感极性判断的正确率如图4所示。

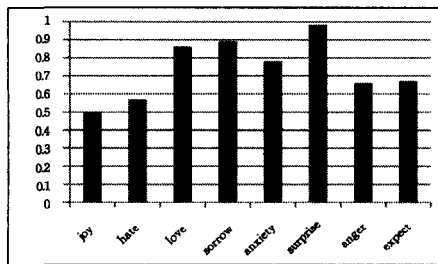


图4 8类基本情感类别单标签情感极性判断正确率

图4反映出高兴情感的类别判断的正确率最低, 是因为在语料库中很多文本在拥有其他高强度情感的同时, 都或多或少地拥有一些高兴情感, 从而导致在这些文本中识别高兴情感比较困难。憎恨情感的类别判断中的正确率也较低, 是因为语料库中收集的拥有憎恨情感的文本的数据较少, 从而导致对模型训练不够充分, 影响了这类文本数据在情感类别判断中的正确率, 今后将进一步完善 Ren\_CECps 中文情感语料库。

**结束语** 本文对文本篇章的多标记情感分类问题进行了深入研究, 提出由词到句再到篇的多标记情感类别识别路线, 并采用主题特征与三支决策相融合的情感分类方法。以 Ren\_CECps 中文情感语料库为实验对象, 依据改进的 LDA 模型

初步判断句的情感类别, 利用篇章中所含句子的情感特征, 并结合三支决策理论, 实现对语料库中文本篇章所具有的惊讶、悲伤、喜爱、高兴、憎恨、期待、焦虑、生气8类情感类别的多标记分类, 实验取得了令人满意的结果。

本文只是利用主题特征与三支决策理论来研究文本篇章的多标记情感分类, 目前的研究还有许多待提高和改进之处。如何更好地识别词和句的情感类别, 分析篇章中的情感主题与情感对象等问题, 也是今后进一步研究的方向。

### 参考文献

- [1] 赵妍妍, 秦兵, 刘挺. 文本情感分析[J]. 软件学报, 2010, 21(8): 1834-1848  
Zhao Y Y, Qin B, Liu T. Sentiment analysis[J]. Journal of Software, 2010, 21(8): 1834-1848
- [2] Turney P D. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews[C]// Proceedings of the 40<sup>th</sup> Annual Meeting of the Association for Computational Linguistics. 2002: 417-424
- [3] Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques[C]// Isabelle P, ed. Proc. of the EMNLP 2002. Morristown: ACL, 2002: 79-86
- [4] 徐琳宏, 林鸿飞, 杨志豪. 基于语义理解的文本倾向性识别机制[J]. 中文信息学报, 2007, 21(1): 96-100  
Xu Lin-hong, Lin Hong-fei, Yang Zhi-hao. Text Orientation Identification Based on Semantic Comprehension[J]. Journal of Chinese Information Processing, 2007, 21(1): 96-100
- [5] 徐军, 丁字新, 王晓龙. 使用机器学习方法进行新闻的情感自动分类[J]. 中文信息学报, 2007, 21(6): 95-100  
Xu Jun, Ding Yu-xin, Wang Xiao-long. Sentiment Classification for Chinese News Using Machine Learning Methods[J]. Journal of Chinese Information Processing, 2007, 21(6): 95-100
- [6] 王素格. 基于 Web 的评论文本情感分类问题研究[D]. 上海: 上海大学, 2008  
Wang Su-ge. A Study on Emotional Comment Text Classification Based on Web[D]. Shanghai: Shanghai University, 2008
- [7] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation[J]. Journal of Machine Learning Research, 2003: 993-1022
- [8] Yao Y Y. An outline of a theory of three-way decisions[C]// Proceeding of the RSCCTC 2012, LNCS(LNAI). 2012, 7413: 1-17
- [9] 任福继, 等. Document for Ren-CECps 1. 0[OL]. <http://al-www.is.tokushima-u.ac.jp/member/ren/Ren-CECps1.0/Ren-CECps1.0.html>, 2009
- [10] Wang Lei, Ren Fu-ji, Miao Duo-qian. A Novel Method for Recognizing Emotions of Weblog Sentences[C]// Proceeding of 2013 IEEE/SICE International Symposium on System Integration. 2013: 358-363
- [11] Tsoumakas G, Katakis I. Multi-Label Classification: An Overview[J]. International Journal of Data Warehousing and Mining, 2007, 3(3): 1-13
- [12] Tsoumakas G, Vlahavas I. Random K-Labelsets: An Ensemble Method for Multilabel Classification[C]// Proceedings of the 18th European Conference on Machine Learning (ECML2007). Warsaw, Poland, 2007: 406-417