

《通用计算机科学杂志》,第 18 卷,第 8 期(2012 年),973-992
提交:22/9/11,接受:14/12/11,出现:28/4/12 J.UCS

通过 Twitter 发现消费者洞察 情绪分析

维拉斯·查姆勒特瓦特、帕塔拉西尼·巴塔拉科索尔、蒂帕科恩·伦卡西里
(泰国曼谷朱拉隆功大学 wilas@chamlertwat.in.th、
pattarasinee.b@chula.ac.th、tippakorn@acc.chula.ac.th)

楚查特·哈鲁猜亚萨
(国家电子和计算机技术中心,泰国巴吞他尼 choochart.haruechaiyasak@nectec.or.th)

摘要:研究消费者行为的传统方法,例如营销调查和焦点小组,需要大量的时间和资源。此外,有些产品,例如

智能手机等产品生命周期较短。作为一种替代解决方案,我们提出了一个基于情绪分析的微博情绪分析系统(MSAS),用于自动分析来自 Twitter 微博服务的客户意见。MSAS 包含五个主要功能:(1) 收集 Twitter 帖子,(2) 过滤主观帖子,(3) 检测每篇帖子的极性,(4) 对产品功能进行分类,(5) 总结和可视化总体结果。我们使用智能手机产品领域作为案例研究。对 100,000 条与智能手机相关的收集帖子进行的实验表明,该系统有助于表明客户对产品功能(如应用程序、屏幕和相机)的情绪。智能手机行业专家的进一步评估证实,该系统产生了一些有效的结果。

关键词:微博,情感分析,自然语言处理,信息提取,信息可视化类别:H.3.5,H.4.3,M.0

1 介绍

过去几年,我们目睹了社交媒体的蓬勃发展。人们通过众多社交网络应用在互联网上构建了一个全球通信网络。社交网络上每天都会产生大量前所未有的内容,尤其是在微博服务上 [Jung, 08]。微博在轻量级通信方面的使用越来越流行,这提高了其作为新通信媒介的潜力 [Jung, 09]。Twitter 是最受欢迎的微博服务之一,它声称每天有超过 2 亿条推文或帖子 [Twitter blog, 11]。这是一场内容生成和分发方式的革命,即通过不受控制地创建、共享和发现消息。Facebook 创始人马克·扎克伯格 (Mark Zuckerberg) 表示,这种行为已成为一种新的社会规范 [Cashmore, 10]。

作为消费者,大多数互联网用户已经将社交媒体融入日常生活。在任何用户设备和界面上分享任何信息都变得更加容易和便捷 [Jung, 10a]。此外,我们的沟通方式已经

第 974 章

改变了。消费者不会单方面相信生产者的信息,但人们需要对话来向他人传达他们的信息。显然,Twitter 极大地降低了创建内容的门槛,因此分享日常生活变得毫不费力。人们可以随时随地更新自己的状态并与网络中的朋友和关注者分享自己的观点[Jung,10b]。

民营企业作为生产者,也广泛采用社交媒体作为营销策略,将自己的产品或服务商业化,不仅要发声,还要用自己的话来倾听顾客的真实声音。尽管大多数人对回答有关产品或服务偏好的调查犹豫不决,但他们在社交网络中表达自己的想法,并在塑造其他消费者的意见方面发挥着巨大的影响力[Zabin,08]。微博是获取这些有用信息的丰富资源。它包含对市场上许多不同主题积极或消极的许多用户情绪表达[Jung, 11a]。

与传统的网站或博客相比,微博是一种社交网络应用,可以产生巨大的用户意见。与短期技术趋势不同,微博是许多人所关注的迷你博客的演变。由于微博中的交流简短而随意,人们可以实时更新彼此的活动[Jung,11b]。它快速、即时地满足了当今用户的行为。因此,在微博中可以很容易地找到对产品或服务的赞赏。这些信息可能会导致消费者做出购买或忽视的决定。微博影响力的规模和速度对研究人员和从业者提出了挑战,他们需要了解其背后的科学原理并设计营销传播的应用程序。

对大规模社交数据的分析可以从消费者的意见中证实企业的观点,对于支持高层管理人员解决现实问题至关重要。这些消费者的声音可以影响品牌认知、品牌忠诚度和品牌宣传。通过社交媒体监控和情绪分析,企业将能够利用消费者洞察来提高产品质量、提供更好的服务,甚至识别新的商机和其他相应的活动。

在智能手机市场,技术的进步不断超越,更新版本以闪电般的速度涌入市场。

根据 HTC 的数据,智能手机的平均保质期已从 2007 年的三年缩短至 2011 年的六至九个月左右[Ferreira, 11]。生产商没有足够的时间通过传统方式研究市场。参考技术采用生命周期,每种智能手机型号都只有有限的时间来证明其产品概念。我们需要一种替代方法来检索改进下一代产品的必要信息[Jung, 12]。

由于许多人在 Twitter 上交流有关智能手机的信息,因此情绪分析可能是发现消费者洞察的合适技术。例如,Twitter 帖子 #1 说:“哦,我喜欢我的新 HTC Sensation 中的‘语音搜索’功能。省去了我输入长句子的麻烦。”Twitter 帖子 #2 说:“我讨厌 htc peep! ! !UberSocial 是最好的,因为它更快。”这个例子可以提示 HTC 保留一些功能,同时终止或改进一些功能。

本研究将尝试调查微博情感分析的结果对于理解某些行业的消费者洞察是否有效,例如

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 975

智能手机市场,或者不是。我们提出了微博情绪分析系统

(MSAS)来完成任务。将应用机器学习和基于词典的技术来获得适当的结果。MSAS 有助于揭示客户对智能手机功能的正面或负面看法。结果以一些可视化格式进行总结和显示。MSAS 的最终结果由智能手机行业的三位专家进行评估。

本文的其余部分安排如下。在下一节中,我们将回顾消费者洞察和情绪分析方面的一些相关研究工作。在第3节中,我们将详细介绍我们提出的系统MSAS。第4节介绍了使用大量Twitter帖子对系统进行分析的结果。第5节总结了本文,并介绍了未来工作的一些注意事项。

2 相关作品

2.1 消费者洞察

消费者洞察是关于消费者是谁以及他们的想法或感受的研究 [Stone, 04]。这可能是从直接营销到数据库营销再到客户关系管理或客户体验管理的革命。企业可以利用这种理解来支持从改变营销策略到改善运营和互动。良好的消费者洞察是良好客户关系管理的基础。因此,企业将能够按照自己的意愿定位自己,以满足客户或利益相关者的需求。

洞察并非消费者有意识的行为或想法,大多数都受到各种外部因素的影响,从经济和社会状况到品牌营销方式。企业可以从投诉或表扬或通过要求进一步的信息来收集这些信息。有些任务可以通过市场研究来实现,而有些则需要从客户数据库、销售和客服人员的反馈中挖掘。

尽管衡量客户满意度变得非常重要,但根据消费者的意见使用信息来设定员工目标可能会扼杀企业的创造力并限制创新。在大型企业中,测量结果会因降低成本和增强控制而产生误导。这个焦点已经变成了内部的;导致员工和整个组织不再首先考虑客户。结果,他们的客户忠诚度和整体满意度下降。消费者洞察过程有助于改变我们对消费者、员工和组织的看法,而不是支持一种适得其反的思考方式。

用于挖掘大量文本数据以获取营销情报的应用程序可分为以下三种类型 [Glance,05]:

- 早期警报 - 当发生罕见但严重甚至致命的情况时通知订阅者。

- 嗡嗡声跟踪 - 跟踪讨论主题的趋势并了解正在形成的新主题。
- 情绪挖掘 - 提取积极与消极的总体衡量标准

观点。

意见观察者是一个早期的框架,用于分析和比较市场上竞争产品的消费者意见[Liu,05]。该系统提供了一个单一的

976 Chamlerwat W.,Bhattarakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现……

可视化可以一目了然地展示产品。用户可以清楚地看到消费者心目中每种产品在各种产品特征方面的优势和劣势。对于产品制造商来说,这种比较使其能够轻松收集营销情报和产品基准信息。第二,一种新的

提出了基于语言模式挖掘的技术,从特定类型的评论中的优点和缺点中提取产品特征。

2.2 情感分析

近年来,许多研究人员对情感分析或所谓的意见挖掘进行了研究。情感分析是一种对自然语言文本的计算研究,旨在识别情感极性、强度以及这些情感所适用的主题 [Liu, 11]。情感分析源于对处理大量数据的自动意见披露和摘要系统的需求,以允许机器理解人类生成的内容 [Lake, 11]。

在商业领域,该技术已用于挖掘品牌监控、民意调查、金融交易、营销和许多现实世界问题的意见。

传统上,企业可能会进行消费者调查来调查客户的需求。调查期间出现了一些偏见。偏差问题被称为布拉德利效应 [Bobo, 09]。人们不愿意提供反映不受欢迎的态度或观点的准确答案。尽管研究人员提出了精心设计的调查,可以提供质量评估,但它们可能成本高昂且耗时。情绪分析是通过不明确向人们提出任何问题来无偏见地分析现有数据的另一种选择。这种方法

情绪分析比传统的调查回答更能反映人们的真实想法。然而,抽样人群存在一个缺点,即可能无法确定目标响应者群体。因此,情绪分析不能完全取代传统方法,但它可以作为一种补充解决方案。

从技术上讲,情绪分析是自然语言处理 (NLP) 研究的一部分。主观性分类和情绪分类可能是该领域研究最广泛的主题。主观性分类是将主观句子与客观句子分开或将观点与事实区分开的过程,而情绪分类是确定情绪取向的过程,即该句子表达的是积极还是消极的感觉。此外,一些研究对确定情绪极性的强度(强度)以衡量语义强度感兴趣。基于特征的情绪分析是一项深入研究,指的是确定实体不同特征上表达的情绪。例如,基于特征的智能手机屏幕情绪分析是研究人们在屏幕上的表达是积极的还是消极的。

人们在多个不同文本粒度级别上研究了各种情绪分析方法 [Agarwal, 11]。粗粒度级别从文档级别分类任务 [Turney, 02] 开始,到句子的细粒度级别 [Hu, 04] 和短语级别 [Wilson, 05]。从根本上讲,用于情绪分析的方法可以分为两类:机器学习方法和基于词典的方法。

1)机器学习方法是监督学习方法。它被称为训练过程,教导代理将输入分类为输出。一旦足够

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 977

通过算法学习标注有情感值的训练数据,对相应领域数据的情感分析将提供有希望的结果[Pang, 02]。

我们简要描述三种流行的分类算法:朴素贝叶斯(NB)是一种基于应用贝叶斯定理的简单概率分类器,特别适合输入维度较高的情况。

- 最大熵(MaxEnt)或多项Logit模型通常用作朴素贝叶斯的替代模型。特别是,在朴素贝叶斯分类器中学习是计算特征和类别同时出现的数量的简单问题,而在最大熵分类器中,通常使用最大后验(MAP)估计来最大化权重,必须使用以下方法来学习迭代过程。支持向量机(SVM)被认为是性能最佳的机器之一。它是分析数据和识别模式的方法,用于分类和回归分析。SVM的本质是找到一个超平面,将文档向量尽可能地从一类分离到另一类。

2)基于词典的方法通常是无监督的方法。它是基于预定情感词典分数提供的特征来估计极性(无论是积极的还是消极的)的规则。这些方法可以在没有任何参考语料库和事先训练的情况下发挥作用。情感词典通常以不考虑领域信息的方式正确估计术语的通用极性[Zhe,10]。

意见词是用来表达积极或消极情绪的词。表达良好感觉的单词(例如 great 或 Excellent)具有正极性,而表达不良感觉的单词(例如 bad 或 awesome)具有负极性。用于识别情绪及其倾向的意见词词典称为“意见词典”。一些有趣的意见词典是 SentiWordNet 和 Pageranking WordNet。

SentiWordNet 是公共语言资源之一,它为 WordNet 的每个同义词集或同义词集分配三个数值分数:积极性、消极性、客观性 [Esuli,07b]。Pageranking WordNet 是随机游走模型在确定 WordNet 同义词集时的一种应用,即根据它们拥有给定语义属性的强度来确定 [Esuli, 07a]。

微博情感分析最近受到学者的关注。微博包含在非常有限的空间内表达积极或消极感受的重要信息,例如,“htc hd2 很棒,1GHz 处理器真的很有帮助。”

然而,这并不总是像识别“我喜欢 Android”或“他讨厌 iPhone”那么容易。今天人们所说的语言很复杂,充满了俚语、歧义、讽刺、成语和反讽,如以下示例所示:“HTC 电池 15 分钟内耗尽,但充电需要一年。我好幸运!”。

关于这个问题的早期研究可能是使用远程监督的 Twitter 情感分类 [Go, 09]。这项工作表明,当使用表情数据进行训练时,机器学习算法的准确率超过 80%,并且 SVM 优于其他分类,而一元模型优于二元模型和词性 (POS)

978 Chamlerwat W.,Bhattarakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现……

特征。此外,有一项研究证实,POS特征对于微博领域的情绪分析毫无用处[Kouloumpis,11]。

后来有趣的工作是基于词典和基于机器学习的方法的结合。在本研究中,我们使用基于词典的方法来执行实体级情感分析,该方法可以提供高精度但召回率较低。然后,为了提高召回率和 F 分数,还应用基于机器学习的方法通过利用基于词典的方法结果中的信息来自动识别观点 [Zhang,11]。

3 建议的解决方案

在本文中,我们提出了一种基于情感分析的方法来确定微博帖子是正面情绪还是负面情绪。我们用 Twitter 来代表微博服务。我们的解决方案中应用了基于机器学习和基于词典的方法。在本节中,我们首先概述 Twitter 的特征。然后,将详细介绍所提出的方法。

3.1 Twitter 特点

Twitter 由 Twitter Inc. 拥有和运营,是Friendfeed、Tumblr和identi.ca等现有同类服务中最受欢迎的微博客服务。

Twitter 用户可以在其用户个人资料上发布短消息(称为推文),并在按时间顺序反向排列(称为时间线)的单个列表上阅读其他人的消息。推文是基于文本的帖子,长度限制为 140 个 UTF-8 字符,内容涉及用户日常生活中发生的小事的任何更新。更新的短暂性使用户可以实时快速发布,立即到达受众。默认情况下,推文是公开可见的,但所有者可以将隐私设置为仅向其朋友显示。用户之间的关系,或者所谓的关注,是不对称的。用户可以关注其他人并查看他们的推文,但其他用户不需要回报。订阅者被称为追随者。当两个订阅者互相关注时,他们就是朋友。

用户可以通过网站twitter.com和许多第三方应用程序(包括基于 Web 的应用程序、桌面客户端和移动电话)直接与 Twitter 互动。由于 Twitter 提供应用程序编程接口(API)以允许与其他服务集成,因此用户可以通过短信服务(SMS)、即时消息(IM)、简易信息聚合(RSS)、电子邮件和许多其他工具(包括代表用户通过自动代理发布的帖子)直接向关注者发送消息。尽管来自世界各地的开发人员,Twitter 周围的生态系统仍然非常广泛。

根据每条消息字符数限制,Twitter 采用一系列惯例,允许用户在推文中添加结构。

1) 地址 要回

复或提及,用户可以在用户名前面使用 @ 符号(也称为 @reply)在推文中引用其他用户。例如,“@alice 太棒了,恭喜!”此操作在用户之间建立链接,并允许线程

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 979

推特是用户之间对话的桥梁。此外,用户还可以使用字母d加用户名来向特定用户发送私人消息。2)主题标签 为了自由地将推文归类,用户可以在单词前使用主题标签或#符号进行标记。例如,“等不及新款#iphone4了”。使用主题标签的做法可能源自HTML

锚点或计算机编程

中在专门词汇前加上标点符号[Boyd, 10]。3)外部链接 要在狭窄的空间内分享链接,用户可以使用URL缩短服务来生成一个唯一的缩写URL,该URL会重定向到所需的网站。例如,http://bit.ly/17zKu来自http://en.wikipedia.org/wiki/Microblogging。自2010年3月起,Twitter为在其上发布的链接提供了链接缩短服务t.co,以保护用户免受恶意网站的侵害,并跟踪推文内链接的点击情况。

4) 转发 要

在 Twitter 上转发其他用户的消息,用户可以通过两种方式进行。

首先,通过复制帖子并在其前面添加 RT @用户名来应用传统的转发(RT)。例如,Alice 发布“新年快乐”,然后 Bob 转发“RT @Alice:新年快乐”。其次,通过单击网络中的转发链接,新的 Twitter 转发要容易得多,然后该推文将被转发给所有用户的关注者。

几项研究试图分析 Twitter 的使用情况。用户意图的主要类型是日常聊天、对话、分享信息和报道新闻,而用户的主要类别是信息源、朋友和信息寻求者[Java,07]。微博是一种补充其他互动形式的对话机制[Zhao,09]。客户可以利用微博作为

口碑相传和讨论的工具,而企业可以将微博作为其整体营销战略的一部分 [Jansen, 09]。

3.2 方法论

产品评论中的一种情绪分析形式是生成基于特征的摘要 [Hu, 04]。产品的特征包括产品属性、组件以及产品的其他方面,例如尺寸、功率和显示。在本文中,我们提出了一种称为微博情绪分析系统(MSAS) 的解决方案来分析 Twitter 上的情绪。我们提出的系统 MSAS 可以确定对产品特征的正面或负面情绪;然后汇总结果以生成用户摘要。

如图 1 所示,MSAS 流程可分为两个阶段:准备阶段和分析阶段。

准备阶段是将所有必需的数据(包括帖子、模型、相关词汇)按如下方式排列在特定领域上的阶段。

980 Chamlertwat W.,Bhattarakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现……

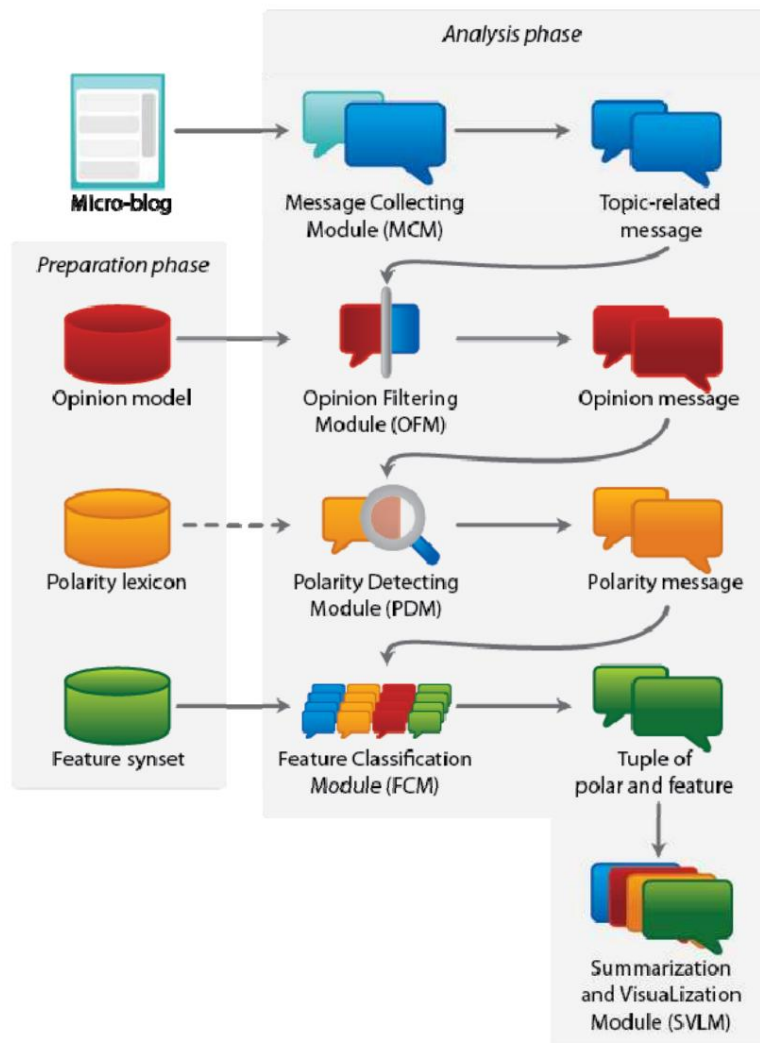


图1:微博情感分析系统 (MSAS)

1)微博:这是系统的输入源。如前所述, Twitter 可以成为微博客的一个很好的代表,因为它提供了大量的数据,这些数据可以作为公共时间线轻松收集。如表 1 所示,我们使用智能手机领域作为案例研究。

Chamlertwat W.,Bhattarakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 981

编号	创建于	消息状态 ID 22285927850	文本
1	2010-08-28 00:49:42		我很喜欢这款三星 Galaxy,但如果你买了它,不要马上插上电源来充满电。而是等电池耗尽...
2	2010-08-27 22:05:01	22274155397	Motorola Milestone 拥有“世界上最好的手机显示屏”。参考:displaymate.com 我用三个摄像头都无法拍摄特写物体。但我的 iPhone 4 可以。这有多差劲?
3	2010-08-28 02:32:26	22292491844	

表1:数据库中存储的收集推文示例

2)意见模型:该模型是基于分类算法创建的。该模型用于过滤有观点的帖子和无观点的帖子。表2显示

一些被归类为观点帖子和非观点帖子的推文示例。将使用极性词典进一步分析每条推文的情绪。

类型	情绪	信息
观点	积极的	哇,太棒了!享受定制吧。:) Nexus one 可能是定制之父,有多种类型的 rom 可供选择
	负面	我的 Nexus One 对移动网络信号变化的位置变化响应不佳。需要一个应用程序来手动刷新网络。
非观点	不适用	寻找 Verizon 新发布的摩托罗拉手机 Devour 以及 HTC Hero 即将上市。 AT&T 终于推出 Android 操作系统了!!!

表 2:观点和观点推文示例

3)极性词典:这是用于基于词典的情感分析方法的词典。我们从 SentiWordNet 3.0 [Baccianella, 10] 中获取了我们的观点词典。

但是,我们丢弃了客观术语,因为它不包含正面或负面的极性。表 3 显示了 SentiWordNet 3.0 中的一些术语示例。在此表中,POS 指词性;对 (POS, ID) 唯一地标识了 WordNet (3.0) 同义词集。值 PosScore 和 NegScore 是积极性和消极性分数。

SynsetTerms 是一组同义术语,其含义编号写在# 符号之后。最后,光泽是这个词的意思。

982 Chamlertwat W.,Bhattarakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现……

销售点编号	PosScore	NegScore	同义词集术语	光泽度
01586866	0.75	0	愉快#2	(指人)举止或行为令人愉悦;“我不喜欢它,可能不是一个令人愉快的人”
01587077	0	0.875	令人讨厌的#1 可怕的#3	冒犯性的甚至 (对人)恶意的;“心情不好”;“一次严重的事故”;“令人厌恶的震惊”;“难闻的气味”;“令人讨厌的伎俩”;“他会在我的葬礼上说脏话吗?” 埃兹拉·庞德

表3:SentiWordNet v3.0的数据格式

4)特征同义词集:这是公司想要研究的多个产品或服务特征的集合。由于不同的客户对每个功能的引用可能不同,因此我们必须通过研究产品将它们分为功能类别

所有制造商网站上的规范。表 4 显示了智能手机功能类别和功能术语的一些示例。

类别	特征
网络	2g、3g、连接、wifi、无线、蜂窝
屏幕	显示器、ppi、像素、电容、分辨率、监视器、液晶屏
容量	容量、磁盘、存储、ram、rom、内存

表 4:智能手机功能类别示例

一旦我们准备好了一切,分析阶段就是由以下四个相关模块组成的检查阶段。

1)消息收集模块 (MCM) :该模块利用 Twitter 搜索 API 来收集我们想要研究的推文。
Twitter 有一个应用程序编程接口 (API),用于通过查询词以编程方式访问推文 [Twitter 网站,11]。搜索结果是匿名的,无需身份验证。在我们收集数据时,Twitter API 已经

任何请求的响应中最多只能包含 100 条推文。数据可以以 XML 或 JSON 格式检索。对于智能手机,我们通过关注相关关键词来使用查询词,例如 iphone、nexus、htc、motorola、nokia、blackberry 和 palm。

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 983

2) 观点过滤模型 (OFM):该模块使用机

器学习方法将给定的推文分类为“主观或观点”和“客观或非观点”。首先,我们开始收集训练数据并手动将推文标记为观点或非观点。然后,我们执行以下任务:特征选择和分类模型构建。在我们之前的研究中,我们发现过滤观点推文的最佳性能是使用支持向量机 (SVM) 和信息增益 (IG) 特征选择。该模型的准确率达到 84.5% [Chamlertwat,11]。

3) 极性检测模块(PDM):该模块确定给定消

息的极性是正还是负。它还可以表明情感词的强度。该模块计算

通过平均推文中找到的每个情感词的分数的来计算每条推文的极性分数。我们按如下方式预处理每个 Twitter 帖子:a) 标记化 – 我们通过观察单词分隔符 (例如空格和标点符号)来分段

每个帖子。例如,帖子“刚刚意识到 iPhone 上的屏幕与我的 Android 手机大小相同”被标记为 16 个单词,“Just”、“realized”、“that”、“the”、“screen”、“on”、“the”、“iPhone”、“is”、“the”、“相同”、“size”、“as”、“my”、“android”、“phone”。所有标记化的单词形成一个词袋。b) 停用词删除 我们通过查找停用词词典从词袋中删除停用词。停用词删除有助于通过清理无意义的术语来减少术语维度。c) 链接删除 我们删除地址和外部链接,因为它们对于检测极性和特征没有用处。d) 术语规范化 我们将动词和形容词的形式更改回基本形式,并用其含义替换缩写。例如,动词“talking”被替换为“talk”。e) 俚语处理 我们通过检测三个或更多字符的序列来替换重复字符的序列,例如,单词“gooooooooood”被简化为“good”

然后,我们通过应用以下规则对每条推文进行评分:(1)

初始化总极性分数: $s \leftarrow 0$ (2)使用

SentiWordNet 检查每个标记,如果标记是正

面的,则 $s \leftarrow s + w$;其中 w 是情绪词的分数的,如果标记是负面的,则 $s \leftarrow s - w$

(3) 如果tweet包含否定,则 $s \leftarrow -s$;其中否定是“不”或“不”一词,它将扭曲整体推文的极性

4)特征分类模块 (FCM):该模块负责概述每条推文的

品牌或产品特征。例如,“Twitterific 非常棒。我所有的推文都是通过 iPhone 进行的”。这条推文分类如下:

twitterific,发推文 → 功能漂亮 → 情绪

(得分 = 0.20833333333333334) 很棒 → 情绪 (得分 =

0.10693039146587664) iPhone → 品牌

5)总结和可视化模块 (SVLM):

984 Chamlerwat W.,Bhattarakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现……

在最后一个过程中,系统将通过基于特征的意见摘要的可视化,提供人们发布的每个产品功能的情感极性概述。我们采用以下三种不同的可视化技术。a) 雷达图 在一张图中比较一个有趣问题的许多属性。

显然,图表覆盖的面积越大,整体值就越大。b) 条形图 – 突出显示单独的数量,尤其是几类数据之间的数量差异。c) 折线图 – 显示值的变化模式。此外,我们可以在同一张图表中使用多条线来显示多条数据。

4 实证评估

4.1 数据集

在本研究中,实验是在智能手机领域进行的,因为 Twitter 上有大量与智能手机相关的对话和想法。智能手机是一种具有广泛功能的移动电话,例如个人数字助理、互联网浏览器和其他应用程序。随着技术进步以可承受的价格提高了智能手机的功能,全球智能手机市场正在不断增长。智能手机成为许多人日常生活的一部分。因此,我们在实验中捕获了与人们如何谈论智能手机相关的推文。通过 Twitter Search API 收集了 2010 年 3 月至 2010 年 6 月期间公共时间线上的所有推文,并使用以下标准进行初步筛选。

- (1) 推文必须包含与智能手机品牌相关的相关关键词如表5所示。
- (2)重复的内容将被忽略。
- (3) 非英语推文也被过滤掉。
- (4) 包含多个产品的推文将不被考虑,因为它可能是比较句,超出了我们的范围。

操作系统	品牌	关键词
苹果iOS	iPhone	IPhone
谷歌安卓	摩托罗拉、HTC、三星、谷歌 Nexus	摩托罗拉里程碑、摩托罗拉 droid、摩托罗拉 cliq、摩托罗拉 devour、摩托罗拉 backflip、htc、三星 galaxy、三星 omnia、三星 wave、 联系一号
RIM 黑莓	黑莓	bb 粗体、bb 火炬、bb 曲线、黑莓
塞班操作系统	诺基亚	诺基亚 n、诺基亚 e、诺基亚 c
掌上操作系统	棕榈	palm pre、palm pixi

表 5:智能手机的相关术语和关键词

4.2 处理

如图 1 所示,我们使用消息收集模块 (MCM) 收集了大约 1,000,000 条推文。然而,我们随机选择了 100,000 条推文进行实验,以减少处理时间并确保每个品牌沿时间线均匀分布。每个收集时间段的推文集合均等分为 10 组。因此,从每组中随机选择 10,000 条推文放入意见过滤模块 (OFM)。为了构建分类模型,我们手动注释了 600 条随机选择的推文,其中一半包含观点和非观点推文。根据我们之前的研究,我们发现对于意见过滤模型,信息增益 (IG) 特征选择和分类算法支持向量机 (SVM) 产生了最佳性能 [Chamlertwat, 11]。使用 OFM 处理 100,000 条推文后,我们得到了大约 20,000 条意见推文的结果。然后,这些推文被发送到极性检测模块 (PDM),以分析和分配每条推文的极性分数。

特征分类模块 (FCM)会进一步提取每条推文中提到的产品特征和品牌,最后通过摘要和可视化模块 (SVLM)将结果展示出来。

4.3 结果

表6的实验结果表明,在研究期间,iPhone是人们谈论最多的品牌,其次是黑莓和HTC。

但经过PDM处理后,排名略有变化,由iPhone变为Blackberry,Nexus,在微博上很少看到关于Palm和Nokia的讨论。

	MCM	正文频分复用		项目数据管理	
黑莓	100,000	4,105	20.63%	548	13.72%
宏达电		2,864	14.39%	428	10.71%
iPhone		8,091	40.65%	2,231	55.84%
摩托罗拉		990	4.97%	157	3.98%
关系		2,305	11.58%	458	11.46%
诺基亚		370	1.86%	48	1.20%
棕榈		73	0.37%	16	0.40%
三星		1,105	5.55%	109	2.78%
合计	100,000	19,903		3,995	人

表 6:推特帖子的数量和百分比

如表 7 所示,人们更喜欢谈论功能、应用程序、相机和电源。令人惊讶的是,称赞的推文数量超过了抱怨的数量。对于黑莓,每个功能的正面和负面推文几乎相等;除了应用程序的正面推文多于负面推文。对于 HTC,每个功能的正面推文数量都多于负面推文。对于 iPhone,最受欢迎的功能是应用程序、网络 and

第 986章

相机。然而,他们对配件有一些担忧。对于摩托罗拉来说,积极的推文是关于相机和电源的。与 HTC 类似,Nexus 对除电源之外的所有功能都有良好的感知。对于诺基亚来说,他们的网络很强大。

对于Palm来说,数据太少,无法说它是好是坏。最后,对于三星来说,它的相机非常出色。

	黑莓 HTC iPhone		摩托罗拉					
	+	-	+	-	+	-	+	-
屏幕	1	0	1	0	4	1	0	0
应用 150 106 43 20 1071 473							3	3
网络 20 21 52 15					104 69 18 12			
系统	0	0	21 1		7	1	12 15	
相机 66 60 138 25 176 73 36								7
容量	2	4	1	1	4	3	0	1
电源53		51 65 39 101 123 35 12						
传感器	1	0	0	0	0	0		0
配件	9	2	3	0	6 14 2			1
尺寸	2	0	3	0	0	0		0
	关系		诺基亚		棕榈		三星	
	+	-	+	-	+	-	+	-
屏幕	0	1	0	0	0	0	7	0
应用	三十七	12	2	4	7	1	10 10	
网络 33		14 17		5	1	0	3	2
系统	4	0	0	0	0	0		0
相机 175 139			6	3	2	0	62	7
容量	1	0	1	1	0	0		1
电源16		19	4	0	4	0	4	1
传感器	1	0	0	0	0	0		0
配件	1	3	5	0	1	0		2
尺寸	0	2	0	0	0	0		0

表 7:FCM 中各品牌的情绪推文数量

总结结果,SVLM 过程将通过三种不同的技术将这些数字可视化。

雷达图 - 比较一个品牌的每个功能,如图 2 所示。例如,我们正在从 10 个产品方面考虑 HTC。相机显然是最好的

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 987

一。权力可能是主要关注的特征,因为在一个有趣的部分中既有正面的意见,也有负面的意见。



图2:各品牌功能比较

条形图 – 比较每个品牌的相同功能,如图 3 所示。例如,我们正在考虑网络功能。iPhone 是讨论最多的,同时正面和负面意见的数量都很高。HTC 的表现相当出色,因为正面意见比负面意见多出两到三倍。



图 3:各特征品牌比较

第 988章

折线图 - 比较所选品牌的每个功能,如图 4 所示。例如,我们选择将黑莓与摩托罗拉和 Nexus 进行比较。



图4 :品牌与特征比较

4.4 验证

我们通过邀请智能手机行业的三位专家为每个智能手机品牌评分来验证系统结果。评分基于 1 到 5 的李克特量表,其中 5 表示非常同意,1 表示非常不同意。专家了解我们的实验并查看系统结果后,我们要求他们对系统结果的信心进行评分。信心分数如表 8 所示。

品牌	专家的置信度评分					阿韦拉格 e 重量
	非常同意(5)	同意 (3) (2)	未定	不同意 (4)	非常不同意 (1) 0	
黑莓	0	1	2	0		3.33
宏达电	1	1	1	0	0	4
iPhone	2	1	0	0	0	4.67
摩托罗拉	0	2	1	0	0	3.67
关系	1	2	0	0	0	4.33
诺基亚	0	0	2	1	0	2.67
棕榈	0	0	2	1	0	2.67
三星	0	0	3	0	0	3
平均的	0.5	0.875	1.375	0.25	0	

表 8:专家对每个品牌的置信度得分

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 989

如表8所示,所有专家基本都同意系统结果。排名前三的得分分别属于未定、同意和非常同意。获得最高信任度得分的前三大品牌是iPhone、Nexus和Blackberry。这些品牌是市场上总销售额最高的品牌之一。获得较低信任度得分的品牌是诺基亚和Palm。两者的权重得分都只有2.67,低于平均水平。原因可能是诺基亚和Palm的推文数量与其他品牌相比相对较低。推文数量少直接反映了这些品牌的低知名度。此外,还有一些问题,例如:并非所有HTC和三星机型都具有良好的相机功能。关于网络功能,它不是智能手机的直接属性,而是还取决于网络运营商。

5 讨论

在本实验中,系统结合了基于词典和基于机器学习的方法。与 Pang 的工作 [Pang, 02] 一样,我们在意见过滤模型 (OFM) 中使用机器学习方法来获得高召回率,而基于词典方法的 SentiWordNet 在极性检测模块 (PDM) 中使用,可以提供高精度 [Baccianella, 10]。

尽管之前已有一些关于微博情感分析的研究,但没有研究强调最终结果的利用。Twitter 情绪 [Go, 09] 的结果仅提供了消费者对产品情绪的概述,但它无法指定对任何确切功能的情绪。

然而,MSAS 的结果不一定能直接反馈产品特性本身。例如,iPhone 的结果显示苹果的配件可能存在一些问题,但事实上大多数配件都不是苹果生产的。苹果允许第三方公司为他们的 iPhone 制造配件。因此,该系统的结果必须首先符合现实世界的情况。

六,结论

在本文中,我们报告了一项利用微博情绪分析系统 (MSAS)来发现消费者洞察的探索性研究。我们的工作再次证实,微博 (尤其是 Twitter)上的情绪分析可以提供支持

为智能手机行业的生产商提供信息,以便他们做出一些决定

下一代产品。我们的系统 MSAS 可以收集有关以下方面的信息:

该产品在不打扰消费者的情况下对产品功能进行了评论,结果也得到了该领域专家的认可。最后,我们得出结论,微博情绪分析对于消费者研究非常有用,特别是在客户花时间在社交媒体上的行业。我们从社交媒体收集的数据越少,分析结果的误差就越大。

990 Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现……

7 未来工作

我们未来的工作是提高分析结果的质量。将会对算法进行更多的研究,以提高系统的性能和准确性。
我们还将研究一些解决方案来处理一些有趣且具有挑战性的问题,例如随着时间的推移监控每个用户的意见和情绪。其他问题是改进算法,以便 (1)处理同一帖子中包含多个特征和多个极性的推文,以及 (2)处理同一帖子中的多个否定。

致谢

本研究得到朱拉隆功大学 90 周年基金会 (Ratchadaphiseksomphot Endowment Fund) 的支持。

参考

[Agarwal, 11] Agarwal, A.,Xie, B.,Vovsha, I.,Rambow, O. 和 Passonneau, R.:Twitter 数据的情感分析。在社交媒体语言研讨会 (LSM 11) 论文集上。计算语言学协会,第 30-38 页,2011 年。

[Baccianella, 10] Baccianella, S.,Esuli, A. 和 Sebastiani, F.:SENTIWORDNET 3.0:用于情绪分析和观点挖掘的增强词汇资源。《第七届国际语言资源与评估会议 (LREC 10) 》论文集,2010 年。

[Bobo, 09] Bobo, LD 和 Dawson, MC:变化已经到来:种族、政治和奥巴马总统之路。杜波依斯评论:种族社会科学研究 6(1):第 1-14 页,2009 年。

[Cashmore, 10] Cashmore P.:Facebook 创始人谈隐私:公共是新的“社会规范”
2010 年 1 月 11 日 摘自: <http://mashable.com/2010/01/10/facebook-founder-on-privacy/>

[Chamlertwat, 11] Chamlertwat, W.,Bhattachakosol, P. 和 Haruechaiyasak, C.:通过意图分类改进 Twitter 上的情绪分析。2011 年第三届计算机工程与应用国际会议 (ICCEA 2011) 论文集,2011 年。

[Esuli, 07a] Esuli, A. 和 Sebastiani, F.:PageRanking WordNet Synsets:意见挖掘的应用。载于 ACL-07 会议记录,计算语言学协会第 45 届年会,第 424-431 页,2007 年。

[Esuli, 07b] Esuli, A. 和 Sebastiani, F.:SentiWordNet:用于意见挖掘的公开词汇资源。第五届语言资源与评估会议论文集,第 417-422 页,2007 年。

[Ferreira, 11] Ferreira, A.:Android 操作系统改变了智能手机的生命周期。2011 年 2 月 16 日 检索自 <http://www.theusdvista.com/business/android-os-changes-smartphone-life-cycle-1.2000033>

[Glance, 05] Glance, N.,Hurst, M.,Nigam, K.,Siegler, M.,Stockton, R. 和 Tomokiyo, T.:从在线讨论中获取营销情报。第 11 届 ACM SIGKDD 数据挖掘知识发现国际会议 (KDD 05) 论文集。ACM,第 419-428 页,2007 年。

Chamlertwat W.,Bhattachakosol P.,Rungkasiri T.,Haruechaiyasak C.:发现... 991

[Go, 09] Go, A.,Bhayani, R. 和 Huang, L.:使用远程监督进行 Twitter 情绪分类。斯坦福大学,加利福尼亚州帕洛阿尔托,美国,2009 年。

[Hu, 04] Hu, M. 和 Liu, B.:挖掘和总结客户评论。在第十届 ACM SIGKDD 知识发现和数据挖掘国际会议 (KDD 04) 的论文集上。ACM,第 168-177 页,2004 年。

[Jung, 08] Jung, JJ: 异构虚拟组织间互操作性的分类对齐。应用专家系统 36(4): 2721-2731, 2008。

[Jung, 09] Jung, JJ: 语境化查询抽样以发现 Web 上的语义资源描述。信息处理与管理 45(2): pp. 283-290, 2009。

[Jung, 10a] Jung, JJ: 用于分布式环境中查询转换的本地映射组合。应用专家系统 37(12): pp. 8401-8405, 2010。

[Jung, 10b] Jung, JJ: 整合社交网络以实现移动服务平台中的情境融合。《通用计算机科学杂志》16(15): pp. 2099-2110, 2010。

[Jung, 11a] Jung, JJ: 基于情境同步促进社会协作:一项实证研究。Expert Systems with Applications 38(5): pp. 4809-4815, 2011。

[Jung, 11b] Jung, JJ: 面向服务架构中基于服务链的业务联盟形成。Expert Systems with Applications 38(3): pp. 2206-2211, 2011。

[Jung, 12] Jung, JJ:大规模信息源中基于语义的查询采样的进化方法。信息科学 182(1):第 30-39 页,2012 年。

[Kouloumpis, 11] Kouloumpis, E.,Wilson, T. 和 Moore, J.:Twitter 情绪分析:好、坏和 OMG!第五届国际 AAAI 网络日志和社交媒体会议 (ICWSM-11),2011 年。

[Lake, 11] Twitter 情绪分析。西密歇根大学,美国密歇根州卡拉马祖,2011 年。

[Liu, 05] 刘斌、胡敏、程建军:《观点观察家:分析和比较网络上的观点》。《第 14 届万维网国际会议 (WWW 05) 》论文集。ACM,第 342-351 页,2005 年。

[Liu, 11] Liu, B.:意见挖掘和情感分析、网络数据挖掘、以数据为中心的系统和应用程序,第 2 部分,第 459-526 页,2011 年。

[Pang, 02] Pang, B.,Lee, L. 和 Vaithyanathan, S.:赞许? 使用机器学习技术进行情感分类。在 ACL-02 自然语言处理经验方法会议论文集 - 第 10 卷 (EMNLP 02) 中,第 79-86 页,2002 年。

[Stone, 08] Stone, M.,Bond, A. 和 Foss, B.:消费者洞察如何使用数据和市场研究来拉近与客户的距离。Kogan Page,2008。

[Turney, 02] Turney, P.:赞成还是反对? :语义取向应用于无监督的评论分类。第 40 届计算语言学协会年会论文集 (ACL 02),第 417-424 页,2002 年

[Twitter 博客,11] 你的世界,更加互联互通。2011 年 8 月 1 日 摘自 <http://blog.twitter.com/2011/08/your-world-more-connected.html>

[Twitter 网站,11] 使用 Twitter 搜索 API。2011 年 10 月 10 日检索自 <https://dev.twitter.com/docs/using-search>

第 992 章

[Wilson, 05] Wilson, T., Wiebe, J. 和 Hoffmann, P.: 在短语级情绪分析中识别语境极性。《人类语言技术与自然语言处理实证方法会议论文集》(HLT 05), 第 347-354 页, 2005 年。

[Zabin, 08] Zabin, J., Jefferies, A.: 社交媒体监控与分析: 从在线对话中获取消费者洞察。Aberdeen Group 基准报告, 2008 年。

[Zhai, 11] Zhai, Z., Liu, B., Xu, H., and Jia, P.: 用于意见挖掘的聚类产品特征。在第四届 ACM 网络搜索和数据挖掘国际会议 (WSDM 11) 论文集上。ACM, 第 347-354 页, 2011 年。

[Zhang, 11] Zhang, L., Ghosh, R., Dekhil, M., Hsu, M. 和 Liu, B.: 结合基于词典和基于学习的 Twitter 情绪分析方法。HP 实验室, HPL-2011-89, 2011 年。

[Zhe, 10] Zhe, X.: 一种融合多种算法和多种特征的情感分析模型。美国俄亥俄州立大学, 哥伦布, 俄亥俄州, 2010 年。