(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 112651237 A
(43)申请公布日 2021.04.13

(21)申请号 201910961379.5

(22)申请日 2019.10.11

(71)申请人 武汉渔见晚科技有限责任公司
地址 430070 湖北省武汉市洪山区创意天地08创意工坊

(72)发明人 刘垚 邹更 任钰欣 黄梓杰

(74)专利代理机构 武汉科皓知识产权代理事务所(特殊普通合伙) 42222
代理人 罗飞

(51)Int.Cl.
G06F 40/284(2020.01)
G06F 16/35(2019.01)

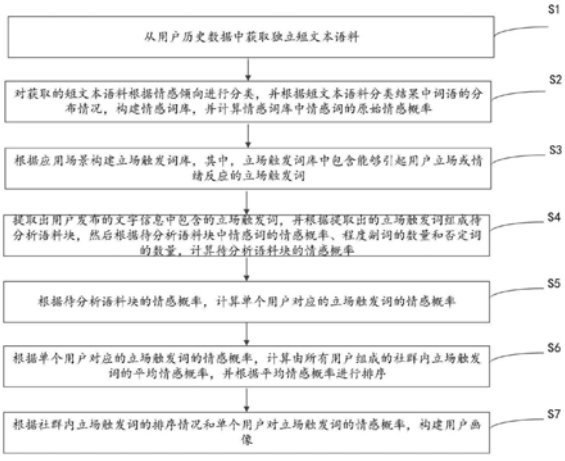
权利要求书3页 说明书14页 附图4页

(54)发明名称

一种基于用户情绪立场的用户画像建立方法及装置、用户画像的可视化方法

(57)摘要

本发明公开了一种基于用户情绪立场的用户画像建立方法及装置、用户画像的可视化方法，其中的用户画像建立方法包括：从用户历史数据中获取独立短文本语料；对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库；根据应用场景构建立场触发词库；计算待分析语料块的情感概率；根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率；根据单个用户对应的立场触发词的情感概率，计算社群内立场触发词的平均情感概率，并根据平均情感概率进行排序；根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像。本发明的方法可以提高用户情绪分析的准确性和直观性。



CN 112651237 A

1.一种基于用户情绪立场的用户画像建立方法,其特征在于,包括:

从用户历史数据中获取独立短文本语料;

对获取的短文本语料根据情感倾向进行分类,并根据短文本语料分类结果中词语的分布情况,构建情感词库,并计算情感词库中情感词的原始情感概率;其中,情感词库中包含正面情感词和负面情感词;

根据应用场景构建立场触发词库,其中,立场触发词库中包含能够引起用户立场或情绪反应的立场触发词;

提取出用户发布的文字信息中包含的立场触发词,并根据提取出的立场触发词组成待分析语料块,然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量,计算待分析语料块的情感概率;

根据待分析语料块的情感概率,计算单个用户对应的立场触发词的情感概率;

根据单个用户对应的立场触发词的情感概率,计算由所有用户组成的社群内立场触发词的平均情感概率,并根据平均情感概率进行排序;

根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率,构建用户画像。

2.如权利要求1所述的方法,其特征在于,对获取的短文本语料根据情感倾向进行分类,并根据短文本语料分类结果中词语的分布情况,构建情感词库,并计算情感词库中情感词的情感概率,包括:

对获取的短文本语料根据情感倾向进行分类,分为正向语料、中立语料和负向语料;

将分类后的语料进行分词,并去冗余后得到语料词库;

统计语料词库中的每一个词语在正向语料、中立语料和负向语料中的分布情况;

根据词语的分布情况,结合卡方校验筛选出与正向、负向有关的词汇作为情感倾向的标志词候选;

对标志词候选进行筛选,删除与对应情感倾向不匹配的词汇,构建情感词库;

查找每一个正面情感词对应的所有原始语料,计算出正面情绪概率的平均值,作为正面情感词的原始情感概率,对于情感词库中的负面情感词,将1减去正面情绪概率的平均值的结果,作为负面情感词的原始情感概率。

3.如权利要求1所述的方法,其特征在于,根据提取出的立场触发词组成待分析语料块,然后根据待分析语料块中情感词、程度副词和否定词的数量,计算待分析语料块的情感概率,包括:

将提取出的立场触发词所在的句子以及前后n个句子组成待分析语料块,其中,n为大于等于1的正整数;

查找待分析语料块中出现的正面情感词和负面情感词,并获取每个正面情感词和负面情感词的原始情感概率;

根据每一个正面情感词和负面情感词,根据预设范围内否定词和程度副词的数量,确定否定系数和程度权重;

根据情感词的原始情感概率、否定系数和程度权重,计算每个情感词的情感概率修正值;

根据情感词的情感概率修正值、正面情感词的数量和负面情感词的数量,计算待分析

语料块的情感概率。

4.如权利要求1所述的方法,其特征在于,根据待分析语料块的情感概率,计算单个用户对应的立场触发词的情感概率,包括:

当立场触发词在用户发布的数据中未出现时,立场触发词的情感概率为空;

当立场触发词在用户发布的数据中出现的一次时,将立场触发词出现的语料块对应的情感概率作为立场触发词的情感概率;

当立场触发词在用户发布的数据中出现的两次及以上时,将立场触发词出现的所有语料块对应的情感概率的平均值作为立场触发词的情感概率。

5.如权利要求1所述的方法,其特征在于,在根据待分析语料块的情感概率,计算单个用户对应的立场触发词的情感概率之后,所述方法还包括:

对单个用户对应的立场触发词的情感概率进行归一化处理,获得单个用户对应的每一个立场触发词的情感概率修正值。

6.如权利要求5所述的方法,其特征在于,根据单个用户对应的立场触发词的情感概率,计算由所有用户组成的社群内立场触发词的平均情感概率,包括:

将每个用户对应的立场触发的立场触发词的情感概率修正值求平均值,计算出社群内立场触发词的平均情感概率;

根据每个立场触发词的平均情感概率的大小进行排序。

7.一种基于用户情绪立场的用户画像建立装置,其特征在于,包括:

语料获取模块,用于从用户历史数据中获取独立短文本语料;

情感词库构建模块,用于对获取的短文本语料根据情感倾向进行分类,并根据短文本语料分类结果中词语的分布情况,构建情感词库,并计算情感词库中情感词的原始情感概率;其中,情感词库中包含正面情感词和负面情感词;

立场触发词库构建模块,用于根据应用场景构建立场触发词库,其中,立场触发词库中包含能够引起用户立场或情绪反应的立场触发词;

待分析语料块情感概率计算模块,用于提取出用户发布的文字信息中包含的立场触发词,并根据提取出的立场触发词组成待分析语料块,然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量,计算待分析语料块的情感概率;

单个用户立场触发词情感概率计算模块,用于根据待分析语料块的情感概率,计算单个用户对应的立场触发词的情感概率;

平均情感概率排序模块,用于根据单个用户对应的立场触发词的情感概率,计算由所有用户组成的社群内立场触发词的平均情感概率,并根据平均情感概率进行排序;

用户画像构建模块,根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率,构建用户画像。

8.一种用户画像的可视化方法,其特征在于,包括:对权利要求1至6任一项所示方法所构建的用户画像进行可视化显示。

9.如权利要求8所述的方法,其特征在于,对用户画像进行可视化显示,包括:

将立场触发词依据平均情感概率的大小映射至预设形状的词块;

构建单个用户对立场触发词的情感概率与色彩特征之间的对应关系;

根据情感概率与色彩特征之间的对应关系,对用户画像进行可视化显示。

10.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该程序被执行时实现如权利要求1至6中任一项权利要求所述的方法。

一种基于用户情绪立场的用户画像建立方法及装置、用户画像的可视化方法

技术领域

[0001] 本发明涉及数据分析技术领域，具体涉及一种基于用户情绪立场的用户画像建立方法及装置、用户画像的可视化方法。

背景技术

[0002] 用户在网络平台中的用户行为常常被用来描述一个用户的特征，这种用户特征被称为用户画像，而根据目的不同，构建用户画像的侧重点也会不同。例如电商平台侧重于用户的消费能力，购买偏好建立用户画像，而社交平台会基于用户的兴趣特点和社交关系建立用户画像，不同的用户画像会帮助平台对用户进行分类，更好的为用户实现定制化服务。

[0003] 本申请发明人在实施本发明的过程中，发现现有技术的方法，至少存在如下技术问题：

[0004] 现有技术中，在对用户进行情感分析时，其采用的情感词库中包含了大量的非网络用语和非日常用语的词汇，同时缺乏现今网络常用语的词汇，使得基于现有的情感词库的情感分析准确性和实用性有所限制。

[0005] 由此可知，现有技术中的方法存在分析结果不够准确的技术问题。

发明内容

[0006] 有鉴于此，本发明提供了一种基于用户情绪立场的用户画像建立方法及装置、用户画像的可视化方法，用以解决或者至少部分解决现有技术中的方法存在的结果不够准确的技术问题。

[0007] 本发明第一方面提供了一种基于用户情绪立场的用户画像建立方法，包括：

[0008] 从用户历史数据中获取独立短文本语料；

[0009] 对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库，并计算情感词库中情感词的原始情感概率；其中，情感词库中包含正面情感词和负面情感词；

[0010] 根据应用场景构建立场触发词库，其中，立场触发词库中包含能够引起用户立场或情绪反应的立场触发词；

[0011] 提取出用户发布的文字信息中包含的立场触发词，并根据提取出的立场触发词组成待分析语料块，然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量，计算待分析语料块的情感概率；

[0012] 根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率；

[0013] 根据单个用户对应的立场触发词的情感概率，计算由所有用户组成的社群内立场触发词的平均情感概率，并根据平均情感概率进行排序；

[0014] 根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像。

[0015] 在一种实施方式中，对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库，并计算情感词库中情感词的情感概率，包括：

[0016] 对获取的短文本语料根据情感倾向进行分类，分为正向语料、中立语料和负向语料；

[0017] 将分类后的语料进行分词，并去冗余后得到语料词库；

[0018] 统计语料词库中的每一个词语在正向语料、中立语料和负向语料中的分布情况；

[0019] 根据词语的分布情况，结合卡方校验筛选出与正向、负向有关的词汇作为情感倾向的标志词候选；

[0020] 对标志词候选进行筛选，删除与对应情感倾向不匹配的词汇，构建情感词库；

[0021] 查找每一个正面情感词对应的所有原始语料，计算出正面情绪概率的平均值，作为正面情感词的原始情感概率，对于情感词库中的负面情感词，将1减去正面情绪概率的平均值的结果，作为负面情感词的原始情感概率。

[0022] 在一种实施方式中，根据提取出的立场触发词组成待分析语料块，然后根据待分析语料块中情感词、程度副词和否定词的数量，计算待分析语料块的情感概率，包括：

[0023] 将提取出的立场触发词所在的句子以及前后n个句子组成待分析语料块，其中，n为大于等于1的正整数；

[0024] 查找待分析语料块中出现的正面情感词和负面情感词，并获取每个正面情感词和负面情感词的原始情感概率；

[0025] 根据每一个正面情感词和负面情感词，根据预设范围内否定词和程度副词的数量，确定否定系数和程度权重；

[0026] 根据情感词的原始情感概率、否定系数和程度权重，计算每个情感词的情感概率修正值；

[0027] 根据情感词的情感概率修正值、正面情感词的数量和负面情感词的数量，计算待分析语料块的情感概率。

[0028] 在一种实施方式中，根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率，包括：

[0029] 当立场触发词在用户发布的数据中未出现时，立场触发词的情感概率为空；

[0030] 当立场触发词在用户发布的数据中出现的一次时，将立场触发词出现的语料块对应的情感概率作为立场触发词的情感概率；

[0031] 当立场触发词在用户发布的数据中出现的两次及以上时，将立场触发词出现的所有语料块对应的情感概率的平均值作为立场触发词的情感概率。

[0032] 在一种实施方式中，在根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率之后，所述方法还包括：

[0033] 对单个用户对应的立场触发词的情感概率进行归一化处理，获得单个用户对应的每一个立场触发词的情感概率修正值。

[0034] 在一种实施方式中，根据单个用户对应的立场触发词的情感概率，计算由所有用户组成的社群内立场触发词的平均情感概率，包括：

[0035] 将每个用户对应的立场触发的立场触发词的情感概率修正值求平均值，计算出社

群内立场触发词的平均情感概率；

[0036] 根据每个立场触发词的平均情感概率的大小进行排序。

[0037] 基于同样的发明构思，本发明第二方面提供了一种基于用户情绪立场的用户画像建立装置，包括：

[0038] 语料获取模块，用于从用户历史数据中获取独立短文本语料；

[0039] 情感词库构建模块，用于对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库，并计算情感词库中情感词的原始情感概率；其中，情感词库中包含正面情感词和负面情感词；

[0040] 立场触发词库构建模块，用于根据应用场景构建立场触发词库，其中，立场触发词库中包含能够引起用户立场或情绪反应的立场触发词；

[0041] 待分析语料块情感概率计算模块，用于提取出用户发布的文字信息中包含的立场触发词，并根据提取出的立场触发词组成待分析语料块，然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量，计算待分析语料块的情感概率；

[0042] 单个用户立场触发词情感概率计算模块，用于根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率；

[0043] 平均情感概率排序模块，用于根据单个用户对应的立场触发词的情感概率，计算由所有用户组成的社群内立场触发词的平均情感概率，并根据平均情感概率进行排序；

[0044] 用户画像构建模块，根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像。

[0045] 基于同样的发明构思，本发明第三方面提供了一种用户画像的可视化方法，包括：对第一方面所述方法所构建的用户画像进行可视化显示。

[0046] 在一种实施方式中，对用户画像进行可视化显示，包括：

[0047] 将立场触发词依据平均情感概率的大小映射至预设形状的词块；

[0048] 构建单个用户对立场触发词的情感概率与色彩特征之间的对应关系；

[0049] 根据情感概率与色彩特征之间的对应关系，对用户画像进行可视化显示。

[0050] 基于同样的发明构思，本发明第四方面提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被执行时实现第一方面所述的方法

[0051] 本申请实施例中的上述一个或多个技术方案，至少具有如下一种或多种技术效果：

[0052] 本发明提供的一种基于用户情绪立场的用户画像建立方法，首先，从用户历史数据中获取独立短文本语料；接着构建情感词库，并计算情感词库中情感词的原始情感概率；接着根据应用场景构建立场触发词库；然后计算待分析语料块的情感概率；接下来根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率；接着计算由所有用户组成的社群内立场触发词的平均情感概率，并根据平均情感概率进行排序；最后根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像。

[0053] 由于本发明提供的方法，可以根据短文本语料分类结果中词语的分布情况，构建情感词库，构建的情感词库中根据情感倾向将其分为正面情感词和负面情感词，可以更好的从用户的文字内容中识别新词的情感倾向、较为准确地分析其中的情绪反应，并构建立场触发词库，然后建立单个用户与立场触发词库的映射，计算出单个用户对应的立场触发

词的情感概率;再计算由所有用户组成的社群内立场触发词的平均情感概率,并进行排序,从而可以根据排序结果进行用户画像的构建,从而可以准确对每个用户对立场触发词的情绪立场,从而可以快速准确地了解每个用户对普遍事物的观点差异和情绪反应特点,解决了现有技术中的方法存在分析结果不够准确的技术问题。

[0054] 进一步地,基于构建的用户画像,本发明还提供了一种用户画像的可视化方法,对用户画像进行可视化显示,提高了直观性。

[0055] 进一步地,根据颜色渐变公式,构建单个用户对立场触发词的情感概率与色彩特征之间的对应关系,然后根据情感概率与色彩特征之间的对应关系,对用户画像进行可视化显示,可以改善显示效果。

附图说明

[0056] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0057] 图1为本发明一种基于用户情绪立场的用户画像建立方法的流程示意图;

[0058] 图2为本发明实施例中一种基于用户情绪立场的用户画像建立装置的结构框图;

[0059] 图3为本发明具体实施例中单个用户的用户画像可视化效果示意图;

[0060] 图4为本发明具体实施例中不同用户的用户画像可视化效果示意图;

[0061] 图5为本发明具体实施例中群体用户的用户画像可视化效果示意图;

[0062] 图6为本发明实施例中一种计算机可读存储介质的结构框图。

具体实施方式

[0063] 本发明的目的在于针对现有技术中的方法存在的结果不够准确的技术问题,提供一种基于用户情绪立场的用户画像建立方法及装置、用户画像的可视化方法,该方法基于用户行为记录,通过系统的解析,抽取与立场触发词库中词汇相对应的情绪表达,进行用户画像的构建,并进一步与制定的相应的色彩特征建立映射(即建立立场触发词与色彩特征之间的对应关系),将结果可视化显示,从而达到提高分析准确性以及显示的直观性的技术效果。

[0064] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0065] 本申请发明人通过大量的研究与实践发现,在用户的行为中,很多数据都会反映出用户的情绪,与情绪相对应的是用户对于许多事物的观点及立场,基于用户对于普遍事物的情绪反应,从而可以构建出用户的情绪图谱。在情感分析方面,现有的情感词库中包含了大量的非网络用语和非日常用语的词汇,同时缺乏现今网络常用语的词汇,使得基于现有的情感词库的情感分析准确性和实用性有所限制。

[0066] 本发明提供了一种基于用户情绪立场的用户画像建立方法、用户画像可视化方

法,其中采用了一种新的情感词库构建方法,总体具有如下优点或者有益技术效果:

[0067] 1.采用本发明中构建的情感词库,可以更好地从用户的文字内容中识别新词的情感倾向、较为准确地分析其中的情绪反应;

[0068] 2.使得用户可以快速地了解自己和他人对普遍事物的观点差异和情绪反应特点;

[0069] 3.实现对某个用户群体的情绪立场特点进行可视化分析;

[0070] 4.同时也有助于对用户进行更多维度的分类。

[0071] 实施例一

[0072] 本实施例提供了一种基于用户情绪立场的用户画像建立方法,请参见图1,该方法包括:

[0073] 步骤S1:从用户历史数据中获取独立短文本语料。

[0074] 具体来说,用户历史数据包括用户的历史行为数据,例如用户的留言信息、发表的评论信息等。独立短文本语料表示具有一定的含义的文本,可以通过现有的工具实现。

[0075] 步骤S2:对获取的短文本语料根据情感倾向进行分类,并根据短文本语料分类结果中词语的分布情况,构建情感词库,并计算情感词库中情感词的原始情感概率;其中,情感词库中包含正面情感词和负面情感词。

[0076] 具体来说,情感倾向即用户的立场或者态度,例如“喜欢”表示正向或者正面情感,“吃”表示中立,“讨厌”表示负向或者负面情感。词语即分类后的语料中包含的词语,可以通过分词操作得到。原始语料的正面情绪概率,为采用开源情感分析API计算出的该语料中包含的情感,属于正面情绪的概率。查找每一个情感词对应的所有原始语料,计算它们的正面情绪概率的平均值,则可以得到情感词的原始情感概率。

[0077] 步骤S3:根据应用场景构建立场触发词库,其中,立场触发词库中包含能够引起用户立场或情绪反应的立场触发词。

[0078] 具体来说,根据应用场景以及平台特色而搜集的可引起用户立场或情绪反应的名词,如歌曲,电影,名人明星,热门概念等。用户可以较为明确地指出对一首歌曲,一部电影,一个人物或者一个热门概念如“转基因”等的态度是正面的还是负面的。而由立场触发词构成的集合就是立场触发词库。

[0079] 步骤S4:提取出用户发布的文字信息中包含的立场触发词,并根据提取出的立场触发词组成待分析语料块,然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量,计算待分析语料块的情感概率。

[0080] 具体来说,用户发布的文字信息,可以是评论、文章等内容。在具体实施时,可以结合立场触发词所在的句子以及上下文信息,组成待分析语料块。

[0081] 步骤S5:根据待分析语料块的情感概率,计算单个用户对应的立场触发词的情感概率。

[0082] 具体来说,可以根据待分析语料块中立场触发词出现的次数、时间等,确定单个用户对应的立场触发词的情感概率,从而构建了单个用户与立场触发词之间的映射。

[0083] 步骤S6:根据单个用户对应的立场触发词的情感概率,计算由所有用户组成的社群内立场触发词的平均情感概率,并根据平均情感概率进行排序。

[0084] 具体来说,步骤S5中得到的是对于单个用户来说,立场触发词库中立场触发词的情感概率,本步骤则是所有用户组成的社群内立场触发词的平均情感概率。可以用出现过

该立场触发词的全部语料的情绪概率求和，再除以出现过该词汇的语料文本总数量来计算。

[0085] 步骤S7：根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像。

[0086] 具体来说，可以根据立场触发词的平均情感概率从大到小进行排序，然后根据单个用户对立场触发词的情感概率，构建用户画像。

[0087] 在一种实施方式中，对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库，并计算情感词库中情感词的情感概率，包括：

[0088] 对获取的短文本语料根据情感倾向进行分类，分为正向语料、中立语料和负向语料；

[0089] 将分类后的语料进行分词，并去冗余后得到语料词库；

[0090] 统计语料词库中的每一个词语在正向语料、中立语料和负向语料中的分布情况；

[0091] 根据词语的分布情况，结合卡方校验筛选出与正向、负向有关的词汇作为情感倾向的标志词候选；

[0092] 对标志词候选进行筛选，删除与对应情感倾向不匹配的词汇，构建情感词库；

[0093] 查找每一个正面情感词对应的所有原始语料，计算出正面情绪概率的平均值，作为正面情感词的原始情感概率，对于情感词库中的负面情感词，将1减去正面情绪概率的平均值的结果，作为负面情感词的原始情感概率。

[0094] 具体来说，在进行候选标志词筛选时，利用卡方检验，判断每个词对三类语料的标志性作用，即句子中是否含有该词汇，与正向、中立、负向三种情感倾向是否有关，然后利用卡方检验的结果，筛选出候选标志词。其中，标志词候选是指初步筛选出的标志词集合。

[0095] 对候选标志词进行词性筛选和人工筛选，去除与对应情感倾向不匹配的词汇，将剩余的词汇作为情感词，并且根据每个词汇对应的情感倾向分为正面情感词和负面情感词。

[0096] 接着分别计算正面情感词和负面情感词的原始情感概率，对于正面情感词库，查找词库中每一个词汇对应的所有原始语料，计算它们的正面情绪概率的平均值，作为该词汇的原始情感概率。为便于计算，再利用归一化公式，将正面情感词库中所有词汇的原始情感概率的值域统一到0～1之间，作为该词汇的情感分数。(需要说明的是，原始语料的正面情绪概率，是指采用开源情感分析API计算出的该语料中包含的情感，属于正面情绪的概率)

[0097] 对于负面情感词库，查找词库中每一个词汇对应的所有原始语料，用1减去所有原始语料的正面情绪概率的平均值，将结果作为该词汇的原始情感概率。为便于计算，利用归一化公式，将负面情感词库中所有词汇的情感概率的值域统一到0～1之间，作为该词汇的情感分数。

[0098] 其中，归一化的公式为：$x' = (x-X_{min})/(X_{max}-X_{min})$ 其中，x表示需要进行归一化的数据，x'是x进行归一化后的取值，X_{max}是所有需要进行归一化的数据中的最大值，X_{min}是所有需要进行归一化的数据中的最小值。

[0099] 通过上述方法可以构建具有情感倾向的情感词库，并且其中采用卡方检验的方

法，可以更为准确地筛选出候选标志词，从而使得情感词库更为准确。

[0100] 在一种实施方式中，根据提取出的立场触发词组成待分析语料块，然后根据待分析语料块中情感词、程度副词和否定词的数量，计算待分析语料块的情感概率，包括：

[0101] 将提取出的立场触发词所在的句子以及前后n个句子组成待分析语料块，其中，n为大于等于1的正整数；

[0102] 查找待分析语料块中出现的正面情感词和负面情感词，并获取每个正面情感词和负面情感词的原始情感概率；

[0103] 根据每一个正面情感词和负面情感词，根据预设范围内否定词和程度副词的数量，确定否定系数和程度权重；

[0104] 根据情感词的原始情感概率、否定系数和程度权重，计算每个情感词的情感概率修正值；

[0105] 根据情感词的情感概率修正值、正面情感词的数量和负面情感词的数量，计算待分析语料块的情感概率。

[0106] 具体来说，待分析语料块则为由将该立场触发词的所在句子及前后n句一共2n+1个句子构成。然后对待分析语料块的情感特征进行分析。

[0107] 在具体的实施过程中，预设范围可以根据需要选取，例如可以选取情感词前后出现的若干词。举例来说，实现过程如下：

[0108] a.从每一个正面情感词汇和负面情感词汇向前寻找两个词汇，如果这两个词汇中含有k个否定词(k的值域为[0,2])，则该情感词汇的否定系数$N_i=(-1)^k$。

[0109] b.如果这两个词汇中不含程度副词，则该情感词汇的程度系数$L_i=1$；

[0110] c.如果含有1个程度副词，其程度权重为L，则该情感词汇的程度系数$L_i=L$，如果含有2个程度副词，其程度权重分别为L_1,L_2，则该情感词汇的程度系数$L_i=L_1\times L_2$；

[0111] d.计算每个情感词的情感概率修正值$P_{i_index}=P_i\times N_i\times L_i$。

[0112] 然后根据计算出的情感词的情感概率修正值、正面情感词的数量和负面情感词的数量，计算待分析语料块的情感概率。

[0113] 具体实施时，将情感词分数情感概率修正值根据情感类型，分为正面情感词和负面情感词两组，即将情感概率修正值大于或等于1的作为正面情感词，小于1的作为负面情感词，并分别按照数值从大到小排列。其中，正面情感词的数量为N_p，负面情感词的数量为N_n。然后按照如下方法计算语料块的情感分数S：

[0114] • 如果$N_p\geq N_n$，且N_n不等于0，首先计算所有负面情感词汇的平均值$\bar{P}_n$，并计算前N_n个正面情感词汇的平均值$\bar{P}_{p1}$以及剩余的正面情感词汇的平均值$\bar{P}_{p2}$，然后计算$\bar{P}_{p1}$和$\bar{P}_{p2}$的平均值$\bar{P}_p$。语料块的情感分数$S=\bar{P}_p/\bar{P}_n$。

[0115] • 如果$N_p<N_n$，且N_p不等于0，首先计算所有正面情感词汇的平均值$\bar{P}'_p$，计算前N_n个负面情感词汇的平均值$\bar{P}_{n1}$以及剩余的负面情感词汇的平均值$\bar{P}_{n2}$，然后计算$\bar{P}_{n1}$和$\bar{P}_{n2}$的平均值$\bar{P}'_n$。语料块的情感分数$S=\bar{P}'_p/\bar{P}'_n$

[0116] • 如果N_n等于0，则该语料块中的负面情感词汇数量为0，文中全部都是正面情感词汇。计算所有正面情感词汇的平均值$\bar{P}'_p$，语料块的情感分数$S=(\bar{P}'_p\times L_{max})/2$，其中$L_{max}$是

程度副词词库中程度权重的最大值。

[0117] • 如果N_p等于0，则该语料块中的正面情感词汇数量为0，文中全部都是负面情感词汇。计算所有负面情感词汇的平均值$\bar{P}_n$，语料块的情感分数$S=(\bar{P}_n/L_{max})\times 2$，其中$L_{max}$是程度副词词库中，程度权重的最大值。

[0118] 通过结合情感词的情感概率修正值、正面情感词的数量和负面情感词的数量的方式，可以提高计算的客观性和准确性。

[0119] 在一种实施方式中，根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率，包括：

[0120] 当立场触发词在用户发布的数据中未出现时，立场触发词的情感概率为空；

[0121] 当立场触发词在用户发布的数据中出现的一次时，将立场触发词出现的语料块对应的情感概率作为立场触发词的情感概率；

[0122] 当立场触发词在用户发布的数据中出现的两次及以上时，将立场触发词出现的所有语料块对应的情感概率的平均值作为立场触发词的情感概率。

[0123] 具体来说，对于在用户数据中出现两次及以上的立场触发词，其情感分数依照包含情感触发词的语料块出现的先后时间顺序加权平均计算得到。

[0124] 举例来说，假设某用户对于一个立场触发词共出现在n个语料块，时间跨度为T，则T除以10，得到10个区间，t1～t10，每个区间的权值分别为1～10，n个语料块分布于10个区间中，每个语料块的情感概率乘以对应区间的权值，最后取加权平均，得到对应的立场触发词的情感概率。

[0125] 在一种实施方式中，在根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率之后，所述方法还包括：

[0126] 对单个用户对应的立场触发词的情感概率进行归一化处理，获得单个用户对应的每一个立场触发词的情感概率修正值。

[0127] 具体来说，为了便于后续计算，本实施方式还对单个用户对应的立场触发词的情感概率进行了归一化处理。

[0128] 在具体的实施过程中，对于待分析用户群体的每一个用户j，其对应的每一个立场触发词k：

[0129] 分别对所有值不为空的立场触发词的正面情感概率(根据计算出的立场触发词的情感概率根据情感倾向划分)进行归一化处理，使所有正面情感概率的修正值分布在0～1之间。对于每一个情感概率得分S_{jk}，其修正值S_{jk_index}的计算方法如下，其中S_{min}是所有正面情感概率得分中的最小值，S_{max}是所有正面情感概率得分中的最大值，归一化处理公式为：$S_{jk\_index}=(S_{jk}-S_{min})/(S_{max}-S_{min})$。

[0130] 分别对所有值不为空的立场触发词的负面情感概率进行归一化处理，使所有负面情感概率的修正值分布在0～1之间，再取它的负数，使最终的负面情感概率得分修正值分布在-1～0之间。对于每一个情感概率得分S'_{jk}，其修正值S'_{jk_index}的计算方法如下，其中S'_{min}是所有负面情感概率得分中的最小值，S'_{max}是所有负面情感概率得分中的最大值，归一化处理公式为：$S'_{j'k\_index}=-(S'_{jk}-S'_{min})/(S'_{max}-S'_{min})$。

[0131] 在一种实施方式中，根据待分析语料块的情感概率，计算单个用户对应根据单个用户对应的立场触发词的情感概率，计算由所有用户组成的社群内立场触发词的平均情感

概率,包括:

[0132] 将每个用户对应的立场触发的立场触发词的情感概率修正值求平均值,计算出社群内立场触发词的平均情感概率;

[0133] 根据每个立场触发词的平均情感概率的大小进行排序。

[0134] 具体来说,对于立场触发词库中的每一个立场触发词k,对社群内使用过该立场触发词的全部用户文本的情感概率得分修正值S_{jk_index}求平均值,则可以得到社群内立场触发词的平均情感概率。然后按照每一个立场触发词的社群总体平均情感概率,将立场触发词按照从高到低的顺序排列。

[0135] 为了更清楚地说明本发明的方法的具体实现过程,下面通过具体示例来进行介绍:

[0136] 1.建立用户与立场触发词库的映射示例

[0137] (1)用户的原始文本(即用户发布的文字信息):

[0138] 每部动画片都会用精心编排的剧情,来向孩子们传达一些道理。天马行空的故事情节是孩子们都爱的,毕竟长大后的圣诞节再没有圣诞老人。坚持梦想不放弃,自然是以往动画片的主题,然而在神奇马戏团之动物饼干中,作为男主的爸比,面对自己极有可能无法还原的身体,几近崩溃、万般恐惧。就连一贯秉承"坚持到底"的妈咪,面对这马上就要成为的既定事实,都慌了神儿,没了主意。倒是天真烂漫的小女儿,用崇拜而依赖的眼神望向爸比,"我喜欢现在Daddy……"无法还原的爸比成了女儿的亲密玩伴、消遣玩具、可爱萌宠,当然还是无可取代的万能英雄,这让男主找回生存的意义。父母和孩子是共同成长的,你做孩子10几年,我当妈不比你多一天,我们一起努力好吧~

[0139] (2)在原始文本中,立场触发词为“神奇马戏团之动物饼干”,取该词所在句子,以及前后n句(以n=2为例)组成待分析语料块。

[0140] 每部动画片都会用精心编排的剧情,来向孩子们传达一些道理。天马行空的故事情节是孩子们都爱的,毕竟长大后的圣诞节再没有圣诞老人。坚持梦想不放弃,自然是以往动画片的主题,然而在神奇马戏团之动物饼干中,作为男主的爸比,面对自己极有可能无法还原的身体,几近崩溃、万般恐惧。就连一贯秉承"坚持到底"的妈咪,面对这马上就要成为的既定事实,都慌了神儿,没了主意。倒是天真烂漫的小女儿,用崇拜而依赖的眼神望向爸比,"我喜欢现在Daddy……"无法还原的爸比成了女儿的亲密玩伴、消遣玩具、可爱萌宠,当然还是无可取代的万能英雄,这让男主找回生存的意义。

[0141] (3)利用前文所示的情感分析方式计算语料块的情感概率,得到:

[0142] 情感概率S_{index}=0.979658

[0143] 2.立场触发词库与用户情绪的数据示例(部分)

[0144]

立场触发词	用户1的情感概率(得分)	用户2的情感概率(得分)
大话西游之月光宝盒	-0.9113407	0.912788
肖申克的救赎	0.984257	0.980452
霸王别姬	0.98767	0.969589
泰坦尼克号	0.926384	0.97395
这个杀手不太冷	0.986618	0.986399
盗梦空间	0.873364	0.966804

三傻大闹宝莱坞	0.928555	0.983658

[0145] 实施例二

[0146] 基于同样的发明构思，本实施例提供了一种基于用户情绪立场的用户画像建立装置，请参见图2，包括：

[0147] 语料获取模块201，用于从用户历史数据中获取独立短文本语料；

[0148] 情感词库构建模块202，用于对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库，并计算情感词库中情感词的原始情感概率；其中，情感词库中包含正面情感词和负面情感词；

[0149] 立场触发词库构建模块203，用于根据应用场景构建立场触发词库，其中，立场触发词库中包含能够引起用户立场或情绪反应的立场触发词；

[0150] 待分析语料块情感概率计算模块204，用于提取出用户发布的文字信息中包含的立场触发词，并根据提取出的立场触发词组成待分析语料块，然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量，计算待分析语料块的情感概率；

[0151] 单个用户立场触发词情感概率计算模块205，用于根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率；

[0152] 平均情感概率排序模块206，用于根据单个用户对应的立场触发词的情感概率，计算由所有用户组成的社群内立场触发词的平均情感概率，并根据平均情感概率进行排序；

[0153] 用户画像构建模块207，根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像。

[0154] 在一种实施方式中，情感词库构建模块202具体用于：

[0155] 对获取的短文本语料根据情感倾向进行分类，分为正向语料、中立语料和负向语料；

[0156] 将分类后的语料进行分词，并去冗余后得到语料词库；

[0157] 统计语料词库中的每一个词语在正向语料、中立语料和负向语料中的分布情况；

[0158] 根据词语的分布情况，结合卡方校验筛选出与正向、负向有关的词汇作为情感倾向的标志词候选；

[0159] 对标志词候选进行筛选，删除与对应情感倾向不匹配的词汇，构建情感词库；

[0160] 查找每一个正面情感词对应的所有原始语料，计算出正面情绪概率的平均值，作为正面情感词的原始情感概率，对于情感词库中的负面情感词，将1减去正面情绪概率的平均值的结果，作为负面情感词的原始情感概率。

[0161] 在一种实施方式中，待分析语料块情感概率计算模块204具体用于：

[0162] 将提取出的立场触发词所在的句子以及前后n个句子组成待分析语料块，其中，n为大于等于1的正整数；

[0163] 查找待分析语料块中出现的正面情感词和负面情感词，并获取每个正面情感词和负面情感词的原始情感概率；

[0164] 根据每一个正面情感词和负面情感词，根据预设范围内否定词和程度副词的数量，确定否定系数和程度权重；

[0165] 根据情感词的原始情感概率、否定系数和程度权重，计算每个情感词的情感概率修正值；

[0166] 根据情感词的情感概率修正值、正面情感词的数量和负面情感词的数量，计算待分析语料块的情感概率。

[0167] 在一种实施方式中，单个用户立场触发词情感概率计算模块205具体用于：

[0168] 当立场触发词在用户发布的数据中未出现时，立场触发词的情感概率为空；

[0169] 当立场触发词在用户发布的数据中出现的一次时，将立场触发词出现的语料块对应的情感概率作为立场触发词的情感概率；

[0170] 当立场触发词在用户发布的数据中出现的两次及以上时，将立场触发词出现的所有语料块对应的情感概率的平均值作为立场触发词的情感概率。

[0171] 在一种实施方式中，所述装置还包括归一化处理模块，用于在根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率之后：

[0172] 对单个用户对应的立场触发词的情感概率进行归一化处理，获得单个用户对应的每一个立场触发词的情感概率修正值。

[0173] 在一种实施方式中，平均情感概率排序模块206具体用于：

[0174] 将每个用户对应的立场触发的立场触发词的情感概率修正值求平均值，计算出社群内立场触发词的平均情感概率；

[0175] 根据每个立场触发词的平均情感概率的大小进行排序。

[0176] 由于本发明实施例二所介绍的装置，为实施本发明实施例一中基于用户情绪立场的用户画像建立方法所采用的装置，故而基于本发明实施例一所介绍的方法，本领域所属人员能够了解该装置的具体结构及变形，故而在此不再赘述。凡是本发明实施例一的方法所采用的装置都属于本发明所欲保护的范围。

[0177] 实施例三

[0178] 基于同一发明构思，本申请还提供了一种用户画像的可视化方法，具体包括对实施例一种所构建的用户画像进行可视化显示。

[0179] 在一种实施方式中，对用户画像进行可视化显示，包括：

[0180] 将立场触发词依据平均情感概率的大小映射至预设形状的词块；

[0181] 构建单个用户对立场触发词的情感概率与色彩特征之间的对应关系；

[0182] 根据情感概率与色彩特征之间的对应关系，对用户画像进行可视化显示。

[0183] 具体实施过程中，可以将立场触发词映射到一个几何形状体即“立场词块”，根据实施例一中根据社群内立场触发词的的排列顺序，对立场词块进行排列，构成一个整体的几何图案。

[0184] 例如以正方形为立场词块的几何形状体，300个词库就是300个小正方形，长为30个小正方形，宽为10个小正方形，最终组成一个规则的的图形(例如长宽比为3:1的长方形，其中第1个立场触发词位于第1行第1列，第30个立场触发词位于第1行第30列，第300个立场触发词位于第10行第30列)。

[0185] 构建单个用户对立场触发词的情感概率与色彩特征之间的对应关系，并根据构建的对应关系，对用户画像进行可视化显示。

[0186] 具体实施过程中，可以利用颜色渐变公式，将用户的情绪反应以颜色的色系不同及强度对每个小正方形进行着色。

[0187] 渐变色计算方法：两种颜色的渐变，就是对A、B两种颜色的RGB通道分别进行颜色

渐变公式的计算，Gradient＝A+(B-A)×p，对于每种色彩通道来说，A是颜色A在该通道上的值，B是颜色B在该通道上的值。p是目标颜色在AB之间所处位置的百分比。

[0188] 例如，选择红色系作为立场为支持或正面情感的代表色，正面情感越强烈，情感概率得分修正值S_{jk_index}越接近于1颜色越深。正面情感越弱，每一个情感概率得分S_{jk_index}越接近与0，因此正面情感的小正方形的颜色，正面情感最强的为红色(255,0,0)，最弱的为浅灰色(245,245,245)，此时对应关系(颜色的计算方法为)R_{jk}＝255+(255-245)*S_{jk_index}，G_{jk}＝255+(255-0)*S_{jk_index}，B_{jk}＝255+(255-0)*S_{jk_index}。

[0189] 类似地，选择蓝色系作为立场为反对或负面情感的代表色，负面情感越强烈，情感概率得分修正值S_{jk_index}越接近于1，颜色越深。负面情感越弱，情感概率得分修正值S_{jk_index}越接近与0。因此负面情感的小正方形的颜色，负面情感最强的为蓝色(0,0,255)，最弱的为浅灰色(245,245,245)，此时的对应关系(颜色的计算方法)为R_{jk}＝245+(245-0)*S_{jk_index}，G_{jk}＝245+(245-0)*S_{jk_index}，B_{jk}＝255+(255-245)*S_{jk_index}。

[0190] 对于情感概率为空的立场词块，其颜色为白色(255,255,255)。

[0191] 如图3所示，为本发明具体实施例中单个用户的用户画像可视化效果示意图。示例图为一个用户对立场触发词库的情绪反应，例如可以将红色系作为立场为支持或正面情感的代表色，其最强值的颜色为(255,0,0)，最弱值的RGB颜色为(245,245,245)；蓝色系作为立场为反对或负面情感的代表色，其最强值的颜色为(0,0,255)，最弱值的RGB颜色为(245,245,245)；情绪反应缺失的颜色为白色(255,255,255)。

[0192] 此外还可以进行不同用户画像的比较以及群体用户画像分析。

[0193] 对于不同用户画像的比较

[0194] 选取两个用户，分析两个用户对于每个立场词块的情绪反应，如果两个用户的情绪反应一致，则使用一种颜色进行表示，如果不一致则使用另一种颜色进行表示。依据一致程度选取同一色系的过渡色进行展示。对于其中一方没有情绪反应、或两方都没有情绪反应的立场词块，用白色表示。具体可以参见图4，其中的单个用户对立场触发词的情感概率与色彩特征之间的对应关系可以采用上述类似的方式实现，在此不再赘述。示例图为两个用户对立场触发词库的情绪反应比较，例如，情绪反应一致为绿色系，其最强值的RGB颜色为(0,201,13)，最弱值的RGB颜色为(245,245,245)；情绪反应不一致为紫色系，其最强值的RGB颜色为(115,9,170)，最弱值的颜色为(245,245,245)；情绪反应缺失的颜色为白色(255,255,255)

[0195] 群体用户画像分析

[0196] 对于某个用户群体，统计每个立场词块的平均情绪反应，统计时进行正向或负向的数量统计，不进行数值平均。对于超过u%(u为预设值)的用户有情绪反应的立场词块，统计每个立场词块正向或负向情绪所占用户总数的百分比，一致性越高，则呈现一种颜色，一致性越低则呈现另一种颜色，以相应的过渡色反应用户群体观点的一致性的程度。如果对于某立场词块有情绪反应的用户数量少于u%，则以白色表示。具体参见图5，其中的单个用户对立场触发词的情感概率与色彩特征之间的对应关系可以采用上述类似的方式实现，在此不再赘述。

[0197] 图5示为用户群体对立场触发词库的情绪反应比较，例如，情绪反应一致为绿色系，其最强值的颜色为(0,201,13)，最弱值的RGB颜色为(245,245,245)；情绪反应不一致为

紫色系,其最强值的颜色为(115,9,170),最弱值的RGB颜色为(245,245,245);情绪反应缺失的颜色为白色(255,255,255)。

[0198] 由于本发明提供的方法,可以根据短文本语料分类结果中词语的分布情况,构建情感词库,构建的情感词库中根据情感倾向将其分为正面情感词和负面情感词,可以更好的从用户的文字内容中识别新词的情感倾向、较为准确地分析其中的情绪反应,并构建立场触发词库,然后建立单个用户与立场触发词库的映射,计算出单个用户对应的立场触发词的情感概率;再计算由所有用户组成的社群内立场触发词的平均情感概率,并进行排序,从而可以根据排序结果进行用户画像的构建,从而可以准确对每个用户对立场触发词的情绪立场,从而可以快速准确地了解每个用户对普遍事物的观点差异和情绪反应特点,解决了现有技术中的方法存在分析结果不够准确的技术问题。

[0199] 进一步地,基于构建的用户画像,本发明还提供了一种用户画像的可视化方法,对用户画像进行可视化显示,提高了直观性。

[0200] 进一步地,根据颜色渐变公式,构建单个用户对立场触发词的情感概率与色彩特征之间的对应关系,然后根据情感概率与色彩特征之间的对应关系,对用户画像进行可视化显示,可以改善显示效果。

[0201] 实施例四

[0202] 请参见图6,基于同一发明构思,本申请还提供了一种计算机可读存储介质300,其上存储有计算机程序311,该程序被执行时实现如实施例一中所述的方法。

[0203] 由于本发明实施例三所介绍的计算机可读存储介质为实施本发明实施例一中基于用户情绪立场的用户画像建立方法所采用的计算机设备,故而基于本发明实施例一所介绍的方法,本领域所属人员能够了解该计算机可读存储介质的具体结构及变形,故而在此不再赘述。凡是本发明实施例一中方法所采用的计算机可读存储介质都属于本发明所欲保护的范围。

[0204] 本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0205] 本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0206] 尽管已描述了本发明的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

[0207] 显然,本领域的技术人员可以对本发明实施例进行各种改动和变型而不脱离本发明实施例的精神和范围。这样,倘若本发明实施例的这些修改和变型属于本发明权利要求

及其等同技术的范围之内，则本发明也意图包含这些改动和变型在内。

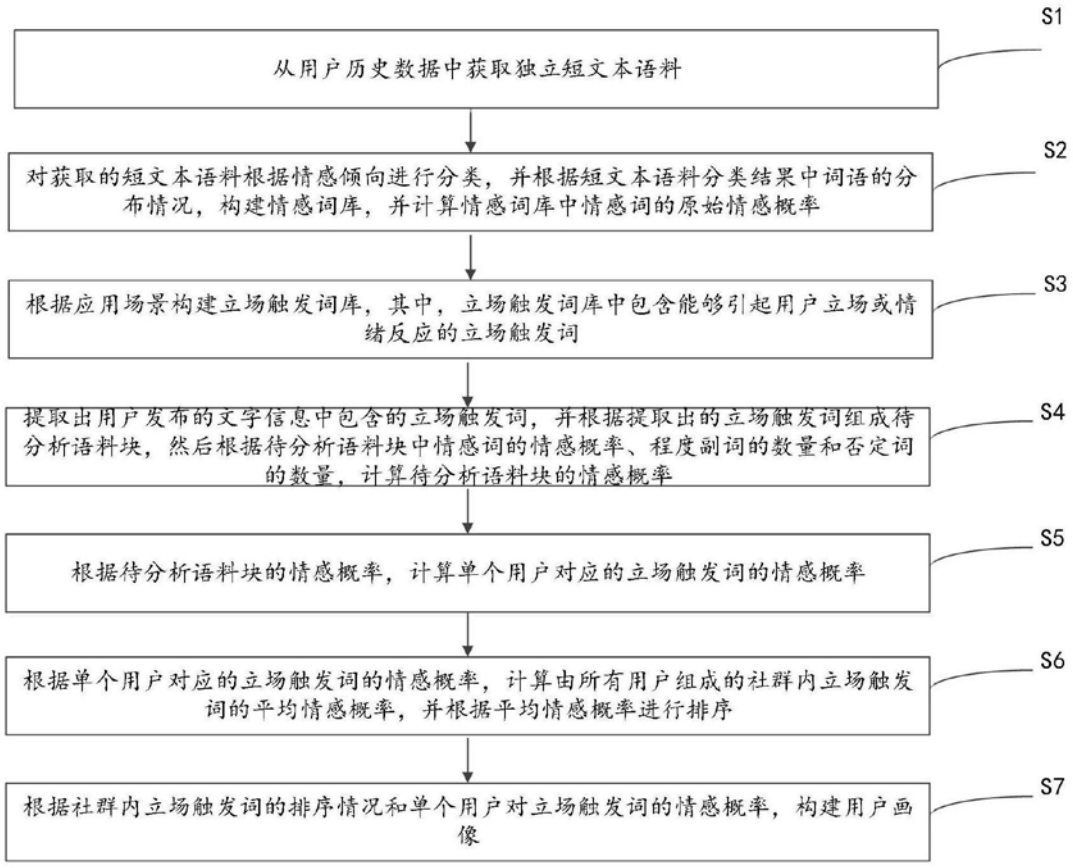

S1
从用户历史数据中获取独立短文本语料
S2
对获取的短文本语料根据情感倾向进行分类，并根据短文本语料分类结果中词语的分布情况，构建情感词库，并计算情感词库中情感词的原始情感概率
S3
根据应用场景构建立场触发词库，其中，立场触发词库中包含能够引起用户立场或情绪反应的立场触发词
S4
提取出用户发布的文字信息中包含的立场触发词，并根据提取出的立场触发词组成待分析语料块，然后根据待分析语料块中情感词的情感概率、程度副词的数量和否定词的数量，计算待分析语料块的情感概率
S5
根据待分析语料块的情感概率，计算单个用户对应的立场触发词的情感概率
S6
根据单个用户对应的立场触发词的情感概率，计算由所有用户组成的社群内立场触发词的平均情感概率，并根据平均情感概率进行排序
S7
根据社群内立场触发词的排序情况和单个用户对立场触发词的情感概率，构建用户画像

图1

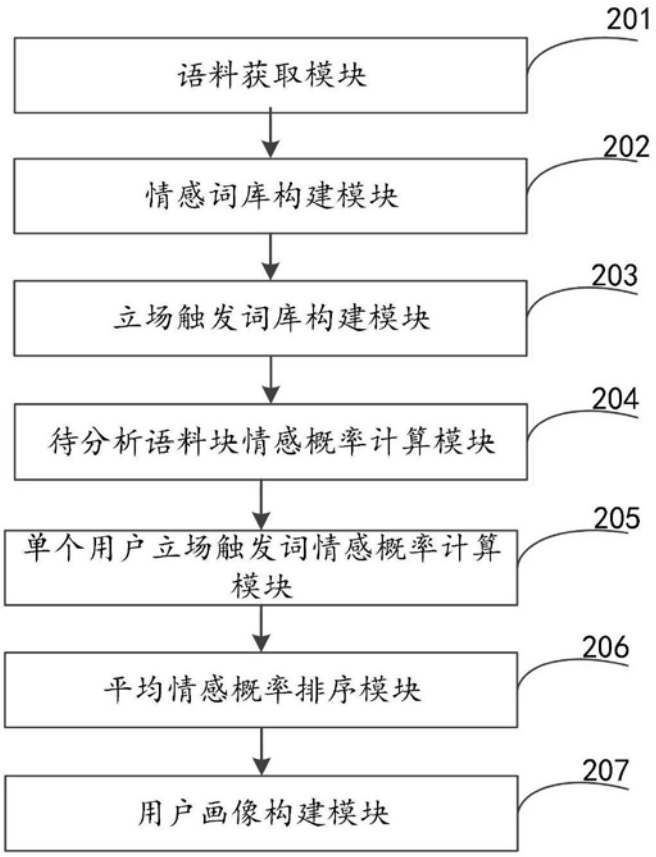

201
语料获取模块
202
情感词库构建模块
203
立场触发词库构建模块
204
待分析语料块情感概率计算模块
205
单个用户立场触发词情感概率计算模块
206
平均情感概率排序模块
207
用户画像构建模块

图2

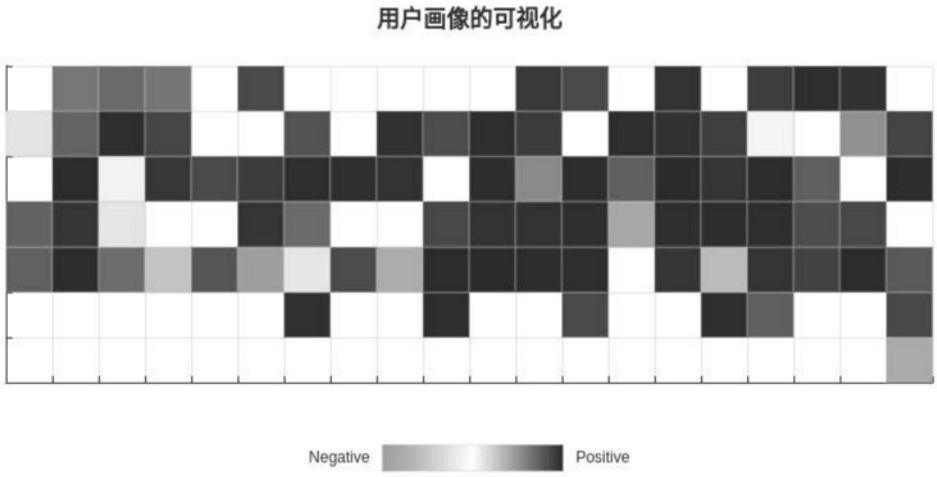

用户画像的可视化
Negative
Positive

图3

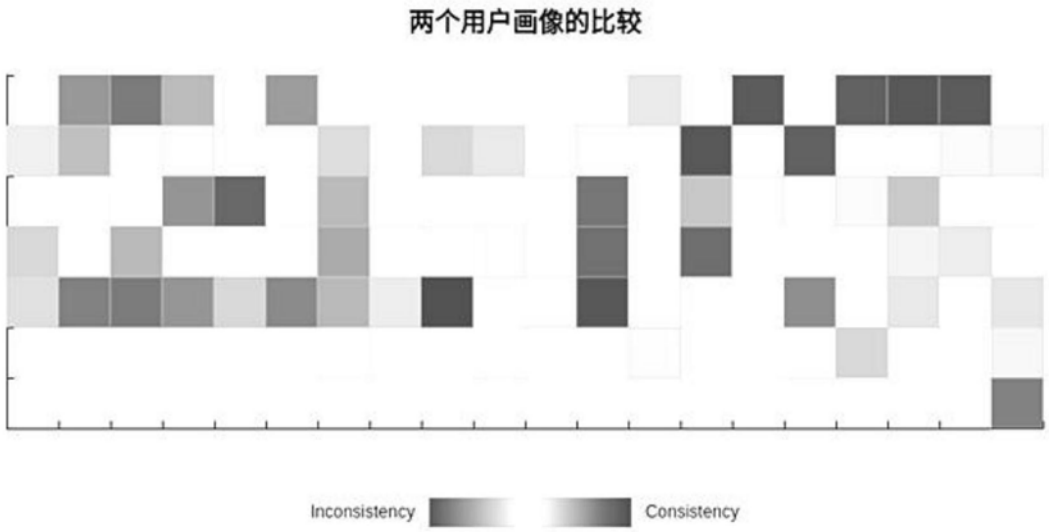

两个用户画像的比较
Inconsistency
Consistency

图4

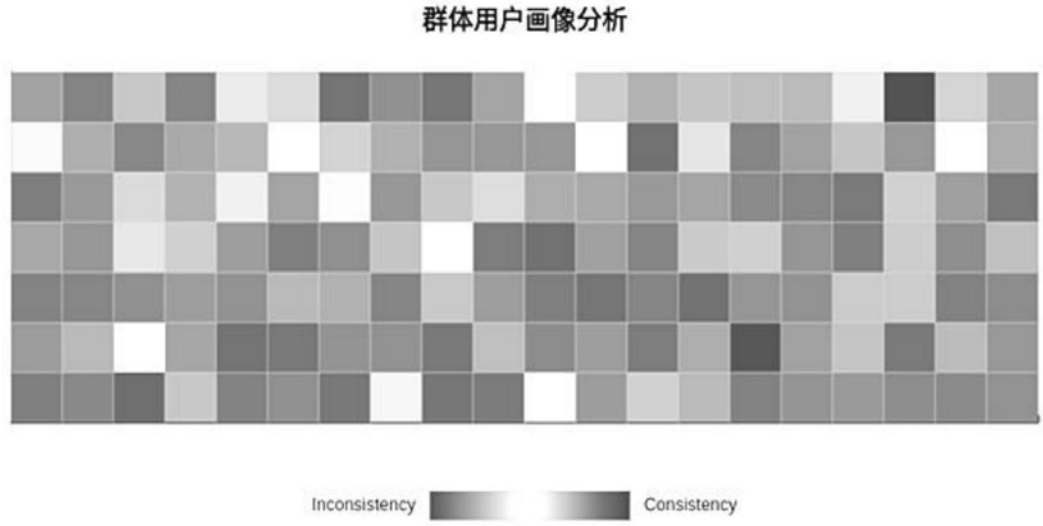

群体用户画像分析
Inconsistency
Consistency

图5

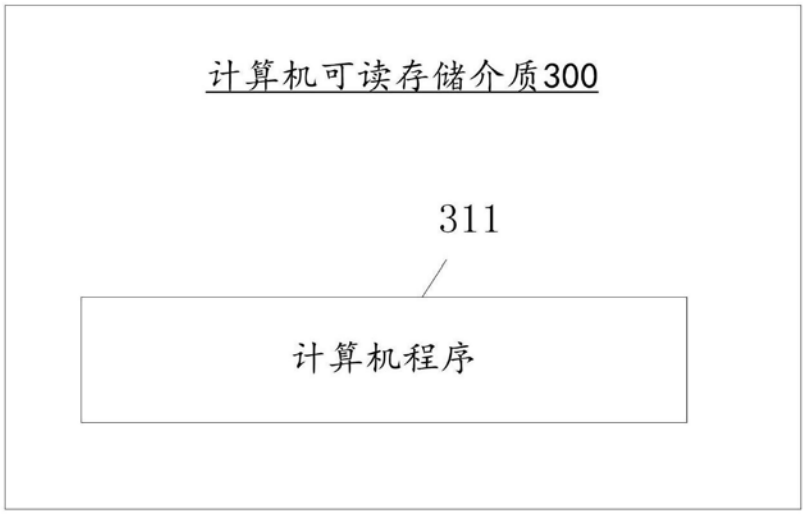


图6