



(12) 发明专利申请

(10) 申请公布号 CN 111914172 A

(43) 申请公布日 2020. 11. 10

(21) 申请号 202010742956.4

(22) 申请日 2020.07.29

(71) 申请人 上海梅斯医药科技有限公司  
地址 201612 上海市松江区漕河泾开发区  
松江高科技园莘砖公路258号34幢  
1801室

(72) 发明人 张发宝 李欣梅

(74) 专利代理机构 上海硕力知识产权代理事务  
所(普通合伙) 31251  
代理人 刘秋香

(51) Int. Cl.  
G06F 16/9535 (2019.01)  
G06F 16/9536 (2019.01)

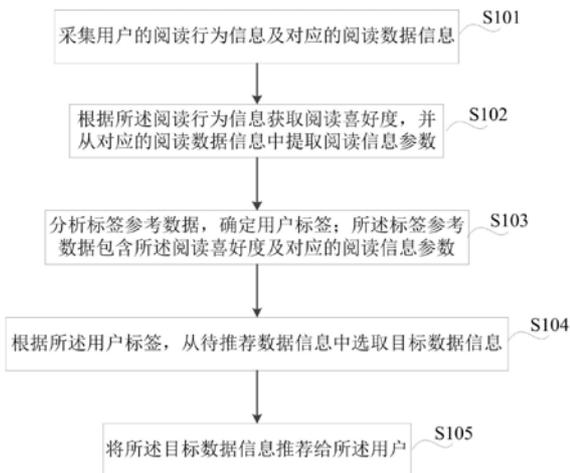
权利要求书3页 说明书11页 附图4页

(54) 发明名称

一种基于用户标签的医学信息推荐方法及系统

(57) 摘要

本发明公开了一种基于用户标签的医学信息推荐方法及系统,其中,智能医学信息推荐方法包括:采集用户在医学网站的阅读行为信息及对应的阅读数据信息;根据阅读行为信息获取阅读喜好度,并从对应的阅读数据信息中提取阅读信息参数;分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;根据用户标签,从待推荐医学数据信息中选取目标医学数据信息;将目标医学数据信息推荐给用户。通过本发明,可以基于用户标签向用户推送其感兴趣的信息,实现信息的精准推送。



1. 一种基于用户标签的医学信息推荐方法,其特征在于,包括:  
采集用户在医学网站的阅读行为信息及对应的阅读数据信息;  
根据所述阅读行为信息获取阅读喜好度,并从对应的阅读数据信息中提取阅读信息参数;  
分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;  
根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;  
将所述目标医学数据信息推荐给所述用户。
2. 根据权利要求1所述的一种基于用户标签的医学信息推荐方法,其特征在于,所述根据所述阅读行为信息获取阅读喜好度具体包括:  
根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度。
3. 根据权利要求1所述的一种基于用户标签的医学信息推荐方法,其特征在于,还包括:  
获取所述用户在所述医学网站的好友信息;  
从所述用户的好友中选取亲密度最高的好友作为目标好友;  
将所述目标好友浏览过且点赞的医学数据信息作为目标医学数据信息;  
将所述目标医学数据信息推荐给所述用户。
4. 根据权利要求1所述的一种基于用户标签的医学信息推荐方法,其特征在于,  
所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项。
5. 根据权利要求1所述的一种基于用户标签的医学信息推荐方法,其特征在于,所述分析标签参考数据,确定用户标签;具体包括:  
对从所述阅读数据信息中提取的各阅读信息参数进行遍历组合,获得参数分析组;每个参数分析组至少包含一个阅读信息参数;  
根据所述用户对所述阅读数据信息的喜好度,基于所述阅读数据信息,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;  
选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数作为目标阅读信息参数;  
获取所述目标参数分析组中的各目标阅读信息参数的影响权重;  
根据所述目标参数分析组中的各目标阅读信息参数的影响权重,确定所述用户的标签。
6. 根据权利要求1-5任意一项所述的一种基于用户标签的医学信息推荐方法,其特征在于,所述标签参考数据还包括所述用户的注册信息;所述注册信息包括昵称、年龄、专业类别、工作领域、医院科室、兴趣爱好之中的任意一项或多项。
7. 根据权利要求6所述的一种基于用户标签的医学信息推荐方法,其特征在于,所述分析标签参考数据,确定用户标签;具体包括:  
对从所述阅读数据信息中提取的各阅读信息参数、及所述用户的各注册信息进行遍历组合,获得参数分析组;每个参数分析组包含至少一个阅读信息参数和至少一项注册信息;

根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数及各注册信息作为目标参数;

获取所述目标参数分析组中的各目标参数的影响权重;

根据所述各目标参数的影响权重,确定所述用户的标签。

8. 一种基于用户标签的医学信息推荐系统,其特征在于,包括:

阅读采集模块,用于采集用户在医学网站的阅读行为信息及对应的阅读数据信息;

信息处理模块,用于根据所述阅读行为信息获取阅读喜好度,并从对应的阅读数据信息中提取阅读信息参数;

标签确定模块,用于分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;

信息选取模块,用于根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;

信息推荐模块,用于将所述目标医学数据信息推荐给所述用户。

9. 根据权利要求8所述的一种基于用户标签的医学信息推荐系统,其特征在于,

所述信息处理模块,根据所述阅读行为信息获取阅读喜好度具体包括:

所述信息处理模块,根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度。

10. 根据权利要求8所述的一种基于用户标签的医学信息推荐系统,其特征在于,还包括:

好友确定模块,用于获取所述用户在所述医学网站的好友信息;并从所述用户的好友中选取亲密度最高的好友作为目标好友;

所述信息选取模块,还用于将所述目标好友浏览过且点赞的医学数据信息作为目标医学数据信息;并通过所述信息推荐模块将所述目标医学数据信息推荐给所述用户。

11. 根据权利要求8所述的一种基于用户标签的医学信息推荐系统,其特征在于,所述标签确定模块具体包括:

组合子模块,用于对从所述阅读数据信息中提取的各阅读信息参数进行遍历组合,获得参数分析组;每个参数分析组至少包含一个阅读信息参数;

模型分析子模块,用于根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

选取子模块,用于选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数作为目标阅读信息参数;

所述模型分析子模块,还用于获取所述目标参数分析组中的各目标阅读信息参数的影响权重;

标签确定子模块,用于根据所述目标参数分析组中的各目标阅读信息参数的影响权重,确定所述用户的标签。

12. 根据权利要求8-10任意一项所述的一种基于用户标签的医学信息推荐系统,其特征在于,

所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项；

所述标签参考数据还包括所述用户的注册信息；所述注册信息包括昵称、年龄、专业类别、工作领域、医院科室、兴趣爱好之中的任意一项或多项。

13. 根据权利要求12所述的一种基于用户标签的医学信息推荐系统，其特征在于，所述标签确定模块具体包括：

组合子模块，用于对从所述阅读数据信息中提取的各阅读信息参数、及所述用户的各注册信息进行遍历组合，获得参数分析组；每个参数分析组包含至少一个阅读信息参数和至少一项注册信息；

模型分析子模块，用于根据所述阅读数据信息，及用户对所述阅读数据信息的喜好度，对所有的参数分析组进行回归分析，获得每个参数分析组的联合影响权重；

选取子模块，用于选取联合影响权重最大的参数分析组作为目标参数分析组，并将所述目标参数分析组中包含的各阅读信息参数及各注册信息作为目标参数；

所述模型分析子模块，还用于获取所述目标参数分析组中的各目标参数的影响权重；

标签确定子模块，用于根据所述各目标参数的影响权重，确定所述用户的标签。

## 一种基于用户标签的医学信息推荐方法及系统

### 技术领域

[0001] 本发明涉及数据处理分析领域,尤其涉及一种基于用户标签的医学信息推荐方法及系统。

### 背景技术

[0002] 医学专业网站上往往包含有大量的医学数据信息,由于医学数据信息的专业性,不同的医护人员可能感兴趣的医学信息也不一样。且医学网站的浏览信息,不同于新闻、娱乐、等商业网站,有反复浏览、反复查看、针对某一词汇反复检索、收藏后反复查看、浏览时间偏长等等特征。商业网站、论坛常用的推荐算法,用于医学网站,推荐不准确。比如,目前,常使用的推荐算法有基于内容推荐法、协同过滤推荐法以及混合推荐法,也包含一些改进算法,例如,基于维度划分为用户维度相关改进法、产品维度相关改进法等这类现有的信息推送方法在向用户推送医学信息时,经常存在推送的信息不准确的情况,导致用户的使用体验较低,亟需改进。

### 发明内容

[0003] 为了克服上述技术缺陷,本申请公开了一种基于用户标签的医学信息推荐方法及系统,具体的,本申请的技术方案如下:

[0004] 一方面,本发明公开了一种基于用户标签的医学信息推荐方法,包括:

[0005] 采集用户在医学网站的阅读行为信息及对应的阅读数据信息;

[0006] 根据所述阅读行为信息获取阅读喜好度,并从对应的阅读数据信息中提取阅读信息参数;

[0007] 分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;

[0008] 根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;

[0009] 将所述目标医学数据信息推荐给所述用户。

[0010] 优选地,所述根据所述阅读行为信息获取阅读喜好度具体包括:根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度。

[0011] 优选地,所述基于用户标签的医学信息推荐方法还包括:获取所述用户在所述医学网站的好友信息;从所述用户的好友中选取亲密度最高的好友作为目标好友;将所述目标好友浏览过且点赞的医学数据信息作为目标医学数据信息;将所述目标医学数据信息推荐给所述用户。

[0012] 优选地,所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项。

[0013] 优选地,所述分析标签参考数据,确定用户标签;具体包括:

[0014] 对从所述阅读数据信息中提取的各阅读信息参数进行遍历组合,获得参数分析

组;每个参数分析组至少包含一个阅读信息参数;

[0015] 根据所述用户对所述阅读数据信息的喜好度,基于所述阅读数据信息,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0016] 选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数作为目标阅读信息参数;

[0017] 获取所述目标参数分析组中的各目标阅读信息参数的影响权重;

[0018] 根据所述目标参数分析组中的各目标阅读信息参数的影响权重,确定所述用户的标签。

[0019] 优选地,所述标签参考数据还包括所述用户的注册信息;所述注册信息包括昵称、年龄、专业类别、工作领域、医院科室、兴趣爱好之中的任意一项或多项。

[0020] 优选地,所述分析标签参考数据,确定用户标签;具体包括:

[0021] 对从所述阅读数据信息中提取的各阅读信息参数、及所述用户的各注册信息进行遍历组合,获得参数分析组;每个参数分析组包含至少一个阅读信息参数和至少一项注册信息;

[0022] 根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0023] 选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数及各注册信息作为目标参数;

[0024] 获取所述目标参数分析组中的各目标参数的影响权重;

[0025] 根据所述各目标参数的影响权重,确定所述用户的标签。

[0026] 另一方面,本发明还公开了一种基于用户标签的医学信息推荐系统,包括:

[0027] 阅读采集模块,用于采集用户在医学网站的阅读行为信息及对应的阅读数据信息;

[0028] 信息处理模块,用于根据所述阅读行为信息获取阅读喜好度,并从对应的阅读数据信息中提取阅读信息参数;

[0029] 标签确定模块,用于分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;

[0030] 信息选取模块,用于根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;

[0031] 信息推荐模块,用于将所述目标医学数据信息推荐给所述用户。

[0032] 优选地,所述信息处理模块,根据所述阅读行为信息获取阅读喜好度具体包括:

[0033] 所述信息处理模块,根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度。

[0034] 优选地,所述基于用户标签的医学信息推荐系统还包括:好友确定模块,用于获取所述用户在所述医学网站的好友信息;并从所述用户的好友中选取亲密度最高的好友作为目标好友;所述信息选取模块,还用于将所述目标好友浏览过且点赞的医学数据信息作为目标医学数据信息;并通过所述信息推荐模块将所述目标医学数据信息推荐给所述用户。

[0035] 优选地,所述标签确定模块具体包括:

[0036] 组合子模块,用于对从所述阅读数据信息中提取的各阅读信息参数进行遍历组合,获得参数分析组;每个参数分析组至少包含一个阅读信息参数;

[0037] 模型分析子模块,用于根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0038] 选取子模块,用于选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数作为目标阅读信息参数;

[0039] 所述模型分析子模块,还用于获取所述目标参数分析组中的各目标阅读信息参数的影响权重;

[0040] 标签确定子模块,用于根据所述目标参数分析组中的各目标阅读信息参数的影响权重,确定所述用户的标签。

[0041] 优选地,所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项;所述标签参考数据还包括所述用户的注册信息;所述注册信息包括昵称、年龄、专业类别、工作领域、医院科室、兴趣爱好之中的任意一项或多项。

[0042] 优选地,所述标签确定模块具体包括:

[0043] 组合子模块,用于对从所述阅读数据信息中提取的各阅读信息参数、及所述用户的各注册信息进行遍历组合,获得参数分析组;每个参数分析组包含至少一个阅读信息参数和至少一项注册信息;

[0044] 模型分析子模块,用于根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0045] 选取子模块,用于选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数及各注册信息作为目标参数;

[0046] 所述模型分析子模块,还用于获取所述目标参数分析组中的各目标参数的影响权重;

[0047] 标签确定子模块,用于根据所述各目标参数的影响权重,确定所述用户的标签。

[0048] 本发明至少包含以下一项技术效果:

[0049] (1) 本申请从每次的阅读行为信息中获取对阅读的数据信息的喜好度,然后从该阅读数据信息中提取阅读信息参数,通过分析大量地阅读数据的喜好度和对应的阅读信息参数,确定用户的标签,知道哪些阅读信息参数对该用户的阅读行为有很大影响,从而在后续待推荐的数据信息中,选取与该用户匹配的医学数据信息作为目标医学数据信息,推荐给用户,实现医学数据信息的精准推送。

[0050] (2) 本申请中的阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源等,从多个方面和维度对阅读信息进行了考量,细化了对用户阅读行为可能产生影响的参数,为后续的目标数据推荐提供了重要参考依据。

[0051] (3) 本申请中在用户阅读行为信息和阅读数据信息的基础上,还增加了用户注册信息,从多个方面来分析,确定用户的标签。如此,可使得推荐的数据信息更加贴合用户实际,也不会太单一。

[0052] (4) 本申请中对各阅读信息参数进行了分析,由于用户的阅读行为可能受一个或多个因素的影响,因此采用回归分析法可以准确分析出一个或多个阅读信息参数对用户阅

读行为(阅读喜好度)的影响,最后选取联合影响最大的参数分析组作为目标,从而获得对用户喜好度影响较大的参数因子及对应的权重,从而可准确的确定用户的标签。

[0053] (5)本申请中,除了基于用户自身的标签进行医学信息推荐外,还会将亲密度最高的好友浏览过的且点赞过的医学数据信息推荐给该用户,提高医学信息数据推荐的质量、贴合度、以及拓展推荐宽度。

### 附图说明

[0054] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简要介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域的普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0055] 图1为本申请智能医学信息推荐方法的一个实施例的流程图;

[0056] 图2为本申请智能医学信息推荐方法的另一实施例的流程图;

[0057] 图3为本申请智能医学信息推荐系统的一个实施例的结构框图;

[0058] 图4为本申请智能医学信息推荐系统的另一实施例的结构框图。

### 具体实施方式

[0059] 为了使本发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明作进一步地详细描述,显然,所描述的实施例仅仅是本发明一部份实施例,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做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0060] 本发明公开了一种基于用户标签的医学信息推荐方法,具体的,实施例一如图1所示,包括:

[0061] S101,采集用户在医学网站的阅读行为信息及对应的阅读数据信息;

[0062] 具体的,阅读行为信息包括用户的阅读点击行为(用户是否点进去浏览阅读)、以及点击阅读后的阅读时长、阅读点赞信息等等;阅读数据信息则是指用户阅读的具体内容相关的信息,比如阅读的某篇医学文章的文章内容、字数、标题等等。

[0063] S102,根据所述阅读行为信息获取阅读喜好度,并从对应的阅读数据信息中提取阅读信息参数;

[0064] 具体的,比如根据用户是否点进去看,以及点进去后的浏览时长、是否点赞、打赏等阅读行为信息来确定用户对当前阅读的数据信息的喜好度。比如,用户点进去看,且浏览的时间比较长,并有点赞或打赏行为,那么则说明用户对该阅读数据信息非常喜欢,非常感兴趣。而如果用户浏览到标题后,根本没有点进去看,或者点进去后阅读的时常非常短,则从侧面说明该阅读数据信息并不是用户感兴趣的内容。而从对应的阅读数据信息中提取阅读信息参数,具体的阅读信息参数包括标题长度、信息类别、点击量(浏览量)、点赞数、全文总字数、信息来源之中的任意一项或多项。这些都可能是影响到用户阅读行为的因素。

[0065] S103,分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;

[0066] 具体的,通过分析包含阅读喜好度及对应的阅读信息参数的标签参考数据,看影

响到用户阅读行为的是哪些阅读信息参数,各自影响的比重是多少。比如通过分析发现,用户喜欢浏览简短且浏览量较大的医学研发类文章,那么在阅读信息参数中,点击量、全文总字数、及文章类别是影响用户喜好度最大的几个因素,也就是说点击量越大、全文总字数越短的医学研发类文章则更受该用户的喜欢。由此,可以给用户打上标签:全文简短、点击量大、关注医学研发。

[0067] S104,根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;

[0068] 具体的,获取到用户标签后,便可从大量待推荐医学数据信息中选取目标医学数据信息,比如根据上述的用户名标签,从待推荐的数据信息中选取点击量大、且全文简短的医学研发类数据信息作为目标医学数据信息。

[0069] 较佳的,还可进一步获取各目标医学数据信息与该用户标签的相关度,根据相关度来对选取出的目标医学数据信息进行排序,进而择优推荐。比如推荐排序靠前的预设数量的目标医学数据信息。而关于目标医学数据信息与用户标签的相关度,则可根据该目标医学数据信息的阅读信息参数,以及各阅读信息参数影响该用户阅读行为的比重(权重),进而根据预设的计算公式计算出该目标医学数据信息与该用户标签的相关度。

[0070] S105,将所述目标医学数据信息推荐给所述用户。

[0071] 具体的,这些目标医学数据信息便是用户可能想去浏览阅读的内容,因此,确定好目标医学数据信息后,便可将这些目标医学数据信息推荐给相应的用户,从而让用户可以浏览到更多感兴趣的内容,达到精准推送的效果。

[0072] 通过上述实施例,可以通过用户阅读行为和对应的阅读数据信息,确定用户标签,再基于用户标签向用户推送其感兴趣的信息,实现信息的精准推送。

[0073] 本发明的另一实施例,如图2所示,包括:

[0074] S201,采集用户在医学网站的阅读行为信息及对应的阅读数据信息;

[0075] S202,根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度;

[0076] 具体的,根据用户是否点进去浏览可以初步判断用户是否感兴趣,再根据用户点进去后的阅读时长及点赞信息可以进一步判断用户的感兴趣程度。当然,由于不同的文章,全文字数不同,因此,阅读时长本身便会存在差异,对此,我们可进一步获取阅读的数据信息的长度(全文总字数),基于用户的阅读速度,估算所述用户阅读该数据信息的参考时长,再将所述用户阅读该数据信息的实际阅读时长与参考阅读时长进行比较,若实际阅读时长超出参考阅读时长,则进一步说明用户比较感兴趣,且超出的时长越长,则感兴趣的程度越高。此外,用户是否点赞也可以作为用户是否感兴趣的一个重要参考因素。

[0077] 医学网站的医学信息,不同于新闻、娱乐、等商业网站的数据信息,医学网站的专业信息有反复浏览、反复查看、针对某一词汇反复检索、收藏后反复查看、浏览时间偏长等等特征,这些特征是专业信息所独有的,一般的娱乐新闻或商业新闻往往一看而过,不会存在反复查看浏览的情况。

[0078] S203,从对应的阅读数据信息中提取阅读信息参数;

[0079] 具体的,前面我们分析了用户的喜爱度,也就是说用户是否感兴趣及感兴趣的程度,那么后面则需要的分析是哪些因素影响到用户的阅读行为,也就是用户的喜爱度,以及影响的比重是多少。鉴于此,我们从阅读数据信息中提取阅读信息参数,这些参数都可能是

影响到用户阅读行为的,不同的用户,各因素影响的权重不同。具体的,所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项。

[0080] S204,对从所述阅读数据信息中提取的各阅读信息参数进行遍历组合,获得参数分析组;每个参数分析组至少包含一个阅读信息参数;

[0081] 具体的,比如,我们以提取的阅读信息参数为信息类别、点击量及全文字数为例,对这三个阅读信息参数进行遍历组合,获得各参数分析组如下表所示:

[0082] 表1-阅读信息参数组合表

阅读信息 参数 组别	信息类别	点击量	全文字数
A 组	√		
B 组		√	
C 组			√
D 组	√	√	
E 组	√		√
F 组		√	√
G 组	√	√	√

[0084] S205,根据所述用户对所述阅读数据信息的喜好度,基于所述阅读数据信息,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0085] 根据所述用户对所述阅读数据信息的喜好度及对应的阅读数据信息,对每个参数分析组进行回归分析,看看哪组对用户的喜好的联合影响权重最大。比如,上表1中的A组、B组、C组中由于只包含了一个阅读信息参数,则只需分析其包含的阅读信息参数对用户喜好度的影响,而D组、E组、F组均包含了两个阅读信息参数,则分析这两个阅读信息参数共同作用下对用户喜好度的影响,而对于G组包含了三个阅读信息参数,则分析这三个阅读信息参数共同作用下对用户喜好度的影响。

[0086] 具体的,如何分析这些阅读信息参数对用户喜好度的影响呢,本实施例采用了回归分析的方法,该回归分析指在相关变量中将一个变量视为因变量,其他一个或多个变量视为自变量,建立多个变量之间线性或非线性数学模型数量关系式并利用样本数据进行分析的统计分析方法。

[0087] 本实施例采用的回归分析的运行结果会包括:基本统计量、回归方程、均值多重比较。其中均值多重比较这部分结果是可以看出,各因素各自对目标的影响大小(量化结果)、各因素共同作用下对对目标的影响,这个回归分析和均值多重比较的结果,可以大大提高对信息推荐的准确度。

[0088] 应用在本实施例中,则是将用户的对阅读数据信息的喜好度作为因变量,将参数

分析组中的各阅读信息参数作为自变量进行分析,建立各阅读信息参数对用户喜好度的影响模型,通过该模型,则可以确定出该参数分析组中的各阅读信息参数共同作用下对该用户的喜好度的影响权重,以及各协同作用的阅读信息参数所起到的权重比值。

[0089] S206,选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数作为目标阅读信息参数;

[0090] 具体的,本步骤中选取联合影响权重最大的参数分析组作为目标参数分析组,也就是选取出了确定用户喜爱度的主要阅读信息参数。

[0091] S207,获取所述目标参数分析组中的各目标阅读信息参数的影响权重;

[0092] 具体的,获取到目标参数分析组后,可进一步获取到该目标参数分析组中的各目标阅读信息参数的影响权重,后续便可从待推荐的数据信息中提取目标阅读信息参数的值进行计算,获得该用户的喜爱度打分,从而确定是否是用户感兴趣的数据信息并是否进行推荐。

[0093] S208,根据所述各阅读信息参数对所述用户的喜好度的各自影响权重及共同影响情况,确定所述用户的标签。

[0094] 具体的,前一步骤中,我们确定了目标参数分析组中各目标阅读信息参数的影响权重,从而便可建立对应的模型或公式来作为或者来确定用户的标签。

[0095] S209,根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;

[0096] 具体的,确定了用户的标签,便可知道用户喜欢阅读哪种类型的数据信息。比如向用户推荐阅读标题长(超过预设值),生物制药领域的资讯,且在这些资讯中根据点击量和点赞量综合考虑推荐排序,将点击量或点赞量更大的生物制药资讯排在前面,作为主要目标医学数据信息。通过用户标签中各目标阅读信息参数的影响权重,可计算出用户对该篇数据信息的喜爱度打分,进而便可根据打分的高低来进行排序推荐。

[0097] S210,将所述目标医学数据信息推荐给所述用户。

[0098] 本发明方法的另一实施例,在上述任一实施例的基础上,标签参考数据除了包含阅读喜好度和阅读信息参数外,还包括该用户的注册信息;具体的,注册信息包括昵称、年龄、专业类别、工作领域、医院科室、兴趣爱好之中的任意一项或多项。本发明实施例流程如下,包括:

[0099] S301,采集用户在医学网站的阅读行为信息及对应的阅读数据信息;

[0100] S302,根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度;

[0101] S303,从对应的阅读数据信息中提取阅读信息参数;

[0102] 具体的,所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项;

[0103] S304,对从所述阅读数据信息中提取的各阅读信息参数、及所述用户的各注册信息进行遍历组合,获得参数分析组;每个参数分析组包含至少一个阅读信息参数和至少一项注册信息;

[0104] S305,根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0105] S306,选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标

参数分析组中包含的各阅读信息参数及各注册信息作为目标参数；

[0106] S307, 获取所述目标参数分析组中的各目标参数的影响权重；

[0107] S308, 根据所述各目标参数的影响权重, 确定所述用户的标签。

[0108] 具体的, 本实施例在上述实施例的基础上, 并不是像上一实施例一样仅根据回归分析的各目标阅读信息参数对阅读行为影响权重来确定用户标签, 而是还结合参考了用户注册信息来进行分析, 确定用户标签, 如此, 一方面, 从多个维度来综合考量, 使得分析出的用户标签更加准确。

[0109] S309, 根据所述用户标签, 从待推荐医学数据信息中选取目标医学数据信息；

[0110] S310, 将所述目标医学数据信息推荐给所述用户。

[0111] 本实施例中, 增加了了用户名的注册信息来辅助确定用户标签, 当然, 该确定的用户标签并不是一成不变的, 还可定期更新用户标签, 鉴于仅根据用户的历史阅读行为信息及阅读信息数据信息来推荐目标医学数据信息, 可能存在推荐的数据信息较为片面单一的情况, 而进一步结合了用户的注册信息的话, 则对用户的定位更准确、更全面。

[0112] 本申请的另一个方法实施例, 在上述任一方法实施例的基础上, 增加了基于好友维度来进行医学信息推荐的方式, 具体的, 依据好友信息进行医学信息推荐包括:

[0113] S401, 获取所述用户在所述医学网站的好友信息；

[0114] S402, 从所述用户的好友中选取亲密度最高的好友作为目标好友；

[0115] S403, 将所述目标好友浏览过且点赞的医学数据信息作为目标医学数据信息；

[0116] S404, 将所述目标医学数据信息推荐给所述用户。

[0117] 本实施例中, 除了基于用户自身标签进行医学信息推荐外, 还可以基于用户所在的好友圈最关注的信息去推广所在圈内的用户。具体的, 获取该用户的好友信息, 根据其与其好友的亲密度来选取目标好友, 比如该用户与好友A经常会经常在论坛沟通交流; 都浏览过很多同样的医学数据信息; 其专业或兴趣爱好较为相近等等, 可以从多个维度考虑该用户与好友的亲密度。亲密度越高, 则该好友喜欢的医学信息数据越可能也是该用户感兴趣的。因此, 本实施例, 除了基于用户自身的标签进行医学信息推荐外, 还会将亲密度最高的好友浏览过的且点赞过的医学数据信息推荐给该用户, 提高医学信息数据推荐的质量、贴合度、以及拓展推荐宽度。

[0118] 基于相同的技术构思, 本发明还公开了一种基于用户标签的医学信息推荐系统, 具体的, 本系统的一个实施例, 如图3所示, 包括:

[0119] 阅读采集模块10, 用于采集用户在医学网站的阅读行为信息及对应的阅读数据信息; 具体的, 阅读行为信息包括用户的阅读点击行为(用户是否点进去浏览阅读)、以及点击阅读后的阅读时长、阅读点赞信息等等; 阅读数据信息则是指用户阅读的具体内容相关的信息, 比如阅读的某篇文章及该文章的相关信息。

[0120] 信息处理模块20, 用于根据所述阅读行为信息获取阅读喜好度, 并从对应的阅读数据信息中提取阅读信息参数; 具体的, 阅读喜好度是指对阅读的这篇数据信息的喜好程度, 不同的阅读数据信息, 用户对其阅读喜好度不同; 同一篇阅读数据信息, 不同的用户来阅读, 可能不同的用户对其的喜好度也不一样。对于阅读喜好度的确定, 一般可以从用户在医学网站的阅读行为信息中获取, 比如, 如果用户点进去阅读, 且阅读时长比该阅读数据信息的参考阅读时长要长不少, 则说明该用户对该阅读数据信息进行了精读, 可能反复读了

几遍,或者边读边琢磨,这从一定程度上反应给出用户对手该数据信息非常感兴趣。

[0121] 优选的,所述阅读信息参数包括标题长度、信息类别、点击量、点赞数、全文总字数、信息来源之中的任意一项或多项。

[0122] 进一步优选的,所述信息处理模块20,根据所述用户的阅读点击行为、点击阅读后的阅读时长、反复浏览次数、反复浏览时长、以及阅读点赞信息,获得所述用户对对应的阅读数据信息的喜好度。

[0123] 标签确定模块30,用于分析标签参考数据,确定用户标签;所述标签参考数据包含所述阅读喜好度及对应的阅读信息参数;

[0124] 具体的,通过分析包含阅读喜好度及对应的阅读信息参数的标签参考数据,看影响到用户阅读行为的是哪些阅读信息参数,各自影响的比重是多少,从而为该用户打上对应的标签。

[0125] 信息选取模块40,用于根据所述用户标签,从待推荐医学数据信息中选取目标医学数据信息;

[0126] 具体的,获取到用户标签后,便可从待推荐医学数据信息中选取该用户可能感兴趣的数据信息进行推荐。优选地,还可进一步获取各目标医学数据信息与该用户标签的相关度,根据相关度来对选取出的目标医学数据信息进行排序,进而择优推荐。比如推荐排序靠前的预设数量的目标医学数据信息。而关于目标医学数据信息与用户标签的相关度,则可根据该目标医学数据信息的阅读信息参数,以及各阅读信息参数影响该用户阅读行为的比重(权重),进而根据预设的计算公式计算出该目标医学数据信息与该用户标签的相关度。

[0127] 信息推荐模块50,用于将所述目标医学数据信息推荐给所述用户。

[0128] 具体的,比如可以在用户浏览该网站时,将目标医学数据信息放在用户浏览的首页,吸引用户的注意,提高用户的阅读兴趣,及对该网站的吸合度。

[0129] 本申请智能医学信息推荐系统的另一实施例,在上述系统实施例的基础上,如图4所示,所述标签确定模块具体包括:

[0130] 组合子模块31,用于对从所述阅读数据信息中提取的各阅读信息参数进行遍历组合,获得参数分析组;每个参数分析组至少包含一个阅读信息参数;

[0131] 模型分析子模块32,用于根据所述阅读数据信息,及用户对所述阅读数据信息的喜好度,对所有的参数分析组进行回归分析,获得每个参数分析组的联合影响权重;

[0132] 选取子模块33,用于选取联合影响权重最大的参数分析组作为目标参数分析组,并将所述目标参数分析组中包含的各阅读信息参数作为目标阅读信息参数;

[0133] 所述模型分析子模块32,还用于获取所述目标参数分析组中的各目标阅读信息参数的影响权重;

[0134] 标签确定子模块34,用于根据所述目标参数分析组中的各目标阅读信息参数的影响权重,确定所述用户的标签。

[0135] 具体的,如何分析这些阅读信息参数中哪些对用户喜好度影响大呢,或者哪些共同作用对用户喜好度的影响大呢,本实施例采用了回归分析的方法,比如可采用多因素回归分析的方法,该回归分析方法中将一个变量视为因变量,其他一个或多个变量视为自变量,建立多个变量之间线性或非线性数学模型数量关系式并利用样本数据进行分析的统计

分析方法。

[0136] 该回归分析的运行结果会包括：基本统计量、回归方程、均值多重比较。其中均值多重比较这部分结果是可以看出，各因素各自对目标的影响大小（量化结果）、多个因素因素协同作用下对目标的影响，这个回归分析和均值多重比较的结果，可以大大提高对信息推荐的准确度。

[0137] 应用在本实施例中，则是将用户的对阅读数据信息的喜好度作为因变量，将参数分析组中各阅读信息参数作为自变量进行分析，选取联合影响最大的参数分析组作为目标参数分析组，进而根据该目标参数分析组中各目标阅读信息参数对用户喜好度的影响，建立用户标签模型，通过该模型，则可以根据阅读数据信息中各目标阅读信息参数对该用户的喜好度的影响权重，计算出用户的喜好度评分，后续可据喜好度评分得出该篇阅读数据信息是否是该用户感兴趣的内容。

[0138] 本申请智能医学信息推荐系统的另一实施例，在上述第一个系统实施例的基础上，所述标签参考数据还包括所述用户的注册信息；所述注册信息包括昵称、年龄、专业类别、工作领域、医院科室、兴趣爱好之中的任意一项或多项。

[0139] 所述标签确定模块30具体包括：

[0140] 组合子模块31，用于对从所述阅读数据信息中提取的各阅读信息参数、及所述用户的各注册信息进行遍历组合，获得参数分析组；每个参数分析组包含至少一个阅读信息参数和至少一项注册信息；

[0141] 模型分析子模块32，用于根据所述阅读数据信息，及用户对所述阅读数据信息的喜好度，对所有的参数分析组进行回归分析，获得每个参数分析组的联合影响权重；

[0142] 选取子模块33，用于选取联合影响权重最大的参数分析组作为目标参数分析组，并将所述目标参数分析组中包含的各阅读信息参数及各注册信息作为目标参数；

[0143] 所述模型分析子模块32，还用于获取所述目标参数分析组中的各目标参数的影响权重；

[0144] 标签确定子模块34，用于根据所述各目标参数的影响权重，确定所述用户的标签。

[0145] 本实施例中，增加了了用户名的注册信息来辅助确定用户标签，当然，该确定的用户标签并不是一成不变的，还可定期更新用户标签，鉴于仅根据用户的历史阅读行为信息及阅读信息数据信息来推荐目标医学数据信息，可能存在推荐的数据信息较为片面单一的情况，而进一步结合了用户的注册信息的话，则对用户的定位更准确、更全面。

[0146] 本申请的最后一个实施例，在上述任一实施例的基础上，所述基于用户标签的医学信息推荐系统还包括：

[0147] 好友确定模块，用于获取所述用户在所述医学网站的好友信息；并从所述用户的好友中选取亲密度最高的好友作为目标好友；

[0148] 所述信息选取模块，还用于将所述目标好友浏览过且点赞的医学数据信息作为目标医学数据信息；并通过所述信息推荐模块将所述目标医学数据信息推荐给所述用户。

[0149] 本实施例中，除了基于用户自身标签进行医学信息推荐外，还获取该用户的好友信息，根据其的好友的亲密度来选取目标好友，比如该用户与好友A经常会经常在论坛沟通交流；都浏览过很多同样的医学数据信息；其专业或兴趣爱好较为相近等等，可以从多个维度考虑该用户与好友的亲密度。亲密度越高，则该好友喜欢的医学信息数据越可能也是该

用户感兴趣的。因此,本实施例,除了基于用户自身的标签进行医学信息推荐外,还会将亲密度最高的好友浏览过的且点赞过的医学数据信息推荐给该用户,提高医学信息数据推荐的质量、贴合度、以及拓展推荐宽度。

[0150] 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,则可对这些实施例作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。

[0151] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明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。

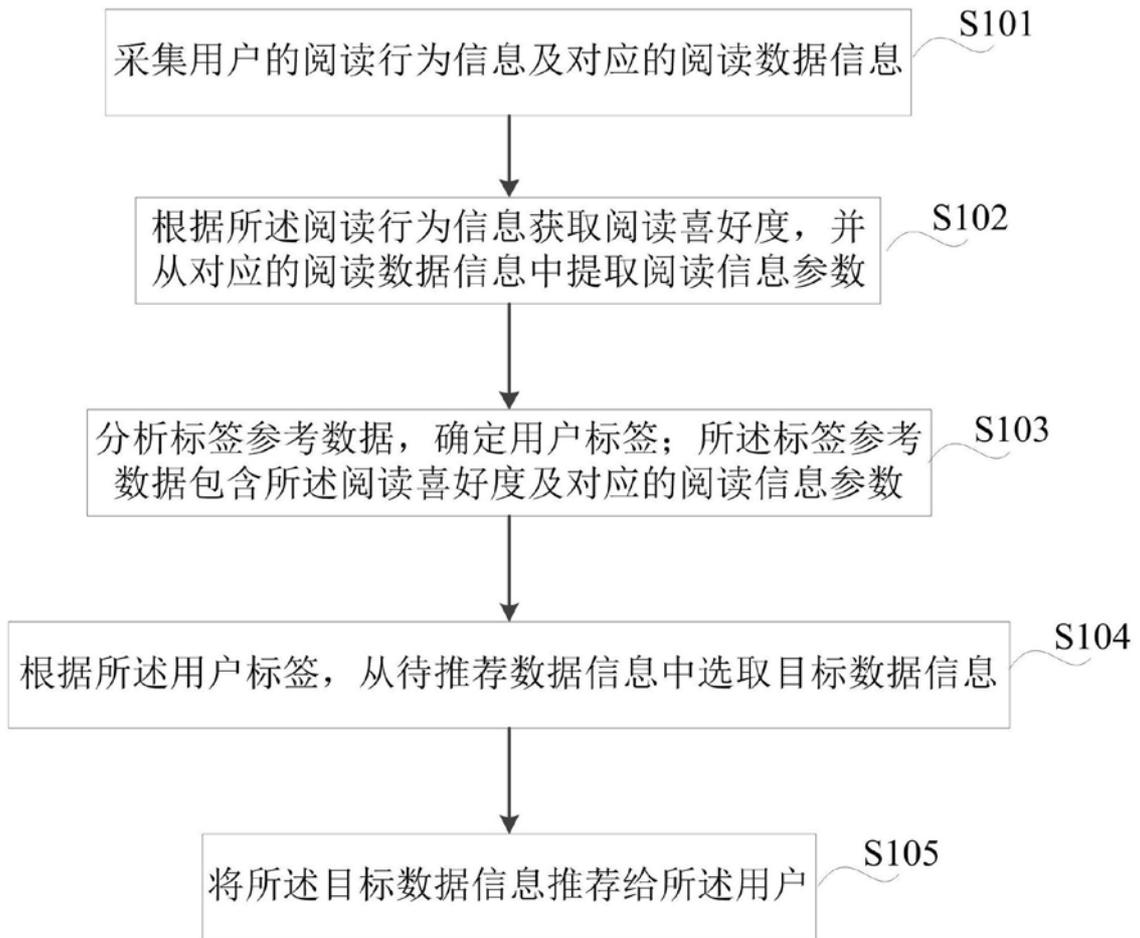


图1

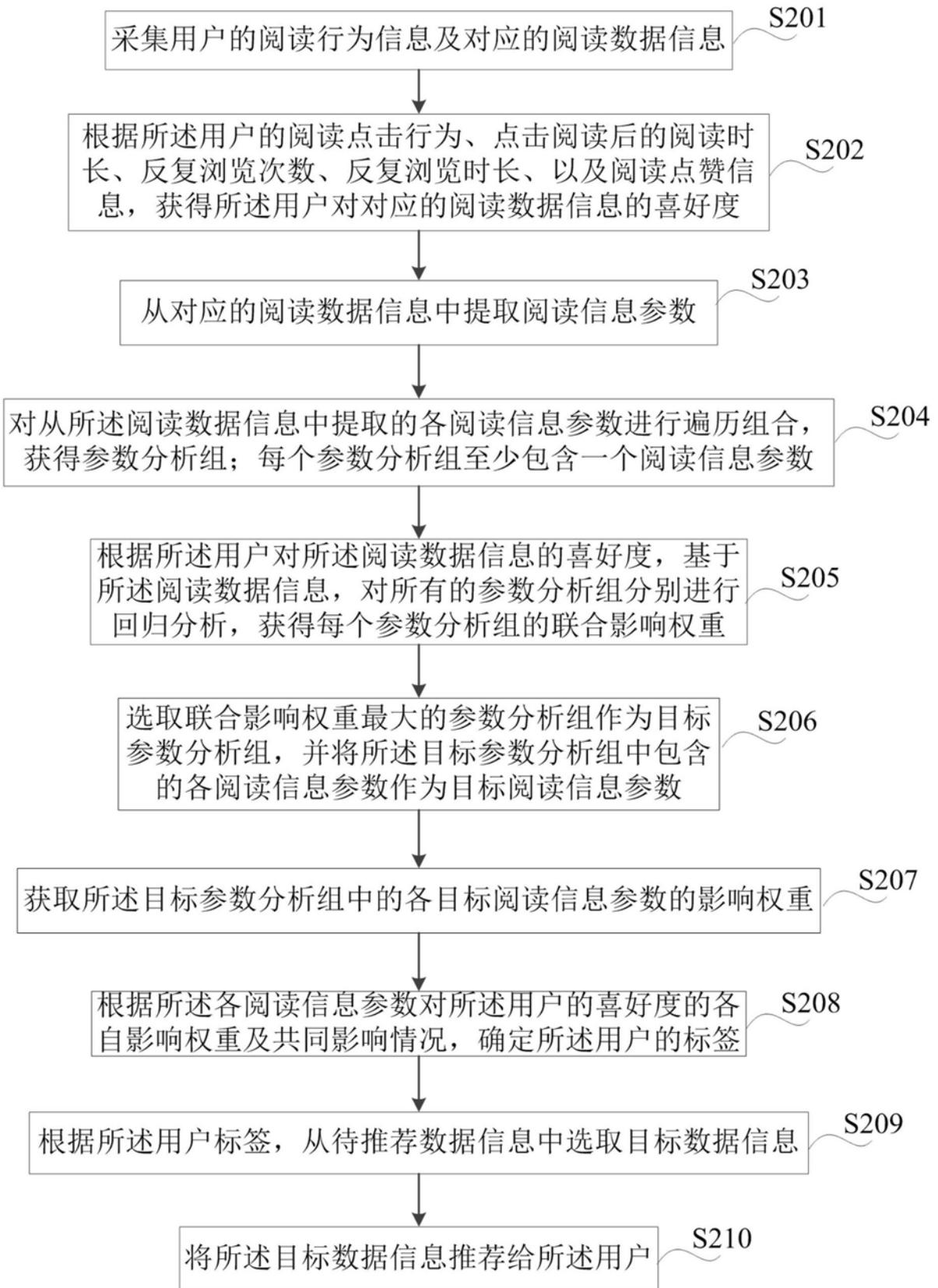


图2



图3

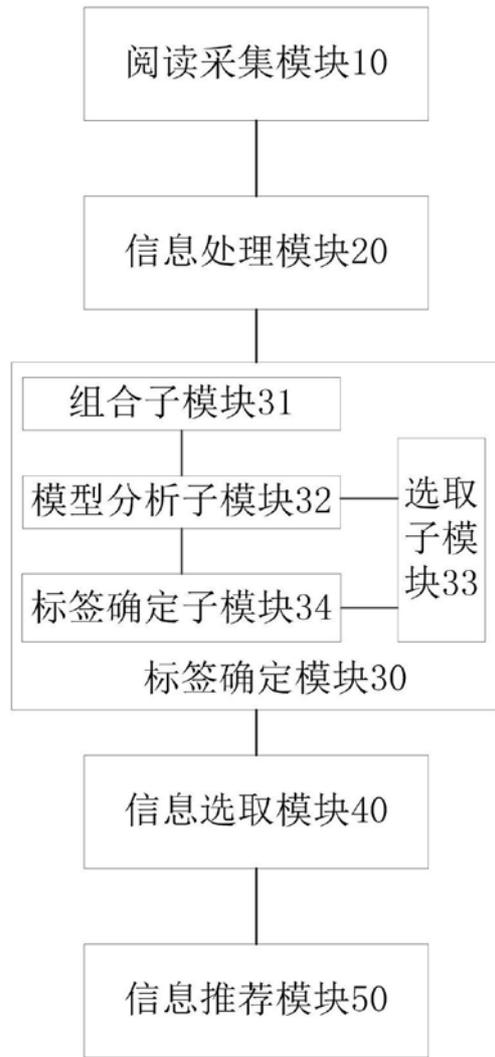


图4