



(12)发明专利申请

(10)申请公布号 CN 111125543 A

(43)申请公布日 2020.05.08

(21)申请号 201911304248.6

(22)申请日 2019.12.17

(71)申请人 掌阅科技股份有限公司

地址 100124 北京市朝阳区四惠大厦2029E

(72)发明人 王海璐 冯丽洋 李帆 曹雯潇

(74)专利代理机构 北京市浩天知识产权代理事务
所(普通合伙) 11276

代理人 梁倩

(51)Int.Cl.

G06F 16/9536(2019.01)

G06F 16/9537(2019.01)

G06F 16/335(2019.01)

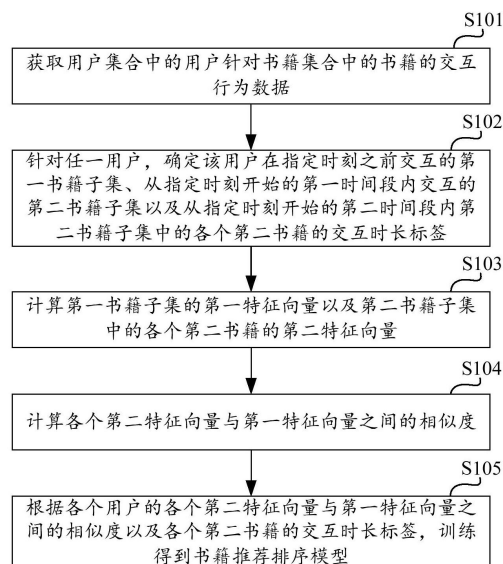
权利要求书3页 说明书22页 附图5页

(54)发明名称

书籍推荐排序模型的训练方法、计算设备及
存储介质

(57)摘要

本发明公开了一种书籍推荐排序模型的训练方法、计算设备及存储介质,该方法包括:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定第一书籍子集、第二书籍子集和各个第二书籍的交互时长标签;计算第一书籍子集的第一特征向量和各个第二书籍的第二特征向量;计算各个第二特征向量与第一特征向量之间的相似度;根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签训练得到书籍推荐排序模型。该方案在模型训练过程中充分考虑了用户与书籍之间的交互时长,得到的书籍推荐排序模型能够准确地基于交互时长对书籍进行排序,有助于提高用户对推荐书籍的交互时长。



1. 一种书籍推荐排序模型的训练方法,包括:

获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;

针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签;其中,所述第二时间段包含所述第一时间段;

计算所述第一书籍子集的第一特征向量以及所述第二书籍子集中的各个第二书籍的第二特征向量;

计算各个第二特征向量与所述第一特征向量之间的相似度;

根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

2. 根据权利要求1所述的方法,所述针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签进一步包括:

对该用户的交互行为数据进行数据分析,得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍;

将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中;

从该用户在所述第一时间段内交互过的书籍中筛除所述第一书籍子集中的各个第一书籍,将筛除后得到的书籍作为第二书籍添加至第二书籍子集中;

对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析,确定所述第二书籍子集中的各个第二书籍的交互时长标签。

3. 根据权利要求1所述的方法,所述计算所述第一书籍子集的第一特征向量进一步包括:

对所述第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取,得到各个第一书籍的书籍特征向量;

根据各个第一书籍的书籍特征向量,计算所述第一书籍子集的第一特征向量。

4. 根据权利要求3所述的方法,所述方法还包括:对该用户在指定时刻之前产生的交互行为数据进行数据分析,确定所述第一书籍子集中的各个第一书籍对应的交互时长;

所述根据各个第一书籍的书籍特征向量,计算所述第一书籍子集的第一特征向量具体为:根据各个第一书籍对应的交互时长,确定各个第一书籍对应的权重值;将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算,得到所述第一书籍子集的第一特征向量。

5. 根据权利要求1所述的方法,计算所述第二书籍子集中的各个第二书籍的第二特征向量进一步包括:

对所述第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量。

6. 一种书籍推荐方法,包括:

获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定所述待推荐用户在当前时刻之前交互的第三书籍子集;

对所述书籍集合中的书籍进行筛选,得到第四书籍子集;

计算所述第三书籍子集的第三特征向量以及所述第四书籍子集中的各个第四书籍的第四特征向量；

计算各个第四特征向量与所述第三特征向量之间的相似度；

将各个第四特征向量与所述第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中，利用所述书籍推荐排序模型对各个第四书籍进行排序，得到所述待推荐用户对应的书籍推荐排序列表。

7. 一种计算设备，包括：处理器、存储器、通信接口和通信总线，所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信；

所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行以下操作：

获取用户集合中的用户针对书籍集合中的书籍的交互行为数据；

针对任一用户，确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签；其中，所述第二时间段包含所述第一时间段；

计算所述第一书籍子集的第一特征向量以及所述第二书籍子集中的各个第二书籍的第二特征向量；

计算各个第二特征向量与所述第一特征向量之间的相似度；

根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签，训练得到书籍推荐排序模型。

8. 一种计算设备，包括：处理器、存储器、通信接口和通信总线，所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信；

所述存储器用于存放至少一可执行指令，所述可执行指令使所述处理器执行以下操作：

获取待推荐用户针对书籍集合中的书籍的交互行为数据，确定所述待推荐用户在当前时刻之前交互的第三书籍子集；

对所述书籍集合中的书籍进行筛选，得到第四书籍子集；

计算所述第三书籍子集的第三特征向量以及所述第四书籍子集中的各个第四书籍的第四特征向量；

计算各个第四特征向量与所述第三特征向量之间的相似度；

将各个第四特征向量与所述第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中，利用所述书籍推荐排序模型对各个第四书籍进行排序，得到所述待推荐用户对应的书籍推荐排序列表。

9. 一种存储介质，所述存储介质中存储有至少一可执行指令，所述可执行指令使处理器执行以下操作：

获取用户集合中的用户针对书籍集合中的书籍的交互行为数据；

针对任一用户，确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签；其中，所述第二时间段包含所述第一时间段；

计算所述第一书籍子集的第一特征向量以及所述第二书籍子集中的各个第二书籍的

第二特征向量；

计算各个第二特征向量与所述第一特征向量之间的相似度；

根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签，训练得到书籍推荐排序模型。

10. 一种存储介质，所述存储介质中存储有至少一可执行指令，所述可执行指令使处理器执行以下操作：

获取待推荐用户针对书籍集合中的书籍的交互行为数据，确定所述待推荐用户在当前时刻之前交互的第三书籍子集；

对所述书籍集合中的书籍进行筛选，得到第四书籍子集；

计算所述第三书籍子集的第三特征向量以及所述第四书籍子集中的各个第四书籍的第四特征向量；

计算各个第四特征向量与所述第三特征向量之间的相似度；

将各个第四特征向量与所述第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中，利用所述书籍推荐排序模型对各个第四书籍进行排序，得到所述待推荐用户对应的书籍推荐排序列表。

书籍推荐排序模型的训练方法、计算设备及存储介质

技术领域

[0001] 本发明涉及信息处理技术领域,具体涉及一种书籍推荐排序模型的训练方法、计算设备及存储介质。

背景技术

[0002] 电子书形式的书籍由于具有获取方便等优势,受到了大量用户的喜爱。书籍阅读平台大多是按照书籍全文的相似度进行书籍推荐的。现有技术中,一般是将与用户曾经阅读过的书籍在书籍内容上具有较高相似度的若干个书籍作为推荐书籍并向用户展现。然而,这种书籍推荐方式,没有充分地考虑所推荐的书籍是否能够吸引用户花费较长时间来进行阅读,用户很有可能在阅读某本推荐书籍时仅阅读了几分钟后续就不再继续阅读,导致书籍推荐的推荐效果不佳。

发明内容

[0003] 鉴于上述问题,提出了本发明以便提供一种克服上述问题或者至少部分地解决上述问题的书籍推荐排序模型的训练方法、计算设备及存储介质。

[0004] 根据本发明的一个方面,提供了一种书籍推荐排序模型的训练方法,该方法包括:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍的交互时长标签;其中,第二时间段包含第一时间段;计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二特征向量;计算各个第二特征向量与第一特征向量之间的相似度;根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0005] 根据本发明的另一方面,提供了一种书籍推荐方法,该方法包括:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定待推荐用户在当前时刻之前交互的第三书籍子集;对书籍集合中的书籍进行筛选,得到第四书籍子集;计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与第三特征向量之间的相似度;将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0006] 根据本发明的另一方面,提供了一种计算设备,包括:处理器、存储器、通信接口和通信总线,处理器、存储器和通信接口通过通信总线完成相互间的通信;存储器用于存放至少一可执行指令,可执行指令使处理器执行以下操作:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍的交互时长标签;其中,第二时间段包含第一时间

段;计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二特征向量;计算各个第二特征向量与第一特征向量之间的相似度;根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0007] 根据本发明的另一方面,提供了一种计算设备,包括:处理器、存储器、通信接口和通信总线,处理器、存储器和通信接口通过通信总线完成相互间的通信;存储器用于存放至少一可执行指令,可执行指令使处理器执行以下操作:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定待推荐用户在当前时刻之前交互的第三书籍子集;对书籍集合中的书籍进行筛选,得到第四书籍子集;计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与第三特征向量之间的相似度;将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0008] 根据本发明的另一方面,提供了一种存储介质,存储介质中存储有至少一可执行指令,可执行指令使处理器执行以下操作:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍的交互时长标签;其中,第二时间段包含第一时间段;计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二特征向量;计算各个第二特征向量与第一特征向量之间的相似度;根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0009] 根据本发明的又一方面,提供了一种存储介质,存储介质中存储有至少一可执行指令,可执行指令使处理器执行以下操作:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定待推荐用户在当前时刻之前交互的第三书籍子集;对书籍集合中的书籍进行筛选,得到第四书籍子集;计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与第三特征向量之间的相似度;将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0010] 根据本发明提供的技术方案,基于针对书籍的交互行为数据,能够便捷地确定每个用户的第一书籍子集、第二书籍子集以及第二书籍子集中的各个第二书籍的交互时长标签;依据第一书籍子集的特征向量以及各个第二书籍的特征向量,能够精准地计算各个第二书籍的特征向量与第一书籍子集的特征向量之间的相似度;根据各个用户的各个第二书籍的特征向量与第一书籍子集的特征向量之间的相似度以及各个第二书籍的交互时长标签进行训练,模型训练过程中充分考虑了用户与书籍之间的交互时长,使得训练得到的书籍推荐排序模型能够快速、准确地基于交互时长对书籍进行排序;利用书籍推荐排序模型对待推荐书籍进行排序时,能够优先地将用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,使得用户能够便捷地从中找到感兴趣的书籍,从而有助于提高用户对所推荐的书籍的交互时长,极大地提升了书籍推荐效果。

[0011] 上述说明仅是本发明技术方案的概述,为了能够更清楚了解本发明的技术手段,而可依照说明书的内容予以实施,并且为了让本发明的上述和其它目的、特征和优点能够更明显易懂,以下特举本发明的具体实施方式。

附图说明

[0012] 通过阅读下文优选实施方式的详细描述,各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的,而并不认为是对本发明的限制。而且在整个附图中,用相同的参考符号表示相同的部件。在附图中:

[0013] 图1示出了根据本发明实施例一的一种书籍推荐排序模型的训练方法的流程示意图;

[0014] 图2示出了根据本发明实施例二的一种书籍推荐排序模型的训练方法的流程示意图;

[0015] 图3示出了根据本发明实施例三的一种书籍推荐方法的流程示意图;

[0016] 图4示出了根据本发明实施例四的一种书籍推荐方法的流程示意图;

[0017] 图5示出了根据本发明实施例七的一种计算设备的结构示意图;

[0018] 图6示出了根据本发明实施例八的一种计算设备的结构示意图。

具体实施方式

[0019] 下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例,然而应当理解,可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反,提供这些实施例是为了能够更透彻地理解本公开,并且能够将本公开的范围完整的传达给本领域的技术人员。

[0020] 实施例一

[0021] 图1示出了根据本发明实施例一的一种书籍推荐排序模型的训练方法的流程示意图,如图1所示,该方法包括如下步骤:

[0022] 步骤S101,获取用户集合中的用户针对书籍集合中的书籍的交互行为数据。

[0023] 用户集合中包含有书籍阅读平台中的多个用户,这多个用户可以为书籍阅读平台中所有的用户,也可以为部分用户,此处不做具体限定。书籍集合中包含有书籍阅读平台提供的多个书籍。用户针对书籍的交互行为数据为用于描述用户与书籍之间发生交互的数据,具体可包括:用户针对书籍的书籍阅读数据、书籍评论数据、书籍下载数据等。在步骤S101中,可从书籍阅读平台中获取其存储的用户集合中的用户针对书籍的书籍阅读数据、书籍评论数据、书籍下载数据等交互行为数据。

[0024] 步骤S102,针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍的交互时长标签。

[0025] 通过对交互行为数据进行数据分析可获知用户集合中的每个用户在哪个时间与哪些书籍发生过交互。针对任一用户,可通过对该用户的交互行为数据进行数据分析,得到该用户交互过的各个书籍以及各个书籍对应的交互时间,进而确定该用户在指定时刻之前交互的第一书籍子集以及在从指定时刻开始的第一时间段内交互的第二书籍子集,并统计

在从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍对应的交互时长,根据统计得到的交互时长确定各个第二书籍的交互时长标签。其中,第二时间段包含第一时间段。本领域技术人员可根据实际需要对指定时刻、第一时间段、第二时间段以及各个第二书籍的交互时长标签进行设置。例如,指定时刻可为2019年11月5日的00:00:00(即零点零分零秒),第一时间段可为从指定时刻开始的1天内,第二时间段可为从指定时刻开始的7天内。

[0026] 其中,第一书籍子集中的各个第一书籍是指该用户在指定时刻之前交互过的书籍。第二书籍子集与第一书籍子集没有交集,第二书籍子集中的各个第二书籍是指该用户在第一时间段内交互过而在指定时刻之前没有交互过的那些书籍,也就是说,第二书籍是指该用户在第一时间段内增量交互的、之前没有交互过的书籍。

[0027] 步骤S103,计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二特征向量。

[0028] 第一特征向量是指第一书籍子集整体对应的特征向量,而第二特征向量是指第二书籍子集中每个第二书籍的书籍特征向量。其中,第一特征向量可通过对第一书籍子集中的各个第一书籍的书籍特征向量进行运算处理而得到。具体地,对第一书籍子集中的各个第一书籍进行特征提取,得到各个第一书籍的书籍特征向量,对各个第一书籍的书籍特征向量进行运算处理,将运算结果作为第一书籍子集的第一特征向量。对第二书籍子集中的各个第二书籍进行特征提取,得到各个第二书籍的书籍特征向量作为各个第二书籍的第二特征向量。

[0029] 步骤S104,计算各个第二特征向量与第一特征向量之间的相似度。

[0030] 在得到了第一特征向量与各个第二特征向量之后,就可根据第一特征向量与各个第二特征向量,采用余弦相似度计算方式或者欧几里德距离计算方式等相似度计算方式,对各个第二特征向量与第一特征向量之间的相似度进行计算。

[0031] 步骤S105,根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0032] 其中,可根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签进行模型训练,训练得到书籍推荐排序模型。利用书籍推荐排序模型,能够对任一用户的各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表,各个第二书籍在书籍推荐排序列表中的排序位置能够很好地符合在第二时间段内该用户与各个第二书籍之间的交互时长的排序,将交互时长较长的第二书籍排在靠前的位置。那么利用书籍推荐排序模型对待推荐书籍进行排序时能够优先地将用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,从而有助于提高用户对所推荐的书籍的交互时长,进而有助于提升用户在书籍阅读平台中的留存率。本领域技术人员可根据实际需要确定具体训练方式,此处不做限定。

[0033] 利用本实施例提供的书籍推荐排序模型的训练方法,基于针对书籍的交互行为数据,能够便捷地确定每个用户的第一书籍子集、第二书籍子集以及第二书籍子集中的各个第二书籍的交互时长标签;依据第一书籍子集的特征向量以及各个第二书籍的特征向量,能够精准地计算各个第二书籍的特征向量与第一书籍子集的特征向量之间的相似度;根据各个用户的各个第二书籍的特征向量与第一书籍子集的特征向量之间的相似度以及各个

第二书籍的交互时长标签进行训练,模型训练过程中充分考虑了用户与书籍之间的交互时长,使得训练得到的书籍推荐排序模型能够快速、准确地基于交互时长对书籍进行排序;利用书籍推荐排序模型对待推荐书籍进行排序时,能够优先地将用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,使得用户能够便捷地从中找到感兴趣的书籍,从而有助于提高用户对所推荐的书籍的交互时长,极大地提升了书籍推荐效果。

[0034] 实施例二

[0035] 图2示出了根据本发明实施例二的一种书籍推荐排序模型的训练方法的流程示意图,如图2所示,该方法包括如下步骤:

[0036] 步骤S201,获取用户集合中的用户针对书籍集合中的书籍的交互行为数据。

[0037] 其中,可从书籍阅读平台中获取用户集合中的各个用户针对书籍集合中的书籍的书籍阅读数据、书籍评论数据、书籍下载数据等交互行为数据。交互行为数据至少包含:用户ID、用户所交互的书籍ID、交互起始时间、交互终止时间等数据。具体地,用户ID可为用户在书籍阅读平台中的账号,例如手机号、用户名、邮箱、微信号等第三方平台账号。

[0038] 步骤S202,针对任一用户,对该用户的交互行为数据进行数据分析,得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍。

[0039] 具体地,可对该用户在指定时刻之前产生的交互行为数据进行数据分析,从交互行为数据中提取出与该用户发生过交互的书籍ID,从而得到该用户在指定时刻之前交互过的书籍。对该用户在指定时刻开始的第一时间段内产生的交互行为数据进行数据分析,从交互行为数据中提取出与该用户发生过交互的书籍ID,从而得到该用户在从指定时刻开始的第一时间段内交互过的书籍。例如,可将第一时间段设置为从指定时刻开始的1天内。

[0040] 步骤S203,将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中。

[0041] 步骤S204,从该用户在第一时间段内交互过的书籍中筛除第一书籍子集中的各个第一书籍,将筛除后得到的书籍作为第二书籍添加至第二书籍子集中。

[0042] 假设通过数据分析得到,该用户在指定时刻之前交互过的书籍包括书籍4、书籍5、书籍6和书籍7,该用户在第一时间段内交互过的书籍包括书籍1、书籍2、书籍3、书籍4和书籍5。那么在步骤S203中将书籍4、书籍5、书籍6和书籍7分别作为第一书籍添加至第一书籍子集中,即第一书籍子集中包含有书籍4、书籍5、书籍6和书籍7;在步骤S204中,由于该用户在第一时间段内交互过的书籍中的书籍4和书籍5为第一书籍子集中的第一子集,所以从书籍1、书籍2、书籍3、书籍4和书籍5筛除书籍4和书籍5,筛除后的得到的书籍包括书籍1、书籍2和书籍3,将书籍1、书籍2和书籍3分别作为第二书籍添加至第二书籍子集中,即第二书籍子集中包含有书籍1、书籍2和书籍3。可见,第二书籍子集与第一书籍子集没有交集。

[0043] 步骤S205,对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析,确定第二书籍子集中的各个第二书籍的交互时长标签。

[0044] 第二时间段包含第一时间段。例如,可将第二时间段设置为从指定时刻开始的7天内。在步骤S205中对该用户在第二时间段内产生的交互行为数据进行数据分析,具体可分析包含有第二书籍的书籍ID的交互行为数据,统计在第二时间段内各个第二书籍对应的交互时长,其中,每个第二书籍对应的交互时长可根据该用户在第二时间段内产生的针对该第二书籍的交互行为数据中的交互起始时间和交互终止时间进行计算得到。然后根据统计

得到的各个第二书籍对应的交互时长确定各个第二书籍的交互时长标签。以根据各个第二书籍对应的交互时长所包含的分钟数量确定各个第二书籍的交互时长标签为例,假设第二书籍子集中包含有书籍1、书籍2和书籍3,经统计得到,在第二时间段内,书籍1对应的交互时长为100分钟,书籍2对应的交互时长为60分钟,书籍3对应的交互时长为30分钟,那么可将书籍1对应的交互时长标签确定为100,将书籍2对应的交互时长标签确定为60,将书籍3对应的交互时长标签确定为30。

[0045] 步骤S206,计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二特征向量。

[0046] 对于第一特征向量的计算,可利用神经网络或者embedding(嵌入)方法等,分别对第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取,得到各个第一书籍的书籍特征向量,其中,部分内容可为前若干个章节或段落的内容。第一书籍的书籍特征向量能够反映第一书籍的内容信息,书籍特征向量具体可为embedding向量;然后根据各个第一书籍的书籍特征向量,计算第一书籍子集的第一特征向量。在根据各个第一书籍的书籍特征向量计算第一特征向量的过程中,可结合用户与第一书籍之间的交互时长等因素进行计算。具体地,可对该用户在指定时刻之前产生的交互行为数据进行数据分析,确定第一书籍子集中的各个第一书籍对应的交互时长;接着根据各个第一书籍对应的交互时长,确定各个第一书籍对应的权重值,交互时长越长,权重值越大;然后将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算,得到第一书籍子集的第一特征向量。第一特征向量能够综合、准确地反映用户喜欢交互的书籍的内容信息。

[0047] 例如,第一书籍子集中包含有书籍4、书籍5、书籍6和书籍7,若经对该用户在指定时刻之前产生的交互行为数据进行数据分析得到书籍4对应的交互时长为30分钟,书籍5对应的交互时长为50分钟,书籍6对应的交互时长为20分钟,书籍7对应的交互时长为100分钟,那么书籍4对应的权重值4可为 $30/(30+50+20+100)$,书籍5对应的权重值5可为 $50/(30+50+20+100)$,书籍6对应的权重值6可为 $20/(30+50+20+100)$,书籍7对应的权重值7可为 $100/(30+50+20+100)$,即权重值4为0.15,权重值5为0.25,权重值6为0.1,权重值7为0.5。在确定了各个第一书籍对应的权重值之后,将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权求和,将加权求和的结果作为第一书籍子集的第一特征向量。

[0048] 对于各个第二书籍的第二特征向量的计算,可利用神经网络或者embedding方法等,分别对第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量,其中,第二特征向量能够反映第二书籍的内容信息,第二特征向量具体可为embedding向量。

[0049] 可选地,各个第一书籍的书籍特征向量和各个第二书籍的第二特征向量还可为其他能够反映书籍的内容信息的向量,此处不做具体限定。

[0050] 步骤S207,计算任一第二特征向量与第一特征向量之间的夹角的余弦值,利用余弦值确定该第二特征向量与第一特征向量之间的相似度。

[0051] 在本实施例中,采用余弦相似度计算方法来确定各个第二特征向量与第一特征向量之间的相似度。具体地,利用余弦计算公式,计算任一第二特征向量与第一特征向量之间的夹角的余弦值,然后利用计算得到的余弦值确定该第二特征向量与第一特征向量之间的相似度。

[0052] 步骤S208,根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0053] 其中,在模型训练过程中,可采用文档列表(Listwise)等机器学习排序模型作为初始训练模型。书籍推荐排序模型的训练步骤可包括:针对任一用户,将该用户的各个第二特征向量与第一特征向量之间的相似度输入至初始训练模型中,利用初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;接着根据各个第二书籍在书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益(NDCG),然后根据归一化折损累计增益更新初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。其中,可根据各个第二书籍的交互时长标签来确定NDCG计算过程中所涉及的各个第二书籍的相关性分数,若各个第二书籍分别为书籍1、书籍2和书籍3,书籍1对应的交互时长标签为100,将书籍2对应的交互时长标签为60,将书籍3对应的交互时长标签为30,那么可将书籍1的相关性分数确定为100,将书籍2的相关性分数确定为60,将书籍3的相关性分数确定为30。对于NDCG的具体计算方式可参考现有技术中的NDCG的计算方式,此处不再赘述。本领域技术人员可根据实际需要对迭代结束条件进行设置,此处不做具体限定。

[0054] 其中,NDCG越接近于1,说明在书籍推荐排序模型输出的书籍推荐排序列表中各个第二书籍的排序位置越能很好地符合各个第二书籍的交互时长标签,即越能很好地符合用户与各个第二书籍之间的交互时长的排序,将交互时长较长的第二书籍排在靠前的位置。那么在利用书籍推荐排序模型进行书籍推荐时,就能够很好地将用户可能感兴趣的且未来交互时长可能较长的书籍优先排列在靠前的位置,有助于提高用户对所推荐的书籍的交互时长,进而有助于提升用户在书籍阅读平台中的留存率。

[0055] 可选地,为了使训练得到的书籍推荐排序模型所输出的书籍推荐排序列表中的书籍能够更加符合用户的兴趣度等特征,还可获取用户集合中的各个用户的用户特征向量,用户特征向量可用于反映用户的活跃度、阅读模式喜好(例如喜欢阅读付费书籍还是免费书籍)、兴趣度等各方面的特征。那么在步骤S208中,则需要根据各个用户的各个第二特征向量与第一特征向量之间的相似度、各个第二书籍的交互时长标签以及各个用户的用户特征向量,训练得到书籍推荐排序模型。具体地,针对任一用户,将该用户的各个第二特征向量与第一特征向量之间的相似度以及该用户的用户特征向量输入至初始训练模型中,利用初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;接着根据各个第二书籍在书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益,根据归一化折损累计增益更新初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。

[0056] 利用本实施例提供的书籍推荐排序模型的训练方法,通过对用户针对书籍的交互行为数据进行数据分析,能够准确、快速地确定每个用户的第一书籍子集、第二书籍子集以及第二书籍子集中的各个第二书籍的交互时长标签;结合各个第一书籍对应的交互时长,对各个第一书籍的书籍特征向量进行加权运算,所得到的第一书籍子集的特征向量能够综合、准确地反映用户喜欢交互的书籍的内容信息;根据各个用户的各个第二书籍的特征向量与第一书籍子集的特征向量之间的相似度以及各个第二书籍的交互时长标签进行训练,有效地提升了书籍推荐排序模型基于交互时长对书籍进行排序的准确率。

[0057] 实施例三

[0058] 图3示出了根据本发明实施例三的一种书籍推荐方法的流程示意图,如图3所示,该方法包括如下步骤:

[0059] 步骤S301,获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定待推荐用户在当前时刻之前交互的第三书籍子集。

[0060] 其中,可从书籍阅读平台中获取待推荐用户针对书籍集合中的书籍的书籍阅读数据、书籍评论数据、书籍下载数据等交互行为数据。在获取了待推荐用户的交互行为数据之后,对待推荐用户的交互行为数据进行数据分析,得到待推荐用户在当前时刻之前交互过的书籍,将待推荐用户在当前时刻之前交互过的书籍作为第三书籍,从而得到第三书籍子集。

[0061] 步骤S302,对书籍集合中的书籍进行筛选,得到第四书籍子集。

[0062] 其中,第四书籍子集中的第四书籍为待推荐书籍。考虑到书籍集合中不仅可能包含有待推荐用户已经交互过的书籍,还可能包含有不符合待推荐用户的阅读喜好的书籍,因此需要对书籍集合中的书籍进行筛选,将经筛选得到的书籍作为第四书籍,从而得到第四书籍子集。

[0063] 步骤S303,计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量。

[0064] 第三特征向量是指第三书籍子集整体对应的特征向量,而第四特征向量是指第四书籍子集中每个第四书籍的书籍特征向量。其中,第三特征向量可通过对第三书籍子集中的各个第三书籍的书籍特征向量进行运算处理而得到。具体地,对第三书籍子集中的各个第三书籍进行特征提取,得到各个第三书籍的书籍特征向量,对各个第三书籍的书籍特征向量进行运算处理,将运算结果作为第三书籍子集的第三特征向量。对第四书籍子集中的各个第四书籍进行特征提取,得到各个第四书籍的书籍特征向量作为各个第四书籍的第四特征向量。

[0065] 步骤S304,计算各个第四特征向量与第三特征向量之间的相似度。

[0066] 在得到了第三特征向量与各个第四特征向量之后,就可根据第三特征向量与各个第四特征向量,采用余弦相似度计算方式或者欧几里德距离计算方式等相似度计算方式,对各个第四特征向量与第三特征向量之间的相似度进行计算。

[0067] 步骤S305,将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0068] 其中,书籍推荐排序模型基于上述实施例一或实施例二训练获得。将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。那么在向待推荐用户进行书籍推荐时,可在书籍推荐页面中将各个第四书籍按照其在书籍推荐排序列表中的排列顺序进行呈现,从而优先将待推荐用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,有助于提高用户对所推荐的书籍的交互时长,进而有助于提升用户在书籍阅读平台中的留存率。

[0069] 利用本实施例提供的书籍推荐方法,基于用户针对书籍的交互行为数据,能够便

捷地确定用户的第三书籍子集、第四书籍子集以及第四书籍子集中的各个第四书籍的交互时长标签;依据第三书籍子集的特征向量以及各个第四书籍的特征向量,能够精准地计算各个第四书籍的特征向量与第三书籍子集的特征向量之间的相似度;通过将各个第四书籍的特征向量与第三书籍子集的特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,能够优先将用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,使得用户能够便捷地从中找到感兴趣的书籍,从而有助于提高用户对所推荐的书籍的交互时长,极大地提升了书籍推荐效果。

[0070] 图4示出了根据本发明实施例四的一种书籍推荐方法的流程示意图,如图4所示,该方法包括如下步骤:

[0071] 步骤S401,获取待推荐用户针对书籍集合中的书籍的交互行为数据以及待推荐用户的用户画像数据。

[0072] 其中,可从书籍阅读平台中获取待推荐用户针对书籍集合中的书籍的交互行为数据,从用户画像数据库中获取待推荐用户的用户画像数据,用户画像数据包含有用户的性别、年龄、阅读喜好(包括喜欢阅读的书籍类型、喜欢的作者、喜欢阅读的书籍题材、喜欢的阅读模式等)等各方面的数据。

[0073] 步骤S402,对待推荐用户的交互行为数据进行数据分析,得到待推荐用户在当前时刻之前交互过的书籍。

[0074] 其中,可对待推荐用户的交互行为数据进行数据分析,从交互行为数据中提取出与待推荐用户发生过交互的书籍ID,从而得到待推荐用户在当前时刻之前交互过的书籍。

[0075] 步骤S403,将待推荐用户在当前时刻之前交互过的书籍作为第三书籍添加至第三书籍子集中。

[0076] 步骤S404,在书籍集合中查找书籍画像数据与用户画像数据相匹配的书籍。

[0077] 具体地,可将书籍集合中的每个书籍的书籍画像数据分别与待推荐用户的用户画像数据进行匹配,计算每个书籍的书籍画像数据与用户画像数据之间的匹配值,接着对书籍画像数据与用户画像数据之间的匹配值按照从高到低的顺序进行排序,得到各个书籍画像数据对应的各个书籍的排列结果,将排列结果中排列靠前的n个书籍确定为相匹配的书籍,其中,n大于1。例如,可将n设置为1000。

[0078] 步骤S405,从相匹配的书籍中筛除第三书籍子集中的各个第三书籍,将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0079] 考虑到步骤S404所查找到的相匹配的书籍中可能包含有待推荐用户已经交互过的书籍(即第三书籍),则在确定第四书籍的过程中,需要从相匹配的书籍中筛除第三书籍子集中的各个第三书籍,筛除后得到的书籍为既与用户画像数据相匹配且待推荐用户又未曾交互过的书籍,那么将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0080] 步骤S406,计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量。

[0081] 对于第三特征向量的计算,可利用神经网络或者embedding方法等,分别对第三书籍子集中的各个第三书籍的全文或部分内容进行特征提取,得到各个第三书籍的书籍特征向量,其中,部分内容可为前若干个章节或段落的内容。第三书籍的书籍特征向量能够反映第三书籍的内容信息,书籍特征向量具体可为embedding向量;然后根据各个第三书籍的书

籍特征向量,计算第三书籍子集的第三特征向量。在根据各个第三书籍的书籍特征向量计算第三特征向量的过程中,可结合待推荐用户与第三书籍之间的交互时长等因素进行计算。具体地,可对待推荐用户的交互行为数据进行数据分析,确定第三书籍子集中的各个第三书籍对应的交互时长;接着根据各个第三书籍对应的交互时长,确定各个第三书籍对应的权重值,交互时长越长,权重值越大;然后将各个第三书籍的书籍特征向量与各个第三书籍对应的权重值进行加权运算,得到第三书籍子集的第三特征向量。第三特征向量能够综合、准确地反映待推荐用户喜欢交互的书籍的内容信息。

[0082] 对于各个第四书籍的第四特征向量的计算,可利用神经网络或者embedding方法等,分别对第四书籍子集中的各个第四书籍的全文或部分内容进行特征提取,得到各个第四书籍的第四特征向量,其中,第四特征向量能够反映第四书籍的内容信息,第四特征向量具体可为embedding向量。

[0083] 可选地,各个第三书籍的书籍特征向量和各个第四书籍的第四特征向量还可为其他能够反映书籍的内容信息的向量,此处不做具体限定。

[0084] 步骤S407,计算任一第四特征向量与第三特征向量之间的夹角的余弦值,利用余弦值确定该第四特征向量与第三特征向量之间的相似度。

[0085] 在本实施例中,采用余弦相似度计算方法来确定各个第四特征向量与第三特征向量之间的相似度。具体地,利用余弦计算公式,计算任一第四特征向量与第三特征向量之间的夹角的余弦值,然后利用计算得到的余弦值确定该第四特征向量与第三特征向量之间的相似度。

[0086] 步骤S408,将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0087] 其中,书籍推荐排序模型基于上述实施例一或实施例二训练获得。

[0088] 利用本实施例提供的书籍推荐方法,通过对用户针对书籍的交互行为数据进行数据分析,能够准确、快速地确定用户的第三书籍子集、第四书籍子集以及第四书籍子集中的各个第四书籍的交互时长标签;结合各个第三书籍对应的交互时长,对各个第三书籍的书籍特征向量进行加权运算,所得到的第三书籍子集的特征向量能够综合、准确地反映用户喜欢交互的书籍的内容信息;通过将各个第四书籍的特征向量与第三书籍子集的特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,实现了对待推荐书籍的快速、精准排序,能够优先将用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,使得用户能够便捷地从中找到感兴趣的书籍,极大地改善了书籍推荐效果。

[0089] 实施例五

[0090] 本发明实施例五提供了一种非易失性存储介质,存储介质存储有至少一可执行指令,该可执行指令可执行上述任意方法实施例中的书籍推荐排序模型的训练方法。

[0091] 可执行指令具体可以用于使得处理器执行以下操作:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍的交互时长标签;其中,第二时间段包含第一时间段;计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二

特征向量;计算各个第二特征向量与第一特征向量之间的相似度;根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0092] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对该用户的交互行为数据进行数据分析,得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍;将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中;从该用户在第一时间段内交互过的书籍中筛除第一书籍子集中的各个第一书籍,将筛除后得到的书籍作为第二书籍添加至第二书籍子集中;对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析,确定第二书籍子集中的各个第二书籍的交互时长标签。

[0093] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取,得到各个第一书籍的书籍特征向量;根据各个第一书籍的书籍特征向量,计算第一书籍子集的第一特征向量。

[0094] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对该用户在指定时刻之前产生的交互行为数据进行数据分析,确定第一书籍子集中的各个第一书籍对应的交互时长;根据各个第一书籍对应的交互时长,确定各个第一书籍对应的权重值;将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算,得到第一书籍子集的第一特征向量。

[0095] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量。

[0096] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:计算任一第二特征向量与第一特征向量之间的夹角的余弦值,利用余弦值确定该第二特征向量与第一特征向量之间的相似度。

[0097] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:针对任一用户,将该用户的各个第二特征向量与第一特征向量之间的相似度输入至初始训练模型中,利用初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;根据各个第二书籍在书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益,根据归一化折损累计增益更新初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。

[0098] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:获取用户集合中的各个用户的用户特征向量;根据各个用户的各个第二特征向量与第一特征向量之间的相似度、各个第二书籍的交互时长标签以及各个用户的用户特征向量,训练得到书籍推荐排序模型。

[0099] 实施例六

[0100] 本发明实施例六提供了一种非易失性存储介质,存储介质存储有至少一可执行指令,该可执行指令可执行上述任意方法实施例中的书籍推荐方法。

[0101] 可执行指令具体可以用于使得处理器执行以下操作:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定待推荐用户在当前时刻之前交互的第三书籍子集;对书

籍集合中的书籍进行筛选,得到第四书籍子集;计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与第三特征向量之间的相似度;将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0102] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对待推荐用户的交互行为数据进行数据分析,得到待推荐用户在当前时刻之前交互过的书籍;将待推荐用户在当前时刻之前交互过的书籍作为第三书籍添加至第三书籍子集中。

[0103] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:获取待推荐用户的用户画像数据;在书籍集合中查找书籍画像数据与用户画像数据相匹配的书籍;从相匹配的书籍中筛除第三书籍子集中的各个第三书籍,将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0104] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对第三书籍子集中的各个第三书籍的全文或部分内容进行特征提取,得到各个第三书籍的书籍特征向量;根据各个第三书籍的书籍特征向量,计算第三书籍子集的第三特征向量。

[0105] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对待推荐用户的交互行为数据进行数据分析,确定第三书籍子集中的各个第三书籍对应的交互时长;根据各个第三书籍对应的交互时长,确定各个第三书籍对应的权重值;将各个第三书籍的书籍特征向量与各个第三书籍对应的权重值进行加权运算,得到第三书籍子集的第三特征向量。

[0106] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:对第四书籍子集中的各个第四书籍的全文或部分内容进行特征提取,得到各个第四书籍的第四特征向量。

[0107] 在一种可选的实施方式中,可执行指令进一步使处理器执行以下操作:计算任一第四特征向量与第三特征向量之间的夹角的余弦值,利用余弦值确定该第四特征向量与第三特征向量之间的相似度。

[0108] 在一种可选的实施方式中,书籍推荐排序模型基于实施例五的存储介质训练获得。

[0109] 实施例七

[0110] 图5示出了根据本发明实施例七的一种计算设备的结构示意图,本发明具体实施例并不对计算设备的具体实现做限定。如图5所示,该计算设备可以包括:处理器(processor) 502、通信接口(Communications Interface) 504、存储器(memory) 506、以及通信总线508。其中:处理器502、通信接口504、以及存储器506通过通信总线508完成相互间的通信。通信接口504,用于与其它设备比如客户端或其它服务器等的网元通信。处理器502,用于执行程序510,具体可以执行上述书籍推荐排序模型的训练方法实施例中的相关步骤。具体地,程序510可以包括程序代码,该程序代码包括计算机操作指令。处理器502可能是中央处理器CPU,或者是特定集成电路ASIC(Application Specific Integrated Circuit),或者是被配置成实施本发明实施例的一个或多个集成电路。计算设备包括的一个或多个处理器,可以是同一类型的处理器,如一个或多个CPU;也可以是不同类型的处理器,如一个或

多个CPU以及一个或多个ASIC。存储器506,用于存放程序510。存储器506可能包含高速RAM存储器,也可能还包括非易失性存储器(non-volatile memory),例如至少一个磁盘存储器。

[0111] 程序510具体可以用于使得处理器502执行以下操作:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内第二书籍子集中的各个第二书籍的交互时长标签;其中,第二时间段包含第一时间段;计算第一书籍子集的第一特征向量以及第二书籍子集中的各个第二书籍的第二特征向量;计算各个第二特征向量与第一特征向量之间的相似度;根据各个用户的各个第二特征向量与第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0112] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:对该用户的交互行为数据进行数据分析,得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍;将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中;从该用户在第一时间段内交互过的书籍中筛除第一书籍子集中的各个第一书籍,将筛除后得到的书籍作为第二书籍添加至第二书籍子集中;对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析,确定第二书籍子集中的各个第二书籍的交互时长标签。

[0113] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:对第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取,得到各个第一书籍的书籍特征向量;根据各个第一书籍的书籍特征向量,计算第一书籍子集的第一特征向量。

[0114] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:对该用户在指定时刻之前产生的交互行为数据进行数据分析,确定第一书籍子集中的各个第一书籍对应的交互时长;根据各个第一书籍对应的交互时长,确定各个第一书籍对应的权重值;将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算,得到第一书籍子集的第一特征向量。

[0115] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:对第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量。

[0116] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:计算任一第二特征向量与第一特征向量之间的夹角的余弦值,利用余弦值确定该第二特征向量与第一特征向量之间的相似度。

[0117] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:针对任一用户,将该用户的各个第二特征向量与第一特征向量之间的相似度输入至初始训练模型中,利用初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;根据各个第二书籍在书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益,根据归一化折损累计增益更新初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。

[0118] 在一种可选的实施方式中,程序510进一步使得处理器502执行以下操作:获取用

户集合中的各个用户的用户特征向量;根据各个用户的各个第二特征向量与第一特征向量之间的相似度、各个第二书籍的交互时长标签以及各个用户的用户特征向量,训练得到书籍推荐排序模型。

[0119] 程序510中各步骤的具体实现可以参见上述书籍推荐排序模型的训练实施例中的相应步骤对应的描述,在此不赘述。所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的设备的具体的工作过程,可以参考前述方法实施例中的对应过程描述,在此不再赘述。

[0120] 实施例八

[0121] 图6示出了根据本发明实施例八的一种计算设备的结构示意图,本发明具体实施例并不对计算设备的具体实现做限定。如图6所示,该计算设备可以包括:处理器602、通信接口604、存储器606、以及通信总线608。其中:处理器602、通信接口604、以及存储器606通过通信总线608完成相互间的通信。通信接口604,用于与其它设备比如客户端或其它服务器等的网元通信。处理器602,用于执行程序610,具体可以执行上述书籍推荐方法实施例中的相关步骤。具体地,程序610可以包括程序代码,该程序代码包括计算机操作指令。处理器602可能是中央处理器CPU,或者是特定集成电路ASIC,或者是被配置成实施本发明实施例的一个或多个集成电路。计算设备包括的一个或多个处理器,可以是同一类型的处理器,如一个或多个CPU;也可以是不同类型的处理器,如一个或多个CPU以及一个或多个ASIC。存储器606,用于存放程序610。存储器606可能包含高速RAM存储器,也可能还包括非易失性存储器,例如至少一个磁盘存储器。

[0122] 程序610具体可以用于使得处理器602执行以下操作:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定待推荐用户在当前时刻之前交互的第三书籍子集;对书籍集合中的书籍进行筛选,得到第四书籍子集;计算第三书籍子集的第三特征向量以及第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与第三特征向量之间的相似度;将各个第四特征向量与第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用书籍推荐排序模型对各个第四书籍进行排序,得到待推荐用户对应的书籍推荐排序列表。

[0123] 在一种可选的实施方式中,程序610进一步使得处理器602执行以下操作:对待推荐用户的交互行为数据进行数据分析,得到待推荐用户在当前时刻之前交互过的书籍;将待推荐用户在当前时刻之前交互过的书籍作为第三书籍添加至第三书籍子集中。

[0124] 在一种可选的实施方式中,程序610进一步使得处理器602执行以下操作:获取待推荐用户的用户画像数据;在书籍集合中查找书籍画像数据与用户画像数据相匹配的书籍;从相匹配的书籍中筛除第三书籍子集中的各个第三书籍,将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0125] 在一种可选的实施方式中,程序610进一步使得处理器602执行以下操作:对第三书籍子集中的各个第三书籍的全文或部分内容进行特征提取,得到各个第三书籍的书籍特征向量;根据各个第三书籍的书籍特征向量,计算第三书籍子集的第三特征向量。

[0126] 在一种可选的实施方式中,程序610进一步使得处理器602执行以下操作:对待推荐用户的交互行为数据进行数据分析,确定第三书籍子集中的各个第三书籍对应的交互时长;根据各个第三书籍对应的交互时长,确定各个第三书籍对应的权重值;将各个第三书籍

的书籍特征向量与各个第三书籍对应的权重值进行加权运算,得到第三书籍子集的第三特征向量。

[0127] 在一种可选的实施方式中,程序610进一步使得处理器602执行以下操作:对第四书籍子集中的各个第四书籍的全文或部分内容进行特征提取,得到各个第四书籍的第四特征向量。

[0128] 在一种可选的实施方式中,程序610进一步使得处理器602执行以下操作:计算任一第四特征向量与第三特征向量之间的夹角的余弦值,利用余弦值确定该第四特征向量与第三特征向量之间的相似度。

[0129] 在一种可选的实施方式中,书籍推荐排序模型基于实施例七的计算设备训练获得。

[0130] 程序610中各步骤的具体实现可以参见上述书籍推荐实施例中的相应步骤对应的描述,在此不赘述。所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的设备的具体的工作过程,可以参考前述方法实施例中的对应过程描述,在此不再赘述。

[0131] 通过本实施例提供的方案,在模型训练过程中充分考虑了用户与书籍之间的交互时长,使得训练得到的书籍推荐排序模型能够快速、准确地基于交互时长对书籍进行排序;利用书籍推荐排序模型对待推荐书籍进行排序时,能够优先地将用户可能感兴趣的且未来交互时长可能较长的书籍排列在靠前的位置,使得用户能够便捷地从中找到感兴趣的书籍,从而有助于提高用户对所推荐的书籍的交互时长,极大地提升了书籍推荐效果。

[0132] 在此提供的算法和显示不与任何特定计算机、虚拟系统或者其它设备固有相关。各种通用系统也可以与基于在此的示教一起使用。根据上面的描述,构造这类系统所要求的结构是显而易见的。此外,本发明也不针对任何特定编程语言。应当明白,可以利用各种编程语言实现在此描述的本发明的内容,并且上面对特定语言所做的描述是为了披露本发明的最佳实施方式。

[0133] 在此处所提供的说明书中,说明了大量具体细节。然而,能够理解,本发明的实施例可以在没有这些具体细节的情况下实践。在一些实例中,并未详细示出公知的方法、结构和技术,以便不模糊对本说明书的理解。

[0134] 类似地,应当理解,为了精简本公开并帮助理解各个发明方面中的一个或多个,在上面对本发明的示例性实施例的描述中,本发明的各个特征有时被一起分组到单个实施例、图、或者对其的描述中。然而,并不应将该公开的方法解释成反映如下意图:即所要求保护的本发明要求比在每个权利要求中所明确记载的特征更多的特征。更确切地说,如权利要求书所反映的那样,发明方面在于少于前面公开的单个实施例的所有特征。因此,遵循具体实施方式的权利要求书由此明确地并入该具体实施方式,其中每个权利要求本身都作为本发明的单独实施例。

[0135] 本领域那些技术人员可以理解,可以对实施例中的设备中的模块进行自适应性地改变并且把它们设置在与该实施例不同的一个或多个设备中。可以把实施例中的模块或单元或组件组合成一个模块或单元或组件,以及此外可以把它分成多个子模块或子单元或子组件。除了这样的特征和/或过程或者单元中的至少一些是相互排斥之外,可以采用任何组合对本说明书(包括伴随的权利要求、摘要和附图)中公开的所有特征以及如此公开的任何方法或者设备的所有过程或单元进行组合。除非另外明确陈述,本说明书(包括伴随的权

利要求、摘要和附图)中公开的每个特征可以由提供相同、等同或相似目的的替代特征来代替。

[0136] 此外,本领域的技术人员能够理解,尽管在此所述的一些实施例包括其它实施例中包括的某些特征而不是其它特征,但是不同实施例的特征的组合意味着处于本发明的范围之内并且形成不同的实施例。例如,在权利要求书中,所要求保护的实施例的任意之一都可以以任意的组合方式来使用。

[0137] 应该注意的是上述实施例对本发明进行说明而不是对本发明进行限制,并且本领域技术人员在不脱离所附权利要求的范围的情况下可设计出替换实施例。在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。单词“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本发明可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。单词第一、第二、以及第三等的使用不表示任何顺序。可将这些单词解释为名称。

[0138] 本发明公开了:A1.一种书籍推荐排序模型的训练方法,包括:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签;其中,所述第二时间段包含所述第一时间段;计算所述第一书籍子集的第一特征向量以及所述第二书籍子集中的各个第二书籍的第二特征向量;计算各个第二特征向量与所述第一特征向量之间的相似度;根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0139] A2.根据A1所述的方法,所述针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签进一步包括:对该用户的交互行为数据进行数据分析,得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍;将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中;从该用户在所述第一时间段内交互过的书籍中筛除所述第一书籍子集中的各个第一书籍,将筛除后得到的书籍作为第二书籍添加至第二书籍子集中;对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析,确定所述第二书籍子集中的各个第二书籍的交互时长标签。

[0140] A3.根据A1所述的方法,所述计算所述第一书籍子集的第一特征向量进一步包括:对所述第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取,得到各个第一书籍的书籍特征向量;根据各个第一书籍的书籍特征向量,计算所述第一书籍子集的第一特征向量。

[0141] A4.根据A3所述的方法,所述方法还包括:对该用户在指定时刻之前产生的交互行为数据进行数据分析,确定所述第一书籍子集中的各个第一书籍对应的交互时长;所述根据各个第一书籍的书籍特征向量,计算所述第一书籍子集的第一特征向量具体为:根据各个第一书籍对应的交互时长,确定各个第一书籍对应的权重值;将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算,得到所述第一书籍子集的第一特征向量。

[0142] A5.根据A1所述的方法,计算所述第二书籍子集中的各个第二书籍的第二特征向量进一步包括:对所述第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量。

[0143] A6.根据A1-A5任一项所述的方法,所述计算各个第二特征向量与所述第一特征向量之间的相似度进一步包括:计算任一第二特征向量与所述第一特征向量之间的夹角的余弦值,利用所述余弦值确定该第二特征向量与所述第一特征向量之间的相似度。

[0144] A7.根据A1-A6任一项所述的方法,所述根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型进一步包括:针对任一用户,将该用户的各个第二特征向量与所述第一特征向量之间的相似度输入至初始训练模型中,利用所述初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;根据各个第二书籍在所述书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益,根据所述归一化折损累计增益更新所述初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。

[0145] A8.根据A1-A7任一项所述的方法,所述方法还包括:获取所述用户集合中的各个用户的用户特征向量;所述根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型具体为:根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度、各个第二书籍的交互时长标签以及各个用户的用户特征向量,训练得到书籍推荐排序模型。

[0146] 本发明还公开了:B9.一种书籍推荐方法,包括:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定所述待推荐用户在当前时刻之前交互的第三书籍子集;对所述书籍集合中的书籍进行筛选,得到第四书籍子集;计算所述第三书籍子集的第三特征向量以及所述第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与所述第三特征向量之间的相似度;将各个第四特征向量与所述第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用所述书籍推荐排序模型对各个第四书籍进行排序,得到所述待推荐用户对应的书籍推荐排序列表。

[0147] B10.根据B9所述的方法,所述获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定所述待推荐用户在当前时刻之前交互的第三书籍子集进一步包括:对所述待推荐用户的交互行为数据进行数据分析,得到所述待推荐用户在当前时刻之前交互过的书籍;将所述待推荐用户在当前时刻之前交互过的书籍作为第三书籍添加至第三书籍子集中。

[0148] B11.根据B9所述的方法,所述方法还包括:获取所述待推荐用户的用户画像数据;所述对所述书籍集合中的书籍进行筛选,得到第四书籍子集具体为:在所述书籍集合中查找书籍画像数据与所述用户画像数据相匹配的书籍;从所述相匹配的书籍中筛除所述第三书籍子集中的各个第三书籍,将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0149] B12.根据B9所述的方法,所述计算所述第三书籍子集的第三特征向量进一步包括:对所述第三书籍子集中的各个第三书籍的全文或部分内容进行特征提取,得到各个第三书籍的书籍特征向量;根据各个第三书籍的书籍特征向量,计算所述第三书籍子集的第三特征向量。

[0150] B13. 根据B12所述的方法, 所述方法还包括: 对所述待推荐用户的交互行为数据进行数据分析, 确定所述第三书籍子集中的各个第三书籍对应的交互时长; 所述根据各个第三书籍的书籍特征向量, 计算所述第三书籍子集的第三特征向量具体为: 根据各个第三书籍对应的交互时长, 确定各个第三书籍对应的权重值; 将各个第三书籍的书籍特征向量与各个第三书籍对应的权重值进行加权运算, 得到所述第三书籍子集的第三特征向量。

[0151] B14. 根据B9所述的方法, 计算所述第四书籍子集中的各个第四书籍的第四特征向量进一步包括: 对所述第四书籍子集中的各个第四书籍的全文或部分内容进行特征提取, 得到各个第四书籍的第四特征向量。

[0152] B15. 根据B9-B14任一项所述的方法, 所述计算各个第四特征向量与所述第三特征向量之间的相似度进一步包括: 计算任一第四特征向量与所述第三特征向量之间的夹角的余弦值, 利用所述余弦值确定该第四特征向量与所述第三特征向量之间的相似度。

[0153] B16. 根据B9-B14任一项所述的方法, 所述书籍推荐排序模型基于A1至A8中任一项所述方法训练获得。

[0154] 本发明还公开了: C17. 一种计算设备, 包括: 处理器、存储器、通信接口和通信总线, 所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信; 所述存储器用于存放至少一可执行指令, 所述可执行指令使所述处理器执行以下操作: 获取用户集合中的用户针对书籍集合中的书籍的交互行为数据; 针对任一用户, 确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签; 其中, 所述第二时间段包含所述第一时间段; 计算所述第一书籍子集的第一特征向量以及所述第二书籍子集中的各个第二书籍的第二特征向量; 计算各个第二特征向量与所述第一特征向量之间的相似度; 根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签, 训练得到书籍推荐排序模型。

[0155] C18. 根据C17所述的计算设备, 所述可执行指令进一步使所述处理器执行以下操作: 对该用户的交互行为数据进行数据分析, 得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍; 将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中; 从该用户在所述第一时间段内交互过的书籍中筛选所述第一书籍子集中的各个第一书籍, 将筛选后得到的书籍作为第二书籍添加至第二书籍子集中; 对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析, 确定所述第二书籍子集中的各个第二书籍的交互时长标签。

[0156] C19. 根据C17所述的计算设备, 所述可执行指令进一步使所述处理器执行以下操作: 对所述第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取, 得到各个第一书籍的书籍特征向量; 根据各个第一书籍的书籍特征向量, 计算所述第一书籍子集的第一特征向量。

[0157] C20. 根据C19所述的计算设备, 所述可执行指令进一步使所述处理器执行以下操作: 对该用户在指定时刻之前产生的交互行为数据进行数据分析, 确定所述第一书籍子集中的各个第一书籍对应的交互时长; 根据各个第一书籍对应的交互时长, 确定各个第一书籍对应的权重值; 将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算, 得到所述第一书籍子集的第一特征向量。

[0158] C21.根据C17所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:对所述第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量。

[0159] C22.根据C17-C21任一项所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:计算任一第二特征向量与所述第一特征向量之间的夹角的余弦值,利用所述余弦值确定该第二特征向量与所述第一特征向量之间的相似度。

[0160] C23.根据C17-C22任一项所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:针对任一用户,将该用户的各个第二特征向量与所述第一特征向量之间的相似度输入至初始训练模型中,利用所述初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;根据各个第二书籍在所述书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益,根据所述归一化折损累计增益更新所述初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。

[0161] C24.根据C17-C23任一项所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:获取所述用户集合中的各个用户的用户特征向量;根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度、各个第二书籍的交互时长标签以及各个用户的用户特征向量,训练得到书籍推荐排序模型。

[0162] 本发明还公开了:D25.一种计算设备,包括:处理器、存储器、通信接口和通信总线,所述处理器、所述存储器和所述通信接口通过所述通信总线完成相互间的通信;所述存储器用于存放至少一可执行指令,所述可执行指令使所述处理器执行以下操作:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定所述待推荐用户在当前时刻之前交互的第三书籍子集;对所述书籍集合中的书籍进行筛选,得到第四书籍子集;计算所述第三书籍子集的第三特征向量以及所述第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与所述第三特征向量之间的相似度;将各个第四特征向量与所述第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用所述书籍推荐排序模型对各个第四书籍进行排序,得到所述待推荐用户对应的书籍推荐排序列表。

[0163] D26.根据D25所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:对所述待推荐用户的交互行为数据进行数据分析,得到所述待推荐用户在当前时刻之前交互过的书籍;将所述待推荐用户在当前时刻之前交互过的书籍作为第三书籍添加至第三书籍子集中。

[0164] D27.根据D25所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:获取所述待推荐用户的用户画像数据;在所述书籍集合中查找书籍画像数据与所述用户画像数据相匹配的书籍;从所述相匹配的书籍中筛除所述第三书籍子集中的各个第三书籍,将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0165] D28.根据D25所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:对所述第三书籍子集中的各个第三书籍的全文或部分内容进行特征提取,得到各个第三书籍的书籍特征向量;根据各个第三书籍的书籍特征向量,计算所述第三书籍子集的第三特征向量。

[0166] D29.根据D28所述的计算设备,所述可执行指令进一步使所述处理器执行以下操

作:对所述待推荐用户的交互行为数据进行数据分析,确定所述第三书籍子集中的各个第三书籍对应的交互时长;根据各个第三书籍对应的交互时长,确定各个第三书籍对应的权重值;将各个第三书籍的书籍特征向量与各个第三书籍对应的权重值进行加权运算,得到所述第三书籍子集的第三特征向量。

[0167] D30.根据D25所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:对所述第四书籍子集中的各个第四书籍的全文或部分内容进行特征提取,得到各个第四书籍的第四特征向量。

[0168] D31.根据D25-D30任一项所述的计算设备,所述可执行指令进一步使所述处理器执行以下操作:计算任一第四特征向量与所述第三特征向量之间的夹角的余弦值,利用所述余弦值确定该第四特征向量与所述第三特征向量之间的相似度。

[0169] D32.根据D25-D30任一项所述的计算设备,所述书籍推荐排序模型基于C17至C24中任一项所述计算设备训练获得。

[0170] 本发明还公开了:E33.一种存储介质,所述存储介质中存储有至少一可执行指令,所述可执行指令使处理器执行以下操作:获取用户集合中的用户针对书籍集合中的书籍的交互行为数据;针对任一用户,确定该用户在指定时刻之前交互的第一书籍子集、从指定时刻开始的第一时间段内交互的第二书籍子集以及从指定时刻开始的第二时间段内所述第二书籍子集中的各个第二书籍的交互时长标签;其中,所述第二时间段包含所述第一时间段;计算所述第一书籍子集的第一特征向量以及所述第二书籍子集中的各个第二书籍的第二特征向量;计算各个第二特征向量与所述第一特征向量之间的相似度;根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度以及各个第二书籍的交互时长标签,训练得到书籍推荐排序模型。

[0171] E34.根据E33所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对该用户的交互行为数据进行数据分析,得到该用户在指定时刻之前交互过的书籍以及在从指定时刻开始的第一时间段内交互过的书籍;将该用户在指定时刻之前交互过的书籍作为第一书籍添加至第一书籍子集中;从该用户在所述第一时间段内交互过的书籍中筛选所述第一书籍子集中的各个第一书籍,将筛选后得到的书籍作为第二书籍添加至第二书籍子集中;

[0172] 对该用户在从指定时刻开始的第二时间段内产生的交互行为数据进行数据分析,确定所述第二书籍子集中的各个第二书籍的交互时长标签。

[0173] E35.根据E33所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对所述第一书籍子集中的各个第一书籍的全文或部分内容进行特征提取,得到各个第一书籍的书籍特征向量;根据各个第一书籍的书籍特征向量,计算所述第一书籍子集的第一特征向量。

[0174] E36.根据E35所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对该用户在指定时刻之前产生的交互行为数据进行数据分析,确定所述第一书籍子集中的各个第一书籍对应的交互时长;根据各个第一书籍对应的交互时长,确定各个第一书籍对应的权重值;将各个第一书籍的书籍特征向量与各个第一书籍对应的权重值进行加权运算,得到所述第一书籍子集的第一特征向量。

[0175] E37.根据E33所述的存储介质,所述可执行指令进一步使所述处理器执行以下操

作:对所述第二书籍子集中的各个第二书籍的全文或部分内容进行特征提取,得到各个第二书籍的第二特征向量。

[0176] E38.根据E33-E37任一项所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:计算任一第二特征向量与所述第一特征向量之间的夹角的余弦值,利用所述余弦值确定该第二特征向量与所述第一特征向量之间的相似度。

[0177] E39.根据E33-E38任一项所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:针对任一用户,将该用户的各个第二特征向量与所述第一特征向量之间的相似度输入至初始训练模型中,利用所述初始训练模型对各个第二书籍进行排序,得到该用户对应的书籍推荐排序列表;根据各个第二书籍在所述书籍推荐排序列表中的排序位置以及各个第二书籍的交互时长标签,计算归一化折损累计增益,根据所述归一化折损累计增益更新所述初始训练模型的权重参数;循环迭代执行上述步骤,直至满足迭代结束条件,得到书籍推荐排序模型。

[0178] E40.根据E33-E39任一项所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:获取所述用户集合中的各个用户的用户特征向量;根据各个用户的各个第二特征向量与所述第一特征向量之间的相似度、各个第二书籍的交互时长标签以及各个用户的用户特征向量,训练得到书籍推荐排序模型。

[0179] 本发明还公开了:F41.一种存储介质,所述存储介质中存储有至少一可执行指令,所述可执行指令使处理器执行以下操作:获取待推荐用户针对书籍集合中的书籍的交互行为数据,确定所述待推荐用户在当前时刻之前交互的第三书籍子集;对所述书籍集合中的书籍进行筛选,得到第四书籍子集;计算所述第三书籍子集的第三特征向量以及所述第四书籍子集中的各个第四书籍的第四特征向量;计算各个第四特征向量与所述第三特征向量之间的相似度;将各个第四特征向量与所述第三特征向量之间的相似度输入至经过训练的书籍推荐排序模型中,利用所述书籍推荐排序模型对各个第四书籍进行排序,得到所述待推荐用户对应的书籍推荐排序列表。

[0180] F42.根据F41所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对所述待推荐用户的交互行为数据进行数据分析,得到所述待推荐用户在当前时刻之前交互过的书籍;将所述待推荐用户在当前时刻之前交互过的书籍作为第三书籍添加至第三书籍子集中。

[0181] F43.根据F41所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:获取所述待推荐用户的用户画像数据;在所述书籍集合中查找书籍画像数据与所述用户画像数据相匹配的书籍;从所述相匹配的书籍中筛除所述第三书籍子集中的各个第三书籍,将筛除后得到的书籍作为第四书籍添加至第四书籍子集中。

[0182] F44.根据F41所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对所述第三书籍子集中的各个第三书籍的全文或部分内容进行特征提取,得到各个第三书籍的书籍特征向量;根据各个第三书籍的书籍特征向量,计算所述第三书籍子集的第三特征向量。

[0183] F45.根据F44所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对所述待推荐用户的交互行为数据进行数据分析,确定所述第三书籍子集中的各个第三书籍对应的交互时长;根据各个第三书籍对应的交互时长,确定各个第三书籍对应的权

重值;将各个第三书籍的书籍特征向量与各个第三书籍对应的权重值进行加权运算,得到所述第三书籍子集的第三特征向量。

[0184] F46.根据F41所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:对所述第四书籍子集中的各个第四书籍的全文或部分内容进行特征提取,得到各个第四书籍的第四特征向量。

[0185] F47.根据F41-F46任一项所述的存储介质,所述可执行指令进一步使所述处理器执行以下操作:计算任一第四特征向量与所述第三特征向量之间的夹角的余弦值,利用所述余弦值确定该第四特征向量与所述第三特征向量之间的相似度。

[0186] F48.根据F41-F46任一项所述的存储介质,所述书籍推荐排序模型基于E33至E40中任一项所述存储介质训练获得。

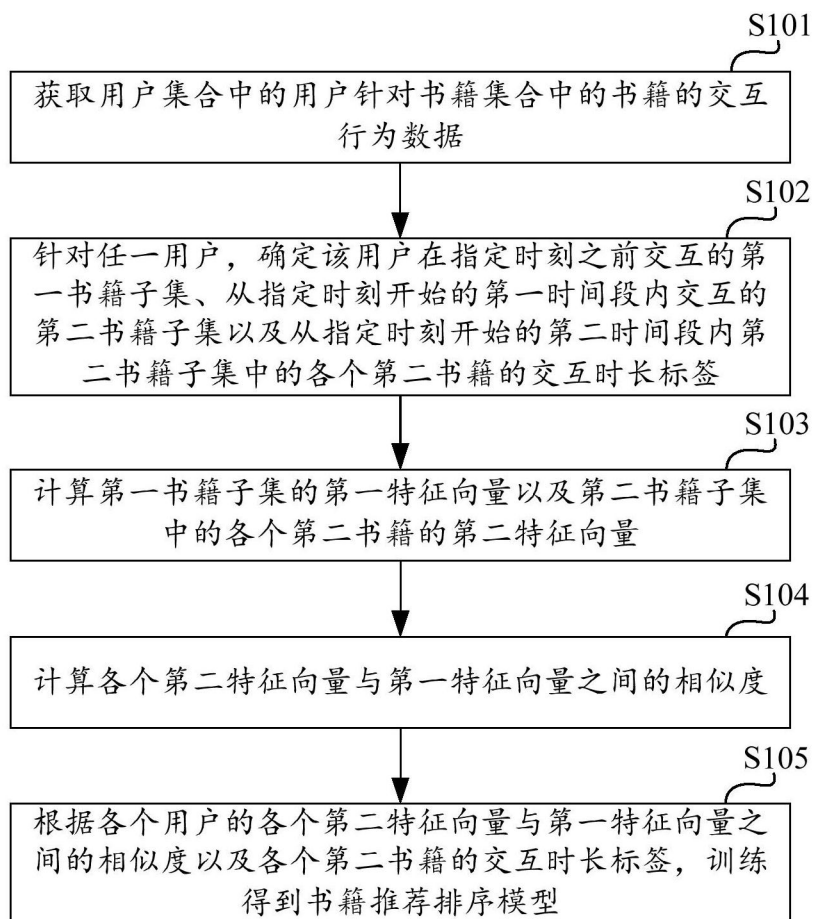


图1

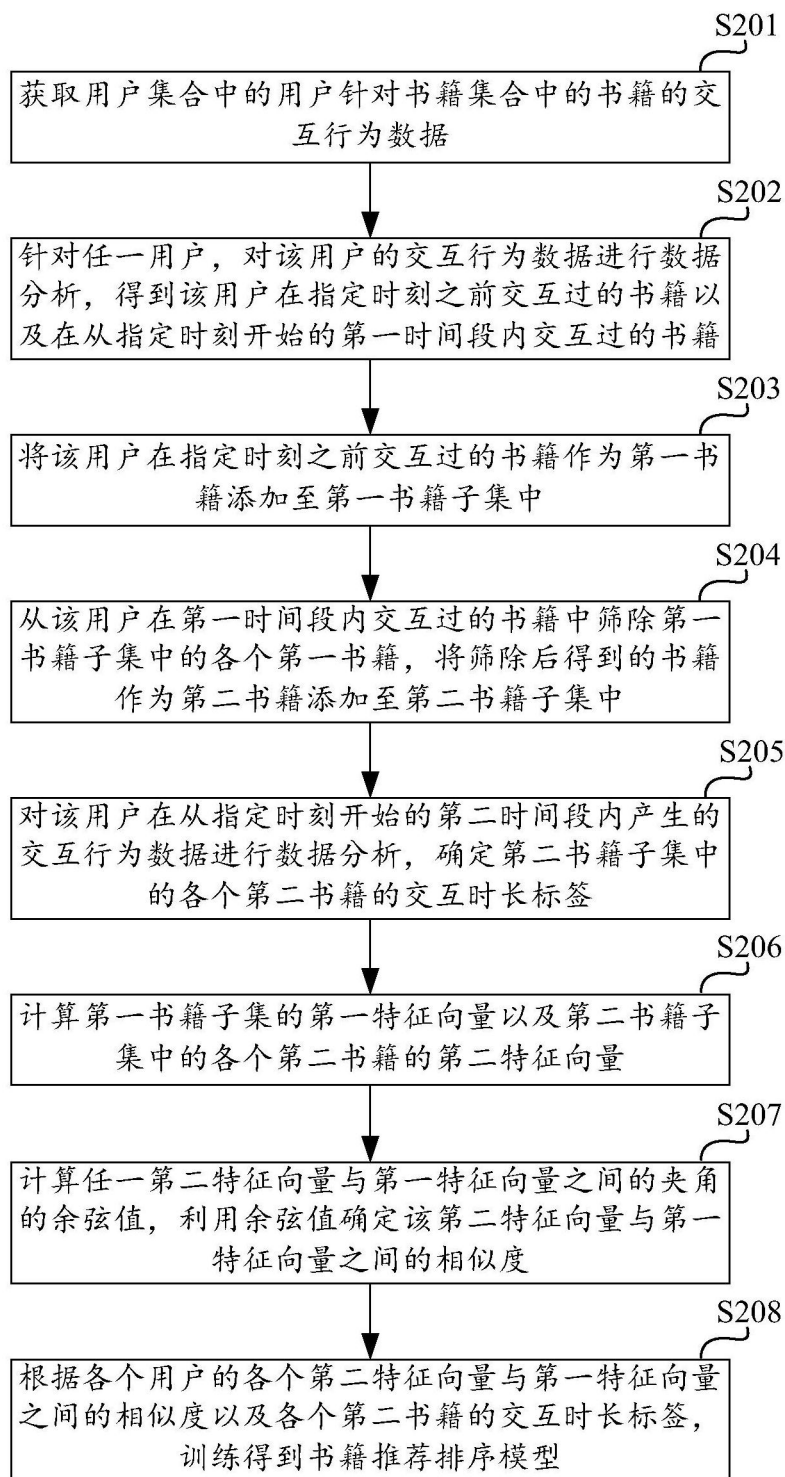


图2

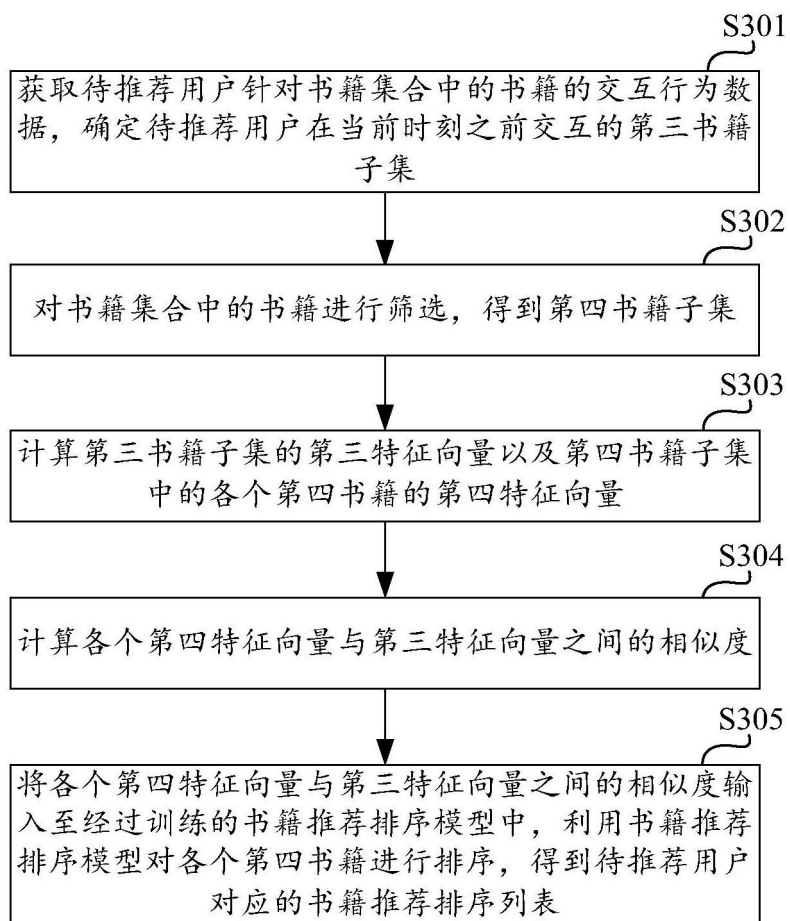


图3

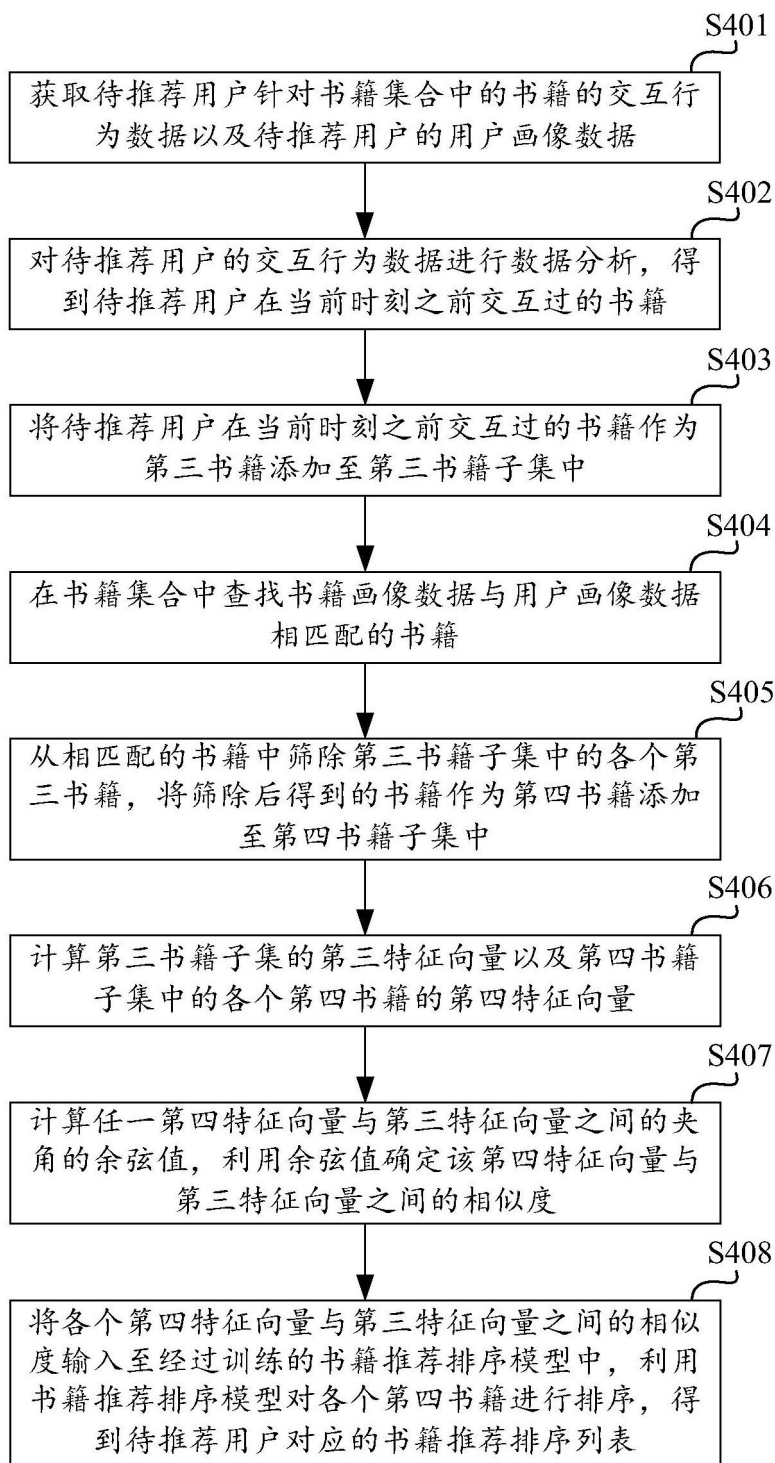


图4

