



(12) 发明专利申请

(10) 申请公布号 CN 119295012 A

(43) 申请公布日 2025. 01. 10

(21) 申请号 202411414419.1

G06F 18/21 (2023.01)

(22) 申请日 2024.10.11

G06F 18/23 (2023.01)

(71) 申请人 河南中医药大学

地址 450000 河南省郑州市郑东新区金水
东路156号(72) 发明人 任妍 何明举 黄静 段涛
马亚玲(74) 专利代理机构 河南豫龙律师事务所 41177
专利代理师 张电臣

(51) Int. Cl.

G06Q 10/10 (2023.01)

G06Q 10/04 (2023.01)

G06Q 50/26 (2024.01)

G06F 16/2458 (2019.01)

G06F 18/15 (2023.01)

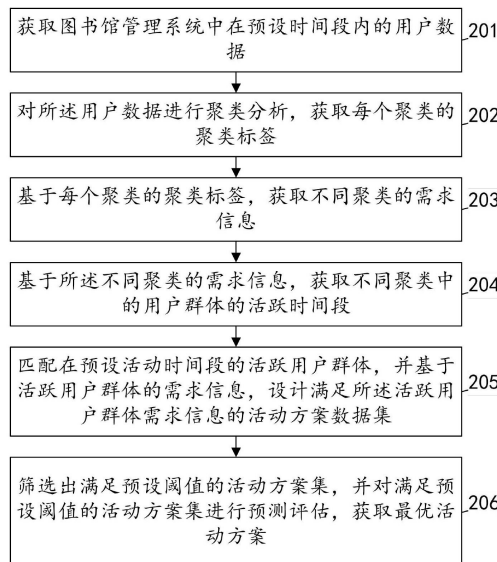
权利要求书2页 说明书9页 附图6页

(54) 发明名称

基于大数据分析的图书馆活动设计方法及
系统

(57) 摘要

本申请提供了一种基于大数据分析的图书馆活动设计方法及系统,所述方法包括:获取图书馆管理系统中在预设时间段内的用户数据;对所述用户数据进行聚类分析,获取每个聚类的聚类标签;基于每个聚类的聚类标签,获取不同聚类的需求信息;基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案,通过本申请提供的图书馆活动方案,极大提高了学生对图书馆活动的参与度,为学生提供丰富多样的学习和文化体验。



1. 基于大数据分析的图书馆活动设计方法,其特征在于,包括:
获取图书馆管理系统中在预设时间段内的用户数据;
对所述用户数据进行聚类分析,获取每个聚类的聚类标签;
基于每个聚类的聚类标签,获取不同聚类的需求信息;
基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;
匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;
筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案。
2. 根据权利要求1所述的基于大数据分析的图书馆活动设计方法,其特征在于,所述对所述用户数据进行聚类分析之前,还包括对所述用户数据进行清洗操作,获取清洗用户数据,将所述清洗用户数据中的非数值型数据转化为数值型数据,获取类型转换数据;将所述类型转换数据归一化到同一尺度,获取归一化数据。
3. 根据权利要求2所述的基于大数据分析的图书馆活动设计方法,其特征在于,所述对所述用户数据进行聚类分析,获取每个聚类的聚类标签,具体步骤为:
提取归一化数据的特征信息,基于轮廓系数确定所述特征信息的最佳聚类数量;
随机初始化最佳聚类数量的聚类中心;
将所述特征信息中的每个数据点分配给最近的聚类中心;获取每个聚类的均值,并更新聚类中心;重复执行此步骤,直至聚类中心不再变化或者达到最大迭代次数,获取聚类结果;
基于主成分分析法将聚类结果可视化,获取每个聚类的聚类标签。
4. 根据权利要求3所述的基于大数据分析的图书馆活动设计方法,其特征在于,所述基于轮廓系数确定所述特征信息的最佳聚类数量,具体包括:
基于特征数量设置聚类数范围;
对所述聚类数范围的每个聚类数,基于聚类算法对所述特征信息的数据点进行聚类;
获取每个数据点与同一簇内其他数据点的平均距离 $a(i)$,以及所述每个数据点与最近簇中数据点之间的平均距离 $b(i)$;
基于预设的公式获取每个数据点的轮廓系数 $s(i)$;
获取不同聚类数下的平均轮廓系数,基于所述轮廓系数的变化趋势,选择平均轮廓系数最高的聚类数作为最佳聚类数量。
5. 根据权利要求1所述的基于大数据分析的图书馆活动设计方法,其特征在于,所述基于每个聚类的聚类标签,获取不同聚类的需求信息,具体内容包括:
基于所述聚类标签提取关键特征;
将每个聚类中的数据点转化为预设格式数据;
基于每个聚类的关键特征及预设格式数据,对每个聚类的用户数据进行关联规则挖掘,生成满足支持度阈值的频繁项集;
从所述频繁项集中生成关联需求,筛选出满足最小置信度的需求,获取不同聚类的需求信息。
6. 根据权利要求1所述的基于大数据分析的图书馆活动设计方法,其特征在于,所述基

于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段,具体内容包括:

获取每个聚类中每个用户的时间数据;

基于每个聚类的聚类标签和时间数据,计算每个小时内的借阅次数;

计算所述借阅次数的均值和标准差,将超过预设阈值的时间段作为活跃时间段。

7.根据权利要求1所述的基于大数据分析的图书馆活动设计方法,其特征在于,所述筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案,具体内容包括:

确定活动方案集的评估指标;

基于每个评估指标预设的阈值筛选出符合阈值的活动方案,获取候选方案集;

基于层次分析法为每个候选方案的不同评估指标分配权重;

基于预设的评估模型获取每个候选方案不同指标的评估值,基于不同指标的权重及评估值,获取不同指标的综合评分;

将所有候选活动方案按照综合评分从高到低排序,选择评分最高的方案作为最优方案。

8.一种基于大数据分析的图书馆活动设计系统,用于实现权利要求1-7的基于大数据分析的图书馆活动设计方法,其特征在于,包括:

用户数据获取模块,用于获取图书馆管理系统中在预设时间段内的用户数据;

用户特征获取模块,用于对所述用户数据进行聚类分析,获取每个聚类的聚类标签;

需求信息获取模块,用于基于每个聚类的聚类标签,获取不同聚类的需求信息;

活跃时间获取模块,用于基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;

活动方案数据集获取模块,用于匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;

最优活动方案获取模块,用于筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案。

9.一种电子设备,其特征在于,包括:

一个或多个处理器;存储器;以及一个或多个计算机程序,其中所述一个或多个计算机程序被存储在所述存储器中,所述一个或多个计算机程序包括指令,当所述指令被所述设备执行时,使得所述设备执行如权利要求1至7任一项所述的基于大数据分析的图书馆活动设计方法的步骤。

10.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有计算机程序,当其在计算机上运行时,使得计算机执行如权利要求1至7任一项所述的基于大数据分析的图书馆活动设计方法的步骤。

基于大数据分析的图书馆活动设计方法及系统

技术领域

[0001] 本申请涉及图书馆活动设计技术领域,具体说是一种基于大数据分析的图书馆活动设计方法及系统。

背景技术

[0002] 当前,图书馆作为校园内承载着知识传承与学术交流的重要场所,其现状与潜能之间存在着显著的差距,图书馆虽深受一部分热爱学习、渴望知识的学生青睐,成为他们探索未知、深化理解的场所,但部分学生因缺乏对图书馆功能的全面了解,或是未能深刻认识到图书馆资源对于个人成长的巨大助力,错失了图书馆这一学习场所。学生对图书馆认识上的不足导致了图书馆参与度的两极分化,不利于形成全校范围内浓厚的学习氛围。

[0003] 图书馆作为校园文化生活的重要组成部分,本应是一个充满活力与创意的交流空间,能够定期举办多样化的阅读推广活动、学术讲座、文化沙龙等,以吸引并满足不同学生的兴趣和需求。然而,现状却是活动稀少甚至几乎为零,减少了学生了解图书馆、接近图书馆的途径,限制了图书馆在促进学生全面发展、拓宽知识视野方面所能发挥的作用。因此,如何克服上述存在的技术问题和缺陷成为需要重点解决的问题。

发明内容

[0004] 为了克服现有技术中活动稀少甚至几乎为零,减少了学生了解图书馆、接近图书馆的问题,本申请提供了一种基于大数据分析的图书馆活动设计方法及系统,采用了如下的技术方案:

[0005] 第一方面,本申请提供了一种基于大数据分析的图书馆活动设计方法,包括:

[0006] 获取图书馆管理系统中在预设时间段内的用户数据;

[0007] 对所述用户数据进行聚类分析,获取每个聚类的聚类标签;

[0008] 基于每个聚类的聚类标签,获取不同聚类的需求信息;

[0009] 基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;

[0010] 匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;

[0011] 筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案。

[0012] 进一步地,所述对所述用户数据进行聚类分析之前,还包括对所述用户数据进行清洗操作,获取清洗用户数据,将所述清洗用户数据中的非数值型数据转化为数值型数据,获取类型转换数据;将所述类型转换数据归一化到同一尺度,获取归一化数据。

[0013] 进一步地,所述对所述用户数据进行聚类分析,获取每个聚类的聚类标签,具体步骤为:

[0014] 提取归一化数据的特征信息,基于轮廓系数确定所述特征信息的最佳聚类数量;

[0015] 随机初始化最佳聚类数量的聚类中心;

[0016] 将所述特征信息中的每个数据点分配给最近的聚类中心;获取每个聚类的均值,并更新聚类中心;重复执行此步骤,直至聚类中心不再变化或者达到最大迭代次数,获取聚类结果;

[0017] 基于主成分分析法将聚类结果可视化,获取每个聚类的聚类标签。

[0018] 进一步地,所述基于轮廓系数确定所述特征信息的最佳聚类数量,具体包括:

[0019] 基于特征数量设置聚类数范围;

[0020] 对所述聚类数范围的每个聚类数,基于聚类算法对所述特征信息的数据点进行聚类;

[0021] 获取每个数据点与同一簇内其他数据点的平均距离 $a(i)$,以及所述每个数据点与最近簇中数据点之间的平均距离 $b(i)$;

[0022] 基于预设的公式获取每个数据点的轮廓系数 $s(i)$;

[0023] 获取不同聚类数下的平均轮廓系数,基于所述轮廓系数的变化趋势,选择平均轮廓系数最高的聚类数作为最佳聚类数量。

[0024] 进一步地,预设的公式为 $s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$,其中所述轮廓系数的值在-1和1之间,所述轮廓系数的值越接近1表示聚类效果越好,所述轮廓系数的值接近0表示数据点在簇的边缘,负值表示聚类效果较差。

[0025] 进一步地,所述基于每个聚类的聚类标签,获取不同聚类的需求信息,具体内容包括:

[0026] 基于所述聚类标签提取关键特征;

[0027] 将每个聚类中的数据点转化为预设格式数据;

[0028] 基于每个聚类的关键特征及预设格式数据,对每个聚类的用户数据进行关联规则挖掘,生成满足支持度阈值的频繁项集;

[0029] 从所述频繁项集中生成关联需求,筛选出满足最小置信度的需求,获取不同聚类的需求信息。

[0030] 进一步地,所述基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段,具体内容包括:

[0031] 获取每个聚类中每个用户的时间数据;

[0032] 基于每个聚类的聚类标签和时间数据,计算每个小时内的借阅次数;

[0033] 计算所述借阅次数的均值和标准差,将超过预设阈值的时间段作为活跃时间段。

[0034] 进一步地,所述筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案,具体内容包括:

[0035] 确定活动方案集的评估指标;

[0036] 基于每个评估指标预设的阈值筛选出符合阈值的活动方案,获取候选方案集;

[0037] 基于层次分析法为每个候选方案的不同评估指标分配权重;

[0038] 基于预设的评估模型获取每个候选方案不同指标的评估值,基于不同指标的权重及评估值,获取不同指标的综合评分;

[0039] 将所有候选活动方案按照综合评分从高到低排序,选择评分最高的方案作为最优方案。

- [0040] 第二方面,本申请还提供了一种基于大数据分析的图书馆活动设计系统,包括:
- [0041] 用户数据获取模块,用于获取图书馆管理系统中在预设时间段内的用户数据;
- [0042] 用户特征获取模块,用于对所述用户数据进行聚类分析,获取每个聚类的聚类标签;
- [0043] 需求信息获取模块,用于基于每个聚类的聚类标签,获取不同聚类的需求信息;
- [0044] 活跃时间获取模块,用于基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;
- [0045] 活动方案数据集获取模块,用于匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;
- [0046] 最优活动方案获取模块,用于筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案。
- [0047] 第三方面,本申请提供了一种电子设备,包括:
- [0048] 一个或多个处理器;存储器;以及一个或多个计算机程序,其中所述一个或多个计算机程序被存储在所述存储器中,所述一个或多个计算机程序包括指令,当所述指令被所述设备执行时,使得所述设备执行如第一方面所述的方法。
- [0049] 第四方面,本申请提供了一种计算机可读存储介质,该计算机可读存储介质中存储有计算机程序,当其在计算机上运行时,使得计算机执行如第一方面所述的方法。
- [0050] 第五方面,本申请提供了一种计算机程序,当所述计算机程序被计算机执行时,用于执行第一方面所述的方法。
- [0051] 在一种可能的设计中,第五方面中的程序可以全部或者部分存储在与处理器封装在一起的存储介质上,也可以部分或者全部存储在不与处理器封装在一起的存储器上。
- [0052] 本申请具有以下有益效果:
- [0053] 1. 本申请通过获取图书馆管理系统中在预设时间段内的用户数据;对所述用户数据进行聚类分析,获取每个聚类的聚类标签;基于每个聚类的聚类标签,构建每个聚类的用户画像,获取不同聚类的需求信息。本申请通过聚类分析将用户划分为不同的群体,进而识别出不同群体用户的特征,使得图书馆在进行活动制定的过程中能更精准地理解用户的阅读习惯和兴趣,进而为图书馆活动设计提供个性化的数据支撑。
- [0054] 2. 本申请通过基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案。本申请通过获取不同用户群体的活跃时间段,使得在进行图书馆活动设计的过程中能够更精准地针对特定用户群体;基于活跃用户群体的需求信息设计的图书馆活动方案,能够更好地吸引目标用户的参与兴趣,提升用户在活动中的体验感;通过筛选满足预设阈值的活动方案可以避免不必要的资源浪费。通过数据分析使得图书馆活动方案的设计决策更加科学、客观。本申请极大提高了学生对图书馆活动的参与度,为学生提供丰富多样的学习和文化体验。

附图说明

- [0055] 图1为本申请实施例可以应用于其中的示例性系统架构图;

- [0056] 图2为本申请实施例的基于大数据分析的图书馆活动设计方法流程图；
- [0057] 图3为本申请实施例的获取每个聚类的聚类标签的流程图；
- [0058] 图4为本申请实施例的确定最佳聚类数量的流程图；
- [0059] 图5为本申请实施例的获取不同聚类的需求信息的流程图；
- [0060] 图6为本申请实施例的确定不同聚类中的用户群体的活跃时间段的流程图；
- [0061] 图7为本申请实施例的获取最优活动方案的流程图；
- [0062] 图8为本申请实施例的系统流程图；
- [0063] 图9是本申请实施例的计算机设备示意图。

具体实施方式

[0064] 除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同;本文中在申请的说明书中所使用的术语只是为了描述具体的实施例的目的,不是旨在于限制本申请;本申请的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形,意图在于覆盖不排他的包含。本申请的说明书和权利要求书或上述附图中的术语“第一”、“第二”等是用于区别不同对象,而不是用于描述特定顺序。

[0065] 在本文中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是,本文所描述的实施例可以与其它实施例相结合。

[0066] 为了使本技术领域的人员更好地理解本申请方案,下面将结合附图,对本申请实施例中的技术方案进行清楚、完整地描述。

[0067] 如图1所示,系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0068] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

[0069] 终端设备101、102、103可以是具有显示屏并且支持网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、MP3播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器、膝上型便携计算机和台式计算机等等。

[0070] 服务器105可以是提供各种服务的服务器,例如对终端设备101、102、103上显示的页面提供支持的后台服务器。

[0071] 需要说明的是,本申请实施例所提供的基于大数据分析的图书馆活动设计方法一般由服务器/终端设备执行,相应地,基于大数据分析的图书馆活动设计系统一般设置于服务器/终端设备中。

[0072] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需

要,可以具有任意数目的终端设备、网络和服务端。

[0073] 继续参考图2,图中示出了本申请的一种基于大数据分析的图书馆活动设计方法流程图,所述方法包括以下步骤:

[0074] 步骤201,获取图书馆管理系统中在预设时间段内的用户数据。

[0075] 在本申请实施例中,获取所述用户数据的字段包括用户ID、借阅时间、书籍ID、书籍类型,其中用户ID表示唯一标识一个用户的编号;借阅时间表示用户借阅书籍的具体日期;书籍ID表示唯一标识一本书的编号;书籍类型表示书籍的分类,如小说、科技、历史等;书籍标题表示书籍的名称。例如:通过学校的图书馆管理系统,获取三个月内的用户数据;本申请实施例中的预设时间段根据实际需求设置。通过在预设时间段内获取图书馆管理系统中的用户数据,可以帮助了解用户的阅读偏好,进而为图书馆活动设计提供个性化的数据支撑。

[0076] 步骤202,对所述用户数据进行聚类分析,获取每个聚类的聚类标签。

[0077] 在本申请实施例中,对所述用户数据进行聚类分析之前,还包括对所述用户数据进行清洗操作,获取清洗用户数据,将所述清洗用户数据中的非数值型数据转化为数值型数据,获取类型转换数据;将所述类型转换数据归一化到同一尺度,获取归一化数据。

[0078] 在本申请实施例中,将所述类型转换数据归一化到同一尺度,获取归一化数据,可以使用Z-score标准化或Min-Max缩放对所述类型转换数据进行归一化处理。

[0079] 在本申请实施例中,对所述用户数据进行聚类分析,获取每个聚类的聚类标签请参考图3,具体步骤为:

[0080] 步骤301,提取归一化数据的特征信息,基于轮廓系数确定所述特征信息的最佳聚类数量。

[0081] 步骤302,随机初始化最佳聚类数量的聚类中心。

[0082] 步骤303,将所述特征信息中的每个数据点分配给最近的聚类中心;获取每个聚类的均值,并更新聚类中心;重复执行此步骤,直至聚类中心不再变化或者达到最大迭代次数,获取聚类结果。

[0083] 步骤304,基于主成分分析法将聚类结果可视化,获取每个聚类的聚类标签。比如某些聚类可能偏好特定类型的书籍或在特定时间段借阅。

[0084] 在本申请实施例中,所述提取所述归一化数据的特征信息,包括:将所述借阅时间作为所述特征信息的周期性特征,将所述书籍ID作为所述书籍类型的分布特征,将所述书籍类型作为用户偏好类型的特征。

[0085] 在本申请实施例中,所述基于轮廓系数确定所述特征信息的最佳聚类数量请参考图4,具体包括:

[0086] 步骤401,基于特征数量设置聚类数范围;

[0087] 步骤402,对所述聚类数范围的每个聚类数,基于聚类算法对所述特征信息的数据点进行聚类;

[0088] 步骤403,获取每个数据点与同一簇内其他数据点的平均距离 $a(i)$,以及所述每个数据点与最近簇中数据点之间的平均距离 $b(i)$ 。

[0089] 步骤404,基于预设的公式获取每个数据点的轮廓系数 $s(i)$;

[0090] 步骤405,获取不同聚类数下的平均轮廓系数,基于所述轮廓系数的变化趋势,选

择平均轮廓系数最高的聚类数作为最佳聚类数量。

[0091] 所述预设的公式为 $s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$, 其中所述轮廓系数的值在-1和1之间, 所述轮廓系数的值越接近1表示聚类效果越好, 所述轮廓系数的值接近0表示数据点在簇的边缘, 负值表示聚类效果较差。

[0092] 步骤203, 基于每个聚类的聚类标签, 获取不同聚类的需求信息。

[0093] 在本申请实施例中, 所述基于每个聚类的聚类标签, 获取不同聚类的需求信息, 请参考图5, 具体步骤包括:

[0094] 步骤501, 基于所述聚类标签提取关键特征; 关键特征包括但不限于基本特征、行为特征、兴趣特征等; 其中基本特征可以为平均年龄、性别比例、教育程度等; 行为特征可以为借阅频率、借阅书籍的种类、借阅时长等; 兴趣特征可以为常借书籍类型、用户反馈(如评论、评分)等。

[0095] 步骤502, 将每个聚类中的数据点转化为预设格式数据, 预设格式数据可以设置为 Data1: {年龄:25, 性别:男, 借阅书籍:小说} Data1;2: {年龄:30, 性别:女, 借阅书籍:科幻}。

[0096] 步骤503, 基于每个聚类的关键特征及预设格式数据, 对每个聚类的用户数据进行关联规则挖掘, 生成满足支持度阈值的频繁项集;

[0097] 步骤504, 从所述频繁项集中生成关联需求, 筛选出满足最小置信度的需求, 获取不同聚类的需求信息。

[0098] 步骤204, 基于所述不同聚类的需求信息, 获取不同聚类中的用户群体的活跃时间段, 请参考图6, 具体内容包括:

[0099] 步骤601, 获取每个聚类中每个用户的时间数据。

[0100] 步骤602, 基于每个聚类的聚类标签和时间数据, 计算每个小时内的借阅次数; 例如可以采用: 活动数据.groupby(['聚类标签', '小时']).count() 来统计每个小时的借阅次数。

[0101] 步骤603, 计算所述借阅次数的均值和标准差, 将超过预设阈值的时间段作为活跃时间段。

[0102] 步骤205, 匹配在预设活动时间段的活跃用户群体, 并基于活跃用户群体的需求信息, 设计满足所述活跃用户群体需求信息的活动方案数据集。

[0103] 在本申请实施例中, 将预设活动时间段作为查询条件, 筛选出满足预设活动时间段的活跃用户群体, 并获取满足活跃用户群体的需求信息; 基于活跃用户群体的需求信息, 确定活动方案, 其中活动方案包括活动类型(如讲座、阅读分享、培训等)、活动主题及内容、活动时间、活动地点及活动形式(线上或者线下)。

[0104] 步骤206, 筛选出满足预设阈值的活动方案集, 并对满足预设阈值的活动方案集进行预测评估, 获取最优活动方案。

[0105] 在本申请实施例中, 筛选出满足预设阈值的活动方案集, 并对满足预设阈值的活动方案集进行预测评估, 获取最优活动方案, 请参考图7, 具体内容包括:

[0106] 步骤701, 确定活动方案集的评估指标, 所述评估指标可以为预期参与人数、用户满意度、成本效益比、资源利用率等;

[0107] 步骤702, 基于每个评估指标预设的阈值筛选出符合阈值的活动方案, 获取候选方

案集;

[0108] 步骤703,基于层次分析法为每个候选方案的不同评估指标分配权重;

[0109] 步骤704,基于预设的评估模型获取每个候选方案不同指标的评估值,基于不同指标的权重及评估值,获取不同指标的综合评分。

[0110] 举例地,假设评估指标为预期参与人数、成本效益比、用户满意度,则计算这些指标的综合评分方式为:综合评分= $w_1 \times \text{预期参与人数} + w_2 \times \text{成本效益比} + w_3 \times \text{用户满意度}$,其中, w_1 、 w_2 、 w_3 是预期参与人数、成本效益比、用户满意度的权重。

[0111] 步骤705,将所有候选活动方案按照综合评分从高到低排序,选择评分最高的方案作为最优方案。

[0112] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,该计算机程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,前述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)等非易失性存储介质,或随机存储记忆体(Random Access Memory,RAM)等。

[0113] 应该理解的是,虽然附图的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,其可以以其他的顺序执行。而且,附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,其执行顺序也不必然是依次进行,而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0114] 继续参考图8,本实施例所述的基于大数据分析的图书馆活动设计系统包括:用户数据获取模块801,用户特征获取模块802,需求信息获取模块803,活跃时间获取模块804,活动方案数据集获取模块805,最优活动方案获取模块806。其中:

[0115] 用户数据获取模块801,用于获取图书馆管理系统中在预设时间段内的用户数据;

[0116] 用户特征获取模块802,用于对所述用户数据进行聚类分析,获取每个聚类的聚类标签;

[0117] 需求信息获取模块803,用于基于每个聚类的聚类标签,获取不同聚类的需求信息;

[0118] 活跃时间获取模块804,用于基于所述不同聚类的需求信息,获取不同聚类中的用户群体的活跃时间段;

[0119] 活动方案数据集获取模块805,用于匹配在预设活动时间段的活跃用户群体,并基于活跃用户群体的需求信息,设计满足所述活跃用户群体需求信息的活动方案数据集;

[0120] 最优活动方案获取模块806,用于筛选出满足预设阈值的活动方案集,并对满足预设阈值的活动方案集进行预测评估,获取最优活动方案。

[0121] 为解决上述技术问题,本申请实施例还提供计算机设备。具体请参阅图9,图9为本实施例计算机设备基本结构框图。

[0122] 所述计算机设备9包括通过系统总线相互通信连接存储器9a、处理器9b、网络接口9c。需要指出的是,图中仅示出了具有组件9a-9c的计算机设备9,但是应理解的是,并不要求实施所有示出的组件,可以替代的实施更多或者更少的组件。其中,本技术领域技术人员

可以理解,这里的计算机设备是一种能够按照事先设定或存储的指令,自动进行数值计算和/或信息处理的设备,其硬件包括但不限于微处理器、专用集成电路(Application Specific Integrated Circuit,ASIC)、可编程门阵列(Field-Programmable GateArray,FPGA)、数字处理器(Digital Signal Processor,DSP)、嵌入式设备等。

[0123] 所述计算机设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述计算机设备可以与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互。

[0124] 所述存储器9a至少包括一种类型的可读存储介质,所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器(例如,SD或DX存储器等)、随机访问存储器(RAM)、静态随机访问存储器(SRAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、可编程只读存储器(PROM)、磁性存储器、磁盘、光盘等。在一些实施例中,所述存储器9a可以是所述计算机设备9的内部存储单元,例如该计算机设备9的硬盘或内存。在另一些实施例中,所述存储器9a也可以是所述计算机设备9的外部存储设备,例如该计算机设备9上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。当然,所述存储器9a还可以既包括所述计算机设备9的内部存储单元也包括其外部存储设备。本实施例中,所述存储器9a通常用于存储安装于所述计算机设备9的操作系统和各类应用软件,例如基于大数据分析的图书馆活动设计方法的程序代码等。此外,所述存储器9a还可以用于暂时地存储已经输出或者将要输出的各类数据。

[0125] 所述处理器9b在一些实施例中可以是中央处理器(Central Processing Unit,CPU)、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器9b通常用于控制所述计算机设备9的总体操作。本实施例中,所述处理器9b用于运行所述存储器9a中存储的程序代码或者处理数据,例如运行所述基于大数据分析的图书馆活动设计方法的程序代码。

[0126] 所述网络接口9c可包括无线网络接口或有线网络接口,该网络接口9c通常用于在所述计算机设备9与其他电子设备之间建立通信连接。

[0127] 本申请还提供了另一种实施方式,即提供一种非易失性计算机可读存储介质,所述非易失性计算机可读存储介质存储有基于大数据分析的图书馆活动设计方法的程序,所述基于大数据分析的图书馆活动设计可被至少一个处理器执行,以使所述至少一个处理器执行如上述的基于大数据分析的图书馆活动设计方法的步骤。

[0128] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本申请各个实施例所述的方法。

[0129] 显然,以上所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例,附图中给出了本申请的较佳实施例,但并不限制本申请的专利范围。本申请可以以许多不同的形式来实现,相反地,提供这些实施例的目的是使对本申请的公开内容的理解更加透彻全面。尽管参照前述实施例对本申请进行了详细的说明,对于本领域的技术人员来而言,其依然可以对前述各具体实施方式所记载的技术方案进行修改,或者对其中部分技术特征进

行等效替换。凡是利用本申请说明书及附图内容所做的等效结构,直接或间接运用在其他相关的技术领域,均同理在本申请专利保护范围之内。

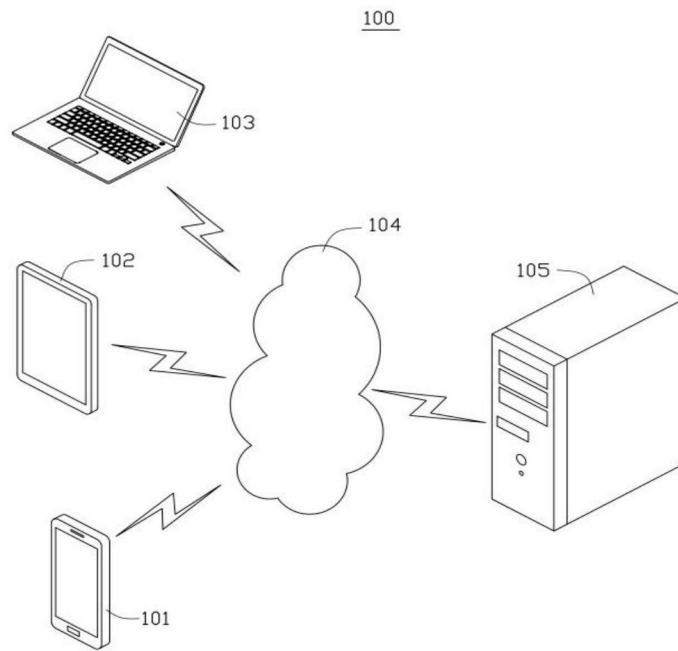


图1

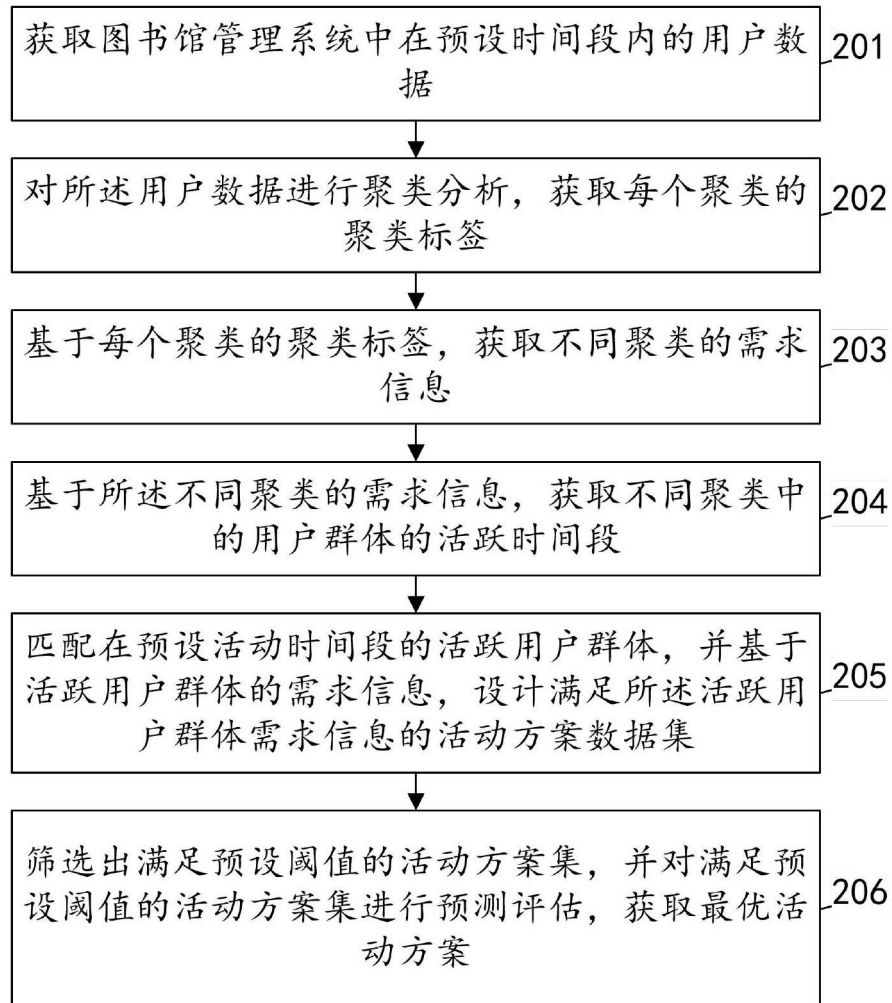


图2

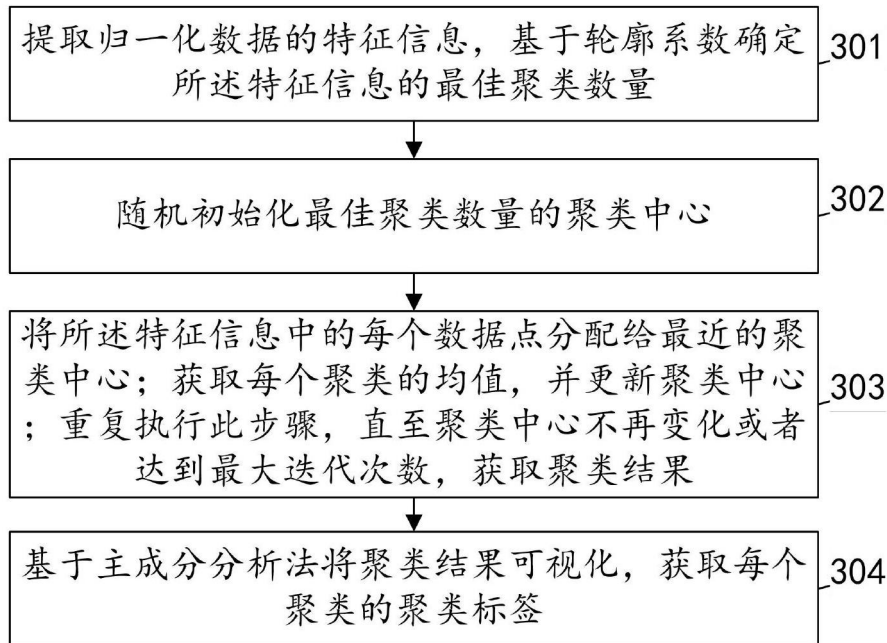


图3

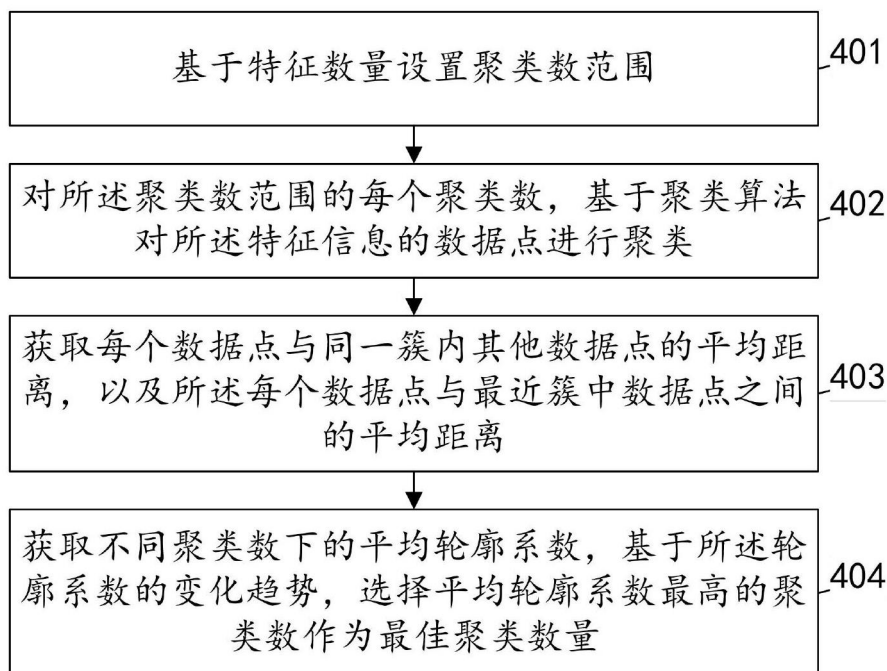


图4

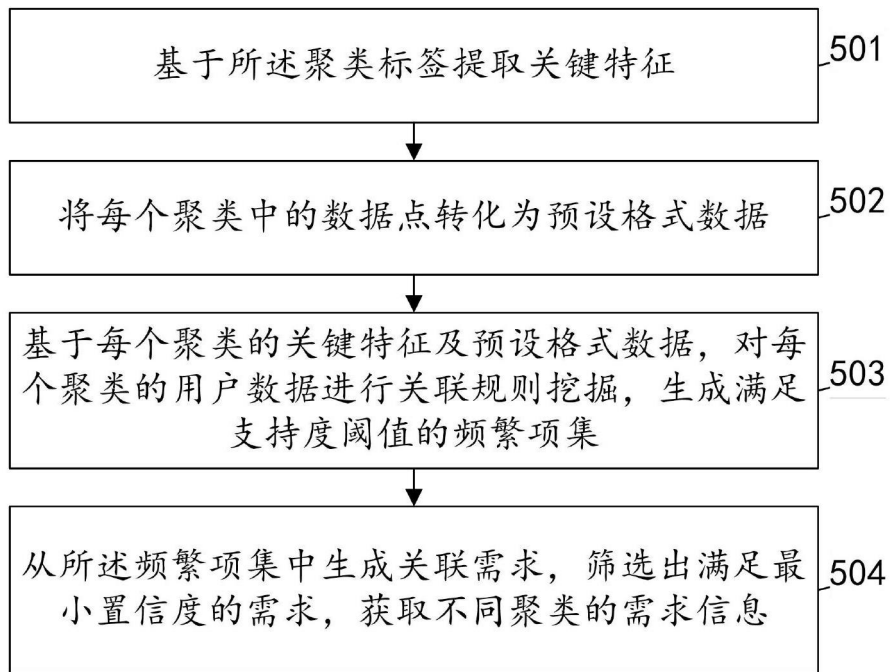


图5

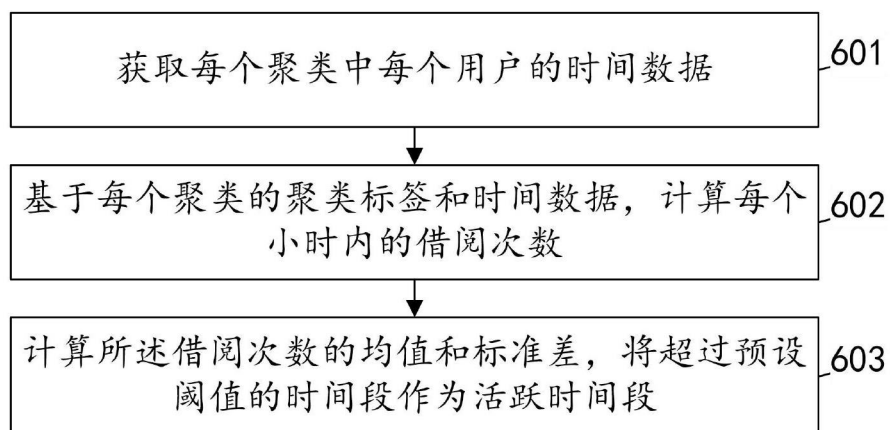


图6

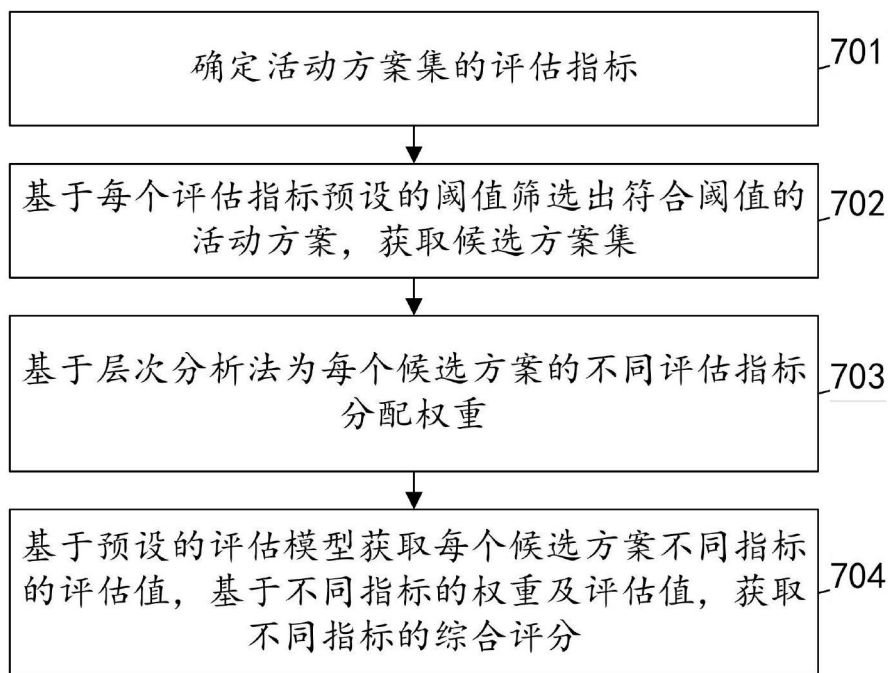


图7

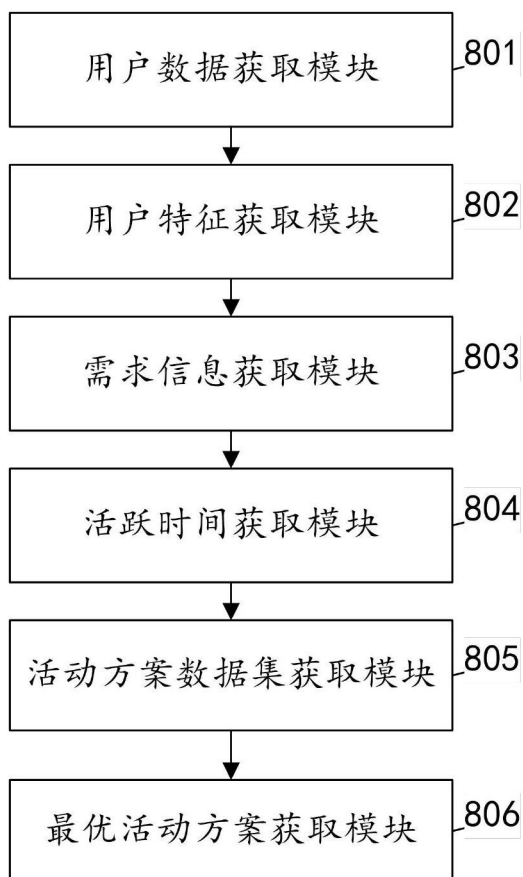


图8

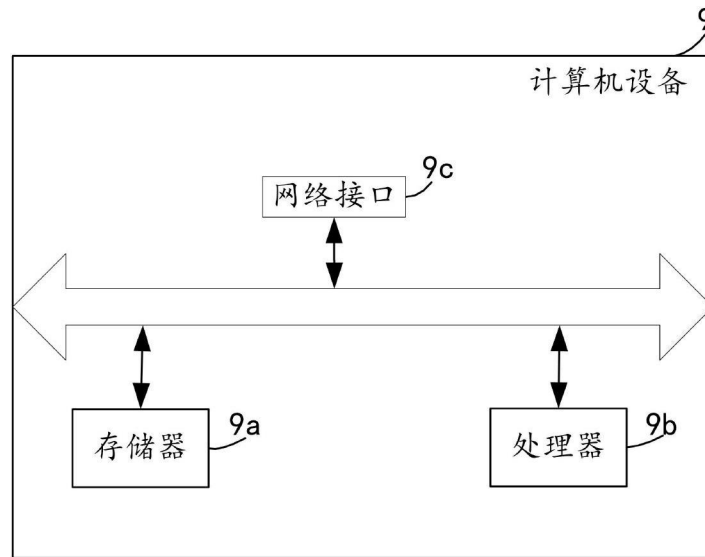


图9