



项目编号:

2018006

CALIS 全国农学文献信息中心研究项目 结题报告

项目名称: 基于数字图书馆知识发现系统中的用户画像研究

项目关键词: 用户画像 图书馆 知识发现系统

项目单位(盖章): 北京农学院图书馆

通信地址: (详细地址含邮编) 北京市昌平区回龙观镇北农路7号

项目主持人: 宁璐

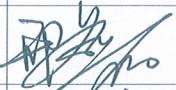
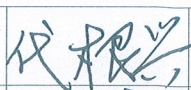
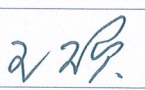
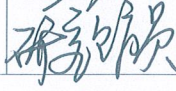
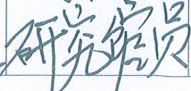

联系电话: (8610) 80795674-201

电子邮件: 13701168920@139.com

提交日期: 2019年05月10日

项目结题验收单

1 专家验收表（主持人所在单位组织 3-5 名专家对项目进行验收、自评。）

项目名称	基于数字图书馆知识发现系统中的用户画像研究					
主持人	宁璐	职务/职称	系统部主任/副研究馆员			
所在单位	北京农学院图书馆（加盖公章）					
专 家 意 见	<p>随着大数据技术逐渐被人们所熟知以及各种各样的用户需求，图书馆改变和提升服务模式势在必行。以前图书馆工作人员要想知道用户的需求，通常以调研方式为主，但是现在随着互联网技术的发展，用户浏览资源的行为及相应数据越发全面并且更易获取，通过对用户属性数据和行为数据的分析，能够更加准确地了解用户的阅读需求，进而实现有目的有针对性的服务。同时，随着图书馆知识发现系统中的数字资源日益增加，这样一来使用户很难在海量资源中快速找到自己感兴趣的资源，图书馆也无法准确、及时的将资源推送给有需求的用户，用户画像作为大数据时代实现精准服务的应用方法之一，为图书馆知识发现系统中的数字资源推荐提供了新的思路。</p> <p>本项目通过构建数字图书馆知识发现系统用户画像，以达到为用户提供精准的资源推荐服务的目的；通过借鉴用户画像在其他等领域的构建方法，以及数据挖掘中分类算法的应用，结合图书馆知识发现系统的特点，运用朴素贝叶斯分类器和自己研究出的算法依据用户搜索行为对其进行用户画像中行为标签的提取，对图书馆用户进行画像的构建；对基于用户画像的数字图书馆知识发现系统主要从系统需求和系统功能模块两方面进行分析，从而提出该系统的设计思路，并使用 Axure 对系统进行原型设计。</p> <p>本项目通过借鉴用户画像在计算机领域和电子商务领域的构建方法，并将其引入到数字图书馆知识发现系统中，并构建用户画像以改变图书馆资源服务模式，实现更加精准的资源服务为高校图书馆拓展信息服务途径、提高信息服务水平提供了借鉴和参考。</p> <p style="text-align: right;">（如需要可增加页数）</p>					
专家签字		陈进		代根兴		王瑛
职务/职称		研究馆员		研究馆员		副研究馆员

题目：基于数字图书馆知识发现系统中的用户画像研究

关键词：用户画像，图书馆，知识发现系统

1 研究背景、目的及意义

在图书馆各式各样的业务中，用户画像在大多数情况下是基于大数据环境生成的，通过数据挖掘、机器学习以及人工智能算法，对图书馆的每一位用户都进行了描绘，包括对图书馆用户的基本属性、阅读属性等的分析，通过挖掘并分析图书馆用户检索过程中生成的大量行为数据，抽象出与该用户的需求和偏好有关的标签化过程^[1]。为用户创建画像的过程就是为用户添加标签的过程，由于用户使用图书馆知识发现系统产生的是一系列数字化操作和行为，因此有必要通过数据挖掘和分析的方法高度细化用户产生的行为信息，以获取与用户相关的特征标识，最终将其输出为用户标签。通过数据挖掘和分析用户多个方面产生的多维标签可以形成用户的精准画像。一般而言，用户画像是用来分析用户的基本属性和基本行为的，因此构建用户画像能够精准知道用户的兴趣和需求，可以用于个性化推荐服务、营销服务和各种用户推广服务。目前，基于用户画像的数字图书馆研究主要包括通过对用户画像建模和分析，进而构建出图书馆服务能力的综合评价指标，并通过该指标为图书馆社区用户提供有效的参考。此外，还可以通过构建用户画像形成可视化的统计描述、多维交叉分析以及用户知识图谱等图书馆智能化服务。到目前来讲，基于数字图书馆知识发现系统中的用户画像处于起步阶段，还需要更多的研究和探索^[2]。

随着大数据技术逐渐被人们所熟知以及各种各样的用户需求，图书馆改变和提升服务模式势在必行。在转型发展过程中，图书馆开始重视对用户使用资源产生的大量数据进行分析，进而了解用户的需求并且利用人工智能技术来创新服务模式，以满足不同用户的需求从而提升图书馆的服务价值。以前图书馆工作人员要想知道用户的需求，通常以调研方式为主，但是现在随着互联网技术的发展，用户浏览资源的行为及相应数据越发全面并且更易获取，通过对用户属性数据和行为数据的分析，能够更加准确地了解用户的阅读需求，进而实现有目的有针对性的服务。同时，随着图书馆知识发现系统中的数字资源日益增加，这样一来使用户很难在海量资源中快速找到自己感兴趣的资源，图书馆也无法准确、及时的将资源推送给有需求的用户，用户画像作为大数据时代实现精准服务的应用方法之一，为图书馆知识发现系统中的数字资源推荐提供了新的思路。当前，用户画像已经成为一个热门话题，在计算机领域和电子商务领域已经被广泛应用。通过深度挖掘与图书馆相关的门户平台、系统等交互产生的数据，构建数字图书馆知识发现系统用户画像，以此掌握用户的资源偏好，并提供精准的资源推荐服务，这是提高图书馆资源利用率的重要途径。本文通过借鉴用户画像在计算机领域和电子商务领域的构建方法，将其引入到数字图书馆知识发现系统中，并构建用户画像以改变图书馆资源服务模式，实现更加精准的资源服务^[3]。

2 研究内容及方法（思路、方法、具体内容）

一、研究思路

1. 分析处理调研数据

此部分主要研究用户画像的数据来源、用户画像的数据采集以及用户画像的信息识别。利用统计分析方法初步给用户分成类,然后结合现有数字图书馆知识发现系统中的服务内容,分析出用户属性数据,由检索数据、阅览数据等分析出用户行为数据。并对调研到的数据进行预处理,去除与构建用户画像相关性不大的数据,利用统计分析软件对图书馆服务现状进行统计分析,并研究数据之间的关系。

2. 构建用户画像

此部分基于对数据之间的关系的分析与研究,确定构建数字图书馆知识发现中用户画像的算法,此算法要做到基于用户的搜索行为对其进行分类,以7天为一周期,找到其在周期内所最感兴趣的领域,确定行为维度的标签,在和属性维度合并一起,共同构建图书馆用户画像。

3. 系统设计思路及原型设计

基于用户画像的构建过程以及各大高等院校用户对图书馆知识发现系统的实际需求提出设计思路,并利用Axure实现原型设计。

二、研究方法

本项目研究拟采用文献分析、统计分析、数据建模等研究方法和手段。

1. 文献分析

文献分析法是依据研究对象,通过查阅文献资料,较全面地了解研究问题的相关内容,以此来对所研究的问题有进一步深入的了解,并揭示其规律或属性的一种方法。本项目研究将通过查阅大量硕士论文、期刊及网络信息等,了解国内外对用户画像的相关研究,尤其是在数字图书馆知识发现系统中的应用研究,为本项目研究提供理论依据。

2. 统计分析

借助统计分析软件对调研来的用户数据进行分析与研究,了解各大高校图书馆服务现状以及用户对图书馆资源使用的情况。为构建图书馆知识发现系统中的用户画像提供数据基础。

3. 数据建模

通过对数字图书馆知识发现系统的用户数据进行收集、挖掘和分析清洗,对图书馆用户的基础数据进行收集,并进行用户行为建模,提出构建用户画像的算法。然后再收集一组测试数据集对算法进行验证。

三、具体研究内容

1. 现有构建用户画像的方法

(1) 基于用户行为偏好的画像方法

构建用户画像的根本就是对其使用行为进行充分剖析。国外学者在用户网络行为分析方面,深度挖掘了网站站点的网络日志,并通过对用户行为聚类来描绘不同的用户群体。以这一研究作为背景,一些国外其他学者根据用户实时变化的动态,对其进行了动态画像的构建,实现了对用户的实时追踪^[13]。国内学者通过对用户画像的不断研究,最终发现用户在互联网中留下的“痕迹”可以作为对其构建画像的基础。基于用户行为偏好的画像方法能够提升画像的准确性。图书馆知识发现系统构建用户画像可以基于用户偏好,通过收集用户在使用系统时产生的历史数据,对其进行“绘制脸谱”,进而为其提供合适的数字资源,实现图书馆知识发现系统服务模式的转变,使其利用率大大提升^[14]。

(2) 基于主题模型的画像方法

基于主题的用户画像构建方法, 主要通过对用户在图书馆知识发现系统中的历史记录, 从用户感兴趣的资源信息中提取出不太明显的主题特征, 从而以概率分布的形式描述用户的特征。在得到每一特征的概率分布情况之后, 将其所对应的属性进行量化。在构建用户画像中经常使用的主题模型包括 LDA (Latent Dirichlet Allocation)、ATM (Author-Topic Model)、ARTM (Author-Recipient-Topic Model) 和 APTM (Author Interest Topic Model) 模型^[15]。

(3) 基于用户兴趣偏好的画像方法

用户兴趣偏好在图书馆知识发现系统中是指用户的检索偏好, 此画像方法是目前备受关注的热点算法之一。一些国外的研究学者利用数学模型将用户的兴趣记录下来, 并以此为基础对其建立了个人兴趣档案, 用来体现出用户的兴趣爱好^[16]。研究学者通过研究发现隐含狄利克雷分布模型可以作为基于用户兴趣偏好的画像方法, 使用该模型需要注意的是无论是用户信息还是用户的样本量都要保证足够大, 如果样本量或单个样本中的数据量其不够充足时, 用户画像的精度就不能够满足要求。

2. 图书馆知识发现系统应用用户画像的可行性

(1) 社会可行性

通过调研全国部分高校图书馆知识发现系统提供服务的现状, 对数据进行统计分析后, 我们发现超过一半的图书馆老师对图书馆知识发现系统应用用户画像听说过甚至有一定的了解, 但是已经将用户画像应用到知识发现系统中的仅有不到 10%。在被调研的对象中, 无论是已经应用用户画像的还是目前没有应用的, 超过 90% 的被调研的老师认为将用户画像应用到图书馆知识发现系统中是十分有必要的, 其中有大约 20% 表示在未来愿意尝试将用户画像应用到本校的图书馆中。部分高校图书馆应用用户画像情况如图 1 所示。通过调研还发现大部分高校图书馆老师了解用户画像都是自己主动去了解或者在同行交流会上获取信息, 还有少部分在相关公司的宣讲会上得知, 由此可见, 将用户画像应用到图书馆知识发现系统中已经成为趋势, 具体分析情况如图 2 所示。

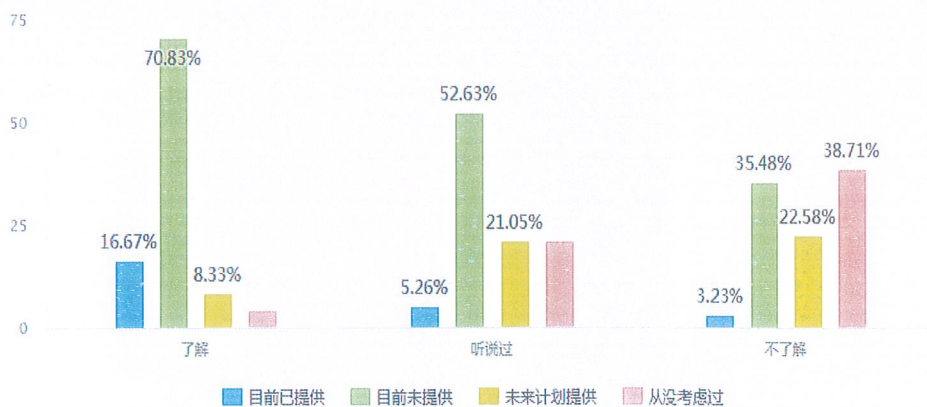


图 1 部分高校图书馆应用用户画像的情况

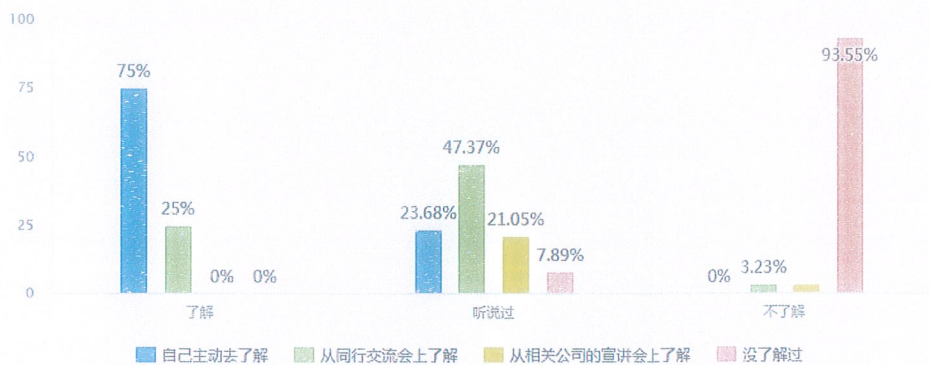


图2 高校图书馆了解用户画像的途径

(2) 技术可行性

到目前为止,我国的图书馆知识发现系统仍然处于快速发展阶段,提供的服务还不够成熟。图书馆不仅要做到能够为用户提供充足的资源,还需要将这些资源能够被充分利用,这就需要对图书馆知识发现系统构建用户画像,实现对用户精准服务。

在国内,电子商务领域应用用户画像较为广泛,图书馆知识发现系统可以从中借鉴成功经验,为己所用。虽然目前国内图书馆知识发现系统中尚未广泛应用该功能,但是天津图书馆、武汉大学图书馆社区知识用户构建的案例,证明了将用户画像应用于图书馆服务是可行的^[17]。

3. 用户画像模型构建步骤

(1) 数据来源

对用户使用图书馆知识发现系统产生的数据进行收集是构建用户画像的基础。用户数据在一般情况下主要分为两种:用户属性数据和用户行为数据,也就是所说的静态数据和动态数据,如图3所示。用户属性数据,也被称为静态数据,主要是指用户的学号、姓名、性别、年龄、院系、专业、教育程度等基本信息,这些数据可以通过统一身份认证获取,如图4所示。用户行为数据是基于用户检索行为的信息,也称为动态数据。这些数据一般包括使用图书馆知识发现系统时检索、浏览、收藏、下载、订阅、分享、访问终端等的的数据,收集用户行为数据的目的是最大程度地恢复用户的信息和行为,如图5所示。

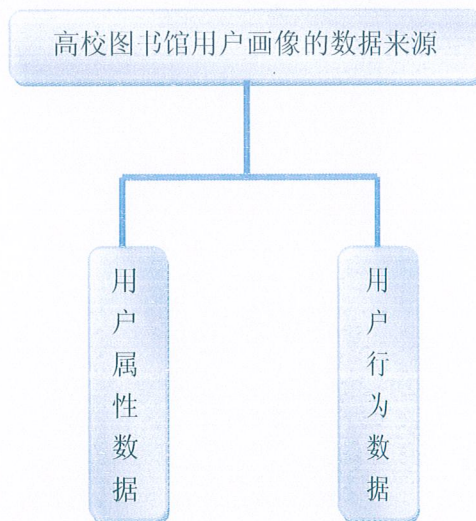


图3 用户数据来源图

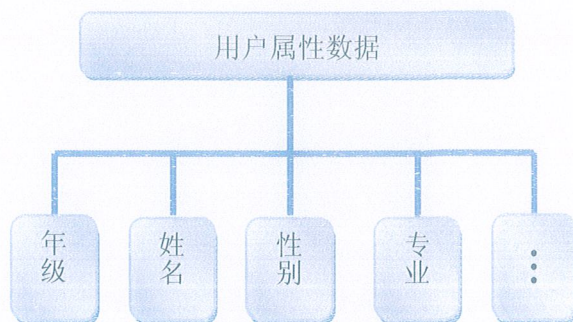


图4 用户属性数据

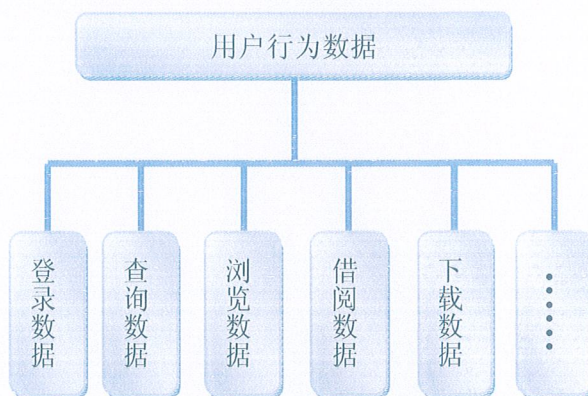


图5 用户行为数据

(2) 数据预处理

用户在使用图书馆知识发现系统时，每一次的检索都是一件不确定事件，因此就会产生大量的有用的或没用的数据，在为其进行画像的时候，并不是与该用户有关的所有数据都需要分析，只需要选取相关性较强的数据对其进行分析即可，避免研究没用的数据降低分析效率。

从数字图书馆知识发现系统的角度来看，与模型的构建相关性较强的信息有：用户的专业、受教育程度、用户使用的资源、资源使用频次、资源使用时间等等，而除此以外的数据信息与构建模型关系不大，可以忽略不计^[18]。在用户画像数据中，需要通过数据挖掘技术分类、聚类、关联分析等对用户的属性数据和行为数据进行预处理，进而获取到对构建用户画像有用的数据信息。

(3) 用户分类算法

在机器学习和统计中，用户分类是根据已知类别所具有的特征来确定待分类用户的所属类别^[33]。实现分类的算法在具体实现中也被称为分类器。目前用来实现用户分类的算法多种多样，例如决策树、随机森林、K邻近值、神经网络、朴素贝叶斯算法等。通过查阅大量文献以及结合采集到的数据的特点，本项目研究发现朴素贝叶斯算法所需参数较少且算法效率相对稳定，与其他算法相比，较适合应用于对图书馆知识发现系统中的用户进行分类。因此，最终本项目研究确定使用该算法基于用户搜索行为对其进行初步分类，在原有用户分类算法的基础上，研究设计出“时间-频次”面积算法来确定最终的用户画像中的动态标签。

(4) 数据标签

用户画像数据标签是一种在用户相关抽象数字统计分析后更具指导性，更生动，更易于理解的信息，也就是说用户画像数据标签是从用户的海量数据中提取关键信息。构建用户画像标签就是对用户信息的高度精炼概括，最后将所有标签综合起来，就可以勾勒出该用户的“画像”。

构建图书馆知识发现系统中用户画像的数据标签可以从两个角度去考虑，其一是用户的一些基本属性信息标签，其二就是用户的行为信息标签。用户属性标签是用来描述个人的基本信息，比如姓名、性别、专业、学历等。行为信息标签是用来对用户的检索行为进行高度概括，用来描述用户的行为特征^[19]。

(5) 用户画像模型

将收集到的动态数据和静态数据在分析和处理后进行标记，形成分类和多级用户画像标签结构。在此基础上，可以“可视化”用户画像模型。用户画像与传统意义上的数据统计大有不同。数据统计就是简单的对用户行为进行求和，并不能够发现数据间存在的关系，而用户画像则是对用户数据进行分析，进而发现数据间存在的细微关系，能够将用户的需求直观地描述出来。因此，需要构建用户画像模型，来对每一位用户进行深度剖析。用户画像模型主要包括三层：分别是数据来源层、数据采集层以及数据挖掘层，如图6所示。

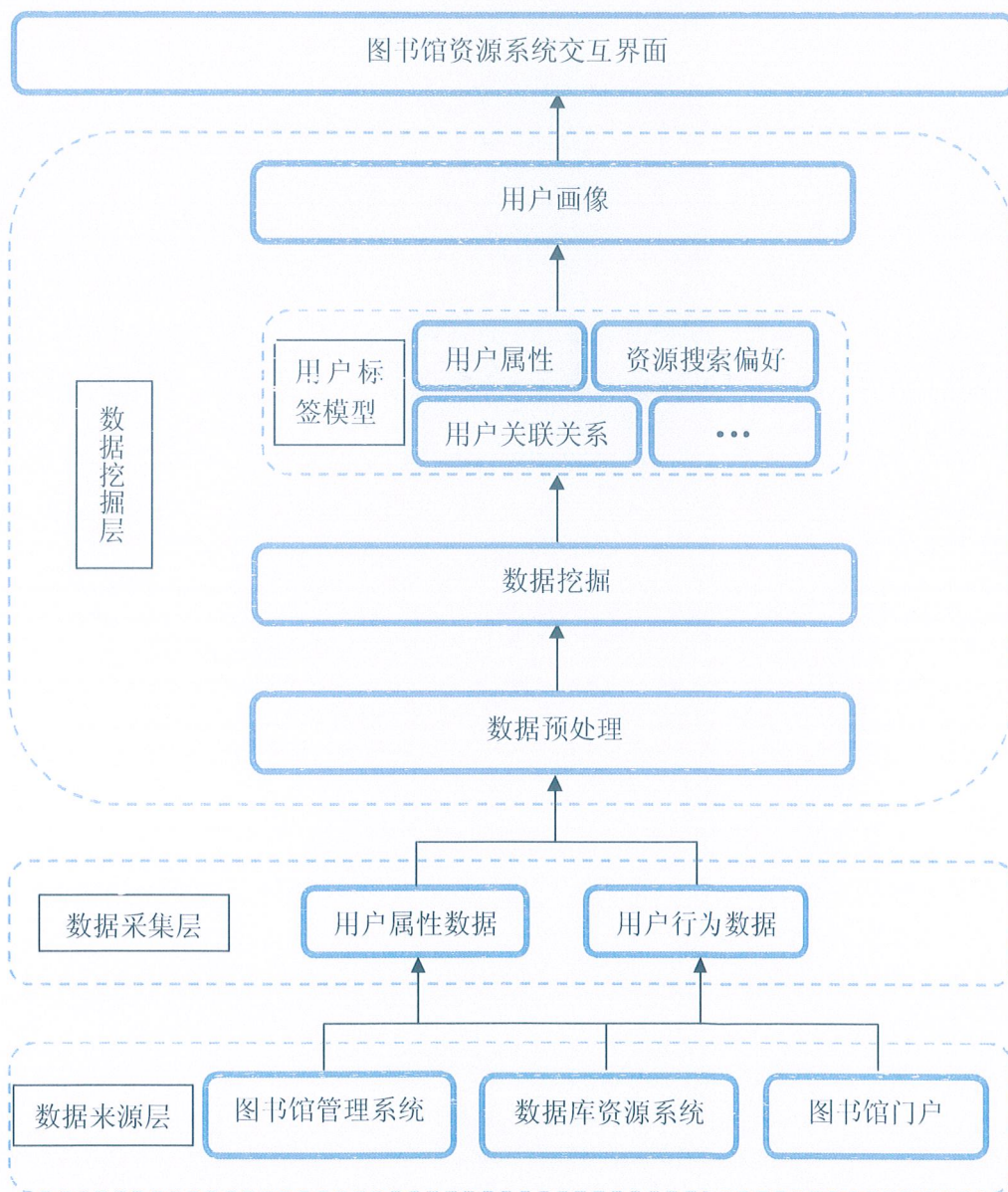


图6 用户画像模型

数据来源层是构建用户画像模型的根本，是整个模型的基础，是用户属性数据和行为数据

的有效整合；数据采集层需要对有效数据进行采集和排序，然后将其存储在数据库中；数据挖掘层是整个模型的重要所在，需要对采集到的数据库进行数据清洗、数据集成、数据分类等预处理，构建用户数据标签模型，再利用数据挖掘技术对用户进行分类和分析，从而形成读者群体用户画像，最后，依据用户的行为和需求，有针对性的对用户进行资源推送。

4. 系统设计思路及原型设计

(1) 系统设计思路

基于用户画像的图书馆知识发现系统要想实现为用户提供其所在研究领域搜索排名靠前的信息资源需要实现以下几个功能：提供按标题、作者、关键词、摘要和期刊名称等查询条件搜索相关数字资源，并对检索结果按搜索次数进行排序处理；记录用户登录、检索、查阅、下载等记录，为推荐服务提供数据来源；对数据库中用户基本属性信息、用户使用系统搜索资源的行为信息等统计，筛选搜索热门关键词、图书、期刊等信息；挖掘数字图书馆知识发现系统数据库中的用户画像，并利用构建的群体用户画像将用户进行分类，然后再针对不同类别的用户进行热门搜索和最新推荐。基于用户画像的图书馆资源推荐系统共分为五个模块，包括用户管理模块、资源搜索模块、资源统计模块、用户画像模块、相关推荐模块，该系统的系统层次结构如图7所示。

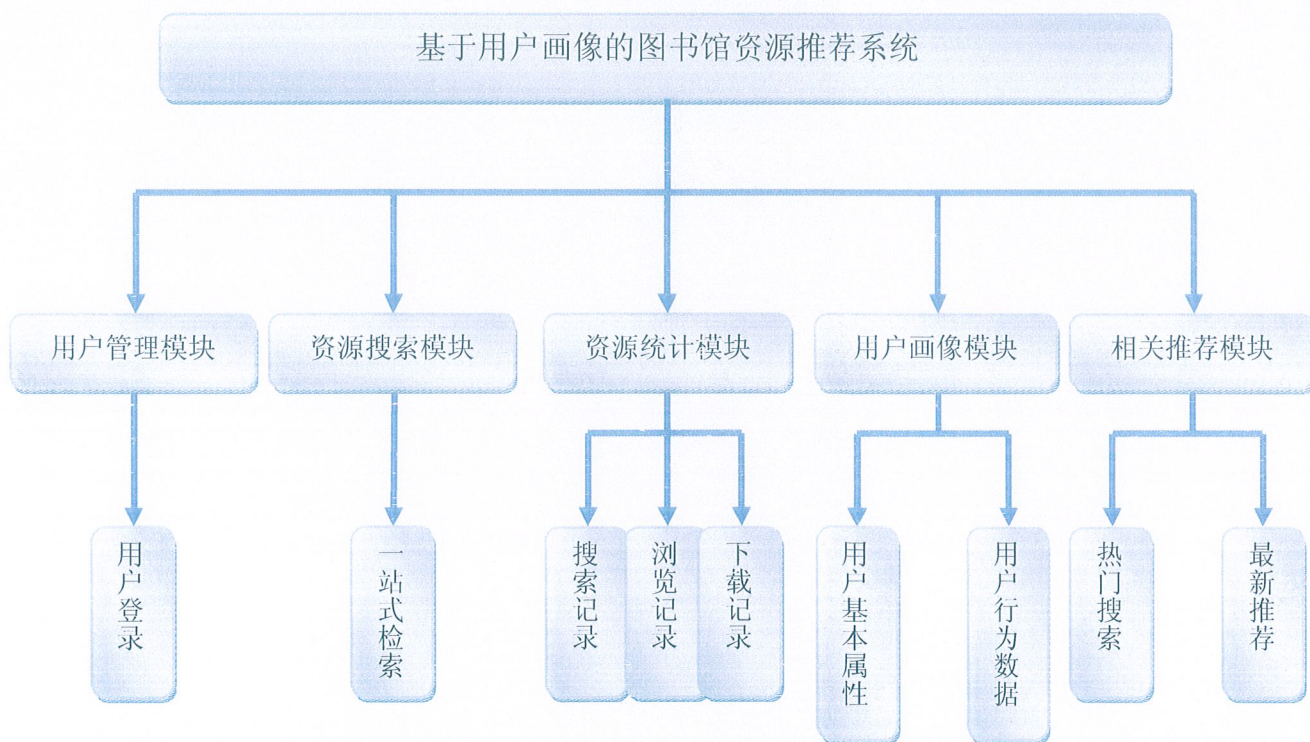


图7 基于用户画像的图书馆资源推荐系统的系统层次结构

(2) 原型设计

当用户成功登录后，就会进入首页界面，此页面不仅为用户提供搜索功能，还为读者用户提供了用户画像，以及针对用户画像所进行的相关推荐和热门推荐。系统首页如图 8 所示。

基于用户画像的数字图书馆知识发现系统

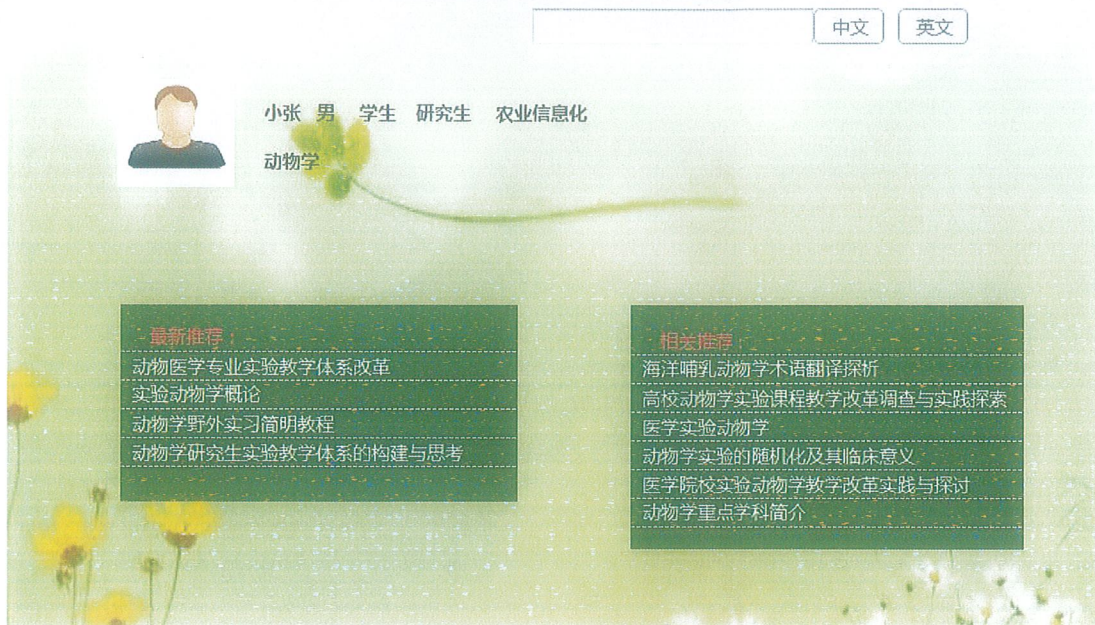


图 8 系统首页

四、创新点

在图书馆知识发现系统中，大数据和个性化图书馆资源推送一直是该领域的研究热点，越来越多的研究学者开始投身到对图书馆的大数据研究中去，尤其是对图书馆知识发现系统中产生的数据的研究。虽然图书馆知识发现系统基于大数据的研究成果在逐年增加，但是对如何将大数据利用好的研究还处于发展阶段。最近几年，用户画像技术的应用成为了大数据应用研究中的新的研究热点，并且得到了十分广泛的关注，但是将“图书馆”和“用户画像”一起作为关键词在中国知网中进行搜索，可查阅到的文献屈指可数，只有数十篇，也仅仅只是对图书馆应用用户画像的可行性和意义做了研究，很少甚至没有学者提出具体的算法去构建图书馆知识发现系统中的用户画像，说明该研究内容在图书馆知识发现系统中处于空白阶段，急需补充。

本项目以图书馆知识发现系统中的用户作为研究对象，运用相关技术分析图书馆用户搜索资源时的特征因素，建立用户画像。重点研究了运用朴素贝叶斯分类算法对用户进行分类，设计了“时间-频次”面积算法，通过该算法确定了用户行为属性标签，在图书馆知识发现系统中建立用户画像模块。本项目的研究，较传统图书馆系统具有一定的创新性。希望研究结果能够揭示一定的规律，并对图书馆工作人员有借鉴意义，也期望该研究能为后续研究提供帮助。

3 结论与建议

一、结论

本项目基于现有的构建用户画像的方法,在原有用户分类算法的基础上,研究设计出“时间-频次”面积算法,并结合图书馆知识发现系统的特点,提出了构建图书馆用户画像的算法。主要研究内容如下:

(1)分别面向高校图书馆和高校师生、教职工发放调查问卷,一方面了解现在部分高校图书馆提供用户画像的现状以及高校图书馆对提供用户画像所持有的态度。另一方面用来收集用户的属性信息和行为信息。通过对问卷结果进行交叉分析,明确了图书馆知识发现系统服务现状和图书馆工作人员对提供用户画像所持有的态度;通过对问卷的统计分析,对数据进行预处理,提取出不同领域的特征词。

(2)利用朴素贝叶斯算法,依据用户每一次检索的特征词将其初步分类,并记录下分类结果以及检索的时间和频次。再利用“时间-频次”面积算法,通过比较,最终确定用户的行为属性标签。最后将行为属性标签与用户的基本信息属性标签相结合,最终确定了某一用户的画像。

(3)对基于用户画像的数字图书馆知识发现系统进行需求分析、功能分析以及数据库的初步设计,并提出了整个系统的设计思路以及预期实现的功能,运用 Axure 对其进行了原型设计。

二、建议

本项目研究中,只是基于朴素贝叶斯算法和“时间-频次”面积算法对用户进行了标签提取、构建画像,对其如何应用只是提出了思路,并没有进行实际的应用。未来对其研究应用应有以下几个方面:

(1)由于本项目研究的重点是构建图书馆知识发现系统中用户画像的算法,因此在数据预处理时采用的是人工分词,可能存在误差,希望在以后的研究中,可以使用一些分词技术利用计算机来完成这项工作,不仅可以节约成本,还能提高准确率。

(2)在本项目研究中,虽然为图书馆知识发现系统中的用户构建了画像,但是由于用户画像模型相对简单,因此构建出的画像正确率较高,但是准确率需要进一步提高。因此在今后的研究中需要多加考虑一些其他因素,不断优化算法,进而完善用户画像。

(3)本项目研究只是基于用户基本属性信息和用户的行为对其进行了用户画像算法的研究,但作为一个完整的系统,并未将用户画像与推荐算法相结合实现为用户推荐资源,增加兴趣相似用户群体的资源共享。

(4)在用户画像应用上,只对应用于资源推荐进行了相关研究,在今后的研究中,还可以对用户画像在图书馆服务的其他应用方面开展更广泛的研究,为用户提供更好的服务。

4 项目成果(发表的文章、开发的软件、取得的实践效果等)

(1)指导研究生完成硕士毕业论文《用户画像技术在数字图书馆知识发现系统中的应用研究》。

(2)发表论文一篇。《Application and Research of User Portrait in Digital Library》(Computer Science and Engineering (ISSN: 2475-8841))。

5 参考文献

[1] 邱均平,张聪.高校图书馆馆藏资源协同推荐系统研究[J].图书情报工作,2013,(22):

132 - 137.

- [2] 徐立宁. 基于动态精准画像的图书馆个性化推荐服务研究[J]. 图书馆学报, 2018(10).
- [3] 王庆, 赵发珍. 基于“用户画像”的图书馆资源推荐模式设计与分析[J]. 现代情报, 2018(3): 105-109, 137.
- [4] 张小可, 等. 贝叶斯网络在用户画像构建中的研究[J]. 移动通信, 2016(22): 22-26.
- [5] 李雅. 基于读者用户画像的高校图书馆精准化服务研究[J]. 农业图书情报学报, 2018(12).
- [6] Billsus D, Pazzani MJ. A Hybrid User Model for News Story Classification[M]. Sprin Vienna: UM99 User Modeling, 1999: 99-108.
- [7] Henczel S. Creating User Profiles to Improve Information Quality[J]. Online, 2004(3): 30-33.
- [8] Rumpler B. A Study of the impact of the user profile in documentary systems[J]. Online information review, 2001(6): 359-365.
- [9] OTTERLO M. Project BLIIPS: making the physical public library more intelligent through artificial intelligence[J]. Qualitative and quantitative methods in libraries, 2017, 5(2): 287-300.
- [10] 曾鸿, 吴苏倪. 基于微博的大数据用户画像与精准营销[J]. 2016(16): 306-308.
- [11] 张丽娟. 基于大数据分析的用户画像助力精准营销研究[J]. 电信技术, 2017(1): 61-62, 65.
- [12] 刘速. 浅议数字图书馆知识发现系统中的用户画像——以天津图书馆为例[J]. 图书馆理论与实践, 2017(6): 103-106.
- [13] ADOMAVICIUS G, TUZHILIN A. Using data mining methods to build customer profiles [J]. Computer, 2001, 34(2): 74-82.
- [14] 杨帆. 以图像分析为基础的图书馆大数据[J]. 图书馆, 2018(04).
- [15] Wei X, Croft W B. LDA-based document models for ad-hoc retrieval[J]. 2006: 178-185.
- [16] PAZZANIMJ, BILLSUSD. Content-based recommendation system[M]. Springer, 2007: 325-341.
- [17] 韦良珍. 大数据环境下用户画像在图书馆的应用[J]. 中华医学图书情报杂志, 2018, 27(3): 33-37.
- [18] 杨燕. 图书馆用户画像及其应用研究[J]. 四川图书馆学报, 2018(1).
- [19] 潘宇光. 高校智慧图书馆读者信息需求的用户画像[J]. 合肥工业大学学报(社会科学版), 2018(4).
- [20] 任力安, 何清, 史忠植. 分类超曲面方法在海量数据分类中的应用[J]. 计算机科学, 2002, 29(9): 33-35.
- [21] 威滕. 数据挖掘实用机器学习技术[M]. 北京: 机械工业出版社, 2006.
- [22] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. Psychol Rev, 1958, 65(6): 386.
- [23] 唐发明. 基于统计学习理论的支持向量机算法研究[D]. 武汉: 华中科技大学, 2005.
- [24] 何清, 李宁, 罗文娟, 等. 大数据下的机器学习算法综述[A]. 中国计算机学会人工智能会议[C]. 2013.
- [25] 周志华. 机器学习[J]. 中国民商, 2016, (3).

- [26] 陈慧香, 邵波. 国外图书馆领域用户画像的研究现状及启示[J]. 图书馆学研究, 2017 (20) .
- [27] 陈沛玲. 决策树分类算法优化研究[D]. 中南大学, 2007.
- [28] Amato G, Straccia U. User Profile Modeling and Applications to Digital Libraries[C]. Springer - Verlag: European Conference on Research and Advanced Technology for Digital Libraries, 1999: 184-197.
- [29] LEWIS C, CONTRINO J. Making the invisible visible: personas and mental models of distance education library users[J]. Journal of library & information services in distance learning, 2016, 10 (1-2) :15-29.
- [30] 黄文彬, 徐山川, 吴家辉, 等. 移动用户画像构建研究[J]. 现代情报, 2016, 36 (10) :54 - 61.
- [31] 王庆福. 贝叶斯网络在用户兴趣模型构建中的研究[J]. 无线互联科技, 2016 (12) : 101 - 102.
- [32] 光峰, 姚程宽, 卢灿举, 等. 数据挖掘经典算法研究[J]. 商丘师范学院学报, 2016, 32 (3) .
- [33] 于淼. 我国当代大学生网络游戏消费动机研究[J]. 吉林大学, 2008.4
- [34] 杨伟光. 面向大数据分析的决策树算法研究[J]. 电子技术与软件工程, 2018, 12.
- [35] 皮亚宸. K 近邻分类算法的应用研究[J]. 通讯世界, 2019, 1.
- [36] 郭勋诚. 朴素贝叶斯分类算法应用研究[J]. 通讯世界, 2019, 1.
- [37] 滕广青, 崔林蔚. 大学图书馆网络用户柔性化细分[J]. 图书馆学研究, 2017, 1.
- [38] 刘海鸥, 孙晶晶, 苏妍嫫, 张亚明. 国内外用户画像研究综述[J]. 情报理论与实践, 2018, 8.