

在线学习资源个性化推荐服务模型的构建

张小雪, 张立国

陕西师范大学教育学院, 西安 710062

【摘要】 该研究通过梳理国内外个性化学习资源建设现状, 提出利用学习者自我评价和 Felder-Silverman 量表进行前测以检测学习者的学习风格, 利用学习者在学习过程中的学习行为和学习路径等分析出隐性学习者特征, 采用协同过滤技术和数据挖掘技术相结合的方法将学习者特征与在线学习资源相匹配, 建构基于协同过滤和数据挖掘的在线学习资源个性化推荐服务模型并进行案例实践, 以期为在线学习者特征的维度确立以及在线学习资源个性化建设等研究提供参考。

【关键词】 在线学习资源; 个性化推荐; 协同过滤; 数据挖掘技术

【中图分类号】 G40-057 **【文献标志码】** A **【文章编号】** 1004-5287(2017)02-0172-04

【DOI】 10.13566/j.cnki.cmet.cn61-4317/g4.201702015

Construction of personalized recommendation service model of online learning resources

Zhang Xiaoxue, Zhang Ligu

School of Education, Shaanxi Normal University, Xi'an 710062, China

【Abstract】 By understanding the status of personalized learning resources construction at home and abroad, the study presents the following ideas: using the learner's self-evaluation and pre-test method of Felder-Silverman scale to test learner's learning styles; using learners' learning behavior in learning process to construct learners' hidden characteristics, adopting the combination of collaborative filtering technology and data mining technology to match learners' characteristics with online learning resources, and constructing personalized recommendation service of online learning resources mainly based on collaborative filtering and data mining and carrying out case practice. We hope to provide reference for studies on establishment of online learner's characteristics dimension and construction of personalized recommendation service of online learning resources.

【Key words】 online learning resources; personalized recommendation; collaborative filtering; data mining technology

近年来,在线学习发展迅速,学习者对学习资源的个性化和实用性的要求越来越高。无论是传统教育还是在线教育,学习资源的质量对教学效果都起着决定

性的作用,在在线教育中更是如此。因此,在线教育研究者开始关注在线教育资源的建设,并利用大数据和学习分析技术为学习者提供适合的学习资源推送

- 林杰,邵阳,游箭. 如何提高 CT 介入放射学的教学效果[J]. 川北医学院学报, 2008, 23(3): 327-328
- 徐秋贞,储成凤,杨小庆. 浅谈医学影像专业的 CT、MRI 教学实践[J]. 现代医用影像学, 2006, 15(4): 163
- 王卫国. 虚拟仿真实验教学中心建设思考与建议[J]. 实验室研究与探索, 2013, 32(12): 5-8
- 赵鹤亮,洪波,赵新斌,等. CT、MRI 仿真操作系统在实验教学中的应用[J]. 河北联合大学学报: 医学版, 2014(1): 122-123
- 刘磊,周作栋,彭永昱. 安卓手机的前景与现状[J]. 成功: 教育版, 2013(1): 295

- 焦健. Eclipse 下 Android 环境的搭建[J]. 信息与电脑: 理论版, 2012(6): 33-34
- 周倩,王军,王亮亮. 基于串口转 wifi 模块的单片机与安卓手机通讯[J]. 福建电脑, 2016, 32(3): 120-121
- 张群,杨絮,张正言,等. 蓝牙模块串口通信的设计与实现[J]. 实验室研究与探索, 2012, 31(3): 79-82
- 牛俊荣. 关于摆位向体位设计转化的思考[J]. 现代医用影像学, 2002, 11(6): 254
- 安晓莉. AT89S51 单片机并行 I/O 端口的扩展[J]. 电子设计工程, 2009, 17(8): 66-67

收稿日期: 2016-09-22

作者简介: 张小雪(1994-),女,新疆伊犁人,硕士研究生在读,主要研究方向:教育技术基本理论,网络与远程教育。

通信作者: 张立国(1965-),男,陕西榆林人,教授,博士,主要研究方向:教育技术基本理论,网络与远程教育。

电话: 18629426083; E-mail: zhangliguo@126.com

服务,进而更好地满足在线学习者的学习需求,以实现人性化、个性化的学习。在梳理文献时发现,目前国内关于资源个性化推送方式的研究主要有两种方式:一种是通过在线学习过程中所产生的数据进行分析来实现资源的个性化推送,这种方式没有充分考虑到学习者原有的认知结构和学习风格,推送结果不尽如人意;另一种是利用学习者前测与学习过程的动态数据分析相结合的方法来实现资源的个性化推送,这种方式增加了学生自评,但是没有运用学习风格量表对学习风格进行进一步确定,使学习资源推送的准确度和适应性不够^[1]。

鉴于在线学习者的特殊性,其学习起点、学习风格各有差异,根据对在线学习者学习前自我评价的文本信息挖掘、基于学习风格量表的前测等隐性学习特征分析和学习过程挖掘学习行为的显性特征分析相结合的方式来确定学习者特征,利用协同过滤技术和数据挖掘技术完成学习资源特征和学习者特征相匹配,提高学习资源个性化推送的精度和效率,满足在线学习者个性化学习的需求。

1 理论基础

1.1 学习风格模型研究

要实现个性化资源推送,确定学习者的学习风格是其关键步骤。目前,主流的研究学习风格的模型有: Myers-Briggs 的人格类型模型的学习风格理论、Kolb 的学习风格模型以及 Felder-Silverman 学习风格模型,等。该研究选择 FLSM(felder & silverman learning style model) 并采用其学习风格测量表 ILS(index of learning styles),主要是基于其 ILS 的信度和效度很高,相比于其他学习风格模型,该学习风格模型划分的维度更为详细和完整,使用也较为普遍。该模型将学习风格分为信息加工、感知、输入、理解四个维度。其中,对知识的加工分为活跃型和沉思型;知识的感知分为感悟型和直觉型;知识的输入分为视觉型和语言型;对知识的理解分为序列型和综合型等 8 种类型^[2]。

该研究正是采用此学习风格模型,可以大致检测出其学习风格偏好弱、适中还是强烈。以信息加工维度为例,通过量表测量可以看出学习者的信息加工方式是偏向活跃还是沉思型。学习者登录用户后首先进行学习前自我评价和学习风格量表的测试,根据测量结果进行编码、挖掘与分析,初步建立学习者显性学习者特征模型并形成个人信息库。

1.2 协同过滤推荐技术

协同过滤推荐(collaborative filtering recommendation)技术是目前计算机领域使用较为广泛和成熟的

一种个性化推荐技术。协同过滤推荐通过分析用户特性和兴趣,等,利用协同推荐算法计算用户之间和数字资源之间的相似性。例如类似音频、视频等难以区分的学习资源,从中找出相似度最高的 n 个用户或资源。在协同过滤算法中把某种特性相似度最高的用户叫做邻居,每个用户会对资源进行评分,然后参照邻居对所获取资源的评分来预测目标用户对资源的评分,将预测评分最高的 m 个资源推荐给目标用户,使其尽可能与用户的学习风格相匹配^[3]。协同过滤推荐技术因其有效和较为准确的推荐结果,得到了计算机领域个性化推荐方式的广泛应用。该研究将计算机科学领域中的协同过滤技术应用于教育领域中的在线学习系统中,并根据学习者的学习风格、学习兴趣 and 个性化需求向其准确推荐适合其学习风格和学习内容的学习资源,挖掘在线学习中学习者的潜在学习兴趣,也是对传统资源检索方式的补充,改善学习者的学习体验,激起学习者的学习兴趣。

2 基于协同过滤技术的在线学习资源个性化推荐服务模型

2.1 信息化环境下学习者特征模型的构建

学习资源个性化推荐系统为了推送符合学习者学习风格和认知风格的学习资源,首先要确定学习者的学习风格和特征以及学习阶段,等。因此,学习者特征模型的建构是该研究的首要任务,如图 1 所示。

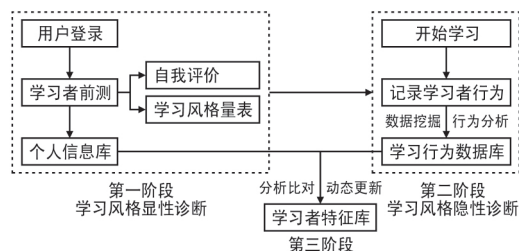


图1 学习者特征模型

关于学习风格的分类与要素分析的研究很多,也较为成熟。在该研究中对学习者学习风格的诊断分为显性诊断和隐性诊断两种方式,李宝等研究者也运用了此种诊断维度。第一阶段显性诊断方式主要是指学习者在进入在线学习前进行学习特征的自我评价和填写学习风格量表,能够得到学习者明显的描述。自我评价主要包括学习者注册时填写的基本信息,包括学习方式、兴趣、喜欢的学习资源类型和学习能力,等,系统会提取文本关键信息,进行编码和分析生成学习者个人信息库。之后在线学习系统采用学习风格测量量表,检测出学习者的学习风格倾向和每种风格的强弱程度,进而推断出学习者属于何种学习风格

类型。

第二阶段隐性诊断方式主要是通过对学习过程中学习者的学习行为和学习路径及学习内容等进行记录、挖掘和分析,等,对收集的动态数据进行分析,对静态的学习风格进行修正,进而确定其在线学习的行为特征,建立学习者行为库。这些记录行为特征信息的数据是和学习者前测数据进行结合和比较分析的,是为学习者在线资源个性化推荐进行动态调整和修正的依据,即第三阶段。在确定学习者的学习风格和学习者特征后,系统开始为学习者推送适合的学习资源,在其学习过程中,由于学习内容或者学习兴趣的改变,系统会随时对学习风格做出适当的调整,保证对其个性化的服务。

2.2 在线学习资源个性化推荐服务模型的构建

学习者学习特征维度确定后,系统已经完成基于某个时间对学习特征特征的把握,然而这只是实现在线资源推送服务的前期基础,要实现在线资源个性化推荐服务还需完成学习资源特征库的构建以及学习者特征与资源特征的相似匹配,如图2所示。

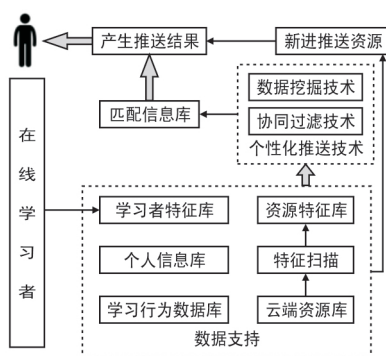


图2 在线学习资源个性化推荐服务模型

2.2.1 数据支持 数据支持模块是一个大的信息数据库,包括学习者特征库和资源信息库,学习者特征库包括个人信息库和学习行为数据库。学习资源是个性化推荐系统的基础,在资源特征库中,云端资源库是它的数据来源。我们将在线学习资源从难度系数和资源类型两个维度进行编码,按照难度分为4种,初始资料、介绍资料、拓展资料和专业资料^[4],这在某一类知识主题中都是以一个树状的层级结构呈现的,符合学习者的认知方式。按照资源类型可以分为文本类、视频音频类和图片图表类,等,以适应不同的学习风格类型。利用数据挖掘技术中的分类方法对云端数据库中的学习资源特征扫描、分类并形成资源特征库,这样有利于资源的管理,也有利于根据学习者对不同类型的认知偏好和对知识的掌握程度进行推送,实现个性化学习。

2.2.2 个性化推荐技术 个性化推荐技术模块是推荐系统的核心模块,是实现在线学习资源个性化推荐的中枢。学习者特征模型的建构在一定程度上缓解了协同过滤算法中的数据稀疏和冷启动的问题。通过协同过滤技术和数据挖掘技术中的关联规则和Web数据挖掘等方法将学习资源库中的“孤立”资源以不同的维度关联起来,通过形成邻居和用户评分,实现优势互补,完成学习者特征与资源特征的相似度匹配,产生匹配信息库。根据学习者的学习需要系统会依据学习风格、学习状态、学习路径等以多种媒体呈现的方式为学习者在线推送其可能感兴趣的学习资源^[5]。

新近推送资源模块是对协同推荐技术的补充。“最新资源”是传统的网站和学习平台进行资源推送的一种常见方式,可以让用户和学习者及时了解各种资源的更新情况。此外,新近资源推荐模块也可以缓解协同过滤技术中的“冷启动”问题,用户接触到新资源的机会越多,就会增加新近资源的评分率。系统通过判断每个学习者的兴趣方向和专业类别,等,为其推荐相关领域的各种类型资源,避免了新近资源因缺少评分不能和学习者特征进行匹配而缺乏被推荐的机会。由此,增加新近资源模块就可以达到缓解协同过滤推荐算法“冷启动”的目的,使其形成动态的、完整的推荐系统。

3 在线学习资源个性化推荐技术的实现

下面我们通过具体的步骤与算法来说明基于协同过滤算法的个性化推荐技术的实现。

第一步:形成学习者—学习资源评分矩阵。

如表1所示,横向 Res_n 表示学习资源的个数,纵向 User_1 表示学习者的个数; E_{ij} 表示系统中第*i*个学习者对第*j*个资源的显性评分。 E_{ij} 的值表示学习者对资源的喜爱程度,一般为整数且在1~5分之间。

表1 学习者—学习资源评分矩阵

	Res_1		Res_i	Res_j	Res_n
User_1	E_{11}		E_{1i}	E_{1j}	E_{1n}
.....					
User_i	E_{i1}		E_{ii}		
User_j	E_{j1}			E_{jj}	
User_n	E_{n1}		E_{ni}	E_{nj}	E_{nn}

第二步:计算矩阵稀疏度。

首先我们给学习者—学习资源评分矩阵稀疏度设定一个最小值 α ,然后计算实际的稀疏程度。

稀疏度计算公式为: $Sparsity = \frac{E_n}{U_n \times R_n}$,其中 E_n

表示学习者对资源评价的数量, U_n 表示学习者的数量, R_n 表示学习资源的数量。当 $Sparsity < \alpha$ 时, 说明在线学习资源库评价矩阵过于稀疏, 不能和学习者特征进行匹配从而为其推荐相应的学习资源。

第三步: 以隐性评分对协同过滤算法进行修正。

根据上一步的结果就需要对学习者的特征进行挖掘, 分析学习者的学习行为, 形成对学习资源的隐性评分。我们用 R、D、T 3 个字母分别代表推荐、下载、学习时长 ($> 15 \text{ min}$) 等 3 种学习行为, 各种学习行为的发生对资源评分采用 5 分制, 如表 2 所示。

表 2 学习者行为对学习资源的隐式评分

学习者行为	分值	学习者行为	分值
R	3.5	R + D	4.5
D	4	D + T	4.5
T	4	R + T + D	5
T + R	4.5		

每个学习资源的隐性评分分值作为第一步骤的动态补充, 参与到下一步骤的计算。

第四步: 寻找相似的学习者, 形成邻居。

运用协同过滤技术最核心的步骤是寻找与当前学习者特征和对资源评分相似的邻居集^[9]。计算当前学习者 i 和其他学习者 j 之间相似性的过程如下: 首先得到学习者 i 和学习者 j 共同评分过的所有资源集合 $R_{ij} = R_i \cup R_j$, 然后通过相似性度量方法计算两个学习者之间的相似性。这里我们使用修正的余弦相似性公式, 如下:

$$Sim(i, j) = \frac{\sum_{s \in R_{ij}} (E_{is} - \bar{E}_i) (E_{js} - \bar{E}_j)}{\sqrt{\sum_{s \in R_{ij}} (E_{is} - \bar{E}_i)^2} \sqrt{\sum_{s \in R_{ij}} (E_{js} - \bar{E}_j)^2}} \quad (1)$$

通过使用公式 (1) 计算出用户间的相似性 $Sim(i, j)$ 之后, 找出与当前学习者相似性最高的前 m 个用户集合 A 作为其邻居, 即 $A = \{User_1, User_2, User_3, \dots, User_n\}$ 。

第五步: 产生推荐结果。

一般根据上步得出的邻居集中学习者对学习资源的评分, 就可以预测出当前学习者对该资源的评价, 并产生最终的推荐结果。文章采用如下推荐方式:

$$P_{i,k} = \bar{E}_i = \sum_{x \in A} Sim(i, x) (E_{x,k} - \bar{E}_x) / \sum_{x \in A} |Sim(i, x)| \quad (2)$$

公式 (2) 中 $P_{i,k}$ 表示为推荐系统预测的学习者 i 对资源 k 的评分, $\bar{E}_i$ 为学习者 i 所有评分分值的平均值, A 为上一步得出的相近邻居集, 最后取出相似度最高的 N 个资源, 得出推荐结果, 推送给在线学习者。

4 在线学习资源个性化推荐服务模型的案例

为了测试该在线资源个性化推荐模型的性能, 该研究在陕西师范大学已搭建好的《教育技术学导论》课程平台上进行了试用和反馈调查 (如图 3 所示)。这门课程是面向本科生, 采用混合式教学模式, 学习者在该平台主要是进行课前预习、课后协作讨论和反思性学习。



图 3 《教育技术学导论》课程应用平台

在该平台中, 我们实现了学习者风格测试和根据课程进度主动为学习者推送相应学习资源的功能。经过 4 周的使用, 该研究就课程资源的推荐应用效果对使用该课程平台的 21 名本科生进行了满意度调查。问卷调查包括学习者风格测试准确度, 学习资源的实用性和针对性, 对学习者的兴趣的引导, 对反思性学习的促进, 对混合式学习活动的帮助等多个方面以及对此个性化推荐模块的建议。大多数学生认为, 通过自我测评和学习风格量表能够对自己的学习风格有较为准确的把握, 最后我们用加权平均法得到的学习者对此资源推荐模块的平均满意度为 76.8%。

调查结果显示, 学生的整体满意度较高, 学习体验较好, 系统能够主动为学习者推荐符合课程内容和进度以及记忆学习者风格的学习资源, 节省了自己在网上收集有关学习资料的时间, 同时推荐的一些拓展资源也能够引导学生主动学习和发现新的学习兴趣。但从学生对个性化推荐模块提出的建议来看, 推荐的学习资源类型不是特别丰富, 还不能及时地根据学习者的在线协作学习活动和反思性学习为其提供支持性材料, 提高该模型的适应性, 增加推荐资源的种类, 使其能够支持多种在线学习活动, 提高在线学习效率, 这也将是我们为了进一步提升该在线学习资源个性化推荐系统性能的主要努力方向。

该研究从在线学习者进入学习系统学习前的自我评价和采用 Felder-Silverman 量表来推断学习者显性的学习风格, 并在学习过程中动态挖掘和分析学习者隐性学习行为特征, 不断来修正学习者学习特征模型的构建。从学习资源库中获取在线资源的分类和特征, 利用优化的协同过滤算法使之与学习者特征进行匹配, 为不同学习风格的在线学习者推荐不同类型

· 技术与教育 ·

基于跨语种的医学信息资源获取方法研究

侯震, 侯丽, 孙月萍, 李姣

中国医学科学院医学信息研究所, 北京 100020

【摘要】: 越来越多的人通过互联网获取医学信息资源, 而语种的差异成为了医学信息传播中的主要障碍。提高医学信息资源跨语种传播, 可以使优质的医学信息资源得到最大限度的利用。文章介绍了医学信息资源跨语种传播的主要方法, 以及利用该方法传播美国癌症研究中心相关数据的实例。

【关键词】: 医学信息; 信息资源管理; 多语种

【中图分类号】: G434 **【文献标志码】**: A **【文章编号】**: 1004-5287(2017)02-0176-03

【DOI】: 10.13566/j.cnki.cmet.cn61-4317/g4.201702016

Methods to acquire medical information resources based on cross-language

Hou Zhen, Hou Li, Sun Yueping, Li Jiao

Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing 100020, China

【Abstract】: A growing number of people obtain medical information via the Internet, and different language versions of medical information become the main obstacle to the spreading of medical information. By improving cross-language transmission of medical information resources, we can make the most use of high-quality medical information. This paper introduces the key techniques of spreading medical information across languages and makes a case study of applying these techniques to transmit medical data from American National Cancer Institute.

【Key words】: medical information; information resource management; multi-language

1 语种对获取医学信息资源的影响

Physician Data Query(PDQ) 是美国癌症研究中心

研发并维护综合癌症信息数据库^[1]。其包含有癌症的预防、诊断、治疗、遗传学、护理、替代治疗方法等诸多方面的信息。该数据库定期对全世界癌症研究的

的学习资源, 优化在线学习体验。随着移动终端的普及, 移动学习快速发展, 碎片化学习方式更需要人性化、个性化的服务, 增加学习者与资源的智能交互, 提高学习效率。因此, 该研究结合协同过滤技术和数据挖掘来构建在线学习资源个性化推荐服务模型, 在后续的研究中还需对协同过滤算法和数据挖掘技术的结合进行更深入地探讨和完善, 以适应学习者在移动环境下、智慧环境下的新型学习方式。

参考文献

[1] 孙歆, 王永固, 邱飞岳. 基于协同过滤技术的在线学习资

源个性化推荐系统研究[J]. 中国远程教育, 2012(8): 18-83

[2] 姜强, 赵蔚, 杜欣. 基于 Felder-Silverman 量表用户学习风格模型的修正研究[J]. 现代远距离教育, 2010: 62-67

[3] 李宝, 张文兰. 智慧教育环境下学习资源推送服务模型的构建[J]. 远程教育杂志, 2015(3): 41-49

[4] Yang Y J, Wu Chuni. An attribute-based ant colony system for adaptive learning object recommendation[J]. Expert Systems with Applic, 2009, 36(2): 3034-3047

[5] 叶树鑫, 何聚厚. 协作学习中基于协同过滤的学习资源推荐研究[J]. 计算机技术与发展, 2014, 24(10): 63-67

基金项目: 社科基金项目“面向知识服务的公众健康知识组织体系构建研究”(14BTQ032); 国家人口与健康科学数据共享平台国际合作(NCMI-MEC01-201605)

收稿日期: 2016-11-02

作者简介: 侯震(1991-), 男, 北京人, 研究实习员, 学士, 主要研究方向: 多媒体技术在医学教育中的应用。

通信作者: 李姣(1981-), 女, 辽宁本溪人, 副研究员, 博士, 主要研究方向: 医学信息学, 数据挖掘, 知识发现, 医学信息可视化。

电话: 010-52328740; E-mail: li.jiao@imicams.ac.cn