

基于混合机制的个性化学习资源推荐系统研究

——以国家开放大学学习网为例

杨伟 韩立平

(国家开放大学 北京 100039)

[摘要]随着信息技术的迅速发展,学习者可获得的学习资源形式更加多样,数量更加庞大。如何利用大数据、AI,向学习者推荐个性化的学习资源成为人们关注的焦点。根据学习者的特征,塑造个性化的学习路径将真正满足学习需求,从而提高学习水平。因此,本文以国家开放大学学习网为例,通过对已有的不同的算法进行研究,从而基于混合推荐机制,设计了国开学习网的个性化学习资源推荐系统。

[关键词]混合推荐;个性化;学习资源推荐系统;国开学习网

[中图分类号] G640

[文献标识码] A

[文章编号] 2096-711X(2020)13-0065-03

doi: 10.3969/j.issn.2096-711X.2020.13.030

[本刊网址] <http://www.hbxb.net>

一、研究背景

教育信息化的推进,使在线教育平台如雨后春笋般涌现出来,给学习者提供海量学习资源的同时,也使学习者面临着如何科学选择学习资源的问题。平台应满足学习者这一诉求,提供个性化学习路径。《教育信息化十年发展规划(2010-2020)》也明确提出,应努力为学习者提供个性化的信息化学习环境与服务。而平台解决这一问题的有效途径就是建立个性化学习资源推荐系统,向学习者提供精准的学习资源服务。推荐系统是指导用户发现可以满足自身需求信息的工具,最早用于电子商务领域,随着在线教育的发展,建立学习平台的学习资源推荐系统显得更加迫切,尤其对于实施远程开放教育的国家开放大学来说。

为满足开放大学的发展需求,解决原有的电大四级平台存在的教学问题,2015年秋季,国家开放大学决定将中央电大在线平台的所有统设课程迁移到国开学习网,为教师、学生和教学管理者提供一站式服务。学习网内嵌入了一个Moodle平台,可以实现课程管理、学生管理及教师管理。但通过观察平台中的学习服务,发现平台提供的学习资源缺乏针对性,缺少个性化的学习资源推荐,多数界面为静态化技术解决。国开开办的学生来自全国各地,在年龄、职业、学历等各方面差异显著,因此应把握学习者的差异化特征,提供个性化学习资源。同时,国开正在大力探索引进外部优质资源,学习网中的学习资源将更加丰富,可以满足学习者多元化的学习需求。建立此系统,将增强学习者的学习兴趣,提高学习者的学习自主性。学习者通过发现适合自己的学习资源,将会拓展相关学习,建立新型的学习路径,提高综合素养。同时,也利于学校深入分析学习者各项数据,真正了解学生的特征。系统通过获取学习者数据,并进行分析,能形成更加完善的用户画像,为平台各项功能的改善提供数据基础。

二、相关研究

个性化推荐的概念在1994年由Resnick最早提出,之后,人们开始对其进行研究。研究主要分类理论设计与实践应用两个方面,多数为研究设计。S. Bhaskaran等人提出了基于信任的混合推荐策略,其使用AprioriAll算法对学习者的浏览日

志序列进行挖掘,分析学习者的特征,实现对学习内容的自定义推荐。De-Marcos认为学习资源推荐是多目标组合优化问题,对学习者的建模以优化、组合适当的学习资源的内容和序列是其中最关键的技术。刘敏、郑明月提出从学习风格、在线学习偏好、学习者知识结构以及在线学习行为等方面对学习者的分析,从而为其提供个性化学习资源。巩晓月将深度学习模型与概率矩阵算法进行融合,提出了深度学习的概率矩阵分解模型,该模型将提高推荐的准确性。只有少数研究者进行了实践,建立了推荐系统并推广应用。Mampadi等人开发了一个能根据学习者的学习风格为其提供不同的用户界面的个性化网络学习资源系统,许薇设计并实现了基于个性化学习资源推送的平台,从学习过程隐形评分设计、测试题设计、资源评分等方面优化了协同过滤算法,设计了模型。

但关于国开学习网的个性化学习资源推荐系统研究较少,大部分研究都是关于国开学习网应提供个性化的学习服务,没有从技术层面解决问题。关德君提出国开学习网应充分应用大数据。刘娜提出要挖掘学习者的学习需求,提供个性化、全程跟踪式的学习服务。仅有刘超、薛羽等对国开学习网提出了基于智能分析学习行为数据的辅助教学系统设计,系统主要由数据收集系统、分析系统、推送系统和评价系统四部分组成。可见,以国开学习网为例,提出建立个性化学习资源推荐系统的研究较少,本次研究是一次相对创新的尝试。

三、针对在线学习资源不同的推荐算法研究

(一) 基于内容的推荐算法

基于内容的推荐是一种非常简单且实用的推荐算法,此算法通过利用已知的用户特征和内容特征进行匹配。针对在线学习领域,此算法的难点主要是对学习资源的分类,传统的方式会通过课程名称的相似度分析、基于TF-IDF算法的课程简介分析等算法对文章进行分类。针对国开学习网,课程的资源其实已有了一些明确的分类标签,例如将课程初步分为学部制:文法教学部、经管教学部、理工教学部、外语教学部、农林医药教学部、教育教学部、艺术教学部。不同的课程就可以通过学部名称作为一种固定的分类模式以供推荐。并且开放大学在入学时就可以利用用户填写的一些基础信息进行用户画像

收稿日期: 2020-4-18

作者简介: 杨伟(1992—),女,回族,河北沧州人,国家开放大学学习资源部科员、实习研究员,研究方向:开放教育资源。

的构建,这些基础信息包括但不限于学生性别、年龄、最高学历、从事职业、地区等。因此基于内容的推荐算法特别方便在国开学习网进行开发与使用。但是此算法的缺点是不能实现推荐的个性化,推荐内容可能会无法进行创新,因此会产生学习资源孤岛,不利于学生掌握知识的完整性、知识扩展与推荐的个性化。

(二) 协同过滤算法

协同过滤算法指的是个体借助集体智慧,通过拥有具有某种共同属性的集体的喜好来推荐个体用户感兴趣的内容。协同过滤算法可以解决基于内容推荐所产生的内容缺乏新意,推荐不个性化的问题。通过集体智慧填补用户学习资源孤岛、实现用户的个性化推荐。协同过滤算法可以分为基于用户的协同过滤和基于项目的协同过滤两种,基于用户的协同过滤是计算不同用户之间的相似度矩阵,同类用户之间可以相互推荐之前未访问的内容;基于项目的系统过滤是计算物品之间的相似度,根据物品的相似度和用户的历史行为给用户生成推荐列表。目前对于国开学习网来说更合适与基于项目的协同过滤,但是随着国家开放大学逐渐引入国内或国外的优质课程资源,因此在海量学习资源的场景下基于项目的协同过滤算法不一定完全适应于未来的国开学习网。

(三) 基于深度学习的推荐算法

随着人工智能技术的不断发展,各行各业都开始通过应用AI技术完善业务质量。在人工智能技术中深度学习是当前使用最为广泛的算法。深度学习可以通过用户画像挖掘用户和项目内容的隐特征,借此辅助信息来完善传统的推荐算法。但是由于深度学习模型复杂,并且模型的训练周期较长,而推荐系统又有一定的实时性要求因此,基于深度学习的推荐算法通常作为一种辅助的手段来支持推荐系统。

四、基于混合推荐机制的国开学习网推荐系统

通过对不同推荐算法的对比发现,仅依靠一种推荐算法是难以满足当前的国开学习网的推荐学习资源的所有需求,并且随着学校的不断发展,用户的数据量、课程量都会呈现指数级增长,需要设计一种更加完善的推荐系统来满足学生的在线学习需求以及系统不断发展的技术要求。因此,本文提出了一种基于混合推荐机制的推荐系统,首先介绍了混合推荐机制,之后以国开学习网为例设计了整个在线学习平台的推荐系统。

(一) 混合推荐机制

混合推荐机制是指不单采用某一种推荐算法,而是将多种推荐算法混合在一起,通过不同的混合方式可以弥补单一一种算法的不足,大幅提高推荐的质量。目前混合推荐机制可以分成七类:

1. 加权混合 通过线性加权不同推荐算法得到最终的分值。

2. 切换混合 根据系统中的数据情况对不同的推荐算法因情况进行切换。例如在国开学习网的初期课程数量相对用户数量较少,因此可以采用基于用户的协同过滤,随着国内外优质课程的不断引入,当课程数量多于用户数量时可以切换成基于项目的协同过滤。

3. 分区混合 将多种推荐混合并将不同的推荐结果分不同的区显示给用户,可以提高推荐内容的多样性消除知识孤岛效应。

4. 特征组合混合 特征组合利用多个推荐算法的特征数据来作为原始输入,利用其中一个算法作为主算法最终生成推荐结果。

5. 特征增强混合 第一个算法可能事先预处理第二个算法依赖的数据,生成中间可用的特征或者数据(中间态),再供第

二个算法使用最终生成推荐结果。

6. 级联混合 在第一阶段中,级联混合推荐代理生成候选者排名,而第二推荐代理被用作决胜局,并提供更完善的结果。

7. 元级别混合将两个不同的推荐组件组合在一起,其中第一个组件生成一个模型,该模型又用作第二个组件的输入。

(二) 国开学习网推荐系统设计

以国开学习网为例,利用混合推荐机制来设计一个高可用、高质量的在线学习资源的推荐系统。系统可以分成以下六部分:

1. 用户行为日志收集模块 通过离线处理的方式采集用户的行为日志,用以构建用户画像,用户的行为日志包括用户编号、用户所观看的课程编号、用户观看某课程的时间、课程所属的类型、访问时间。使用MySQL关系型数据库将用户的行为日志进行存储。

用户画像构建模块 在获取了用户的行为日志之后可以对刻画用户画像。本推荐系统通过对用户进行相似性分析与人工标注的方式对不同的用户进行分析与刻画。首先基于用户的行为日志,针对所观看课程的名称、种类与时长组成特征向量计算不同用户间的相似度。相似度计算采用余弦相似度,计算公式为:

$$\text{Cos}(x, y) = \frac{x \cdot y}{\|x\| * \|y\|}$$

x, y 分别表示两个不同用户的特征向量。由于此计算过程较长因此采用离线处理的方式。

2. 学习资源分类与特征提取模块 随着学校的不断发展,课程的数量将不断上涨,对于一些国开引入外部学习资源其课程分类可能会出现不完整、与国开学习网内资源课程类型不统一等问题。因此若想实现基于内容的推荐算法,需要对学习资源进行分类,方便进行相似内容的推荐与基于项目的协同过滤。学习资源的分类通过TF-IDF的方法对资源的简介、标题进行分词。

3. 资源召回模块 召回阶段利用混合推荐机制对海量的学习资源进行召回。在不同的阶段将采用不同的混合方式。当用户数量与资源数量相差较多时通过切换混合来选择使用基于何种协同过滤计算进行推荐;当用户更加注重短期偏好时,通过加权混合的方式提高短期用户感兴趣资源的权重;当推荐数量不足的时候,可以采用分区混合的方式将热门资源与推荐资源进行混合召回;之后,对于一些非本校学生的访客,可能存在用户冷启动的问题,利用切换混合为没有特征的用户召回热门数据,对于存在少量特征的用户采用基于项目的协同过滤。

4. 资源排序模块 在排序阶段对召回的资源进行打分与排序,打分策略即为通过不同的加权混合推荐算法所打的分值在此阶段可以对类似的学习资源进行去重与打压,尽量保持学习资源的多样性。

5. 人工干预模块 此模块是针对一些特殊的学习资源在刚上线期间没有或很少用户访问的情况下还需要尽量向用户推送的内容。例如在防疫期间,可以向用户推送优质的有关防疫知识的课程等。

国开学习网平台对于学习者资源推荐的质量,将影响到学习者的学习兴趣和学习效果。因此,基于混合推荐机制,本文设计了国开学习网的个性化学习资源推荐系统。以期引发相关的技术探索,并能进行实践,将系统真正用于国开学习网中。

(下转第69页)

补充方式,在量化的基础上提升资助的温度,加强对学生的
人文关怀,通过多渠道、多主体、多方式力求对学生家庭状况有全
面真实的把握。

需要注意的是,对于“隐藏困难生”而言,要格外注意方式
方法,探寻其隐藏背后的原因,是那骄傲的自尊、是不愿为人所
知的家庭或身体上的缺失、是想把资助让给更需要人的那份无
知和善良,真正走进学生心里去感同身受,引导学生以科学理
性的态度看待。

(五) 制定严明的监督惩罚机制

有效的监督惩罚机制便是高校资助育人的“最后一公里”。
长期监督机制的缺失使部分学生形成了认识误区,他们错误地

认为受资助是自身不可剥夺的权利。

对证明材料造假、未按时还款失信的不良行为必须要有成
文规定。文件的制定本身带有约束力、威慑力和执行力,针对
行为严重程度,确定惩罚不同等级,较轻以批评教育为主,重则
收回资助金、记入学生不良信用记录、给予严肃处理处分。

在开展具体资助工作、资助政策宣讲、榜样实例分享等关
键事件和重要时间节点上讲明、讲清失信、造假、不作为等违
规行为的代价,将违规必罚的理念通晓到每一位同学,形成校
园内互相监督的氛围,使心存侥幸的同学不敢轻易触碰底线和
红线。

参考文献:

- [1] 张志勇.新形势下高校资助育人工作探索[J].学校党建与思想教育,2019(11):84-86.
- [2] 教育部要求进一步加强高校学生资助工作机构建设[DB/OL].http://www.china.com.cn/chinese/PI-c/1220288.htm,2016-5-26.
- [3] 朋辈心理咨询——技术伦理与视角第2版[M].北京:中国人民大学出版社,2013.
- [4] 教育部等六部门关于做好家庭经济困难学生认定工作的指导意见[DB/OL].http://www.moe.gov.cn/srcsite/A05/s7505/201811/t20181106_353764.html,2018-10-30.
- [5] 杜志欣,付靖崧,李悦宁.高校贫困生资助工作的困境及策略——以肇庆学院为例[J].高教探索,2019(2):116-121.

Research on the Realistic Dilemma and Path of Subsidized Education in the University

WANG Hui

(School of Material Science and Engineering, Tiangong University, Tianjin 300387, China)

Abstract: As the highest level of education, universities are the main positions for subsidized education, whose effects concern the underlying problems of who, how and for whom to cultivate, which is the key to foster virtue through education. Now the realistic dilemma of subsidized education centralizes the deficient of proportion and professionalization of practical subject, narrowing the target of funded education, the scanty of policy propaganda, incomplete mechanism of accurate funding. Around the above problems, treating all the student as the target of funded education and realizing subsidized students' shift from object to subject, ensuring the mechanism of accurate funding, propaganda and superintendent, provide reference to step out of dilemma.

Key words: university; subsidized education; dilemma; countermeasures

(责任编辑:桂杉杉)

(上接第 66 页)

参考文献:

- [1] 教育部关于印发《教育信息化十年发展规划(2011-2020年)》的通知[EB/OL].http://old.moe.gov.cn/publicfiles/business/htmlfiles/moe/s5892/201203/133322.html,2012-3-13.
- [2] 胡晶,熊伟.学习网 moodle 平台的主要功能及其应用[J].河北广播电视大学学报,2016(2):15-18.
- [3] 刘敏,郑明月.智慧教育视野中的学习分析与个性化推荐[J].中国电化教育,2019(9):38-46.

Research on Personalized Learning Resource Recommendation System Based on Hybrid Mechanism ——Taking the Learning Network of the Open University of China as an Example

YANG Wei, HAN Li-ping

(The Open University of China, Beijing 100039, China)

Abstract: With the rapid development of information technology, the form of learning resources available to learners is more diverse and the number is even greater. How to use big data and AI to recommend personalized learning resources to learners has become the focus of attention. According to the characteristics of learners, shaping a personalized learning path will truly meet the learning needs, thereby improving the level of learning. Therefore, this article takes the Learning Network of the Open University of China as an example, through the study of existing different algorithms, based on the hybrid recommendation mechanism, designed a personalized learning resource recommendation system of the China's network development.

Key words: hybrid recommendation; personalization; learning resource recommendation system; learning network of the Open University of China

(责任编辑:章樊)