

基于 Moodle 平台“教”和“学”行为的个性化分析

□魏莉¹ 韩艳辉² 张智鹏³ 许佩彤⁴

(1. 山西广播电视大学, 山西 太原 030027; 2. 国家开放大学, 北京 100039;
3. 北京开放大学 石景山分校, 北京 100041; 4. 开平开放大学, 开平 529300)

摘要: 以开放源代码为代表的 Moodle 在线学习平台已经成为“互联网+”时代的重要教学环境之一, 它提供行为痕迹记录、实时统计和追踪等功能, 积累了大量在线教育过程中的案例和数据, 为基于大数据分析技术研究个性化教育提供了数据基础。选择一门基于 Moodle 学习平台开发的在线课程为样本, 利用教师和学习者在线行为日志数据, 分析网上教师和学习者在线行为的个性化规律及相关度测算, 构建个性化学习分类模型, 利于开展个性化学习服务, 提升学习效果。

关键词: 大数据分析; Moodle 平台; 个性化教育

中图分类号: G434 **文献标识码:** A **文章编号:** 1008—8350 (2019) 04—0010—05

在大数据时代, 数据挖掘可以有效地用于提取潜在的模式和关联性, 已被广泛应用于教育、医疗、交通等诸多领域。随着在线教育的盛行、在线教育资源快速增长以及教育数据挖掘技术、信息技术和人工智能的发展, 在线教育平台能够使得更多的人有机会接受教育, 并且体验到更加智能化的学习支持服务。智慧教育的真谛就是通过人机协同的数据智慧和教学智慧, 以“精准、个性和优化”为原则, 让教师能够施展高效的教学方法, 让学习者能够获得适宜的个性化学习服务和学习发展体验。

在线课程平台众多, 如 edX、Udacity、Coursera、中国大学 MOOC、学堂在线和智慧树等, 教师可以快速灵活地开发在线课程, 并与学习者开展互动交流, 学习者可以自由选择学习内容和学习时间。但相比

传统教育, 在异步学习平台上开展远程教育存在较大的局限性, 比如学习者和教师之间如何开展互动, 如何在学习平台开展个性化的教学服务。如果教学支持服务的个性化没有得到有效的实施, 将可能出现高辍学率。比如有的学习者认为学习内容不够有挑战性, 可能会影响学习积极性; 如果学习内容太难, 可能因为跟不上进度而放弃。面对这些挑战, 要想实现教育个性化, 就必须根据学习者的在线学习行为了解他们的现状和学习需求, 并根据学习者的个性化特点及时地给予反馈。

国家开放大学学习平台是在开源的 Moodle 学习平台基础上开发的, 是结合教师和学习者的需求定制开发的在线学习平台, 能够记录所有的在线行为痕迹, 积累了教学和学习过程中大量的案例和数据, 为通过大数据分析教学的个性化特征提供了数据基础。

一、国内外网上学习行为分析趋势

目前, 国内外关于网上学习行为的分析所使用的数据普遍是基于 Moodle 平台和慕课(MOOC)平台的。Moodle 学习平台提供的教学资源种类多于 MOOC 平台, 教学模式更丰富, 互动方式更多样, 提供学习情况追踪等功能, 从学习分析方面来看, 优势相对较明显。

收稿日期: 2019—09—25

基金项目: 国家开放大学 2016 年度青年课题(G16A0403Q); 国家开放大学“十二五”规划 2015 年度专项课题(G15A0008S)

作者简介: 魏莉(1984—), 女, 河北邢台人, 山西广播电视大学, 讲师;

韩艳辉(1972—), 男, 河北沧州人, 国家开放大学, 副教授;

张智鹏(1981—), 男, 山东德州人, 北京开放大学石景山分校, 讲师;

许佩彤(1980—), 女, 广东开平人, 广东省开平开放大学, 讲师。

国外关于在线教育数据的挖掘方法探究很广泛。目前马来西亚开放大学 Katrina Sin 和印度 Bharathiyar 大学的 Loganathan 分析总结了处理大数据的挑战,归纳了教育数据挖掘领域常用的技术,如回归分析法、近邻技术、聚类和分类,以及开源的工具如 Hadoop、Orange 和 Weka 等。巴基斯坦古吉拉特大学计算机中心和旁遮普大学信息部的 Shafiq Aslam 等人介绍了数据挖掘算法在教育挖掘领域中的应用,并介绍了 SVM、EM、PageRanker、朴素贝叶斯和 Apriori 算法的重要应用价值。^[1-3]

国内对于在线学习行为数据分析技术的研究文献众多。国家开放大学魏顺平^[4-6]等人做了一系列基于 Moodle 平台的学习分析技术研究,以实例从管理者、辅导教师、学习者等不同用户视角展示了学习分析技术的应用过程。他采用演绎和归纳相结合的方法,参考 Moodle 学习平台中的数据表结构,构建学习分析数据模型,并详细介绍数据结构及其处理方法,提供学习分析的技术细节,为本领域的其他研究提供参考。郭富强基于 SOM 网络建立教学行为分析模型,解决了数据分析中多维向量的有效聚类的问题。贾积有^[7]等人对北京大学 6 门 MOOC 的学习行为及效果进行线性回归分析,发现成绩与平时测验成绩和论坛活跃程度(发帖、回帖)有比较强烈的正相关性,并使用 SQL 语言、SPSS 和 WEKA 等数据挖掘软件对这些海量数据进行了细致的分析。宗阳^[8]等人基于 Moodle 平台在线行为数据进行挖掘分析,通过对论坛文本数据进行情感计算,分析关于学业情绪与在线学习行为之间的关系。

国内外关于个性化教育相关研究主要有以下两个方面。一方面是预测学习效果,大数据被广泛用于预测分析,具体到在线学习行为分析,如预测学习者的学习效果和未来表现。英国开放大学教育技术研究所发现,学习分析的方法可以解析、评估历史行为的状况和规律,发现潜在的问题,也支持对学习者和教育者预测未来成果,如预测学习成绩、未来表现等。蒋卓轩等人从北京大学 MOOC 上选择典型的学习行为特征,对学习者的学习效果进行分析,可预测出潜在的认真学习者。

另一方面是关于个性化资源推荐。美国新媒体联盟 NMC 的地平线报告中,连续几年都预测学习分

析技术将是未来的研究主流。曹帅等人基于 2011-2015 年“学习分析技术与知识国际会议”论文认为,学习分析技术与个性化学习的实现是未来重点研究的课题之一。美国教育部教育技术办公室 Marie Bienkowski 等人研究利用学习分析改善学习支持服务,并改善可衡量的目标,探索个性化学习方案及自适应学习系统,学习者监控自己的学习行为,教师针对学习者的行为进行课程完善和教学干预,值得借鉴。马德坚^[9]分析了个性化教育发展的过程,发现了大数据对实现个性化教育的潜能优势,为实践个性化教育提供了理论指导,但是缺乏数据支撑和具体的实现方案。刘海鸥等人基于个性化学习数据构建画像模型和精准服务模式,缺少数据实证分析。刘敏^[10,11]等人提出智慧教育视野的个性化资源推荐,但是学习分析和资源推荐的过程中人工干预过多。

从上述研究我们可以看出,国内外学者已经在探索将大数据分析的研究方法应用于在线教育领域。通过构建在线学习分析模型,对在线学习行为深入挖掘,能够较准确地反映学习者的在线学习表现或者预测学业成绩并推荐学习资源。但是,针对如何将教师和学习者的在线行为数据进行综合考虑,如何通过分析在线行为痕迹发现学习者和教师的个性化特征,及时的改进教学方式,如何使教师因材施教、有的放矢地开展教学指导,如何具体实现个性化教育仍然有待探索实践。

二、个性化教育模型设计与实践

本研究通过收集 Moodle 平台上的教师和学习者的大型数据集,选取国家开放大学 Moodle 平台开设的一门课程作为研究实例,用大数据方法进行分析挖掘,分析教师和学习者在时间方面的个性化规律,分析学习者的学习风格和学习能力等个性化特征,教师可以依据分析结果显示出的学习者个性化特点实施教学干预。

《媒体辅助英语教学》是一门在线课程,由国家开放大学韩艳辉老师率领团队创立了“一站式、跨区域协作教学模式”,该课程由国家开放大学整合全国广播电视大学(开放大学)的优秀教师组建在线教师团队。该课程在教学平台上组织实施教学活动,引导学习者自主学习;提供共享交流平台,推动

学习者协作学习;提供软件(云应用)学习资源,提升学生实践能力;多种测评及时反馈,激发学生主动学习;全程记录学导互动,不断提升教学效果。课程总时长约十周,配备专业的“辅学团队”,团队有分工和协作。此课程不设终结性考试,采用百分之百形成性考核的方式,可见教学过程完整的记录到在线平台上,使得日志数据能够较完整、客观、真实和准确地反映学习者的学习表现。本研究对象是2018年秋季学期在线课程,共计46名教师,学习者共计103名,截至2018年12月30日17:00,全体师生的在线行为记录一共109869条,其中教师行为记录54195条,学习者行为记录55674条。

(一) 时间个性化模型

在线教育中师生的异地和时间异步特征,使得师生不能像传统的面授教育那样交流,容易导致在线学习者的情感缺失。当学习者困惑以及受到挫败时,如果教师或者学习同伴能即时地给予指导和帮助,就能够减少学习者在虚拟网络学习空间中的孤独感。学习者的求知欲望和交往诉求以及教师的支持响应,都会对课程整体的学习氛围、交往黏度产生直接的影响。分析教师和学习者个性化时间规律,统计行为高峰时间^[12],有利于解决教师和学习者异步交流的问题,缩短学习者等待教师回复的响应时间,有助于增加和激发学习者的学习兴趣。

将整个学期的教师和学习者的所有在线行为时间作为研究范围,统计一天中的六个时间段,教师和学习者的在线行为频次,如图1所示。我们可以看出学习者们经常学习到深夜,教师们也辅学到深夜。

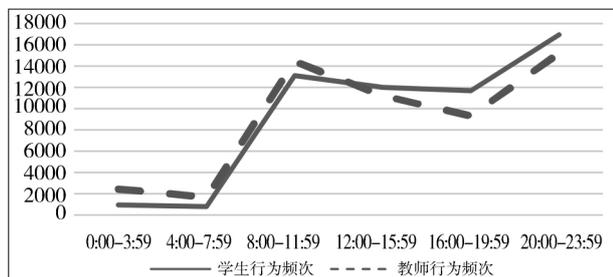


图1 教师和学习者在一天中的六个时间段的在线行为频次

分析一天的六个时间段中教师在线行为频次和学习者在线行为频次的相关性,用软件SPSS 20对双变量做皮尔逊相关性分析,结果为在0.01水平上显著正相关,相关系数高达0.976,说明教师和学习者的在线时间规律高度一致。

学习者的在线时间规律高度一致。

将整个学期的教师和学习者的所有在线行为时间作为研究范围,统计教师和学习者一周七天中的在线行为频次,如图2所示。

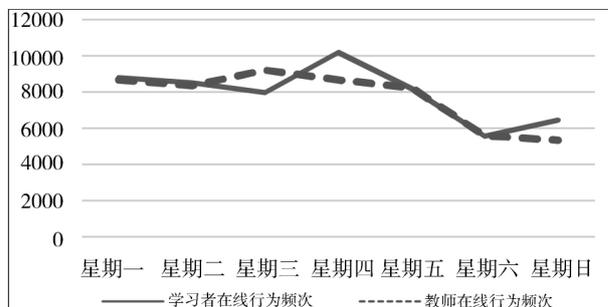


图2 教师和学习者在一周七天中的在线行为频次

分析一周七天中教师在线行为频次和学习者在线行为频次的相关性,使用软件SPSS 20对双变量做皮尔逊相关性分析,结果为在0.05水平上显著正相关,相关系数达到0.835,说明教师和学习者的在线时间规律一致性较高。

通过在教学过程中分析学习平台的数据,发现学习者和教师的个性化时间规律,学习者的在线学习时间主要集中在晚上,周一至周五明显高于周末,并且周四达到峰值,辅学团队即时调整辅学的时间节奏,这样师生的交流就比较同步,学习者可以即时得到教师的反馈。可见分析学习者个性化学习时间规律,调整辅学时间效果明显,这也侧面说明研究远程学习者学习时间个性化特征的必要性。

(二) 在线行为频次个性化模型

本研究将整个学期的教师和学习者的所有在线行为作为研究范围。该课程以模块为单位,每个模块配备一至两位辅学教师负责学术型辅学工作,基于模块计算,师生比约为1:50。本学期46名教师团队的活动记录和论坛发帖的数据统计显示,主要是13名老师起到核心作用。统计各学习模块的学习者在线行为频次和教师在线行为频次,如图3所示。学习模块二的内容为理论学习部分,图3数据显示这个模块的学习者在线行为频次出现低谷。从学习模块三开始是实践内容,教师依据数据分析结果和学习内容特定增加辅学频次后,激发了学习者的参与度,学习者在线行为频次明显增加,在模块四区间达到一个小高峰,课程整体的学习氛围高涨。

统计各学习模块的学习者在线行为频次和教

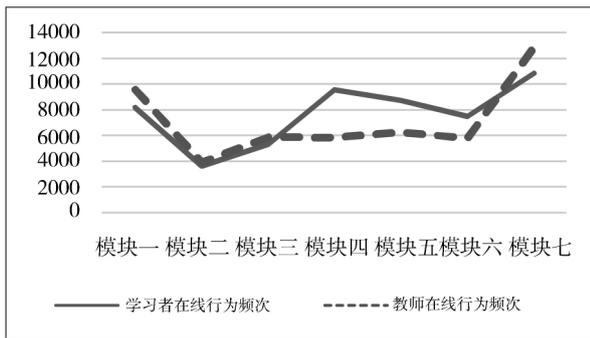


图3 按学习模块统计教师和学习者在线行为频次

师在线行为频次,使用软件 SPSS20 对双变量做皮尔逊相关性分析,结果为在 0.05 水平上显著正相关,相关系数达到 0.811。说明教师和学习者的在线行为频次具有很高的 consistency。可见分析学习者每个学习模块的频次,能够准确地了解学习者的在线行为个性化特征,及时调整教学策略,激发学习积极性,使得师生形成良性互动。这也侧面说明研究远程学习者学习行为频次个性化特征的必要性。

(三) 学习者在线投入风格个性化模型

学习者的在线投入时间和风格存在较大的差异性和变化性,仅仅通过问卷调查很难全面、深入、准确地把握学习者的真实情况,需要充分利用大数据分析技术对学习者的海量数据信息进行深度的分析挖掘。本课程共计 103 名学习者,经过数据清洗,选取实际参与学习的 81 人(因 8 人没有实质参与实践模块学习,另有 14 人因其他问题未取得成绩)。筛选学习者个性化特征要素,确定通过对学习者、学习时间和学习结果聚合作为学习者分类变量,以“学习者”为主分类变量,将“时间”划分为一个学习模块,每个模块将成为“学习者的属性”,以每个学习者的“成绩”为分类变量。需要对每个变量的取值都要进行归一化处理,将其全部转化为 [0,1] 之间的取值。时间按学习模块划分,以“学习者”“模块”为分类项,以“在线行为频次”为汇总项,进行分类计数,得到一个 $m \times n$ 数据矩阵,可以了解学习者在各学习模块的学习投入情况。

将整个学期的教师和学习者的所有在线行为作为研究范围,从“时间”中抽取出“模块”信息,然后以“学习者”“周次”为分类变量,以“行为频次”为汇总项,进行分类计数。开始时间为 2018 年 10 月 15 日,结束时间为 2018 年 12 月 16 日,共计 7 个学

习模块,统计每个学习者在这 7 个学习模块的在线行为频次,然后进行归一化处理。聚合学习结果特征作为分类变量,统计每个学习者的学习成绩。这个矩阵表格信息量较大,需要采用聚类算法,帮助我们看清学习者在学习投入风格上的情况。

由聚类所生成的簇是一组数据对象的集合,这些对象与同一个簇中的对象彼此相似,与其他簇中的对象相异。聚类是通过识别研究对象之间的距离在一个 n 维空间,其中 n 是变量的数目,对相似的对象分组。根据矩阵数据,采用 SPSS 20 的 K-means 算法,对学习者进行聚类分析,预设分为 2-5 类,结果分为 3 类时最佳。每个聚类中的案例数目如表 1 所示。

表 1 每个聚类中的案例数目

类别	案例数目	案例比例	
聚类	1	8	10%
	2	27	33.75%
	3	45	56.25%
有效	80		
遗漏	1		

根据表 1 的聚类结果,可绘制折线图,如图 4 所示。

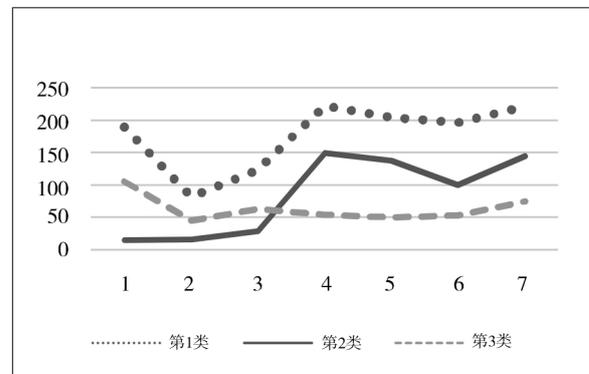


图4 聚类结果折线图

从图中可以直观地看到,第 1 类学习者属于持续积极型,占 10%;第 2 类学习者属于迎头赶上型,占 33.75%;第 3 类学习者属于持续平稳型,占 56.25%。由此可见,个性化的学习支持服务非常重要,本门课程的辅导教师每周汇总数据分析结果,能很好地发现学习者投入风格的个性化特点,使得积极的学习者不断求知,热情不减;使得“后进生”跟上进度,迎难而上;使得相对“沉默”的大多数学习者稳步学习,踏实进取。

本实验说明,在学习过程中利用聚类方法实现“人以群分”,对学习者的画像,有利于因材施教。可以利用分类结果预判学习者的个性学习需求,为学习者提供适合其个性发

展需要的服务 给予不同的教学指导,有利于改进学习效果。

三、结语

以大数据统计思想为基点,基于国家开放大学学习网上的教学和学习行为数据记录,能够比较准确和科学地描述学习者的个性化时间和在线学习行为特征,分析网上教学和学习行为的个性化规律及相关度测算,利用数据挖掘技术构建个性化学习分类和聚类模型,判断学习者的个性化学习风格,预测学习者的学习需求。在课程的教学过程中分析挖掘当前的学习行为数据,可以准确把握学习者的学习状况,有利于教师及时地进行学习干预,调整教学指导时间、方式和节奏,提升个性化教育质量。在线教育蕴藏着丰富的教育数据信息,深度挖掘的成果可以作为教学人员和教学管理人员科学决策的依据。在线学习空间需要优化资源供给,改革教学模式、重塑评价方式、创新服务模式,提升监管水平,需要基于在线行为数据深入探索。智慧教育盛行,需要深入探索学习分析技术在个性化教育方面的应用,助力教学的实践与改革。

参考文献:

- [1] Marie Bienkowski, FengMingyu, Barbara Means. Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief[R]. Washington, D.C. U.S. Department of Education. Office of Educational Technology, 2012.
- [2] Mike Sharples, Patrick McAndrew, etc. Innovating Pedagogy 2013: Open University Innovation Report2 [R]. Milton Keynes: The Open

University, 2013.

[3] Shafiq Aslam, Imran Ashraf. Data Mining Algorithms and their applications in Education Data Mining [J]. International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 7, July 2014(1): 50-56.

[4] 魏顺平. 学习分析技术: 挖掘大数据时代下教育数据的价值 [J]. 现代教育技术, 2013(2): 5-11.

[5] 魏顺平, 赵璇, 程罡. 数据驱动的在线学习表现模型构建研究 [J]. 天津大学学报, 2015(3): 35-41.

[6] 魏顺平. 学习分析数据模型及数据处理方法研究 [J]. 中国电化教育, 2016(2): 8-11.

[7] 贾积有. MOOC 学习行为及效果的大数据分析——以北大 6 门 MOOC 为例 [J]. 网络化与数字化, 2014(9): 23-29.

[8] 宗阳, 陈丽, 郝勤华, 胡红梅. 基于在线学习行为数据的远程学习者学业情绪分析研究——以 Moodle 平台为例 [J]. 开放学习研究, 2017(6): 11-20.

[9] 马德坚. 大数据支持下的个性化教育实现 [J]. 软件导刊·教育技术, 2016(2): 52-53.

[10] 刘敏, 郝明月. 智慧教育视野中的学习分析与个性化资源推荐 [J]. 中国电化教育, 2019(9): 38-47.

[11] 章文捷, 韩艳辉, 杨永建. 提升在线课程学习者参与度和学习效果的研究 [J]. 福建广播电视大学学报, 2016(3): 21-25.

[12] 魏莉. 远程教育时间管理的分析与策略 [J]. 山西广播电视大学学报, 2015(1): 14-16.

Personalized Analysis of “Teaching” and “Learning” Behaviors Based on Moodle Platform

Wei Li¹, Han Yanhui², Zhang Zhipeng³, Xu Peitong⁴

(1. Shanxi TV University, Taiyuan, Shanxi, 030027; 2. The Open University of China, Beijing, 100039;

3. Shijingshan Branch, Beijing Open University, Beijing, 100041; 4. Kaipig Open University, Kaiping, 529300)

Abstract: Based on open source coding, Moodle, the online learning platform has become one of the important teaching environments in the era of “Internet +”. It provides such functions as behaviors tracing, real-time statistics producing and tracking, and accumulates a large number of cases and data in online education process, establishing a data foundation for researching personalized education based on big data analytics technology. This paper selects an online course based on Moodle as a sample, uses the online behavior log data of teachers and learners, analyzes the personalized rules and correlation calculation of online behaviors of teachers and learners, and constructs a personalized learning classification model, which is conducive to offering personalized learning services and improving learning effect.

Key words: big data analysis; Moodle platform; personalized education

本文责编: 董娜