

基于平台数据挖掘的大学生在线学习行为画像的构建

李晓红, 金伟锋

浙江中医药大学药学院, 浙江 杭州 311402

摘 要 : 通过对网络教学环境下大学生在线学习行为特点的分析, 构建大学生在线学习行为画像。首先, 在已有研究理论基础上, 将学生的在线学习行为归纳为四个维度七个指标。然后, 采用相关性分析指标之间以及指标与总成绩之间的相关性, 确定核心指标。最后, 以七个指标对大学生的学习行为进行聚类分析, 将学习行为分为四个类别, 分析四个类别的特点, 探讨大学生在线学习行为的特点和规律。

关 键 词 : 在线学习行为; 数据挖掘; 交互式行为; 聚类分析

Construction of College Students' Online Learning Behavior Portrait Based on Platform Data Mining

Li Xiaohong, Jin Weifeng

college of pharmaceutical science, ZCMU, Hangzhou, Zhejiang 311402

Abstract : This study analyzes the characteristics of college students' online learning behaviors in the context of network teaching and constructs a portrait of their online learning behaviors. Firstly, based on existing research theories, students' online learning behaviors are summarized into four dimensions including seven indicators. Then, correlation analysis is used to analyze the relationships between the indicators and the total scores. Finally, cluster analysis is performed on the seven indicators of students' learning behaviors, dividing them into four categories, analyzing the characteristics of these categories, and exploring the characteristics and patterns of learning behaviors.

Keywords : online learning behavior; data mining; interactive behavior; cluster analysis

一、研究背景

随着现代信息技术在教育教学中的深度融合和广泛应用, 在线学习以其独特的优势成为高等院校学生学习的重要方式之一。在传统教学环境下, 教师以面对面的形式, 通过观察、测验、谈话等方式对学生的学习历程和态度进行分析, 发现问题并及时干预。然而在线学习环境下, 学生的在线学习过程与学习结果以数据的形式呈现给教师, 除了直观的分数的之外, 还需要教师从大量数据中挖掘学生学习行为的特点和规律, 发现问题并解决问题。因此, 在网络学习环境中, 融合学习历程研究方法和学习分析技术, 充分发挥智能技术和教师在问题发现、分析和解释方面的互补作用, 是一个重要的研究课题。

(一) 研究现状

关于在线学习行为影响因素、学习效果的研究成为热点, 很多研究者从教育学理论出发, 从学生学习投入度的角度, 以问卷的形式, 将质性研究与量化研究相结合^[1,2], 提出学习者的学习行为受到感知有用性和感知易用性的影响。也有一些研究者基于某门课程的在线学习数据, 利用统计分析方法, 构建在线学习行为模型等^[3-5]。目前, 基于网络学习空间行为数据分析, 对网络学习者相关的研究主要有以下三个方面。

1. 学习投入研究

朱红, 汪雅霜等^[6-8]从行为、情感和认知三个方面进行学习投入度量表的实证研究。张思等^[9-11]以教师网络学习空间为例, 提

出网络学习空间中学习者投入模型, 该研究中学习者的学习投入状况涉及行为、认知和情感等多方面要素, 该研究主要体现在结构化数据方面, 而如发帖内容等与主体的关联性、发帖内容的新颖性等内容由于目前研究者能力的局限没有考虑。

2. 学习交互分析

朱珂^[12]提出了网络学习空间中的学习者交互模型, 研究了交互性和沉浸感对学习持续使用意愿的影响。考虑了学习投入的操作交互、信息交互、概念交互三个维度, 建立学习交互模型, 以及相应的度量指标。该模型的概念交互采用人工编码的方式进行处理, 很大程度上制约了交互分析模型的应用范围。近年来, 也有个别研究者探讨在线学习环境下情绪交互作用对学习质量的影响, 为提高教师引导学生提高学习质量给出了参考策略^[13-14]。

3. 行为模式分析

Ferguson等^[15]对慕课中的学习者行为进行聚类分析, 并将其分为七种的主要学习行为参与模式。该研究表明, 不能简单地从一个学习环境中采取聚类方法并将其应用于另一个学习环境, 即使这些环境显然是相似的。另外, 短暂访问课程的学习者以及彻底完成课程的学习者这两个极端之间的模式很可能会最大程度的提高学习质量和学习环境。

基于以上在线学习的理论研究, 本文从独立学习行为、系统交互行为、资源交互行为、社会交互行为四个维度, 建立新的在线学习行为模型及相应的度量指标。为教师及时了解学生的学习行为动态和学习规律, 制定更好的教学和学习决策提供参考。

二、在线学习行为画像的构建

本文所有数据来源于2023年大学生服务外包创新创业大赛A14题：超星学习通部分线上学习数据。

（一）学习行为画像维度及指标

基于在线平台的运行数据和师生在线活动，从独立学习行为、系统交互行为、资源交互行为、社会交互行为四个维度内含的作业平均成绩、考试平均成绩、在线时长、活动参与度、任务完成度、讨论数和回答问题数七个指标，构建学生学习行为指标体系，挖掘在线学习环境下促进学习的作用机制和内在逻辑，刻画学生的在线学习行为结构模型，反映学生学习行为。

1. 独立学习行为

指学习者在学习过程中主要通过内化思考来进行学习，而与空间中的其它要素发生直接关系，如独立完成作业、独立完成考试等，在一定程度上反应了学习者对学习内容的掌握程度，学习动机的强烈程度和自我管理能力。

2. 系统交互行为

该指标是指学习者通过与学习系统进行交互来完成学习行为，体现为在线时长、登录时间间隔等，在一定程度上反映学生对在线学习的关注度和持续性，本文中以在线时长来研究系统交互行为。

3. 资源交互行为

该指标是指学习者通过与教师在学习平台上发布的学习资源进行交互来完成学习行为，比如视频观看、资源下载和浏览、问卷、文献读后感、小笔记等等，体现了学生的学习兴趣和创新型思维的培养。本文中，体现为任务完成度和活动参与度来。

4. 社会交互行为

该指标是指学习者通过与在线学习空间中的社交网络进行交互来完成学习，包括与教师、同伴进行各类交流活动，如讨论（发帖）、回答问题（回帖）等。

（二）多维度学习行为分析模型

采用相关性分析、聚类分析从不同角度和层面对数据进行处理，从而得到不同的分析结果。

1. 学习行为的相关性分析

本文采用 Pearson 相关系数对上述所有标签与综合成绩进行数据分析和建模，衡量指标与综合成绩的线性关系强弱，以此确定核心指标。经计算，与综合成绩存在显著相关的因素有考试提交率、作业成绩、活动参与度、考试平均成绩、在线时长，其中平时考试成绩、作业成绩、参与度的相关系数均超过0.4，存在强相关性，可作为核心指标；在线时长与参与度和考试提交率存在较强相关性，说明学生在线参与活动主要是参与教师发布的活动以及在线考试；讨论数与其它各种指标因素不存在显著的相关性或者相关性很低，因此认为讨论数不是影响综合成绩的核心指标。

2. 基于学习行为的 K-means 聚类分析

为了更好的对学习者的学习行为模式进行挖掘，以考试提交率、作业成绩、活动参与度、考试平均成绩、在线时长，任务完成度，讨论数为指标，采用 K-means 聚类分析的方法对网络学习行为数据进行分析，将学生在线学习行为按七个指标分为四种类型，分别定义为创新型、稳健型、参与型、且过型。对比发现，

不同类型的学习者，在学习行为方面有很大的区别。

创新型学习者和稳健型学习者自主学习行为得分比较高，知识掌握得比较好，能够独立地思考并完成学习任务。参与型学习者虽然在自主学习行为方面低于稳健型学习者，但是在交互行为方面却远超稳健型，与创新型相当。究其原因可能是这部分学生虽然学习效果不尽人意，但是对于学习活动仍然表现出较高的积极性，愿意主动地参与各种学习活动和交流，比较活跃。且过型学生各指标得分均为负分，学习者的学习成效比较差，这与学习者的个人兴趣、知识水平、个人自我管理能力有关。

三、结束语

与传统课堂教学相比，在线学习已成为今后教育发展的趋势。其教学过程更加复杂，学生的学习行为表现出四大特性——时空不受限、兴趣是动力、分数是主导、因此，这种在线学习行为分析会更加多样性和综合性。

本文通过对学生学习行为的分析，利用数据分析方法建立多维度的学习者画像，一方面可以直观清晰地了解学生的学习特点和规律，丰富线上学习资源库以满足不同类型学生的需求；另一方面，也为教师及时掌握学生的学习状态，有针对性的进行学习干预和指导提供参考。

本文的研究存在一定的局限性，一方面对于学习者学习行为的分类没有考虑到具体学科，另一方面，对于具体的参与活动以及学习资源的内容和数量等，并不了解，所以有些结论难免带有一定的主观性。

参考文献

- [1] 孙志梅, 张屹. 智慧课堂环境下独立学院学生课堂投入度及影响因素研究——以《大学生计算机基础》课程为例[J]. 中国教育信息化, 2020, (16): 11-15.
- [2] 钟园博. 在线学习环境下大学生学习投入度及影响因素研究[D]. 重庆师范大学, 2023.
- [3] Nistor N, Derntl M, Klammer R. Learning Analytics: Trends and issues of the empirical research of the years 2022-2014 [C]. Design for Teaching & Learning in a Networked World European Conference on Technology Enhanced Learning, 2015: 453-459.
- [4] 李曼丽, 徐舜平, 孙梦嫒. MOOC 学习者课程学习行为分析: 以电路原理课程为例[J]. 开放教育研究, 2015, 21(2): 63-69.
- [5] 陈圆圆, 刘盛峰, 董克. 基于数据挖掘的成人学习者在线学习行为与学习效果分析[J]. 安徽广播电视大学学报, 2019, (1): 38-42.
- [6] 朱红. 高校学生参与度及其成长的影响机制——十年首都大学生发展数据分析[J]. 清华大学教育研究, 2010, 31(06): 35-43+63.
- [7] 王纾. 研究型大学学生投入对学习收获的影响机制研究——基于2009年“中国大学生学情调查”的数据分析[J]. 清华大学教育研究, 2011, 32(04): 24-32.
- [8] 汪雅霜. 大学生学习投入度的实证研究——基于2012年“国家大学生学习情况调查”数据分析[J]. 中国高教研究, 2013, (01): 32-36.
- [9] 张思, 刘清堂, 雷诗捷, 等. 网络学习空间中学习者学习投入的研究——网络学习行为的大数据分析[J]. 中国电化教育, 2017, (4): 24-30, 40.
- [10] 王晶. 清华大学网络学堂学生使用情况统计与分析[J]. 现代教育技术, 2014, (7): 106-119.
- [11] 梁云真, 赵呈领等. 网络学习空间中交互行为的实证研究[J]. 中国电化教育, 2016, (7): 22-28.
- [12] 朱珂. 网络学习空间中学习者交互分析模型及应用研究[J]. 电化教育研究, 2017, (5): 43-48.
- [13] 俞建慧. 在线学习环境下基于交互视角的情绪分析及干预研究[D]. 浙江师范大学, 2022. DOI: 10.27464/d.cnki.gzsfu.2022.000154.
- [14] 刘君玲. 在线协作学习中情绪交互的过程机理与干预策略研究[D]. 陕西师范大学, 2020. DOI: 10.27292/d.cnki.gsxfu.2020.002665.
- [15] Ferguson R, Clow D. Examining engagement: Analysing learner subpopulations in massive open online courses (MOOCs) [A]. ACM, Proceedings of the Fifth International Conference on Learning Analytics And Knowledge [C]. ACM, 2015: 51-58.