

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

陈圆圆, 刘盛峰, 董 克, 陈彦彦, 宋 阳

(安徽广播电视大学 教育科学学院, 合肥 230022)

摘要:针对当前成人学习者在线学习行为研究不足的问题,依托安徽继续教育在线学习平台,以本科课程《大学英语 1》为例,利用 SPSS22.0 统计软件以及数据挖掘分析成人学习者的在线学习行为,探索在线学习行为与学习效果之间的相关性。研究结果表明,作业成绩、平时成绩、在线考试成绩、已学视频时长和学习进度因素与综合成绩呈强相关,而年龄、专业与综合成绩的关联度呈弱相关,性别与综合成绩无关联性。基于此,进行了教学反思,并提出意见和建议。

关键词:在线学习;学习效果;数据挖掘;成人学习者

中图分类号:G728

文献标志码:A

文章编号:1008-6021(2019)01-0038-05

一、引言

根据中国互联网络信息中心(CNNIC)公布的第 41 次《中国互联网络发展状况统计报告》,截至 2017 年 12 月,我国网民规模达 7.72 亿,普及率达到 55.8%,手机网民占比达 97.5%^[1]。随着现代教育技术日新月异地发展,在线学习不仅已然成为传统学习方式的有效补充,也是促使终身学习理念得以实现的最有效途径。在线学习在时间与空间上的自由度和灵活度为成人提供了公平受教育的机会,对构建学习型社会起着不可或缺的作用。在线学习方式是由多媒体网络学习资源、网上学习社区及网络技术平台构成的全新的学习环境,相对于其他的学习模式来说,具有个性化学习、丰富的交互性和协作性、自主式学习以及教学资源的共享性特征,可以让学生得到全新的学习体验。鉴于此,本文拟开展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方面的研究。

一些学者对于在线学习的基本理论、模型构建以及学习者的学习行为、学习效果等方面进行了系统的研究。如以学习投入理论为视角,探讨学习投入、学

习行为、学习效果之间的关联^[2]。利用逻辑回归的机器学习方法建立学习行为和学习效果之间的模型,检验了逻辑回归对于在线学习效果研究的价值^[3]。近些年大数据在各个教育领域得到广泛应用,对在线学习行为进行数据挖掘分析,可以更加客观地反映学习过程中存在的问题,预测学习者未来学习趋势并帮助研究者找到合理的对策,进而实现教师、学习者和学习资源的有效交互。傅钢善、王改花以“现代教育技术”网络学习系统学习的 2 801 名学习者研究对象,运用数据挖掘方法和统计学方法对在线学习行为进行定量分析,探讨网络学习行为特征与学习效果的关系^[4]。还有学者对学习行为及其所在情境中产生的数据进行记录和分析,构建数字化学习环境下个性化行为分析模型,基于学习行为分析模型对学习结果预测框架进行分类设计等^[5]。

我们注意到,由于在线学习过程的复杂性和随机性,研究者们得出的结论有所不同,对影响在线学习效果的因素也存在不同看法。学习者在线学习行为,如观看视频资源有效时长、文本阅读次数、作业成绩、

收稿日期:2018-10-08

基金项目:安徽省质量工程智慧课堂试点项目(项目编号:2017zhkt122);安徽省高等学校人文社会科学研究项目(项目编号:SK2018A0755)

作者简介:陈圆圆(1983—),女,安徽铜陵人,讲师,硕士。研究方向:英语文化与翻译、英语教育。

专业类别、年龄、论坛发帖数量等因素是否与最终的学习效果正相关? 为了探明这个问题,有必要通过对真实有效的在线学习行为数据进行分析,做可视化处理,探讨在线学习行为与学习效果的关联度,进而优化网络课程结构,提升在线学习满意度,达到理想的学习效果。

远程教育专家霍姆伯格认为,远程教育的主要目标群体是有工作、家庭和社会责任的成人。他们之所以选择远程教育,是为了解决工学矛盾、利用工余时间提升自己的专业技能或拓宽知识面^[6]。目前以大数据和在线学习为主题在中国知网上检索到 248 篇期刊论文,其中针对成人的在线学习研究占比很小,而且大部分成果集中在理论研究层面,基于网络学习平台的有针对性的数据采集分析研究相对较少,不利于挖掘在线学习行为的影响因素及其与学习效果之间的关联。安徽继续教育在线学习平台是由安徽省教育厅主导依托安徽广播电视大学搭建的安徽省统一的成人高等教育教学与管理平台,入驻园区的各普通高校共享共建课程网络教学资源,共同推进符合成人学习特点的远程化教学模式改革,构建适应现代远程教育发展的支持服务体系^[7]。经过几年的建设,基于安徽继续教育在线平台的教学模式已逐渐成形,并且不断发展成熟。在线平台教学资源的创建与应用,使学习行为不受时空限制,对实现高等学历继续教育教学效果最大化具有非常重要的现实意义。

二、研究设计与实施

(一) 研究方法

本研究采用量性研究和数据挖掘方法。量性研究即“通过数字资料来研究现象的因果关系,认为获得数字的研究可达到测量精确,能较客观地描述问题和现象,并用统计学方法分析资料和设对照组来避免研究中的偏差。”^[8]数据挖掘一般是指通过统计、在线处理分析、情报检索和模式识别等算法从大量数据中搜索隐藏于数据背后的信息的过程。以安徽继续教育在线学习平台为支撑,利用 SPSS22.0 统计软件数据挖掘成人学习者的在线学习记录,并对其进行量性分析,一方面可以发现在线学习行为规律与内在驱动力,另一方面可以反思课程教学设计的不足之处,探索优化在线学习效果的方法与路径。

(二) 研究实施过程

1. 研究框架

成人学习者在线学习行为特点与特征,主要包括学习进度、已学视频时长、作业成绩、考试成绩、专业、年龄等指标,本研究从以下三个方面进行分析:(1) 成人在线学习行为数据挖掘分析;(2) 不同性别、不同年龄、不同专业在线学习者学习过程中的差异性表现;(3) 在线学习行为与综合成绩之间的关联。

2. 研究对象

本研究的对象为 2018 春季参加《大学英语 1》网络课程学习的成人本科生。《大学英语 1》课程针对成人学习者的特点,基于安徽继续教育在线平台开展网上教学,通过设定模块安排、任务驱动、多元交互、立体评价的教学设计安排教学活动和在线考试,为在线网络学习课程。在线学习者共计 136 人,其中男生 51 人,女生 85 人;文科专业 109 人,理科专业 27 人。与全日制高校相比,本研究的对象是成人学生,存在年龄跨度大、在线学习时间不稳定等情况。

三、结果与分析

利用关联规则对所采集的数据进行挖掘,研究结果可为优化成人在线学习效果提供理论依据。下述内容将分别从成人在线学习行为数据挖掘分析、成人在线学习差异表现以及在线学习行为与学习效果之间的关联度三个方面进行研究。

(一) 成人在线学习行为数据挖掘分析

成人在线学习行为数据详见表 1。

表 1 成人在线学习行为数据挖掘分析(N=136)

| 学习行为 | 均值 | 标准差 | 极小值 | 极大值 |
|--------|----------|----------|-----|-------|
| 学习进度 | 0.923 4 | 0.232 2 | 0 | 1 |
| 已学视频时长 | 20.720 0 | 5.938 0 | 0 | 22.83 |
| 在线作业成绩 | 15.930 0 | 6.733 0 | 0 | 24 |
| 平时成绩 | 49.050 7 | 14.030 4 | 0 | 60 |
| 在线考试成绩 | 25.535 3 | 11.779 3 | 0 | 40 |

1. 学习进度

学习进度范围在 0~100%之间,均值为 92.34%,该课程要求学习者学习进度至少达到 90%。数据显示,4 名学生学习进度是 0,8 名学生学习进度有数据,但是不足 90%,这 12 名学生考试成绩均不及格,同时,也是造成及格率(85.3%)不够理想的一个因素。

2. 已学视频时长

学习进度范围在 0~22.83 小时之间,均值为 20.72 小时。根据教学文件,学习者完成视频资源

90%的进度即可达到应学视频资源的时长要求,即 20.55 小时。从均值 20.72 小时来看,有 18 人未达到应学视频时长要求,其中 13 人不及格,包括 4 位网上学习活动为 0 的学习者,该课程共有 20 人未通过考试,可见学习视频资源与最终获取该课程的学分有直接关联。

3. 在线作业成绩

在线作业成绩范围在 0~24 分之间,均值为 15.93 分。在线作业成绩未达平均值的有 22 人,10.3 分以下 13 人,其中 11 人作业成绩为 0,10 人考试不及格,可见在线作业是与最终能否通过课程考试有直接关系的。同时,值得教师进一步思考和了解的是,这些学习者没完成在线作业的原因是什么?是不熟悉教学文件的要求,还是未得到提醒,抑或没有能力完成?了解之后可以有针对性地加以改进。

4. 平时成绩

平时成绩范围在 0~60 分之间,均值为 49.05 分。平时成绩包括在线作业成绩和网上学习行为成绩两个部分,前者占 60%,后者占 40%。四次作业的权重均为 25%,总分 24 分;网上学习行为权重为 40%,考核分值按学习任务完成的难易程度进行分配,总分为 36 分。网上学习行为成绩包括视频资源的学习时长和文本辅导的浏览学习。有 30 人未达到平均值,其中 13 人考试不及格。经分析,主要症结在于在线作业成绩不理想。

5. 在线考试成绩

在线考试成绩范围在 0~40 分之间,均值为 25.54 分。38 人未达平均值,其中 21 人未参加在线考试,18 人考试不及格。经匹配发现有 8 人既没有完成在线作业也没有参加在线考试,这也是造成整个班级及格率不理想的一个重要因素。课程责任教师应该追踪到底是什么因素导致这样的无效学习。

6. 综合成绩

综合成绩是由平时成绩(60%)和在线考试成绩(40%)两项成绩合成。学习者的综合成绩平均分为 74.67 分,不及格人数为 20 人,60 分至 69 分为 12 人,70 分至 79 分为 21 人,80 分至 89 分为 57 人,90 分以上 26 人,及格率达 85.2%(见图 1)。大部分学习者考试成绩较为理想,也反映出他们学习积极性较高。

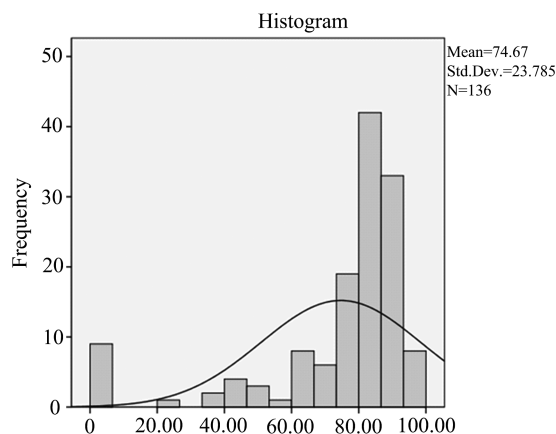


图 1 综合成绩分布图

(二) 成人在线学习差异性表现

《大学英语 1》在线学习人数共计 136 人,他们的性别差异、年龄差异和专业差异是否与在线学习效果存在关联度,或者说从学习者的性别、年龄以及专业数据背后可以挖掘出哪些与在线学习相关的信息,这都是值得思考的问题。

1. 性别差异

在线学习男生 51 人,女生 85 人,男女生比例(男:37.5%,女:62.5%)并不悬殊,具有一定的代表性。男女生的平均分分别为 74.40 分和 74.84 分,几近无差别,足以说明在线学习英语时性别差异影响因素不显著,男女在线学习能力和资质相当。

2. 年龄差异

表 2 学生年龄分析

| | 均值 | 标准差 | 最小值 | 最大值 |
|----|-------|------|-----|-----|
| 年龄 | 28.49 | 6.71 | 20 | 48 |

学习者年龄差异见表 2,年龄分布较为集中,大部分学生的专科文凭获取时间大约在 2010 年前后。经知网文献筛查,与在线学习研究相关的论文从 2014 年(354 篇)起开始激增,2017 年(614 篇)达到高峰,从一个侧面反映出直到 2010 年前后在线学习和在线学历继续教育模式尚不普及。学习者承担着工作和家庭的双重责任,同时希冀空闲时间提升个人学历和知识,因此在有选择的情况下绝大多数人目的明确,会认真把握学习机会。

3. 专业差异

表 3 专业差异

| 专业类别 | 人数 | 平均分 | 标准差 |
|------|-----|----------|-----------|
| 文科 | 109 | 77.161 5 | 21.236 66 |
| 理科 | 27 | 64.614 8 | 30.547 31 |

专业差异见表 3。本课程学习者来自 7 个专业, 总体可以分文理两类, 文理学生比例差异明显。数据分析表明: 文科生比理科生平均分高出 12.55 分, 就标准差数值来看, 理科生与文科生的标准差相差较大, 是各组数据分析中标准差最为显著的一组。可见, 专业差异对在线学习效果的影响呈正相关。

(三) 在线学习行为数据与综合成绩之间的关联度

数据挖掘分析在线学习行为的各方面因素, 旨在

表 4 在线学习特征与综合成绩之间的关联度分析 (* $P<0.05$, ** $P<0.01$)

| | 学习进度 | 作业成绩 | 平时成绩 | 在线考试成绩 | 学视频时长 | 学习行为 | 专业 | 年龄 |
|------|---------|---------|---------|---------|---------|---------|---------|---------|
| 综合成绩 | 0.871** | 0.829** | 0.939** | 0.908** | 0.869** | 0.872** | 0.211** | 0.056** |

四、结论与讨论

(一) 结论

(1) 成人在线学习者的学习行为包括学习进度、已学视频时长、在线作业成绩、平时成绩以及在线考试成绩等学习痕迹构成。(2) 作业成绩、平时成绩、在线考试成绩、已学视频时长和学习进度因素与综合成绩呈强相关; 年龄、专业与综合成绩的关联度呈弱相关, 其中文科专业和理科专业学生的学习效果差异性明显; 性别与综合成绩无关联性。

(二) 讨论

远程教育专家 Moore 提出, 远程教育的交互包括学习者与学习资源交互, 师生交互, 以及生生交互^[9]。学习者与学习资源交互具有单向性特点, 对学习者的自觉性和自主性要求较高。师生交互是最有价值的交互行为, 能有效产生教学临场感。教师是学习资源的设计者, 他们知道如何督促学习者学习, 提高学习者的兴趣以及通过测评关注学习效果。生生交互指的是学习者之间在线互相交流和协作。

研究表明, 学习者与学习资源之间的交互表现较满意。因为本课程是纯网络教学模式, 对在线学习周期、学习总时长、作业完成次数均有最低限制, 所以大多数学习者能按教学文件要求完成学习进度与任务。通过问卷调查与电话回访发现, 93% 的学习者对课程的随学随考的考核方式表示满意; 93% 的学习者认为该课程的课程资源丰富, 能满足在线学习需求; 87% 的学习者认为通过该课程学习有所收获; 94% 的学习者反馈平台的在线学习功能可以满足学习的需求。但仍有少部分学习者(17.6%)未能按时完成学习, 其原因可能有: 第一、学习意识淡薄, 学习动力不足。在

探索在线学习特征与学习效果之间的关联度, 通过对成人学习者学习行为的相关分析, 得出以下结果(见表 4)。结果表明: 在线学习效果与各类分析指标, 如作业成绩、卷面成绩、学习行为、已学视频时长、年龄以及专业等均呈正相关。作业成绩、平时成绩、在线考试成绩、以及已学视频时长等均呈强相关, 这也是因为这些指标在教学文件里有明确的硬性规定, 与学习者的最终成绩直接挂钩。而年龄、专业与综合成绩的关联度呈弱相关, 性别与最终得分无关联性。

线学习对学习者的自觉性要求较高, 如果缺乏自律, 极有可能半途而废; 第二、学习时间不充裕。大部分学习者都存在工学矛盾, 如何兼顾家庭、工作和学习对很多人来说确实是个难题。第三、自学能力较弱, 单向学习资源的能力不强, 导致学习受挫、容易放弃。

研究发现, 在线学习在师生交互和生生交互方面表现不尽如人意。在线学习缺少面授教学的情感交流效果, 教学临场感的缺失容易造成学习者的孤独感与无助。由于师生交互和生生交互不做考核要求, 学习者在论坛的交流痕迹明显稀少, 但是据统计, 积极参与互动交流的学习者成绩均在 90 分以上; 遇到问题(如视频无法观看)没有及时解决的学习者最终由于未完成设定的学习指标而导致成绩不及格, 可见师生交互的重要性。教师并非是学习资源的搬运工, 而应该是学习资源的生产者、引导者、监督者与协助者, 应在学习者的全程学习过程中发挥积极作用。

基于以上数据分析, 对于提高该课程的质量、在线学习效果以及学习效率提出以下几点改进建议。(1) 优化学习资源。进一步完善课程资源的设计, 适度降低学习者的学习难度; 提高试题库与教学视频匹配度, 有针对性地录制期末复习指导视频; 提高播放视频资源的流畅度, 适当切割长视频, 满足学生碎片化时间学习要求, 方便学习者手机客户端观看。(2) 提高师生交互。适当安排网上集中答疑和若干次面授教学, 为学习者提供提问、解惑的窗口, 增强同步交互性; 尝试混合式教学模式, 开展《大学英语 1》智慧课堂试点教学, 强化教学临场感。(3) 跟踪学习者学情。及时跟踪学习者的学习进度, 发挥各班级辅导员的促学作用, 鼓励学习者积极学习; 并针对理科专业

群体做问卷调查,了解他们的学习困难和学习需求,提高理科生的学习成绩。

综上,利用数据挖掘研究与在线学习效果相关的影响因素有着积极意义,有助于促进课程改革,优

化教学设计。通过分析和总结在线学习行为,可以为下一步预测学习者在线学习效果提供充分依据,进而深入研究并检验数据挖掘的结果,为今后的在线学习深入研究做铺垫。

参考文献:

- [1] 中国互联网信息中心.第 41 次中国互联网络发展状况统计报告[EB/OL].(2018-01-31)[2018-09-10].http://www.cac.gov.cn/2018-01/31/c_1122346138.htm
- [2] 马志强,苏珊,张彤彤.基于学习投入理论的网络学习行为模型研究:以“网络教学平台设计与开发”课程为例[J].现代教育技术,2017,27(1):74.
- [3] 宗阳,孙洪涛,张亨国,等.MOOCs 学习行为与学习效果的逻辑回归分析[J].中国远程教育,2016(5):14.
- [4] 傅钢善,王改花.基于数据挖掘的网络学习行为与学习效果研究[J].网络教育与远程教育,2014(9):53.
- [5] 武法提,牟智佳.基于学习者个性行为分析的学习结果预测框架设计研究[J].中国电化教育,2016(1):41.
- [6] 龚志武.关于成人学生网上学习行为影响因素的实证研究[J].中国电化教育,2004(8):32.
- [7] 钱丰收,董克.继续教育在线课程资源评价模型研究[J].蚌埠学院学报,2017(1):146.
- [8] 胡雁.定性研究[M].北京:人民卫生出版社,2002:42.
- [9] MOORE M.Three Types of Interaction[J].The American Journal of Distance Education, 1989(2):1-6.
- [10] 刘少伟,赵长在,雷少飞.基于数据分析的学习者在线学习行为研究[J].中国成人教育,2016(23):11.
- [11] 彭文辉,曾东薇.近十年我国网络学习行为期刊论文的内容分析研究[J].中国远程教育,2015(1):42-48.

Analysis of Online Learning Behavior and Learning Effect of Adult Learners Based on Data Mining

CHEN Yuanyuan, LIU Shengfeng, DONG Ke, CHEN Yanyan, SONG Yang

(School of Education and Science, Anhui Radio and TV University, Hefei 230022, China)

Abstract: Relying on the Anhui continuing education online learning platform, taking the undergraduate course “College English 1” as an example, this paper tries to analyze the online and learning behavior of adult learners and explore the correlation between online learning behavior and learning effect with the help of SPSS22.0 statistical software and data mining. The results show that the results of the online homework, the usual grades, the online test scores, the duration of the video watched and the progress of the online learning are strongly correlated with the comprehensive scores, while the correlation between age, major and comprehensive scores is weak. Besides, gender and comprehensive scores are not related. Based on the above analysis, the inadequacies of the curriculum design are reflected, and suggestions are put forward to provide a basis for a better understanding of the online learning effect of adult learners.

Keywords: online learning; learning effect; data mining; adult learners

[责任编辑 李潜生]