

# 人类行为动力学视角的成人学习者 在线学习质量影响因素研究

董克<sup>1</sup>, 贾礼远<sup>2</sup>, 姚梦月<sup>1</sup>

(1. 蚌埠学院 经济与管理学院, 安徽 蚌埠 233032; 2. 安徽开放大学 信息与建筑工程学院, 安徽 合肥 230022)

**摘要:**随着社会竞争的日益激烈,终身学习日趋常态化,成人学习者的在线学习规律和质量越来越受到研究者的关注。通过采集成人学习者的在线学习行为数据,并对此进行定量分析:首先,采用人类行为动力学研究方法,实证研究了成人学习者的在线学习时间间隔服从幂律分布的特征,且在线学习行为具有阵发和胖尾的特征;其次,应用多元回归分析方法,找出影响成人学习者在线学习的内在因素;最后,利用聚类分析算法,对成人学习者样本群体进一步进行细分,探索细分状态下影响不同学习者群体学习的动因。研究结果可为教师进行在线教学设计以及充分挖掘学习者的学习兴趣提供理论指导。

**关键词:**成人学习者;学习行为;人类行为动力学;K-means 算法;多元回归模型

**中图分类号:**G728

**文献标识码:**A

**文章编号:**2097-0625(2024)04-0058-06

## 一、引言

随着互联网和人工智能技术的快速崛起,以及社会竞争环境的不断加剧,人们对知识的追求与渴望愈发强烈。如今,人们获取知识的途径更加广泛,大量线上平台崛起,如网易公开课、中国大学慕课以及 B 站等,为人们的日常学习提供了方便。通过互联网进行线上学习从而获得知识与学历,已成为一种更为灵活、高效的学习方式。据不完全统计,截至 2024 年 3 月,中国上线慕课数量超过 7.68 万门,注册用户 4.54 亿,学习人数达 12.77 亿人次,在校生获得慕课学分认定 4.15 亿人次<sup>[1]</sup>。这种通过互联网进行学习的趋势不仅改变了人们的学习方式,更逐渐演变成一种风尚,使网络学习成为提升自我、增强个人竞争力的重要途径。

成人学习者作为社会进步与发展的重要推动力,其在线学习的需求与问题日益受到关注。截至 2024 年 6 月,国家开放大学累计招收开放教育高等学历学

生 2 572 万人,占全国受过高等教育人数的 10.3%;累计毕业生 1 961 万人,占恢复高考以来我国大学毕业生总量的 8.5%。目前,学校在籍生 456.7 万人,接近高等教育在学总规模的十分之一。然而,针对成人学习者在线学习特点与策略的研究尚显不足,如何有效提升成人学习者的在线学习质量,已成为学术界与实践界共同关注的焦点。因此,深入探索成人在线学习的规律与策略,对于促进成人教育的发展、提升社会整体知识水平具有重要战略意义。

## 二、文献综述

2005 年,Barabás 通过实证研究名人通信信件,揭示了人类活动的时间间隔遵循幂律分布,这一发现对传统的 Poisson 过程描述人类活动的观念构成了挑战,从而激发了对人类行为动力学的研究热潮。时至今日,学术界已对人类行为动力学展开了深入的探索<sup>[2-3]</sup>,并得出了若干重要结论。韩筱璞等基于人类兴趣可受感知影响而改变的现象,构建了一个基于兴

**收稿日期:**2024-07-30

**基金项目:**全国教育科学“十四五”规划教育部重点项目“成人学习者网络学习行为数据挖掘与网络教学质量监控机制构建”(项目编号:DKA220345);安徽省高校自然科学基金项目“基于学习者画像的在线学业监控与预警机制研究”(项目编号:KJ2020A1217)

**作者简介:**董克(1982—),男,安徽蚌埠人,教授,博士。研究方向:教育管理。

趣变迁的人类行为动力学模型,并通过数值模拟进行了验证<sup>[4]</sup>。曹盼盼等则聚焦于写信间隔时间的分析,进一步强调了幂律分布在人类社会行为中的普遍适用性<sup>[5]</sup>。郭进利针对博客评论的时间间隔分布进行了实证研究,其结果同样支持幂律分布的结论<sup>[6]</sup>。易明等则从多个维度研究了在校学生的图书借阅行为,揭示了其中显著不同的人类动力学特征<sup>[7]</sup>。樊超等则深入剖析了 MOOC 平台用户的数据,发现用户在线学习的时间间隔与持续时间均呈现出幂律分布的特点<sup>[8]</sup>。李月等则运用人类行为动力学模型,深入挖掘了在线学习行为时间间隔的规律性<sup>[9]</sup>。

聚类算法作为一种无监督学习算法,其核心分类主要包括基于划分的聚类、基于层次的聚类、基于密度的聚类以及基于网格的聚类方法。此类算法以其原理简明、易于理解与实现等特性,在数据挖掘领域得到了广泛应用<sup>[10]</sup>。彭涛等的研究中,研究人员创新性地运用了双角度聚类建模方法,对百万余名学习者的学习行为数据进行了深入剖析,从而全面且深刻地描绘出学习者的行为特征轮廓<sup>[11]</sup>。丁泽宇等则聚焦于 K-means 算法,通过该算法对护理本科生网络学习行为特征进行细致分析,进一步探讨网络学习行为与学习效果之间的潜在联系<sup>[12]</sup>。而刘笑言等则采用了聚类分析技术,对相关数据进行有效划分,旨在探索多种机器学习算法对学生学习绩效产生的具体影响因素<sup>[13]</sup>。

当前,关于学习行为分析的研究,主要聚焦两大领域:一是学习行为分析的方法与技术探索。这一领域涵盖了多种分析方法,包括但不限于教育数据挖掘、统计分析法、社会网络分析法、话语分析法以及内容分析法等<sup>[14]</sup>。研究者们基于不同视角,对在线学习行为数据进行深入研究<sup>[15-17]</sup>。此外,部分学者还借助数学模型,如聚类方法<sup>[18]</sup>、回归方法<sup>[19]</sup>及滞后序列分析法<sup>[20]</sup>,对行为数据进行精准剖析,成效显著。二是学习行为分析的应用实践。例如,魏顺平等站在辅导教师的立场,依托在线学习平台后台数据,深入分析了学习行为特征<sup>[21]</sup>;而姜强等则通过运用学习分析仪表盘法,对学习行为进行了系统研究,并成功将其应用于自适应学习系统之中<sup>[22]</sup>。

在详尽梳理现有文献资料的基础上,可以发现当前关于学习行为的研究领域显著偏重于在校学生,而对成人学习者的相关研究较少。鉴于这一现状,为了更深刻地揭示成人学习者在线学习行为与其学习效

果之间的内在联系与规律,本研究采用人类行为动力学的理论视角,对成人学习者的在线学习行为数据进行全面而深入地挖掘分析。与既有研究相比,本研究的独特之处在于专注成人学习者群体,并综合运用了多元化的数据分析方法,以实现数据的深度挖掘,进而探索并阐明影响成人学习者在线学习效果的深层次动因。这一研究成果不仅为教育工作者在制定科学合理的在线教学方案时提供有力的理论支撑,同时也为成人学习者在自主开展在线学习过程中提供了有价值的参考和指导。

### 三、研究对象

本研究基于国家开放大学在线学习平台,针对某门课程筛选出 425 名成人学习者的在线学习行为数据作为研究样本。为确保数据的有效性,剔除了在线行为数据为 0 的 101 名学习者,最终样本数量为 324 名,数据有效率为 76.2%。研究的时间跨度为 2024 年 3 月 1 日至 7 月 15 日,共计 136 天,涵盖了成人学习者在此期间的完整学习周期。本研究全面收集了该周期内成人学习者的在线行为及活动数据,具体包括以下四个维度:(1)课程资源利用情况。包括线上课程资源的完成度;(2)学习行为特征。涵盖课程行为时长、访问次数、访问时长、行为总数以及浏览数等关键指标;(3)作业完成情况。涉及提交的形成性考核作业数、已评阅的形成性考核作业数;(4)师生互动情况。详细记录了学生的学习参与度,如学生发帖量、发帖字数、回帖数以及教师的回帖数,从而全面反映学习过程中的交互情况。

### 四、研究结果

#### (一)基于人类行为动力学的在线学习行为规律

图 1 展示了成人学习者群体在指定学习周期内,其在线学习行为次数的分布情况。鉴于各学习中心的数据导入平台时间存在差异,导致学期初学习者的在线学习行为次数相对较低。随后,这一数据逐渐上升,在 6 月 20 日这一时间节点达到了峰值,此现象可归因于教师精心规划的学习活动安排,包括考前对学生进行的集中线上辅导以及历年期末试卷的上传,有效激发了学习者的参与热情。随着时间的推移,学习者的兴趣与活跃度呈现出下降趋势。综上,成人学习者群体的学习行为展现出明显的周期性与阵发性特点,这种现象可能与教师的线上教学活动策划及学习者个人学习兴趣的波动密切相关。

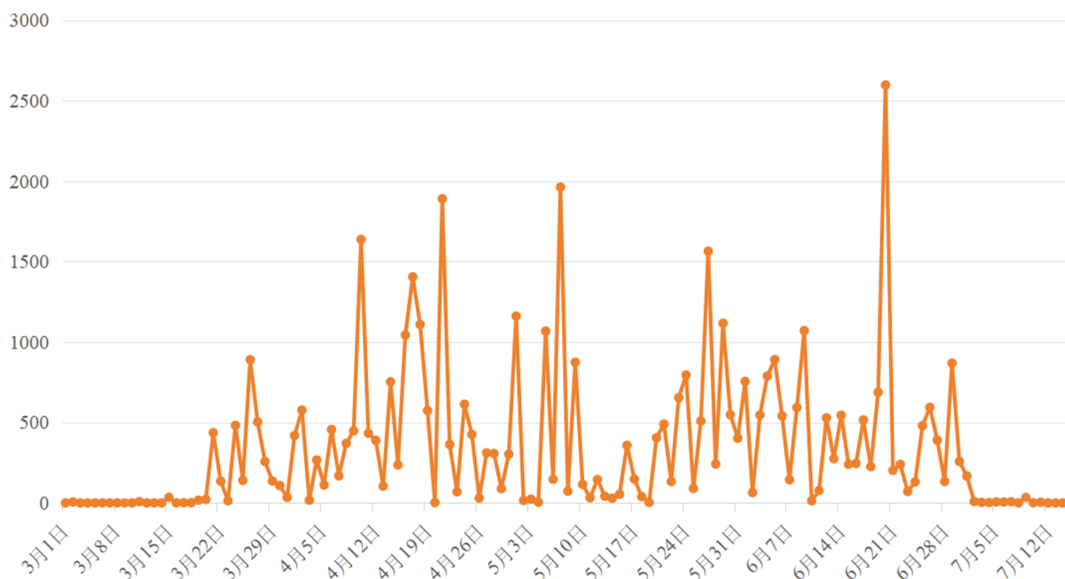


图 1 成人学习者群体在线学习行为次数分布

图 2 展示了在双对数坐标体系下,针对成人学习者群体所进行的一项在线学习行为时间间隔分布的统计结果。在排除图 2 初始的下垂头部以及尾部的异常扩展部分后,该学习行为间隔时间的数据分布呈现出接近直线的形态,这一现象显著地揭示了该群体学习行为间隔时间的幂律分布特性,并伴随有胖尾现象的存在。

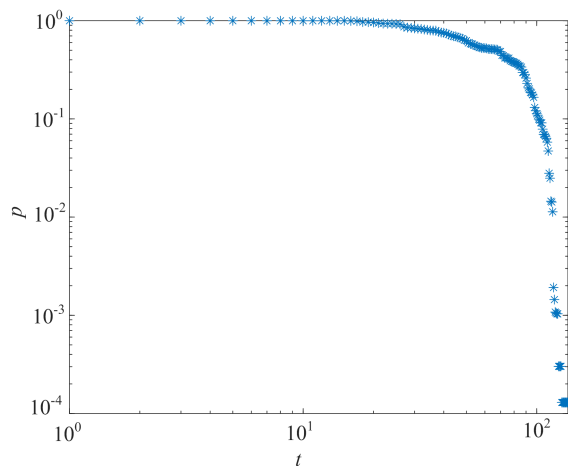


图 2 双对数坐标下成人学习者群体的学习行为时间间隔分布

对图 2 中的数据点进行直线拟合,可得拟合曲线的函数表达式为:

$$p = 3.3670t^{-0.4458} \quad (1)$$

其中,拟合优度的值为 0.871 5,这一数值反映了拟合曲线与原始数据之间的接近程度。为了更直观地展示这一拟合效果,在双对数坐标下绘制成人学习者群

体的学习行为时间间隔分布图,并叠加了拟合曲线,如图 3 所示。

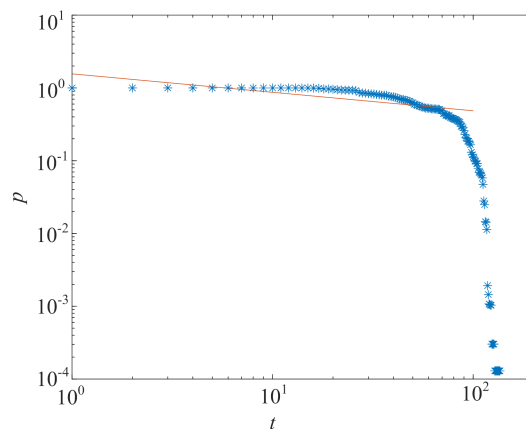


图 3 成人学习者群体的学习行为时间间隔分布与拟合曲线

研究显示,成人学习者的在线学习行为与普通高校学生的学习行为在统计规律上展现出相似性,二者均遵循幂律分布模式<sup>[7-8]</sup>,并具备周期性、阵发性等共性特征。在时间间隔分布上,两者均显现出显著的“胖尾”效应。进一步分析发现,学习兴趣的减弱与学习行为发生概率之间存在负相关关系,而教学活动的介入则与学习行为发生概率之间形成正相关关系。为深入探究直接作用于成人学习者的学习质量的因素,本文计划采用多元线性回归模型对数据进行深度挖掘,以期识别并揭示影响成人学习者学习质量的关键因素。



## (二)多元线性回归模型

为建立多元线性回归模型,将  $n$  个自变量  $X_1, X_2, \dots, X_n$  与因变量  $Y$  进行关联,具体模型形式如下:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon, \quad i = 1, 2, \dots, n \quad (2)$$

分别对自变量进行  $F$  检验,找出  $F$  的最小值,即:  $F_{\min}^1 = \min\{F_1^1, F_2^1, \dots, F_n^1\}$ 。若  $F_{\min}^1 < F_0$ , 其中  $F_0$  为临界值,则删除对应的自变量,直至删除所有对因变量  $Y$  无显著影响的自变量。

基于采集的成人学习者线上学习相关数据,我们首先将相关指标划分为以下四类:(1)课程资源,线上课程资源完成度  $X_1$ ;(2)学习行为,课程行为时长  $X_2$ ,访问次数  $X_3$ ,访问时长  $X_4$ ,行为总数  $X_5$ ,浏览数  $X_6$ ;(3)作业情况,提交形考数  $X_7$ ,评阅形考数  $X_8$ ;(4)师生互动,学生发帖数  $X_9$ ,发帖字数  $X_{10}$ ,回帖数  $X_{11}$ ,教师回帖  $X_{12}$ 。将上述分类的指标代入到之前建立的多元线性回归模型中,经过逐步回归分析,我们最终得到以下模型结果:

$$Y = 85.734 + 0.117X_1 - 0.079X_5 + 2.390X_7 + 0.896X_{11} + \epsilon \quad (3)$$

经数据分析,显示线上课程资源完成度  $X_1$  与最终形考成绩之间存在明确的正向关联,表明完成度越高,形考成绩越优。与此同时,行为总数  $X_5$  则对最终形考成绩展现出负向影响,这一发现需引起关注。此外,提交形考作业数  $X_7$  与最终形考成绩之间存在显著的正向关系,意味着作业提交数量的增加对提升形考成绩具有积极作用。另一显著正向影响因素为回帖数  $X_{11}$ ,其增加亦能显著提升最终形考成绩。综上所述,线上课程资源完成度  $X_1$ 、提交形考作业数  $X_7$  以及回帖数  $X_{11}$  均为影响最终形考成绩的关键因素,且均呈现出显著的正向效应。

相较于普通高校的学生,成人学习者具备多元化的职业背景,其教育目标亦呈现多样化特点。他们更加重视学习的实用性与应用性,旨在不仅获得学历证书,更期望汲取能直接应用于职场实践的知识与技能。为此,我们通过运用聚类分析方法,旨在精准描绘出不同学生群体的分布轮廓,并深入剖析其行为特征。这一系列举措,旨在为构建更为个性化和精细化的成人教育体系奠定坚实基础。具体地,我们采用聚类技术,对成人学习者群体实施了进一步的细分处

理,旨在挖掘并识别那些潜藏于群体内部、对成人学习效果产生深远影响的内在因素,从而为优化教学策略、提升学习质量提供有力支持。

## (三)K-means 聚类算法

K-means 聚类算法是由 J. B. MacQueen 于 1967 年提出,是数据挖掘中无监督学习的一种关键算法。其核心原理在于从数据集中选取  $K$  个簇中心  $m_1, m_2, \dots, m_k$ ,随后通过使用迭代计算的方式,不断优化簇的划分,以确保同一个簇内的样本数据  $x_i$  的相似度最大化,而不同簇的样本数据则保持较高的区分度。具体的算法执行流程如下:

Step1:从数据集  $W$  中任意选择  $k$  个对象  $m_1, m_2, \dots, m_k$  作为初始聚类中心对象;

Step2:根据每个聚类对象的均值(中心对象),计算每个对象与这些中心对象的距离;并根据最小距离重新对相应对象进行划分  $w_i, i = 1, 2, \dots, k$ ;

Step3:重新计算每个有变化聚类的均值(中心对象);

Step4:循环执行 Step2 和 Step3,直到每个聚类不再发生变化为止。这种划分使得式  $E = \sum_{j=1}^k \sum_{x_i \in w_j} \|x_i - m_j\|^2$  的值最小。

应用 K-means 聚类算法,对成人学习者学习数据进行聚类分析。其中聚类个数  $k$  的值为 3,将成人学习者群体的数据分为三个组,每组分别为 255、51、18 个学习者个体(见图 4)。

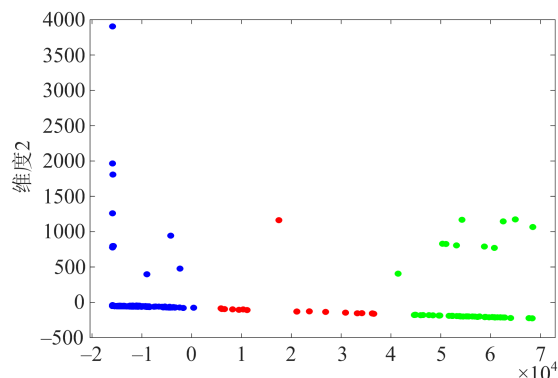


图 4 成人学习者在线学习行为聚类

根据表 1 的数据分析,课程行为时长与最终形考成绩之间并未展现出直接的必然关联。然而,值得注意的是,簇 2 与簇 3 的访问次数及行为总数的统计值均显著高于簇 1。此外,研究还发现,那些积极参与

发帖与回帖的学习者,其最终形考成绩也普遍较高。进一步地,教师的积极回帖行为被证实能够显著影响成人学习者的学习行为模式,进而间接地对其最终形考成绩产生积极影响。

表 1 在线教学维度及最终形考成绩均值

指标	簇 1	簇 2	簇 3
课程行为时长	73.98	36.71	56.00
访问次数	78.00	377.29	366.06
行为总数	143.44	455.14	330.44
发帖数	0.50	2.10	1.06
发帖字数	56.77	215.67	71.67
回帖数	0.98	2.24	2.56
教师回帖数	0.21	0.86	0.33
课程完成度	39.68	97.12	95.93
最终形考成绩	90.98	94.93	96.69

五、结论与讨论

(一) 结论

研究结果显示,成人学习者的学习行为展现出周期性与阵发性等显著特征,且其学习质量、学习积极性及教师回帖行为之间存在着密切的内在联系。

(二) 讨论

针对成人学习者在线学习行为数据的深入分析,进行了以下四点反思:

1. 在线学习过程中,课程教师与辅导老师对成人学习者的引导至关重要,其重要性应得到充分的认识与重视。数据统计结果表明,有 101 名成人学习者,即接近四分之一的学员,在一个完整的学习周期内,未能登录学习平台参与学习活动。因此,对于课程教师及辅导老师而言,如何有效引导成人学习者及时登

录学习平台完成学习任务,成为一个亟待解决且意义重大的问题,其重要性不容忽视。

2. 与普通高校相比,针对成人学习者的课程教师在教学重点上需相应调整。鉴于成人学习者的学习目标明确,主要聚焦于学历提升及职业技能增强,课程教师在设计线上课程时需清晰展现成人教育与普通高等教育的差异。具体而言,教师应通过增加与工作实践及行业前沿紧密相关的知识模块,以激发成人学习者的学习兴趣,并确保教学内容紧密贴合其实际需求。

3. 大多数成人学习者在形成性考核作业中展现出了良好的完成能力。然而,值得注意的是,仍有部分学习者在最终形考成绩中未达到及格线。针对此部分学习者,其成绩不理想的原因可能涵盖年龄因素、在线学习操作技能不熟练,以及独立完成线上学习任务所面临的挑战等多方面。因此,如何有效提升这部分学习者的学业成绩,已成为亟待解决的重要问题。

4. 对成人学习者群体进行深入细分,旨在根据各自特性设计差异化的线上教学方案,并定制个性化的学习路径,以更有效地促进学习质量。经过分析,发现不同学习者群体在访问频率与行为总量上对其最终形考成绩的影响存在显著差异。基于此,提出针对特定群体定制专属线上教学方案的策略。如设计一系列时长为 5 分钟左右的微型课程,聚焦于核心知识点,便于学习者灵活学习。通过实施此方案,可以持续监测学习者的学习效果与形考成绩之间的关系。同时,观察到积极参与发帖与回帖活动的学习者往往获得更高的形考成绩,可以鼓励教师主动融入讨论,设计吸引学习者的发帖主题,以此激发学习兴趣,增强学习动力,最终实现更好的学习效果。

参考文献:

[1] 王繁,刘永强,周天华. 人工智能引领高等教育数字化创新发展[J]. 中国高等教育,2024(Z1):9-12.  
[2] 郭进利. 复杂网络和人类行为动力学演化模型[M]. 北京:科学出版社,2013.  
[3] 周涛,韩筱璞,闫小勇,等. 人类行为时空特性的统计力学[J]. 电子科技大学学报,2013,42(4):481-540.  
[4] 韩筱璞,周涛,汪秉宏. 基于自适应调节的人类动力学模型[J]. 复杂系统与复杂性科学,2007,4(4):1-5.  
[5] 曹盼盼,阎春宁. 人类通信模式的幂律分布和 Zipf 定律[J]. 复杂系统与复杂性科学,2009,6(4):51-56.  
[6] 郭进利. 博客评论的人类行为动力学实证研究和建模[J]. 计算机应用研究,2011,28(4):1422-1424.  
[7] 易明,张展豪,李怡. 大学生图书借阅行为的人类动力学分析[J]. 图书馆学研究,2019(22):83-93.

- [8] 樊超,宗利永. MOOC 在线学习行为的人类动力学分析[J]. 开放教育研究,2016,22(02):53-58.
- [9] 李月,姜强,方慧,等. 基于人类动力学的在线学习行为规律分析及建模研究[J]. 电化教育研究,2023,44(1):77-83.
- [10] 章永来,周耀鉴. 聚类算法综述[J]. 计算机应用,2019,39(7):1869-1882.
- [11] 彭涛,单志龙. 基于双路聚类的在线学习行为分析研究[J]. 华南师范大学学报(自然科学版),2021,53(6):122-128.
- [12] 丁泽宇,弓伟,冀璇,等. 护理本科生网络学习行为特征聚类及可视化研究[J]. 中华护理教育,2023,20(7):773-779.
- [13] 刘笑言,刘创,周建军,等. 基于机器学习的高职生学习绩效影响因素研究[J]. 中国职业技术教育,2024(14):54-63.
- [14] 顾小清,刘妍,胡艺龄. 学习分析技术应用:寻求数据支持的学习改进方案[J]. 开放教育研究,2016(5):34-45.
- [15] FERGUSON R. Learning Analytics: Drivers, Developments and Challenges[J]. International Journal of Technology Enhanced Learning, 2012, 4(5):304-317.
- [16] FERGUSON R, SHUM S B. Learning Analytics to Identify Exploratory Dialogue Within Synchronous Text Chat[A]// Proceedings of the 1st International Conference on Learning Analytics and Knowledge[C]. ACM, 2011:99-103.
- [17] 陈圆圆,刘盛峰,董克,等. 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):9-14.
- [18] PARK Y, YU J H, JO I H. Clustering Blended Learning Courses by Online Behavior Data: A Case Study in a Korean Higher Education Institute[J]. The Internet and Higher Education, 2016, 29:1-11.
- [19] 李爽,王增贤,喻忱,等. 在线学习行为投入分析框架与测量指标研究:基于 LMS 数据的学习分析[J]. 开放教育研究,2016(2):77-88.
- [20] 杨现民,王怀波,李冀红. 滞后序列分析法在学习行为分析中的应用[J]. 中国电化教育,2016(2):17-23.
- [21] 魏顺平,韩艳辉,王丽娜. 基于学习过程数据挖掘与分析的在线教学反思研究[J]. 现代教育技术,2015(6):89-95.
- [22] 姜强,赵蔚,李勇帆,等. 基于大数据的学习分析仪表盘研究[J]. 中国电化教育,2017(1):112-120.

## Factors Affecting Online Learning Quality of Adult Learners from the Perspective of Human Behavior Dynamics

DONG Ke<sup>1</sup>, JIA Liyuan<sup>2</sup>, YAO Mengyue<sup>1</sup>

(1. School of Economics and Management, Bengbu University, Bengbu Anhui 233032, China;

2. School of Information and Architecture Engineering, Anhui Open University, Hefei Anhui 230022, China)

**Abstract:** With the increasingly fierce social competition and the normalization of lifelong learning, the online learning rules and quality of adult learners has attracted more and more attention from researchers. The online learning behavior data of adult learners are collected and quantitatively analyzed. Firstly, the human behavior dynamics research method is used to empirically study the characteristics that the online learning time interval of adult learners obeys the power law distribution, and the online learning behavior has the characteristics of flare-up and heavy tail. Secondly, the multiple regression analysis method is used to find out the internal factors that affect adult learners' online learning. Finally, the clustering analysis algorithm is used to further subdivide the sample group of adult learners, and explore the motivation that affects the learning of different learners in the subdivision state. The research results can provide theoretical guidance for teachers to design online teaching and fully tap learners' interests in learning.

**Keywords:** adult learner; learning behavior; human behavior dynamics; K-means algorithm; multiple regression model

[责任编辑 许炎]