

# 以学习分析技术理解学习失败

## ——基于在线课程退课行为的视角

刘桐 顾小清

(华东师范大学 教育学部教育信息技术学系, 上海 200062)

**[摘要]** 理解学习失败是教育研究的热点。然而,已有研究主要依赖自我报告、社会调查等方法,存在主观性较强等问题,不适合在线学习情境。为解决这一问题,本研究利用细粒度学习行为数据,通过引入“学习路径相似度”概念,以时域建模的方式,找出不同类型退课学习者群体,结合未退课学习者数据,依据学习路径的图拓扑参数和行为发生概率,在学习意义层面描述并解释不同类型学习者之间的差异,从而理解学习失败发生的过程,为探明“学习何以发生”提供新的依据。研究发现:1)退课学习者存在“体验型”与“孤立型”两类学习者;2)图拓扑参数计算结果表明,不同类型学习者在学习投入、学习资源切换频率、学习路径聚集度、小世界网络特性和学习资源重要性等方面存在明显差异;3)行为发生概率结果表明,不同类型学习者存在类似的视频交互行为,“体验型”学习者存在课程甄选行为,“孤立型”学习者缺少社交性交互行为。文章最后从分析方法、辩证理解退课、创新课程形式和社会性学习支持等方面提出对策和建议。

**[关键词]** 学习分析; 学习失败; 退课行为; 学习路径; 学习者聚类

**[中图分类号]** G434

**[文献标识码]** A

**[文章编号]** 1007-2179(2022)03-0093-09

### 一、引言

自下向上以数据驱动的形式理解学习现象,解读学习发生机理,从而连接以人为中心的学习设计,是当前技术赋能学习的核心议题,也是学习分析技术的未来走向(顾小清等,2020)。退学是线下学习失败的主要表现(Ajjawi et al., 2020)。然而,随着学习形式的不断进化,研究者认识到,在线学习情境中,退出某门课程学习也是一种学习失败,并指出学习倦怠、学习抱怨、学业情绪等可作为预测学

习失败的重要指标(Jones & Rienties, 2022)。由于在线学习的自愿性与灵活度强,随着受众群体逐步扩大,退课现象日益严重。据统计,edX 在线学习平台的课程结业率只有 5%,学堂在线平台的课程结业率也仅有 4.5%(Feng et al., 2019)。因此,探明在线学习退课发生的原因具有重要的现实意义(Lee & Choi, 2011; Jin, 2020; Chaker & Bachelet, 2020)。

诸多学者围绕学习者退课现象开展研究。常李艳等(2021)使用文献分析方法,发现学习者个人、

[收稿日期] 2022-02-22

[修回日期] 2022-04-15

[DOI 编码] 10.13966/j.cnki.kfjyyj.2022.03.010

[基金项目] 2020 年上海市科学技术委员会科研计划项目“教育数据治理与智能教育大脑关键技术研究及典型应用”(20511101600)。

[作者简介] 刘桐,博士研究生,华东师范大学教育信息技术学系,研究方向:学习分析、教育数据挖掘(messil9@163.com);顾小清,教授,华东师范大学教育信息技术学系,研究方向:教育大数据与学习分析、技术支持的教学创新(xqgu@ses.ecnu.edu.cn)。

[引用信息] 刘桐,顾小清(2022). 以学习分析技术理解学习失败——基于在线课程退课行为的视角[J]. 开放教育研究,28(3): 93-101.

课程以及社会环境因素是造成学习者退课的重要影响因素。休和张永森(Hew & Cheung, 2014)使用问卷调查方法,发现扩展知识面、满足好奇心和个人挑战等是学习者参加在线课程学习的主要驱动力,而获取结业证书不是多数学习者的最初目标。然而,以上研究大多采用传统的社会学方法,存在一定主观性。随着学习过程数据的积累以及智能技术的发展,学习路径分析技术可逐步弥补传统方法的不足。有学者从细粒度数据层面,通过专家知识与真实数据双驱动的形式,使用集成学习方法(Narayanasamy & Elçi, 2020)、深度学习方法(Yin et al., 2020)等知识发现技术,寻找学习者在线学习的行为模式与规律。

在此背景下,本研究考虑在线课程参与形式自由化、课程内容碎片化、学习时间灵活化等特点,利用细粒度学习数据,综合使用数据挖掘算法与工具,从时域角度建模学习者退课过程,动态捕捉学习路径的状态变化,解析学习者退课的行为特征和规律,为理解学习失败提供事实依据。

## 二、文献综述

分析学习路径的动态演变是理解学习失败的切入点。通过查找近三年有关“学习路径”的国内外文献发现,当前研究主要聚焦于学习路径顶层设计、学习路径模式发现和学习路径自动生成。

学习路径的顶层设计在理论层面探讨“应然”问题。钟绍春(2020)提出,以教学目标为逻辑起点,以学科知识和能力图谱为核心,通过筛选学生群体的特点,提供选择数量少、覆盖度高的路径集。唐烨伟等(2019)设计了基于学习者画像的精准个性化学习路径规划框架,通过分析学习结果,为学习者提供针对其个性学习需求、学习内容和学习活动。

行为模式挖掘关注学习活动的自组织过程。乐惠骁等(2019)利用优秀慕课学习者的平台记录数据,挖掘其学习序列长度、行为共现等方面的行为模式及其差异。模式识别、机器学习等智能算法也被用于学习行为挖掘。有学者使用层次聚类方法,发现学习者在线游戏学习环境的行为路径模式,并与最终学习结果进行关联挖掘(Feng & Yamada, 2021)。也有学者基于学习者参与讨论、访问课程资源以及登录时间等行为序列数据,使用主成分分

析算法降低数据维度,提取数据特征,再利用贝叶斯网络与决策树算法,预测和预警学习者退课(Xing et al., 2016)。

学习路径自动生成是个性化路径推荐的主要途径。任维武等(2020)通过引入“契合度”概念,提出基于强化学习方法的自适应学习路径生成机制,实现学习者与学习资源之间的合理匹配。李振等(2019)通过先验知识子图生成、学习目标子图生成和学习路径生成与优化三个步骤,输出覆盖学习目标的知识元学习序列。高嘉骐等(2021)基于知识图谱,通过评估学习者状态,提出学习路径自动生成方法。吴东山等(Son et al., 2021)使用遗传算法和蚁群优化算法等元启发式算法,生成符合学习者学习背景和学习目标的学习路径,解决“信息迷航”等问题。

可见,学术界对学习路径的内涵存在两种理解:一是学习路径是已然存在的先验预设。此种观点将学习路径等同于知识关联,认为依赖于知识之间的层级、包含与被包含等逻辑关系,通过差异化的知识点、概念、资源等联结方式形成的结构化序列,就是学习路径。肖奎等(2018)认为,学习路径是从起始节点(知识基础)到达目标节点(学习目标)的过程。二是学习路径是学习者的行为轨迹。此种观点认为学习路径是学习过程的客观行为记录,它随着学习主体所处情境、行为习惯以及学习任务的差异而呈现不同样态,构建过程也极具生成性。詹森等(Janssen et al., 2008)认为学习路径是学习者达成学习目标的一系列行为操作。

以上两种理解存在本质差异,前者秉持学习路径是严谨完备的专家知识体系子集的观点,后者认为学习路径是学习者学习过程的外显行为表现轨迹。相较而言,行为轨迹更加契合学习分析技术方法。本研究将学习路径定义为:学习者与学习资源交互过程中生成的行为序列,在数据层面具有细粒度、生成性、可计算等特点,同时具有时序化、网络化特征。

综上,分析挖掘学习路径以及终点事件的生成过程是当前研究的热点,但较多研究仅利用学习成绩对学习者群体做简要划分,并依此为分析依据,难免会忽视学习者的学习动机、学习方式以及学习习惯等方面多样性,为理解学习失败带来认知偏差。因此,

本研究依照从下向上的思路,从细粒度数据中挖掘学习者的不同群体,分析不同群体的内在模式规律,从另一个角度理解学习失败。

### 三、研究设计

#### (一) 研究问题

如何从学习者的过程性数据中识别并提取合理且可解释的模式与特征,是本研究的核心问题,具体包括:

1) 如何提取退课学习者群体的行为特征?换言之,从学习路径差异的角度,使用何种方法鉴别退课学习者的类型?

2) 不同类型学习者的学习路径有何差异?不同类型学习者群体之间的行为模式以及行为差异是什么?造成差异的原因是什么?

#### (二) 数据介绍

本研究调取了学堂在线平台“心理学概论”课程的数据仓库,涉及选课用户的视频操作和交流讨论等学习活动的162 996条数据记录(Yu et al., 2020),涵盖1 634名学生(退课学生1 303名,未退

课学生331名,退课率为79.4%)。本研究主要使用四个数据字段,分别是学习者ID、学习行为(如查看问题、点击视频、回复讨论等)、学习资源ID(如视频资源、发布问题、参与讨论等)和系统时间戳。

#### (三) 研究方法与过程

本研究首先利用学习者是否退课的标签数据,区分退课与未退课学习者,划分退课学习数据集与未退课学习数据集;其次,截取学习者ID列表、学习资源ID列表和学习者操作行为列表,形成退课和未退课学习者的行为数据集,并通过数据清洗和索引编码处理,形成数值化的行为编码序列,由此表示学习路径;再次,计算学习路径相似度,构建学习路径相似度矩阵,作为聚类算法的输入,发现不同类型的退课学习者群体;最后,比较不同类型退课学习者群体与未退课学习者群体在学习路径图拓扑参数与行为模式等方面差异(见图1)。

1. 学习路径编码。本研究使用索引编码方法处理学习路径数据。具言之,本研究利用资源ID与具体行为(如暂停、播放、快进、关闭等),作为

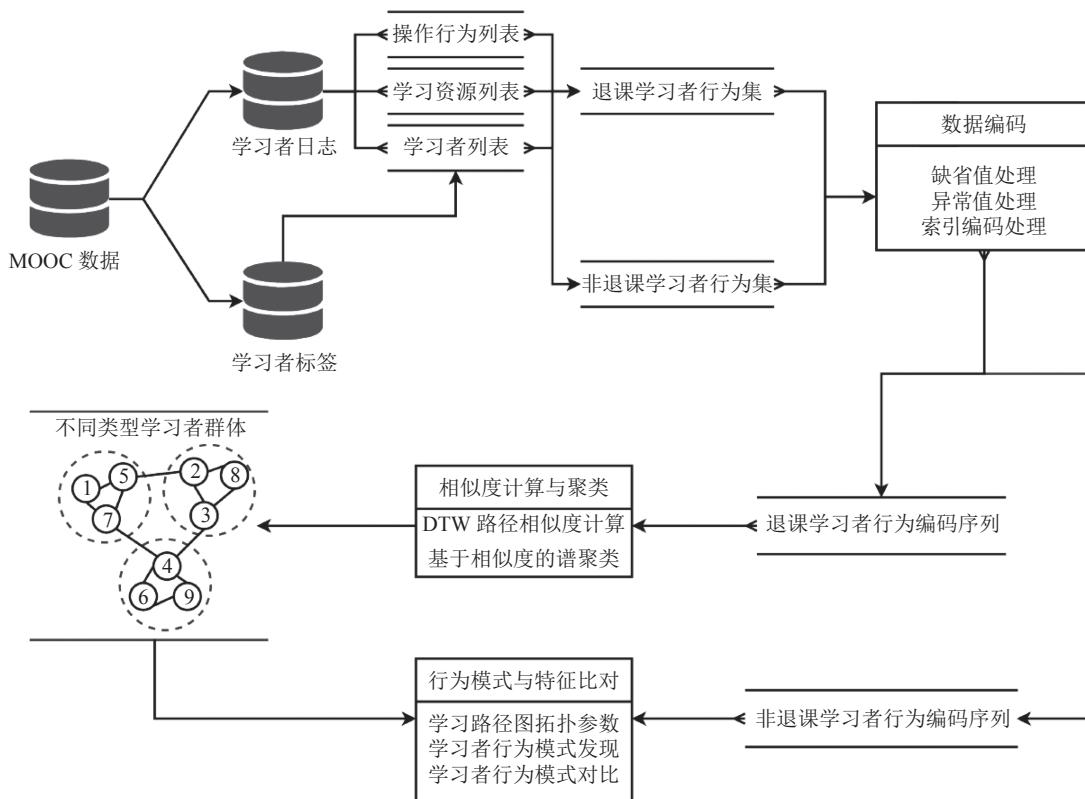


图1 研究技术路线

学习路径节点的索引依据,保证每个资源的操作行为都具有独立的编码。例如,学习路径编码为[1, 17, 20, 9, 11, 21],“1”表示加载A视频,“17”表示暂停A视频,“9”表示加载B视频,“11”表示关闭B视频。

2.学习路径相似度计算。群体特征挖掘的主要难点在于如何基于学习路径差异,从群体中发现相似的个体。因此,如何计算不同学习者路径之间的相似度成为首要解决的技术问题。由于学习者路径呈现离散化、时域难以对齐等问题,无法直接计算学习路径的空间距离,因此本研究使用动态时间规划算法计算学习路径之间的相似度。计算方法见公式(1)。动态时间规划算法的目标是通过离散序列数据点距离计算方法,按次序求解两个序列的点间距离,计算构建点距离矩阵,并在矩阵中找到经过的元素值和最小路径,使得最终累积距离最小。累积距离就是两条学习路径的相似度。

$$DTW(X, Y) = \min_{\phi} d_{\phi}(X, Y) \quad (1)$$

3.基于路径相似度的聚类计算。学习路径由学习资源串联而成,随学习进程逐步呈现出有向无环图的特征。本研究使用基于图论的谱聚类方法获取不同类型的学习者群体。与K-Means等算法相比,谱聚类方法对数据分布的适应性更强,计算量较小,也易于实现。同时,为了衡量聚类效果与可信度,本研究选取方差比标准作为聚类效果的度量衡,该标准通过计算聚类簇的组内聚合度和组间分散度之间的比值,衡量聚类簇之间的距离与差异,从而得到不同聚类群体,计算方法见公式(2)。

$$S = \frac{Tr(B_k)}{Tr(W_k)} \times \frac{N-k}{k-1} \quad (2)$$

其中,k代表聚类类别数,N代表数据数目,Tr( $B_k$ )代表组间协方差,Tr( $W_k$ )代表组内协方差。该分值越大,说明聚类效果越好。

4.基于图拓扑参数的学习路径差异比较。为进一步解释聚类结果的教育意义,基于学习路径的图特征,研究选取五个图的拓扑参数指标,比较学习路径差异:1)平均路径长度,指学习路径总长度的平均值;2)平均边长度,指学习路径中连接任意两个节点之间最短路径的平均值;3)平均聚集系数,指评估学习路径中节点的聚集程度;4)平均边长度/平均聚集系数,指平均边长度与平均聚集系数的

比值,评估学习路径是否具有小世界网络的特征;5)特征向量中心度,用于评估学习路径中节点的重要程度。

5.基于步长矩阵的行为模式差异比较。行为模式用于发现学习者操作之间的相关性与关联关系,区别于聚合统计后的概率计算。本研究选取步长矩阵作为学习者行为模式发现的工具,通过计算独立路径点的行为发生概率,在细粒度层面预测学习者的具体操作,从而得出学习者行为转换规律,并结合数据可视化方法,直观地发现行为模式及差异。

## 四、研究结果

### (一) 学习路径相似度与聚类结果

研究问题一聚焦鉴别不同类型的退课学习者。如上文所述,行为序列的距离度量方式会影响学习路径相似度结果,聚类簇数量是影响最终聚类效果的重要参数。因此,在相似度计算时,本研究选取三个常用的距离计算方式,包括欧氏距离、汉明距离、L1距离;在谱聚类计算中,选择2、3、4三个聚类簇数量作为聚类算法的超参数候选集合,进行综合计算和对比实验。

数据清洗、编码、相似度计算和聚类计算均使用Python数据处理脚本,对比实验结果见表一。可以看出,在汉明距离和两个聚类簇的参数下,聚类的方差比分数最高(1 032.52),说明在此参数下聚类效果最好。经过进一步分割计算,本研究发现退课学习者群体中存在两个聚类群体,分别有1 101人和199人。

表一 不同超参数的聚类度量结果

方差比分数		聚类参数		
		2	3	4
距离参数	欧氏距离	185.38	300.27	365.45
	汉明距离	<b>1 032.52</b>	888.22	1 030.21
	L1距离	185.38	300.27	327.95

### (二) 学习路径群体特征差异比较

研究问题二聚焦于在鉴别不同类型学习者群体后,从教育意义层面比较学习者的群体差异。综合考虑未退课学习者与两类退课学习者,本研究利用图拓扑参数与行为步长矩阵,分别比较探析学习

路径的图特征差异和学习行为模式差异。

### 1. 基于图拓扑参数的学习路径差异比较

1) 学习路径投入度。学习者平均学习路径长度指学习者行为操作长度的平均值,一定程度上表示学习者与学习资源之间交互的次数,更能反映学习者在线学习的投入程度。结合表中数据可以发现,三类学习者的平均学习路径长度分别是21.41、172.37和293.14。与退课学习者相比,未退课学习者的平均学习路径长度显著长于退课学习者,说明未退课学习者投入时间更多,与学习资源的交互频次更长。退课学习者中,类别1退课学习者的平均学习路径长度最短,类别2退课学习者的平均学习路径长度明显长于类别1退课学习者。由此可见,退课学习者的学习路径差异仍较大。

2) 学习路径稳定度。平均边长度一般用来评估网络中两节点之间的距离,能够体现网络节点的联通性。例如,六度分隔理论体现了人际社会的“弱纽带”性质。本研究的三类学习者的平均边长度分别为2.379、2.184和6.454。可见,两类退课学习者的边长度都短于未退课学习者。学习路径节点由学习资源构成,连接边由操作行为构成,平均边长度代表学习者在不同学习资源之间转换时所执行的操作次数。与未退课学习者相比,退课学习者在不同学习资源之间的切换频率更高,单一资源使用的持续时间较短。值得一提的是,结合平均学习路径长度,本研究推断类别1退课学习者资源使用数量低于未退课学习者与类别2退课学习者。

3) 学习路径聚集度。平均聚类系数用于描述网络中顶点之间结集成团程度,发现图结构中的小集团。本研究中,三类学习者的平均聚集系数为0.545、0.622和0.284。由此可见,类别2退课学习者的学习路径聚集系数较高,未退课学习者的学习

路径聚集系数最低。结合前两个图的拓扑参数,本研究发现,类别1退课学习者由于投入的时间较少,利用资源较少,势必会出现高聚集性,而类别2退课学习者虽然投入的时间较多,但只利用局部有限的课程学习资源,缺少与其它可用资源交互,导致学习路径呈现较高聚集性。

4) 学习路径均衡度。平均聚类系数与平均边长度通常被同时考虑,用来判断网络是否具有小世界网络特征。小世界网络最早由沃茨和斯特罗加茨在1998年提出,通常意指网络中包含节点集团和近距离网络,同时子网络中大部分节点之间都具有连接,因此,小世界具有高平均聚类系数和低平均边长度的特征(Watts & Strogatz, 1998)。三类学习者平均聚类系数与平均边长度的比值分别是4.365、3.511和22.725(见表二)。结合小世界网络的属性,相较于未退课学习者,类别2退课学习者的学习路径更具小世界网络特征。结合前三个图的拓扑参数结果,本研究从两方面理解此现象:一方面,类别2退课的学习路径具有小世界网络特征,此类学习者可能面临“信息迷航”,学习内容与活动形式较为局限,难以找到可进一步利用的课程学习资源,由此导致退课行为的发生;另一方面,此类学习者的自主学习目的可能更强,在深入学习所需资源后难以达到目的,便退出了课程。

5) 学习路径重要度。特征向量中心度通过迭代计算某一节点与相邻节点的中心性分数,从而体现此学习路径节点的重要程度。特征向量中心度为非负值,数值越大代表节点的重要程度越高。结合图2可知,三类学习路径节点特征向量值分布在0.1与0.27之间,与未退课学习者相比,退课学习者的特征值上限较高,未退课学习者的特征节点数量多于类别2退课学习者,且远多于类别1退课学习者;与未退课学习者相比,类别2退课学习者

表二 不同类型学习者学习路径图拓扑参数

参数值		图拓扑参数			
		平均学习路径长度	平均边长度	平均聚类系数	平均聚类系数/平均边长度
学习者类型	退课_类别1	21.41	2.379	0.545	4.365
	退课_类别2	172.37	2.184	0.622	3.511
	未退课	293.14	6.454	0.284	22.725

的高特征向量值分布较平均,未退课学习者的高特征向量值分布较集中,处于0.15与0.18之间,说明大部分学习资源在后两种学习者学习过程中具有较高的重要性。但由于类别2退课学习者与未退课学习者的特征值分布出现了两级分化的特征,部分资源的特征值接近于0,说明存在部分重要性较低的学习资源。

## 2. 基于步长矩阵的行为模式差异比较

行为步长矩阵(behavior steps matrix)用于计算每步中出现不同学习行为的概率,借此解释学习者每步操作中倾向于使用的资源。鉴于每类学习者的平均学习路径长度存在差异,本研究利用可视化方法呈现行为步长矩阵,用颜色深浅代表行为发生概率的高低,由此发现三种学习者行为模式(见图3)。

1)类别1退课学习者注重课程体验。在完成约50步操作行为后,大部分这类退课学习者就结束了课程学习,且学习过程以加载视频、播放视频、暂停视频等行为操作为主。类别1退课学习者的平均学习路径长度最短,接触的学习资源也较少,侧重于课程内容体验。例如,在观看课程视频介绍或少部分课程视频后,“体验型”学习者可能发现课程内容不符合他们的预期,或未产生继续学习的兴趣,就转学其它在线课程。大部分退课学习者属于此类型(占85%)。

2)三类学习者存在相似的行为操作模式。他们都侧重视频操作交互行为,这与在线课程设计本身有极强的相关性,因为当前在线课程多以视频为主。同时,讨论区的发帖、回帖、回答学习问题、检查学习问题正误等操作也出现在学习者学习路径中,但占比较少。而在视频交互操作行为中,类别1退课学习者较多采用“加载视频—播放视频—暂停视频”的行为模式,类别2退课学习者与未退课学习者较多采用“加载视频—播放视频—

暂停视频—回答问题—检查问题答案正误—播放视频”等连续学习的行为模式。

3)类别2退课学习者的学习路径缺少社会性学习行为。虽然类别2退课学习者的学习行为模式具有较连续性的特征,但最终还是未完成课程内容的学习(见图3)。未退课学习者参与发帖、回帖等社交行为操作较多,类别2退课学习者的行为操作多发生于视频交互操作、课程问题回答等可独立完成的学习情景中,属于“孤立型”学习者。由此可见,此类学习者一方面缺少主动求助、分享等行为,另一方面因为较少参与课程讨论,因此获得同伴或助教的支持也会减少。

## 五、结论与讨论

通过引入“学习路径相似度”概念,本研究基于在线课程学习者的行为数据、利用时间动态规划算法和谱聚类算法,发现退课学习群体存在“体验型”与“孤立型”两类学习者。为解释聚类结果的教育意义,本研究综合考虑上述两类退课学习群体和未退课学习者,从学习路径平均长度、稳定性、聚集度、均衡度和重要度五方面比较,发现退课学习者的学习投入较低,聚集性较高,且存在内部差异,“孤立型”学习者的学习路径更具小世界网络特征。整体来看,退课学习者在利用资源和学习主动性方面与未退课学习者存在较大差异。借助行为步长矩阵,本研究发现三类学习者存在类似的视频交互行为,“体验型”学习者存在课程甄选行为,“孤立型”学习者缺少社会性交互学习行为。

(一)基于学习路径相似度计算,为理解学习失败提供新视角

与传统学习失败研究方法不同,本研究依照从下向上的思路,利用学习过程数据记录,挖掘分析退课行为的群体特征与模式,为理解学习失败提供

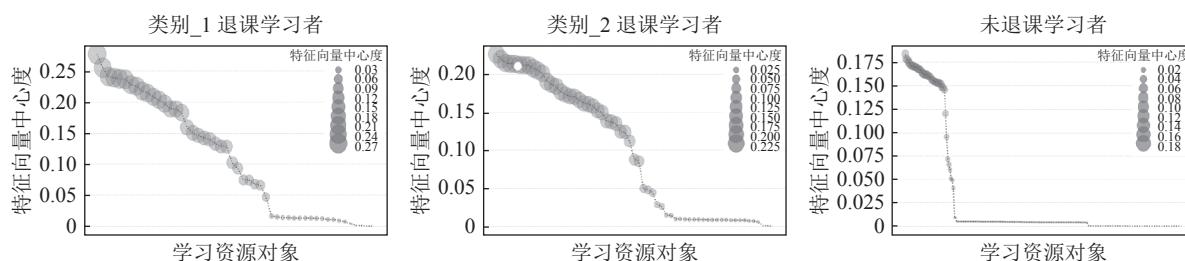


图2 三类学习者学习路径特征向量中心度分布

新的研究视角。具言之,传统研究方法大多采用自我报告或社会调查的方法,是从上向下的设计思路,少部分行为分析研究利用计数的方式,统计学习过程中的行为转换关系,未能有效利用过程性的行为信息。本研究使用动态时间规划算法和谱聚类算法,不仅考虑学习行为出现的频次和转换概率,同时利用行为点层面的序列演变信息,计算学习者行为序列的空间距离关系,构建学习路径相似度矩阵,并基于子图切分的谱聚类算法,识别不同的学习者群体。整体来看,本研究所用方法,考虑了学习者交互过程的整体信息,提升了数据利用效率;还利用了学习路径的图特征,为解释算法的教育意义打下了基础。

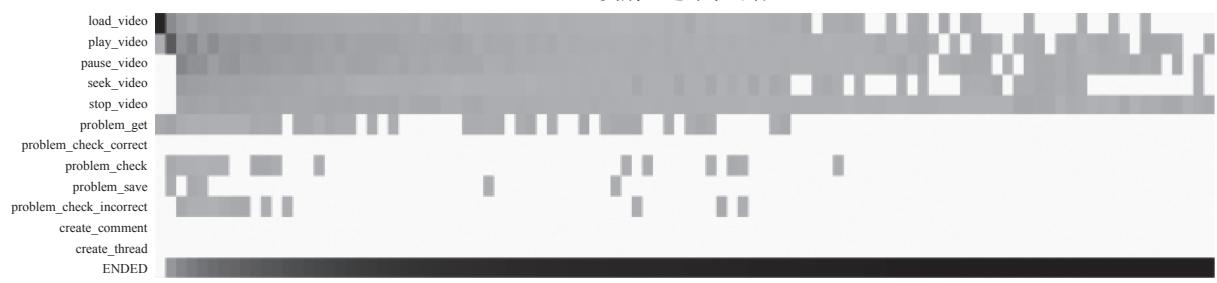
## (二) 辩证理解学习者退课,优化课程建设结构

### “体验型”学习者的高占比是导致所谓“高

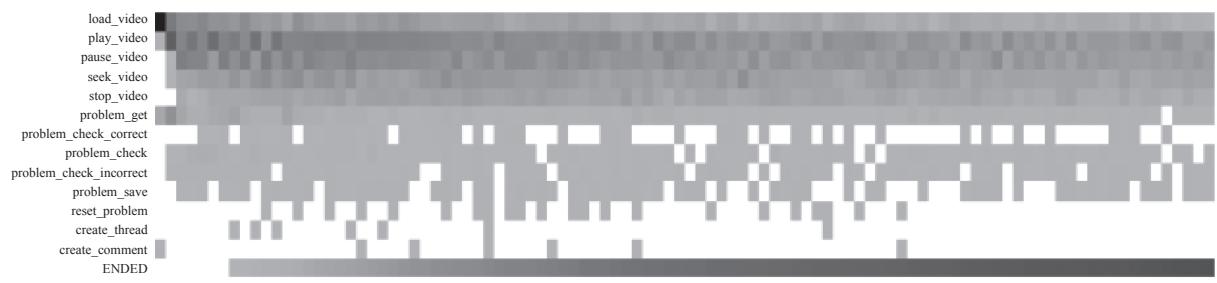
辍学率”的重要原因。在线课程建设的初心是打破优质教育资源的壁垒,为学习者提供包容开放的在线学习环境。退课行为虽然降低了资源的利用率,却是学习者根据自身学习需求积极选择的结果。因此,合理利用“体验型”学习者的行为模式,建立基于学习者静态属性、学习者行为交互数据、课程属性标签以及课程内容标签等多维度过程数据的个性化推荐机制,可为后续课程建设提供改进优化的思路。相应的对策包括:

一是定位潜在学习者。发现课程的潜在受众,是节省双方时间成本的有效方式。“体验型”学习者的行为交互序列长度较短,课程选择短暂,难以有效探知课程本身的学习内容。因此,本研究建议在历史行为数据的基础上,通过课程标签、内容标签等形式,建立学习者与课程之间的匹配关系,

类别 1 退课学习者



类别 2 退课学习者



未退课学习者

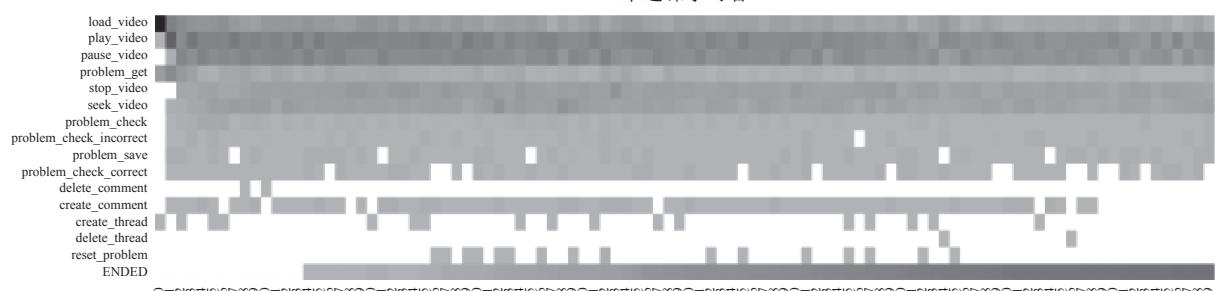


图 3 三类学习者行为步长矩阵

向学习者解释推荐的理由与依据,为学习者推荐相关的介绍性资源,从而挖掘潜在的课程学习者。

二是提供个性化学习支持,避免信息迷航。当前部分在线课程成绩与学习者修业学分挂钩,学习者在非主动选择的情况下,如果难以获取有效的学习支持,学习体验下降,就会出现“孤立型”学习者。因此,结合此类学习者学习路径小世界网络的特征,及早发现信息迷航行为,为学习者及时提供个性化学习解答与帮助,将有助于学习者尽早摆脱疑问,引导学习者进入下一阶段的学习。

### (三)创新形式,激发学习者学习兴趣

结合研究所发现的共性行为模式可以发现,当前学习者的课程学习以观看视频为主,穿插问题回答、课程讨论等传统在线学习形式,且不同类别学习者的行为交互持续度不高。为追寻这一现象的原因,本研究进一步统计学习者人口信息,发现大部分学习者处于26~30岁间,且以本科生和研究生为主。作为受过高等教育,在网络环境中成长的“数字原住民”,这些学习者的学习需求更明确,对学习资源的质量要求更高,且更喜欢主动选择信息。因此,单纯的视频学习容易带来学习过程的单调性,难以吸引学习者持续的深度参与。

因此,创新课程形式,是激发学习者学习兴趣,突破学习瓶颈的可行路径。首先,课程资源的质量是基础,本研究建议在提升视频资源清晰度的基础上,不断改进课程活动的组织排列方式、学习评价方式,增强学习资源的可用性与易用性。其次,教师可增设在线学习的反馈渠道,加强与学习者的交流,例如,可利用视频弹幕的方法,鼓励学习者学习过程中随时发言,降低学习反馈的门槛与成本,用创新的交流方式,为后续教学活动提供异步反馈。

### (四)提供支持,建立同伴沟通渠道

综合来看,“孤立型”学习者是潜在的课程完 成者。此类学习者的学习投入较高,但缺少社交性学习互动,获得的学习支持较少。有学者认为缺少社会性支持是大部分退课学习者的主要原因(Onah et al., 2014)。因此,随着课程学习的推进,为学习者提供多样化的学习支持,是提升课程完结率的重要手段。主要对策如下:

其一,建立有效的人机交互方式。例如,可利用智能对话机器人,为学习者提供适应性的情感支

持,避免负面学业情绪对学习过程的消极影响,同时可作为资源推荐的窗口,为学习者提供适应性的学习资源支持;其二,建立通畅的学习沟通渠道,促进学习者与讲授者、学习同伴互动,合力建设学习分享空间,并与学业评价相挂钩,推动学习者主动开展学习交流,提升同伴交流频率,增加学习互动的机会。

### [参考文献]

- [1] Ajjawi, R., Dracup, M., Zacharias, N., Bennett, S., & Boud, D.(2020). Persisting students' explanations of and emotional responses to academic failure[J]. Higher Education Research & Development, 39(2): 185-199.
- [2] Chaker, R., & Bachelet, R.(2020). Internationalizing professional development: using educational data mining to analyze learners' performance and dropouts in a french mooc[J]. International Review of Research in Open and Distributed Learning, 21(4): 199-221.
- [3] 常李艳,刘婧,黄崇(2021). MOOC 研究:近 20 年文献的系统分析 [J]. 图书馆论坛, 41 (11): 38-51.
- [4] Feng, W. , Tang, J. , & Liu, T. X. (2019). Understanding dropouts in MOOCs[A]. The 31th AAAI Conference on Artificial Intelligence[C]. January 27-February1, 2019, Hawaii, USA: 517-524.
- [5] Feng, X., & Yamada, M.(2021). An analytical approach for detecting and ex-plaining the learning path patterns of an informal learning game[J]. Educational Technology & Society, 24(1): 176-190.
- [6] 高嘉骐,刘千慧,黄文彬(2021). 基于知识图谱的学习路径自动生成研究 [J]. 现代教育技术, 31 (7): 88-96.
- [7] 顾小清,胡艺龄(2020). 理解、设计和服务学习:学习分析技术的回顾与前瞻 [J]. 开放教育研究, 26 (2): 40-42.
- [8] Hew, K. F., & Cheung, W. S.(2014). Students' and instructors' use of massive open online courses ( MOOCs ) : motivations and challenges[J]. Educational Research Review, 112: 45-58.
- [9] Janssen, J., Berlanga, A., Vogten, H., & Koper, R.(2008). Towards a learning path specification[J]. International Journal of Continuing Engineering Education and Life Long Learning, 18(1): 77-97.
- [10] Jin, C.(2020). MOOC student dropout prediction model based on learning behavior features and parameter optimization[J]. Interactive Learning Environments, 2020: 1-19.
- [11] Jones, M., & Rienties, B.(2022) . Designing learning success and avoiding learning failure through learning analytics: the case of policing in England and Wales[J]. Public Money & Management, 42(1): 32-39.
- [12] Lee, Y., & Choi, J.(2011). A review of online course dropout research: implications for practice and future research[J]. Educational Technology Research and Development, 59(5): 593-618.
- [13] 乐惠骁,范逸洲,贾积有,汪琼(2019). 优秀的慕课学习者如何学习——慕课学习行为模式挖掘 [J]. 中国电化教育, (2): 72-79.
- [14] 李振,周东岱(2019). 教育知识图谱的概念模型与构建方法研究 [J]. 电化教育研究, 40 (08): 78-86+113.
- [15] Narayanasamy, S. K., & Elçi, A.(2020). An effective predic-

- tion model for online course dropout rate[J]. International Journal of Distance Education Technologies, 18(4): 94-110.
- [16] Onah, D. , Sinclair, J. , & Boyatt, R. (2014). Dropout rates of massive open online courses: behavioural patterns[A]. The 6th International Conference on Education and New Learning Technologies [C]. July 7-9, 2014 Barcelona, Spain: 5825-5834.
- [17] 任维武,郑方林,底晓强(2020).基于强化学习的自适应学习路径生成机制研究 [J].现代远距离教育, (6): 88-96.
- [18] Son, N. T., Jaafar, J., Aziz, I. A., & Anh, B. N.(2021). Meta-heuristic algorithms for learning path recommender at MOOC[J]. IEEE Access, 9: 59093-59107.
- [19] 唐烨伟,茹丽娜,范佳荣,庞敬文,钟绍春(2019).基于学习者画像建模的个性化学习路径规划研究 [J].电化教育研究, 40 (10): 53-60.
- [20] Watts, D. J., & Strogatz, S. H.(1998). Collective dynamics of 'small-world' networks[J]. Nature, 393(6684): 440-442.
- [21] Xing, W., Chen, X., Stein, J., & Marcinkowski, M. (2016) . Temporal predication of dropouts in MOOCs: reaching the low hanging fruit through stacking generalization[J]. Computers in Human Behavior, 58: 119-129.
- [22] 肖奎,吴天吉(2018).一种面向学习路径推荐的知识网络构建方法 [J].计算机工程与科学, 40 (8): 1521-1526.
- [23] Yin, S., Lei, L., Wang, H., & Chen, W.(2020). Power of attention in MOOC dropout prediction[J]. IEEE Access, 8: 202993-203002.
- [24] Yu, J., Luo, G., Xiao, T., Zhong, Q. , Wang, Y., Feng, W., Luo, J., Wang, C., Hou, L., Li, J., Liu, Z., & Tang, J. (2020). MOOC Cube: a large-scale data repository for NLP applications in MOOCs[A]. The 58th Annual Meeting of the Association for Computational Linguistics [C]. July 5-10, 2020 Online: 3135-3142.
- [25] 钟绍春(2020).课堂教学新模式构建方向与途径研究 [J].中国电化教育, (10): 40-48.

(编辑:李学书)

## Understanding Learning Failure with Learning Analytics: A Perspective of Online Course Dropout Behavior

LIU Tong & GU Xiaoqing

(Department of Education Information Technology, Faculty of Education, East China Normal University, Shanghai 200062, China)

**Abstract:** Understanding learning failure is a research hotspot in education. However, the current research relies more on self-report and social survey research methods, which have problems such as strong subjectivity and are challenging to apply to online learning situations. To solve the problem, this study uses fine-grained learning behavior data, introduces the concept of "learning path similarity", and uses time-domain modeling to find out different types of dropout learners; Moreover, combining the learning data from whom did not dropout, according to the graph topology parameters of the learning path and the probability of behavior occurrence, describe and explain the differences between different types of learners at the level of learning meaning, so as to understand the occurrence process of learning failure, and provide new insights into "why learning occurs". The study found that: 1) Among the learners who dropped out of the course, there were two types of learners: "experiential type" and "isolated type"; 2) The calculation results of graph topology parameters showed that different types of learners had significant differences in learning engagement, learning resource switching frequency, aggregation degree of learning paths, the characteristics of small-world networks and the importance of learning resources; 3) The results of behavior occurrence probability show that different types of learners have common video interaction behaviors, and "experiential" learners have course selection behaviors. "isolated" learners lack patterns such as social interaction behaviors. Finally, the article puts forward countermeasures and suggestions from the aspects of analysis methods, dialectical understanding of dropout, innovative curriculum forms, and social learning support.

**Key word:** learning analysis; learning failure; dropout behavior; learning path; learner clustering