

中国 MOOCs 学习者价值研究

——基于 RFM 模型的在线学习行为分析

宗 阳 郑勤华 陈 丽

(北京师范大学 北京 100875)

【摘要】MOOCs 自兴起以来,它的高注册人数和高辍学率特征并存,现阶段中国 MOOCs 大多为以视频为主要学习资源的 xMOOC,对学习者的学习支持力度相对薄弱,这要求学习者具有较高的自主学习能力。在 MOOCs 中,怎样识别学习者价值,采用什么方法对学习者的学习进行分类,进而给学习者提供有效的个性化学习支持是重点研究问题。在 RFM 模型基础上依据 xMOOC 在线学习过程特点,提出中国 MOOCs 学习者的价值衡量模型 ML-RFT,采用模型指标原始值与所有学习者平均值比较的方法将 MOOCs 学习者分为八类。借鉴 RFM 衡量顾客价值的方法,构建 ML-RFT 模型,对 MOOCs 平台上一门 xMOOC 进行学习者的价值识别分类,并有针对性地对每类学习者提供个性化的学习支持策略。研究表明采用 RFM 方法可以对 MOOCs 学习者价值进行有效识别区分,学习者分类结果可以成为 MOOCs 教师或管理人员对在线学习者提供个性化学习支持服务的重要依据。

【关键词】RFM 模型,ML-RFT 模型,xMOOCs,在线学习行为,学习者价值;

【中图分类号】G64

【文献标识码】B

【文章编号】1001-8700(2016)02-0021-08

DOI:10.13927/j.cnki.yuan.2016.0014

一、引言

随着研究和实践的深入,MOOCs(大规模在线开放课程)在经历了2012年的MOOCs之年和2013年的反MOOCs之年的转变之后,发展更趋于多样化。SPOCs(小型私有化在线课程)、DLMOOC(深度学习MOOC)、MOOL(大规模开放网络实验室)、DOOC(分布式开放协作课)等新名词大量涌现,以开放在线教育为主要特征的后MOOCs时代已经到来^[1]。MOOCs从2012年大规模兴起以来,课程注册人数逐年上升,但是同时伴随而来的是高辍学率的出现。MOOCs在给MOOC学习者带来学习便利的同时,也对学习者的自主学习能力以及学习毅力等提出了更高的要求。有研究表明MOOCs辍学率

高的原因主要集中在缺乏毅力、可投入的时间有限、语言困难和网络障碍方面,在找出导致辍学率高的原因后,该研究针对MOOC课程提出了针对整体学习者或课程的策略,如适当增加付费政策、丰富证书层次以及限制选课人数等^[2]。在线学习中,优质的学习支持服务是提高在线教育服务水平,确保在线教育质量的关键。

MOOCs中教师与学习者时空分离,如何帮助学习者适应在线学习环境,保障学习效果,是以MOOCs为代表的在线教育需要解决的难题。随着信息技术的飞速发展,在线学习越来越突出体现了“以学习者为中心”的教学理念。这种理念对于学生来说其实就是一种个性化的学习,根据学生需要及实际情况,让学生自由选择适合自己的学习内容

【基金项目】全国教育科学“十二五”规划2014年度国家重点课题(编号:ACA140009)“教育信息化与大型开放式网络课程(MOOCs)战略研究”的研究成果。

【作者简介】宗阳,北京师范大学远程教育研究中心硕士研究生;郑勤华,北京师范大学远程教育研究中心副教授,管理学博士,硕士生导师;陈丽,北京师范大学远程教育研究中心教授,理学博士,博士生导师。

和学习手段,有助于充分挖掘学习者的学习潜力^[3]。对于教师和管理者来说,怎样给学习者提供及时、合理、有效的个性化学习支持服务是保证在线教学质量的关键。适应性、个性化的质量学生支持服务需要建立在充分了解相关学生信息的基础上^[4]。这要求教师和管理者在提供学习支持服务的过程中,主动了解每个学习者的学习特征并时刻关注其变化,为他们提供实时、个性化的支持服务,最大程度满足学习者的一切合理需求。

现阶段在中国 MOOCs 学习中,由于在线学习过程的复杂多样性以及对在线学习过程分析方法的成熟性,中国 MOOCs 教师对学习者的学习评价大都停留在总结性评价(基本为考试成绩)上。学习者的个性化学习应该体现在具体在线学习过程中,对学习者的学习仅仅进行总结性评价显然不适应新时代在线学习理念,更无法为学习者提供有效的个性化学习支持服务。对在线学习者进行有效的个性化学习支持需要建立在对学习者充分了解的基础之上,只有充分了解每个或每一类学习者的学习特征,才能根据学习者不同学习特点提供有效的个性化学习支持。对在线学习者的学习过程行为进行分析,了解在线学习者学习特点,识别学习者价值,进而根据学习者价值进行学习者分类是在学习过程中对学习者提供有效个性化支持服务的前提条件。

对学习者个体来说,学习是一个投入产出的过程^[5],需要学习者投入宝贵的学习时间和必要的财力才能获得满意的产出。对于教师或管理者来说,学习者的价值体现在对教学效果的保障上。本研究中学生价值主要用来表征学习者在线学习过程中的学习积极性和学习投入,如果价值较高表示学习者学习积极性较高,学习投入较多,正常情况下会获得满意的学习结果。在中国 MOOCs 环境下,怎样识别学习者价值并对学习者进行分类,进而为不同价值类别学习者提供有效的个性化学习支持是本研究的核心问题。

有研究表明当前中国 MOOCs 教学模式主要仍为 xMOOC(基于行为主义学习理论的 MOOC),这与国外各大主流 MOOCs 平台,包括 Coursera、edX 等平台上的课程教学模式相一致^[6]。xMOOC 教学模

式下的中国 MOOCs 呈现的教学资源多为视频,教学活动多为在线作业或测试。基于此,本研究所说 MOOCs 均指 xMOOC,xMOOC 中在线学习者学习行为路径相对比较清晰,国内一般的在线学习平台都对学习者的学习行为有比较全面的记录,通过分析平台记录的数据可以深入了解学习者的学习行为,进而衡量学习者价值,教师或管理者可以据此为学习者提供个性化的学习支持。

中国 MOOCs 学习者进行在线学习过程可以简单概括为:首先登录 MOOCs 在线学习平台,然后选择要学习的课程资源并且花费时间进行在线观看学习。该学习过程与顾客在线购物的过程比较类似,在线购物时,顾客会首先登录电商网站,然后选择要购买的商品,并花费金钱进行支付。对电商网站忠诚度较高或价值较高的顾客会经常访问并且在电商网站上花费很多金钱购买商品,同样,对 MOOCs 价值较高、学习积极性高、学习投入很多的学习者也会经常访问 MOOCs 学习平台,并且花费很多时间精力学习课程资源。本研究借鉴电商领域应用广泛的 RFM 模型来对学习者进行价值识别和分类。

二、文献综述

RFM 模型是电商领域衡量客户价值和客户创利能力的重要工具和手段,该模型通过分析顾客的最近购买时间(Recency,R)、某一时间区段的购买频率(Frequency,F)及某一时间区段的购买总金额(Monetary,M)这三个指标来衡量客户对企业的忠诚度以及客户的个性化需求情况,这样不仅可以帮助企业进行个性化精准营销,也可以提高客户的再次购买频率^[7]。RFM 的研究主要集中在客户关系管理分析模式中,在众多的客户关系管理中,研究和应用的领域主要集中在电子商务等具有直销行业特点的行业。在远程在线教育领域,台湾学者 Hui-Chu Chang 在 RFM 模型基础上提出 EL-RFM 模型来衡量 e-learning 领域学习者的学习行为^[8],该研究是 e-learning 领域 RFM 模型应用的开创性研究。EL-RFM 三个指标具体内涵如下表 1 所示。

模型每个指标的权值赋值方法依据 Hughes 的排序等分划分方法,Hughes 的排序等分划分方法认

为 R、F、M 三个质变同等重要,每个指标赋予相同的权重,这样将三个指标分别按照从小到大和从多到少进行排序,然后将其平均划分为五等分^[9]。

表 1 EL_RFM 模型定义

指标	意义	定义
EL_R	资源获得与上传时间间隔	当新的学习资源上传后,资源上传时间和学习者第一次获得该资源的时间间隔能被记录。R 值取用资源获得时间间隔平均值
EL_F	登录次数	在一个特定时间内学习者登录学习平台的次数
EL_M	花费时间	学习者花费在做练习或者讨论问题上的时间

EL_RFM 模型在 Hughes 等分划分基础上,将 5 等分换成 3 等分,前 33% 的学习者赋值 3,其次 2,最小为 1。这样学习者的学习行为被分为像 (3,3,3) ... (1,1,1) 共 27 类,基于此分类,研究者在 Moodle 平台上进行了测试实验,基于 Moodle 平台上记录的行为数据得到了关于学习者的价值分类,然后将其与学习者线下课堂学习行为进行对比,得到两者之间的对应关系,具体如下表 2 所示:

表 2 EL_RFM 分类与线下课堂学习行为对应关系

EL_RFM 值	线下课堂学习行为
(3,3,3)	该类学习者是最认真的学习者。少数几个学生几乎不在课上回答问题,但是他们在在线讨论室里讨论非常活跃。
(1,3,1)	该类学生属于中等水平,他们仅仅在考试前进行学习。
(3,1,1)	该类学生从来不会旷课,但是在课上不用心学习,这类学生仅仅下载学习资源但是并不学习。
(1,1,1)	该类学生经常迟到或旷课。

EL_RFM 模型在 RFM 模型基础上,将在线学习者比作顾客,将在线学习行为比如资源获得与上传时间间隔、登录次数、学习时间比作电商领域常用的行为指标,比如最近购买时间、购买频率和购买金额。该模型应用到中国 MOOCs 学习者价值识别上存在一定的应用弊端,具体体现在以下两点:

1. 该模型依赖特定实验环境,不具有在线教育领域普适推广特性,更不适用于 MOOCs 学习领域。该模型需要设定特定的实验环境,R 指标的计算是指资源获得时间与资源上传时间的间隔,这与

MOOCs 学习形式不符,资源获得与在线学习发生概念之间相差较远。

2. 该模型三个指标表征意义不统一,这样导致模型分类结果不准确,只能反映部分学习行为。该模型三个指标具体表征的是在线学习者的学习,R 和 M 指标相较 F 指标较为贴近在线学习者的学习行为,F 指标使用学习者登录平台次数来表征,与实际应该表征的学习次数意义相差较远,登录平台不代表进行学习。R 指标使用资源获得与上传之间的时间间隔来表征,资源获得和学习发生在意义上也相差较远。M 指标表示学习者做练习或讨论的时间,学习者在在线平台上做练习和讨论问题更多可表征学习者的意义建构,但是中国 MOOCs 平台上大多讨论行为发生较少,在以资源呈现为主要特征的 xMOOC 下用讨论或做练习时间来表征学习不现实也不准确。

因此,有必要从 MOOCs 学习过程实际出发,构建新的 RFM 模型来对 MOOCs 学习者价值进行识别分类,进而可以用来指导在线教育实践。

三、ML-RFT 模型构建

MOOCs 学习者的学习流程可简单概括为:学习者登录课程平台观看学习课程资源(视频为主),然后做相应的练习或测试,最后退出课程学习平台。可以看出 MOOCs 学习者学习行为特征比较统一,将 RFM 模型应用于 MOOCs 学习者可以弥补 RFM 模型在远程在线教育领域应用研究的不足。借鉴 RFM 分类方法,本研究构建 ML-RFT(MOOC-Learning RFT 简称)学习者分类模型来实际挖掘分析学习者在线学习行为进而衡量 MOOCs 学习者价值。

本研究将 RFM 模型下顾客的在线购物行为类比为 MOOCs 学习者的在线学习行为,将 RFM 三个指标维度修订为 ML_R,ML_F 和 ML_T,得出 MOOCs 的在线学习者价值衡量模型 ML-RFT。具体定义如下表 3 所示。

ML-RFT 模型基本可以表征 MOOCs 学习者的主要学习过程,MOOCs 中价值较高、投入较多、学习积极的学习者应该具有如下三个特征:

(1) 学习者在学习过程中,会规划好自己的学

习时间,分阶段去学习课程资源。

(2) 学习者在学习过程中,会经常登录课程平台查看学习资源进行学习。

(3) 学习者在学习过程中,会花费较多的时间来观看或浏览学习资源进行学习。

表3 ML-RFT 模型

指标	指标定义	指标表征	指标意义
ML_R MOOC Learning Recen- cy	最近一次学习时间	学习者距离最近一次学习的时间周期,为最近一次学习时间距离设定分析时间之间的时间间隔	ML_R 值与学习者价值之间呈负相关关系。ML_R 值越低即越接近分析时间的学习者,其学习积极性越高,学习潜力价值越高。
ML_F MOOC Learning Frequency	一段时间内的学习频率	学习者一段时间内学习的次数,时间单位可以根据自己分析的实际需求设为季度,月,星期,天数或者任何时间单位。	ML_F 值与学习者价值之间呈正相关关系。ML_F 值越高代表学习者学习投入越多,学习积极性越高。
ML_T MOOC Learning Time	一段时间内的学习时间	学习者一段时间内的总的学习时间,时间单位可以根据自己分析的实际需求设为季度,月,星期,天数或者任何时间单位。	ML_T 值和学习者价值呈正相关关系。ML_T 值越高,表示学习者学习意愿越强烈,学习投入越多,学习潜力越大。

基于 ML-RFT 模型,本研究认为该模型中的三个指标维度分别代表了中国 MOOCs 学习者学习的三个相对独立的学习过程,应该被赋予相同的指标权重。Hughes 的 RFM 模型计算方法比较简单明确,不会产生由于指标权重和阈值设定的主观性导致的分类不准确。因此本研究借鉴 Hughes 的模型计算方法,给 ML-RFT 三个指标赋予同等权重,通过比较各指标 ML-RFT 值和全部 ML-RFT 的平均值,观察其上升下降情况(大于等于用↑表示,小于用↓表示,数据库中为了分析方便分别用 1,0 进行记录),以此来判断学习者的价值,进而将学习者分为(ML_R↓ML_F↑ML_T↑)、(ML_R↓ML_F↓ML_T↑)、(ML_R↓ML_F↑ML_T↓)、(ML_R↓ML_F↓ML_T↓)、(ML_R↑ML_F↑ML_T↑)、(ML_R↑ML_F↑ML_T↓)、(ML_R↑ML_F↓ML_T↑)和(ML_R↑ML_F↓ML_T↓)共八类。具体分类方法流程如下:

首先通过编写数据库函数分别计算每个学习者 ML_R、ML_F、ML_T 三个指标值,然后对每个指标计算平均值,平均指标值计算公式如下:

首先通过编写数据库函数分别计算每个学习者 ML_R、ML_F、ML_T 三个指标值,然后对每个指标计算平均值,平均指标值计算公式如下:

$$\overline{ML_R} = \frac{\sum ML_R_i}{n} \quad (i=0, 1, \dots; n \text{ 为学习者人数})$$

$$\overline{ML_F} = \frac{\sum ML_F_i}{n} \quad (i=0, 1, \dots; n \text{ 为学习者人数})$$

$$\overline{ML_T} = \frac{\sum ML_T_i}{n} \quad (i=0, 1, \dots; n \text{ 为学习者人数})$$

公式 1 ML-RFT 模型指标平均值计算公式

在确定各学习者指标值和平均值后,将学习者单个指标维度指标值与均值进行比较,这样将中国 MOOCs 学习者划分为八种学习者价值类型,每类学习者具有不同的学习行为特征,具体如下表 4 所示。

表4 中国 MOOCs 学习者价值分类特征

学习者类型	学习特征描述
类型 1 (ML_R↓ML_F↑ML_T↑) (0, 1, 1)	该类学习者学习频繁、学习投入较多且最近学习时间间隔短,可以判定该类学习者价值很高,学习潜力很大,是 MOOCs 课程中优质学习者类群,是 MOOCs 教学效果的重要保证。
类型 2 (ML_R↓ML_F↓ML_T↑) (0, 0, 1)	该类学习者最近学习时间间隔短、学习时间较多,但是学习频率较低,学习者价值贡献不及第一类,但是,这类学习者具有很高的潜在价值。
类型 3 (ML_R↓ML_F↑ML_T↓) (0, 1, 0)	该类学习者最近学习时间间隔较短,学习频率高,属于比较积极的学习者,但学习时间较短,这类学习者存在两种可能,一种可能是学习者自身学习能力价值有限,也可能是学习者自身学习能力很强,但是对课程内容不感兴趣,导致学习时间较少。

类型 4 (ML_R ↓ ML_F ↓ ML_T ↓) (0 0 0)	该类学习者最近学习时间间隔较短,但学习频率和学习时间的都较低,这种学习者存在两种可能,如果属于新注册学习者,那么他们可能会成为是 MOOCs 规模扩大或盈利重要学习者来源,属于重要发展类群;如果属于老注册学习者,则是无价值学习者类群。
类型 5 (ML_R ↑ ML_F ↑ ML_T ↑) (1 1 1)	该类学习者学习频率很高、学习时间也很长,但最近学习时间间隔较长。该类学习者虽然存在退课的风险,但是如果采用合适的个性化学习支持策略可能会成为 MOOCs 学习效果保障的潜在重要影响因素。
类型 6 (ML_R ↑ ML_F ↑ ML_T ↓) (1 1 0)	该类学习者学习频率较高,但长时间没有进行学习,而且学习时间较短。这类学习者可能在开课前期频繁浅尝辄止地浏览学习资源,并没有深入认真学习,并且已经很久没有进行学习。
类型 7 (ML_R ↑ ML_F ↓ ML_T ↑) (1 0 1)	该类学习者学习时间较长,但学习者的学习频率较低,并且也是很久没有进行学习。这类学习者学习存在偶然性,可能在我们分析的时间段内登录平台学习了较长时间。
类型 8 (ML_R ↑ ML_F ↓ ML_T ↓) (1 0 0)	该类学习者学习频率较低,学习时间较短,并且很久没有进行学习,这类学习者是 MOOCs 辍学率高和课程完成率的重要影响因素。

通过分析每类学习者的不同行为特征,MOOCs 课程教师或管理者能够方便有针对性对 MOOCs 学习者提供有效个性化学习支持服务,进而提高 MOOCs 教学效果。

四、基于 ML - RFT 模型的 MOOCs 学习者价值案例分析

本研究选取 365 大学平台上一门典型 xMOOC 类型课程《思想道德修养与法律基础》来对 ML - RFT 模型进行应用研究^[10]。本门案例课程于 2015 年 10 月 1 日正式开课,平台上的学习者可以在开课期间自由加入该课程学习也可以由任课教师或管理员来添加学习者。该课程建议学习时长是每周学习 3 小时,对学习者没有先修知识要求。该课程属于典型的 xMOOC 类型课程,它的教学内容包括绪论在内共九章内容,所有学习资源均为内容讲解视频,共包括 9 章 28 节 157 个视频,每章或每节后对应一个练习测试题作业,共 30 个测试作业。按照学习进度,老师会提前上传视频学习资源及作业测试题。

本研究通过课程后台导出学习者行为数据进行指标挖掘,分析时间为 2015 年 12 月 14 日,这距离开课已经有两个多月的时间,后台数据记录选课

人数共 2527 人,其中最早加入课程时间为 2015 年 10 月 14 日,最晚加入课程时间为 2015 年 12 月 9 日。

基于 ML - RFT 分类方法模型,本研究针对本门课程实际教学情况,选取视频的观看作为学习者的学习表征,将三个指标维度具体定义如下:

ML_R 最近一次学习时间:在数据分析的时候该课程还在进行中,并且不断有学习者加入进来。本研究选取学习者最近学习(观看视频)时间与数据分析时间(2015 - 12 - 14 10:30:20)之间的时间间隔来表征 ML_R。

ML_F 学习频率:本研究以天为时间单位,从学习者加入课程开始到数据分析时间(2015 - 12 - 14 10:30:20)每天的平均学习次数来表征 ML_F。

ML_T 学习时间:本研究以天为时间单位,从学习者加入课程开始到数据分析时间(2015 - 12 - 14 10:30:20)每天的平均学习时间来表征 ML_T。

确定具体指标计算方法后,本研究第一轮通过编写数据库函数来挖掘本门课程 2527 名学习者三个指标维度上的整体指标值情况,具体数据情况如下表 5 所示。

表 5 案例课程学习者整体指标情况

维度指标	最大值	最小值	均值
ML_R 最近学习间隔	1549380 秒/约 18 天	0	403122 秒/约 4.7 天
ML_F 学习频率	约为每天学习 247 次	0	约为每天学习 3 次
ML_T 学习时间	每天学习 28214 秒/约 7.8 小时	0	约为每天学习 673 秒/约 11.2 分钟

通过第一轮数据分析发现,三个指标维度最小值均为0,即学习者注册课程后并未观看视频进行学习,本研究将三个指标均为0的学习者抽取出来进行单独分析,发现他们加入课程时间集中在11和12月份,这与数据分析时间比较接近,可能学习者刚加入课程还没有来得及去观看视频进行学习,这

与实际情况比较相符。因为ML-RFT模型每个指标确定都是以学习者学习发生为前提,所以本研究在第二轮数据分析中去除三个指标均为0的学习者共1050人,这样得到有效学习者样本共1477人。去除无关干扰样本后的学习者三个指标维度上整体指标情况如下表6所示。

表6 案例课程去除无关样本后整体指标情况

维度指标	最大值	最小值	均值
ML_R 最近学习间隔	1549380秒/约18天	38133秒/约10小时	654142秒/约7.6天
ML_F 学习频率	约为每天学习247次	约为每天学习0.03次	约为每天学习5.2次
ML_T 学习时间	每天学习28214秒/约7.8小时	约为每天学习0.26秒	约为每天学习1113秒/约18.5分

在第一轮数据处理分析的基础上,本研究将每个学习者单个指标与总体均值进行比较,如果大于等于均值将指标在数据库中标记为1,小于均值记为0,这

样得到学习者(1,1,1)、(1,1,0)……(0,0,0)共8类表征形式。通过第二轮数据分析,得到该课程学习者的具体分类,分类结果参见下表7和图1所示。

表7 案例课程学习者分类结果

学习者类型	人数	所占比例
类型1 (ML_R ↓ ML_F ↑ ML_T ↑) (0,1,1)	220	14.9%
类型2 (ML_R ↓ ML_F ↓ ML_T ↑) (0,0,1)	52	3.52%
类型3 (ML_R ↓ ML_F ↑ ML_T ↓) (0,1,0)	48	3.25%
类型4 (ML_R ↓ ML_F ↓ ML_T ↓) (0,0,0)	469	31.75%
类型5 (ML_R ↑ ML_F ↑ ML_T ↑) (1,1,1)	53	3.59%
类型6 (ML_R ↑ ML_F ↑ ML_T ↓) (1,1,0)	40	2.71%
类型7 (ML_R ↑ ML_F ↓ ML_T ↑) (1,0,1)	46	3.11%
类型8 (ML_R ↑ ML_F ↓ ML_T ↓) (1,0,0)	549	37.17%

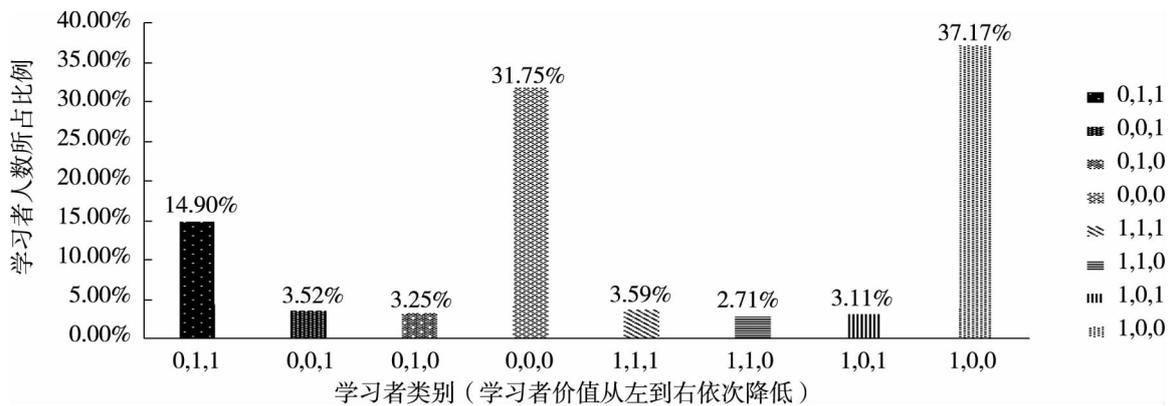


图1 案例课程学习者类别分布柱状图

在对学习者分类基础上,本研究针对两个极端价值类群最高价值类群(0,1,1)和最低价值类群

(1,0,0)分别进行三个指标值抽取分析,得到最高价值类群学习者整体指标情况如下表8所示:

表8 最高价值类群(0,1,1)学习者整体指标情况表

维度指标	最大值	最小值	均值
ML_R 最近学习间隔	653396秒/约7.56天	38133秒/约10小时	305947秒/约3.5天
ML_F 学习频率	约为每天学习247次	约为每天学习5.22次	约为每天学习20.3次
ML_T 学习时间	每天学习28214秒/约7.8小时	约为每天学习1116.7秒/约0.3小时	约为每天学习4722秒/约1.3小时

最低价值类群学习者情况表如下表 9 所示:

表 9 最低价值类群(1 0 0) 学习者整体指标情况表

维度指标	最大值	最小值	均值
ML_R 最近学习间隔	1549380 秒/约 18 天	655301 秒/约 7.6 天	654142 秒/约 12 天
ML_F 学习频率	约为每天学习 5.2 次	约为每天学习 0.03 次	约为每天学习 1.13 次
ML_T 学习时间	每天学习 1109 秒/约 18.5 分	约为每天学习 0.26 秒	约为每天学习 229 秒/约 3.8 分

五、结果讨论

通过案例分析结果可以看出,类型 8(1 0 0) 类群学习者占最大比重,达到 37.17%,其次是类型 4(0 0 0) 类群学习者,达到 31.75%。并且类型 4(0 0 0) 类型学习者如果考虑注册时间因素,区分一下新老注册学习者,其中可能会有一部分归类到类型 8(1 0 0) 类群中。而我们看到学习者价值最高的需要重要保持的学习者类群(0 1 1) 比例达到 14.9%,这是人数排名第三的类群。剩下五类学习者人数一共才占 16.18%,并且他们各自占比大都在 3% 左右。通过两个极点类群学习者整体指标分析,可以看出最有价值学习者在学习间隔、学习频率、学习时间上与最无价值学习者有显著差异。综合上述结果,可以看出本研究案例课程中学习者具有以下特点:

1. 学习者价值类别人数分布呈现两极分化的趋势。本案例中最有价值和无价值学习者类群占绝大多数,而中间一般价值学习者类群人数仅占很少的部分。

2. 不通学习者类型之间可以相互转化。本案例中,需要通过分析学习者注册课程时间来进一步判断学习者类别的第四类学习者占 31.75%,说明学习者类型之间相互转化的潜力很大,针对价值较低的学习者提供适当的学习策略会实现学习者价值的提升转变,从而有效提升 MOOCs 学习效果。

3. 各学习者价值类别群体学习者学习行为差异显著。通过实际追踪两类学习者的群体学习间隔、学习频率和学习时间,可以看出不同类别学习者行为差异显著,应用该模型方法可以对学习者价值进行有效区分。

本研究认为针对 8 种价值类群学习者,MOOCs 课程教师或者管理员可以依据对学习效果的影响重要程度具体再将学习者聚类为:重要发展类群、一般发展类群、一般保持类群、重要挽留类群、一般挽留类群和一般价值类群共 6 类。针对每一类群级别可以有针对性地提供个性化的学习支持策略,具体学习者价值类群及策略建议参见下表 10 所示。

表 10 MOOCs 学习者价值类群及学习策略建议

学习者价值类型	策略建议
类型 1(0 1 1) (ML_R ↓ ML_F ↑ ML_T ↑) 重要发展类群	该类学习者属于 MOOCs 中需要重要保持发展的类群。该类学习者是最有价值的类群,继续维持与这类学习者的关系是课程成功的重要保障。
类型 2(0 0 1) (ML_R ↓ ML_F ↓ ML_T ↑) 一般发展类群	该类学习者属于 MOOCs 中需要一般发展的类群。分析、了解然后满足该类学习者的需求,利用针对性的学习策略来吸引他们,提高他们的学习频率,就会实现向重要保持类群的转变,进而大大提高课程学习效果。
类型 3(0 1 0) (ML_R ↓ ML_F ↑ ML_T ↓) 一般发展类群	该类学习者属于 MOOCs 中需要一般发展的类群。对该类学习者需要进一步了解他们的实际学习可能,确定登录次数高,学习时间短的具体原因,这样再针对性进行学习支持,适当维持与这些学习者的关系对于提高课程学习效果会提供一定的帮助。
类型 4(0 0 0) (ML_R ↓ ML_F ↓ ML_T ↓) 一般保持类群	该类学习者在 MOOCs 中一般价值类群,它存在两种归类可能,如果是新注册学习者,可能学习者还没来得及观看学习资源,后续演变则很可能成为重要发展类群,如果是老注册学习者,则表明该学习者只是最近偶然学习了一次,很有可能属于无价值学习者类群。对于该类学习者类群需要进一步使用注册时间细分,确定之后再针对性地提供教学策略。

类型 5(1, 1, 1) (ML_R ↑ ML_F ↑ ML_T ↑) 重要挽留类群	该类学习者属于 MOOCs 中需要重要挽留的类群。对于该类学习者, 或者因为资源已经学完, 或者恰巧只是最近不进行学习, 但是总体学习频率和学习时间较高, 至少前期学习价值很大, MOOCs 教师或管理者应该尽量挽留, 可通过适当的营销手段(比如定期邮件或短信提醒)提高该类学习者的学习忠诚度, 或推荐一些相关课程, 让学习者继续参加 MOOCs 学习。
类型 6(1, 1, 0) (ML_R ↑ ML_F ↑ ML_T ↓) 一般挽留类群	该类学习者属于 MOOCs 中一般挽留类群。这类学习者的学习行为与远程在线教育的本质 - 教与学的时空分离有关, 受学习者自身影响很大, 登陆次数很多, 但是学习时间很短, 可以从提高学习者自主学习能力的角度进行教学策略设计。
类型 7(1, 0, 1) (ML_R ↑ ML_F ↓ ML_T ↑) 一般挽留类群	这类学习者属于 MOOCs 中的一般挽留类群。该类学习者学习的发生存在偶然性, 受学习者自身影响很大, 可以从提高学习者自主学习能力的角度进行教学策略设计。
类型 8(1, 0, 0) (ML_R ↑ ML_F ↓ ML_T ↓) 一般价值类群	这类学习者属于 MOOCs 中一般价值类群。该类学习者会降低课程完成率, MOOCs 管理者或教师可以考虑这类学习者到底是不是适合该课程的学习, 甚至可以实施劝退的策略建议。

六、总结

本研究的主要贡献在于为 MOOCs 学习者分类提供了一个新的思路, 使用本研究提出的 ML - RFT 模型, 能够有效地将 MOOCs 学习者进行自动化的分类, 分类后能为教育机构的个性化学习干预提供可能。本研究虽然仅仅用一门课程的数据做了分析, 但研究并非为了固化特定的分类数据标准, 而是提出一套标准化的模型以支持适应性的学习者价值分类。未来, 本研究会从以下两个方面进行扩展研究: 一是基于 ML - RFT 价值分类模型进行特定类别学习者时间序列追踪, 发现学习者价值类别演化规律。二是进行更多案例课程的研究, 细致分析每类学习者学习行为特征, 并尝试建立学习者分类特征与个性化学习支持服务之间的对应关系模型。

MOOCs 为中国高等教育变革提供了契机, 高质量的学习资源和适应性的学习支持服务为中国新型高等教育体系的构建提供了可能。在以 MOOCs 为代表的在线学习环境中, 只有真正挖掘出有价值的学习者, 进而有针对性地对提供个性化适应性的学习支持服务, 才能有效提高在线学习质量。

【参考文献】

- [1][6]郑勤华,李秋菊,陈丽. 中国 MOOCs 教学模式调查研究[J]. 开放教育研究, 2015(6): 71 - 79.
- [2]罗晔,李阿利. MOOC 高辍学率原因及对策分析[J]. 学理论, 2015(5): 139 - 140.
- [3]郝智慧. 基于个性化学习的远程学习支持服务研究[J]. 知识经济, 2015(17): 99 - 100.
- [4]冯晓英,史颖,张传俭,陈丽. 面向个性化学习支持服务的开放大学学生模型研究[J]. 北京广播电视大学学报, 2015(2): 19 - 25.
- [5]刘冕. 学习价值问题研究[D]. 东北师范大学, 2006.
- [7]王渊. 基于 RFM 模型的协同过滤方法及其在个性化推荐中的应用[D]. 杭州电子科技大学, 2014.
- [8]Chang H C. Developing EL - RFM model for quantification learner's learning behavior in distance learning [C]. // Education Technology and Computer (ICETC), 2010 2nd International Conference on. IEEE, 2010: V4 - 452 - V4 - 454.
- [9]Hughes A M. Strategic Database Marketing [J]. Butterworthheinmann Oxford April, 2000.
- [10] http://xuetang.365dx.com/courses/course-v1:365DX+10610183X+2015_T2/about.

(本文责任编辑: 闫兵)