

影响MOOC学习质量的关键因素探讨

——基于学习者行为视角

李 丽, 张芳芳, 刘 源, 蚁文洁

(深圳大学管理学院, 广东 深圳 518060)

【摘要】 本文在深圳大学MOOC——“礼行天下”的相关数据的基础上,提取了反映学习者行为的关键数据。采用K-Means聚类法将学习者分为3类:只是注册者、学习热爱者、狂热学习者。并对聚类结果的有效性、稳健性及现实意义进行了检验。然后以这3类分类结果为基础,通过统计图表,对学习者的行为特征进行了描述性分析。最后,通过回归分析的方法,对影响MOOC学习质量的关键因素进行了探讨。结果显示:不注重平时课程测验、不经常访问学习页面、不积极参加互动讨论现象是影响MOOC学习质量的关键因素。其中,平时课程测验的影响最为显著。在此研究基础上,本文提出了提高MOOC学习质量的建议,为MOOC团队建设更好的课程提供借鉴。

【关键词】 MOOC;学习者行为;K-Means聚类法;回归分析

【中图分类号】 G642 **【文献标识码】** A **【文章编号】** 2095-5065 (2016) 08-0042-11

1 前言

1.1 研究背景

随着互联网的发展,在线课程因其优越的特性,受到了越来越多学习者的关注。大型开放式网络课程——MOOC (Massive Open Online Course),广泛活跃在各大网络课程平台上。

2012年5月,哈佛大学宣布与麻省理工学院合作成立非营利性组织edX,并向世界各地的名牌大学发出邀请,共同在开源的平台上提供优质开放的网络课程。一时间,各大MOOC平台兴起。国外以Coursera、edX、Udacity为代表,国内则有清华大学学堂在线、爱课程网中国大学MOOC、智慧树等。由于MOOC具有开放性、大规模参与、高品质、高灵活性等特点,加之各大平台的大力推广,近几年MOOC的影响力愈发深远。

和传统课程教育相比,从时间和空间上来看,MOOC学习受到的限制很小。MOOC学习者可以在任意时间、任意地点通过网络来实时学习课程,时间和空间上都具有很大的便利性;从课程规模来看,一门好的MOOC,注册人数多达上万人次,甚至会更多,远远超过传统课程;从学习者的学习动机来看,MOOC学习者的学习动机更加多元化,当

收稿日期:2016-6-27

作者简介:李丽(1962—),女,吉林长春人,博士,教授,研究方向为智能管理与决策;

张芳芳(1990—),女,河南洛阳人,硕士研究生在读,研究方向为智能管理与决策;

刘源(1994—),男,浙江舟山人,本科在读,研究方向为电子商务;

蚁文洁(1993—),女,广汕头人,本科在读,研究方向为电子商务。

然这也和MOOC学习者在年龄、教育背景、社会背景等方面没有严格的限制有关。

总之, MOOC来源于传统课程, 又有所超越。它的出现, 不仅为更多的学习者提供了丰富的教育资源, 而且提高了学习者的学习品质。这样一种创举, 引发了社会教育业发展的变革。

1.2 研究问题及意义

MOOC发展虽然迅速, 但在实际运行中, 还是存在着诸多问题。相比于传统课程, MOOC较少得到权威高校的学分认证或是行业认可, 学习者多数只能依兴趣爱好驱动上课的热情, 加之周围缺少监督, 导致课程完成率很低。维基百科(2013)上有数据统计, Coursera的课程完成率仅为7%~9%; 国内各主流MOOC平台的数据也显示, 大部分课程的完成率都不足10%。“选课人多、听课人少、合格率低”成了MOOC一个普遍的难题。

为了进一步提高MOOC质量, 改善学习者的学习效果, 本文以提高MOOC学习质量为出发点, 通过对学习行为的研究, 找到了影响MOOC质量的关键因素。

MOOC借助于互联网, 精确地记录了各学习者平时的学习行为。利用好这些数据, 不仅有助于把握学习者的学习行为情况, 还能为MOOC的设计提供借鉴。

1.3 研究数据与方法

以往的相关研究多数集中在实证分析上, 利用各平台上课程的数据对学习行为进行的研究较少。本文所涉及的数据, 主要来自深圳大学开设的“礼行天下”这门MOOC。该课程数据涉及全国共39所大学的588位学习者的学习信息。以此数据为基础, 本文首先采用K-Means聚类法对学习行为进行分类, 然后以得到的分类结果为基础, 选取几个典型的学习者行为进行描述性分析, 最后运用多元回归法对学习效果和学习行为之间的关系进行了分析探讨。

2 文献回顾

关于MOOC的介绍, 最早出现在2012年的《科学》杂志上, Stein预测了MOOC将对未来教育模式产生重大影响^[1]; 2013年, 《自然》杂志上出现了另一篇详细介绍MOOC的发展、现状和趋势的文章^[2]; 2015年, SIGKDD大会成功举办了一场以“预测MOOC中学习者退课行为”为主题的数据挖掘竞赛, 并对知识发现和数据挖掘等问题作了探讨。国内高校的MOOC在2013年9月首次亮相。

近几年来, MOOC的发展虽然已经取得了很大的成就, 但是由于MOOC发展时间尚短, 相关的课程设计、考核体制、奖惩措施、管理方案以及研究工作尚不成熟。目前国内的相关课程报告尚未呈体系, 即使从世界范围来看, MOOC的相关研究成果也不够完善。目前, 基于MOOC平台的数据对学习行为的研究尚不多, 已有的成果大致可归为以下五大类。

2.1 对学习行为分类的研究

蒋卓轩等^[3]根据学习者的行为和特点, 将不同学习者归为“打酱油者(Bystander)”“虎头蛇尾者(Anti-climax)”“始终摇摆者(Waverer)”“坚定完成者(Stalwart)”“只为学习者(Just-learner)”5类; Kizilcec等^[4]运用聚类分析法将3门MOOC的注册学习者分为完成者、旁听者、脱离者、尝鲜者4类, 并对模型的稳健性进行了检验; 教育技术咨询专家Hill等^[5]将MOOC学习者分成5种类型: 爽约者、旁观者、顺便访问者、被动参与者、主动参与者。

2.2 对学习行为选课、退课情况的分析

金娣^[6]以Coursera上的一门课程为例, 对学生的流失情况进行了调查, 对影响MOOC学习完成率的因素进行了探讨, 并从课程管理和学习者自我管理两方面提出了改进建议; 吴江等^[7]从选课、退课、课程参与行为及成绩4个方面对爱课程平台上的一门课程的学习者行为数据进行了统

计分析,对MOOC学习者的行为特点及行为背后的原因进行了探讨;Yang D等^[8]建立了一个“存留者”模型,对MOOC学习者的退课行为进行了分析,并结合标准的社交网络分析技术及学生参与论坛讨论的情况,找出了影响学生退课的关键因素。

2.3 对学习课程参与情况的分析

贾积有等^[9]通过文本分析并建立计量模型,研究了影响课程完成度的因素,建立了线性回归模型;Reeve^[10]提出了在学习过程中相互关联的4个维度,即行为参与度、情感参与度、认知参与度和代理参与度;潘丽佳^[11]在问卷分析的基础上,对学习参与者在MOOC属性与学习绩效之间是否会起到特定的中介作用作了探讨;李曼丽等^[12]运用了Tobit和Logit两个定量分析模型,分别对MOOC学习者的课程参与情况和课程完成情况作了分析,发现学习动机越强的学生在论坛中越活跃,注册越早的学生越有可能完成课程,个体课程参与度对其他学习者有明显的正向影响。

2.4 建立模型对最终学习效果的预测

Taylor等^[13]建立了一个逻辑回归模型,运用最大似然估计法结合高斯—牛顿法,对MOOC学习情况作了预测,并运用决策树作了分析;蒋卓轩等^[3]运用了线性判别分析、逻辑回归和线性核支持向量机这3种分类模型,对时序变化作了分析;Girish等^[14]建立了一个隐式马尔可夫链模型,来识别学生的行为模式,以预测学习者退课的情况,并作适当的干预以期减少退课情况的发生。

2.5 基于学习者实际生活背景差异的分析

Nesterko S等^[15]运用HarvardX上的18门课程的数据,对处于不同地理位置的学习者注册课程、课程完成率等情况作了分析,并结合当地的实际人口状况,形成了具有相似特点的聚类;薛宇飞等^[16]对不同文化背景的学习者的共性和特质进行了深层次的分析,选取了edX平台上来自美国、印度、英国的5001名学生的数据,勾画了跨文化背景的学习行为特征,比较了不同文化背景

学习者在MOOC学习上的异同。

3 MOOC学习者行为研究

3.1 学习者行为的定义

吴江等^[7]提出,MOOC学习者在MOOC学习过程中主要有选课、课程参与(包括观看视频、提交测验与作业、参与讨论、参加考试)、退课等行为。蒋卓轩等^[3]则从学习者参与的角度,主要选取了“看视频”和“提交测验”这两个基本的学习行为作为学习者分类的依据,同时对学习者参与课程的时间、成绩和论坛的相关行为作了分析。

本文结合以往的研究成果及相关的的生活经验,将MOOC学习者行为定义为,用户从了解某一MOOC开始,直到该门MOOC结束,所有的和该课程有关的行为,包括课程注册、课程参与、退课等。

本文主要选取课程参与过程中的几个典型学习行为:观看视频时长、访问学习页面数、论坛讨论数和任务完成率,同时对期末考试成绩和综合成绩作了相关分析。

3.2 学习者分类

传统的教学理念,会按照学习者的成绩把学习者分成优、良、中、差4种。基于MOOC学习者的不同生活环境、不同专业背景、不同年龄层次等特点,这种简单的按照成绩的高低来区分学习者的方法,很难体现学习者的真实学习情况和学习水平的高低。

MOOC具有开放性、规模性、无地域限制等特点,任何人在任何地点都可以轻松参与MOOC的学习。对于学习者而言,其学习动机、学习时间、学习背景、学习状态以及学习目标等方面都存在着一定的差异,导致了其学习行为也会大相径庭。

因而,我们在对学习进行分类时,就需要综合多方面的指标,制定合理的分类评价体系,

综合考虑学习者多方面的实际情况。

Kizilcec等^[4]运用K-Means聚类法,对MOOC学习者的行为进行了研究,探讨了MOOC学习者分类问题。对于MOOC学习者分类的问题,需要对聚类结果的有效性、稳健性和现实意义这3个方面作评估。基于以上准则,本文首先对聚类标准进行了分析,为之后对学习者的行为分类奠定了基础。

(1) 聚类的有效性。聚类轮廓系数是评价聚类准确性的有效方法。如果某样本的轮廓系数为正,则说明对该样本现行的分类方法是最优的。轮廓系数的取值范围为-1~1。若某个类别的轮廓系数取1,则说明该类里面的样本全是相同的。 K 是K-Means聚类方法中的分类参数。对比 K 取4、5、6时的情况, K 取3时的轮廓系数最好。本文将588位学习者归为3类,平均轮廓系数为0.765,有99.32%的样本的轮廓系数为正,说明聚类结果较好。轮廓系数详见表1。

表1 轮廓系数统计

聚 类	统 计			
	个 案 数 目	平 均 值	最 小 值	最 大 值
1	297.000	0.896	0.123	0.945
2	31.000	0.294	-0.076	0.508
3	260.000	0.673	-0.018	0.803
总 计	588.000	0.765	-0.076	0.945

(2) 聚类的稳健性。在运用K-Means聚类法进行聚类时,由于 K 值是人为设定的(通常根据实际经验来确定),具有一定的模糊性,因而有必要对其稳健性进行检验。将 K 分别取3、4、5、6进行测试,结果显示:虽然聚类的数目增加了,但在聚类结果中,第一类别的数目,一直稳定在296左右,第二、三类别中的样本数随着聚类数目的增加,而逐渐被拆分。例如, K 由3变为4时,第二类学习者就会被进一步细分为两类,其差别主要是观看视频时长。检验表明,对该数据运用K-Means聚类法,具有较好的稳健性,即某些参数的变动,不会对整体的聚类趋势产生较大的影响。

(3) 现实意义。聚类结果需要有一定的现实意义,能反映出不同类别学习者学习行为的特点。各类学习者学习行为的特点,如表2所示。本文根据学习者的行为特征,将学习者的学习行为分为如下3类。

表2 最终聚类中心及聚类结果

项 目	聚 类		
	1	2	3
视频观看时长	30.33	1099.17	581.69
讨论数	0.24	16.35	9.52
访问数	9.16	283.19	127.38
任务完成率	0.12	0.94	0.92
个案数目	297.00	31.00	260.00

只是注册者(Only Register, OR)。这类学习者的学习行为偏少,平均仅有9次的学习页面访问数和12%的任务完成率,观看视频时间、参与讨论也偏少。这类学习者在注册后,几乎没有任何学习行为。此类学习者人数最多,占有所有学习者的50.51%。

学习热爱者(Crazy for Knowledge, CFK)。这类学习者具有看视频时间很长(平均超过1000分钟)、任务完成率高、论坛讨论活跃、访问学习页面的次数多等特点。该类学习者除了完成课程指定的学习任务外,还花了很多课外时间观看视频、参与讨论等。但是,该类学员占比很低,仅5.27%。

达标学习者(Up to Standard, UTS)。这类学习者的任务完成率很高,但相比学习热爱者而言,其自主学习投入时间较少。这类学习者多以完成课程学习任务为目标,在达到了课程考核标准后,就很少花费时间再进行学习。这类学员所占比例较大,有44.22%。

3.3 学习者行为分析

MOOC学习者的学习行为众多,主要的学习行为有选课、课程参与(例如,视频观看、作业提交、学习页面的浏览、完成随堂测验等)、退课等。本文主要从课程参与的角度,选取观看

视频时长、论坛讨论、学习页面访问情况和任务完成率,这几个具有代表性且容易得到数据的指标,来反映学习者行为的特征。

(1) 观看视频时长。MOOC以网络为媒介,以视频为载体,将各课程通过网络视频呈现给广大学生。无论是课程的讲解、探讨,还是作业的分析,都离不开网络视频。因而,观看学习视频的情况,成为了分析MOOC学习者学习行为的一个重要方面。观看视频的长短、频率以及时间段,直接反映了学习者的实际学习状况。

由图1可见,有近一半的学员观看视频时间为零,学习者以600分钟左右的观看时长居多,超过1000分钟的很少。平均观看视频时长为330.48分钟,标准差为338.355,数据波动大,表明学习者之间有明显的差异。在3类学习者中,“只是注册者”大多观看视频时长为零,“学习热爱者”观看时间则大多超过了800分钟,“达标学习者”则以582分钟为均值,90.91为标准差,近似呈现出正态分布情况。若除去观看视频时长为零的那部分数据,在参与了视频学习的这部分学员中,观看视频时长小于500分钟的并不多,大多数学员还是按要求完成了指定的视频观看时间。这一方面反映出,有不少学生是从头至尾的“打酱油

者”,完全没有参与到课程中来;但另一方面,若参与了课程学习,大多数学生还是趋向于达到课程的要求,看视频半途而废的情况并不多。

(2) 论坛讨论。论坛交流是学习者课后交流学习情况,分享学习心得,解答疑难的重要途径。参与论坛讨论是学习者的自主行为。当然,课程设计团队也可以把参与论坛的情况与综合成绩挂钩。在本文研究的课程中,论坛的讨论情况和综合成绩挂钩。结合前面用到分类方法,将各类学习者的论坛讨论情况反映到了频率分布直方图中,如图2所示。可以发现,实际参与讨论的人数不到注册人数的1/3,论坛参与度还不够高。其中,“只是注册者”完全没有参与论坛,“达标学习者”仅有很少一部分没参与论坛,“狂热学习者”都参与了。在参与论坛讨论的学习者中,“达标学习者”和“狂热学习者”较活跃,大多参与了10次以上。另外,参与论坛的次数超过40以上的很少。

(3) 访问学习页面数。由于MOOC的开放性、便利性,以及学习时间的灵活性,学习者可根据自身实际情况来学习课程。图3反映了该课程在开课期间,学习者各时段的学习页面访问情况。可以发现,对于学习时间而言,晚上8点到12

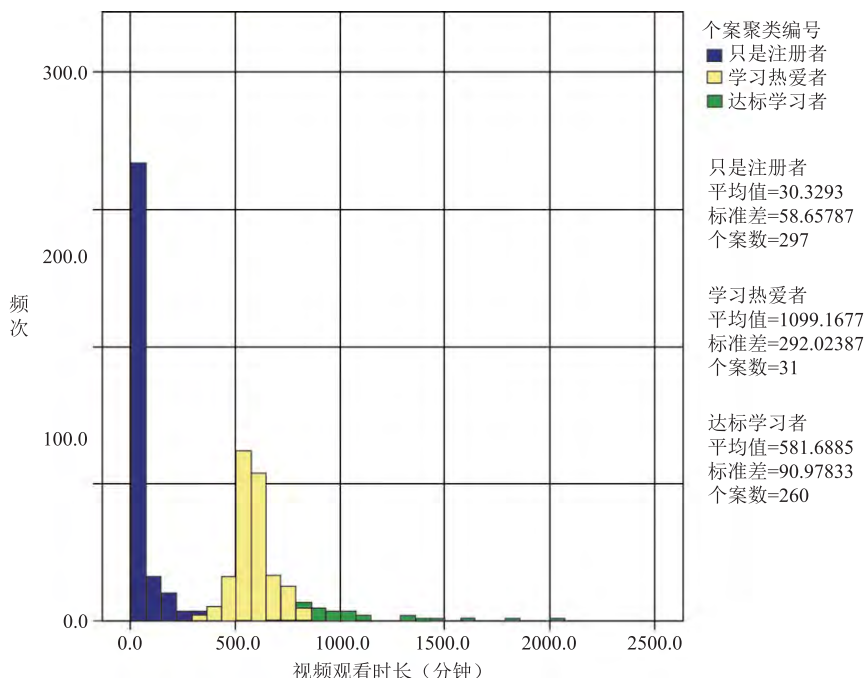


图1 各类别学习者观看视频时长直方图

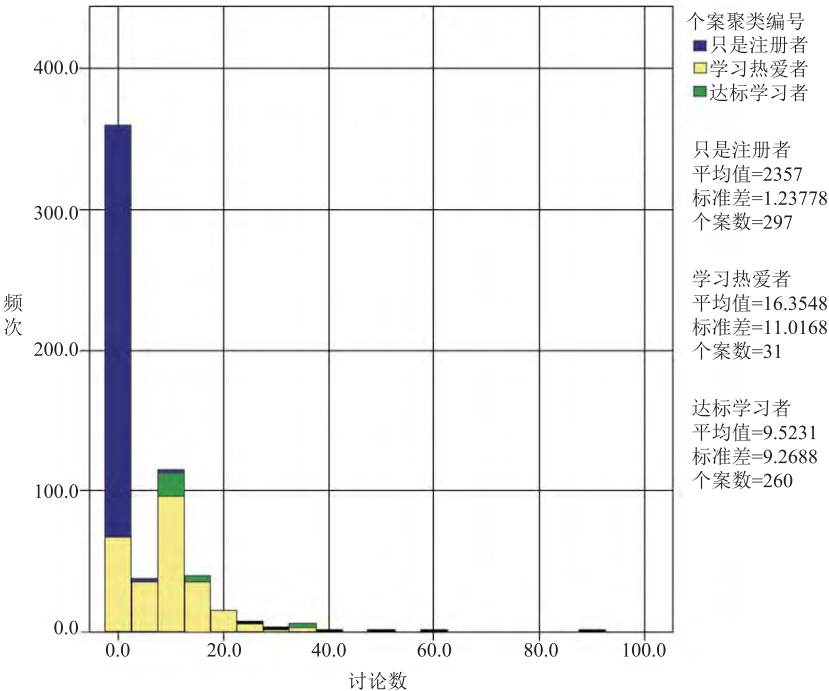


图2 “礼行天下”课程各类学员讨论情况分布直方图

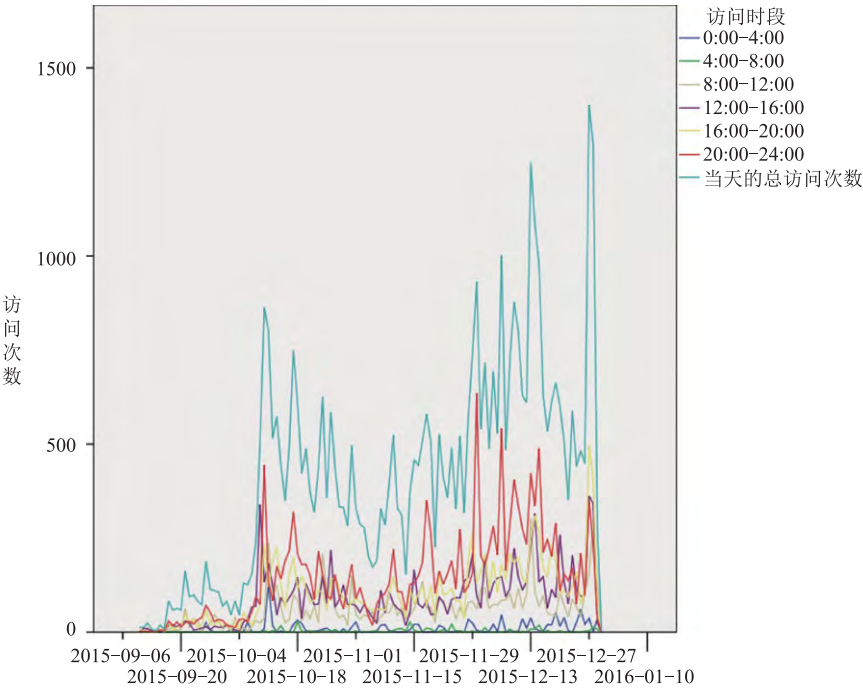


图3 各时段学员学习页面访问次数情况折线图

点是大多数学员选择的学习时间,凌晨12点到早上8点这个时间段进行学习的较少,其他的时间段较均等;对于访问量而言,在开课的初期,即9月份到10月中旬,这段时间学习者对学习网站的访问量较少。从10月份中旬到11月初,这段时间的访问量骤增,可能是由于章节考核测试时间的限制引起的。整个11月份的访问量较平稳,比10月下旬略低,但高于开课初期;12月份,进入期末考试时间,访问量进入了高潮。显然,学习页面的访问量受考试的影响程度较大。

(4) 任务完成情况。图4反映了3类学习者的任务完成率情况。“只是注册者”这部分人群最多,大多数人任务完成率不超过30%,多数为零,但也有很少一部分完成率达到了100%。这类学员平均任务完成率为12.34%,标准差为0.2545,较大,这反映出学员存在较大差异。

“学习热爱者”大多任务完成率为100%,但由于这部分人群较少,仅31人,个别任务完成率较低的会导致整体的标准差较大。“达标学习者”平均任务完成率较高,有91.54%,但也有少部分的完成率很低,存在一定差异。

总体而言,3类学习者中,任务完成率为零和100%的这两类最多,居于中间的不多,一方面反映出有不少学员在注册了课程后,完全就没有再参与了;另一方面也反映出一旦学习者参与了部分课程的学习,就很少会半途而废。

4 成绩

学习成绩是学习效果最直观反映。期末考试成绩反映了学生对课程知识的掌握情况,综合成绩则涵盖了平时成绩和期末成绩(期末成绩占40%),可以综合反映出学习者的课程参与度、学习状态、学习效果等情况。因而,对这两个指标的分析,有助于我们从多维度来了解MOOC学习者的学习状况。

对比图5和图6可以发现,期末考试成绩和综合成绩两者的趋势及分布大致一致,但也有些出入。比如,高分段人群中,有的学习者虽然期末考试成绩较低,但是综合成绩较高。这反映出有不少学习者平时表现比期末考试成绩更优秀。

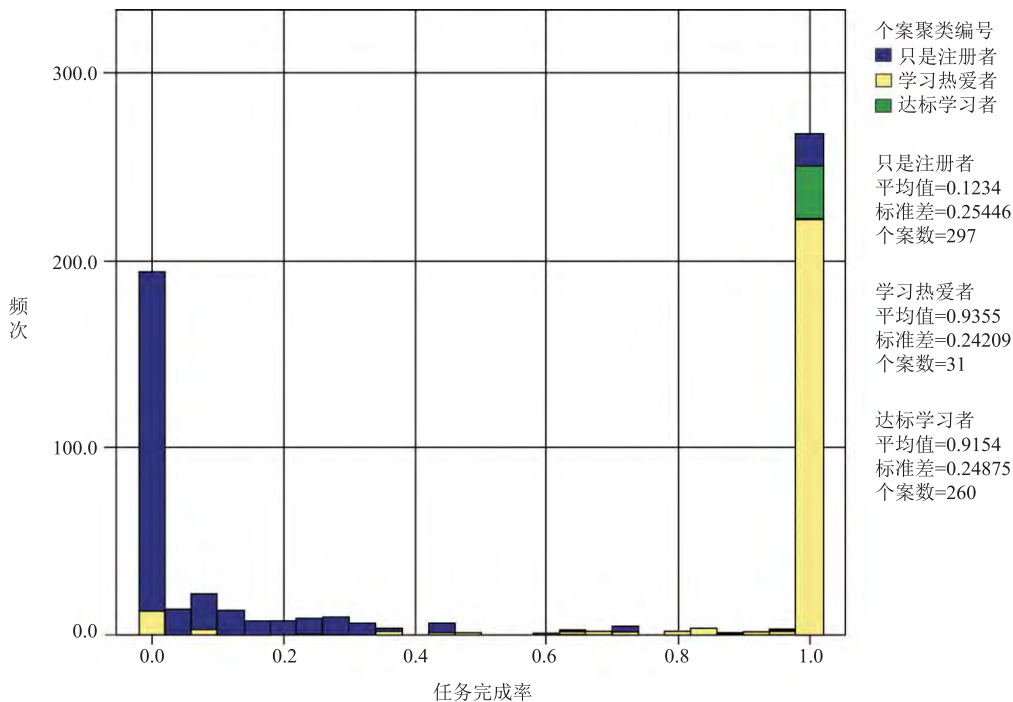


图4 各类学习者任务完成率情况

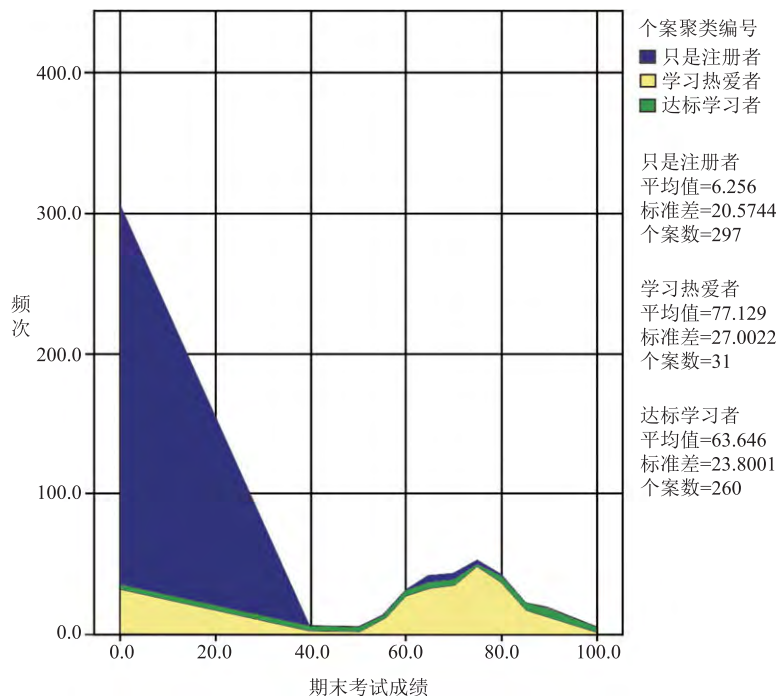


图5 期末考试成绩情况

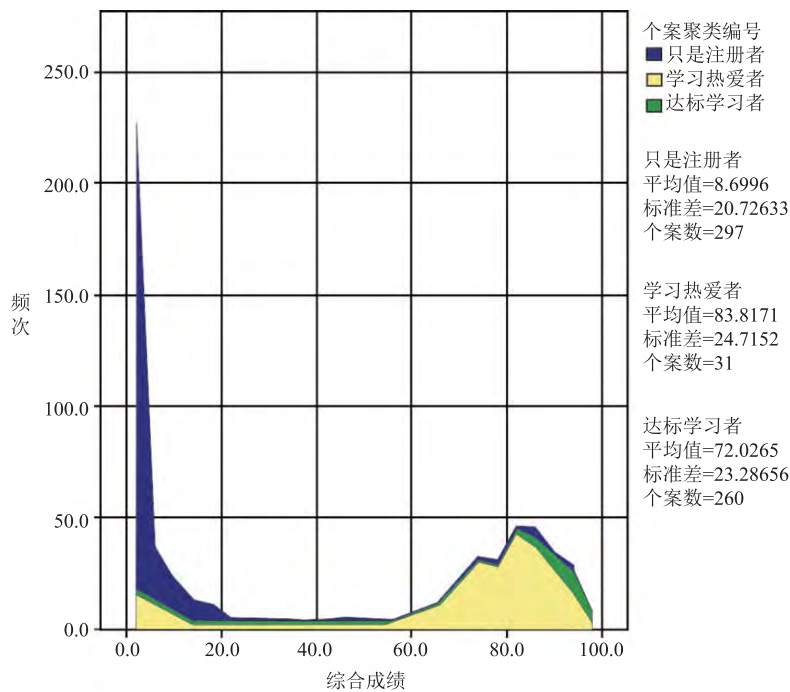


图6 综合成绩情况

另外,就各类学员来看,“学习热爱者”无论是在期末考试中,还是综合成绩评定中,都取得了相当优秀的成绩,且两者的出入不大;“达标学习者”的两项成绩大多都在60分以上,但也有一部分学习者成绩不理想,可能是对课程的得分不够重视导致,另外,“达标学习者”期末考试的成绩不如平时成绩优异;“只是注册者”则无论是期末考试成绩还是综合成绩都没有及格,由于平时成绩几乎为零,故期末考试的成绩要优于综合成绩。

5 多元回归分析

5.1 模型构建

本节将根据前面的内容和结论,试图从学习行为和学习成绩之间寻找一个定量化的关系。基于前面对学习行为特征的分析以及实际数据的获取情况,本文主要选取视频得分(X_1)、讨论得分(X_2)、作业得分(X_3)、访问数得分(X_4)、课程平时测验得分(X_5)、任务完成率(X_6)为解释变量,选取期末考试成绩(Y)为被解释变量。

由于在回归过程中发现解释变量之间有严重的多重共线性,故采用逐步回归法,剔除变量,最终确定的多元回归模型如下:

$$\begin{aligned} Y = & -0.684400 + 0.867199X_2 + 0.511568X_4 + 2.701398X_5 \\ & (0.325278) (0.188271) (0.245842) (0.096571) \\ t = & (-2.104048) (4.606111) (2.080885) (27.97330) \\ R = & 0.873334, \bar{R}^2 = 0.872683, F = 1342.180 \end{aligned}$$

5.2 现实意义检验

模型估计结果说明,在假定其他变量不变的情况下,讨论得分多得1分,平均来说期末考试成绩将增加0.87分;访问数得分多得1分,平均来说期末考试成绩将增加0.51分;课程测验得分多得1分,平均来说期末考试成绩将增加2.70分。这与现实经验判断具有一致性。

5.3 统计检验

(1) 拟合优度:由表3可知, $R^2=0.873334$,修正后的可决系数 $\bar{R}^2=0.872683$,这说明模型对样本拟合得较好。

(2) F检验:针对 $H_0: \beta_2=\beta_4=\beta_5=0$,给定显著性水平 $\alpha=0.05$,在F分布表中查出自由度为 $k-1=3$ 和 $n-k=584$ 的临界值 $F_{\alpha}(3,584)=2.6$ 。由表3得 $F=1342.180 > F_{\alpha}(3,584)$,应拒绝原假设,说明回归方程显著,即“讨论得分”“访问数得分”“课程测验得分”等变量联合起来确实对“期末考试成绩”有显著影响。

(3) t检验:分别针对 $H_0: \alpha_j=0, (j=2,4,5)$,给定0.05显著性水平,查t分布表得自由度为 $n-k=584$ 的临界值为 $t_{\frac{\alpha}{2}}(n-k)=1.960$ 。由表3数据可知, α_2 、 α_4 、 α_5 对应的t统计量分别为4.606111、2.080885、27.97330,其绝对值都大于1.960,这说明在0.05显著性水平下,分别都应当拒绝 $H_0: \alpha_j=0, (j=2,4,5)$ 。也就是说,在保持其他变量不变的情况下,解释变量“讨论得分”“访问数得分”“课程测验得分”都对被解释变量“期末考试成绩”有显著影响,OLS回归结果如表3所示。

5.4 结果分析

通过对“讨论得分”“访问数得分”“课程测验得分”与被解释变量“期末考试成绩”建立多元线性回归模型,并通过现实意义以及统计的检验验证了其合理性。从模型中可得出“课程测验得分”对“期末考试成绩”的影响更显著,而“讨论得分”“访问数得分”对“期末考试成绩”影响相对较小。这说明,平时进行的小测验能够有效地促进学习者进行自主学习,在未来进行MOOC设计或者学习时,应注意平时测试的重要性。

6 结论与建议

6.1 结论

本文通过对“礼行天下”这门MOOC学习者

表3 OLS回归结果

Dependent Variable : Y
Method : Least Squares
Date : 03/09/16 Time : 20:43
Sample : 588
Included observations : 588

Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	-0.684400	0.325278	-2.104048	0.0358
X ₂	0.867199	0.188271	4.606111	0.0000
X ₄	0.511568	0.245842	2.080885	0.0379
X ₅	2.701398	0.096571	27.97330	0.0000
R-squared	0.873334	Mean dependent var		14.14762
Adjusted R-squared	0.872683	S.D.dependent var		15.07061
S.E. of regression	5.377421	Akaike info criterion		6.209074
Sum squared resid	16887.33	Schwarz criterion		6.238848
Log likelihood	-1821.468	Hannan-Quinn criter.		6.220675
F-statistic	1342.180	Durbin-Watson stat		1.916053
Prob (F-statistic)	0.000000			

学习数据的统计分析，探讨了MOOC学习者的行为特征及其发生的原因。

(1) 分析表明，现在的学习者虽然倾向于通过MOOC进行学习，参与MOOC学习的人数较多，但是很多学习者往往很难坚持下来，“选而不学”、不注重平时成绩、不积极参加互动讨论等现象仍比较严重。

(2) 数据结果显示，有超过一半的学习者属于“只是注册者”，几乎没有参与到课程中来；而参与到课程中的学员，大多都在朝着完成课程学习要求的方向努力，中途荒废的学习者并不多。因而，实时提醒学习者参与课程（例如，每周发课程学习状况的邮件），督促学习者参与到课程中来，就十分有必要了。

(3) 回归分析表明，“课程测验得分”“讨论得分”和“访问数得分”对期末考试成绩有显著性影响。其中“课程测验得分”的影响尤为显著。这表明，平时成绩是对学习效果的一个重要写照。加强平时学习，加大对平时测验的考核力度，对提高学习质量也有较显著的影响。

6.2 相关建议

随着MOOC平台的进一步建设，有效提

高MOOC学习平台的质量，也就成了打造精品MOOC的一个重要方面。基于前文的内容和结论，本文从合理设置课程和提高学习者学习效果这两个角度，为MOOC建设团队提供如下建议。

(1) 从课程设置的角度。应加大对“只是注册者”的监督管理，制定相应的奖惩机制，对于“选而不学”的学习者，进行提示提醒；

优化课程综合成绩评价标准的设计，加大平时成绩的比重，使之更加合理地体现学习者的学习情况；加强对互动讨论环节的设置，使之对学习者的讨论能够及时做出回应，调动学员的积极性；在视频播放过程中，适当加入问题，提高学习者观看视频的注意力，避免空有学习时间数据，而无实际学习行为现象的发生。

(2) 从学习者行为角度。应加强“平时学习”观念的培养，使学习者养成注重平时学习的习惯；加大对讨论区发言重视程度的研究，使学习者产生归属感、成就感，从而更加乐于学习和分享；采取相应措施（惩罚或者激励政策），让学习者尽快参与到课程的学习中来，可减少课程的荒废率。

7 展望

本文通过对MOOC学习者行为的研究,进一步探讨了影响MOOC学习质量的关键因素。但是里面所涉及的数据还不够广、不够全,仅选取了一门课程,共588位学习者学习数据为例,说服力不够。同时,对学习行为分析中,主要选取的是参与课程过程中的行为,对于选课行为、退课行为还缺少分析。

随着MOOC的快速发展,未来对MOOC的研究工作也将提出更高的要求。从学习者行为这一研究领域来看,未来的工作还可以集中在跨文化背景下不同学习者学习行为研究、基于学习者行为的学习效果建模预测研究等。

【参考文献】

- [1] Stein L A . Casting a wider net [J] . Science , 2012 , 338 (6113) : 1422-1423 .
- [2] Waldrop M . Online learning : Campus 2.0[J].Nature , 2013 , 495 (7440) : 160-163 .
- [3] 蒋卓轩,张岩,李晓明.基于MOOC数据的学习行为分析与预测[J].计算机研究与发展,2015,52(3):1-14.
- [4] Kizilcec R F , Piech C , Schneider E . Deconstructing disengagement : Analyzing learner subpopulations in massive open online courses[C] . Proceedings of the Third International Conference on Learning Analytics and Knowledge. New York : ACM , 2013 : 170-179 .
- [5] Hill P . Emerging Student Patterns in MOOCs : A (Revised) Graphical View [EB/OL] . <http://mfeldstein.com/emerging-student-patterns-in-moocs-a-revised-graphical-view/> .
- [6] 金娣. MOOC学习者流失成因的分析与对策——以Coursera平台《怪诞行为学》课程为例. 吉林教育 : 综合 , 2015 (17) : 4-5 .
- [7] 吴江,马磐昊. MOOC学习者行为分析研究——以爱课程平台的一门慕课为例[J]. 知识管理论坛 , 2015 (3) : 52-61 .
- [8] Yang D , Sinha T , Adamson D , et al. Turn on , tune in , drop out : Anticipating student dropouts in massive open online courses[C/OL] . Neural Information Processing System Workshop on Data Driven Education. <http://lytics.stanford.edu/data-driven-education/papers/yangetal.pdf> .
- [9] 贾积有,缪静敏,汪琼. MOOC学习行为及效果的大数据分析——以北大6门MOOC为例. 工业和信息化教育 , 2014 (9) : 23-29 .
- [10] Reeve , J.A self-determination theory perspective on student engagement. In Handbook of research on student engagement .2012 : 149-172 .
- [11] 潘丽佳. MOOC设计、学习者参与度和学习绩效的关系研究[D]. 杭州:浙江大学,2015 .
- [12] 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 (2) : 63-69 .
- [13] Taylor C , Veeramachaneni K . O ' Reilly U M. Likely to stop? Predicting Stopout in Massive Open Online Courses [J/OL] . [2015-10-20] . <http://arxiv.org/pdf/1408.3382v1.pdf> .
- [14] Girish B . Predicting student retention in massive open online courses using hidden Markov models [R/OL] . Berkeley : University of California , Berkeley , 2013 .
- [15] Nesterko S , Dotsenko S , Hu Q , et al . Evaluating the geographic data in MOOCs[C/OL] . Neural Information Processing System Workshop on Data Driven Education . [2014-06-01] . <http://nesterko.com/files/papers/nips2013-nestero.pdf> .
- [16] 薛宇飞,黄振中,石菲. MOOC学习行为的国际比较研究——以“财务分析与决策”课程为例[J]. 开放教育研究 , 2015 (6) : 80-85