

# 挖掘有意义学习行为特征:学习结果预测框架

武法提 田 浩

(北京师范大学 教育技术学院,北京 100875)

**[摘要]** 及时有效地对学习结果进行预测,是学习分析的核心议题,也是为学习者提供个性化学习支持服务、保证学习者学习成功的关键。其中,如何寻找兼具预测效力与教学意义的学习行为特征是值得关注的问题。本研究以 Cloudbag 教育云平台中的 108 位学生为研究对象,基于特征工程的方法构建学习结果预测框架。基于文献调研和对教师的访谈,本研究将学习结果预测框架分为学生与学生交互、学生与教师交互、学生与内容交互、学生与系统交互四个维度,共包含 10 个特征变量;结合学生在平台中的学习行为数据,对特征变量进行量化和筛选,通过相关分析、信息增益(率)分析筛选出八个有效的特征变量,构成最终的特征集合;使用八种机器学习算法对学习结果进行预测,结果表明:随机森林算法对学习结果的预测效果优于另外七种算法,其特征集合对学习结果的预测准确率可以达到 73.15%。本研究最后从有效学习行为指标和有效学习行为特征等方面对研究结果进行总结和反思,期望能够为混合式学习环境下学习分析和评估提供研究支持。

**[关键词]** 学习分析;学习行为特征;学习结果预测;学习结果

**[中图分类号]** G442

**[文献标识码]** A

**[文章编号]** 1007-2179(2019)06-0075-08

## 一、引言

智能时代的教育面临变革的需求,大力发展智能教育在《新一代人工智能发展规划》(国务院,2017)和《教育信息化 2.0 行动计划》(教育部,2018)等政策文件中均有提及。不难设想,以教育云平台为代表的智能学习环境将成为支撑智能教育发展的重要实践。然而,使用教育云平台进行混合学习和在线学习时,学生自主性较差、教师难以实时掌握学生学情等问题涌现。但相比传统学习环境,教育云平台可以记录更丰富、细粒度的学习行为数据,数据驱动的学习分析技术成为破解上述难题的关键手段:基于学习者行为数据可以预测学习结果,

并根据预测结果对学习过程进行针对性干预,改善学习者学习体验;也可以帮助教师综合了解学生学情,实时调整教学策略和教学内容。可见,预测学习者学习结果已成为学习分析领域的关键议题。为了保证预测效果可以切实改善教育教学,预测分析应从单纯关注教育大数据转为探索有意义的教学行为指标(Merceron et al., 2015)。因此,本研究旨在构建基于学习行为的学习结果预测框架,提取有意义的学习行为特征,并使用教育云平台的学习行为数据对框架进行验证。

目前学习结果预测的相关研究中,针对预测指标和预测效果的成果较多。比如,在预测指标方面,沙希里等(Shahiri & Husain, 2015)发现已有的学习

[收稿日期] 2019-08-26

[修回日期] 2019-10-21

[DOI 编码] 10.13966/j.cnki.kfjyyj.2019.06.008

[基金项目] 北京师范大学教育学部学科建设综合专项科研培育项目“场景驱动的个性化学习服务模型及其应用研究”(2019KYPY005)。

[作者简介] 武法提,博士,教授,博士生导师,北京师范大学教育学部,研究方向:智能学习系统与学习分析(wft@bnu.edu.cn);田浩,博士研究生,北京师范大学教育学部,研究方向:学习分析(tianhao@mail.bnu.edu.cn)。

预测研究使用最多的指标是学生平均绩点、每次任务的得分和课堂测验得分;其次是学生人口学变量和期末考试成绩;最后是学生参与课外活动和社会网络的表现。加德纳等(Gardner & Brooks, 2018)指出 MOOC 学习预测指标可以分为五类,分别是点击数据、论坛发帖内容、作业完成情况、课程元数据以及学习者的人口学信息。范逸洲等(2018)将预测指标分成倾向性指标、人机交互指标和人际交互指标三类。在预测效果方面,研究者对学习成绩这一连续变量的预测一般使用回归算法(Strang, 2017; 牟智佳等, 2017; Soffer & Cohen, 2019),而对成绩等级、及格与否等离散变量的预测多使用分类算法(舒莹等, 2019; Xing & Du, 2019)。更多的研究者使用多种算法进行预测,以规避算法对数据质量的敏感性。沃拉等(Vora & Rajamani, 2019)针对学生的社会人口学变量预测其学业表现,将深度信念网络和支持向量机算法结合,最终取得 71.5% 的预测准确率,并且这一混合算法要优于任意单一算法。勒梅等(Lemay & Doleck, 2019)使用逻辑回归、朴素贝叶斯、关联规则算法等八种机器学习算法,依据学生观看 MOOC 视频的数据,预测其每周任务得分,结果表明关联规则算法预测效果最佳,可以取得 70.2% 的预测准确率。祖海尔(Zohair, 2019)使用小规模数据预测学生学习成绩,发现当预测指标为连续变量时,线性判别分析的预测准确率最高;当预测指标编码为离散变量时,支持向量机算法的预测准确率最高。

通过上述分析可知,目前学习结果预测研究已呈现研究方法多样性、样本数据多元性等特点,但是对如何构建有意义学习行为特征的关注还不多,不足如下:1)在数据源方面,大多局限于学习平台的直观行为数据,对学习者的其他类型(如认知、情绪等)的数据关注不够;2)在数据指标方面,大多重视交互数量,忽视了交互质量;3)在研究目的方面,存在过于追求预测高精度的误区,忽略了对利益主体(学生、教师、管理者)的解释力;4)在研究范式方面,过分重视数据的作用,直接从数据中归纳学习规律,忽视了研究者和教师的能动性。针对上述问题,本研究通过文献调研确定了基于学习行为的学习结果预测框架维度;通过对教师的访谈明晰各维度的学习行为特征及其教学涵义;并通过多种预测算法,

检验该框架对学习结果的预测效果,为基于教育云平台的学习行为分析和后续干预提供借鉴。

## 二、基于学习行为的学习结果预测框架

已有研究将学习结果预测指标分为两类:静态指标和动态指标(见图 1)。其中,静态指标指客观存在的、不随学习过程变化的数据,比如学习者人口学信息和课程元数据。动态指标根据指标变化周期的长短,可分为倾向性指标和学习行为指标:倾向性指标表示学习者意志品质,其变化周期较长,包括学习者的学习偏好、学习习惯、认知风格等信息,这些指标与学习过程关系密切、相对稳定,但有可能随着长期学习而发生改变;学习行为指标指在线学习过程实时产生的行为数据,其变化周期较短,行为数据由交互过程产生(曹良亮等, 2012)。因此,按照交互主体,本研究将学习行为指标分为四类,分别是学生与学生的交互行为、学生与教师的交互行为、学生与内容的交互行为以及学生与系统的交互行为。

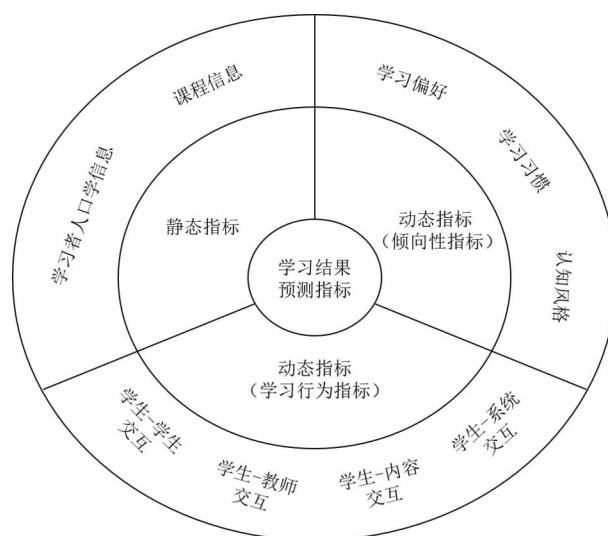


图 1 学习结果预测指标分类

在学习初期,静态指标往往具备较好的预测效果,但随着学习过程的推进,动态指标将逐渐成为预测的核心依据(Park & Tennyson, 1986)。然而,动态指标中的倾向性指标表征的是学习者的内隐属性,需要通过学习行为指标间接体现。基于此,本研究以学习行为指标下的四类交互行为作为分类依据,通过文献调研和对教师的访谈,确定各类交互中典型的学习行为特征,由此构建基于学习行为的学习

结果预测框架(以下简称预测框架,见表一)。

学习发生在社会情境中,因此,学生与学生的交互可以将学习变成合作的、社会建构的活动,而非单独的、孤立的任务(Mcneil et al., 2000)。该类交互主要涉及学生情绪、社交网络规模、社交质量、发帖数量。学生情绪关注学生在发帖文本中表现出的情绪倾向,社交网络规模关注平台论坛中与学生交互的人数,社交质量关注学生同伴的情绪状况,发帖数量关注学生在论坛发言的总次数。

教师在学习过程中起到的作用体现在三个方面,分别是完成课程设计并进行知识传授、对学习者进行评价、为学生提供帮助和鼓励(Moore, 1989)。因此,学生与教师的交互主要涉及学生是否积极参与各类任务、学生的拖延程度以及熟练掌握程度。积极参与主要关注学生回复老师提问贴的速度,拖延程度关注学生任务完成过程中的拖沓情况,熟练掌握反映学生完成作业任务的速度。

学生与内容的交互可以看作学生与教师的间接交互,因为学习内容是由教师精心制作编排的(陈丽等,2016)。学生与内容的交互包括选择和控制内容两方面。选择内容指学习者能够选择自己所需的学习资源;控制内容指自主控制资源的阅读、播放、呈现形式等(王志军等,2016)。本研究使用的教育云平台目前尚无法记录课件的播放、暂停等细化操作,因此该类交互目前主要关注资源下载和课件浏览的次数。

虽然以上三种交互都是在系统的支持下完成的,但“学生与系统的交互”仍有必要作为单独的交

互类型。学习者必须首先学会如何与系统交互,进而才能借助工具开展其他类型的交互(Hillman et al., 1994)。学生与系统的交互体现了系统设计者和开发者的思想,即界面是否符合学习者的使用习惯、运用的技术是否容易被学习者接受,以及通过该种方式展示的学习内容能否顺利传达给学习者。考虑到平台无法实现对系统访问路径的记录,因此本研究重点关注系统登录次数。

### 三、学习结果预测框架的实证探索

#### (一)平台与数据概况

本研究使用“Cloudbag 教育云平台”对预测框架进行实证探索。该平台包括“教学中心”“互动答疑”“电子教材”“课件点播”“资源下载”和“作业考试”六大功能模块,可以支持教师备课、上传资源、发布作业考试、发布公告,并支持学生观看课件、下载资源、进行考试、参与互动答疑等。本研究选取贵阳市某中学 108 位学生 2016 年 3 月至 2017 年 11 月产生的近 9000 条平台数据记录并进行分析(见表二)。数据分两部分:一是学生在上述时间段八次闭卷考试的卷面成绩,所有学生使用同一试卷同时进行线下考试;二是学生使用平台时产生的学习行为数据。

#### (二)学习结果测量

本研究将学习结果定义为学生考试成绩的等级,成绩等级的划分依据为:求取每位学生八次线下考试成绩的平均分,将分数距离平均分为五等份,由高到低命名为 A 到 E,根据正态分布规律可知 A 到 E

表一 基于学习行为的学习结果预测框架

交互类型	特征	编码	特征含义	相关指标
学生与学生交互	学生情绪	L_EM	学生在帖子文本中表现的情绪状况	帖子字数、帖子内容
	网络规模	L_NS	与学生在论坛中有互动记录的人数	帖子创建者、帖子回复者
	社交质量	L_SQ	学生同伴的平均情绪状况	帖子字数、帖子内容
	发帖数量	L_PN	学生论坛发言的总数	回答贴数量、回复贴数量
学生与教师交互	积极参与	T_EG	学生回复老师提问贴的速度	发帖时间、回帖时间
	拖延程度	T_PR	学生提交作业的拖沓程度	作业发布时间、作业截止时间、作业提交时间
	熟练掌握	T_PF	学生完成作业的速度	作业开始作答时间、作业提交时间
学生与内容交互	资源下载	R_DL	学生下载资源的总次数	各资源下载次数
	课件浏览	R_CV	学生浏览课件的总次数	各课件浏览次数
学生与系统交互	系统登录	S_LG	学生登录平台及访问各功能模块的总次数	平台登录次数、各功能模块登录次数

表二 教育云平台中学生行为数据统计结果

学习行为指标	最小值	最大值	平均值	标准差	缺失比例(%)
登录平台总次数	5	721	96.11	96.08	0
登录“课外阅读”模块次数	0	1	0.01	0.10	99.1
登录“课件点播”模块次数	2	226	52.21	45.32	0
观看课件总次数	0	449	101.41	100.27	0.9
观看课件总时长	0	26945	5759.41	6670.85	1.9
登录“教学中心”模块次数	9	488	89.77	80.42	0
登录“作业考试”模块次数	1	813	58.82	89.12	0
登录“资源下载”模块次数	0	98	21.93	17.75	0.9
资源下载总数	0	147	13.35	22.18	7.4
登录“互动答疑”模块次数	1	375	47.55	47.99	0
提问贴总数	0	5	0.32	0.79	80.6
问题被点赞次数	0	0	0	0	100
回答贴总数	0	163	15.22	17.40	2.8
回答被点赞次数	0	0	0	0	100
回答被加精次数	0	3	0.30	0.66	82.7
回复贴总数	0	161	5.83	15.57	12.0
考试平均分	430.31	708.25	619.22	51.75	0

五等级的学生占比分别为 3.59%、23.84%、45.14%、23.84% 和 3.59%，对应人数为 4 人、26 人、48 人、26 人、4 人。

### (三) 数据处理与分析

本研究使用 SPSS 20.0 和 WEKA 3.8 工具处理与分析数据,具体流程为:数据清洗、特征量化与标准化、特征筛选、考察特征对学习结果的预测效果。

### (四) 数据分析结果

#### 1. 数据清洗

数据清洗分两步:第一步是个案清洗,由于本研究的研究对象数量有限,因此不做个案的筛选,仅把每个个案的异常行为指标值替换为缺失值;第二步是变量清洗,删除缺失值超过 80% 的指标,包括登录课外阅读模块次数、提问贴总数、问题被点赞次数、回答被点赞次数及回答被加精次数五项指标。

#### 2. 特征量化与标准化

上述预测框架确定了四组共 10 个特征变量。为了便于后续分析,本研究对 10 个特征变量进行量化处理。根据特征的含义,前两组特征应综合多个数据指标计算得出,后两组特征可以通过对数据库

原始数据进行简单统计得出(见表三),各自的缺失值按其他学生在对应特征上的平均值予以填充。

至此,十个特征变量全部提取完毕。由于不同特征的量纲不同,导致实际数值之间差异较大,不具备直接比较的意义。为了方便后续分析,本研究统一将各特征的取值转化为标准分数(Z-Score)。

### 3. 特征选择

#### 1) 相关分析

在完成特征的量化后,本研究对特征进行筛选,删除预测效果不佳的特征。皮尔森相关系数(Pearson Correlation)可以衡量变量之间的线性相关关系,本研究计算 10 个特征与成绩之间的相关系数,若特征与成绩之间的相关性低(相关系数绝对值小于 0.1)或与其他特征之间存在高度共线性(方差膨胀因子大于 10),表明该特征质量不佳,应考虑将其删除。基于此,本研究使用 SPSS 20.0 计算特征变量与成绩之间的相关性及相关性之间的共线性(见表四)。

10 个特征变量中,只有“学生情绪(L\_EM)”与成绩的相关系数绝对值小于 0.1( $0.037 < 0.1$ ),且

表三 学习行为特征变量计算方法

维度	特征	计算方法
学生与学生交互	学生情绪	$i$ 为学生在“互动答疑”模块的发帖总数; $word_n$ 表示第 $n$ 条贴子的字数; $affection_n$ 为第 $n$ 条贴子文本的情感极性(通过百度情感倾向分析 API 计算得出)。计算结果介于 0-1 之间,数值越接近 1,表明学生情绪越积极,越接近 0,表明情绪越消极。
	网络规模	$L_{NS} = n_s + n_i$ $n_s$ 为学生在“互动答疑”模块回复的人数; $n_i$ 为在“互动答疑”模块回复该学生的人数。
	社交质量	$i$ 为与该学生在“互动答疑”模块有过互动记录的人数; $frequency_n$ 表示该学生与第 $n$ 个人的互动频次; $affection_n$ 为第 $n$ 个人的情绪值(可由“学生情绪”特征的计算方法求得)。计算结果介于 0-1 之间,数值越接近 1 表明同伴情绪越积极;越接近 0,表明同伴情绪越消极。
	发帖数量	$L_{PN} = n_a + n_r$ $n_a$ 为学生在“互动答疑”模块发表回答的总数; $n_r$ 为学生在“互动答疑”模块发表回复的总数。
学生与教师交互	积极参与	$i$ 为学生在“互动答疑”模块对教师提问的回答总数; $ta_n$ 为第 $n$ 条回答的发布时间; $tq_n$ 表示第 $n$ 条回答对应的提问发布时间。
	拖延程度	$i$ 为学生在“作业考试”模块提交的作业总数; $tb_n$ 表示第 $n$ 次作业的发布时间; $ts_n$ 表示第 $n$ 次作业的提交时间; $te_n$ 表示第 $n$ 次作业的截止时间。计算结果介于 0-1 之间,数值越接近 1,表明拖延状况越严重,越接近 0 表明拖延状况较轻。
	熟练掌握	$i$ 为学生在“作业考试”模块提交作业总数; $ts_n$ 表示第 $n$ 次作业的开始作答时间; $tm_n$ 表示第 $n$ 次作业的提交时间。
学生与内容交互	资源下载	资源下载总数,数值可直接从平台数据库获得。
	课件浏览	课件浏览总数,数值可直接从平台数据库获得。
学生与系统交互	系统登录	系统登录总数,数值可直接从平台数据库获得。

表四 相关分析与共线性分析

特征	L_EM	L_NS	L_SQ	L_PN	T_EG
相关系数	-0.037	0.232 *	0.308 **	0.285 **	-0.201 *
方差膨胀因子	3.134	1.208	1.06	1.331	1.312
特征	T_PR	T_PF	R_DL	R_CV	S_LG
相关系数	-0.292 **	0.367 **	0.446 **	0.692 **	0.749 **
方差膨胀因子	1.132	1.268	1.504	1.968	2.351

注: \* 表示显著性水平小于 0.05; \*\* 表示显著性水平小于 0.01; \*\*\* 表示显著性水平小于 0.001。

结果不显著。同时,共线性诊断结果表明,各特征之间不存在高度共线性。由此,相关分析表明“学生情绪”这一特征对成绩的解释力不佳。

2) 信息增益(率)分析

信息增益(Information Gain)能够反映特征携带的信息量,体现对整个系统的重要程度,因此信息增益越大,表明特征越重要。本研究使用 WEKA 3.8 工具,选择信息增益评估器计算每个特征的信息增益,并使用排序算法(Ranker)对各特征的信息增益排序。不过有研究者认为,信息增益偏向于优先选择取值较多的特征,容易出现过拟合现象,于是将信息增益率(Information Gain Ratio,即信息增益与特征的熵值之比)作为补偿手段,弥补信息增益存在的问题。基于此,本研究也选择信息增益率评估器计算每个特征的信息增益率,使用排序算法对每个特征变量的信息增益率排序(见表五)。结果表明,

“学生情绪(L\_EM)”和“积极参与(T\_EG)”表现差。

表五 信息增益(率)排序

信息增益		信息增益率	
排序	特征	排序	特征
1	S_LG	1	S_LG
2	R_CV	2	R_DL
3	L_PN	3	L_NS
4	T_PF	4	R_CV
5	L_SQ	5	T_PR
6	R_DL	6	L_SQ
7	T_PR	7	T_PF
8	L_NS	8	L_PN
9	T_EG	9	T_EG
10	L_EM	10	L_EM

结合相关分析和信息增益(率)分析可以看出,“学生情绪”和“积极参与”两个特征预测任务表现不佳,其它八个特征变量构成本研究最终的特征集合。

4. 特征变量对学习结果的预测效果分析

本研究同时使用贝叶斯网络、朴素贝叶斯、支持向量机、序列最小优化、逻辑回归、关联规则、决策树和随机森林八种分类算法,分别使用特征选择前和特征选择后的特征集合预测学习结果(见表六),预

测算法的测试采用五折交叉检验方法。

预测结果显示,删除两项特征后,除贝叶斯网络和支持向量机算法的预测准确率保持不变,另外六种算法的准确率均有不同程度的提升,即用更少的有效特征取得了更好的预测效果,表明特征选择的有效性。其中,随机森林算法取得了最好的预测效果,准确率达 73.15%。本研究的最终目的并非追求预测的高准确率,而是旨在构建有意义的预测框架,进而刻画可被教师理解的学习过程;同时,本研究面向混合学习环境,尚无法有效监测学生线下学习过程。基于以上两点考虑,70%左右的准确率可以认为是比较理想的预测结果,反映了预测框架可以有效地预测学习结果。

#### 四、讨论与总结

##### (一)对学习行为指标的反思

教育云平台中,学生主动与老师、同伴互动交流的意愿较低。本研究获取了教育云平台记录的 17 项学习行为指标,缺失率较高的五项指标分别为登录课外阅读模块次数、提问贴总数、问题被点赞次数、回答被点赞次数和回答被加精次数。除登陆课外阅读模块次数外,其余四项指标均属于“互动答疑”功能模块,表明学生对平台提供的协作、互动功能的使用尚不够充分。另外,相比回答和回复操作,学生主动发帖提问的次数较少。教师访谈结果表明,“在线学习中,学生较少自发进行线上有意义讨论,更多的是老师发布提问,学生回复并发表自己的观点”。可见,学生使用平台的主动性仍有待加强(李爽等,2016)。

##### (二)对学习结果预测特征变量的反思

本研究中,学生情绪未对学习成绩产生显著影响。虽然以往有大量研究证明,积极情绪会对学习成绩带来正向影响(Pekrun et al., 2011; Mega et al., 2014; Jarrell et al., 2016)。也有研究发现,教师期望、同伴竞争等因素会给学习者带来紧张、焦虑等负面情绪(夏洋等,2018),但适当的负面情绪反而能帮助学生更专注于学习过程,投入更多努力以获得较好的学习结果(Pekrun & Linnenbrink - Garcia, 2012)。由此可见,学业情绪对学习成绩的影响是复杂的过程。

另外,研究结果表明,学生在论坛中回答问题的积极性与学习成绩无明显关联。通过访谈教师,本研究发现论坛中的大部分问题并非要求即时作答,学生倾向于选择课后作答,因此问题回答的时效性不强;另外,部分问题属调查类题目(诸如“航天器的失重现象还有什么问题,请提出来!”“第一轮投票:你认为哪位同学的展示最有历史韵味?”),这些问题与学生对知识的掌握情况关联不大,因此学生在论坛中回答问题的积极性与学习成绩之间不存在显著相关。

最后,特征选择的结果显示,相比人际交互,人机交互对学习成绩的影响更大。本研究中,人机交互包括学生与内容交互、学生与系统交互两类,在学习行为指标中体现为各类频次和时长,较高的交互频次(如下载频次、观看频次、登录频次等)和交互时长(如观看时长、登录时长等)反映了学习者花费更多时间和精力用于内容的学习(范逸洲等,2018)。人际交互与成绩的低相关性反映了基础教育对学生协作

表六 学习结果预测效果

算法	特征选择前	特征选择后					
	准确率	准确率	精准率	召回率	F1 值	平均绝对误差	均方根误差
贝叶斯网络	71.30%	71.30%	0.697	0.713	0.704	0.1581	0.3071
朴素贝叶斯	64.81%	65.74%	0.638	0.657	0.645	0.1579	0.3416
支持向量机	62.04%	62.04%	0.601	0.620	0.610	0.1519	0.3897
序列最小优化	59.26%	60.19%	0.584	0.602	0.592	0.2593	0.3452
逻辑回归	62.96%	64.81%	0.666	0.648	0.653	0.1644	0.3494
关联规则	62.04%	65.74%	0.643	0.657	0.644	0.1709	0.3378
决策树	62.04%	62.96%	0.623	0.630	0.625	0.1591	0.3628
随机森林	72.22%	73.15%	0.730	0.731	0.729	0.156	0.2744

学习的要求不高,因此,学生与老师、同伴的互动交流对学习成绩的影响不大。

### (三) 研究创新点

学习分析的关键是探索有意义的教育数据。本研究在获取平台指标数据的同时,采集了学生论坛发帖的文本数据,从文本内容中挖掘学生的学业情绪,这是扩展数据源的一次尝试;其次,本研究通过特征工程的方法,从平台学习行为指标的基础上提取学习行为特征,特征提取结合了文献调研和访谈,保证研究过程和结果符合教育实践的真实需求。

### (四) 研究局限与展望

受限于教育云平台的功能设计、数据埋点设计,以及技术开发等因素,平台目前无法记录更微观的操作行为,比如只能记录视频观看的总次数和总时长,无法获取每个视频的观看时间,以及学生观看视频时的暂停、回看等操作。后续研究可以完善和规范平台数据格式和数据埋点规则。

另外,受教师教学设计的限制,平台的使用仍集中于资源的下载观看、在线作业与考试等功能,教师设计的生生、师生之间的讨论与协作活动较少,导致部分指标缺乏数据而被剔除。未来研究者应与一线教师深入合作,探究将教育云平台常态化融入课堂教学的模式。

最后,本研究构建的学习结果预测框架在现有研究情境下取得了较理想的预测效果,但仍有提升空间。未来研究可以考虑采集更多维度的学生数据(如多模态数据),构建更全面的特征,比如邀请专家对特征量化方法进行迭代修改,以保证特征可以更好地反映学生学习行为特点;使用深度学习算法代替一般的机器学习算法,以取得更优的预测准确率。

### [参考文献]

- [1] 曹良亮, 袁克定(2012). 在线学习者学习行为特点的初步探讨[J]. 中国远程教育, (5): 56-61.
- [2] 陈丽, 王志军(2016). 三代远程学习中的教学交互原理[J]. 中国远程教育, (10): 30-37 + 79-80.
- [3] 范逸洲, 汪琮(2018). 学业成就与学业风险的预测:基于学习分析领域中预测指标的文献综述[J]. 中国远程教育, (1): 5-15 + 44 + 79.
- [4] Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs[J]. *User Modeling and User-Adapted Interaction*, 28

(2): 127-203.

- [5] 国务院(2017). 国务院关于印发新一代人工智能发展规划的通知[A/OL]. [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm).
- [6] Hillman, D. C., Willis, D. J., & Gunawardena, C. N. (1994). Learner-interface interaction in distance education: An extension of contemporary models and strategies for practitioners[J]. *American Journal of Distance Education*, 8(2): 30-42.
- [7] Jarrell, A., Harley, J. M., & Lajoie, S. P. (2016). The link between achievement emotions, appraisals, and task performance: pedagogical considerations for emotions in CBLEs[J]. *Journal of Computers in Education*, 3(3): 289-307.
- [8] 教育部(2018). 教育部关于印发《教育信息化2.0行动计划》的通知[A/OL]. [http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425\\_334188.html](http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425_334188.html).
- [9] Lemay, D. J., & Doleck, T. (2019). Grade prediction of weekly assignments in MOOCs: mining video-viewing behavior[J]. *Education and Information Technologies*: 1-10.
- [10] 李爽, 王增贤, 喻忱, 宗阳(2016). 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*, 22(2): 77-88.
- [11] Mcneil, S. G., Robin, B. R., & Miller, R. M. (2000). Facilitating interaction, communication and collaboration in online courses[J]. *Computers & Geosciences*, 26(6): 699-708.
- [12] Mega, C., Ronconi, L., & De Beni R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement[J]. *Journal of educational psychology*, 106(1): 121-131.
- [13] Merceron, A., Blikstein, P., & Siemens, G. (2015). Learning analytics: From big data to meaningful data[J]. *Journal of Learning Analytics*, 2(3): 4-8.
- [14] Moore, M. G. (1989). Three types of interaction[J]. *American Journal of Distance Education*, 3(2): 1-7.
- [15] 牟智佳, 武法提(2017). MOOC学习结果预测指标探索与学习群体特征分析[J]. *现代远程教育研究*, (3): 58-66 + 93.
- [16] Park, O. C., & Tennyson, R. D. (1986). Computer-based response-sensitive design strategies for selecting presentation form and sequence of examples in learning of coordinate concepts[J]. *Journal of Educational Psychology*, 78(2): 153-158.
- [17] Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ)[J]. *Contemporary educational psychology*, 36(1): 36-48.
- [18] Pekrun, R., & Linnenbrink-Garcia, L. (2012). *Academic Emotions and Student Engagement*[M]. Springer US.
- [19] Shahiri, A. M., & Husain, W. A. (2015). Review on predicting student's performance using data mining techniques[J]. *Procedia Computer Science*, 72: 414-422.
- [20] 舒莹, 姜强, 赵蔚(2019). 在线学习危机精准预警及干

预:模型与实证研究[J]. 中国远程教育, (8): 27-34 + 58 + 93.

[21] Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses[J]. Journal of Computer Assisted Learning, 35(3): 378-389.

[22] Strang, K. D. (2017). Beyond engagement analytics; Which online mixed-data factors predict student learning outcomes? [J]. Education and information technologies, 22(3): 917-937.

[23] Vora, D. R., & Rajamani, K. A. (2019). A hybrid classification model for prediction of academic performance of students: A big data application[J]. Evolutionary Intelligence, 2019: 1-14.

[24] 王志军, 陈丽, 韩世梅(2016). 远程学习中学习环境的交互

互性分析框架研究[J]. 中国远程教育, (12): 37-42 + 79-80.

[25] 夏洋, 徐忆(2018). 英语专业课堂环境因素对学生消极学业情绪的影响研究[J]. 外语与外语教学, (3): 65-76 + 144-145.

[26] Xing, W., & Du, D. (2019). Dropout prediction in MOOCs: Using deep learning for personalized intervention[J]. Journal of Educational Computing Research, 57(3): 547-570.

[27] Zohair, L. M. A. (2019). Prediction of Student's performance by modelling small dataset size[J]. International Journal of Educational Technology in Higher Education, 16(1): 1-18.

(编辑:魏志慧)

## Mining Meaningful Features of Learning Behavior: Research on Prediction Framework of Learning Outcomes

WU Fati & TIAN Hao

(School of Educational Technology, Beijing Normal University, Beijing 100875, China)

**Abstract:** *As the core issue of learning analytics, timely and effective prediction of learning outcomes is the key to ensuring learners' success and an essential means to provide learners with personalized learning support services. How to find learning features that have both predictive effectiveness and teaching significance has become a problem worthy of attention. This study built a learning outcome prediction framework by feature engineering method, based on 108 students in the CLOUDBAG Platform. Firstly, through literature review and teacher interviews, the framework was divided into four dimensions: student-student interaction, student-teacher interaction, student-content interaction, and student-interface interaction, which contained a total of ten learning features. Secondly, using students' learning behavior data in the platform, ten features were quantified and filtered, and eight effective features were selected through correlation analysis and information gain (rate) analysis to form the final feature set. Finally, learning outcomes were predicted by eight machine learning algorithms. The results showed that the random forest algorithm was better than the other seven algorithms, and the prediction accuracy can reach 73.15%. At the end of the paper, research results were summarized and reflected from the aspects of effective learning behavior indicators and effective learning features. It was expected to provide research support for learning analytics and evaluation in a blended learning environment.*

**Key words:** *feature engineering; learning analytics; learning features; learning outcome prediction; random forest*