

Online Learning Behavior and Learning Outcomes

—An Empirical Study Based on Learning Analysis

Jia Li, Rong Guan

School of Statistics and Mathematics, Central University of Finance and Economics, Beijing
Email: lijsky0814@163.com, guanrong721@qq.com

Received: Oct. 9th, 2019; accepted: Oct. 17th, 2019; published: Oct. 24th, 2019

Abstract

Network teaching platform has attracted more and more attention due to its unique spatial and temporal convenience. By mining massive behavioral data left by learners on the technology platform, learners' learning status can be mastered, thus playing an intervention and guidance role in the learning process. In view of this, this paper adopted the behavior data set of the freshmen majoring in computer engineering from the University of Genoa, Italy, who took part in the experiment course of "digital circuit", and explored the relationship between learning behavior and learning effect by analyzing the students' performance warning. The conclusion shows that in the early warning analysis of whether the students can pass the final test, the effect is better when the classification tree model is used to predict the grades in middle and late stage of the course learning. In the warning analysis of how many points can students get, the effect is better when the regression tree model is used to predict the grades in preintermediate stage of the course learning.

Keywords

Learning Analysis, Online Learning, Learning Effect, Achievement Early Warning Analysis, Intervention Warning

在线学习行为与学习效果

——基于学习分析的实证研究

李 佳, 关 蓉

中央财经大学统计与数学学院, 北京
Email: lijsky0814@163.com, guanrong721@qq.com

收稿日期: 2019年10月9日; 录用日期: 2019年10月17日; 发布日期: 2019年10月24日

摘要

网络教学平台凭借独特的时空便利性受到越来越多人的关注, 通过挖掘学习者在技术平台上留下的海量行为数据, 可以掌握学习者的学习状况, 从而对学习过程起到干预和指导作用。鉴于此, 文中采用了意大利热那亚大学计算机工程专业大一学生参加“数字电路”实验课程的行为数据集, 通过对学生进行成绩预警分析, 探索了学习行为与学习效果之间的关系。结论表明: 在对学生能否通过最终测验进行不及格预警分析时, 用分类树模型于课程学习的中后期阶段进行预测效果较好; 在对学生的最终测验成绩进行分数预警分析时, 用回归树模型于课程学习的前中期阶段进行预测效果较好。

关键词

学习分析, 在线学习, 学习效果, 成绩预警, 干预警示

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

传统的学习环境随着信息技术的广泛应用正在经历转型和重塑。学习者只要在当前的学习空间下选择适合的学习终端, 即可开启数字化学习。利用大数据和人工智能技术, 可以收集学习者留在技术平台上的学习行为数据, 进行学习行为分析与评价, 挖掘出对教学者和学习者有用的教学信息, 进而提高教育管理效率和学习效率。因此, 以学习分析为主题的研究趋势日益形成, 并成为当前教育应用研究的热点问题之一。

国内外对于在线学习行为分析这一主题的探究大体有三个角度, 分别是个人特征、平台架构、建模分析。在探讨个人特征对学习行为的影响时, DeBoer [1]研究了学习者自身的年龄、性别、地区等人口统计学特征对其课程通过率的影响。李阳[2]通过调查问卷, 对西北大学的学生进行调查研究, 探讨了不同群体特征的差异及与人格特征、学习效果之间的内在联系, 并根据讨论结果提出了针对性建议。宗阳[3]基于 Moodle 平台案例课程, 通过构建异步学习平台中远程学习者学业情绪分析模型, 探究了远程学习者在线学习行为与学业情绪之间的关系, 结论表明学习者的学业情绪与作业成绩呈显著正相关。陈晋音[4]认为, 学习者的性格特征与学习效率之间是存在联系的, 通过分析在线学习的行为特征, 可以实现个性化学习方法推荐。赵蔚[5]收集了 Moodle 平台中的学习过程数据, 运用 SSAS 顺序分析和聚类分析算法、SPSS 分层聚类分析, 挖掘不同学习风格、学习偏好的学习者的学习路径。

除了个人特征角度, 有的学者以平台架构为导向, 对学习行为进行了分析。魏顺平[6]以 Moodle 平台上开展的某在线培训课程记录的日志数据为样本, 思考了如何对在线教学绩效进行评估, 并验证了其提出的绩效评估模式的有效性。刘双[7]拟设计一个行为数据采集的工具, 用以实时收集学生的学习行为数据, 并利用 xAPI 技术进行数据标准化处理, 为后续的数据分析打好基础。洪丹丹[8]为了增强在线学习平台记录下的统计数据的可读性, 充分发挥用户行为数据的作用, 以统计理论为基础, 结合了先进的 ECharts 可视化技术与 JFinal 轻量级框架对某在线学习平台课程后的算法指标模型、数据库模型、前端数据可视化与后端数据应用等技术进行了研究。谭召[9]则是基于自身实际的 Java 教学经验, 专门设计了一种基于 SSH 框架的 Java 在线教学平台, 该平台包括用户管理模块、媒体学习环境模块、试题练习环境模

块、程序训练环境模块、学情反馈模块等等, 该在线学习平台虽然仅以 Java 语言作为教学内容, 但其平台设计思路及框架也适用于其他类型的学习平台。谢修娟[10]提出, 将决策树 C4.5 算法运用到网络学习平台的决策分析当中, 并据此设计了基于决策树理论的在线教学辅助系统, 以帮助学生实时的获取学习反馈, 进而提升学习效率, 端正学习态度。

此外, 还有部分学者致力于对学习行为数据进行建模分析, 以期对学生的行为起到规范、预警等作用。S. Jiang 等人[11]用 logistic 回归作为分类器, 根据学习者一周的学习行为记录, 对学生最终通过测验获取证书的概率做了预测。Kloft [12]采用机器学习中的决策支持向量机分析了学习者的课程点击序列, 也是对学生的中途辍学率进行判定。Halawa [13]通过学习者的学习特征判定其是否对学习失去兴趣, 对退出率给出了较准确的预测。李阳[14]从结构维度、功能维度等多个维度首先对在线学习行为进行了划分, 给出了在线学习行为分析模型的总体架构, 在此基础上, 采用大数据技术深入研究了其行为分析模型的适应性, 丰富了行为科学及人工智能理论。王慧芬[15]以华东师范大学校内的远程学习网上的学习行为日志数据为案例, 利用 Hadoop 数据分析工具, 从学习行为数据中挖掘学习者的信息进行行为特征分析与相关分析, 进而提供学习平台优化策略。吴林静[16]基于数据挖掘技术, 提出了一种在线学习行为分析模型, 主要是将网络平台中的在线学习行为分成了独立学习行为、资源交互行为、系统交互行为以及社会交互行为四类, 在各类别行为下分别运用相关分析、分类分析等统计方法展开研究。

在研究了以往文献后, 本文希望能在此基础上继续创新, 对学生的重点学习行为作更精细的分析, 充分利用经典统计学方法和机器学习算法基于学生的学习行为表现进行成绩预警分析, 并择优选用成绩预警模型, 深入把握学习行为规律, 以更好的促进学生的学习成果。本文基于 Vahdat 等[17]提供的 DEEDS 数据集, 对学生的成绩预警机制分别进行探索和设计, 旨在解决如下两个问题: ① 学习者的学习行为是否会对学习效果产生影响? 如何根据学习行为对学习者的学习过程进行干预, 起到预警作用? ② 何时对学生进行成绩预警分析效果最好? 本文的结构安排如下: 第 1 章为引言部分, 介绍研究的背景和现状, 第 2 章介绍研究采用的数据与方法, 第 3 章为实证分析部分, 第 4 章是本文所得出的研究结论。

2. 数据与方法

2.1. 数据和变量

2.1.1. 数据来源

本文使用的数据来源于一个名为 DEEDS (Digital Electronics Education and Design Suite, 数字电路教育和设计组件)的网络学习平台, 该平台是由意大利的两位学者联合开发出来的, 用以进行数字电路课程的教学。数据主要记录了意大利热那亚大学计算机工程专业的新生学习“数字电路”这门课程的学习过程, 是由 Vahdat 等学者在 2015 年收集得到的。Deeds 平台集成了“有限状态机模拟器(d-FsM)”、“数字电路模拟器(d-DcS)”、“微机接口与编程模拟器(d-McE)”三大模拟器, 分别承担不同的职能和作用, 以供学习者完成组装、设计常规实验电路, 设计、测试有限状态机以及使用汇编语言练习微机接口与编程, 实现特定的数字电路功能。学习者正是通过上述模拟器完成了“数字电路”课程的学习。

在课程的实验教学过程中, 设计者采用了让学生根据预设问题自主学习探索的教学方式, 即学生为了更好的完成每一课时下的练习, 会在提供的实验环境下学习相应课程并通过实际操作进行探索。每课时提供的学习时长为 3 小时, 学生要在规定的时间内, 以课程设定的学习目标为导向进行在线课程学习与实操锻炼, 课时学习完成的标志以实验报告的提交为准。研究者在每位学生的个人电脑上安装了一款软件收集学生的 ID 信息并记录其对应的学习活动, 以及该活动的起止时间、学生的鼠标按键点击次数、鼠标移动距离、键盘敲击次数等指标, 最终形成学生的学习行为数据。

2.1.2. 变量及其含义

通过对同类活动进行合并、考虑数据的缺失率, 本文提炼出 8 种不同的学习行为。学习行为名称及详细描述如表 1 所示。由于原始数据集中记录的是学生的每一项细微的动作, 因此部分变量例如鼠标点击次数、键盘敲击次数等存在大量 0 值, 不易直接对原始数据集进行分析操作。此外, 由于数据表中的学习行为变量种类繁多, 且反映了学生学习状况的直接进度, 具有较高的分析价值, 因此, 本文通过数据透视表对学习行为变量下的变量值进行了“频次”和“时长”两方面的汇总, 辅以学生参与课时数的日志记录, 便可计算出学生在该门课程中各学习行为的执行状况。

Table 1. Learning behavior names and detailed descriptions

表 1. 学习行为名称及详细描述

变量名称	详细说明
Aulaweb	在学习管理平台上上传与下载学习资料
Blank	所访问的窗口标题未被记录下来
Deeds	在数字电路模拟器(d-DcS)上进行一系列电路模拟操作
Diagram	绘制相应逻辑电路的时序电路图
Other	上述行为之外, 实验课中可能发生的相关行为, 如使用 Facebook 聊天
Properties	使用属性窗口设置任务所需参数
Study	学习或查看某一特定练习内容
TextEditor	与文本编辑相关的行为, 如在 Word 中填写实验报告

在该门课程的学习中, 每位同学有两次期末测试的机会, 两次测试内容相同但细节不同。每位同学可选择不参加测试、参加一次测试或两次测试均参加。本文将选取每位学生的最高测验成绩作为其最终成绩, 同时剔除学习行为记录或者期末成绩有缺失情况的学生样本, 最终得到 93 位研究对象。

通过上述数据预处理, 获得了整理后的数据集(前 3 条样本如表 2 所示)。其中, Student_id 表示学生的序号; point 表示最终学习成绩; log 表示学生实际参加的课时数; “N+学习行为”表示学习行为发生频次, “T+学习行为”表示学习行为持续时长。学习行为发生频次与学习时长均统计课时平均水平, 且单位为分钟。该门课程共 6 课时, 课时的学习与否由学生自行决定。

Table 2. Partial sample presentation of data set after preprocessing

表 2. 预处理后数据集部分样本展示

Student_id	point	log	NAulaweb	TAulaweb	...	TTextEditor
1	94.5	5	14.8	1	...	33.7
2	44	6	12.8	1	...	43.3
3	85	4	13.3	0.3	...	59.1
...

2.2. 研究方法

本文主要从以下两个方面展开研究: ① 预警分析。基于学生的学习行为, 首先通过 Logistic 回归、分类树模型对期末成绩能否合格进行预测, 并通过 AUC 值及混淆矩阵选出较优模型。其次再利用传统回归和回归树模型对期末成绩具体分数情况作出预测, 并通过 MAPE 值选出较优模型。根据模型结果设计

学习行为表现的预警机制, 以便在课程中途达到干预和警示的目的。② 选取最佳预警时间节点。本文希望辅以 5 折交叉验证的方法分别计算出 Session 3~Session 6 的学习行为对期末成绩的预警模型, 分别选出不同预警目的下的最优预警课时节点, 使得预警模型既能满足预测准确性, 又能满足预警时效性的需求。

3. 实证分析

3.1. 不及格预警模型

在构建不及格预警模型时, 本文分别利用逻辑回归和分类树模型对学生能否通过测验做出判断, 并通过 AUC 值及混淆矩阵选出较优模型。结果显示, Logistic 分类器的 AUC 值为 0.78, 当选定阈值为 0.489 时, 即模型返回的预测概率值大于 0.489 时预测成绩为合格, 真阴性率和真阳性率分别可达到 0.682 和 0.857。分类树模型计算出的 AUC 值高达 0.895, 说明该分类器比 Logistic 回归模型具有更准确的预测效果, 当设定阈值为 0.466 时, 模型得到的真阴性率和真阳性率可分别达到 0.864 和 0.816。虽然真阳性率与 Logistic 回归模型相比略有降低, 但 AUC 值以及真阴性率都有大幅提升, 对于实际问题来说有更加现实的指导意义。此外, 分类树模型的预测误判率为 0.161, 错分 15 个观测值, 相比 Logistic 回归模型有明显减少。因此, 在基于学生的学习行为探索分类预警模型时, 分类树相比于 Logistic 回归模型具有更高的精确度以及更低的错判率, 更适宜用来进行分类预警分析。分类树拟合结果如图 1 所示。

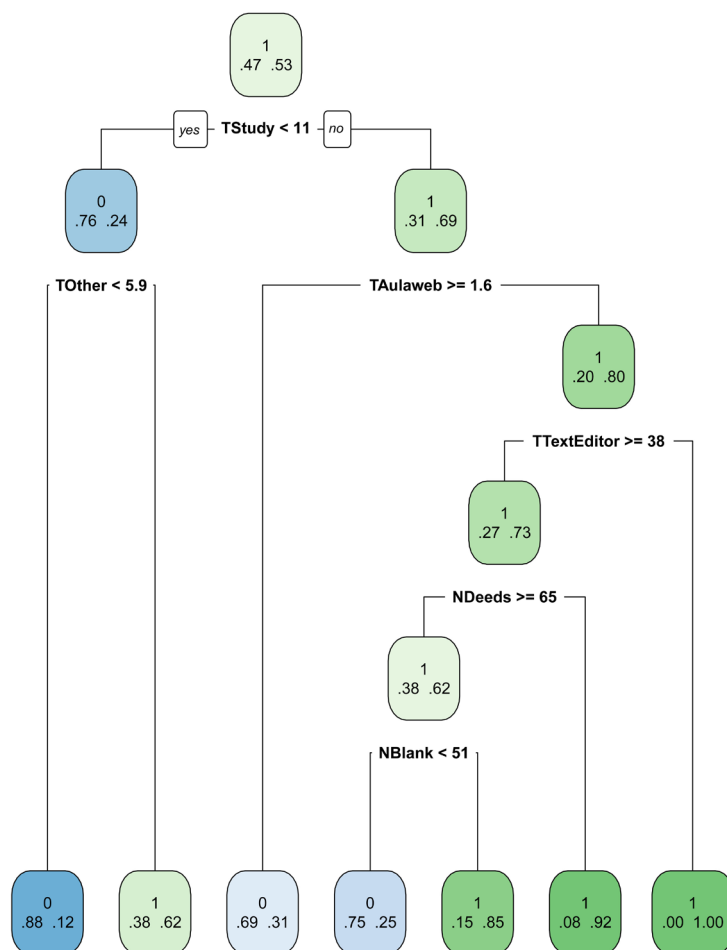


Figure 1. Classification tree model fitting results
图 1. 分类树模型拟合结果

这里, 分类树选择了 TStudy、TOther、TAulaweb、TTextEditor、NDeeds、NBlank 作为分裂的筛选变量。这说明, 学生期末测验的成绩合格率不仅与学生的线上学习与实操锻炼有关, 也与学生学习过程中的专心程度有密切的联系。

首先, 分类树选择了 TStudy 作为最先分裂的变量。这说明, 学生在网络学习平台上的学习时长是非常重要的指标, 线上学习时间越长, 表明学生的学习投入度越高, 成绩合格的概率越大。此外, 在考虑了线上学习因素之后, 其他学习行为, 例如在学习管理平台上上传与下载学习资料和撰写实验报告所花费的时间等, 也对学生的成绩合格率有较大的影响。例如, 当学生每课时的线上学习时长超过 11 分钟时, 花费不超过 1.6 分钟的时间下载学习资料并且在 38 分钟类即可撰写完成实验报告的学生具有非常高的概率会通过最终的期末测验。

由表 3 可知, 在错分的样本中, 将不及格预测为及格应被视为是较严重的错误, 因为这将导致学习者或者监管者放松对该学生的学习干预进度。在本次用分类树进行不及格预警分析时, 错分的样本有 19 号、25 号、49 号、60 号、89 号以及 102 号同学, 不同学生被错分的具体原因有所不同, 例如, 19 号同学被误判为能够及格是因为其进行课程学习的时间和次数过小, 导致分类树的预测出现了偏差。

Table 3. Classification results of learning behavior data

表 3. 分类树对学习行为数据的分类结果

		预测值	
		不合格	合格
真实值	不合格	38	6
	合格	9	40

3.2. 分数预警模型

在构建分数预警模型时, 本文分别利用传统回归模型和回归树模型对学生取得期末成绩的具体分数作出预测, 并通过 MAPE 值选出较优模型。通过相关分析, 本文采用传统线性回归模型对学习行为数据进行了拟合, 并通过了模型的显著性检验(F 检验的 P 值 < 0.001), 预测结果的 MAPE 值为 57.08%, 回归模型适应性的诊断结果较好。其次, 用回归树模型对原始数据集进行预测, 得到的平均绝对百分误差 MAPE 值为 40.65%。从预测具体结果以及 MAPE 值来看, 回归树相比于传统线性回归模型具有更高的拟合度, 预测误差更小, 更适宜用来进行回归预警分析。回归树拟合结果如图 2 所示。

由图 2 可以看出, 回归树筛选出了 TTStudy、TTAulaweb、TDeeds、NProperties、TDiagram、NDiagram 作为重要的影响变量, 从第一个变量(线上学习时长)出发, 根据不同变量的不同取值, 能够发现当 $\log(TTStudy)$ 不低于 2.5, $\log(TDeeds)$ 小于 3.4 且每课时下 NProperties 操作次数不少于 32 次时, 学生有很大概率能通过期末测验, 并根据每课时下 Diagram 操作频次和时长的不同表现而取得不同水平的成绩。

表 4 是回归树预测结果的部分示例。

Table 4. Examples of partial prediction results from regression tree model

表 4. 回归树模型部分预测结果示例

ID	真实成绩	预测成绩
1	94.5	87.1
2	44	33.2
3	85	87.1
4	30	33.2
5	38.5	33.2

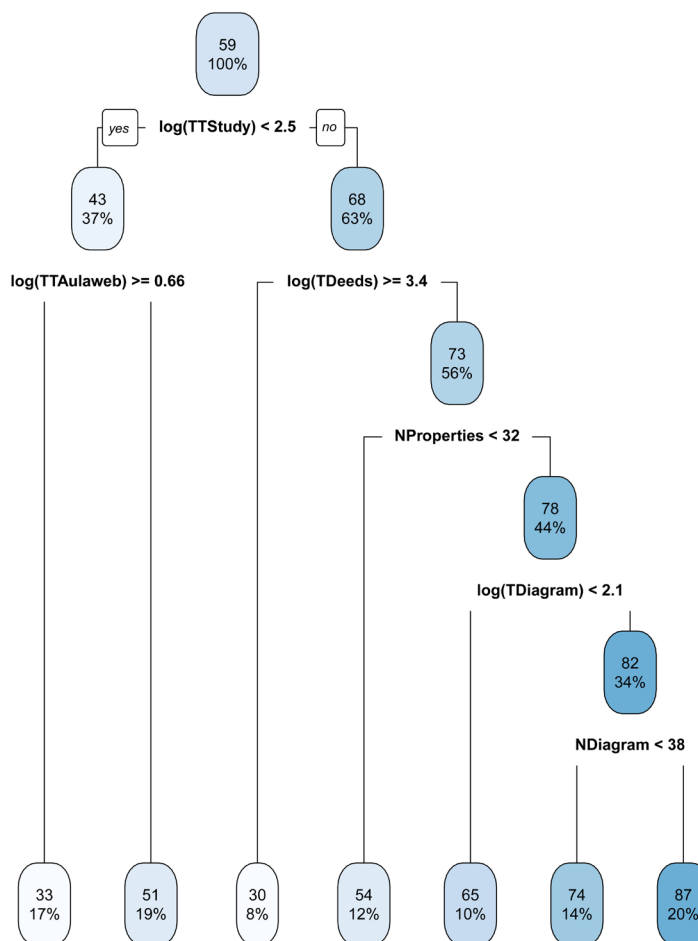


Figure 2. Regression tree model fitting results
图 2. 回归树模型拟合结果

3.3. 最佳预测课时节点选取

成绩预警分析的目的在于通过预先发出警报, 对学生学业成绩进行沟通 and 干预, 促使预警主体提早作出对应的决策, 从而形成由被动转向主动的教学关怀机制, 提升学生的学习效果。由上文分析可知, 在对学生能否通过期末测验进行预警分析时, 适宜选用精确度更高、错判率更低的分类树模型。而在具体分析学生期末测验成绩的高低时, 回归树模型具有更小的预测误差。在解决了模型适用性这一问题的前提下, 本文希望能进一步推测出在何时运用上述预警模型时, 既能使预测误差较小, 又能使干预者尽早的找到正确的预警方向, 实现更快更好的预警效果。

基于这一设想, 本文拟对最佳预测课时节点的选取进行建模分析, 选用五折交叉验证的方式分别对截止到 Session 3、Session 4、Session 5、Session 6 课时下学生的学习情况拟合分类树和回归树模型, 得到每课时节点下模型的平均误判率或平均绝对百分误差的均值, 进行分析。本文在 k 折交叉验证的理论基础上进行创新, 从不同课时节点的角度进行分析, 旨在从精准预测和及早干预中寻求最佳平衡点。

交叉验证得到的平均误判率和平均绝对百分误差的均值如表 5 所示。从表中可以看出, 当采用分类树进行分类预警分析时, 预测学生成绩的平均误判率会随着课时数的增加而不断减小, 但是较晚的课时节点并不利于成绩预警的目的, 不具备实践意义, 因此, 在希望既能及早干预又能精准预测的前提下, 本文认为, 可在 Session 4 或 Session 5 结束后, 即课程学习的中后段对学生能否通过测试进行预测, 并根

据预测结果对存在不合格风险的学生进行教学干预和指导, 学生也可针对自身特点进行调整, 改进学习效果。

Table 5. The results of cross-validation for different class nodes

表 5. 对不同课时节点进行交叉验证的误差结果

预测课时节点	分类树误差 (平均误判率)	回归树误差 (平均绝对百分误差均值)
Session 3	0.464	0.691
Session 4	0.438	0.692
Session 5	0.423	0.743
Session 6	0.324	0.656

对成绩高低的预测也可作类似的分析。当采用回归树进行成绩高低的预警分析时, 预测误差始终较高, 但与分类预警分析相同, 采用全部课时下的学习行为数据作为原始数据集进行预测时, 学生的学习行为表现最充分, 预测误差最低, 预测效果最好, 但这并不符合预警的初衷。对此, 本文认为, 在预测误差几乎平稳且较高的情况下, 及早作出预警模型进行干预是最合理的选择, 可考虑在课程学习前中期对学生最终取得成绩的高低进行预测, 根据模型结果对每位同学进行个性化的指导, 查缺补漏, 调整学习状态, 以取得更好的学习成果。

4. 结论

随着科学技术的不断发展, 在线网络学习平台在教学实验中的应用越来越广泛, 然而国内外对于学习行为分析的研究还处于探索阶段。本文基于真实的网络平台学习实例, 从多种角度尝试学习分析技术, 通过对学习行为进行“频次”和“时长”两方面的量化来分析学生在虚拟网络学习课程中的学习行为总体情况, 探询学习行为与学习效果之间的关联, 通过学习干预和指导来提升学习效果, 得出结论如下:

1) 在对学生能否通过最终测验进行预警分析时, 分类树相比于 Logistic 回归模型具有更高的精确度以及更低的错判率, 更适宜用来进行不及格预警分析。并且, 分类树选择了 TStudy、TOther、TAulaweb、TTextEditor、NDeeds、NBlank 作为重要影响变量。这说明, 学生期末测验的成绩合格率不仅与学生的线上学习与实操锻炼有关, 也与学生在学习过程中的专心程度有密切的联系。因此, 在今后的学习中, 老师和学生应密切关注学生对在线课程的学习时长、电路模拟操作次数以及报告撰写时长; 在学习态度的把控上, 应鼓励学生专心、踏实的进行学习。

2) 在对学生最终测验成绩的具体分数高低情况进行预测时, 回归树相比于传统线性回归模型具有更高的拟合度, 预测误差更小, 更适宜用来进行分数预警分析。并且, 回归树筛选出了 TTStudy、TTAulaweb、TDeeds、NProperties、TDiagram、NDiagram 作为重要的影响变量。由此说明, 在今后的教学中, 老师和学生应重视学生对在线课程的学习、时绘电路图的绘制以及电路模拟操作; 在选择问题解决方式时, 鼓励学生采取试错和查阅学习资料的方式寻求答案。

3) 针对不同的预警目的采取不同的预测课时节点, 既能使预测误差较小, 对学生的指导作用更准确, 又能使干预者尽早的找到正确的预警方向, 实现更快更好的预警效果。研究表明, 在对学生能否通过最终测验进行不及格预警分析时, 可于课程学习的中后期阶段进行预测, 并根据预测结果对存在不合格风险的学生进行教学干预和指导, 学生也可针对自身特点进行调整, 改进学习效果; 在对学生最终测验成绩的具体分数高低情况进行分数预警分析时, 可考虑在课程学习前中期阶段进行预测, 根据模型结果对每位同学进行个性化的指导, 查缺补漏, 调整学习状态, 以取得更好的学习成果。

基于前人研究基础, 本文在研究内容上的创新点主要有如下几处: 首先同时考虑了不及格预警模型和分数预警模型并采取了特定的筛选标准选择出更精准的预警模型; 其次通过交叉验证的方式选择出最佳预测课时节点。但是本文的研究总体来说是一个初步的探索性研究, 研究结论还存在一些不足。例如, 本文用于学习分析的样本数量不够充分, 数据类型较为简单, 因此在解释层面可能会存在缺陷。此外, 本文仅以该门数字电路实验课程为例对不同的学习行为进行了分析, 并未考虑到不同的实验主题下, 学生的学习行为是否会有差异。郑富龙[18]早在 2005 年时就对传统实验和虚拟仿真实验在教学方面的异同进行了对比分析, 作者认为, 教学模式和实验主题的契合度并不是一成不变的, 不同的实验主题应当搭配不同的教学模式, 例如, 当进行综合性实验或者设计性实验时, 教学者其实更适宜采用虚拟仿真实验的教学模式, 这种情况下学习者的学习接受效果更好。

本文的研究旨在为今后的在线网络学习平台的教学活动开展、教学干预、学生的自我认知和调整以及教学结果评价提供一定的参考。为了完善研究发现, 提升研究价值, 未来将开展的研究工作主要包括:

1) 利用其他的数据挖掘方法对学习行为数据进行进一步的探索, 发现更多隐藏在学生操作行为背后的认知过程, 设计出更加科学细致的预警模型, 为教与学提供有效的数据支撑。

2) 提升学习分析角度的多样性, 综合应用多种分析方法增强结论的可靠性和完备性, 例如, 在网络课程学习之余, 研究者可以通过对学习者的访谈或问卷调查等方式来收集其在情感和态度上的有价值的信息, 这种质性研究的方法可以弥补纯量化分析而导致的信息缺失问题, 使得学习分析这一工具在教育过程优化这一进程中发挥更大的作用。

3) 学习分析对学习成果的促进和优化作用被证实是确实存在的, 但长期影响效应的好坏还尚未可知。从前人的研究来看, 学习分析作为预警工具有助于学习者进行自我把控, 调整自己的学习状态和学习习惯, 进而对后续学习产生有利影响。因此, 本文希望能够在更长的一段时间内, 追踪学习者的学习行为表现, 通过技术手段, 分析学习分析这一工具对学习者的在线学习行为的全面影响机制。

参考文献

- [1] DeBoer, J., Stump, G., Seaton, D. and Breslow, L. (2013) Diversity in MOOC Students' Backgrounds and Behaviors in Relationship to Performance in 6.002 x. *Proceedings of the 6th Learning International Networks Consortium Conference*, Cambridge, 2013.
- [2] 李阳, 马力, 官巍. 大学生在线学习行为与人格特征的相关性研究[J]. 中国教育信息化, 2016(17): 18-21.
- [3] 宗阳, 陈丽, 郑勤华, 等. 基于在线学习行为数据的远程学习者学业情绪分析研究——以 Moodle 平台为例[J]. 北京广播电视大学学报, 2017(6): 11-20.
- [4] 陈晋音, 方航, 林翔, 郑海斌, 杨东勇, 周晓. 基于在线学习行为分析的个性化学习推荐[J]. 计算机科学, 2018, 45(S2): 422-426+452.
- [5] 赵蔚, 李士平. 基于学习分析的自我调节学习路径挖掘与反馈研究[J]. 中国电化教育, 2018, 381(10): 15-21.
- [6] 魏顺平. 在线教育管理者视角下的学习分析——在线教学绩效评估模式构建与应用[J]. 现代教育技术, 2014, 24(9): 79-85+93.
- [7] 刘双. 基于 xAPI 的网络学习行为数据采集工具设计与实现[D]: [硕士学位论文]. 武汉: 华中师范大学, 2016.
- [8] 洪丹丹, 李飞, 姚磊, 徐墨, 锁志海. 在线学习行为分析数据可视化快速开发框架设计与实践[J]. 微电子学与计算机, 2018, 35(7): 6-12.
- [9] 谭召, 鲁雁, 武斌. 在线学习平台设计与学习行为分析[J]. 信息与电脑(理论版), 2018(8): 118-120.
- [10] 谢修娟, 李香菊, 操凤平, 孙丽. 基于改进 C4.5 的 E-Learning 教学辅助系统的研究与实现[J]. 佳木斯大学学报(自然科学版), 2018, 36(1): 64-67+82.
- [11] Jiang, S., Williams, A.E., Schenke, K., Warschauer, M. and O'Dowd, D. (2014) Predicting MOOC Performance with Week 1 Behavior. *Proceedings of the 7th International Conference on Educational Data Mining*, Cordoba, 2014.
- [12] Kloft, M., Stiehler, F., Zheng, Z. and Pinkwart, N. (2014) Predicting MOOC Dropout over Weeks Using Machine

Learning Methods. *Proceedings of the EMNLP Workshop on Modeling Large Scale Social Interaction in Massively Open Online Courses*, Doha, October 2014, 60-65. <https://doi.org/10.3115/v1/W14-4111>

- [13] Halawa, S., Greene, D. and Mitchell, J. (2014) Dropout Prediction in MOOCs Using Learner Activity Features. *Proceedings of the European MOOC Summit*.
- [14] 李阳. 大数据环境下在线学习行为分析模型研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2017.
- [15] 王慧芬. 基于 Hadoop 的学习者在线学习行为研究[D]: [硕士学位论文]. 上海: 华东师范大学, 2017.
- [16] 吴林静, 劳传媛, 刘清堂, 程云, 毛刚. 网络学习空间中的在线学习行为分析模型及应用研究[J]. *现代教育技术*, 2018, 28(6): 46-53.
- [17] Vahdat, M., Oneto, L., Anguita, D., Funk, M. and Rauterberg, M. (2015) A Learning Analytics Approach to Correlate the Academic Achievements of Students with Interaction Data from an Educational Simulator. In: Conole, G., *et al.*, Eds., *EC-TEL 2015*, LNCS 9307, Springer, Berlin, 352-366. <https://doi.org/10.1007/978-3-319-24258-3>
- [18] 郑富龙. 仿真实验与传统实验的教学探索[J]. *实验室研究与探索*, 2005(9): 84-86.