

可变形卷积神经网络赋能学生在线学习退课预测

邵永伟, 马 坤, 荆 山, 陈贞翔

(济南大学 信息科学与工程学院, 山东 济南 250022)

摘 要: 针对大规模开放在线课程(MOOC)中高退课率现象, 基于可变形卷积神经网络构建预测模型, 以预测学生的退课结果。通过KDD Cup 2015数据集验证模型效果, 实验结果表明, 该方法优于现有主流的机器学习和深度学习模型。此外, 通过分析学生行为数据以解释退课预测结果, 从而提高预测结果可信度。学习退课预测分析有助于教师对学生进行个性化指导, 并可作为智能助手为学生提供更有针对性的学业帮助。

关键词: 大规模开放在线课程; 退课预测; 学业帮扶; 智能助教; 智能伴学

DOI: 10.11907/rjdk.251112

扫描二维码阅读全文: 

中图分类号: TP18; G434

文献标识码: A

文章编号: 1672-7800(2026)004-0191-08

Student Drop-out Prediction for Online Learning with Deformable Convolutional Networks

SHAO Yongwei, MA Kun, JING Shan, CHEN Zhenxiang

(Information Science and Engineering, University of Ji'nan, Ji'nan 250022, China)

Abstract: To address the issue of high dropout rates in Massive Open Online Courses (MOOCs), this paper has built a predictive model based on deformable convolutional networks to predict student dropouts. With the KDD Cup 2015 dataset, experimental results demonstrate that this method outperforms existing machine learning and deep learning models. Additionally, by analyzing student behaviors to explain dropout predictions, this paper has improved the credibility of the results. This analysis of dropout prediction assists teachers in providing personalized guidance and help students master the learning as an intelligent assistant.

Key Words: massive open online courses; dropout prediction; academic support; intelligent teaching assistant; intelligent companion learning

0 引言

随着网络技术的快速发展, 大规模在线开放课程(MOOC)应运而生^[1]。目前平台上的大多数课程可以免费获取, 学生可以享受视频课程、论坛和在线测试, 并且可以零成本退出课程, MOOC平台面临着高退课率的问题^[2]。统计数据显示, MOOC课程的退课率高达90%^[3-4], 课程完成率仅为4.5%~7%^[5-6], 这意味着大多数学生未能完成课程^[7]。因此, 通过分析学生学习行为特征以预测退课情

况, 并干预学生行为以提高课程完成率, 对于提升平台价值具有重要意义^[8]。

人工智能(AI)的快速发展为高等教育带来了深刻变革, 特别是随着在线教育和生成式模型的推广应用, AI与高等教育的联系日益紧密, 并取得了显著成效。在全球教育越来越重视个性化学习的趋势下, 可通过构建个性化学习路径和课程推荐系统以支持学生的个性化学习。传统教学评价方法过程静态、固定、单一, 难以全面跟踪每位学生的发展状况^[9]。而将AI应用于教学评价中, 可以持续收集学生学习行为数据, 实现学习质量的动态追踪。

收稿日期: 2025-03-22

基金项目: 山东省本科教学改革研究项目(Z2024139); 山东省高等教育学会高等教育研究专项课题(SDGJ2023C05); 济南大学教学研究重点项目(JZ2411); 山东省电化教育馆人工智能教育研究课题(SDDJ202501004)

作者简介: 邵永伟(1999-), 男, 济南大学信息科学与工程学院硕士研究生, 研究方向为计算机技术; 马坤(1981-), 男, 博士, 济南大学信息科学与工程学院教授, 研究方向为教育数据分析、自然语言处理、大数据处理; 荆山(1979-), 男, 济南大学信息科学与工程学院教授, 研究方向为计算机网络、教育数据分析; 陈贞翔(1979-), 男, 济南大学信息科学与工程学院教授, 研究方向为网络空间安全。本文通讯作者: 马坤。

为解决MOOC中的高退课率问题,许多研究人员提出了不同的方法。早期研究主要采用机器学习方法以预测学生的退课情况^[10-11]。近年来,随着深度学习技术在各领域的广泛应用和显著成效,深度学习已成为退课预测的主流^[5,12]。然而,现有的退课预测方法仍存在一些問題:由于学生的学习行为风格各异,现有模型通常侧重于挖掘学生的整体学习行为特征,而忽视了个体的局部空间个性化学习行为特征。为解决该问题,本文提出了一种基于可变形卷积神经网络的退课预测模型。

本文主要有以下3方面贡献:①使用可变形卷积核提取学生个性化学习行为特征,可变形卷积核通过调整采样位置,适应学生学习行为矩阵中的数据分布特征,既提取学生的全局学习行为特征,又挖掘局部空间中的个性化学习行为特征;②提出基于可变形卷积神经网络的退课预测模型,该模型首先利用可变形卷积神经网络提取学生个性化学习行为特征,然后将特征向量输入双向长短时记忆网络(BiLSTM)提取学习行为的时序特征,最终通过全连接层生成退课预测结果;与基线模型在KDD Cup 2015数据集上的比较显示,该模型性能优良;③根据学生学习行为数据分析学生学习风格并进行群体聚类,将课程周期细化为多个子周期,对不同学生群体在各子周期中的学习行为特点进行可视化,以解释退课预测结果。

1 相关工作

1.1 人工智能赋能高等教育

2024年4月,教育部高等教育司公布了首批“人工智能+高等教育”应用场景典型案例,提出积极推动高等教育与人工智能技术的融合,利用智能技术支撑人才培养模式创新、教学方法改革和教育治理能力提升。目前,人工智能在高等教育中已取得初步成果。例如,利用虚拟现实(VR)等技术构建虚拟实验平台,提高了学生的容错率并激发了学习热情。林毅贞等^[13]将AI技术应用于课程作业批改中,并调研了学生对智能批改的态度,结果表示学生对智能批改持欢迎态度。以ChatGPT为代表的生成式大语言模型对学生学习方式产生了重大影响。Guo等^[14]让学生在辩论前与聊天机器人互动,通过设置复杂性和验证质量等指标以评估学生能力。陆小飞等^[15]利用智能机器人进行口语和听力教学,取得了良好的教学效果。

1.2 人工智能赋能在线课程退课预测

1.2.1 传统机器学习方法

早期的退课预测研究主要依赖传统机器学习方法,将退课预测视为二分类任务。例如,Taylor等^[16]从原始学生学习日志中构造了27个特征,使用逻辑回归模型预测退课结果。Kloft等^[17]使用支持向量机模型,分析学生的历史学习行为数据,以预测学生下周参加课程的概率。Liu等^[18]使用K均值算法对退课进行预测,取得了良好效果。

Liang等^[19]利用XuetangX上的学生数据构造特征,使用梯度提升树模型对学生未来10天的退课情况进行预测。然而,传统机器学习方法需要人工构建复杂特征,并且对特征间关系的探究不足。

1.2.2 深度学习的方法

近年来,随着深度学习在各种领域对高维度复杂数据处理方面展现出优异性能,许多研究人员开始使用深度学习的方法解决退课预测任务中存在的问题。例如,Mubarak等^[20]将卷积神经网络和长短时记忆网络结合,用于预测退课结果。Mrhar等^[21]使用人工神经网络预测学生是否退课的同时,利用学生的点击流日志和论坛发帖数据对学生情绪进行分析。Fu等^[22]提出名为“CLSA”的管道模型,该模型集成了卷积神经网络、长短时记忆网络、静态注意力机制,将卷积神经网络生成的高维向量输入到长短时记忆网络(LSTM)中,以获得融入时间序列信息的向量表示,并通过静态注意力机制进一步优化特征权重。Zheng等^[23]将自注意力机制引入到模型中,挖掘学生学习行为之间的相关性;Chen等^[24]将改进的量子粒子群优化算法与深度学习模型结合,从而提出IQPSO-PLSTM模型,该模型首先使用随机森林对数据进行预处理,获取特征的重要性权重,然后通过改进的量子粒子群优化算法对权重进行优化,最终使用Peephole LSTM对带有优化权重的数据进行建模。Nithya等^[25]利用FP-Growth关联规则算法,从学生行为日志中挖掘学生学习行为之间的重要性关系,将其作为特征添加到人工神经网络中进行建模。Hancox等^[26]提出基于时序图的卷积神经网络,并将其用于退课预测,构建带权有向图,图中每个节点代表学生完成的每个事件,权重代表事件完成的时间间隔,将退课预测任务视为二元图分类任务。Niu等^[27]设计了一个卷积神经网络自编码器,对高维度学生行为特征进行降维,在保留学生代表性特征的同时还保留了特征间非线性关系。然而,传统卷积核由于采样位置固定,在面对学生不规则的数据分布时,容易忽视部分学习行为特征。同时,单向LSTM无法有效处理远距离的时间序列信息,限制了对时间特征的全面捕获。本文采用可变形卷积核,通过学习偏置以调整卷积核中像素的位置,使其在每个位置都能捕获学生的特征信息,从而增强信息提取能力。双向LSTM则考虑了两个时间方向的信息,能够有效处理序列信息中的远距离影响^[23]。

1.2.3 大语言模型

大语言模型凭借其卓越的上下文感知和复杂数据处理能力,在自然语言处理任务中表现优异,近年来也被应用于退课预测任务。唐瞻^[28]统计学生在课程周期中的学习行为发生次数,将这些行为进行One-Hot编码,并结合学生人口统计特征,输入BERT预训练模型进行退课预测,以捕获不同时段的行为关系。Won等^[29]通过文字描述学生的人口统计特征和课程成绩特征,将这些特征隐藏在语句中,并将退课结果以语言形式描述,再输入预训练模型,

以捕获类别特征之间的关系。然而,使用大语言模型的退课预测方法更多侧重于学生各种类特征之间的关系,忽视了学生学习行为中蕴含的个人学习风格信息。

2 整体架构

2.1 问题描述

在退课预测任务中,对退课的定义有所不同。本文利用学生在课程周期内的学习行为记录评估是否已退课。如果学生在连续10天内没有学习活动,则标记为退课学生;否则,标记为非退课学生。本文将退课预测问题转化

为二分类问题,其中0代表非退课学生,1代表退课学生。

2.2 模型框架

本文提出的预测模型主要分为4个阶段。首先,根据学生学习行为日志构造学习行为矩阵并将其作为模型输入;其次,可变形卷积神经网络通过动态调整采样网络以学习矩阵数据分布,并提取学生的个性化学习行为特征;使用两种不同尺寸的可变形卷积核以提取多尺度下的个性化学习行为特征;再次,BiLSTM对学生个性化学习行为特征进行时间序列特征提取,得到最终特征;最后,通过全连接层生成退课预测结果。本文模型整体结构如图1所示。

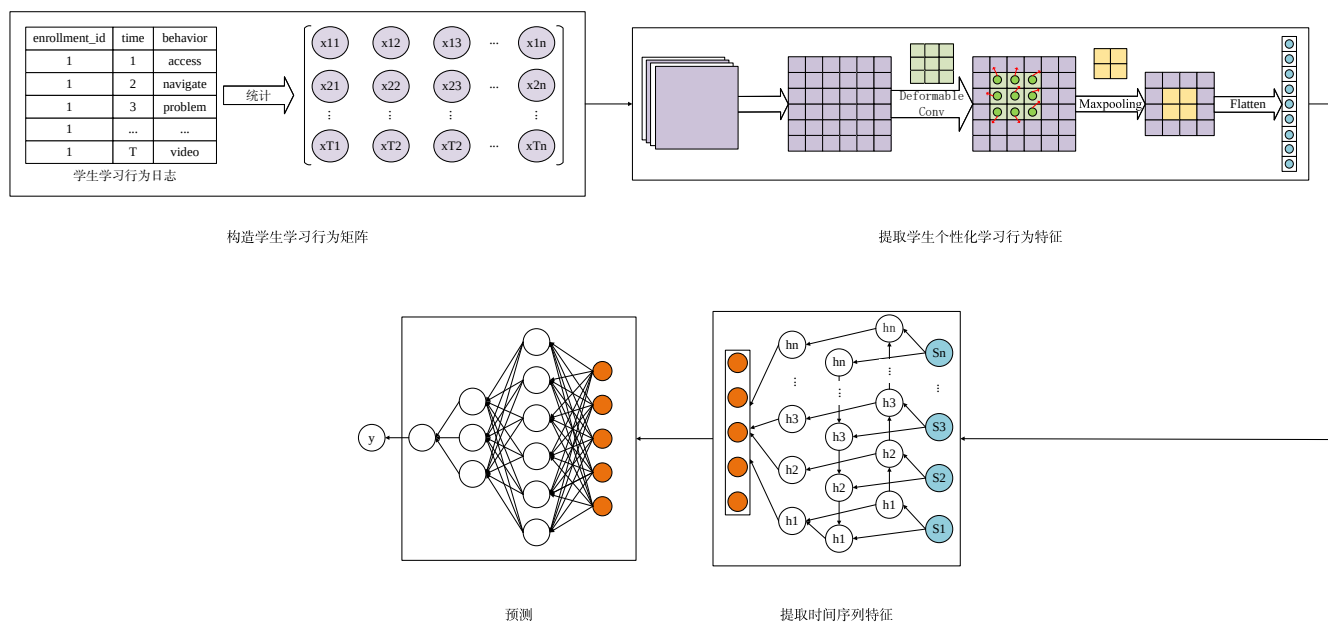


Fig. 1 The model structure

图1 模型结构

2.2.1 学生学习行为矩阵构建

本文通过统计学生在课程周期中各种学习行为的发生次数以构建学生学习行为矩阵。例如,如果学生在课程的第10天有15次“讨论行为”记录,则该学生学习行为矩阵中第10天的讨论行为次数为15。通过统计学生在 \$T\$ 天内 \$n\$ 种学习行为的发生频率,最终得到学生学习行为矩阵 \$X\$,如式(1)所示。

$$X_T = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^n \\ x_2^1 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ x_T^1 & x_T^2 & \cdots & x_T^n \end{bmatrix} \quad (1)$$

其中,水平维度代表学生学习行为特征,表示学生 \$n\$ 种学习行为发生频率统计次数;垂直维度代表课程周期,表示时间维度,共有 \$T\$ 天。

2.2.2 学生个性化学习行为特征提取

每个学生在不同时间段会有不同的学习行为,因而每个学生学习行为矩阵中的数据分布不同。将可变形卷积应用于学生的学习行为矩阵,分析学生学习数据分布,通

过学习卷积核中每个像素偏移量以改变卷积核形状,从而提取学生个性化学习行为特征。将学生学习行为矩阵 \$X\$ 作为输入,构建卷积层并使用卷积核对矩阵进行处理得到学生学习特征,如式(2)所示。

$$M = \text{DeformableConv2D}(X, L) \quad (2)$$

其中,DeformableConv2D表示二维的可变形卷积,\$L\$表示卷积核个数。

将卷积后的矩阵输入到MaxPooling最大池化层,以进行特征降维,同时防止模型过拟合并降低模型参数量。池化和展平操作如式(3)、式(4)所示。

$$N = \text{MaxPooling}(M) \quad (3)$$

$$F = \text{Flatten}(N) \quad (4)$$

2.2.3 时间序列特征提取

作为改良版的循环神经网络(Recurrent Neural Network, RNN),LSTM按照时间步重复接收序列特征 \$F_t\$,连接前一时刻的隐藏层状态 \$h_{t-1}\$,更新输入门 \$i_t\$ 以及候选细胞状态 \$\tilde{C}_t\$。通过遗忘门 \$f_t\$ 选择旧细胞状态信息中的部分信息,输入门 \$i_t\$ 从候选细胞状态信息中选择添加到新细胞状

态中的信息。更新完细胞状态后,根据输入前一时刻的隐藏层状态 h_{t-1} ,最终得到输出 h_t ,LSTM 中的门控和细胞更新公式如式(5)一式(10)所示。

$$i_t = \sigma(W_i \cdot [h_{t-1}, s_t] + b_i) \quad (5)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, s_t] + b_f) \quad (6)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, s_t] + b_c) \quad (7)$$

$$c_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (8)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, s_t] + b_o) \quad (9)$$

$$h_t = o_t \odot \tanh(c_t) \quad (10)$$

其中, σ 和 $\tanh$ 表示激活函数, $\cdot$ 和 $+$ 分别表示矩阵的点乘和矩阵相加运算, $\odot$ 表示逐元素相乘运算, W_i 、 W_f 、 W_c 、 W_o 分别表示输入门、遗忘门、候选细胞状态、输出门的权重矩阵, b_i 、 b_f 、 b_c 、 b_o 分别表示输入门、遗忘门、候选细胞状态、输出门的偏置向量。BiLSTM 是双向神经网络,最终输出 h_t 由正向 LSTM 结果 h_{t_e} 和反向 LSTM 结果 h_{t_f} 拼接而成,表示为式(11)。

$$h_t = [h_{t_e}, h_{t_f}] \quad (11)$$

3 实验方法

3.1 数据集

本文所采用的数据集源自真实世界场景下的 KDD CUP 2015 数据集,该数据集的数据采集自中国颇具影响力的在线大规模开放在线课程(MOOC)教育平台——学堂在线。该数据集包含 79 186 名学生在 39 门课程中的 120 542 条注册学习行为记录。数据集记录了学生在学堂在线平台上的 7 种学习行为: access、discussion、navigate、page_close、problem、video 和 wiki,每种学习行为的具体含义如表 1 所示。

每名学生可以同时注册多门课程,对应唯一的注册 id。学生的原始学习行为记录示例如表 2 所示。其中,en-

rollment_id 表示唯一学生所选课程的唯一注册 id;time 表示学习行为发生时间;source 表示学习行为发生所使用的端口类型,分为 browser 和 server;event 表示学习行为;objects 表示学生访问或导航的对象 id。

采用学生学习行为矩阵构建方法对学生原始日志数据进行处理,处理结果示例如表 3 所示。

3.2 评价指标

本文使用准确率(Accuracy)、精准率(Precision)、召回率(Recall)和 F1 分数评估模型性能。准确率表示模型预测正确的学生数量与所有模型预测学生数量间的比值,精准率表示在模型预测为非退课学生中真正为非退课学生的比例,召回率表示在模型预测为退课学生中被正确预测的比例,F1 分数通过平衡精准率和召回率对模型性能进行全面评估。4 种评价指标的数值越高,表示模型效果越好。

3.3 基线模型

本文选取了 4 个传统机器学习方法和 3 个深度学习模型作为基准模型。

LR^[16]:逻辑回归(Logistic Regression)模型,它是一种广义的线性回归分析模型。

DT^[30]:决策树(Decision Tree)模型,它通过连续的逻辑判断得出最后预测结果。

RF^[31]:随机森林(Random Forest)模型,它利用了 Bagging 集成思想,集成了多个决策树模型的分类器。

NB^[32]:朴素贝叶斯(Naive Bayes)网络,它是一种基于贝叶斯定理与特征条件独立假设的分类方法。

CNN^[33]:卷积神经网络(Convolutional Neural Network),数据首先经过卷积神经网络,之后依次经过池化和全连接操作,从而得到预测结果。

LSTM^[34]:长短时记忆网络(Long Short Term Memory)模型,它是一种改进的循环神经网络模型,可以提取出学生学习行为中的时序特征。

CNN-LSTM:将卷积神经网络和长短时记忆网络融合的模式,该模型使用传统卷积核提取学生的学习行为特征。

3.4 相关设置

将数据集中训练集和测试集的比例按照 8:2 进行划分,模型中两个可变形卷积核大小分别设置为 3×3、2×2,步长为 1。为了解决过拟合问题,将 Dropout 的值设置为 0.4, Batch_size 设置为 256,学习率为 0.000 1,运行轮次设置为 20。使用反向传播算法对模型进行训练,由于模型执行二

Table 1 Explanation of the student learning behavior
表 1 学生学习行为含义说明

行为	含义
access	访问课程中除去观看视频和做课程作业以外的其他对象
discussion	参加课程讨论
navigate	导航至课程其他部分
page_close	关闭网页
problem	做课程作业
video	观看课程视频
wiki	浏览课程维基百科

Table 2 Example of student's original learning behavior record

表 2 学生原始学习行为记录示例

enrollment_id	time	source	events	object
1	2014-06-14T09:38:29	server	navigate	Oj6eQgzrdqBMLaCtaq1IkY6zruSrb71b
1	2014-06-14T09:38:39	server	access	3T6XwoiMKgol57cm29Rjy8FXVFcIomxl
1	2014-06-14T09:38:39	server	access	2cmZrZW2h6l191it03e89FGcABLWhf3W
1	2014-06-14T09:38:48	server	page_close	RMtgC2bTAqEftenUUyia504wsyzeZWf
1	2014-06-14T09:41:49	browser	problem	Oj6eQgzrdqBMLaCtaq1IkY6zruSrb71b

分类任务,因此选择二元交叉熵函数作为损失函数,使用 Adam 优化器以优化模型。本文所有实验均在相同配置环境下运行,实验环境配置如表 4 所示。

Table 3 Example of data processing results
表 3 数据处理结果示例

enrollment_ id	time	access	discussion	navigate	page_ close	problem	video	wiki
1	0	20	0	0	17	0	14	0
1	1	0	0	0	0	0	0	1
1	2	21	0	5	13	27	1	0
1	3	9	5	4	0	0	3	0
1	4	14	0	4	2	14	1	0
1	5	3	41	0	2	0	2	3

Table 4 Environmental configuration
表 4 环境配置

实验环境	实验配置
CPU	16 vCPU Intel(R) Xeon(R) Gold 6430
GPU	NVIDIA RTX 4090, 24 GB GDDR6X
操作系统	Windows 11
编程语言	Python
框架	Pytorch

4 实验结果分析

4.1 性能比较

本文使用 KDD Cup 2015 数据集以验证本文模型性能,并将模型与基线模型性能进行比较,结果如表 5 所示(实验结果中的最佳值予以加粗显示)。

Table 5 Comparison results of the proposed model and baseline models on the KDD Cup2015 dataset

表 5 本文模型与基线模型在 KDD Cup2015 数据集上的性能比较结果

Model	Accuracy	Precision	Recall	F1
LR	0.862 3	0.853 1	0.862 5	0.852 4
DT	0.806 0	0.869 6	0.887 6	0.878 5
RF	0.865 1	0.879 6	0.946 2	0.911 6
NB	0.854 6	0.773 2	0.733 1	0.770 0
CNN	0.860 7	0.865 8	0.934 0	0.914 3
LSTM	0.864 5	0.870 1	0.932 5	0.901 2
CNN-LSTM	0.871 1	0.880 0	0.952 1	0.915 6
Ours	0.875 3	0.895 4	0.960 0	0.923 5

可以看出,由于使用了 Bagging 集成思想,集成了多个决策树的随机森林模型性能表现显著优于其他传统机器学习方法,Accuracy 为 86.51%,Precision 为 87.96%,Recall 为 94.62%,F1 为 91.16%。与朴素贝叶斯方法相比,CNN 方法在 Precision 和 Recall 两个指标上提升明显,分别提升了 9.26% 和 2.009%,这是由于深度学习方法可以自动提取高级特征,因而取得了较好性能。本文模型与传统的机器学习方法 LR 相比,在指标 Recall 和 F1 上分别提升了 9.29% 和 6.99%,说明本文模型在特征学习方面具有显著优势。与 CNN 相比,由于加入 LSTM 后提取了数据中的时序特征,因而 CNN-LSTM 模型在 Accuracy 和 Precision 两个指标

上分别提高了 1.04% 和 1.42%。CNN-LSTM 模型同时考虑了学生的学习行为特征和时序特征,与两个模型相比,其整体性能得到了提升。本文模型使用了可变形卷积核,根据不同的数据分布动态调整采样位置,在局部空间中提取到更多的学生个性化学习行为特征。与 CNN-LSTM 模型相比,模型在 4 个指标上的性能均得到了一定程度的提升,模型在 Accuracy、Precision、Recall 和 F1 上的表现分别为 87.53%、89.54%、96% 和 92.35%。

4.2 学生退课原因分析

本文通过分析学生学习行为的特点,对退课预测结果作进一步解释。这种方法的优点在于帮助学生在退课行为发生前识别自身不足,进而改进学习行为,同时降低课程退课率,减少课程资源浪费。本文采用分层聚类方法,设定聚类数为 3。根据课程注册 ID 数量确定学生数量,依据学生在课程周期中学习行为的发生次数对学生群体进行聚类。聚类分析结果如表 6 所示。

Table 6 Clustering results of student group
表 6 学生群体聚类结果

类别	非退课学生 数量	退课学生数量	非退课学生 占比/%	退课学生 占比/%
0	11 837	5 143	69.71	30.29
1	12 654	90 396	12.28	87.72
2	470	42	91.79	8.21

可以看出,类别 0 学生群体中,非退课学生与退课学生的比例约为 7:3,这一类学生群体在整体学生群体中的人数比例相对平衡。类别 1 学生群体中,非退课学生占比为 12.28%,退课学生占比为 87.72%。这说明大部分学生在这一类群体中选择退出课程,显示出这类学生的学习行为更容易导致课程失败。类别 2 学生群体中,仅有 512 名学生,且非退课学生占比接近 92%。这表明这类学生的学习行为对课程学习有积极作用,但保持这种学习行为的学生较为少见。本文从两个角度进行探讨:一是通过折线图展现每类学生群体学习行为在整个课程周期内的变化,分析其与退课行为之间的关系;二是通过柱状图体现每类学生的每个学习行为在 4 个课程子周期中的占比,分析每类学生群体可能的退课原因。最后,对每类学生群体的退课原因进行总结,并给出针对性的改进意见。

4.2.1 学生学习行为变化趋势分析及其对退课的影响

为了进一步分析学生退课原因,本文通过分析学生的聚类标签和学习行为记录,统计了每类学生在课程周期中每天所有学习行为的平均发生次数和平均持续时间。“访问课程”“导航至课程其他部分”等行为对学习质量影响有限,而“做课程作业”“观看课程视频”等行为在一定程度上反映了学生学习质量。因此,本文选取“problem”、“video”、“wiki”和“discussion”作为有效学习行为,统计了 3 类学生在课程周期中有效学习行为的平均发生次数。统计结果如图 2—图 4 所示。

从图 3 可以看出,类别 2 学生群体在课程周期中的学

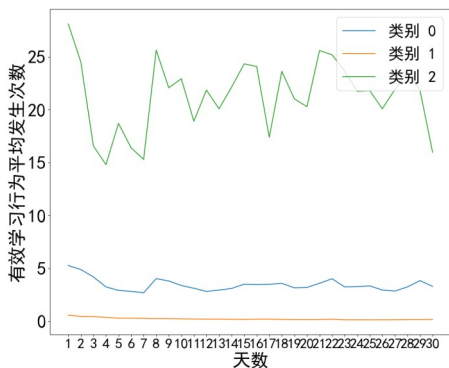


Fig. 2 The average frequency of effective learning behaviors among students

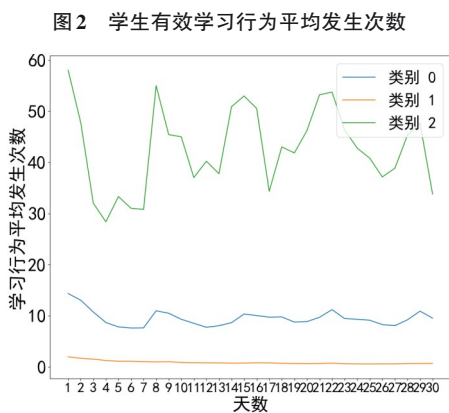


Fig. 3 The average numbers of occurrences of all learning behaviors among students

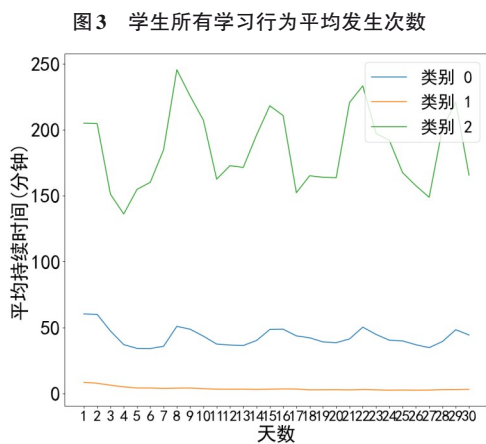


Fig. 4 Average duration of student learning behavior

图4 学生学习行为平均持续时间

习行为发生最频繁,次数变化范围为20~60次,且每天学习时间显著高于其他两类群体,某一天的学习行为持续时间达250 min。这表明,类别2学生对课程有高投入和持续学习兴趣,有效学习行为次数也维持在较高水平,显示出较高的学习效率和质量。类别0学生群体的学习行为次数和持续时间相对平稳,通常在10次左右,学习时间约50 min。这表明,该群体的学习行为较为一致,但活跃度不高,可能需要更多激励或支持以提升学习积极性。类别1学生群体的学习行为数据始终较低,每天的有效学习行为不超过5次,学习时间不足5 min。这表明,该群体对课程的投入

明显不足,进一步验证了表6中的聚类结果。

综合以上分析可知,类别2学生群体在课程周期内表现出较高的学习活跃度和学习质量。相对而言,类别0和类别1的学生群体则需要更多关注和干预,以提升他们的学习动力和参与度。

4.2.2 课程子周期内学习行为分布与学生退课行为关联分析

为了研究学生每个学习行为在不同学习阶段中的占比,本文将课程周期划分为4个子周期,记为w1、w2、w3和w4,分析了3类学生群体在这4个子周期中7种学习行为的平均发生次数占比。统计结果如图5—图7所示。

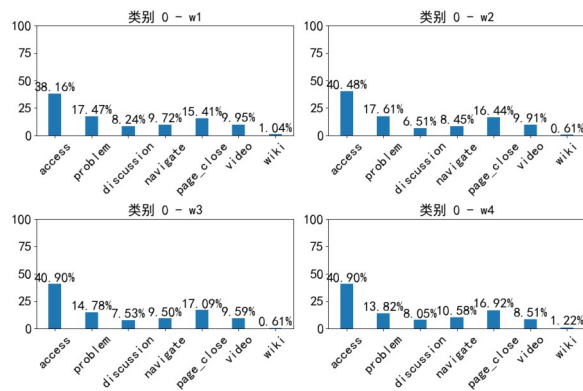


Fig. 5 Proportion of average occurrences of 7 learning behaviors among cluster 0 student groups in four course sub cycles

图5 类别0学生群体在4个课程子周期中7种学习行为发生次数平均占比

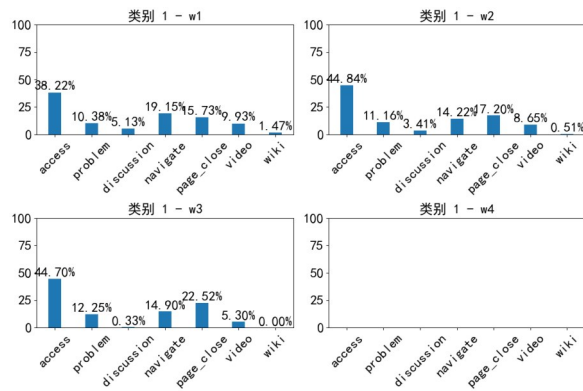


Fig. 6 Proportion of average occurrences of 7 learning behaviors among cluster 1 student groups in four course sub cycles

图6 类别1学生群体在4个课程子周期中7种学习行为发生次数平均占比

类别1学生在学习过程中,整体上以访问课程资源(access)为主,并且该行为在各子周期中均占最高比例。其中,在w1~w3阶段,access行为分别占比38.22%、44.84%和44.70%,但在w4阶段,该群体的所有学习行为数据均缺失,可能表明这部分学生在后期放弃了学习。此外,该群体在w1~w3阶段提交课程作业行为(problem)占比分别为10.38%、11.16%和12.25%,体现出该群体在学习初期保持了一定的作业参与度,但整体作业参与率较低。其他行为

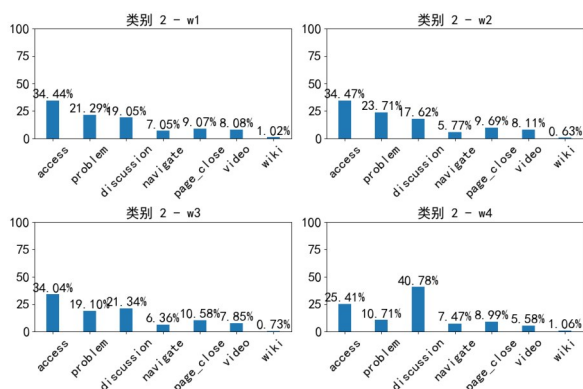


Fig. 7 Proportion of average occurrences of 7 learning behaviors among cluster 2 student groups in four course sub cycles

图7 类别2学生群体在4个课程子周期中7种学习行为发生次数平均占比

方面,导航至课程其他部分行为(navigate)在w1-w3阶段分别为19.15%、14.22%和14.90%,关闭页面行为(page_close)在w1-w3阶段分别占比15.73%、7.20%和22.52%,表明该群体在课程中的页面跳转和退出行为较为频繁。此外,该群体对观看课程视频(video)和浏览课程维基百科(wiki)的兴趣较低。该群体的学生在前期较为活跃,频繁访问课程资源,但对观看课程视频和浏览课程维基百科等辅助学习资源的使用较少,并且在课程后期可能出现放弃学业的现象。

类别2学生群体在4个子周期中整体保持较为稳定的学习模式,其访问课程资源在w1-w3阶段分别为34.44%、34.47%和34.04%,在w4阶段有所下降,为25.41%。值得注意的是,该群体的参与课程讨论(discussion)行为占比相对较高,在w1-w3阶段分别占比9.05%、7.62%和21.34%,在w4阶段更是达到了40.78%,说明该群体在后期更倾向于通过讨论参与学习。此外,该群体的作业提交行为在w1-w4阶段占比分别为21.29%、23.71%、19.10%和10.71%,整体表现较为积极,尤其在前期作业参与较多,但在w4阶段有所下降。此外,该群体的观看课程视频和浏览课程维基百科行为仍然较少,但整体比类别1学生更为稳定。该群体学生具有较为稳定的学习习惯,除频繁访问课程资源外,还较多参与课程讨论,并在前期保持较高的作业参与度。但在后期,访问课程资源和作业提交行为有所下降,而课程讨论行为大幅上升,表明该群体更倾向于在学习后期通过讨论获取知识。

类别0学生群体的整体学习模式介于类别1和类别2之间,其访问课程资源在w1-w4阶段分别占比38.16%、40.48%、40.90%和40.90%,在4个子周期中保持相对稳定。提交课程作业行为在w1-w4阶段分别占比17.47%、17.61%、14.78%和13.82%,整体下降趋势较为缓慢。此外,该群体的参与课程讨论行为在w1-w4阶段占比分别为8.24%、6.51%、7.53%和8.05%,相对较低,说明该群体在学习过程中更倾向于独立学习,而不是通过讨论交流获取知

识。值得注意的是,该群体的关闭页面(page_close)行为在w1-w4阶段分别占比15.41%、16.44%、17.09%和16.92%,表明这部分学生的页面切换和退出行为较为稳定。观看课程视频和浏览课程维基百科行为同样较少。该群体学生在整个课程周期内保持较为稳定的访问课程资源行为,作业提交行为略有下降,并且对课程讨论的参与度相对较低,可能倾向于独立完成学习任务。此外,该群体的页面关闭行为相对较高,可能意味着他们的学习方式较为快速,习惯于频繁切换页面。

通过对3类学生群体在4个课程子周期中的学习行为进行分析可以发现,不同群体在学习过程中呈现出不同的行为模式。类别0学生访问课程资源的学习行为最稳定,作业提交行为逐步下降,讨论参与度较低,可能倾向于独立完成学习任务。类别1学生主要依赖访问课程资源,但作业提交较低,对辅助学习资源(如观看课程视频和浏览课程维基百科)兴趣较小,并且在后期可能放弃学习。类别2学生具有较稳定的学习习惯,积极参与作业提交,并在后期更倾向于通过课程讨论获取知识。

4.2.3 学生退课原因总结及改进意见

结合以上分析可知,3类学生的退课原因各有不同:类别0学生的退课原因可能是学习方式过于独立、缺乏与其他学生的互动、对课程逐渐失去兴趣,从而退课,建议其参与更多的学习活动,如小组讨论、答疑环节,增强学习归属感,以提高学习体验与兴趣;类别1学生的退课原因可能是缺乏持续的学习动力、课程活动参与度低、对课程兴趣不足而退课,建议对此类学生提供更多外部激励,鼓励他们利用更多的学习资源,同时增加自己的学习投入;类别2学生的退课原因可能是后期作业难度加大、学习压力上升,过度依赖课程讨论但未能解决遇到的问题而导致退课,建议此类学生利用更多的学习资源,寻找有效的学习支持(如教师答疑或更具针对性的讨论),避免因为作业压力过大而退课。

5 结语

本文提出了一个基于可变形卷积神经网络的学生退课预测模型,该模型通过可变形卷积核学习学生学习行为矩阵中的数据分布,动态调整采样位置,从而更有效地提取包含学生个性化学习行为的特征。最终,特征向量经过全连接层生成退课预测结果。实验在KDD Cup 2015数据集上开展,与使用传统机器学习和深度学习方法的基线模型相比,本文模型在4个评价指标上表现更佳。此外,结合聚类方法分析学生群体,揭示了每类学生的退课预测原因,提升了预测结果的可解释性。

然而,该模型仍存在一些不足。比如:对学生使用不同资源次数等特征的提取不够充分。未来工作中,可提取更多维度的学生特征,并使用更多数据集以验证模型

性能。

参考文献:

- [1] WATTERS A. MOOC mania: debunking the hype around massive open online courses [EB/OL]. <https://www.slj.com/story/got-mooc-massive-open-online-courses-are-poised-to-change-the-face-of-education>.
- [2] LIPI M, BERTINI M, FRASCONI P, et al. Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(2): 871-882.
- [3] JORDAN K. Initial trends in enrolment and completion of massive open online courses [J]. International Review of Research in Open and Distance Learning, 2014, 15(1): 133-160.
- [4] MEYER R. What it's like to teach a MOOC (and What the Heck's a MOOC?) [EB/OL]. <https://www.theatlantic.com/technology/archive/2012/07/what-its-like-to-teach-a-mooc-and-what-the-hecks-a-mooc/260000/>.
- [5] SINHA T, JERMANN P, LI N, et al. Your click decides your fate: inferring information processing and attrition behavior from MOOC video click-stream interactions [DB/OL]. <https://arxiv.org/abs/1407.7131>.
- [6] FENG W, TANG J, LIU T X. Understanding dropouts in MOOCs [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2019: 517-524.
- [7] MOHAMAD N, AHMAD N B, JAWAWI D N A, et al. Feature engineering for predicting MOOC performance [J]. IOP Conference Series Materials Science and Engineering, 2020, 884(1): 012070.
- [8] ZHANG H F, CHEN Z X, XU S J. Artificial intelligence enables new reform of higher education pattern [J]. Software Guide, 2022, 21(11): 166-171.
张慧峰, 陈竹湘, 徐水晶. 人工智能赋能高等教育模式新变革 [J]. 软件导刊, 2022, 21(11): 166-171.
- [9] KOTSIANTIS S. Educational data mining: a case study for predicting dropout-prone students [J]. International Journal of Knowledge Engineering and Soft Data Paradigms, 2009, 1(2): 101-111.
- [10] XING W, CHEN X, STEIN J, et al. Temporal predication of dropouts in MOOCs: reaching the low hanging fruit through stacking generalization [J]. Computers in Human Behavior, 2016, 58: 119-129.
- [11] DESMARAIS M C, NACEUR R, BEHESHTI B. Linear models of student skills for static data [C]//Montreal: UMAP Workshops, 2012.
- [12] NIU K, ZHOU Y, LU G, et al. PMCT: parallel multiscale convolutional temporal model for MOOC dropout prediction [J]. Computers and Electrical Engineering, 2023, 112: 108989.
- [13] LIN Y Z, HUANG P. Research on the technology of intelligent correction of online homework system for "fundamentals of mechanical design" [J]. Technology Innovation and Application, 2019(29): 145-147, 149.
林毅贞, 黄平. 智能批改《机械设计基础》网上作业系统的技术研究 [J]. 科技创新与应用, 2019(29): 145-147, 149.
- [14] GUO K, ZHONG Y, LI D, et al. Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation [J]. Computers & Education, 2023, 203: 104862.
- [15] LU X F, LIAO J, XU Q. The current status and prospects of the application research of educational robots in foreign language oral teaching [J]. Foreign Language World, 2021(1): 11-19.
陆小飞, 廖剑, 许琪. 教育机器人在外语口语教学中的应用研究现状及前瞻 [J]. 外语界, 2021(1): 11-19.
- [16] TAYLOR C, VEERAMACHANENI K, O'REILLY U M. Likely to stop? predicting stopout in massive open online courses [DB/OL]. <https://arxiv.org/abs/1408.3382>.
- [17] KLOFT M, SITEHLER F, ZHENG Z, et al. Predicting MOOC dropout over weeks using machine learning methods [C]//Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs, 2014: 60-65.
- [18] LIU T Y, XIU L. Finding out reasons for low completion in MOOC environment: an explicable approach using hybrid data mining methods [C]//2017 International Conference on Modern Education and Information Technology, 2017: 376-384.
- [19] LIANG J, YANG J, WU Y, et al. Big data application in education: dropout prediction in Edx MOOCs [C]//2016 IEEE Second International Conference on multimedia big data (BigMM), 2016: 440-443.
- [20] MUBARAK A A, CAO H, HEZAM I M. Deep analytic model for student dropout prediction in massive open online courses [J]. Computers & Electrical Engineering, 2021, 93: 107271.
- [21] MRHAR K, DOUIMI O, ABIK M. A dropout predictor system in MOOCs based on neural networks [J]. Journal of Automation Mobile Robotics and Intelligent Systems, 2020, 14(4): 72.
- [22] FU Q, GAO Z, ZHOU J, et al. CLSA: a novel deep learning model for MOOC dropout prediction [J]. Computers & Electrical Engineering, 2021, 94: 107315.
- [23] ZHENG Y, SHAO Z, DENG M, et al. MOOC dropout prediction using a fusion deep model based on behaviour features [J]. Computers and Electrical Engineering, 2022, 104: 108409.
- [24] CHEN X, LIU H Q, ZHAO F. Prediction of MOOCs dropout based on IQPSO-PLSTM model [C]//Proceedings of the 2023 8th International Conference on Mathematics and Artificial Intelligence, 2023: 7-12.
- [25] NITHYA S, UMARANI S. MOOC dropout prediction using FIAR-ANN model based on learner behavioral features [J]. International Journal of Advanced Computer Science and Applications, 2022, 13(9): 607-617.
- [26] HANCOX Z, RELTON S D. Temporal graph-based CNNs (TG-CNNs) for online course dropout prediction [C]//International Symposium on Methodologies for Intelligent Systems, 2022: 357-367.
- [27] NIU K, LU G, PENG X, et al. CNN autoencoders and LSTM-based reduced order model for student dropout prediction [J]. Neural Computing and Applications, 2023, 35(30): 22341-22357.
- [28] TANG Z. Research on key technologies of MOOCs based on modeling of interactive behaviors [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2023.
唐瞻. 基于交互行为建模的MOOCs关键技术研究 [D]. 南京: 南京邮电大学, 2023.
- [29] WON H S, KIM M J, KIM D, et al. University student dropout prediction using pretrained language models [J]. Applied Sciences, 2023, 13(12): 7073.
- [30] DE VILLE B. Decision trees [J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2013, 5(6): 448-455.
- [31] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45: 5-32.
- [32] ZHANG H. The optimality of naive Bayes [C]//Florida: The Seventeenth International Florida Artificial Intelligence Research Society Conference, 2004.
- [33] SALEKSHAHREZAEI Z, LEEVY J L, KHOSHGOFTAAAR T M. Feature extraction for class imbalance using a convolutional autoencoder and data sampling [C]//2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), 2021: 217-223.
- [34] SHOU Z, CHEN P, WEN H, et al. MOOC dropout prediction based on multidimensional time-series data [J]. Mathematical Problems in Engineering, 2022(1): 2213292.

(责任编辑: 孙娟)