

基于 Stacking 集成模型的混合式学习成绩预测研究

杨钰辉¹ 朱 静^{1,2,3} 张健齐¹

YANG Yuhui ZHU Jing ZHANG Jianqi

摘要

针对传统学生成绩预测模型存在结构单一、过度依赖特征等局限性,导致泛化能力和预测准确性受限,文章提出了一种基于贝叶斯优化的 Stacking 集成学生成绩预测方法。该方法结合 XGBoost、AdaBoost 和逻辑回归,通过 Stacking 算法实现模型的优势互补,提升预测能力,减少对单一特征的依赖。通过分析雨课堂的线上、线下数据,筛选出与成绩相关的关键特征,并采用 SMOTE 算法解决数据不均衡问题。同时,把贝叶斯用于超参数调优,用以平衡训练时间与预测效果。实验结果表明,Stacking 模型的准确率为 77.3%,显著优于 XGBoost (73.7%)、AdaBoost (55.8%) 和逻辑回归 (54.3%)。该方法为混合式教学中的学情把握、教学干预和成绩预测提供了有效支持。

关键词

学生成绩预测;混合式教学;Stacking;贝叶斯优化;SMOTE 算法

doi: 10.3969/j.issn.1672-9528.2025.07.002

0 引言

随着教育数据的不断积累和人工智能技术的快速发展,人工智能在教育领域的应用日益广泛,涵盖智能教学辅助、智能评估、成绩预测及个性化学习等方面。其显著提升了学习效果,并推动了教育改革^[1]。特别是混合式教学模式,已逐渐成为高校主要的教学方式,打破了传统时间与空间的限制,同时也带来了新的挑战——如何从复杂的教育数据中提取有效信息^[2-3]。教育数据为学生成绩预测提供了丰富来源,通过分析学生的学习行为数据,教育管理者可以优化课程设计、提高教学效率,并帮助学生规避学业风险^[4-7]。

然而,随着学生人数和学习行为数据的显著增长,教师难以全面及时地掌握每位学生的学习情况,进而影响整体教学质量^[8-9]。因此,如何高效利用教育数据来辅助教学决策,成为亟待解决的关键问题。混合式教学在全面推广过程中,逐渐展现出其独特的优势,结合线上与线下数据进行成绩预测,不仅为教师提供数据支持,帮助优化教学方案,还展示了巨大的潜力^[10]。

尽管已有研究通过机器学习^[11-12]和深度学习方法在成绩预测中取得了一定成功,但现有方法仍面临泛化能力不足和过度依赖特征选择等问题。例如,任伊敏^[13]利用 LightGBM

和遗传算法进行成绩预测;Korch 等人^[14]通过决策树和随机森林分析学生多维信息;王冠帮等人^[15]基于 K-Means 算法取得较高预测精度,但这些方法的应用效果受限于特定特征或模型设计^[16]。

为进一步提高成绩预测精度,本文提出了一种基于贝叶斯优化的 Stacking 集成模型。该模型结合 XGBoost、AdaBoost 和逻辑回归,通过 Stacking 算法实现模型优势互补,增强模型的泛化能力和预测准确性。贝叶斯优化调优超参数,平衡训练时间与预测效果。实验结果表明,该模型在混合式学习数据上表现优异,优于单一模型的预测性能。本文的研究不仅为混合式教学中的成绩预测提供了新方法,也为教育决策者提供了有力的数据支持,推动个性化教育的发展。

1 基础知识

1.1 集成学习和 Stacking 方法

集成学习通过组合多个弱学习器,生成更强的预测模型,主要方法有 Bagging、Boosting 和 Stacking。与 Bagging 和 Boosting 的线性组合不同,Stacking 采用多种异构学习器,使用不同特征和设计原则,通过元学习器组合它们的输出,从而实现更精准的预测^[17]。

Stacking 模型源于 Wolpert 提出的“Stacked Generalizations”^[18],通过双层结构提升预测精度。第一层包含多个基分类器(如 Model1、Model2、Model3),使用不同算法训练,并通过交叉验证生成预测结果矩阵。第二层使用元分类器对第一层输出进行再学习,生成最终预测结果。该结构能够有效修正基分类器的误差,提升整体性能。具体流程如图 1 所示。

1. 新疆农业大学计算机与信息工程学院 新疆乌鲁木齐 830052

2. 智能农业教育部工程研究中心 新疆乌鲁木齐 830052

3. 新疆农业信息化工程技术研究中心 新疆乌鲁木齐 830052

[基金项目] 科技创新 2030:“新一代人工智能”重大项目 (2022ZD0115805)

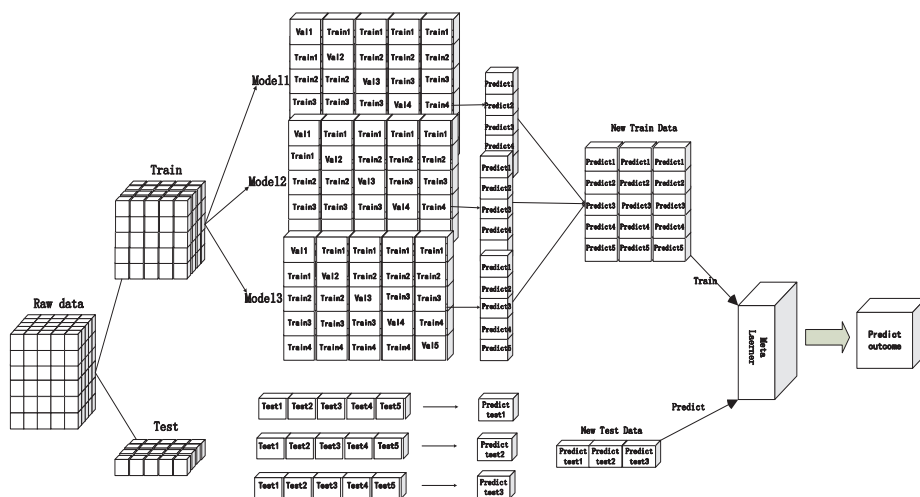


图1 Staking集成算法流程

1.2 基础模型组件

在复杂的预测任务中，单一模型往往存在局限性，而集成学习通过融合多模型的优势，能够显著提升整体预测性能。Stacking集成学习通过多模型协作，优化了泛化能力和预测精度。为提高学生成绩预测的准确性，本研究选择Stacking模型，并对其组件分别进行介绍。

1.2.1 XGBoost

XGBoost (extreme gradient boosting) 是一种基于梯度提升树的改进算法，通过二阶导数优化和正则化技术，既避免了过拟合，又提高了计算效率。该算法通过多棵分类回归树 (CART) 进行迭代训练，每轮拟合前一轮的预测残差，并调整样本权重，使后续树更加关注难以预测的样本。最终，模型通过加权组合各决策树的叶节点输出得到预测结果。

本研究选用XGBoost作为Stacking模型的基学习器，利用其强大的残差拟合和抗过拟合能力，提升集成模型的性能。核心公式为：

$$BOB_j = \sum_{p=1}^n L(y_p, \hat{y}_p) + \sum_{t=1}^M \Omega(f_t) \quad (1)$$

式中： $L(y_p, \hat{y}_p)$ 代表损失函数，用来评估预测值与真实值之间的差异； $\Omega(f_t)$ 代表正则化项，用于控制模型的复杂度，从而减少过拟合的风险。

1.2.2 AdaBoost

AdaBoost (adaptive boosting) 是一种传统的Boosting算法，通过迭代训练弱分类器来提高整体模型性能。每轮迭代中，AdaBoost调整样本权重，将错分样本赋予更高权重，使后续弱分类器更关注这些难以分类的样本。最终通过加权整合所有弱分类器的预测结果，形成一个强分类器。AdaBoost在处理高维数据时表现良好，能增强模型的鲁棒性，但对噪声较为敏感，过多噪声可能影响性能。

本研究中，选择AdaBoost作为基学习器，强化对误差

样本的关注，从而提高模型的鲁棒性和误差修正能力，其损失函数可以表示为：

$$e = \sum_{i=1}^n \omega_i \cdot 1(h(x_i) \neq y_i) \quad (2)$$

式中： e 表示弱分类器的误差率； ω_i 表示每个样本的权重； $h(x_i)$ 表示分类器的预测结果； y_i 表示真实标签。通过不断减少误差率，AdaBoost不断优化模型的预测性能。

1.2.3 逻辑回归

在Stacking模型中，逻辑回归作为元学习器，用于整合多个基学习器的输出并生成最终预测结果。逻辑回归因其简单、高效、易解释的特点，能够通过学习各基学习器的权重来有效整合它们的互补信息，从而提升模型的稳定性和预测精度。

$$P(y = 1|x) = \frac{1}{1 + e^{-(w \cdot x + b)}} \quad (3)$$

式中： x 为基学习器的预测结果； w 为学习到的权重参数； b 为偏置项。通过该公式计算各基学习器组合后结果的概率值，从而生成最终预测。

2 基于Stacking集成学习的学生成绩预测模型

2.1 数据采集

本研究采用某高校1073名学生的多源数据集，数据来源包括线上雨课堂、线下成绩和教务系统。雨课堂数据主要包含学生的课堂参与度（如得分、签到、上课率等），线下数据则包含平时成绩和期末成绩等。尽管这些数据在教育中较为普遍，教师和教育管理人员接触较多，但其应用仍然有限，且教师往往难以实时、精准地获取学生的学习表现，影响个性化教学的实施。

为解决这一问题，本研究将学生成绩划分为四个等级：不及格（60分以下）、中等（60~74分）、良好（75~84分）和优秀（85~100分）。这种分级不仅有助于教师实时把握学生学习状况，也为后续教学干预提供数据支持。研究中的学生特征包括基本信息（如年龄、居住地）和行为特征（如成绩、出勤率等），具体特征如表1所示。

2.2 数据处理

数据处理是提升模型性能的关键步骤。首先，针对缺失值问题，本研究根据特征重要性进行处理：不影响成绩的特征（如学号、姓名等）直接删除；缺失较多且价值较低的特征也予以删除。对缺失较少的特征，采用均值填充、中位数填充、K近邻填充和随机森林填充等方法。经过比较，随机森林填充法表现最佳，因其在准确性和稳健性方面优于其他方法。

表 1 数据集变量及其定义

属性类别	特征名	特征含义	数据类型
学生基本信息	id	学生的学号	int
	sex	学生的性别	int
	age	学生的年龄	int
	clans	学生的民族背景	int
	origin	学生的出生或主要居住地	int
	category	考试语言能力（如汉语、双语）	int
	residence	考生背景（如农村、城镇）	int
	affiliation	学生的政治身份	int
	graduate	学生是应届生还是往届生	int
学生行为特征	pages	在学习平台上观看的总页数	int
	checkins	在课程中的签到次数	int
	attendance	学生按时到课的比例	object
	comments	在课程中发送弹幕的总次数	int
	submissions	学生向课程平台提交作业或其他内容的次数	int
	announcements	学生阅读课程公告的频率	int
	score	学生在在线课程中的总得分	float
	performance	学生在平时课堂中的表现分数	float
	exam	学生在期末考试中的得分	float
	final	学生在学期结束时的最终成绩	float

其次，针对类别不均衡问题，本研究使用 SMOTE 算法对数据进行处理。SMOTE 通过在特征空间中生成新的少数类样本，避免模型对少数类的偏倚，提升了模型对少数类的识别能力。该处理不仅增强了模型的泛化能力，还为教师提供了更精确的成绩预测，从而支持个性化教学和精准的教学干预。

2.3 特征选择

为提高模型的预测精度，本研究使用皮尔逊相关系数进行特征筛选。分析结果显示，期末成绩（exam）、平时成绩（performance）、总成绩（score）、平台观看总页数（pages）、公告阅读（announcements）和到课率（attendance）与学生

成绩高度相关，期末成绩的相关系数达到 0.883 3。相比之下，性别（sex）、居住地（residence）和年龄（age）与成绩的相关性较低。这一发现不仅通过数据分析得出，也与实践经验相符。例如，课件推送和公告阅读等在线学习活动与学生成绩密切相关，反映了课堂互动和实时反馈在学生表现中的重要作用。因此，本研究保留了与成绩高度相关的特征，以提升模型的预测能力，并为个性化教学提供支持。

2.4 模型设计

Stacking 模型的性能主要取决于基学习器和元学习器的选择。通过整合多个基学习器，Stacking 能够更准确地捕捉数据中的多元特性，从而优化整体表现。在本研究中，选用了 XGBoost 和 AdaBoost 作为基学习器，逻辑回归则作为元学习器。为确保模型达到最佳表现，研究中还使用了贝叶斯优化进行超参数调优，以平衡训练时间和预测效果。

在第一层，本研究选择了 XGBoost 和 AdaBoost 作为基学习器。XGBoost 是一种基于梯度提升的算法，特别擅长处理高维数据并能够应对过拟合问题，通过优化损失函数捕捉数据中的复杂非线性关系。

AdaBoost 则通过动态调整弱学习器的权重，专注于分类的难点问题，逐步提升模型的整体精度。这两种模型的结合使得第一层的预测更加多样化。

在第二层，选用逻辑回归作为元学习器。逻辑回归将第一层的输出整合起来，通过加权处理进一步优化整体模型的表现。作为一种线性分类器，逻辑回归简单且稳定，能够有效综合并平衡第一层的预测信息，提升模型的准确性和泛化能力。

这种两层结构确保模型既能捕捉复杂的非线性特征，又能通过元学习器进行整合修正，最终增强预测结果的稳健性和准确性。整个 Stacking 集成学习框架如图 2 所示。

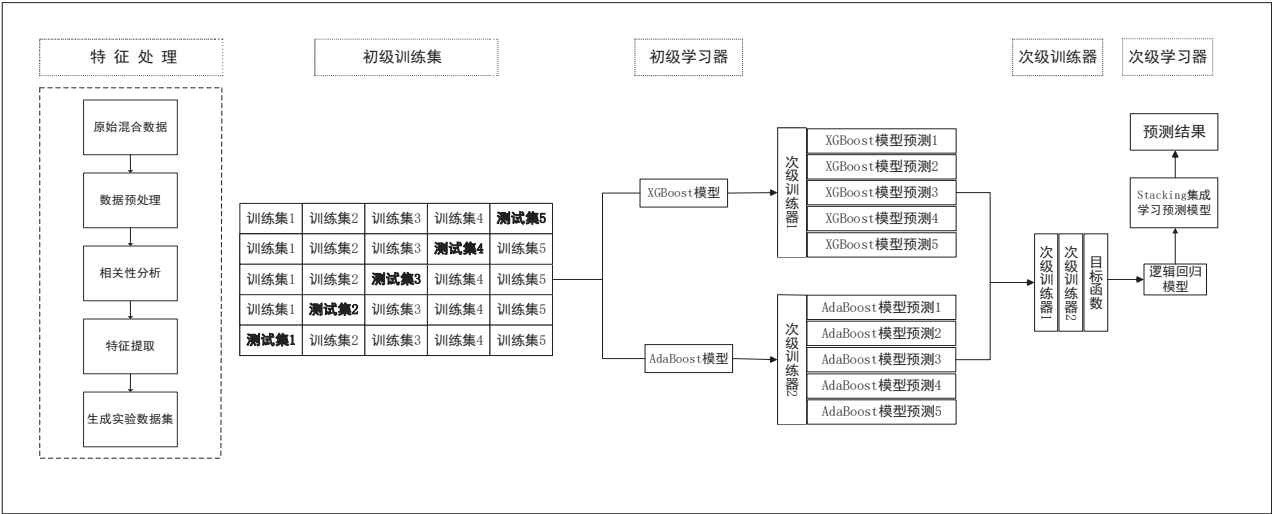


图 2 Staking 集成学习框架图

3 实验分析

3.1 数据集划分

本实验将数据集按 8:2 的比例随机划分为训练集和测试集，旨在比较不同模型在学生成绩预测中的表现，特别验证 Stacking 集成学习模型的有效性。为增强模型稳定性并降低过拟合风险，本研究采用五折交叉验证，通过反复使用不同子集作为训练集和验证集，确保模型充分学习数据特征并准确评估性能。这种交叉验证方法不仅提升了模型稳定性，还为教育工作者提供了更可靠的成绩预测支持，帮助他们及时调整教学策略。

3.2 评价指标

模型评估是实验的重要组成部分。使用测试集对训练好的 Stacking 集成模型进行评估，并计算准确率、召回率和 F_1 分数等多种指标。这些评价指标可以从不同维度衡量模型的预测表现。为确保模型在不同数据划分条件下的稳定性，采用了 5 折交叉验证对这些指标进行了进一步验证，其计算公式分别为：

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$F_1 = 2 \cdot \frac{\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (7)$$

3.3 贝叶斯优化与超参数设置

在建立模型的过程中，挑选适宜的参数对提升模型的表现至关重要。为找到最佳的模型参数，本研究采用了贝叶斯对基学习器 XGBoost 和 AdaBoost 进行超参数调优。通过 5 折交叉验证，动态调整各模型的超参数组合，确保其在训练过程中处于最佳状态。最终的参数设置如表 2 所示。

表 2 超参数设置表

算法	参数
XGBoost	colsample_bytree=0.84; learning_rate=0.05; max_depth=7; n_estimators=200; subsample=0.8;
AdaBoost	learning_rate=0.1; n_estimators=100;
Logistic Regression	C=0.01; max_iter=300;

3.4 结果分析

实验结果表明，Stacking 集成学习模型在学生成绩预测任务中表现优异，显著提升了分类器的性能。通过与 XGBoost、AdaBoost、Logistic Regression 和 Gradient Boosting 等

模型的对比，实验结果显示，Stacking 模型在准确率（77.3%）、精确率（77.6%）、召回率（77.3%）和 F_1 分数（77.3%）上均表现最佳。而 XGBoost 为 73.7%，Gradient Boosting 和 Bagging 为 71.9%，AdaBoost 和 Logistic Regression 则分别为 55.8% 和 54.3%，如表 3 所示。这表明，Stacking 能够更全面地捕捉学生成绩与学习行为间的复杂关系，提供更精准的预测。

该预测结果不仅为教学提供了重要参考，还帮助教师更早地进行干预，实现因材施教。通过精准预测学生学习行为，教师可及时识别表现不佳或潜在问题的学生，调整教学策略，如加强互动、提供辅导或调整课程内容，从而有效提升学生学习效果和教学质量。

表 3 不同预测模型的预测误差对比

单位：%

模型	Accuracy	Precision	Recall	F_1 -score
XGBoost	73.7	73.7	73.7	73.5
AdaBoost	55.8	56.5	55.8	55.3
LogisticRegression	54.3	54.1	54.3	53.6
Bagging	71.9	72.0	71.9	71.7
GradientBoosting	71.9	72.5	71.9	71.8
Stacking	77.3	77.6	77.3	77.3

为了更直观地展示 Stacking 模型效果，绘制了混淆矩阵。该模型在不及格（D）和良好（L）类别上的表现突出，准确率分别为 89.29% 和 76.19%。然而，在中等（M）和优秀（Y）类别的表现略逊色，准确率分别为 70.24% 和 72.62%。这表明，虽然 Stacking 在不及格和良好类别的预测精度较高，但在中等和优秀成绩方面仍有提升空间，如图 3 所示。因此，教师可通过调整教学策略和反馈机制，进一步优化学生的学习成果。

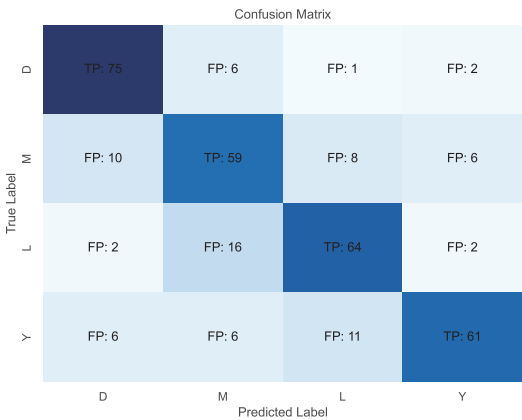


图 3 四分类混淆矩阵

总体而言，Stacking 模型在处理混合式学习行为数据时展现了较强优势，特别是在应对不平衡数据和复杂特征等方

面。通过整合多个基学习器的预测结果, Stacking 有效减少偏差, 提升预测效果, 增强了模型的稳定性与泛化能力。因此, 该模型不仅具有理论价值, 也为实际教学中的个性化干预提供了有力支持。

4 结语

本研究提出了一种基于 Stacking 集成算法的学生成绩预测模型, 结合雨课堂的线上学习数据和线下成绩数据, 旨在提升学生成绩预测的准确性和泛化能力。首先, 通过皮尔逊相关性分析筛选与成绩高度相关的特征, 为后续建模提供数据基础。接着, 采用 SMOTE 算法解决类别不平衡问题, 进一步优化模型表现。最终, 结合 AdaBoost 和 XGBoost 作为基学习器、Logistic Regression 作为元学习器, 通过多层次的学习策略, Stacking 模型有效整合了多个学习器的优势, 显著提升了预测性能。

实验结果表明, Stacking 模型在精度和召回率等关键指标上显著优于单一模型, 尤其在混合式学习环境下表现突出。通过精准的预测有助于教师及时识别潜在问题学生, 并根据个性化需求调整教学策略, 从而提高学生成绩和整体教学效果, 为个性化教学提供理论支持。

未来, 结合 LSTM 或 Transformer 等时序模型与 Stacking 算法, 将进一步提升模型精度和适应性。此外, 随着教育信息化发展, 保障数据隐私、安全及提升模型可解释性将成为研究重点。这些创新不仅会推动教学模式改革, 也为构建公平高效的教育体系提供了理论依据, 助力学生在复杂社会环境中取得更大成功。

参考文献:

- [1] 第三届清华高等教育论坛开幕 [J]. 现代教育技术, 2024, 34(9): 1.
- [2] 何克抗. 关于我国教育技术学研究现状和教育变革着力点的思考 [J]. 电化教育研究, 2018, 39(8): 5-14.
- [3] 林健. 工程教育的信息化 [J]. 高等工程教育研究, 2022, (1): 1-10.
- [4] 王巍, 闫寒冰. 能力精准测评何以可为? 以研训者数据领导力测评实践为例 [J]. 现代教育技术, 2024, 34(8): 112-122.
- [5] 袁天云. 教育大数据采集机制的研究 [D]. 武汉: 华中师范大学, 2019.
- [6] 杨帆. 可视化技术支持下在线自主学习模型构建及实证研究 [D]. 长春: 东北师范大学, 2021.
- [7] 李绿山. 基于学习分析的在线学习可视化监控和反馈研究 [D]. 长春: 东北师范大学, 2021.
- [8] 顾明远. “人工智能+”时代的教育变革创新: 顾明远先生对话讯飞教育技术研究院 [J]. 现代教育技术, 2024, 34(8): 5-12.
- [9] 田雅慧. 基于学习者画像的 MOOC 学情预警研究 [D]. 上海: 华东师范大学, 2020.
- [10] 王磊. 基于大数据的精准教学干预模型及应用研究 [D]. 武汉: 华中师范大学, 2022.
- [11] NAHAR K, SHOVA B I, RIA T, et al. Mining educational data to predict students' performance: a comparative study of data mining techniques[J]. Education and information technologies, 2021, 26: 6051-6067.
- [12] WU B, ZHENG C L, WU W Q. An analysis of the effectiveness of machine learning theory in the evaluation of education and teaching[J/OL]. Wireless communications&mobile computing, 2021[2025-01-19]. <https://doi.org/10.1155/2021/4456222>.
- [13] 任伊敏. 基于机器学习的学生成绩预测与提升策略研究 [D]. 昆明: 云南师范大学, 2023.
- [14] KORCHI A, MESSAOUDI F, ABATAL A, et al. Machine learning and deep learning-based students' grade prediction[J/OL]. Operations research forum, 2023[2024-11-22]. <https://link.springer.com/article/10.1007/s43069-023-00267-8>.
- [15] 王冠帮, 刘鸿雁, 李劲松, 等. 基于 K-Means 的学生成绩预测方法研究 [J]. 信息技术, 2023, 47 (2): 1-6.
- [16] ORJI F A, VASSILEVA J, GREER J. Evaluating a persuasive intervention for engagement in a large university class[J]. International journal of artificial intelligence in education, 2021, 31: 700-725.
- [17] NTI I K, ADEKOYA A F, WEYORI B A. A comprehensive evaluation of ensemble learning for stock-market prediction[J/OL]. Journal of big data, 2020[2024-05-22]. <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-020-00299-5>.
- [18] NIYOGISUBIZO J, LIAO L, NZIYUMVA E, et al. Predicting student's dropout in university classes using two-layer ensemble machine learning approach: a novel stacked generalization[J]. Computers and education: artificial intelligence, 2022, 3: 10.

【作者简介】

杨钰辉 (1997—), 女, 湖南永州人, 硕士, 研究方向: 计算机技术。

(收稿日期: 2025-02-20 修回日期: 2025-06-30)