

基于机器学习的钢-自燃煤矸石混凝土组合梁 栓钉抗剪承载力研究

王庆贺¹,张提睿¹,李永进²,任庆新¹,包龙生³

(1. 沈阳建筑大学土木工程学院,辽宁 沈阳 110168;2. 福建江夏学院工程学院,福建 福州 350108;
3. 沈阳建筑大学交通与测绘工程学院,辽宁 沈阳 110168)

摘要 目的 确定影响栓钉抗剪承载力的主要因素,量化自燃煤矸石骨料取代率对栓钉抗剪承载力的影响。方法 基于机器学习中的随机森林算法,选取影响栓钉抗剪承载力的7个因素作为输入变量、将栓钉抗剪承载力作为输出变量,构建栓钉抗剪承载力预测模型,并利用现有试验结果验证模型的可靠性,进行输入变量的特征重要性分析;收集现有考虑不同骨料取代率的自燃煤矸石混凝土抗压强度与弹性模量,量化自燃煤矸石骨料取代率对栓钉抗剪承载力的影响。结果 基于机器学习建立的模型判定系数 R^2 为0.94、平均绝对误差MAE为10.56、均方误差MSE为180.44、均方根误差RMSE为13.43;栓钉抗剪承载力的影响因素主要为栓钉直径和混凝土抗压强度,而且栓钉直径与混凝土抗压强度对栓钉抗剪承载力的影响存在耦合关系;与普通混凝土试件相比,取代率为25%、50%、75%和100%时栓钉的抗剪承载力平均降低2.49%、4.46%、6.35%和7.31%。结论 基于随机森林算法所建立的模型具有较高的预测精度;栓钉直径、混凝土抗压强度是影响栓钉抗剪承载力的主要因素;栓钉抗剪承载力随自燃煤矸石骨料取代率的增加而降低。

关键词 组合梁;自燃煤矸石混凝土;抗剪栓钉;抗剪承载力;随机森林;机器学习

中图分类号 TU398.9

文献标志码 A

Shear Capacity of Studs in Steel-Spontaneous-Combustion Coal Gangue Concrete Composite Beams Using Machine Learning

WANG Qinghe¹, ZHANG Tirui¹, LI Yongjin², REN Qingxin¹, BAO Longsheng³

(1. School of Civil Engineering, Shenyang Jianzhu University, Shenyang, China, 110168; 2. College of Engineering, Fujian Jiangxia University, Fuzhou, China, 350108; 3. School of Transportation and Geomatics Engineering, Shenyang Jianzhu University, Shenyang, China, 110168)

收稿日期:2022-05-19

基金项目:国家自然科学基金项目(51808351,51808352);沈阳市科技计划项目(21-108-9-13)

作者简介:王庆贺(1987—),男,副教授,博士,主要从事钢-混凝土组合结构等方面研究。

Abstract: Using random forest algorithm in machine learning, the prediction model of shear capacity of studs was proposed with seven factors affecting the shear capacity of studs as input variables and the shear capacity of studs as output variables; the reliability of the prediction model was further verified by the existing results. The feature importance of input variables was analyzed to determine the main factors affecting the shear capacity of studs. The compressive strength f_{cm} and elastic modulus E_{cm} of concrete with different replacement ratios of spontaneous-combustion coal gangue aggregate (SCGA) were collected to predict the influence of the SCGA replacement ratio on the shear capacity of studs. Results showed that the established random forest prediction model had high prediction accuracy, with R^2 of 0.94, MAE of 10.56, MSE of 180.44, $RMSE$ of 13.43; the stud diameter and the concrete compressive strength were the main factors affecting the shear capacity of the studs; there was a coupling relationship between the diameter of the stud and the compressive strength of concrete on the shear capacity of the stud; the shear capacity of studs decreases with increasing SCGA replacement ratios, compared with the replacement ratio of 0%, the shear capacity of studs decreased by 2.49%, 4.46%, 6.35% and 7.31%, respectively, for SCGA replacement ratio of 25%, 50%, 75% and 100%.

Key words: composite beam; spontaneous-combustion coal gangue concrete; shear stud; shear capacity; random forest; machine learning

煤矸石是煤炭开采和加工过程中产生的固体废弃物,在自然条件下经内部化学反应放热可使原状煤矸石转化为自燃煤矸石^[1-2],将破碎的自燃煤矸石骨料引入钢-混凝土组合梁,既可兼容煤矸石混凝土较低的力学性能,又可具备组合梁承载力高等特点。作为钢-混凝土组合梁中应用广泛的柔性抗剪连接件,栓钉具有抗剪性能无方向性、施工质量易于保证、延性高等特点^[3-4]。研究钢-自燃煤矸石混凝土组合梁的栓钉抗剪性能,可为组合梁静力性能研究奠定理论基础。

目前,国内外学者多采用推出试验研究栓钉的抗剪性能。聂建国等^[5]通过钢-高强混凝土组合梁推出试验,对栓钉抗剪连接件在混凝土中的性能进行研究,并提出了计算建议;汪劲丰等^[6]通过24个栓钉抗剪性能推出试件,研究了栓钉高度对栓钉抗剪承载力的影响,结果表明,栓钉的抗剪承载力随栓钉长径比的增加而增大,同时基于推出试验结果拟合得出考虑栓钉高度的抗剪承载力计算公式;Y. H. Ling等^[7]采用有限元分析方法模拟栓钉的失效模式并进行参数分析,结果表明,栓钉直径和混凝土强度对栓钉抗

剪承载力影响较大,并根据参数分析结果修正了栓钉承载力设计公式;Y. Q. Hu等^[8]通过推出试验研究了栓钉高度与混凝土板厚比值及配筋布置对高强混凝土中栓钉抗剪性能的影响,发现栓钉的抗剪刚度主要受栓钉直径、混凝土抗压强度和弹性模量影响,提出并验证了抗剪刚度预测公式。

机器学习是人工智能领域的一个分支,近几年随着人工智能技术的发展,学者不断将其应用于土木工程领域来解决实际问题。V. Q. Tran等^[9]采用6种机器学习模型对再生混凝土的抗压强度进行预测,研究结果表明,水泥含量是影响再生混凝土抗压强度的最主要因素;P. Ladygina等^[10]采用线性回归、K-近邻、随机森林等机器学习模型对混凝土的抗压强度进行预测,分析结果表明,随机森林模型的预测精度最高,判定系数 R^2 为0.89;C. Avci-Karatas^[11]基于极大极小概率机回归和极限学习机概念建立了两种用于预测钢-混凝土组合结构中栓钉抗剪性能的回归模型,并通过评价指标验证了模型的有效性,两个模型 R^2 分别为0.99和0.95,所建立模型可用于栓钉抗剪性能的预测。

由于混凝土的本身特性使得相关试验材料消耗量大、周期长,因此提高试验效率十分重要^[12]。基于此,笔者采用机器学习中的随机森林算法构建栓钉抗剪承载力预测模型,并利用现有试验结果验证模型可靠性;收集考虑不同骨料取代率的自燃煤矸石混凝土抗压强度与弹性模量,量化自燃煤矸石骨料取代率对栓钉抗剪承载力的影响。研究表明,栓钉抗剪承载力的影响因素主要为栓钉直径和混凝土抗压强度;栓钉抗剪承载力随自燃煤矸石骨料取代率的增加而降低。

1 预测模型建立

1.1 随机森林算法原理

Bagging 是并行式集成学习算法的典型代表,它基于自主采样法,对于给定的包含 t 个样本的数据集进行有放回抽样,经过 t 次随机采样操作,可得到一个含 t 个样本的采样集,重复采样过程 N 次,可得到 N 个含 t 个样本的采样集,然后基于每个采样集训练一个基学习器,再将这些基学习器进行结合,就是 Bagging 算法的基本流程^[13]。随机森林是 Bagging 的一个扩展变体,它在以决策树为基学习器构建 Bagging 集成的基础上^[14],进一步在决策树的训练过程中引入了随机特征选择^[13]。随机森林不仅拥有 Bagging 算法的样本扰动优点,而且引入了特征扰动,这

就使得随机森林最终的泛化能力通过基学习器之间差异度的增加而进一步提升,其模型具有可解释性强、适应能力强、不易过拟合等特点。随机森林算法流程见图 1。

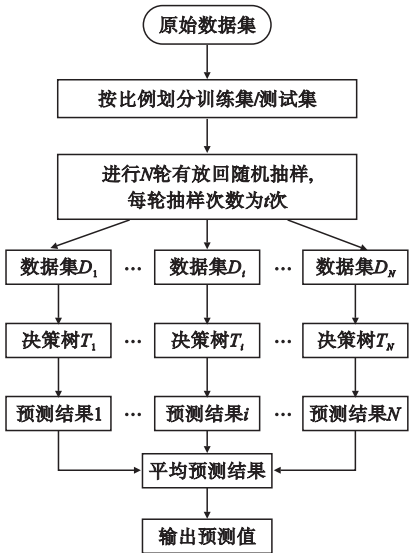


图 1 随机森林算法流程图

Fig. 1 Flow chart of random forest algorithm

1.2 样本特征及数据集划分

采用文献[15]中使用的数据集,该数据集包含 242 组数据,每组数据均包含 8 个变量,在进行机器学习模型建立和训练过程中,将 7 个特征变量作为输入变量:混凝土抗压强度 f_{cm} 、混凝土弹性模量 E_{cm} 、栓钉抗拉强度 f_{um} 、栓钉直径 d_m 、栓钉高度 h_m 、焊环直径 d_{dom} 、焊环高度 h_{wm} ,将栓钉抗剪承载力 P_{cm} 作为输出变量,各变量的统计信息见表 1。

表 1 输入和输出变量的统计信息

Table 1 Statistical information for input and output variables

统计量	f_{cm}/MPa	E_{cm}/GPa	f_{um}/MPa	d_m/mm	h_m/mm	d_{dom}/mm	h_{wm}/mm	P_{cm}/kN
最小值	16.61	15.10	426.00	12.70	69.92	17.00	3.00	61.83
最大值	115.83	46.50	675.00	31.75	200.00	44.45	8.55	318.90
平均值	59.90	34.93	518.78	21.01	107.62	26.97	5.78	1.03
标准差	31.00	6.27	49.48	2.80	29.27	4.12	1.03	55.88
Q_1	32.96	30.76	482.46	19.00	95.00	23.00	5.00	127.46
Q_2	45.35	34.60	519.00	22.00	100.00	29.00	6.00	172.25
Q_3	86.07	40.15	549.20	22.00	125.00	29.71	6.00	215.47

注: Q_1 、 Q_2 、 Q_3 分别表示下四分位数、中四分位数、上四分位数。

将原始数据集按 8:2 的比例划分为训练集和测试集,其中训练集用来训练和优化机器学习模型超参数,测试集用来评估机器学习模型最终预测精度。由于随机森林模型是基于树划分的,与距离计算无关,树结点的选择只关注当前特征属性划分效果,即只与特征内部相对大小有关,而与不同特征间相对大小无关,因此不需要对数据集进行标准化处理。

1.3 模型性能度量

为评估和衡量机器学习模型在数据集上的泛化性能,选取机器学习中最常用的四种性能度量:判定系数 R^2 、平均绝对误差 MAE 、均方误差 MSE 和均方根误差 $RMSE$ 。其中 R^2 用来衡量预测值与真实值之间的接近程度, R^2 越接近于 1,表明模型的预测精度越高; MAE 、 MSE 、 $RMSE$ 用来衡量模型误差,三者互为补充,其值越小,表明模型的预测精度越高。上述四种性能度量的数学表达式分别见式(1)~式(4):

$$R^2 = 1 - \frac{\sum_{i=1}^m (\hat{y}_i - y_i)^2}{\sum_{i=1}^m (\bar{y} - y_i)^2} \tag{1}$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \tag{2}$$

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \tag{3}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \tag{4}$$

式中: y_i 为真实值; \hat{y}_i 为预测值; \bar{y} 为真实值的均值; m 为数据点个数。

2 预测模型验证

为避免模型超参数设置不合理及各变量之间相互影响而引起的模型次优化现象^[16],采用网格搜索法对机器学习模型的超参数进行优化处理^[17]。基于建立的随机森林模型对测试集数据进行预测,预测值与真实值对

比结果见图 2。从图中可以看出,随机森林模型的预测值与测试集的真实值较为接近。测试集内共有 49 组数据,其中误差超过 15% 的只有 5 组,预测结果误差绝大部分在 $\pm 15\%$ 以内。由此可见,随机森林模型对测试集的预测较为准确。

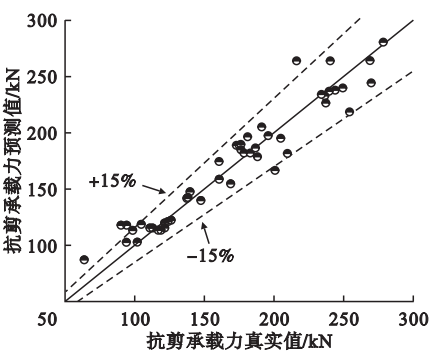


图2 随机森林模型预测结果对比

Fig. 2 Comparison of random forest model prediction results

为进一步验证笔者建立的随机森林模型预测性能,基于模型性能度量,与文献[15]中各模型预测结果进行对比分析,结果见表 2。从表中可以看出,针对判定系数 R^2 而言,所建立的随机森林模型为 0.94,说明其对测试集数据的拟合效果较好,预测能力较强;针对模型误差指标而言,与其他模型相比,随机森林模型的结果数值较低,说明其预测值偏离真实值的程度相对较小。综上表明,笔者建立的随机森林模型具有较高的预测精度。

表2 机器学习模型及其预测性能

Table 2 ML model and prediction performance					
机器学习模型	R^2	MAE	MSE	RMSE	
Linear Regression ^[15]	0.91	12.91	241.92	15.55	
Decision Tree ^[15]	0.86	13.21	397.40	19.94	
Bagged Ensemble Trees ^[15]	0.88	11.60	328.20	18.12	
Super Vector Machine ^[15]	0.95	8.70	146.00	12.08	
Gaussian Process Regression ^[15]	0.92	9.08	231.24	15.21	
Artificial Neural Network ^[15]	0.91	9.86	253.89	15.93	
Random Forest Regression	0.94	10.56	180.44	13.43	

此外,基于随机森林模型对输入变量的特征重要性进行分析,同时与文献[15]中基于决策树模型得到的特征重要性结果进行对比,结果见图3。

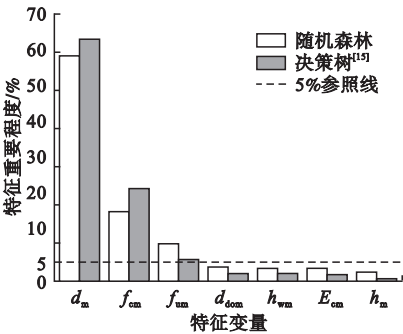


图3 特征重要性分析

Fig. 3 Feature importance analysis

特征重要性分析及对比结果表明,相对于其他输入变量,栓钉直径 d_m 与混凝土抗压强度 f_{cm} 是影响栓钉抗剪承载力的主要因素。栓钉直径对栓钉抗剪承载力的影响最大,其重要度超过50%,混凝土抗压强度对栓钉抗剪承载力的影响次之;焊环直径及高度、栓钉高度、混凝土弹性模量对栓钉抗剪承载力的影响较小,重要度均不足5%。因此,基于奥卡姆剃刀原理^[18],在后续研究中可考虑减少部分对栓钉抗剪承载力影响较小的特

征变量,以此实现对预测模型的有效降维,从而避免模型出现过拟合现象。

3 自燃煤矸石骨料取代率对栓钉抗剪承载力的影响

3.1 参数及范围

为保证预测结果具有一定可靠性,基于统计学遵循的试验随机性、重复性原则,共进行30次自燃煤矸石骨料取代率对栓钉抗剪承载力影响的不同参数预测。

采用H. Q. Liu^[19]收集的考虑不同自燃煤矸石骨料取代率的混凝土抗压强度 f_{cm} 与弹性模量 E_{cm} ,其他参数从原数据集^[15]中随机选取30组用于栓钉抗剪承载力预测,各项参数及其取值范围见表3。为保证数据有效性,将收集的自燃煤矸石混凝土(SCGAC)抗压强度及弹性模量数据与原数据集^[15]中普通混凝土(NAC)数据进行对比,结果见图4。从图中可以看出,虽然有部分数据在原数据集范围之外,但总的来说,收集的考虑不同自燃煤矸石骨料取代率的混凝土抗压强度及弹性模量数据基本在原数据集有效范围之内,在进行参数分析时仍采用全部数据。

表3 参数及取值范围

Table 3 Parameters and their scopes

f_{cm}/MPa	E_{cm}/GPa	f_{um}/MPa	d_m/mm	h_m/mm	d_{dom}/mm	h_{wm}/mm
22.4 ~ 47.4	13.5 ~ 36.2	16 ~ 31.8	436.4 ~ 675	17 ~ 44.5	3 ~ 8.6	70 ~ 200

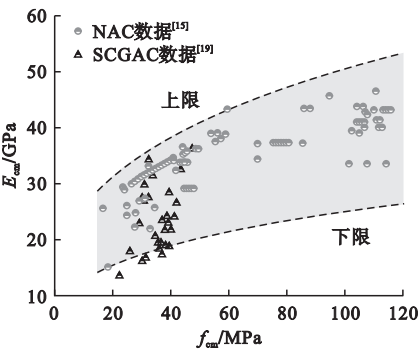


图4 自燃煤矸石混凝土数据与原数据集对比

Fig. 4 Comparison between the data of SCGA concrete and the original database

3.2 预测结果分析

基于前期特征重要性分析结果,图5给出了不同自燃煤矸石骨料取代率下混凝土抗压强度与栓钉直径对栓钉抗剪承载力的影响。其中, x 轴、 y 轴分别为混凝土抗压强度 f_{cm} 与栓钉直径 d_m ; z 轴为与普通混凝土试件相比,栓钉抗剪承载力的降低幅度。不同自燃煤矸石骨料取代率对栓钉抗剪承载力的影响结果见图6。

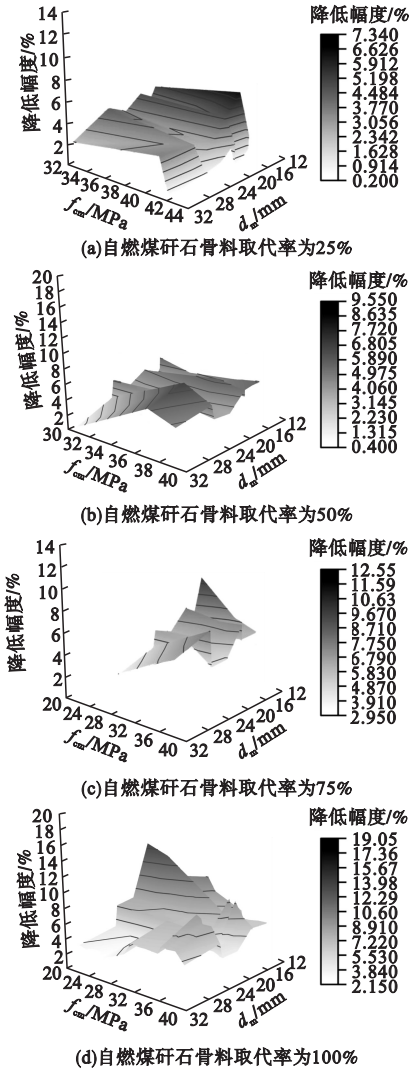


图5 混凝土抗压强度及栓钉直径对栓钉抗剪承载力的影响

Fig. 5 Influence of concrete compressive strength and stud diameter on shear capacity of studs

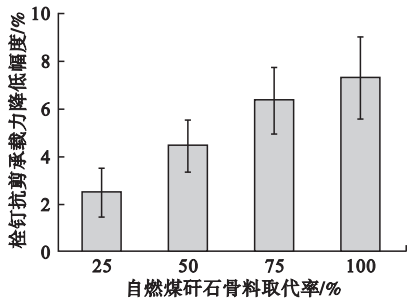


图6 自燃煤矸石骨料取代率对栓钉抗剪承载力的影响

Fig. 6 Influence of replacement ratios of SCGA on shear capacity of studs

从图5和图6中可以看出,栓钉的抗剪承载力随着自燃煤矸石骨料取代率的增加而降低。具体来说,与普通混凝土试件相比,取代率为25%、50%、75%和100%时栓钉的抗剪承载力平均降低了2.49%、4.46%、6.35%和7.31%。分析原因,随着自燃煤矸石骨料取代率的增加,自燃煤矸石混凝土的弹性模量降低、抗压强度降低,使得钢-混凝土组合梁在相同荷载作用下栓钉根部附近混凝土更早发生破坏,从而使得栓钉的抗剪承载力降低。同时,在骨料取代率相同的情况下,随着栓钉直径和混凝土抗压强度的增大,栓钉的抗剪承载力逐渐提高,二者与栓钉抗剪承载力变化呈正相关。而且,二者对栓钉抗剪承载力的影响存在耦合关系。当混凝土抗压强度较小时,栓钉直径对栓钉抗剪承载力的影响不明显;而当混凝土抗压强度较大时,随着栓钉直径的增大,栓钉抗剪承载力的增加幅度也逐渐变大。这说明,要想使栓钉抗剪承载力得到有效提高,需同时考虑适当提高自燃煤矸石混凝土抗压强度和栓钉直径。

4 结 论

(1) 基于随机森林算法建立的栓钉抗剪承载力预测模型对测试集数据拟合效果较好,判定系数 R^2 为 0.94,平均绝对误差 MAE 为 10.56,均方误差 MSE 为 180.44,均方根误差 $RMSE$ 为 13.43。

(2) 栓钉直径与混凝土抗压强度是影响栓钉抗剪承载力的主要因素,以此可以实现对预测模型的有效降维。

(3) 栓钉直径与混凝土抗压强度对栓钉抗剪承载力的影响存在耦合关系,随着混凝土抗压强度和栓钉直径的增大,栓钉的抗剪承载力逐渐提高。

(4) 栓钉抗剪承载力随自燃煤矸石骨料取代率的增加而降低,与普通混凝土试件相比,取代率为25%、50%、75%和100%时栓钉的抗剪承载力平均降低了2.49%、

4.46%、6.35%和7.31%。

参考文献

- [1] 王继宇,耿侃. 自燃煤矸石山发火原因研究进展[C]//2019 中国环境科学学会科学技术年会论文集. 西安:中国环境科学学会,2019:880-885.
(WANG Jiyu, GENG Kan. Research progress on ignition causes of spontaneous-combustion coal gangue mountain [C]//Proceedings of the 2019 CSES Annual Conference on Environmental Science and Technology. Xi'an: Chinese Society for Environmental Sciences, 2019:880-885.)
- [2] ZHOU C, LIU G, WU S, et al. The environmental characteristics of usage of coal gangue in bricking-making: a case study at Huainan, China [J]. Chemosphere, 2014, 95: 274-280.
- [3] 聂建国. 钢-混凝土组合梁结构[M]. 北京:科学出版社,2005.
(NIE Jianguo. Steel-concrete composite beam structures [M]. Beijing: Science Press, 2005.)
- [4] 王倩,刘玉擎. 焊钉连接件抗剪承载力试验研究[J]. 同济大学学报(自然科学版), 2013, 41(5): 659-663.
(WANG Qian, LIU Yuqing. Experimental study of shear capacity of stud connector [J]. Journal of tongji university (natural science), 2013, 41(5): 659-663.)
- [5] 聂建国,谭英,王洪全. 钢-高强混凝土组合梁栓钉剪力连接件的设计计算[J]. 清华大学学报(自然科学版), 1999, 39(12): 94-97.
(NIE Jianguo, TAN Ying, WANG Hongquan. Strength of stud shear connectors in composite steel-HC beams [J]. Journal of Tsinghua university (science and technology), 1999, 39(12): 94-97.)
- [6] 汪劲丰,张爱平,王文浩. 栓钉高度对栓钉连接件抗剪性能的影响[J]. 浙江大学学报(工学版), 2020, 54(11): 2076-2084.
(WANG Jinfeng, ZHANG Aiping, WANG Wenhao. Effects of stud height on shear behavior of stud connectors [J]. Journal of Zhejiang university (engineering science), 2020, 54(11): 2076-2084.)
- [7] LING Y H, ZHENG Z Y, YANG T Y, et al. Behaviour and modeling of the bearing capacity of shear stud connectors [J]. International journal of steel structures, 2019, 19(2): 650-659.
- [8] HU Y Q, QIU M H, CHEN L L, et al. Experimental and analytical study of the shear strength and stiffness of studs embedded in high strength concrete [J]. Engineering structures, 2021, 236(6): 111792.
- [9] TRAN V Q, DANG V Q, HO L S. Evaluating compressive strength of concrete made with recycled concrete aggregates using machine learning approach [J]. Construction and building materials, 2022, 323(3): 126578.
- [10] LADYGINA P, SAMOCHADIN A, VOINOV N, et al. Predicting concrete compressive strength using machine learning [C]. Singapore: International Workshop of Advanced Manufacturing and Automation, 2021.
- [11] AVCI-KARATAS C. Application of machine learning in prediction of shear capacity of headed steel studs in steel-concrete composite structures [J]. International journal of steel structures, 2022, 22(2): 539-556.
- [12] 李嗣昌,舒铮,张春雨. 钢管煤矸石混凝土压弯构件的有限元分析[J]. 沈阳建筑大学学报(自然科学版), 2010, 26(6): 1052-1057.
(LI Guochang, SHU Zheng, ZHANG Chunyu. Finite element analysis of gangue concrete-filled steel tubular beam-columns [J]. Journal of Shenyang jianzhu university (natural science), 2010, 26(6): 1052-1057.)
- [13] 周志华. 机器学习[M]. 北京:清华大学出版社, 2021.
(ZHOU Zhihua. Machine learning [M]. Beijing: Tsinghua University Press, 2021.)
- [14] BREIMAN L. Random forests [J]. Machine learning, 2001, 45(1): 5-32.
- [15] SETVATI M R, HICKS S J. Machine learning models for predicting resistance of headed studs embedded in concrete [J]. Engineering structures, 2022, 254(3): 113803.
- [16] 崔晓宁,王起才,张戎令,等. 基于随机森林的高性能混凝土抗压性能预测[J]. 兰州交通大学学报, 2021, 40(6): 1-6.
(CUI Xiaoning, WANG Qicai, ZHANG Rongling, et al. Prediction of compressive strength of concrete based on random forest [J]. Journal of Lanzhou jiaotong university, 2021, 40(6): 1-6.)
- [17] MEZZATESTA S, TORINO C, MEO P D, et al. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis [J]. Computer methods and programs in biomedicine, 2019, 177: 9-15.
- [18] BLUMER A, EHRENFEUCHT A, HAUSSLER D, et al. Occam's razor [J]. Information processing letters, 1987, 24(6): 377-380.
- [19] LIU H Q, XU Q, WANG Q H, et al. Prediction of the elastic modulus of concrete with spontaneous-combustion and rock coal gangue aggregates [J]. Structures, 2020, 28(3): 774-785.
(责任编辑:杨永生 英文审校:刘永军)