

基于可解释机器学习的螺栓抗剪连接件承载力预测模型

马晨¹, 罗云标^{1,2*}

(1. 天津大学 建筑工程学院, 天津 300072; 2. 天津大学 滨海土木工程结构与
安全教育部重点实验室, 天津 300072)

摘要: 准确预测螺栓抗剪连接件的抗剪承载力对可拆卸再利用钢-混凝土组合梁的设计和应用具有重大意义。本文通过参考公开文献资源, 构建了一个包含200组螺栓抗剪连接件抗剪承载力推出试验结果的数据库。在数据处理阶段, 研究采用可解释机器学习方法, 对数据库的异常样本进行了精准识别, 深入分析了数据间的相关性。在模型构建过程中, 选择推出试验试件的混凝土强度、弹性模量、螺栓抗拉强度、螺栓屈服强度、螺栓直径、预埋长度等核心指标, 并基于四种先进的机器学习算法, 通过超参数调节方法, 构建了螺栓抗剪连接件抗剪承载力的预测模型。经过对比得到最佳的机器学习模型, 其在测试集上的可决系数达到0.987, 显著超越了现有经验模型。为进一步提升预测模型的可信度与透明度, 研究采用可加性解释的SHAP方法, 对最佳机器学习模型的预测结果进行了分析, 揭示了模型预测的内在逻辑与依据。此外, 本研究创新性提出一种基于符号回归和遗传算法、满足量纲平衡的公式生成方法, 通过该方法拟合出了螺栓抗剪连接件抗剪承载力预测公式。该公式具有较高的预测精度, 在测试集上的可决系数达到0.899, 且能够清晰揭示螺栓抗剪连接件关键特征与抗剪承载力之间的显性映射关系, 为实际应用提供了有力支持。

关键词: 抗剪承载力; 可解释机器学习; 螺栓抗剪连接件; 数据驱动; 预测模型; 可决系数; 显性映射关系

中图分类号: TU398, TP181

文献标志码: A

文章编号: 1671-9379(2025)12-0001-09

DOI: 10.13969/j.jzgjz.20240628002

Prediction Model for the Bearing Capacity of Bolted Shear Connectors Based on Interpretable Machine Learning

MA Chen¹, LUO Yunbiao^{1,2*}

(1. School of Civil Engineering, Tianjin University, Tianjin 300072, China; 2. Key Laboratory of Coast Civil Structure Safety of Ministry of Education, Tianjin University, Tianjin 300072, China)

*Corresponding author: luoyunbiao@foxmail.com

Abstract: Accurate prediction of the shear capacity of bolted shear connectors has significant importance for the design and application of reusable steel-concrete composite beams. This study constructed a database containing 200 sets of experimental results from push-out tests on the shear capacity of bolted shear connectors, through the mining of openly accessible literature resources. During the data processing stage, the study employed interpretable machine learning methods to precisely identify abnormal samples and thoroughly analyze the correlations within the dataset. In the process of model development, key features such as concrete strength, elastic modulus, bolt tensile strength, bolt yield strength, bolt diameter, and embedded length of the push-out test specimens were selected. Based on four

收稿日期: 2024-06-28

基金项目: 国家自然科学基金(52238012, 51508381)

作者简介:

马晨(1999—), 男, 硕士研究生, 主要从事组合结构智能化方面的研究。E-mail: mc_2017@tju.edu.cn。

通信作者:

罗云标(1982—), 男, 博士, 副教授, 主要从事组合结构方面的研究。E-mail: luoyunbiao@foxmail.com。

advanced machine learning algorithms and a semi-empirical hyperparameter tuning approach, a prediction model for shear capacity was established. After comparison, the optimal machine learning model was identified, achieving a determination coefficient of 0.987 on the test set, significantly surpassing existing empirical models. To further enhance the credibility and transparency of the prediction model, the study employed the additive explanation method of SHAP to interpret the predictions of the optimal machine learning model, revealing the inherent logic and basis of the model's predictions. Additionally, this study innovatively proposed a method for generating dimensionally balanced formulas based on symbolic regression and genetic algorithms. Through this method, a prediction formula for the shear capacity of bolted shear connectors was fitted. This formula exhibits high prediction accuracy, achieving a determination coefficient of 0.899 on the test set, and clearly reveals the explicit mapping relationship between the key features of bolted shear connectors and their shear capacity, providing strong support for practical applications.

Keywords: shear bearing capacity; interpretable machine learning; bolted shear connector; data-driven; prediction model; determination coefficient; explicit mapping relationship

螺栓抗剪连接件在钢结构建筑,特别是预制装配式建筑中的应用备受关注^[1]。它可以代替传统的栓钉,实现了钢-混凝土组合梁以及钢梁与组合楼板的紧密连接,确保了两者间的协同受力。此外,它还可以广泛应用到主体结构连接、地震减震设备连接以及结构改造加固等领域。其便捷的安装与拆卸特性,使得钢结构建筑在后期维护、改造及扩建过程中能够灵活应对各种需求,极大地提高了建筑的使用效率和生命周期^[2-3]。因此,准确预测螺栓抗剪连接件的抗剪承载力对于确保建筑结构的整体性、安全性和耐久性具有至关重要的意义^[4]。

目前,国内外学者对螺栓抗剪连接件进行了广泛的试验与模拟研究。ATAEI等^[5]发现螺栓的抗拉强度、直径显著影响其抗剪承载力,而螺栓孔间隙大小对螺栓的抗剪承载力影响较小。LIU和BRADFORD等^[6-8]指出螺栓直径和强度等级对其承载力影响较大,而预紧力对其承载力影响较小,并据此提出了计算公式。PAVLOVIC等^[9]通过ABAQUS对比分析了高强螺栓与栓钉连接件的抗剪性能,并提出了螺栓连接件抗剪承载力公式。虽然现有的经验模型^[10-14]在其特定适用范围内具有较好的预测精度,但这些模型之间的相互适应性较差,且一些不满足量纲平衡的预测公式无法反映力学机理。因此,目前在螺栓抗剪连接件的抗剪承载力预测方面仍存在挑战,需要进一步的深入研究。

随着人工智能技术的飞速发展,机器学习与智能计算为结构性能预测等具有挑战性的问题开辟了新路径。这些方法利用历史数据资源,有效揭示了输入物理量(如构件参数)与输出物理量(如性能指标)间的复杂映射关系。例如,RAZAVI SETVATI等^[15]通过多种机器学习算法预测了栓钉连接件的抗剪承载力,ZHU等^[16]则通过元启发式算法优化神经网络以提升群钉抗剪承载力的预测能力,多种机器学习算法和元启发式算法均展现出较高的预测精度。然而,传统机器学习模型在揭示输入与输出间非线性关系的内在机理方面存在局限,常被视为“黑盒”模型,预测过程缺乏透明度和可再利用性。为解

决此问题,可解释机器学习模型应运而生,不仅增强了模型的可信度与应用广度,还促进了输入与输出间映射关系的深入理解。部分方法甚至能生成显式计算公式^[17],故而提高了模型的可解释性。

近年来,可解释性模型逐渐受到关注并在实践中被应用。冯德成等^[18]利用SHAP方法分析了钢筋混凝土矮墙的承载力,戴理朝等^[19]则通过遗传编程算法揭示了锈蚀钢筋混凝土构件抗剪承载力的关键参数映射关系。这些研究为可解释机器学习模型在工程领域的推广与应用奠定了坚实基础。

本文采用可解释机器学习方法,构建了螺栓抗剪连接件抗剪承载力的预测模型,并进行了详尽的解释与分析,旨在实现抗剪性能预测高精度与可解释性的双重目标。研究过程首先收集并分析了200组螺栓抗剪连接件推出试验数据,通过离散性与相关性检验确保数据质量。之后,运用四种机器学习算法构建预测模型,并对比分析螺栓抗剪连接件的抗剪性能。为深入解析模型,采用SHAP方法评估螺栓抗剪连接件构件参数的重要性,实现模型的透明化。最后,结合符号回归与遗传算法,提出一种考虑量纲平衡的公式生成方法,成功推导出螺栓抗剪连接件抗剪承载力的预测公式。建模流程如图1所示,该方法不仅提升了预测精度,还显著增强了模型的可解释性,对相关领域的研究与应用具有重要参考价值。

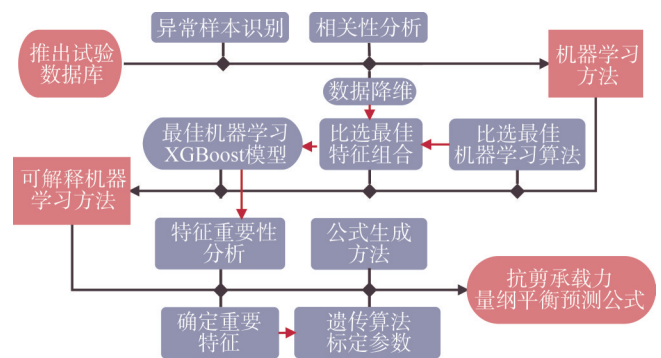


图1 建模流程示意图

Fig. 1 Schematic diagram of the modeling process

1 试验样本数据库与样本质量检验

本文共搜集200组螺栓抗剪连接件抗剪承载力推出试验数据,表1为文献中各试件的特征参数。数据库中包含的抗剪连接件类型如图2所示。根据现有钢-混凝土组合梁中的螺栓抗剪连接件的研究成果,考虑影响其抗剪承载力(P_u)的参数包括:混凝土立方体抗压强度(f_c)、混凝土弹性模量(E_c)、螺栓抗拉强度(f_u)、螺栓屈服

强度(f_y)、螺栓直径(d)、螺栓预埋长度(h)。

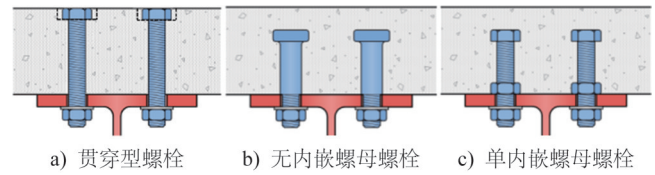


图2 数据库中抗剪连接件类型

Fig. 2 Types of shear connector in the database

表1 螺栓抗剪连接件推出试验数据

Table 1 Databases of bolted shear connector push-out test

数据来源	f_c /MPa		E_c /GPa		d /mm		h /mm		f_u /MPa		f_y /MPa		P_u /kN		样本个数	连接件类型
	min	max	min	max	min	max	min	max	min	max	min	max	min	max		
冷洪君等 ^[21]	32.5	40.7	30.0	32.5	16	20	110	110	434.1	1 038.6	264.8	924.2	55.0	192.2	12	b
邢颖等 ^[14]	45.6	79.2	32.5	36.0	16	22	120	120	989.7	1 303.1	711.2	1 274.3	133.5	269.7	30	a
汪威等 ^[22]	41.2	41.2	32.5	32.5	12	16	115	115	832.0	846.0	665.6*	676.8*	38.8	117.7	20	b
杜浩等 ^[23]	29.4	52.6	30.0	34.5	16	20	50	90	850.0	1 070.0	660.0	950.0	68.1	119.5	24	c
谢榕鲜等 ^[24]	53.1	53.1	35.5	35.5	16	20	80	130	919.7	919.7	735.7*	735.7*	126.8	225.2	7	b
刘中良等 ^[25]	33.7	33.7	31.0	31.0	10	16	120	120	835.0	877.0	663.0	696.0	34.1	98.8	13	b
ATAEI等 ^[5]	28.8	41.5	28.0	32.5	12	20	120	120	955.0	1 319.0	830.0	1 115.0	79.8	196.2	12	a
CHEN等 ^[26]	58.2	58.2	32.3	32.3	16	22	120	120	819.3	848.8	499.8	513.5	96.8	182.0	24	c
LAWAN等 ^[27]	33.0	37.8	31.0*	31.0*	12	16	60	60	807.0	897.0	720.0	768.0	43.2	62.1	18	b
ZHANG等 ^[28]	58.8	62.2	31.1	32.5	22	27	120	120	919.7	1 000.0	735.7*	900.0*	237.5	384.4	22	a
FANG等 ^[29]	61.2	150.3	39.9	43.0	16	27	50	100	937.2	1 435.1	925.5	1 310.0	99.1	356.1	18	a
总计	28.8	150.3	28.0	43.0	10.0	27.0	50.0	130.0	434.1	1 435.1	264.8	1 310.0	34.1	384.4	200	

注:*为原始数据缺失值,包括18个(9%)混凝土弹性模量 E_c 和49个(24.5%)螺栓屈服强度 f_y 缺失,为尽可能保留数据分布规律,减小对后续结果的影响,通过计算同一材料等级下非缺失值的平均值来补全这些缺失值。

预先检查数据库的质量是提升模型准确性的关键。马氏距离^[20]能够表示样本点与分布中心之间的距离,具有尺度无关的特征。在多维随机变量 x 的情境下,若其均值为 μ 且协方差矩阵为 Σ ,马氏距离可以定义为:

$$D_M(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (1)$$

数据库中200个样本的马氏距离分布如图3所示。利用马氏距离可以对样本进行识别,提高数据质量,并深入研究模型的应用范围。观察样本分布,可以发现大部分样本点均匀分布在样本中心附近。样本中,90%的马氏距离均低于75 000的阈值,表明大多数样本具有相近的统计特征。有20/200的样本离中心较远,这些样本集中在四篇文献中。经复核,确认这些样本在数据收集与处理的过程中均保持准确无误,并发现它们源于极端材料条件下的试验设置,比如采用了较小的螺栓直径或较大的混凝土强度等级。为提高模型的泛化能力,保留这些样本,并使其更多地参与模型训练。

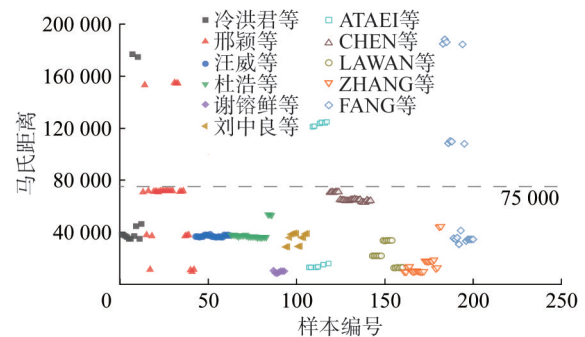


图3 马氏距离图

Fig. 3 Mahalanobis distance diagram

为了更好地评估参数变量之间的相关性,辅助承载力预测模型的参数选择,本文进行了参数变量的相关性分析。图4为各参数之间的相关系数热力图。由图可知,参数间的相关系数越接近1,意味着两个变量之间的正相关性越强。分析结果表明, P_u 与 d 之间的相关系数最大,表明二者变化趋势一致且相关性最强。此外, P_u 与

f_u 、 f_c 相关系数也相对较高。同时,注意到 f_c 与 E_c 、 f_u 与 f_y 的相关系数均接近 1,表明它们之间相关性非常强。

为了筛选出合适的特征,以防止机器学习模型因参数过多而过度复杂,并增强模型的泛化能力,根据参数间相关性大小排序,初步选定 d 、 f_c 、 f_u 、 h 这四个参数进行后续的数据分析。

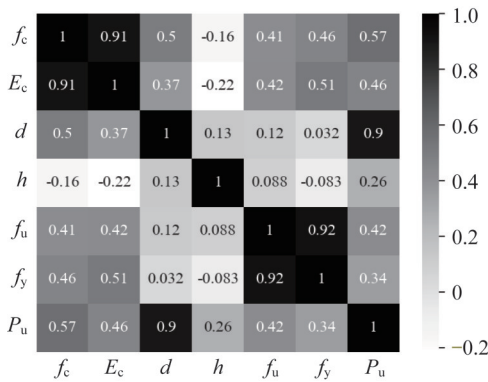


图4 相关系数热力图

Fig. 4 Correlation coefficient heatmap

2 基于机器学习方法的预测模型

为选取最适合预测螺栓抗剪连接件抗剪承载力的机器学习模型,本节首先构建了四种基于不同机器学习算法的预测模型。通过对这些模型进行超参数优化和多种评价指标的综合考量,选择表现最佳的机器学习算法。最后,基于这一最佳算法,进一步建立五个具有不同特征组合的机器学习模型,并通过对比分析,筛选出最佳的输入特征组合,为可解释机器学习方法的实施提供了最佳的机器学习模型。

2.1 机器学习方法及超参数优化

本文的预测任务属于监督学习中的回归问题,因此,对四种适用于此类问题的机器学习算法进行比选。

遗传算法作为一种优化算法,适用于超参数的优化。经过反复尝试,发现当遗传算法的种群大小设定为 500、迭代次数设置为 20、且基因变异概率调整为 0.1 时,能够找到满足本文机器学习模型精度要求的最优超参数组合。因此,在后续进行的遗传算法优化过程中,均采用了这一参数组合,以确保算法的稳定性。

(1) 人工神经网络(ANN)^[18]

数据集被随机分为 80% 的训练集和 20% 的测试集。随后,对训练集和测试集的输入特征与目标变量进行标准化处理,确保数据的均值为 0,标准差为 1。在模型的构建过程中,采用了基于 Python 的 Tensorflow 开源框架。隐藏层的激活函数采用线性整流函数 Relu 函数,以引入非线性因素。模型的训练则采用均方差(Mean

Squared Error, M_{SE}) 为损失函数进行优化。模型的学习率设定为 0.001,最大迭代次数限制为 10 000,并引入早期停止机制,设置惩罚值为 50。图 5 展示出不同隐藏层和神经元个数组别的 ANN 模型精度雷达图。通过测试集可决系数(R^2)和平均绝对百分比误差(M_{APE})来评估模型的预测能力。坐标轴中,“4”代表仅有一个隐藏层,包含 4 个神经元;“8+4”则表示有两个隐藏层,第一层有 8 个神经元,第二层有 4 个神经元。在保证预测精度的同时,为降低模型复杂度,缓解过拟合风险,本文选取了结构简单且表现良好的模型,即具有单个隐藏层包含 8 个神经元的 ANN 模型,以进行后续的比较和分析。

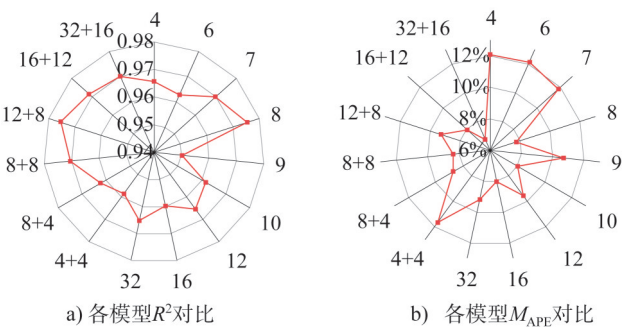


图5 不同隐藏层和神经元个数 ANN 模型精度雷达图
Fig. 5 Radar chart of accuracy for ANN models with different hidden layers and neuron numbers

(2) 支持向量回归(SVR)^[18]

利用 Sklearn 开源机器学习库,通过半经验法和遗传算法优化超参数配置。选择径向基函数作为核函数,设置函数次数为 3,核系数则自动计算。此外,惩罚参数设定为 180,最大迭代次数为 3 000 次,容错系数设置为 0.6。

(3) 随机森林(RF)^[18]

采用与支持向量回归模型相同的超参数优化方法,得到以下的最优超参数配置:决策树的数量为 411 棵,每棵树的最大深度限制为 7,允许使用的最大特征数量为特征总数的平方根,内部节点划分所需的最小样本数量设定为 4,而叶节点所需的最小样本数量则设为 1。

(4) 极端梯度提升树(XGBoost)^[19]

通过基于 Python 的 XGBoost 开源机器学习库,通过超参数优化,得到以下的超参数配置:决策树的数量为 46 棵,每棵树的最大深度限制为 5。在模型训练过程中,采用平方误差作为损失函数,子采样和列采样的比例均被设置为 0.8。

2.2 不同机器学习算法的预测结果对比

为了比较不同机器学习算法在预测螺栓抗剪连接件抗剪承载力方面的精度,本文采用可决系数(R^2)和均方根误差(R_{MSE})作为评估模型预测能力的关键指标。

图6展示了全部200组数据的预测结果,而图7则对比了ANN、SVR、RF和XGBoost四个模型在预测精度上的差异。结果显示:XGBoost模型在测试集的预测中表现最优,其可决系数和均方根误差均优于其他三种模型,其测试集可决系数达到了0.987。这是由于相较于其他算法,XGBoost以串行方式构建决策树,能够高效地捕捉非线性关系,同时引入正则项,有效地控制了模型的复杂度,防止了过拟合,且训练速度快,便于超参数优化,所以在处理本文多特征回归问题时呈现出了较好的预测能力。这四种机器学习算法在测试集上的 R^2 均超过了0.970。同时 M_{APE} 保持在8%以内,这充分说明了这些机器学习算法对于螺栓抗剪连接件抗剪承载力具有出色的学习能力和预测能力。

2.3 不同特征组合的XGBoost模型对比

上述对比分析表明,XGBoost算法在预测中表现出较高的精度。之前的相关性分析显示, d 与 P_u 的变化趋势高度一致,且二者之间的相关性最强。基于螺栓直径,选择了五组不同的特征组合,并应用XGBoost算法进行预测。每组超参数均通过遗传算法进行优化,以获得最佳配置。图8为不同特征组合XGBoost模型预测精度,由图可以看出,包含 f_c 、 f_u 、 d 、 h 的特征组合在XGBoost模型中的表现最佳,其训练集和测试集的可决系数均接近1,且误差较小。相较之下,包含更多特征(f_c 、 f_u 、 d 、 h 、 E_c 、 f_y)的组合在模型训练时出现了欠拟合问题。这是由于过多的输入参数使得模型复杂度过高,以至于无法有效地学习到与目标特征相关的关键信息。这一现象与相关性矩阵的分析结果相契合。

此外,仅包含 d 和 f_c 或 d 和 f_u 的特征组合也能提供较高精度的预测,这与现有研究中的经验模型^[10-14]中所采用的输入特征相吻合(如表2所示)。值得注意的是, f_c 、 f_u 、 d 这三个特征的组合在各项指标上非常接近最佳特征组合,表明它们对螺栓抗剪连接件抗剪承载力的预测具有决定性作用。同时, h 在一定程度上有助于提升预测精度。

3 基于可解释机器学习的预测模型

3.1 基于SHAP方法的特征重要性解释

上述ANN、RF、XGBoost与SVR等机器学习模型在预测精度上表现出色,然而其内部工作机制复杂,难以被直观理解。这些模型无法清晰地表征输入特征与承载力之间的复杂非线性关系,从而导致无法确切判断所得模型的可靠性。为了弥补这一不足,一些可解释机器学习方法^[18]被提出,能够帮助评估模型的可信度。

SHAP方法,作为一种基于博弈论的可解释机器学习方法^[30],旨在描述机器学习模型的可加性特征的重要性。

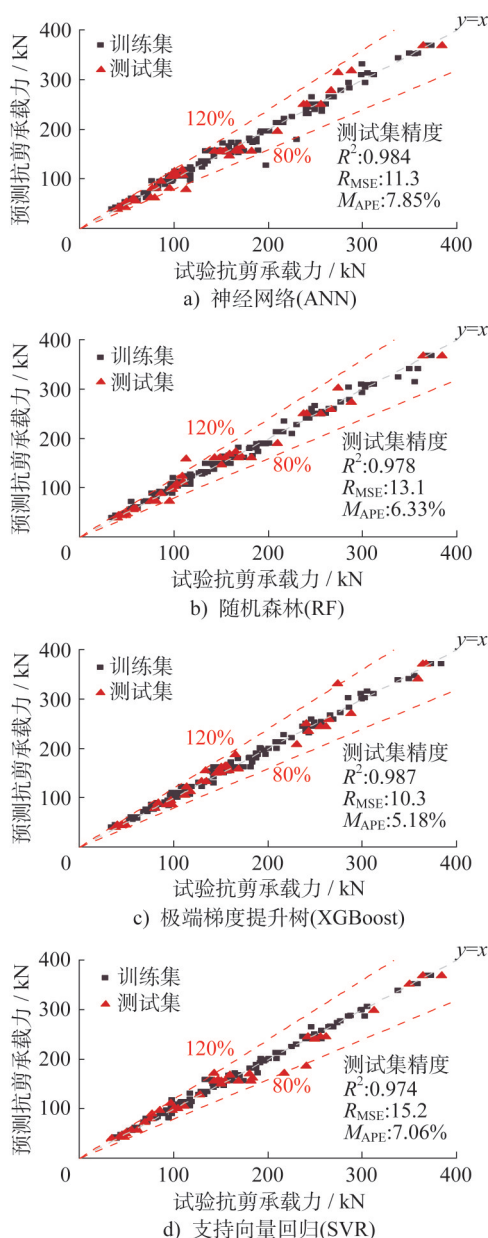


图6 不同机器学习预测模型的预测结果

Fig. 6 Prediction results of different machine learning models

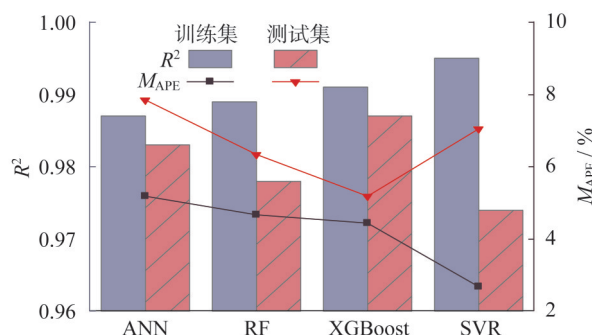


图7 各机器学习模型抗剪承载力预测精度对比

Fig. 7 Comparison of shear bearing capacity prediction accuracy among different machine learning models

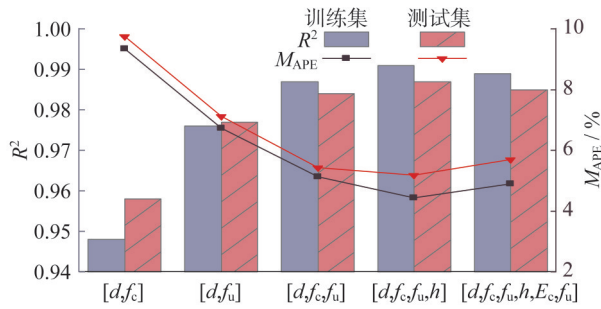


图8 不同特征组合 XGBoost 模型预测精度

Fig. 8 Prediction accuracy of XGBoost model with different feature combinations

表2 螺栓抗剪连接件抗剪承载力的经验模型

Table 2 Empirical model of shear bearing capacity for bolted shear connectors

模型	公式	编号
KWON 等 ^[10]	$0.5A_s f_u$	(2)
LIU 等 ^[11]	$0.66A_s f_u$	(3)
CHEN 等 ^[13]	$(0.0007f_y + 0.53)0.23d^{1.78}f_c^{0.29}$	(4)
邢颖 等 ^[14]	$0.414d^{0.87}f_c^{0.78}f_u^{0.39}$	(5)

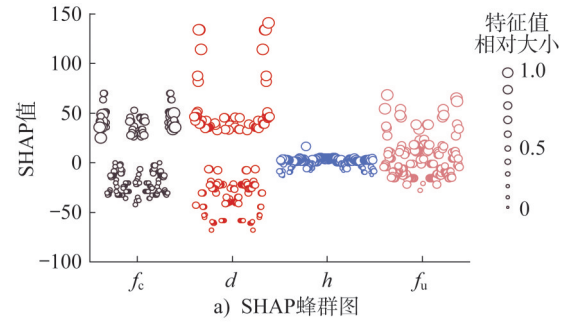
该方法通过对每个特征的边际贡献进行加权求和,进而将模型的预测值表达为二元变量的线性函数,如下式所示:

$$g(x') = \varphi_0 + \sum_{i=1}^M \varphi_i x'_i = f(x) \quad (6)$$

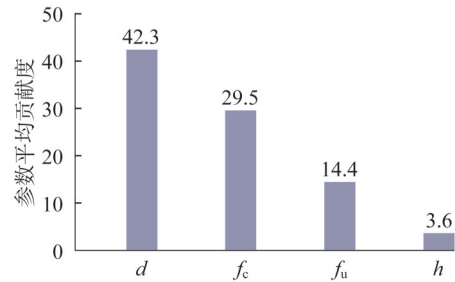
式中: $g(x')$ 为解释模型, $f(x)$ 为机器学习模型, x 为输入特征向量, x' 是简化输入特征,与 x 存在映射关系; M 代表输入特征的个数; φ_0 代表所有样本的预测目标均值; φ_i 为第 i 个特征的贡献值。

通过对 XGBoost 机器学习模型进行解释,可以计算并记录每个样本的 SHAP 值以及各特征的贡献值。这些数据被用于生成各输入特征的 SHAP 值蜂群图(图 9a)),以及各特征的平均贡献度图(图 9b))。SHAP 值蜂群图展示了各特征的 SHAP 值的分布情况,其中左侧纵坐标代表特征的 SHAP 值,而右侧图例则为气泡大小,反映了各特征值的相对大小。

从图 9 可知, d 对抗剪承载力的平均贡献最为显著。当 d 较小时,对应的数据点气泡较小,表明此时 d 对抗剪承载力的平均贡献有限。随着 d 的增加,数据点气泡明显变大,表明 d 对抗剪承载力贡献显著增强。此外, f_c 和 f_u 的贡献度次之。 f_c 的增大能够在一定程度上提高螺栓抗剪连接件的抗剪承载力,但当 f_c 增大到某一阈值后,连接件可能发生螺栓剪断破坏,此时 f_c 的 SHAP 值反而减小。相对而言, f_u 的增大能够有效增加抗剪承载力,其 SHAP 值随着 f_u 增大而增大。最后, h 的贡献最小,这与



a) SHAP 蜂群图



b) 各特征参数对抗剪承载力平均贡献度

图9 XGBoost 模型的 SHAP 全局分析图

Fig. 9 SHAP global analysis diagram of XGBoost model

相关系数分析的结论一致,从而验证了基于 XGBoost 算法的螺栓抗剪连接件抗剪承载力预测模型的合理性。

3.2 满足量纲平衡的公式生成方法

首先,显式预测公式具有高度的可解释性,能够清晰地揭示输入变量(如材料强度、螺栓尺寸等)与结构性能之间的直接关联,有助于深入理解设计背后的物理机制。其次,可以根据实际需求灵活调整显式公式中的参数,进行分析和优化设计。最后,显式公式能够显著提升计算速度,满足工程实践中的多样化需求。

符号回归旨在发现能够最佳拟合给定数据集的隐藏数学表达式或函数。目前,符号回归的实现方法主要包括基于遗传编程算法^[19]和基于长短期记忆网络与符号回归表达式树来满足量纲平衡的公式生成方法^[21]等。本文提出一种新的方法,该方法具有符号回归在特征获取上的优势和遗传算法在参数标定上的高效性,来生成满足量纲平衡的公式。

使用清华大学设计公式智能生成模型 EqDisce^[31]进行初步公式拟合。为拟合出一个简单实用的方程,符号选择加、减、乘、除、乘方和开方六个基础运算符,输入特征选择 f_c 、 f_u 、 d 、 h 。用 R^2 、 M_{AE} 、 R_{MSE} 三个不同的评判标准进行多次符号回归模型训练,得到精度较高的拟合公式及精度如表 3 所示。

结果表明:在满足量纲平衡的条件下,公式(7)~(9)均展现了较高的预测精度,其中公式(7)精度最高,其第二项 $0.097dhf_u$ 可看作公式(2)和(3)引入预埋长度项的修正,从而进一步提升了预测模型的精确度和可靠性。

表 3 拟合公式及精度对比

Table 3 Fitting formula and comparison of accuracy

公式	R^2	R_{MSE}	编号
$0.044d^2f_u + 0.097dhf_u$	0.844	34.19	(7)
$0.38f_u d^2 + f_c d^2$	0.809	37.81	(8)
$0.44d^2f_u$	0.788	39.86	(9)

基于上一节所述的 SHAP 分析方法,绘制螺栓直径 d 与其对应 SHAP 值的散点图,如图 10 所示。观察该图,可以看出在 $d=20$ mm 时,SHAP 值由负转正,这表明当螺栓直径达到或超过 20 mm 时,其对螺栓抗剪连接件抗剪承载力增长的贡献显著增强。此外,注意到在螺栓直径 20 mm 两侧,样本点数量相当。

为进一步提高公式的准确度,基于公式(7),将其适用范围细分为两个区间: $d < 20$ mm 和 $d \geq 20$ mm。将数据集划分为 80% 训练集和 20% 测试集,对训练集使用遗传算法标定分段公式中的常数参数,得到螺栓抗剪连接件的抗剪承载力预测公式如下式所示:

$$P_u = \begin{cases} (0.22d + 0.036h)df_u/1000 & d < 20 \text{ mm} \\ (0.44d + 0.016h)df_u/1000 & d \geq 20 \text{ mm} \end{cases} \quad (10)$$

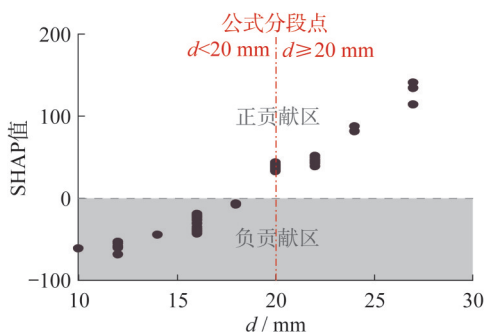


图 10 d 的 SHAP 值图

Fig. 10 SHAP value diagram for d

图 11 为公式(10)的预测结果,从图中可以看出,将预埋长度 h 作为修正项,对公式进行分段处理以及使用遗传算法进行参数标定能够有效提升公式的预测精度。

3.3 抗剪承载力模型对比

表 2 列出了 KWON 等^[10]、LIU 等^[11]、CHEN 等^[13]、邢颖等^[14]给出的螺栓抗剪连接件抗剪承载力的经验模型,通过测试集的 40 组试验值与模型计算值比值的均值(μ)、可决系数(R^2)、均方根误差(R_{MSE})、平均绝对百分比误差(M_{APE})来比较各计算模型的精度,计算结果见表 4。

表 4 中的均值(μ)均在 0.8~1.2 之间,说明以上计算模型均较为合理。邢颖等^[14]的模型预测值总体偏大,原因是其基于试验和有限元结果进行建模,且大直径螺栓

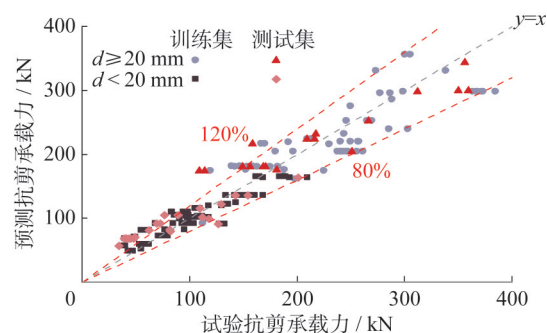


图 11 公式(10)预测结果

Fig. 11 Prediction results of formula (10)

数据量较少,对直径较大的螺栓抗剪连接件抗剪承载力未能很好预测。CHEN 等^[13]的模型 μ 接近 1,但离散度较大,且公式等号两侧不满足量纲平衡条件。各模型对测试集的预测情况如图 12 所示,从图中可以看出,最佳机器学习模型精度最高,本文公式模型次之。结果表明,本文提出的满足量纲平衡的计算公式,其预测精度较现有经验模型高,可解释性强,可为螺栓抗剪连接件抗剪承载力的计算提供科学依据,也可看作是对 KWON 等^[10]和 LIU 等^[11]模型的改进。

表 4 各计算模型测试集精度对比

Table 4 Comparison of accuracy for test sets among various computational models

模型	R^2	R_{MSE}	$M_{APE}/\%$	μ
KWON 等 ^[10]	0.731	47.2	20.0	1.16
LIU 等 ^[11]	0.885	30.8	26.2	0.88
CHEN 等 ^[13]	0.826	38.0	21.0	1.03
邢颖等 ^[14]	0.406	70.1	43.2	0.83
公式(10)	0.899	28.9	25.3	0.90
最佳机器学习模型	0.987	10.3	5.2	1.00

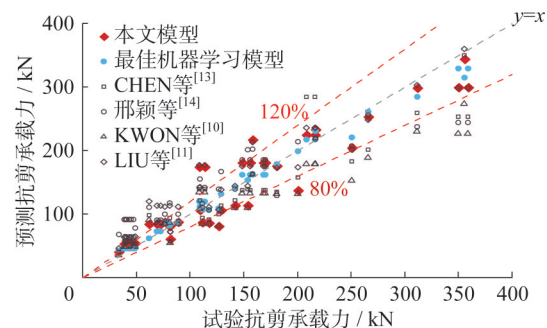


图 12 各计算模型测试集散点图

Fig. 12 Diagram of testing distribution points for different computational models

4 结论

本文旨在运用可解释机器学习方法构建螺栓抗剪连接件抗剪承载力的预测模型,并对其进行解析。首先从文献中搜集了200组螺栓抗剪连接件推出试验数据,并应用神经网络(ANN)、随机森林(RF)、极端梯度提升树(XGBoost)、支持向量回归(SVR)构建了相应的机器学习模型。随后,采用SHAP方法对特征的重要性进行了分析。最后,通过结合符号回归获取特征和遗传算法标定参数来满足量纲平衡的公式生成方法,推导出了螺栓抗剪连接件抗剪承载力实用预测公式。研究表明:

(1) 对于本文收集的三种常用螺栓抗剪连接件,四种机器学习模型均展现出良好的预测能力,其中XGBoost预测精度最高,测试集可决系数达到了0.987。

(2) 利用可解释机器学习方法对最佳模型进行分析,评估了各参数的贡献度,并通过模型特征分析,确定了训练效果最佳且拟合风险最小的参数组合。

(3) 提出了一种新颖的公式生成方法,结合符号回归获取特征和遗传算法标定参数,利用本文建立的数据库拟合出了螺栓抗剪连接件抗剪承载力的预测公式,揭示了关键特征与抗剪承载力之间的显性关系。该公式的预测精度相较于经验模型有明显提升,为螺栓抗剪连接件抗剪承载力的计算提供了可靠依据。

备注:本文所收集的试验样本和编写的程序均可从github.com/Bond-Ma/Bolt-ML获取,XGBoost最佳机器学习模型也可直接下载使用。

参考文献:

- [1] LI X, HE J, ZHOU Y X, et al. A comprehensive review of shear connectors in demountable composite beams [J]. Journal of Constructional Steel Research, 2024, 218: 108723. DOI:10.1016/j.jcsr.2024.108723.
- [2] CHEN J, WANG W, DING F X, et al. Behavior of an advanced bolted shear connector in prefabricated steel-concrete composite beams [J]. Materials, 2019, 12 (18) : 2958. DOI:10.3390/ma12182958.
- [3] WANG W, ZHANG X D, REN Y, et al. Finite element modelling of bolt shear connections in prefabricated steel lightweight aggregate-concrete composite beams [J]. Buildings, 2022, 12 (6) : 758. DOI: 10.3390/buildings 12060758.
- [4] WANG W, ZHANG X D, ZHOU X L, et al. Study on shear behavior of multi-bolt connectors for prefabricated steel-concrete composite beams [J]. Frontiers in Materials, 2021, 8: 625425. DOI:10.3389/fmats.2021.625425.
- [5] ATAEI A, ZEYNALIAN M. A study on structural performance of deconstructable bolted shear connectors in composite beams [J]. Structures, 2021, 29: 519-533. DOI: 10.1016/j.istruc.2020.11.065.
- [6] LIU X P, BRADFORD M A, CHEN Q J, et al. Finite element modelling of steel-concrete composite beams with high-strength friction-grip bolt shear connectors [J]. Finite Elements in Analysis and Design, 2016, 108: 54-65. DOI: 10.1016/j.finel.2015.09.004.
- [7] LIU X P, BRADFORD M A, ATAEI A. Flexural performance of innovative sustainable composite steel-concrete beams [J]. Engineering Structures, 2017, 130: 282-296. DOI:10.1016/j.engstruct.2016.10.009.
- [8] ATAEI A, BRADFORD M A, LIU X P. Experimental study of composite beams having a precast geopolymer concrete slab and deconstructable bolted shear connectors [J]. Engineering Structures, 2016, 114: 1-13. DOI: 10.1016/j.engstruct.2015.10.041.
- [9] PAVLOVIC M, MARKOVIC Z, VELJKOVIC M, et al. Bolted shear connectors vs. headed studs behaviour in push-out tests [J]. Journal of Constructional Steel Research, 2013, 88: 134-149. DOI:10.1016/j.jcsr.2013.05.003.
- [10] KWON G, ENGELHARDT M D, KLINGNER R E. Behavior of post-installed shear connectors under static and fatigue loading [J]. Journal of Constructional Steel Research, 2010, 66(4):532-541. DOI:10.1016/j.jcsr.2009.09.012.
- [11] LIU X P, BRADFORD M A, LEE M S S. Behavior of high-strength friction-grip bolted shear connectors in sustainable composite beams [J]. Journal of Structural Engineering, 2015, 141(6):04014149. DOI:10.1061/(ASCE)ST.1943-541x.0001090.
- [12] LIU S, GAO D P, CHEN B W. Experimental investigation on steel-bamboo composite shear connections [J]. Journal of Building Engineering, 2022, 61: 105207. DOI: 10.1016/j.jobbe.2022.105207.
- [13] CHEN X, YANG Y, XUE Y C, et al. Shear behavior of large studs and novel bolted connectors in steel-UHPC composite beams [J]. Structures, 2022, 45: 2091-2103. DOI: 10.1016/j.istruc.2022.09.019.
- [14] 邢颖,刘雁斌,史才军,等.可恢复组合梁中高强螺栓连接件抗剪性能试验 [J].中国公路学报,2023,36(6):132-142. DOI:10.19721/j.cnki.1001-7372.2023.06.012. XING Ying, LIU Yanbin, SHI Caijun, et al. Experimental study on shear performance of high-strength bolt connectors in recoverable composite beams [J]. China Journal of Highway and Transport, 2023, 36 (6) : 132-142. DOI: 10.19721/j.cnki.1001-7372.2023.06.012. (in Chinese)
- [15] RAZAVI SETVATI M, HICKS S J. Machine learning models for predicting resistance of headed studs embedded in concrete [J]. Engineering Structures, 2022, 254: 113803. DOI: 10.1016/j.engstruct.2021.113803.
- [16] ZHU J S, FAROUK A I B. Development of hybrid models for shear resistance prediction of grouped stud connectors in concrete using improved metaheuristic optimization techniques [J]. Structures, 2023, 50: 286-302. DOI:10.1016/j.istruc.2023.02.040.

- [17] 纪守领,李进锋,杜天宇,等. 机器学习模型可解释性方法、应用与安全研究综述[J]. 计算机研究与发展, 2019, 56(10):2071-2096.
JI Shouling, LI Jinfeng, DU Tianyu, et al. Survey on techniques, applications and security of machine learning interpretability [J]. Journal of Computer Research and Development, 2019, 56(10):2071-2096. (in Chinese)
- [18] 冯德成,吴刚. 混凝土结构基本性能的可解释机器学习建模方法[J]. 建筑结构学报, 2022, 43(4): 228-238. DOI: 10.14006/j.jzjgxb.2020.0491.
FENG Decheng, WU Gang. Interpretable machine learning-based modeling approach for fundamental properties of concrete structures[J]. Journal of Building Structures, 2022, 43(4):228-238. DOI: 10.14006/j.jzjgxb.2020.0491. (in Chinese)
- [19] 戴理朝,王冲,袁平,等. 基于可解释机器学习的锈蚀RC构件抗剪承载力预测模型[J]. 吉林大学学报(工学版), 2024, 54(11): 3231-3243. DOI: 10.13229/j.cnki.jdxbgxb.20230024.
DAI Lizhao, WANG Chong, YUAN Ping, et al. Prediction model for shear capacity of corroded RC beams based on interpretable machine learning[J]. Journal of Jilin University (Engineering and Technology Edition), 2024, 54(11): 3231-3243. DOI: 10.13229/j.cnki.jdxbgxb.20230024. (in Chinese)
- [20] MAHALANOBIS P C. On the generalised distance in statistics [C]//Proceedings of the National Institute of Sciences. Kolkata: [s.n.], 1936.
- [21] 冷洪君. 钢-混凝土组合梁高强螺栓连接件抗剪性能有限元分析[D]. 青岛: 山东科技大学, 2020.
LENG Hongjun. Finite element analysis of shear performance of steel-concrete composite beam of high-strength bolt connectors[D]. Qingdao: Shandong University of Science and Technology, 2020. (in Chinese)
- [22] 汪威. 装配式组合梁高强螺栓连接件力学性能研究[D]. 湘潭: 湘潭大学, 2019.
WANG Wei. Mechanical performance study on high-strength bolt connectors in prefabricated steel-concrete composite beams[D]. Xiangtan: Xiangtan University, 2019. (in Chinese)
- [23] 杜浩,张冰,胡夏闽,等. 钢-混凝土组合梁螺栓连接件受剪性能试验研究[J]. 建筑结构学报, 2017, 38(增刊1): 308-314. DOI: 10.14006/j.jzjgxb.2017.S1.043.
DU Hao, ZHANG Bing, HU Xiamin, et al. Experimental study on the shear behavior of a bolt connector in a steel-concrete composite beam[J]. Journal of Building Structures, 2017, 38(Suppl. 1): 308-314. DOI: 10.14006/j.jzjgxb.2017.S1.043. (in Chinese)
- [24] 谢镛鲜. 装配式组合梁中摩擦型螺栓连接件抗剪性能与设计方法研究[D]. 南宁: 广西大学, 2022.
XIE Rongxian. Research on the shear resistance and design methods of friction bolt connectors for assembled composite beam[D]. Nanning: Guangxi University, 2022. (in Chinese)
- [25] 刘中良,陈俊,霍静思. 装配式组合梁高强螺栓连接件抗剪性能试验研究[J]. 建筑结构, 2017, 47(10): 65-70, 64. DOI: 10.19701/j.jzjg.2017.10.011.
LIU Zhongliang, CHEN Jun, HUO Jingsi. Experimental study on shear behavior of high strength bolt connection for prefabricated composite beam [J]. Building Structure, 2017, 47(10): 65-70, 64. DOI: 10.19701/j.jzjg.2017.10.011. (in Chinese)
- [26] CHEN B C, LIU A R, ZHANG J P, et al. Behavior of T-shaped embedded-nut bolted shear connectors in prefabricated steel-concrete composite beams [J]. Engineering Structures, 2022, 272: 114983. DOI: 10.1016/j.engstruct.2022.114983.
- [27] LAWAN M M, TAHIR M M, MIRZA J. Bolted shear connectors performance in self-compacting concrete integrated with cold-formed steel section [J]. Latin American Journal of Solids and Structures, 2016, 13(4): 731-749. DOI: 10.1590/1679-78252004.
- [28] ZHANG Y J, LIU A R, CHEN B C, et al. Experimental and numerical study of shear connection in composite beams of steel and steel-fibre reinforced concrete [J]. Engineering Structures, 2020, 215: 110707. DOI: 10.1016/j.engstruct.2020.110707.
- [29] FANG Z C, LIANG W B, FANG H Z, et al. Experimental investigation on shear behavior of high-strength friction-grip bolt shear connectors in steel-precast UHPC composite structures subjected to static loading [J]. Engineering Structures, 2021, 244: 112777. DOI: 10.1016/j.engstruct.2021.112777.
- [30] LUNDBERG S M, LEE S. A unified approach to interpreting model predictions [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017.
- [31] 陈培尧,王琛,丁然,等. 土木工程设计公式智能发现方法[J]. 建筑结构学报, 2024, 45(7): 80-88. DOI: 10.14006/j.jzjgxb.2023.0597.
CHEN Peiyao, WANG Chen, DING Ran, et al. An intelligent discovery method for design formulas in civil engineering [J]. Journal of Building Structures, 2024, 45(7): 80-88. DOI: 10.14006/j.jzjgxb.2023.0597. (in Chinese)