DOI: 10.11835/j.issn.2096-6717.2024.091
开放科学(资源服务)标识码OSID:

高性能混凝土强度-组分相关性分析的可解释机器学习模型

靳新斌[a],李艳[a],王志国[a],赵宇欣[a],梁文彪[b]
(长安大学 a.地质工程与测绘学院;b.理学院,西安 710061)

摘　要:高性能混凝土强度与其组分之间存在复杂的相关性,传统的黑盒模型因缺乏可解释性而无法揭示二者之间的实际依存关系。基于高性能混凝土大样本数据集,构建机器学习模型,采用Hyperopt算法对机器学习模型进行优化,并与SHAP可解释技术和特征依赖值算法集成,计算得到高性能混凝土强度与各组分之间非线性关系的可视化表征以及多组分之间的交互耦合作用。结果表明:Hyperopt-极端梯度提升模型为可信任的机器学习模型,具有较高的准确性与稳健性;在高性能混凝土的8种组分(因素)中,龄期、水泥、水和矿渣对抗压强度起关键控制作用,与抗压强度的关系分别满足指数函数、幂函数、高斯函数和指数函数关系,且4种关键因素之间还存在明显的交互耦合效应。龄期与水泥、水、矿渣的交互作用曲线分别符合LogisticCum、LogNormal2D和Power2D函数,水泥与水、水泥与矿渣的交互作用均符合Poly2D函数,矿渣与水的交互作用符合ExtremeCum函数。

关键词:高性能混凝土;Hyperopt超参数优化;可解释机器学习;交互作用
中图分类号:TU528.01　　**文献标志码:**A　　**文章编号:**2096-6717(2026)04-0210-10

Explainable machine learning model for strength-component correlation analysis of high-performance concrete

JIN Xinbin[a], LI Yan[a], WANG Zhiguo[a], ZHAO Yuxin[a], LIANG Wenbiao[b]
(a. School of Geological Engineering and Geomatics; b. School of Sciences, Chang'an University, Xi'an 710061, P. R. China)

Abstract: The strength of high-performance concrete has an intricate relationship with its components. It is evident that traditional black-box models are characterised by an absence of interpretability. As a result, they fail to reveal the actual dependency between concrete strength and its components. Using a large sample dataset of high-performance concrete, machine learning models are constructed and optimized with the Hyperopt

收稿日期:2024-07-10
基金项目:陕西省博士后科研项目(2023BSHEDZZ211);咸阳市科技创新团队项目(L2023CXNLCXTD005);陕西省住建厅科研开发类项目(2021-K40)
作者简介:靳新斌(2002-),男,主要从事建筑安全研究,E-mail:751822306@qq.com。
李艳(通信作者),女,副教授,E-mail:1259578602@qq.com。

Received: 2024-07-10
Foundation items: Shaanxi Postdoctoral Science Foundation Project (No. 2023BSHEDZZ211); Xianyang City Science and Technology Innovation Team Project (No. L2023CXNLCXTD005); Shaanxi Provincial Department of Housing and Urban-Rural Development Project (No. 2021-K40)
Author brief: JIN Xinbin (2002-), main research interest: building safety, E-mail: 751822306@qq.com.
LI Yan (corresponding author), associate professor, E-mail: 1259578602@qq.com.

algorithm. The models are integrated with SHAP explainability techniques and feature dependency algorithms. A visual representation of the nonlinear relationship between the strength of high-performance concrete and its components is calculated. Furthermore, the interaction and coupling effects between multiple components are examined. The results indicate the following: Hyperopt-extreme gradient boosting model is a trusted machine learning model with high accuracy and robustness. Among the eight components(factors), age, cement, water and slag play a key controlling role in compressive strength, and their relationships with compressive strength satisfy the exponential function, power function, Gaussian function, and exponential function, respectively. At the same time, there is a significant interactive coupling effect between the four key factors; the interaction curves between age and cement, water, and slag adhere to the LogisticCum, LogNormal2D, and Power2D functions. The interaction between cement and water, as well as between cement and slag, conforms to the Poly2D function, while the interaction between slag and water conforms to the ExtremeCum function.

Keywords: high-performance concrete; Hyperopt hyperparameter optimization; explainable machine learning; interaction effect

随着现代建筑工程对混凝土性能的要求越来越高，混凝土材料结构体系越来越具有不确定性和复杂性。作为一种添加大量化学外加剂与辅助胶凝材料的复合材料[1-2]，高性能混凝土的强度与各组分之间均具有高度非线性关系[3]，揭示二者的关系对于高性能混凝土的改性及优化配比设计具有重要意义。

基于数据驱动的机器学习技术可以根据数据中输入与输出的映射关系[4-6]建立精确且稳定的强度预测模型，然而，机器学习模型的“黑盒”特性使得预测结果难以解释[7]。为了提高黑盒模型的可解释性，一些学者采用特征重要性分析来解释机器学习模型的预测结果，如Kang等[8]建立了钢纤维混凝土机器学习模型，发现水灰比和硅灰是影响其抗压强度的关键因素。虽然特征重要性分析能提供特征的重要性排序，但无法揭示每个特征对预测结果的具体影响规律。SHAP方法能够从整体数据集的角度给出特征的重要性排序以及特征的定性影响规律，目前已广泛应用于材料特性及影响因素的相关性研究中[9]，如Zhao等[10]采用机器学习技术建立了水泥-注浆材料强度-组分之间的机器学习预测模型，并通过SHAP可解释技术对模型进行了单一影响因素的特征贡献解释。相关研究为复合材料结构体系的配比依赖设计提供了参考依据。然而，混凝土材料各组分之间存在着复杂的水化反应，多个组分之间存在耦合作用及交互影响[11]，并且实际工程需要同时考虑混凝土强度、组分之间的多维关系，SHAP针对单一因素分析的方法难以满足现代建筑工程对多维关系的需求。因此，高性能混凝土同时考虑了各组分主效应及交互作用效应的高维复杂关系，对于其复合设计和关键因素控制十分重要。

尽管机器学习技术对于解决高维复杂问题具有强大的计算能力，但计算结果的精准性和稳健性具有一定数据集依赖性[12]。在机器学习领域，将维度的概念引入对小样本的界定。根据维度的界定，Raudys等[13]将小样本数据集定义为机器学习维度为5～10（即输入因素为5～10）、训练样本为50～100的数据集。当前大量研究都集中在Raudys等[13]提出的范围界定内，如Hassani-Niaki等[14]、Rahmati等[15]采用机器学习方法分别对67组高分子混凝土和107组高强混凝土的力学性能进行了抗压强度预测估计。这些模型性能虽然优异，但研究表明，数据集样本容量较小时，机器学习模型性能欠稳定，故许多学者选择多次试算中性能最优的进行展示[16]。可见，足够样本容量的数据集是构建可信任机器学习模型的基础。

为进一步研究高性能混凝土强度-组分之间的复杂非线性关系，基于开源大样本数据集，采用Hyperopt算法对机器学习模型进行优化，构建稳健、可依赖的高性能混凝土强度预测模型，采用SHAP可解释技术与特征依赖值对影响高性能混凝土强度的各组分敏感性及交互效应进行分析。

1　基于大样本数据集的高性能混凝土强度预测模型构建

1.1　高性能混凝土强度预测模型的构建框架

高性能混凝土强度预测模型的构建逻辑如图1所示。首先，根据文献[2]，收集构建高性能混凝土强度-组分的大样本数据集（1 030组样本），在其中随机抽取121组样本，形成小样本数据集（根据Raudys等[13]对机器学习数据集的界定，将121组样本按7∶3划分训练集与测试集，85组样本数据用于训练，符合小样本数据特征，以备大样本数据集必

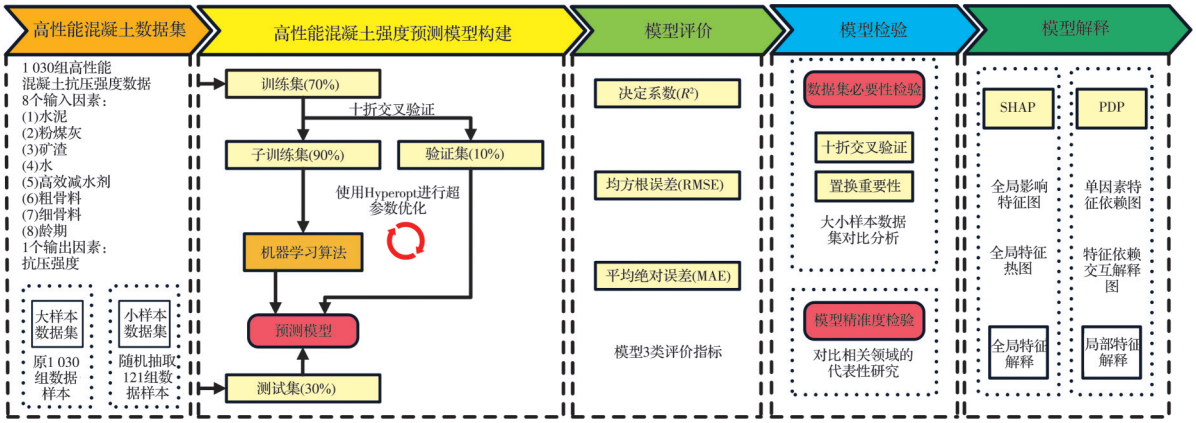


图 1 高性能混凝土预测模型总体框架

Fig. 1 Overall framework of prediction model for high-performance concrete

要性检验使用),按 7∶3 划分为训练集与测试集。随后,将大样本数据集输入至随机森林、梯度提升、轻量级梯度提升、极端梯度提升 4 种机器学习模型中,采用 Hyperopt 算法进行超参数优化,构建 Hyperopt-机器学习模型。采用决定系数(R^2),均方根误差(RMSE,E_{RMS})、平均绝对误差(MAE,E_{MA})3 个评价指标对 4 种机器学习模型的预测效能进行评估,筛选预测效能最优的机器学习模型。在此基础上,采用十折交叉验证与置换重要度技术进行大样本必要性检验,通过对比相关领域代表性研究结果进行模型精准度检验。最后,使用 SHAP 可解释技术与特征依赖值分析方法对高性能混凝土与各组分(因素)之间的相关性进行全局与局部特征解释。

1.2 大样本数据集的特征分析

大样本数据集源自文献[2]收集的 1 030 组高性能混凝土强度-组分数据,该数据集开源于加利福尼亚大学尔湾分校机器学习数据库中,其输入参数为高性能混凝土各组分及外部影响因素,包括水泥、高炉矿渣、粉煤灰、水、高效减水剂、粗骨料、细骨料、龄期等 8 种参数,输出参数为高性能混凝土轴心抗压强度。

研究表明[17],如果模型输入参数(即各影响因素)之间存在相关性,会导致模型的预测和解释性能变差,从而引起“多重共线性”问题。为保证预测模型的可靠性,两个输入参数之间的线性相关系数应低于 0.8[18],基于此,采用 Pearson 线性相关系数对所构建的数据集各输入参数进行线性相关性分析,如图 2 所示。图中正值代表正相关,负值代表负相关,数值大小代表相关系数的大小,数值越大,表明线性相关程度越高。

由图 2 可见,水泥、矿渣、粉煤灰、水、高效减水剂、粗骨料、细骨料和龄期等 8 种输入参数之间的相关系数均小于 0.8,不存在“多重共线性”的问题,满足机器学习模型建模的要求。同时,由图 2 还可以

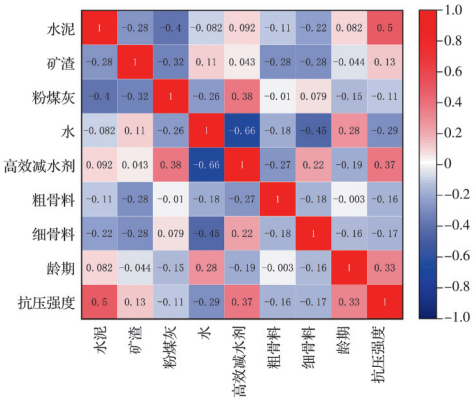


图 2 Pearson 相关分析图

Fig. 2 Pearson correlation analysis diagram

看出,高性能混凝土的抗压强度与水泥之间的相关系数为 0.50,是所有组分中与抗压强度相关系数值最高的,而粉煤灰与抗压强度之间的相关系数最小,相关系数仅 −0.11。8 种因素与抗压强度 Pearson 线性相关系数大小的排序为水泥>高效减水剂>龄期>水>细骨料>粗骨料>细骨料>矿渣>粉煤灰。

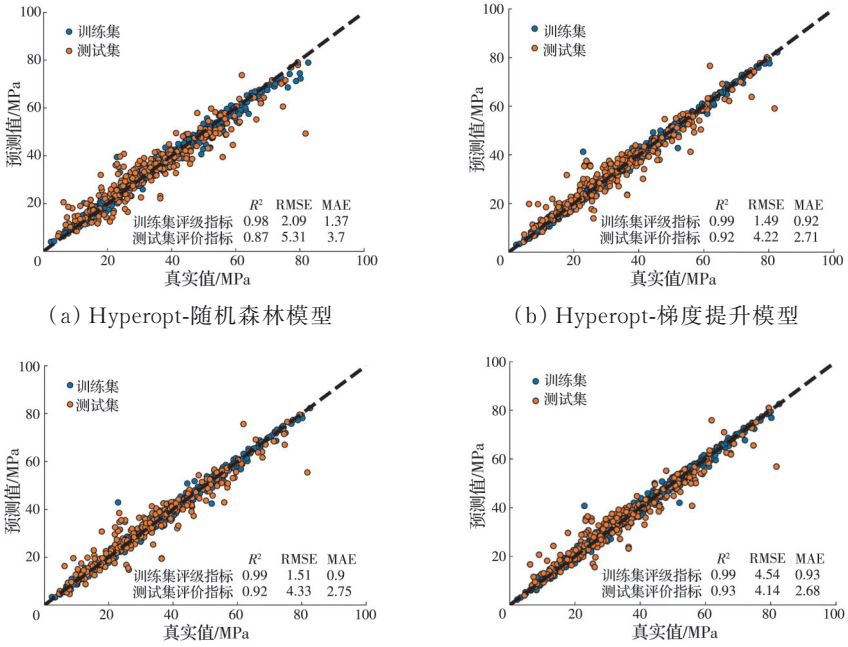
1.3 高性能混凝土强度预测模型的评价指标

决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)是机器学习预测模型构建中常用的评价指标[19-20]。其中 R^2 表示模型的准确程度,其取值在 0~1 之间,越接近于 1 表明模型拟合程度越好;RMSE 反映了预测值与实际值的偏离程度;MAE 表示实际值与预测值之间的差的绝对值的平均值,可以直观地看出模型预测值与真实值的差距大小。RMSE、MAE 越小,表明模型的拟合程度越好。

1.4 Hyperopt-极端梯度提升模型的构建

随机森林、梯度提升、轻量级梯度提升和极端梯度提升是常用的经典机器学习模型,为进一步提升其预测性能,采用 Hyperopt 优化技术对 4 种模型分别进行超参数优化,并使用优化后的机器学习模型对大样本数据集下的高性能混凝土抗压强度进

行预测计算。采用R^2、RMSE、MAE三个评价指标对预测结果进行定量评价，如图3所示。



（a）Hyperopt-随机森林模型　（b）Hyperopt-梯度提升模型

（c）Hyperopt-轻量级梯度提升模型　（d）Hyperopt-极端梯度提升模型

图3　Hyperopt优化后4种机器学习模型的预测性能

Fig. 3　Predictive performance of four machine learning models optimized by Hyperopt

由图3可见，Hyperopt-极端梯度提升模型训练集的R^2、RMSE和MAE三类指标大小为0.99、1.54和0.93，测试集R^2、RMSE和MAE三类指标大小为0.93、4.14和2.68，在4类经过Hyperopt超参数优化的模型中综合预测性能最优；Hyperopt-梯度提升模型与Hyperopt-轻量级梯度提升模型综合预测性能相近，均小幅度劣于Hyperopt-极端梯度提升模型；Hyperopt-随机森林模型训练集R^2、RMSE和MAE三类指标大小为0.98、2.05和1.35，测试集R^2、RMSE和MAE三类指标大小为0.87、5.24和3.67，在4类经过Hyperopt超参数优化的模型中综合预测性能最弱。

依托综合预测性能最优的Hyperopt-极端梯度提升模型，开展模型预测效能检验和强度-组分相关性分析。

2　Hyperopt-极端梯度提升模型的预测效能检验

2.1　模型精准性对比检验

为进一步检验高性能混凝土强度预测模型预测效能的精准性，基于高性能混凝土的强度-组分大容量大样本数据集，归纳总结近年来相关领域研究者提出的代表性预测模型，并将其与Hyperopt-极端梯度提升模型的预测效能进行对比，结果如表1所示。文献[21]及笔者的大量试算结果表明，当训练集样本大于400时，训练集与测试集的比例划分对模型预测结果影响不大。因此，表1中不同预测模型训练集和测试集的划分原则为，在保证训练集大于400的基础上，取多次试算结果中的最优划分比例，即表1对比分析了各类不同预测模型基于大样本数据集的最优预测结果。为保证模型精准性检验结果的客观性和可比性，表1中所有预测模型的预测效能均采用R^2、RMSE和MAE三类评价指标进行评价。

由表1可见，学者们均在高性能混凝土强度预测领域内构建了精确、科学的计算模型，通过改进随机森林模型，文献[3]研究了在超参数不改变的情况下使用数据筛选和模型改进，以提升综合性能。在测试集中，R^2、RMSE和MAE三项评价指标达到0.93、4.43和3.11，但由于改进模型无法结合可解释算法，存在一定局限性。文献[21]提出的Ensemble BPNN模型采用Ensemble算法进行集成，并结合粒子群算法进行超参数优化，在训练集中，R^2、RMSE两项评价指标达到0.96和2.59。在测试集中，R^2、RMSE两项评价指标分别达到0.93和4.01，然而，模型的集成显著增加了其复杂度；文献[17]使用LSTM模型对抗压强度进行计算预测，同时，使用网格搜索进行超参数优化。在训练集中，R^2、RMSE和MAE三项评价指标为0.98、2.53和1.77，在测试集中，R^2、RMSE和MAE三项评价指标为0.93、4.42和3.15，但网格搜索对于时间与计算资源的消耗远高于Hyperopt优化算法[20]；本文

表 1 相关领域的代表性预测模型

Table 1 Representative predictive models in related fields

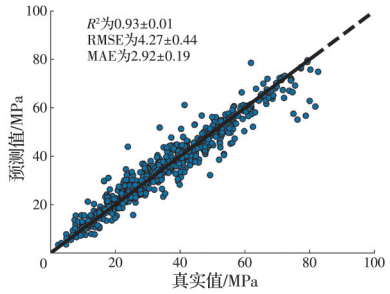
文献	模型	超参数优化方法	数据集		评价指标		
			类型	数量	R^2	RMSE	MAE
Han 等[3]	改进随机森林模型	无优化	训练集	927			
			测试集	103	0.93	4.43	3.31
Tipu 等[21]	Ensemble BPNN	粒子群算法(PSO)	训练集	824	0.96	2.59	
			测试集	206	0.93	4.01	
Cheng 等[22]	Symbiotic Operation Forest	无优化	训练集	927	0.88	5.82	4.47
			测试集	103	0.88	5.43	4.18
Jin 等[17]	LSTM 模型	网格搜索	训练集	773	0.98	2.53	1.77
			测试集	257	0.93	4.42	3.15
Naved 等[23]	Deep neural networks	Adam	训练集	721	0.96	3.84	
			测试集	309	0.89	5.36	
本文	Hyperopt-极端梯度提升	Hyperopt	训练集	721	0.99	1.54	0.93
			测试集	309	0.93	4.14	2.68

Hyperopt-极端梯度提升模型在训练集中 R^2、RMSE 和 MAE 三项评价指标为 0.99、1.49 和 0.92,在测试集中 R^2、RMSE 和 MAE 三项评价指标为 0.93、4.14 和 2.62。对比当前多项研究,在 1 030 组高性能混凝土强度-组分数据集中,构建的 Hyperopt-极端梯度提升模型 R^2 为 0.93,与改进随机森林模型、Ensemble BPNN 模型、LSTM 模型相同,优于 Symbiotic Operation Forest 模型和 Deep Neural Networks 模型。其 RMSE 数值高于改进随机森林模型、Symbiotic Operation Forest 模型、LSTM 模型及 Deep Neural Networks 模型,略低于 Ensemble BPNN 模型,但模型复杂度低于 Ensemble BPNN 模型,总体表现出综合预测性能更优、3 项评价指标数值更高、易用性更好的特点,是高性能混凝土强度-组分相关性分析可依赖的可靠模型。

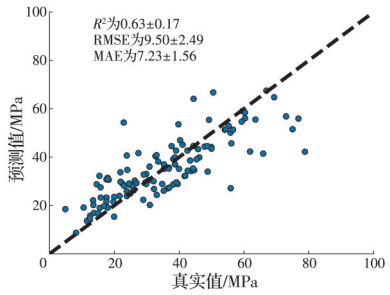
2.2 大样本数据集的必要性检验

为进一步检验大样本数据集对机器学习模型建立的必要性,从样本划分对模型性能影响的角度出发,采用十折交叉验证方法,对比分析十折交叉验证下 R^2、RMSE 和 MAE 三个指标在大小样本数据集中的变化情况。其中大样本数据集十折交叉验证结果见图 4(a),小样本数据集十折交叉验证结果见图 4(b),图中散点为验证集预测结果与真实值的偏差,左上部分为采用 R^2、RMSE 和 MAE 三个评价指标对验证集评价的结果。

对比图 4(a)、(b)可知,十折交叉验证下,大样本数据集的 R^2 值在 0.92~0.94 之间,RMSE 值在 4.26 附近变化,MAE 值在 2.90 附近变化,模型整体性能较稳健。对比小样本数据集,其 R^2 值最低仅 0.43,最高达 0.79,模型整体预测性能变异较大。可见,合理科学的大样本数据集是保障机器学习模型性能稳健性及预测结果可靠性的重要基础。



(a) 大样本数据集结果



(b) 小样本数据集结果

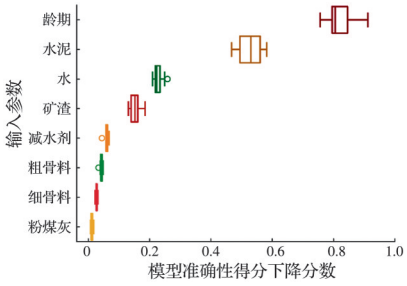
图 4 Hyperopt-极端梯度提升模型十折交叉验证结果

Fig. 4 10-fold cross validation results of Hyperopt-extreme gradient boosting model

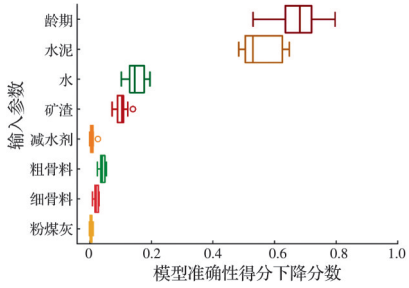
为进一步揭示 Hyperopt-极端梯度提升模型对于样本容量的依赖性机理以及样本数据集容量对于高性能混凝土强度-组分相关性关系解释的影响,采用置换重要度评价各输入参数(高性能混凝土各组分)随机变化对 Hyperopt-极端梯度提升模型预测性能的影响程度,如图 5 所示。图中横轴表示模型预测性能较原模型性能的降低幅度,纵轴为输入参数(即高性能混凝土各组分及因素)。

由图 5 可见,Hyperopt-极端梯度提升模型对大小样本数据集不同依赖性的表征,在小样本数据集

中，减水剂与粉煤灰对Hyperopt-极端梯度提升模型性能的影响基本接近于0。但这两种组分是制备高性能混凝土必备的添加材料，在小样本中Hyperopt-极端梯度提升模型对这两种组分进行了忽略，在置换重要度中，改变减水剂与粉煤灰的数值并不会对模型性能产生影响。而大样本数据集则学习了减水剂与粉煤灰的参数，在改变其数值后，模型预测性能出现一定降低。



（a）大样本数据集中Hyperopt-极端梯度提升模型置换重要度



（b）小样本数据集下Hyperopt-极端梯度提升模型置换重要度

图5　Hyperopt-极端梯度提升模型置换重要度评价

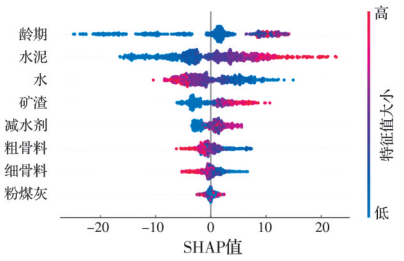
Fig. 5　Permutation importance evaluation of Hyperopt-extreme gradient boosting model

由图5还可看出，小样本数据集中Hyperopt-极端梯度提升模型忽略了减水剂与粉煤灰两类组分，同时过多学习了齡期、水泥等关键组分对高性能混凝土抗压强度的影响，而大样本数据集则有效避免了该问题。以齡期为例，大样本数据集中齡期的改变使Hyperopt-极端梯度提升模型性能下降0.67～0.79，而在小样本数据中模型性能则下降0.53～0.80。可见，大样本数据集能够使模型有效学习到高性能混凝土抗压强度的全部因素特征，为模型稳健预测提供必要保障。

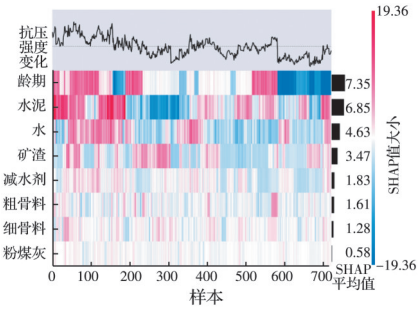
3　高性能混凝土强度-组分相关性特征的可视化分析

3.1　高性能混凝土强度-组分的全局特征解释

基于Hyperopt-极端梯度提升模型，采用SHAP可解释技术进行齡期、水泥、水、矿渣、减水剂、粗骨料、细骨料等各因素的全局特征解释，如图6所示。其中，图6(a)为SHAP全局特征图，图中每一个样本都用红色和蓝色的圆点表示，分别对应输入参数值（高性能混凝土各组分与因素）的大小，样本的水平位置反映输入参数值（高性能混凝土各组分与因素）对高性能混凝土抗压强度的影响程度。同时，如果样本位于纵轴的右侧，则表示该组分对高性能混凝土强度产生正向的影响，若位于纵轴左侧，则表示产生负向的影响。按样本将每一个组分的SHAP值进行可视化展示，即可得到图6(b)全局特征热图。图6(b)中横轴为具体的样本，纵轴为高性能混凝土各组分，图中颜色代表SHAP值的高低，颜色越偏向红色，表示SHAP值越高，即对高性能混凝土的抗压强度正向影响越强；越偏向蓝色，表示SHAP值越低，即对高性能混凝土的抗压强度负向影响越强。纵轴右侧为特征SHAP值的平均值。



（a）高性能混凝土各组分全局影响特征图



（b）高性能混凝土各组分全局特征热图

图6　SHAP全局解释图

Fig. 6　SHAP global explanation diagram

由图6可见，齡期、水泥、矿渣和减水剂对高性能混凝土抗压强度产生正向影响，水、细骨料和粗骨料对高性能混凝土抗压强度产生负向影响，而粉煤灰与高性能混凝土抗压强度之间并未表现出明显的相关性。这一结论与Pearson相关系数结果一致，同时，SHAP全局解释图所得到的强度与组分（因素）之间的关系为典型的非线性关系，弥补了相

关系数仅可描述两因素线性关系的缺憾。分析各组分(因素)的影响机理可知,水泥与水发生缓慢的水化反应,最终产生如水化硅酸钙(C-S-H)等水化产物。这些产物将骨料颗粒结合在一起并随着反应过程不断硬化,使得高性能混凝土的抗压强度更高。然而,在水化反应过程中,水含量的增加或水泥含量的减少导致水灰比增加,进而稀释水泥浆体,从而降低抗压强度,在模型中则表现为水泥产生正向影响,而水产生负向影响;掺入的矿渣与水泥的水化产物$Ca(OH)_2$发生“二次水化反应”,促使水泥进一步水化,产生更多C-S-H凝胶,改善混凝土微观结构,使混凝土强度更高,在模型中表现为矿渣掺量增加产生正向影响;通过降低混凝土的水灰比,减水剂在模型中表现为掺量增加产生正向影响;混凝土的强度取决于水泥基质的强度、骨料的强度以及水泥基质与骨料之间的界面黏结强度。粗、细骨料掺量增加会导致水泥基质与骨料之间的界面黏结强度减小,表现出宏观抗压强度降低[24],此结论与文献[25]的结论一致;粉煤灰受龄期的影响较大,早期活性反应不明显,对混凝土早期抗压强度有降低作用,但后期活性提高,能够有效提升抗压强度。同时,也需要保证适合的粉煤灰掺量,在模型中掺量与抗压强度之间尚未表现出明显的相关性。

由此可见,8种影响因素中,龄期、水泥、水和矿渣4种组分的平均SHAP值较大,分别为7.35、4.85、4.63、3.47 MPa,表明这4种组分(因素)是高性能混凝土强度设计应考虑的关键因素组分。减水剂、细骨料、粗骨料和粉煤灰4种组分的SHAP较小,分别为1.83、1.61、1.28、0.58 MPa,即这4种组分对高性能混凝土抗压强度影响较小。8种影响因素的全局敏感度排序为龄期>水泥>水>矿渣>减水剂>细骨料>粗骨料>粉煤灰。

3.2 高性能混凝土强度-组分单一因素定性定量分析

由SHAP全局相关性特征分析可知,龄期、水泥、水和矿渣是影响高性能混凝土抗压强度的决定性因素,减水剂、细骨料、粗骨料和粉煤灰影响较小。因此,选取龄期、水泥、水和矿渣4个关键因素计算单因素主效应特征依赖值,如图7所示。图中每条单独的浅蓝色曲线对应一个数据样本,其中平均值为深蓝色曲线,即特征依赖曲线。对深蓝色曲线进行非线性拟合即可得到图中特征依赖曲线的拟合曲线。

由图7可见,龄期与抗压强度的关系满足指数函数,其拟合优度R^2为0.98。龄期小于100 d时,对

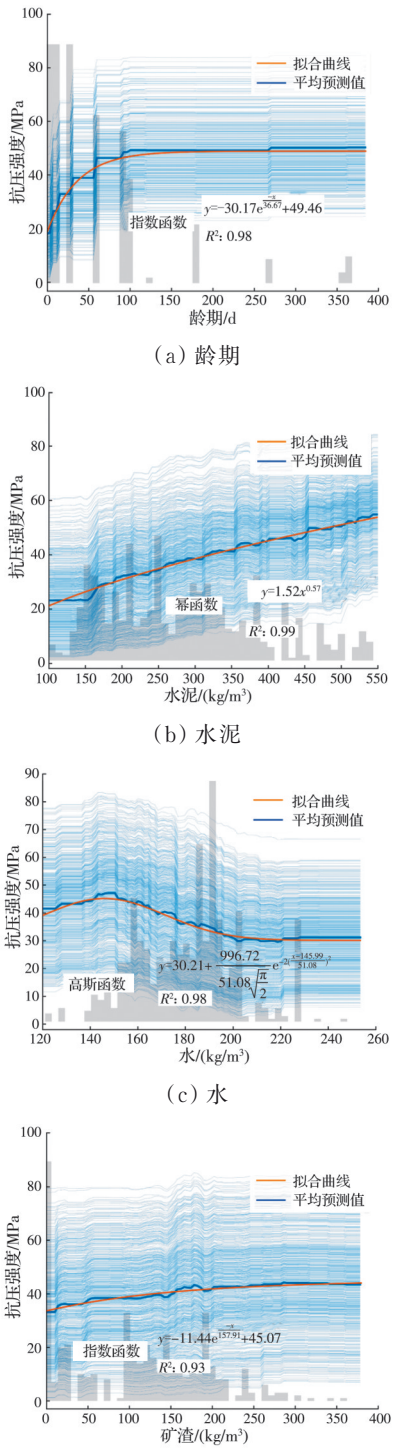


图7 单因素特征依赖图

Fig. 7 Single-factor feature dependency graph

高性能混凝土抗压强度产生明显正向影响,且在初期最明显。但当龄期超过100 d时,高性能混凝土的抗压强度达到相对稳定状态,这一结论与文献[26-27]一致。

水泥与抗压强度的关系满足幂函数,其拟合优度R^2为0.99。随着水泥掺量的增加,抗压强度持续增加。同时,矿渣与抗压强度的关系在趋势上与水泥类似,但其定量关系符合指数函数,拟合优度R^2

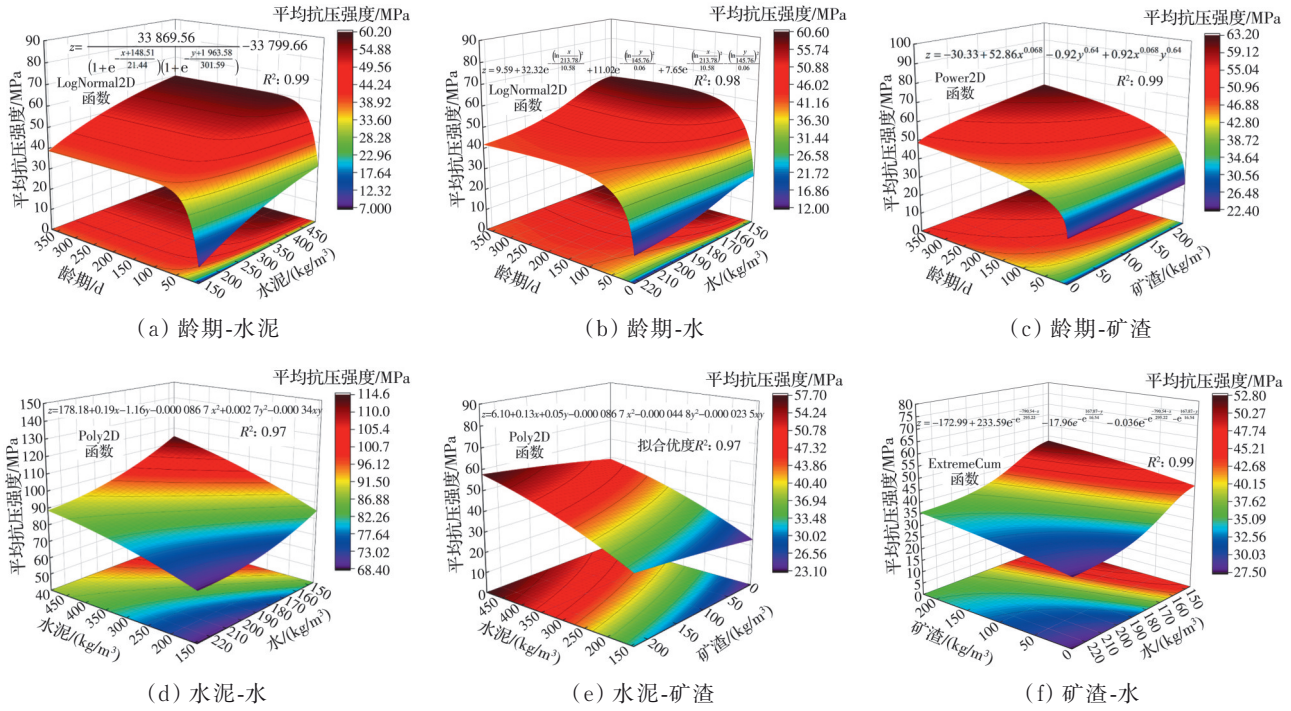
为 0.93。表现为随着矿渣掺量的增加，抗压强度明显增加，但增加幅度明显低于水泥。

水与抗压强度的关系表达式满足高斯函数，其拟合优度 R^2 为 0.98。低掺量时水对高性能混凝土的抗压强度产生一定正向影响，当掺量超过 150 kg/m^3 时，则产生负向影响，而水的掺量超过 200 kg/m^3 时，高性能混凝土的抗压强度则不再降低，这一结论与文献[15]中采用 SHAP 值得到的结果一致。

3.3　高性能混凝土强度-组分多因素耦合效应的定性定量分析

为进一步揭示龄期、水泥、水和矿渣 4 种关键因素的耦合效应，根据两两因素之间交互特征依赖值进行函数拟合，得到两因素作用下平均强度随组分（因素）变化曲面，结果如图 8 所示。图中颜色越偏向红色，表示两因素交互作用下平均抗压强度越高，而颜色越浅，表示两因素交互作用下平均抗压强度越低。

由图 8 可知，4 种因素之间均存在一定的交互作用，其中龄期与水泥、水和矿渣之间的交互作用类似，分别符合 LogisticCum 函数、LogNormal2D 函数与 Power2D 函数，拟合优度 R^2 分别为 0.99、0.98 和 0.99。当养护龄期小于 100 d 时，增加水泥掺量可减少混凝土达到预定强度所需龄期，降低水掺量与增加水泥掺量呈现出相同的效果。相比水泥和水，矿渣对高性能混凝土抗压强度的作用效果更小；而龄期超过 100 d 后，随着龄期的增长，抗压强度基本稳定，此时高性能混凝土的抗压强度取决于水泥、水与矿渣的掺量。



（a）龄期-水泥　（b）龄期-水　（c）龄期-矿渣

（d）水泥-水　（e）水泥-矿渣　（f）矿渣-水

图 8　关键因素交互作用效应

Fig. 8　Key factor interaction effects

水泥与水之间的交互作用符合 Poly2D 函数，拟合优度 R^2 为 0.97。两者交互作用可看作水灰比对高性能混凝土抗压强度的影响，在高水灰比下，高性能混凝土抗压强度呈明显降低，而在低水灰比下，高性能混凝土抗压强度呈明显升高。当水泥掺量为 500 kg/m^3，水掺量为 160 kg/m^3 时，高性能混凝土平均抗压强度达 60 MPa 以上。

水泥与矿渣之间的交互作用符合 Poly2D 函数，拟合优度 R^2 为 0.97。二者均对高性能混凝土抗压强度起正向作用，且矿渣对高性能混凝土抗压强度的作用较水泥稍弱。当水泥掺量为 500 kg/m^3，矿渣掺量为 200 kg/m^3 时，两者共同作用下高性能混凝土平均抗压强度达 55 MPa 以上。

矿渣与水之间的交互作用符合 ExtremeCum 函数，拟合优度 R^2 为 0.99。矿渣对高性能混凝土抗压强度产生正向影响，水对高性能混凝土抗压强度产生负向影响，在含水量较高时，矿渣能有效抑制高含水量对高性能混凝土抗压强度的降低作用，但当水含量较低时，矿渣对高性能混凝土抗压强度的影响不显著。

4　结论

1）机器学习模型具有显著的样本数据集依赖性，对于高维高精度求解需求的复杂问题分析，大

容量样本数据集较小样本数据集具有更高的置信度。基于大容量样本数据集构建的Hyperopt-极端梯度提升模型具有较高的精准度和良好的稳健性,测试集中R^2、RMSE、MAE三项指标分别为0.93、4.14和2.62,是揭示高性能混凝土强度-组分相关性可信任的可解释机器学习模型。

2)高性能混凝土各组分对其轴心抗压强度的敏感性排序为龄期>水泥>水>矿渣>减水剂>细骨料>粗骨料>粉煤灰,其中,龄期、水泥、水和矿渣为关键控制因素。这4种因素与高性能混凝土抗压强度之间的关系表达式分别满足指数函数、幂函数、高斯函数与指数函数,拟合优度R^2分别为0.98、0.99、0.98与0.93。

3)龄期、水泥、水、矿渣这4种关键因素的交互作用分析表明,4种因素均存在一定的交互耦合作用,龄期与水泥、水、矿渣的交互作用曲线分别符合LogisticCum、LogNormal2D和Power2D函数,水泥与水、水泥与矿渣的交互作用均符合Poly2D函数,矿渣与水的交互作用符合ExtremeCum函数。其中,龄期与水泥、龄期与水、龄期与矿渣、矿渣与水交互作用显著,在高性能混凝土配比设计中不宜忽略。

参考文献

[1] Lu C L. Compressive strength prediction of high-performance concrete with utilization of automated least square support vector regression-based algorithm[J]. Multiscale and Multidisciplinary Modeling, Experiments and Design, 2024, 7(3): 1941-1952.

[2] Yeh I C. Modeling of strength of high-performance concrete using artificial neural networks[J]. Cement and Concrete Research, 1998, 28(12): 1797-1808.

[3] Han Q H, Gui C Q, Xu J, et al. A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm[J]. Construction and Building Materials, 2019, 226: 734-742.

[4] 刘凯华, 郑佳凯, 谢维力, 等. 基于机器学习的再生混凝土配合比设计方法[J]. 湖南大学学报(自然科学版), 2023, 50(9): 88-96.
Liu K H, Zheng J K, Xie W L, et al. Mixture design method of recycled aggregate concrete based on machine learning[J]. Journal of Hunan University (Natural Sciences), 2023, 50(9): 88-96. (in Chinese)

[5] Chaabene W B, Flah M, Nehdi M L. Machine learning prediction of mechanical properties of concrete: critical review[J]. Construction and Building Materials, 2020, 260: 119889.

[6] 冯琼, 谢晓扬, 王鹏辉, 等. 基于鲸鱼优化算法-反向传播神经网络的钢筋混凝土耐久性预测[J]. 吉林大学学报(工学版), 2025, 55(7): 2276-2285.
Feng Q, Xie X Y, Wang P H, et al. Prediction of reinforced concrete durability based on whale optimization algorithm-back propagation neural network [J]. Journal of Jilin University (Engineering and Technology Edition), 2025, 55(7): 2276-2285.(in Chinese)

[7] Han F L, Lv Y, Liu Y, et al. Exploring interpretable ensemble learning to predict mechanical strength and thermal conductivity of aerogel-incorporated concrete[J]. Construction and Building Materials, 2023, 392: 131781.

[8] Kang M C, Yoo D Y, Gupta R. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete[J]. Construction and Building Materials, 2021, 266: 121117.

[9] 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.
Feng D C, Wu G. Interpretable machine learning-based modeling approach for fundamental properties of concrete structures[J]. Journal of Building Structures, 2022, 43(4): 228-238. (in Chinese)

[10] Zhao W J, Feng S Y, Liu J X, et al. An explainable intelligent algorithm for the multiple performance prediction of cement-based grouting materials[J]. Construction and Building Materials, 2023, 366: 130146.

[11] Nithurshan M, Elakneswaran Y. A systematic review and assessment of concrete strength prediction models [J]. Case Studies in Construction Materials, 2023, 18: e01830.

[12] 梅生启, 刘晓东, 王兴举, 等. 基于参数相关性分析和机器学习算法的高强混凝土徐变预测[J]. 吉林大学学报(工学版), 2025, 55(5): 1595-1603.
Mei Q S, Liu X D, Wang X J, et al. Prediction of high strength concrete creep based on parametric MIC analysis and machine learning[J]. Journal of Jilin University (Engineering and Technology Edition), 2025, 55(5): 1595-1603.(in Chinese)

[13] Raudys S J, Jain A K. Small sample size effects in statistical pattern recognition: recommendations for practitioners[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(3): 252-264.

[14] Hassani-Niaki M, Pashaian M, Ghorbanzadeh-Ahangari M. Experimental assessment of the efficiency of deep learning method in predicting the mechanical properties of polymer concretes and composites[J]. Journal of Building Engineering, 2023, 76: 107107.

[15] Rahmati M, Toufigh V. Evaluation of geopolymer concrete at high temperatures: an experimental study using machine learning[J]. Journal of Cleaner Production,

2022, 372: 133608.

[16] Lee S, Nguyen N H, Karamanli A, et al. Super learner machine-learning algorithms for compressive strength prediction of high performance concrete[J]. Structural Concrete, 2023, 24(2): 2208-2228.

[17] Jin L B, Zhang Y S, Liu P, et al. Carbon-footprint based concrete proportion design using LSTM and MOPSO algorithms[J]. Materials Today Communications, 2024, 38: 107837.

[18] Avci O, Abdeljaber O, Kiranyaz S, et al. A review of vibration-based damage detection in civil structures: from traditional methods to machine learning and deep learning applications[J]. Mechanical Systems and Signal Processing, 2021, 147: 107077.

[19] Stergiou K, Ntakolia C, Varytis P, et al. Enhancing property prediction and process optimization in building materials through machine learning: a review[J]. Computational Materials Science, 2023, 220: 112031.

[20] 吴禄源, 李建会, 马丹, 等. 基于集成学习与贝叶斯优化的岩石抗压强度预测[J]. 地球科学, 2023, 48(5): 1686-1695.
Wu L Y, Li J H, Ma D, et al. Prediction for rock compressive strength based on ensemble learning and Bayesian optimization[J]. Earth Science, 2023, 48(5): 1686-1695. (in Chinese)

[21] Kumar-Tipu R, Panchal V R, Pandya K S. An ensemble approach to improve BPNN model precision for predicting compressive strength of high-performance concrete[J]. Structures, 2022, 45: 500-508.

[22] Cheng M Y, Khitam A F K. Symbiotic operation forest (SOF): a novel approach to supervised machine learning [J]. Swarm and Evolutionary Computation, 2024, 86: 101532.

[23] Naved M, Asim M, Ahmad T. Prediction of concrete compressive strength using deep neural networks based on hyperparameter optimization[J]. Cogent Engineering, 2024, 11(1): 2297491.

[24] Poon C S, Shui Z H, Lam L, et al. Influence of moisture states of natural and recycled aggregates on the slump and compressive strength of concrete[J]. Cement and Concrete Research, 2004, 34(1): 31-36.

[25] Hariri-Ardebili M A, Mahdavi P, Pourkamali-Anaraki F. Benchmarking AutoML solutions for concrete strength prediction: reliability, uncertainty, and dilemma [J]. Construction and Building Materials, 2024, 423: 135782.

[26] 王升位, 闻一江, 洪项华, 等. 配合比和龄期对塑性混凝土强度、pH值和电导率的影响[J]. 建筑材料学报, 2022, 25(1): 97-101.
Wang S W, Wen Y J, Hong X H, et al. Effect of mix proportion and age on strength, pH value, electrical conductivity of plastic concrete[J]. Journal of Building Materials, 2022, 25(1): 97-101. (in Chinese)

[27] 尚壮壮, 刘元珍, 高宇璇, 等. 再生粗骨料取代率对再生保温混凝土360 d龄期内强度的影响[J]. 西安建筑科技大学学报(自然科学版), 2021, 53(2): 217-222.
Shang Z Z, Liu Y Z, Gao Y X, et al. Study on the effect of recycled coarse aggregate replacement rate on strength of recycled aggregate thermal insulation concrete within 360 days[J]. Journal of Xi'an University of Architecture & Technology (Natural Science Edition), 2021, 53(2): 217-222. (in Chinese)

(编辑　王秀玲)