

基于 Boosting、Bagging、Stacking 与 Voting 四大集成学习模型的粉质砂土冻土抗剪强度预测研究

赵 娜, 王世旺, 王来贵, 米亚宽, 丁明瑞

(辽宁工程技术大学 力学与工程学院, 辽宁 阜新 123000)

摘 要: 为解决传统冻土抗剪强度测试耗时费力、难以满足工程动态预测需求的问题,引入 Boosting 模型、Bagging 模型、Stacking 模型和 Voting 模型对基模型进行集成,探究不同集成策略对于模型性能的影响,提出一种基于简单参数预测冻土抗剪强度的方法。首先以内蒙古地区的粉质砂土为对象,通过室内试验建立以冻结时间、冻结温度、含水率、轴向压力和抗剪强度为特征的 768 组冻土数据集,然后对包含单一模型、集成 Bagging 模型、集成 Boosting 模型的 27 种基模型进行精度比较,选取 9 种精度最高的模型作为 Stacking 模型和 Voting 模型的基础模型。最后,利用决定系数(R^2)、平均绝对误差(MAE)和均方误差(MSE)对 Stacking 和 Voting 模型的学习能力和泛化能力进行综合评估和对比分析,并结合特征重要性结果对各特征的影响程度进行分析。结果表明,Stacking 模型性能最优($R^2=0.998$, MAE=1.7091, MSE=6.4233)。Voting 模型的性能略低于 Stacking 模型($R^2=0.9979$, MAE=1.6864, MSE=6.7903)。两个模型的特征重要性排序为:含水率>冻结时间>冻结温度>轴压。合理的集成学习策略可在较少输入参数条件下实现高精度预测,为寒区工程中冻土稳定性分析与设计提供了一种高效、经济且具有可解释性的新方法。

关键词: 粉质砂土; 抗剪强度; 集成学习; 特征重要性; 可解释性分析

中图分类号: TU43; TP181 **文献标志码:** A **文章编号:** 1000-0240(2026)02-0606-10

0 引言

冻土作为寒区工程建设中的关键地质体,其抗剪强度是衡量冻土稳定性的核心指标,直接影响路基、边坡等工程结构的安全性能。其中,季节等^[1]、唐丽云等^[2]分别将含水率、冻结温度、含石率等作为影响抗剪强度的变量进行研究。杨尚礼等^[3]通过不同温度和细粒含量条件下的试验,系统分析了冻土的力学特性。同时,地热温度和含水率也会对边坡冻土的力学特性产生较大影响^[4]。但传统冻土抗剪强度测试依赖室内试验或现场监测,不仅耗时费力、成本较高,且难以快速响应工程实践中对冻土力学性能的动态预测需求,因此亟需一种高效、精

准的预测方法。

随着机器学习技术的发展,其在处理复杂非线性关系、挖掘数据潜在规律方面的优势逐渐显现,为冻土力学性能预测提供了新的技术路径。例如,机器学习与 SBAS-InSAR 联合对青藏高原地区冻土形变进行监测^[5]。此外,学者们利用机器学习对冻结黏土的应力应变和破坏强度做出预测^[6]。王晨光等^[7]提出基于高度集成的 XGBoost 算法来预测不同粒径分布冻土强度的方法。Zhao 等^[8]利用机器学习选取 2 个宏观参数、38 个细观参数和 19 个微观参数对盐渍冻土进行无侧限抗压强度预测。王冰泉等^[9]利用机器学习对第三级季节冻土冻结深度进行了

收稿日期: 2025-08-31; 修订日期: 2025-11-03

基金项目: 国家重点研发计划项目(2017YFC1503101); 辽宁省重点科技创新基地联合开放基金项目(2020-KF-13-06); 辽宁省教育厅项目(LJ2020JCL013); 辽宁省教育厅基本科研项目(LJ212510147044)资助

作者简介: 赵娜,副教授,主要从事岩体力学及岩层控制研究. E-mail: zhaona24@163.com

通信作者: 王世旺,硕士研究生,主要从事神经网络预测冻土力学参数研究. E-mail: 2021856590@qq.com

预测。Meng 等^[10]利用机器学习对冻结岩土の力学特性进行了预测。Esmacili-Falak 等^[11]通过机器学习对冻土の轴向抗压强度和弹性模量进行了预测。Li 等^[12]采用 ANN 模型对冻土应力应变进行预测和验证。刘凤云等^[13]利用六种机器学习模型,对不同温度、含水率和干密度条件下冻土の导热系数进行了预测。

然而,传统机器学习方法在量化冻土力学参数对抗剪强度影响方面仍存在不足。为此,引入基于机器学习模型的特征可解释性方法,可通过定量评估模型,分析输入变量对输出结果的影响,为冻土本构理论的发展以及工程实践提供有效的量化依据^[14-17]。

因此,本文旨在弥补传统试验方法效率低下の不足,提升冻土抗剪强度获取的效率并改进其预测方法。同时,将 Boosting、Bagging、Stacking 和 Voting 四种集成策略应用于冻土抗剪强度预测模型,以克服现有研究中模型集成策略单一の局限。此外,借助机器学习模型的特征可解释性方法,提高模型在冻土抗剪强度预测过程中的可解释性。

综上,本研究以内蒙古自治区际通道工程新林

北至扎兰屯段边坡の粉质砂土为研究对象,通过室内直剪试验获取 768 组数据,建立了包含冻结时间、轴向压力、含水率、冻结温度和抗剪强度の冻土数据集。在此基础上,对比 Boosting 和 Bagging 集成策略对于单一模型性能的影响,即单一模型为 Boosting 和 Bagging 集成使用的模型。分析 Stacking 和 Voting 集成对异质基础模型的预测性能影响,即基础模型为 Stacking 和 Voting 集成使用的模型,对比不同集成策略对冻土抗剪强度预测精度的影响,通过特征重要性方法识别关键影响因素。本研究为冻土抗剪强度预测提供了一种高效、经济的新方法,同时为寒区工程中冻土稳定性评估与设计提供了科学参考。

1 冻土数据集的建立与分析

具有良好泛化能力的集成学习模型依赖于大量且高质量的训练数据^[18-19],通过梳理国内外关于冻结作用对土体力学性质影响的研究成果可知,不同初始条件均会对冻土抗剪强度产生显著影响。基于此,本文选用冻结时间、轴向压力、含水率和冻结温度作为影响因素,实验整体流程如图 1 所示。

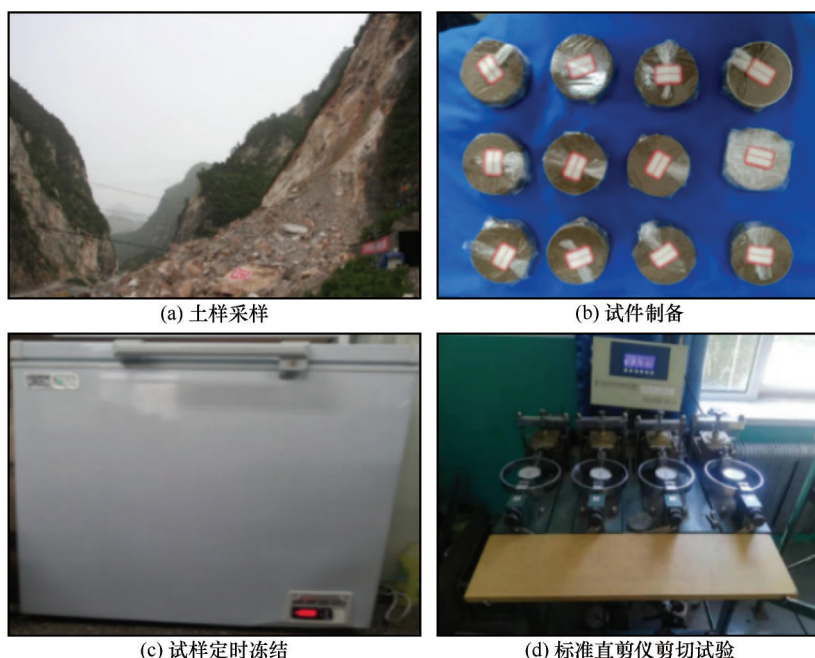


图 1 冻土抗剪强度试验流程示意图

Fig. 1 Schematic diagram of the shear strength test procedure for frozen soil: (a) soil sampling; (b) specimen preparation; (c) timed freezing of specimens; (d) shear test using a standard direct shear apparatus

试验土样取自内蒙古自治区际通道新林北至扎兰屯段边坡,经烘干后测得干密度为 $1.72 \text{ g}\cdot\text{cm}^{-3}$ 。根据试验设计,配置不同含水率的试样;并依据《岩

土工程勘察规范》(GB 50021—2001)制备试件。试件置于冷冻箱中冷冻至设定时间后,采用标准直剪仪在不同轴向压力下进行剪切试验,共获得 768 组

有效数据。为初步分析各特征与抗剪强度之间的线性和非线性关系,计算了其距离相关系数,结果如图2所示。

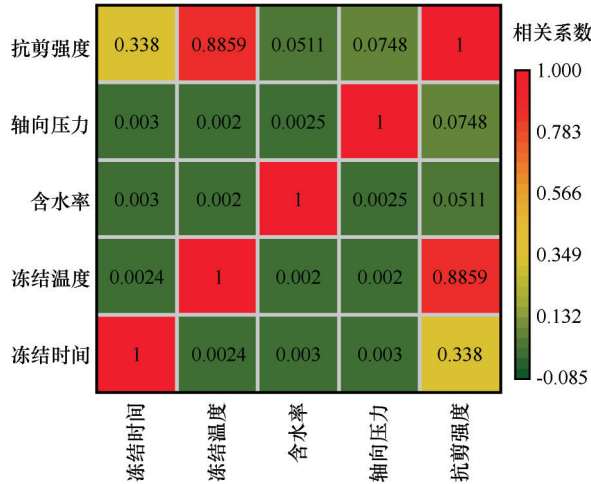


图2 距离相关系数图

Fig. 2 Distance correlation coefficient graph

图2为距离相关系数图,可以综合反映特征之间的线性和非线性关系,根据距离相关系数对建立的冻土数据集分析,冻结温度与抗剪强度的相关性系数为0.8859,表明二者表现出较高的依赖关系。轴向压力与抗剪强度的相关性系数为0.0748,说明两者之间的依赖关系较弱。根据特征间距离相关系数的分析结果,初步探明所选特征与冻土抗剪强度存在一定关联,从而避免了后续建模中引入无关变量,减少了模型的冗余计算。

表1展示了冻土数据集的数据分布情况。通过分析数据的离散程度与异常点分布,可评价数据集

表1 冻土数据集特征分布统计表

Table 1 Statistical distribution of frozen soil dataset features

特征名称	最小值	最大值	均值	下四分位	上四分位	异常值检测结果
冻结温度/℃	-40	-10	-24.9	-30	-15	无异常值
冻结时间/h	2	24	13.0	8	19	无异常值
含水率/%	7.36	36.85	22.1	17.1	31.9	无异常值
轴向压力/kPa	50	400	189.6	100	300	无异常值
抗剪强度/kPa	9.6	232.2	99.6	52.365	130.395	无异常值

注: 异常值采用1.5倍四分位距(1.5 IQR)原则确定。所有特征数据均分布在1.5 IQR范围内,未发现异常值,因此数据可直接用于模型训练。

Note: Outlier detection was performed using the 1.5 IQR rule. All features were within the 1.5 IQR range, indicating no outliers; thus, the data were directly employed for model training.

的整体质量,为下一步数据的归一化提供依据,减少极端值对模型拟合精度的干扰。其中,异常值识别采用1.5倍四分位距(1.5 IQR)原则确定界限。每个特征的IQR值为上四分位值减去下四分位值。五个特征的数据均分布在1.5倍四分位距(1.5 IQR)内,数据集中未发现异常值,因此未进行异常值剔除处理,数据可直接用于模型训练。

2 Boosting与Bagging模型原理

2.1 Boosting模型

Boosting是一种通过迭代训练多个弱学习器并将其加权组合为强学习器的集成方法^[20]。其训练过程为串行方式,每一轮重点学习前一轮预测错误的样本,通过调整样本权重或模型参数,逐步降低模型误差。具体流程如下:

(1)对冻土数据集里的每个样本赋予相同的权重值:

$$w_i = \frac{1}{n} \quad (i = 1, 2, \dots, n) \quad (1)$$

式中: w_i 为第1轮第*i*个样本的权重; n 为冻土数据集总样本数

(2)依据每个样本的权重挑选部分数据,构建冻土数据集D(1)使用带权重的样本训练弱学习器 $h_t(x_i)$,使其在加权样本上的错误率最低,弱学习器在当前权重的错误率:

$$\delta_t = \frac{\sum_{i=1}^n w_i \cdot I(h_t(x_i) \neq y_i)}{\sum_{i=1}^n w_i} \quad (2)$$

式中: δ_t 为模型的犯错率; I 为模型的正确性,分类正确则为0,分类错误为1。

(3)根据错误率计算弱学习器在最终集成中的权重:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \delta_t}{\delta_t} \right) \quad (3)$$

式中: α_t 为第*t*轮弱学习器在最终强学习器中的集成权重; δ_t 为第*t*轮弱学习器的错误率。

对分类错误的样本增加权重,对正确的样本降低权重,更新公式为:

$$w_i^{(t+1)} = \frac{w_i^{(t)} \cdot \exp(-\alpha_t \cdot y_i \cdot h_t(x_i))}{\sum_{j=1}^n w_j^{(t)} \cdot \exp(-\alpha_t \cdot y_j \cdot h_t(x_j))} \quad (4)$$

该操作使后续弱学习器更关注之前被误分类的样本。

(4)迭代完成后,将所有弱学习器按权重组合

成最终的强学习器:

$$H(x) = \text{sign}\left(\sum_{i=1}^T \alpha_i \cdot h_i(x)\right) \quad (5)$$

2.2 Bagging 模型

Bagging 模型通过并行训练的方式集成多个同质模型,其数据划分采用有放回抽样,通过构建多个独立模型并聚合结果提升预测稳定性。Bagging 利用并行训练和结果聚合来减少单一模型的过拟合风险。Bagging 模型流程如下所示:

(1)从冻土数据集中通过有放回随机采样生成多个训练子集,每次从冻土数据集中随机抽取一个样本,放回后继续抽取,重复 n 次,生成新的冻土数据集 D_i ,每个子集称为自助样本,每一个样本被选取的概率是相同的。

(2)用 n 个基础模型分别对 n 个冻土数据集 D_1 、 D_2 、 $\dots$ 、 D_n 进行训练学习,在理想情况下,最终得到 n 个不同的模型 e_1 、 e_2 、 $\dots$ 、 e_n 。

(3)对 n 个模型得到的结果,再利用平均法得到最终结果:

$$H(x) = \frac{1}{T} \sum_{i=1}^T h_i(x) \quad (6)$$

式中: $H(x)$ 为 Bagging 集成模型的预测函数; T 为 Bagging 的基模型总数; $h_i(x)$ 为第 i 个基模型对样本 x 的抗剪强度预测值。

3 Stacking 与 Voting 模型原理

3.1 Stacking 模型

Stacking 是通过分层训练组合多个基模型的集成学习方法。先用基础模型对原始数据生成预测结果,再将这些预测结果作为特征,训练一个元模型来学习如何最优组合基础模型的输出。Stacking 流程如下所示:

(1)将冻土训练集分为 K 折。使用其中 $K-1$ 折作为训练集,1 折作为验证集,确保每一折中数据数量相同。

(2)使用部分冻土训练数据训练出模型 i ($i=1, 2, 3, 4, 5$),用模型 i 预测验证数据 i ,得到 1 折的一维预测序列。对测试集同样使用模型 i 生成一维测试序列。重复该过程,直至遍历所有折次。最终得到 5 个预测序列和 5 个测试序列,将 5 个预测序列合并为一维新训练数据,同时对 5 个测试序列取平均,得到基于测试集的预测序列。

(3)利用步骤(2)得到的新训练数据和新测试数据训练元模型,得到最终预测结果。

3.2 Voting 模型

Voting 模型整体流程如下:

(1)选择不同类型的基模型预测冻土强度,每一个基模型得到一个预测结果。

(2)根据投票规则整合预测结果,第一种硬投票,根据每一个基模型得到的预测结果,依据少数服从多数的规则,将占比最大的结果作为最终结果。第二种软投票,计算每一个基模型的概率均值,均值高的为最终的预测结果。第三种加权投票,为每一个基础模型分配权重,提升模型的预测性能。

在集成学习中,Boosting 与 Bagging 通常使用同质模型进行训练,而 Stacking 与 Voting 则通常采用异质模型。同质模型指使用相同类型的模型构建模型,异质模型指使用不同类型的模型构建模型。通过 Stacking 和 Voting 的异质模型组合,可以在捕捉不同特征模式的同时降低偏差与方差,从而得到最优模型组合。

4 模型构建

4.1 模型框架设计

研究选取九种异质模型作为单一模型,同时选取 Boosting、Bagging、Stacking 和 Voting 四种不同集成策略模型用于集成模型。首先利用 Boosting 模型和 Bagging 模型对九种单一模型进行集成,然后对得出的 27 种模型进行可靠性对比,选取九种性能最高的基础模型作为 Stacking 模型和 Voting 模型的基础进行集成训练,最终对得到的两种集成模型进行性能评价。

4.2 超参数优化

按照前述整体流程和思路进行参数优化,考虑到设备算力的限制,模型采用网格搜索的方法对模型进行参数优化。对于不同的模型优先选取对模型效果影响显著的参数进行调整,对于寻求模型网格搜索的参数组合,模型选用两阶段搜索的方法,第一阶段用较宽的范围确定最优参数的大致区间,第二阶段在粗搜得到的区间内缩小范围,用更小的步长细化参数,模型具体参数如表 2 所示。

4.3 评价指标

为评估所建立模型的预测性能,采用决定系数 (R^2)、平均绝对误差 (MAE) 和均方误差 (MSE) 作为衡量模型总体预测能力的评价指标。相关计算公式如表 3 所示。 R^2 决定系数用于反映模型的拟合

表2 各模型参数取值表
Table 2 Values of each model parameter

模型名称	参数名称	参数取值
Linear regression	alpha	1.0
Ridge	alpha	1.0
Lasso	alpha	1.0
Random_Forest	n_estimators	100
SVR	gamma	1.0
Decision_Tree	max_depth	None
	alpha	0.9
Gradient_Boosting	learning_rate	0.1
	max_depth	3
	n_estimators	100
Xgboost	learning_rate	0.1
	n_estimators	100
	learning_rate	0.1
Lightgbm	max_depth	-1
	n_estimators	100

程度、可靠性及其对数据规律的解释能力,数据越接近 1.0,代表模型的拟合程度越好。平均绝对误差计算预测值与真实值之间绝对误差的平均值,能反映多数普通样本的误差,用于直观判断模型的精度,但对极端误差的反映不足,数值越小代表预测值和真实值之间的误差值越小。均方误差计算预测值与真实值差异平方的平均值,异常值的偏差平方会被急剧放大,根据数值大小可以反映模型在极端误差下的稳健性。评价指标和计算公式如表 3。

表3 评价指标及计算公式
Table 3 Evaluation indicators and calculation formulas

评价指标	计算公式	评判指标
决定系数	$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	R^2 越接近 1,模型解释能力与拟合效果越强
平均绝对误差	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $	MAE 越小,模型对多数样本的预测偏差越低,整体精度越高
均方误差	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$	MSE 越小,模型在极端误差下的稳健性与预测精度越高

5 Bagging 与 Boosting 模型预测结果评估

Stacking 模型通过多个不同类型的基模型对冻土数据进行训练,训练完成后,将基模型的预测结果作为新的特征数据,用于训练元模型,从而进一

步提升模型的预测性能。Voting 模型同样通过多个基模型对冻土数据进行训练,并采用投票或加权投票的方式融合各基模型的预测结果,最终获得综合预测输出。根据 Stacking 和 Voting 模型原理可知,两种模型均属于二次训练机制,对基础模型的预测性能依赖较强,基模型效果较差,则第二阶段的训练结果将明显下降。同时,各基模型的预测结果应尽可能保持差异性,否则第一次与第二次训练的结果将过于相似,难以体现集成效果。因此,本文选取多种异质单一模型作为 Bagging 与 Boosting 模型的基学习器。通过将得到的 9 种异质基础模型进行组合,可增强 Stacking 和 Voting 模型的多样性,使各模型相互补充、取长补短。

为弥补单一模型的缺陷,增强模型的多样性。选用线性回归模型 (LinearRegression、Ridge、Lasso)、树模型 (DecisionTree、RandomForest、XGBoost、GradientBoosting、Lightgbm) 以及支持向量机模型 (SVR),九种模型作为 Bagging 和 Boosting 的训练模型。通过选用九种异质模型可以降低过拟合的风险,并覆盖更广泛的假设空间。利用 Bagging 模型的并行训练特性与 Boosting 模型的串行加权机制,可在保证模型多样性的同时兼顾稳定性与精度。通过对单一模型、Bagging 集成模型及 Boosting 集成模型的可靠性进行对比,选取其中精度最高的九个模型作为 Stacking 和 Voting 的基础模型。模型精确度对比图如图 3 所示。

图 3 展示了九种单一基础模型与其对应的 Bagging 集成和 Boosting 集成模型的性能对比情况。从图中三个评价指标可以看出,经 Bagging 与 Boosting 集成后模型的预测精度整体上均优于单一模型。综合三项指标的对比结果,最终选取 Boost-DecisionTree、Boost-SVR、Bagging-Lasso、Bagging-Ridge、Boost-RandomForest、Bagging-Lightgbm、Boost-GradientBoosting、Boost-Xgboost 和 Bagging-LinearRegression 模型作为最终的 Stacking 集成和 Voting 集成的基础模型。

6 Stacking 与 Voting 模型预测结果评估

6.1 模型预测结果对比

基于前述九种基础模型,分别构建了 Stacking 和 Voting 集成模型,为进一步分析两种模型的预测性能,本文绘制了模型训练集和验证集上的真实值和预测值对比图(图 4~5),其中,训练集对比图用于反映模型对冻土数据的学习能力,验证集对比图用

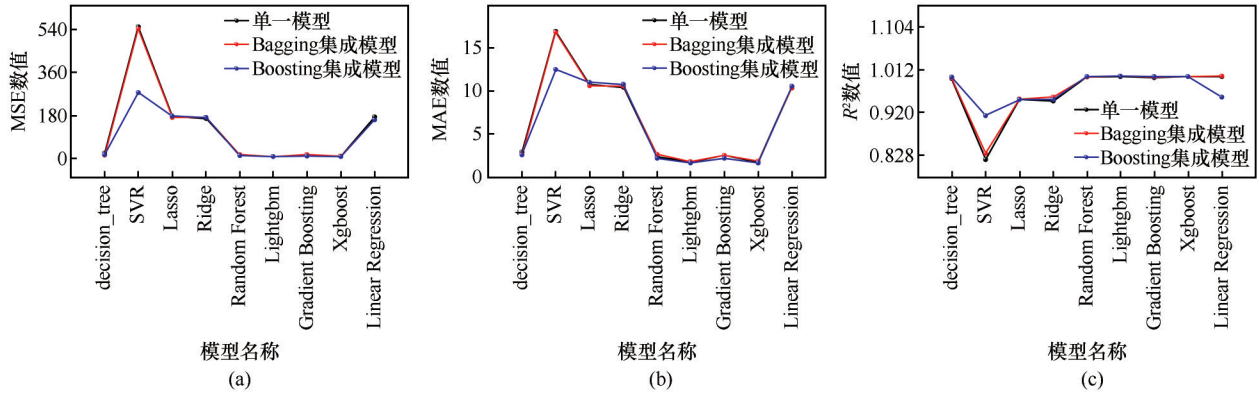


图3 Bagging集成、Boosting集成与单一模型性能对比图

Fig. 3 Comparison of Bagging ensemble, Boosting ensemble, and Single models in prediction performance

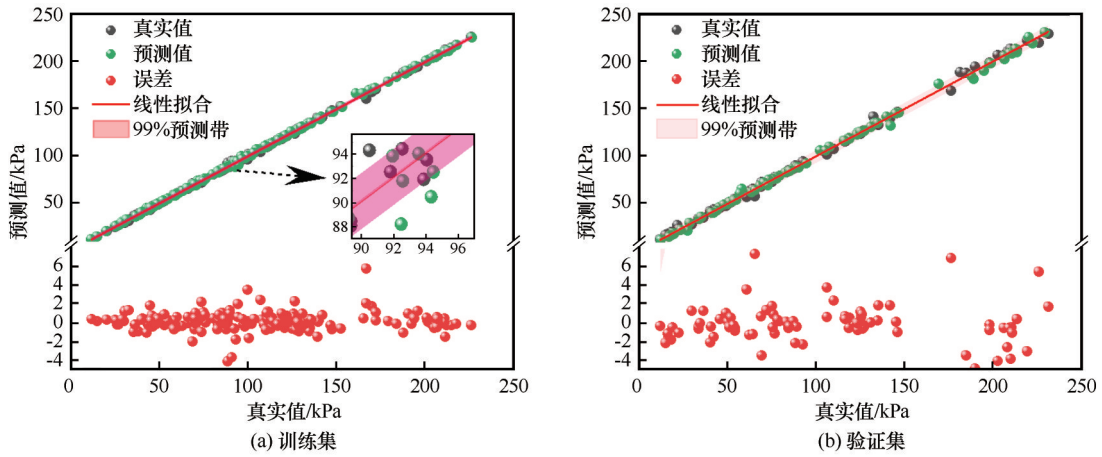


图4 Stacking模型真实值与预测值对比图

Fig. 4 Comparison between the real and predicted values of the Stacking model: (a) training set; (b) validation set

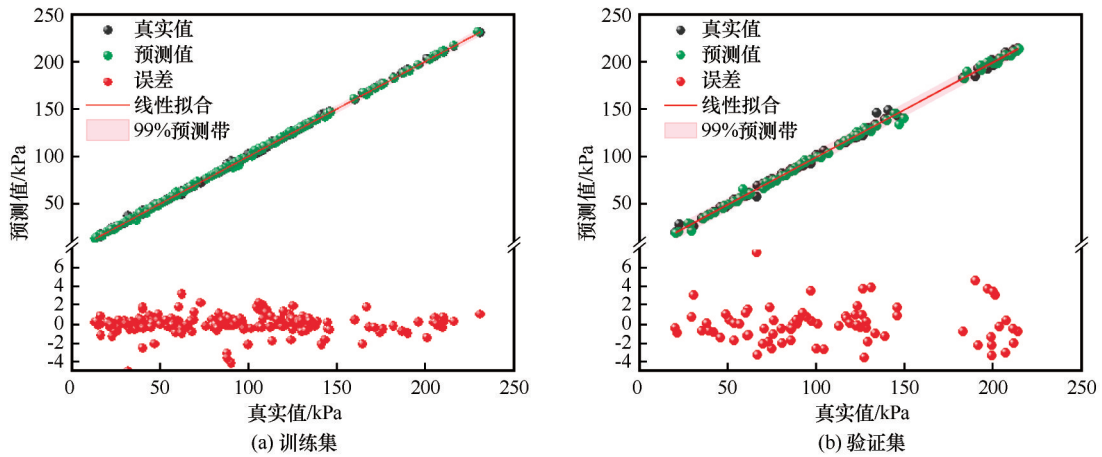


图5 Voting模型真实值与预测值对比图

Fig. 5 Comparison between the real and predicted values of the Voting model: (a) training set; (b) validation set

于评估模型对未知冻土数据的泛化能力。图中黑色点表示真实值,绿色点表示预测值,红色点表示两者之间的误差。

图4(a)为Stacking模型在训练集上的真实值与预测值对比图。从图中可以看出,绿色点与黑色点

的重合度较高,表明模型对冻土数据的学习能力较强。其99%预测区间较窄,说明模型的整体误差较小。图4(b)为Stacking模型在验证集上的真实值与预测值对比图。图中绿色点与黑色点的重合度仍然较高,误差点分布范围无明显扩大,且99%预测

区间宽度与训练集相近,表明模型对冻土数据具有较强的泛化能力。

图 5(a)为 Voting 模型在训练集上的真实值与预测值对比图。从图中可以看出,绿色点与黑色点之间存在一定误差,误差点的离散度较图 4(Stacking 模型)更大。同时,99% 的预测区间略宽于图 4 中的训练集,表明 Voting 模型的学习精度略低于 Stacking 模型。图 5(b)为 Voting 模型验证集的真实值和预测值对比图,采用与图 4(b)相同的方法进行对比分析,综合对比结果可知,Voting 模型的总体泛化能力低于 Stacking 模型。

6.2 模型精度指标分析

在对两种模型的预测结果进行量化比较时,可以发现,Stacking 模型的决定系数(R^2)为 0.9980,略高于 Voting 模型的 0.9979,表明其整体拟合效果更优。在误差指标方面,Stacking 模型的均方误差(MSE)为 6.4233,低于 Voting 模型的 6.7903,说明 Stacking 模型在极端误差下的稳定性更好;而平均绝对误差(MAE)为 1.7091,略高于 Voting 模型的 1.6864,表明 Voting 模型在多数样本的平均偏差上

稍占优势。总体而言,Stacking 模型的拟合精度和泛化能力均优于 Voting 模型,预测结果更为稳健可靠,能够更有效地捕捉冻土抗剪强度数据中的复杂非线性关系。

6.3 模型可解释性分析

模型的 SHAP 值和特征重要性是模型的核心解释工具,可用于揭示模型内部的决策机制,从而解决黑盒问题。下图(图 6~8)是 Stacking 模型和 Voting 模型的 SHAP 值分布图和特征重要性对比图。

图 6 和图 7 分别为 Stacking 与 Voting 模型的 SHAP Value 分析图。SHAP 值能够定量反映各特征对预测结果的影响方向与贡献强度。结果显示,含水率是影响冻土抗剪强度的最主要因素,且呈负相关关系,含水率增加会导致抗剪强度下降。冻结时间的贡献位居第二,表现为正相关,即冻结时间越长,抗剪强度越高。冻结温度影响中等,低温条件有助于提高强度。轴向压力作用最小。相比之下,Stacking 模型的 SHAP 值分布更分散、层次更清晰,表明在捕捉非线性关系和解释模型行为方面优于 Voting 模型。

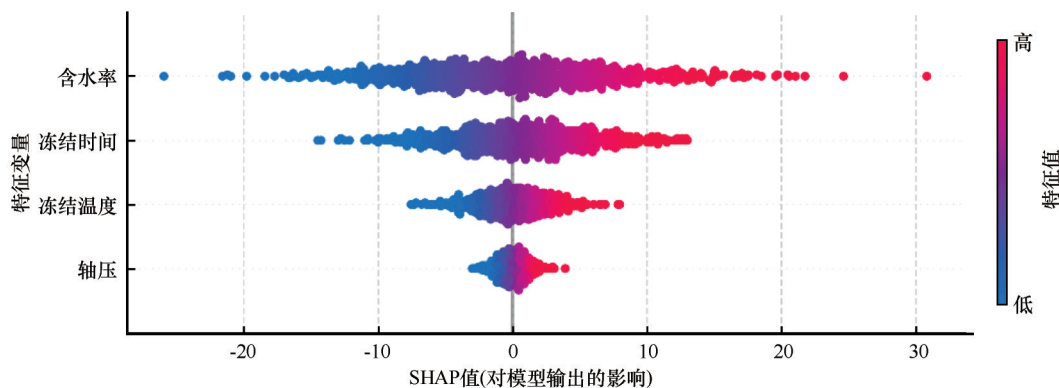


图 6 Stacking 模型 SHAP Value 分析图

Fig. 6 SHAP Value analysis diagram of the Stacking model

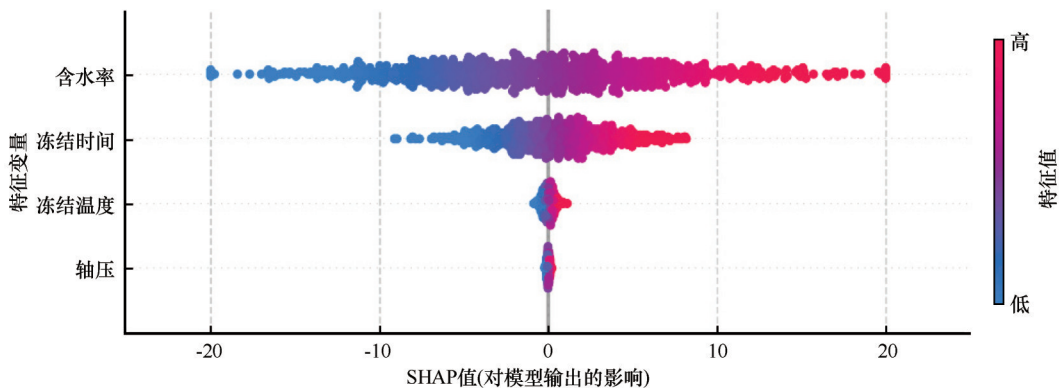


图 7 Voting 模型 SHAP Value 分析图

Fig. 7 SHAP Value analysis diagram of the Voting model

图 8(a)、8(b)分别为 Stacking 模型和 Voting 模型的特征重要性排序图,两者均反映冻土数据集中各特征对冻土抗剪强度预测的相对重要程度,根据两图可以看出,两个模型中,含水率的相对重要性

程度均是最高且远高于其他特征,其次冻结时间作为次要核心因素,这表明含水率和冻结时间在两种集成策略下均被识别成影响抗剪强度的主要驱动因素。

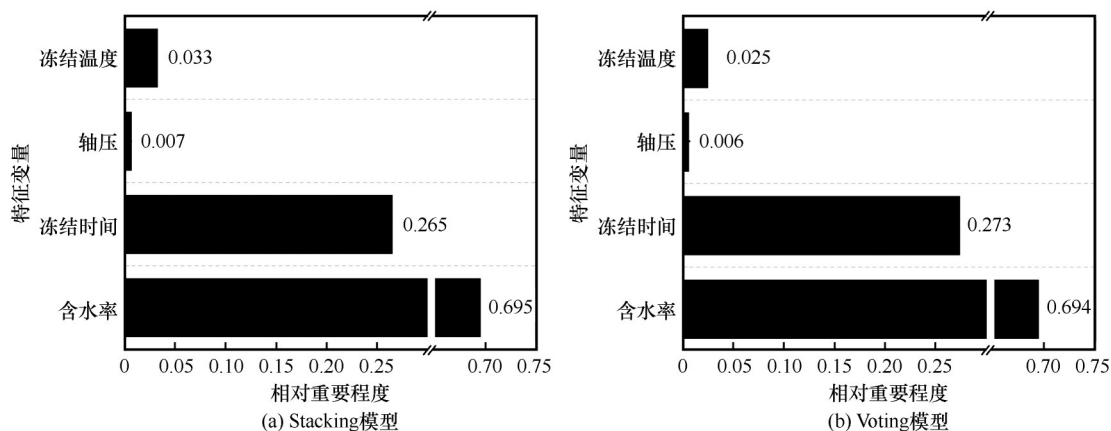


图 8 模型特征重要性分析

Fig. 8 Analysis of the importance of model features: (a) Stacking model; (b) Voting model

7 讨论

本研究通过引入 Boosting、Bagging、Stacking 和 Voting 四种集成学习策略,系统比较了不同集成方法在冻土抗剪强度预测中的表现。结果表明,Stacking 模型的拟合精度与泛化能力均优于 Voting 模型,其决定系数最高 ($R^2=0.998$),误差指标最低,说明分层集成能够更有效地融合异质模型的优势,通过误差互补与特征信息协同提升预测性能。

特征重要性分析结果显示,含水率与冻结时间是影响冻土抗剪强度的主要因素,冻结温度次之,轴向压力影响相对较小。这与以往基于树模型的冻土力学参数预测结论一致^[20-21],进一步验证了集成学习在揭示冻土非线性关系方面的有效性。已有研究表明,增加输入参数的类型与层次有助于提升模型精度^[8],而本研究在仅采用少量宏观参数的情况下,通过 Stacking 集成仍实现了高精度预测,说明合理的模型集成策略可在较低参数维度下实现性能优化。

同时,相关研究表明在冻土力学特性预测中引入时间序列信息有助于描述应力-应变的动态变化规律^[22-23]。本研究聚焦于静态抗剪强度的建模,未来可将时序建模与集成学习相结合,构建可反映冻融循环和加载历时影响的动态预测模型,以进一步提升模型的时空适应性与工程实用性。

8 结论

(1)通过现场取土和室内试验的方法,建立以冻结温度、冻结时间、轴压、含水率和抗剪强度为特征的数据集。确保了数据集的真实性与可靠性。利用距离相关系数对数据集各个影响因素进行线性和非线性依赖分析,其中冻结温度对于抗剪强度具有较强的关联性,轴压对于抗剪强度呈现出较弱的关联性。而各个特征之间相互作用存在复杂的依赖关系。

(2)通过实验结果可见,Bagging 和 Boosting 集成后的模型精度普遍高于单一模型。采用多评价指标对模型进行性能评估,结果表明 Stacking 模型在预测任务中性能优于 Voting 模型,Stacking 模型的均方误差低于 Voting 模型,决定系数更接近于 1,模型在验证集上的泛化能力表现优秀,结果表明,分层训练机制能够更有效地整合异质基模型的优势,实现模型的多样性与高精度预测。同时,通过模型特征重要性分析各特征对冻土抗剪强度预测泛化能力的影响程度。Stacking 模型和 Voting 模型的重要性排序为:含水率>冻结时间>冻结温度>轴压。

(3)将机器学习与冻土工程相融合,依托冻结时间、轴向压力、含水率、冻结温度等基础力学参数,实现了对冻土抗剪强度的智能预测。与传统试验方法相比,该方法显著节省了人力、物力与时间成本,为冻土抗剪强度预测开辟了一条切实可行的

新路径。此外,通过特征重要性分析明确了影响冻土抗剪强度的关键因素,相关结果亦能为冻土力学性能的深入探究提供参考依据。

本文尚存在一定局限性。数据地域性与土体类型单一,数据集仅来源于内蒙古自治区际通道新林北至扎兰屯段的粉质砂土,并未考虑多种土体类型。特征维度不足,现有模型仅纳入4个影响因素,未考虑土体干密度、孔隙率等关键指标。上述局限会在后续研究中进一步探讨。

参考文献(References):

- [1] Ji Jie, Wang Xiaoguang, Xiao Junxiao, et al. Simulation of icy Lunar regolith and experiment on its shear strength [J]. *Journal of Deep Space Exploration*, 2023, 10(2): 199-210. [季节, 王晓光, 肖俊孝, 等. 月球冻土模拟与剪切强度试验分析[J]. *深空探测学报*, 2023, 10(2): 199-210.]
- [2] Tang Liyun, Wang Xin, Qiu Peiyong, et al. Study on shear performance of soil-rock mixture at the freezing-thawing interface in permafrost regions [J]. *Rock and Soil Mechanics*, 2020, 41(10): 3225-3235. [唐丽云, 王鑫, 邱培勇, 等. 冻土区土石混合体冻融交界面剪切性能研究[J]. *岩土力学*, 2020, 41(10): 3225-3235.]
- [3] Yang Shangli, Cui Honghuan, Hu Junhui. Effect of fine particle content on mechanical properties of frozen silty clay [J]. *Journal of Glaciology and Geocryology*, 2024, 46(5): 1612-1622. [杨尚礼, 崔宏环, 胡峻晖. 细粒含量对冻结粉质黏土力学特性的影响[J]. *冰川冻土*, 2024, 46(5): 1612-1622.]
- [4] Li Yan, Jin Huijun, Wen Zhi, et al. Stability of permafrost slopes: a review [J]. *Journal of Glaciology and Geocryology*, 2022, 44(1): 203-216. [李艳, 金会军, 温智, 等. 多年冻土区斜坡稳定性研究综述[J]. *冰川冻土*, 2022, 44(1): 203-216.]
- [5] Yang Yang, Jia Hongguo, Bai Zhenghang, et al. Integrated monitoring of permafrost deformation in the Qinghai-Xizang Plateau using machine learning and SBAS-InSAR [J]. *Bulletin of Surveying and Mapping*, 2025(4): 45-50. [杨洋, 贾洪果, 柏铮航, 等. 联合机器学习与SBAS-InSAR的青藏高原冻土形变监测[J]. *测绘通报*, 2025(4): 45-50.]
- [6] Wang Weijun, Wang Dayan. Prediction and sensitivity analysis of mechanical parameters of frozen clay based on machine learning [J]. *Journal of Glaciology and Geocryology*, 2025, 47(3): 770-780. [王伟军, 王大雁. 基于机器学习的冻结黏土力学参数预测与敏感性分析[J]. *冰川冻土*, 2025, 47(3): 770-780.]
- [7] Wang Chenguang, Qin Haoran, Yang Chaoyue, et al. Strength prediction and influencing factor analysis of frozen soil based on the XGBoost algorithm method [J]. *Metal Mine*, 2025(7): 43-50. [王晨光, 秦浩然, 杨超越, 等. 基于XGBoost算法的冻土强度预测与影响因素分析[J]. *金属矿山*, 2025(7): 43-50.]
- [8] Zhao Huiwei, Bing Hui. Prediction of the unconfined compressive strength of salinized frozen soil based on machine learning [J]. *Buildings*, 2024, 14(3): 641.
- [9] Wang Bingquan, Ran Youhua. Machine learning-based forecasting of future maximum frost depth over the Third Pole [J]. *Journal of Glaciology and Geocryology*, 2023, 45(2): 798-807. [王冰泉, 冉有华. 基于机器学习的第三极季节冻土最大冻结深度未来变化预测[J]. *冰川冻土*, 2023, 45(2): 798-807.]
- [10] Meng Wenzhao, Wu Wei. Machine learning-aided prediction of the mechanical properties of frozen fractured rocks [J]. *Rock Mechanics and Rock Engineering*, 2023, 56(1): 261-273.
- [11] Esmaili-Falak M, Katebi H, Vadiati M, et al. Predicting triaxial compressive strength and Young's modulus of frozen sand using artificial intelligence methods [J]. *Journal of Cold Regions Engineering*, 2019, 33(3): 04019007.
- [12] Li Dongwei, Chen Junhao, Zhou Yan. Elastic-plastic constitutive model of artificial frozen soil based on cascade-correlation artificial neural network [J]. *Journal of China Coal Society*, 2016, 41: 407-411.
- [13] Liu Fengyun, Luo Huairui, Wan Xusheng, et al. Influence factors and model prediction of soil thermal conductivity during freezing [J]. *Journal of Glaciology and Geocryology*, 2024, 46(1): 221-234. [刘凤云, 罗怀瑞, 万旭升, 等. 冻结过程土体导热系数影响因素及预测模型研究[J]. *冰川冻土*, 2024, 46(1): 221-234.]
- [14] Zhong Zilan, Wang Jiarui, Shen Jiaxu, et al. Identification methods and feature importance of pulse-like ground motions based on machine learning [J/OL]. *Journal of Vibration Engineering*, 2025: 1-10. [2025-08-03]. <https://doi.org/10.16385/j.cnki.issn.1004-4523.202411024>. [钟紫蓝, 王嘉睿, 申家旭, 等. 基于机器学习的脉冲型地震动识别方法及特征重要性研究[J/OL]. *振动工程学报*, 2025: 1-10. [2025-08-03]. <https://doi.org/10.16385/j.cnki.issn.1004-4523.202411024>.]
- [15] Zong Huilin, Yuan Xiping, Gan Shu, et al. Spatial distribution mapping of debris flow site in Xiaojiang River Basin based on the GEE platform [J]. *Spectroscopy and Spectral Analysis*, 2025, 45(4): 1045-1060. [宗慧琳, 袁希平, 甘淑, 等. 基于GEE云平台的小江流域泥石流迹地空间分布制图[J]. *光谱学与光谱分析*, 2025, 45(4): 1045-1060.]
- [16] Zhang Ding, Zhou Zonghong. Rock burst prediction based on data preprocessing and improved sparrow algorithm [J]. *Chinese Journal of High Pressure Physics*, 2025, 39(7): 113-126. [张鼎, 周宗红. 基于数据预处理和改进麻雀算法的岩爆预测[J]. *高压物理学报*, 2025, 39(7): 113-126.]
- [17] Han Fengge, Leng Chengbiao, Chen Jiajie, et al. Machine learning method for discriminating granite genetic types based on trace element composition of apatite [J]. *Acta Petrologica Sinica*, 2025, 41(2): 737-750. [韩凤歌, 冷成彪, 陈加杰, 等. 基于磷灰石微量元素组成的机器学习方法判别花岗岩成因类型[J]. *岩石学报*, 2025, 41(2): 737-750.]
- [18] Peng Zesen, Liu Qingzhen, Zhang Yi. Short term load forecasting based on multi-model comprehensive feature selection and LSTM-Attention [J]. *Distributed Energy*, 2022, 7(6): 11-20. [彭泽森, 刘庆珍, 张溢. 基于多模型综合特征选择和LSTM-Attention的短期负荷预测[J]. *分布式能源*, 2022, 7(6): 11-20.]
- [19] Tan Wenkan, Hu Nanyan, Ye Yicheng, et al. Rockburst intensity classification prediction based on four ensemble learning [J]. *Chinese Journal of Rock Mechanics and Engineering*, 2022, 41(Suppl 2): 3250-3259. [谭文侃, 胡南燕, 叶义成, 等. 基于四大集成学习的岩爆烈度分级预测[J]. *岩石力学与工程学报*, 2022, 41(增刊2): 3250-3259.]
- [20] Wang Chenguang, Yang Chaoyue, Qin Haoran, et al. Examination of determinants and predictive modeling of artificially frozen soil strength utilizing the XGBoost algorithm [J]. *Applied Sciences*, 2023, 13(24): 13170.
- [21] Momeni E, He Biao, Abdi Y, et al. Novel hybrid XGBoost

- model to forecast soil shear strength based on some soil index tests [J]. *Computer Modeling in Engineering & Sciences*, 2023, 136(3): 2527-2550.
- [22] Li Kaiqi, He Hailong. Towards an improved prediction of soil-freezing characteristic curve based on extreme gradient boosting model[J]. *Geoscience Frontiers*, 2024, 15(6): 101898.
- [23] Li Kaiqi, Yin Zhenyu, Zhang Ning, et al. A data-driven method to model stress-strain behaviour of frozen soil considering uncertainty[J]. *Cold Regions Science and Technology*, 2023, 213: 103906.

Prediction of frozen silty sand shear strength based on four ensemble learning methods: Boosting, Bagging, Stacking, and Voting

ZHAO Na, WANG Shiwang, WANG Laigui, MI Yakuan, DING Mingrui

(School of Mechanics and Engineering, Liaoning Technical University, Fuxin 123000, Liaoning, China)

Abstract: Traditional laboratory tests for determining the shear strength of frozen soil are time-consuming, labor-intensive, and unable to meet the dynamic prediction requirements of cold-region engineering. To address these limitations, this study aims to develop an efficient and interpretable prediction framework for frozen soil shear strength based on ensemble learning techniques. Specifically, four ensemble strategies—Boosting, Bagging, Stacking, and Voting—were employed to integrate multiple base learners and to investigate the effects of different ensemble mechanisms on model performance. Silty sand collected from the Inner Mongolia region was selected as the research material. Through a series of laboratory direct shear tests, a dataset containing 768 samples was established, with freezing time, freezing temperature, moisture content, axial pressure, and shear strength serving as key features. A total of 27 models, including single base models, Bagging ensembles, and Boosting ensembles, were constructed and compared in terms of predictive accuracy. The nine best-performing models were then selected as base learners for constructing Stacking and Voting ensemble models. Model performance was comprehensively evaluated using the coefficient of determination (R^2), mean absolute error (MAE), and mean squared error (MSE), and the influence of each feature was examined through feature importance analysis. The results demonstrate that the Stacking model achieved superior overall performance ($R^2=0.998$, MAE=1.7091, MSE=6.4233), outperforming the Voting model ($R^2=0.9979$, MAE=1.6864, MSE=6.7903). Both models exhibited consistent feature importance rankings, namely: moisture content > freezing time > freezing temperature > axial pressure. These findings indicate that moisture content and freezing time are the dominant factors controlling the shear strength of frozen soil. In conclusion, this study confirms that an appropriately designed ensemble learning strategy can achieve high-accuracy predictions using a small number of easily measurable parameters. The proposed approach not only reduces the cost and time of traditional testing but also enhances model interpretability. It provides an effective, economical, and scientifically robust method for predicting the shear strength of frozen soil, offering valuable insights for stability assessment and design optimization in cold-region geotechnical engineering.

Key words: silty sand; shear strength; ensemble learning; feature importance; interpretability analysis

(责任编辑: 冉有华)