DOI: 10.11858/gywlxb.20230837

基于 WOA-RF 的边坡稳定性预测模型

张建涛, 刘志祥, 张双侠, 郭腾飞, 袁丛祥

（中南大学资源与安全工程学院, 湖南 长沙 410083）

摘要: 为有效地预测边坡稳定性和预防边坡失稳事故的发生, 提出了鲸鱼优化算法（whale optimization algorithm, WOA）和随机森林（random forest, RF）相结合的混合模型 WOA-RF; 基于所收集的边坡案例, 采用混淆矩阵的分类性能指标和受试者工作特征曲线及线下面积评估混合模型 WOA-RF 的分类和泛化性能; 使用 WOA 对 4 种广泛应用的机器学习模型进行优化, 并将优化后的机器学习模型与 WOA-RF 模型进行对比分析。结果表明: WOA 可以有效地优化超参数和提升模型性能; 最优 WOA-RF 模型在训练集和测试集上的准确率分别为 0.99 和 0.94, 优化后, 准确率、精确率、召回率、精确率和召回率的加权平均值分别提升了 11.9%、19.0%、4.8% 和 11.9%; 对比分析各个模型的预测性能后发现, WOA-RF 模型的各项指标均优于其他模型; 确定了特征重要性排序, 发现容重是影响边坡稳定性的最敏感特征。WOA-RF 模型可有效地预测边坡稳定性, 预测结果可为防护措施的制定提供依据。

关键词: 边坡稳定性预测; 机器学习; 鲸鱼优化算法; 随机森林; 特征重要性

中图分类号: O347; TU457 **文献标志码:** A

矿石开采、道路交通、水利水电以及建筑等工程建设中会形成大量边坡, 破坏了原有的地表地质和生态平衡, 在强降雨、地震、凿岩爆破和开挖等动力荷载作用下, 极易诱发滑坡、崩塌和泥石流等地质灾害, 进而造成人员伤亡和财产损失[1−2]。目前, 边坡失稳已是世界三大地质灾害之一, 边坡稳定性分析与精准预测一直都是边坡防治中的关键和挑战[3]。因此, 建立科学、有效、准确和稳定的边坡稳定性预测模型, 对岩土工程发展和边坡失稳防治具有重大的理论和现实意义。

边坡稳定性分析方法研究一直是岩土工程领域的热点和难点之一。目前, 边坡稳定性分析方法可大致分为传统方法和人工智能方法两大类[4]。传统分析方法主要包括经验法、解析法和数值模拟法, 其基于经验公式或相关力学理论, 对边坡的稳定性进行分析[5]。然而, 边坡作为一个复杂系统, 具有高维性、非线性、不确定性、动态演变性和协同性等特点, 导致稳定性分析具有一定的复杂性和难度, 传统分析方法的应用显现出很大的局限性, 常表现为对于复杂的边坡稳定性问题很难给出合理的解答, 分析结果的可靠度不高[6]。

近年来, 随着人工智能技术的不断发展, 在解决非线性复杂问题上体现出了独特的优势, 为边坡稳定性分析方法的发展提供了全新的思路[7]。国内外许多学者开始将人工智能模型广泛应用于边坡稳定性预测中。张凯等[8] 采用轻量级梯度提升机算法预测边坡稳定性, 分析了归一化处理方法和不同比例的训练集与测试集对模型性能的影响。武梦婷等[9] 基于随机森林（random forest, RF）和极端梯度提升机对边坡稳定性进行评价, 并结合特征的重要程度提出了防治措施。Lin 等[10] 综合比较了 RF、万有引力、支持向量机（support vector machine, SVM）和朴素贝叶斯 4 种算法在边坡稳定性预测中的性能, 发现

* **收稿日期:** 2023-12-25; **修回日期:** 2024-01-19

基金项目: 国家重点研发计划项目（2022YFC2904101）; 国家自然科学基金（52374107, 51974359）

作者简介: 张建涛（2000—）, 男, 硕士研究生, 主要从事机器学习与边坡稳定性研究. E-mail: 215512079@csu.edu.cn

通信作者: 刘志祥（1967—）, 男, 博士, 教授, 主要从事采矿工程与岩石力学研究. E-mail: liulzx@csu.edu.cn

万有引力算法优于其他 3 种算法。Karir 等[11] 分别采用支持向量机、人工神经网络(artifical neural network, ANN)、RF 和极端梯度提升机 4 种算法构建边坡安全系数回归预测模型，并与数值模拟结果进行对比，发现基于树模型的算法(包括 RF 和极端梯度提升机)的预测性能更好。尽管传统的机器学习模型可以独立解决边坡问题，但稳定性较差，特别是在处理复杂目标函数时，收敛能力不足且容易陷入局部最优解。因此，有学者将全局优化算法和机器学习模型融合，构建混合模型，以提升模型的预测性能。陈光耀等[12] 采用正态云模型改进后的果蝇算法优化支持向量机，构建边坡稳定性预测模型，提高了超参数的寻优效率和模型准确率。Qi 等[13] 将萤火虫优化算法分别与逻辑回归、RF 等 6 种机器学习算法融合，构建混合集成模型，并对比各个混合模型的预测性能，结果表明，萤火虫优化算法可实现 6 种算法的性能优化，且混合集成模型可以有效地提升预测性能。然而，机器学习模型目前尚未能对边坡问题进行全面有效地分析，存在一定的局限性。因此，仍需要不断寻找收敛能力更强、稳定性更高的边坡稳定性预测模型。

研究表明，RF 模型在解决边坡稳定性问题上表现出色[9, 11]，但其与优化算法融合的混合集成模型鲜有报道。鉴于此，本研究结合 5 折交叉验证(5-fold cross-validation, 5-fold CV)，构建鲸鱼优化算法(whale optimization algorithm, WOA)与 RF 相结合的混合模型 WOA-RF，以实现边坡稳定性的预测；基于所收集的案例，采用混淆矩阵的分类性能指标、受试者工作特征曲线以及线下面积评价模型的分类和泛化性能，并将 WOA-RF 与使用 WOA 优化的 4 种模型进行对比分析；通过特征重要性分析评估特征的权重系数。本研究旨在为边坡稳定性分析和预测提供新思路。

1　算法原理

1.1　RF

RF 由 Breiman[14] 于 2001 年提出，是一种基于决策树(decision tree, DT)的并行式集成学习算法，通过构建多棵相互独立的决策树，并将其合并，以提高模型的预测和泛化能力。该算法的基本原理如下：首先通过对数据集进行有放回地随机抽样和随机特征选择，训练生成多棵决策树；随后通过投票或取平均值的方法进行集成，以使模型获得更加准确的预测结果和较好的稳定性。目前，该算法已广泛应用于数据挖掘、机器学习和人工智能等领域，成为了一种重要的集成学习算法。

RF 的基本步骤如下：(1) 从训练集中有放回地抽取 m 个随机样本，生成样本集；(2) 基于样本集，随机选择 k 个特征构建决策树($k \ll q$，其中 q 为总特征数)；(3) 重复步骤 (1) 和 (2) 构建 n 棵决策树，并组合成 RF；(4) 采用投票法进行预测，即每棵决策树投票以获得最终的分类结果。

1.2　WOA

WOA 由 Mirjalili 等[15] 于 2016 年提出，是一种基于自然界鲸鱼捕食行为的元启发式全局优化算法。该算法通过模拟鲸鱼群“包围捕食”和“气泡网捕食”2 种捕食行为，在迭代过程中不断调整当前鲸鱼的运动方向和步长，以实现全局搜索和局部收敛的平衡。

WOA 的原理如下。

(1) n 维空间中鲸鱼个体的位置为 $\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)$，模型假设鲸鱼选择“包围捕食”和“气泡网捕食”2 种捕食行为的概率相等，即 $P_1=P_2=0.5$。

(2) 包围捕食时，鲸鱼会游向最优或随机位置的鲸鱼。$\boldsymbol{A}$、$\boldsymbol{C}$ 为系数向量，当 $|\boldsymbol{A}|\geqslant 1$ 时，鲸鱼会游向最优位置的鲸鱼，其位置更新为

$$\boldsymbol{X}_i^{t+1}=\boldsymbol{X}_{\text{best}}^t-\boldsymbol{A}\left|\boldsymbol{C}\boldsymbol{X}_{\text{best}}^t-\boldsymbol{X}_i^t\right| \quad (p<0.5) \tag{1}$$

式中：t 为当前迭代次数；$\boldsymbol{X}_{\text{best}}^t$ 和 $\boldsymbol{X}_i^t$ 分别为当前迭代次数下最优鲸鱼位置向量和第 i 头鲸鱼的位置向量，迭代过程中有更优鲸鱼位置时更新 $\boldsymbol{X}_{\text{best}}^t$；$p$ 为 [0, 1] 内的随机数。当 $|\boldsymbol{A}|<1$ 时，鲸鱼会游向随机位置的鲸鱼，其位置更新为

$$\boldsymbol{X}_i^{t+1}=\boldsymbol{X}_r^t-\boldsymbol{A}\left|\boldsymbol{C}\boldsymbol{X}_r^t-\boldsymbol{X}_i^t\right| \tag{2}$$

式中：$\boldsymbol{X}_r^t$为随机鲸鱼的位置向量。$\boldsymbol{A}$ 和 $\boldsymbol{C}$ 可表示为

$$\boldsymbol{A} = 2\boldsymbol{a}\boldsymbol{r} - \boldsymbol{a} \tag{3}$$

$$\boldsymbol{C} = 2\boldsymbol{r} \tag{4}$$

式中：$\boldsymbol{a}$ 在每个维度的初始值为 2，并随迭代次数线性递减至零，$\boldsymbol{r}$ 为模长在 [0, 1] 内的随机向量。

(3) 气泡网捕食。鲸鱼气泡网捕食时会以螺旋线形状游动，其位置更新为

$$\boldsymbol{X}_i^{t+1} = \left|\boldsymbol{X}_{\text{best}}^t - \boldsymbol{X}_i^t\right| \mathrm{e}^{bl}\cos(2\pi l) + \boldsymbol{X}_{\text{best}}^t \quad (p \geqslant 0.5) \tag{5}$$

式中：b 为常数（默认值为 1），决定螺旋线形状；l为 [−1, 1] 内的随机数。

2　边坡稳定性预测模型

2.1　特征参数的选取

边坡稳定性受地形地貌、地质构造、地层岩性、水文条件以及人为活动等多种内外部、静动态因素的共同影响。根据理论研究和工程实践的惯例，可将主要影响因素归为边坡物理力学参数、几何结构参数和水文条件。基于各参数的敏感性、稳定性、相互独立性以及易获得性等特点，并结合前人的研究成果，选取容重 γ、内聚力 c、内摩擦角 φ、边坡倾角 Φ、边坡高度 H 和孔隙压力比 r_{u} 6 个影响因素作为模型的输入[16−17]，边坡状态 S 作为模型的输出。在模型的输入中，γ、c 和 φ 为边坡的基本物理力学参数，Φ 和 H 为边坡的几何结构参数，r_{u}为边坡失稳的外界触发因素。边坡剖面及模型特征参数如图 1 所示。

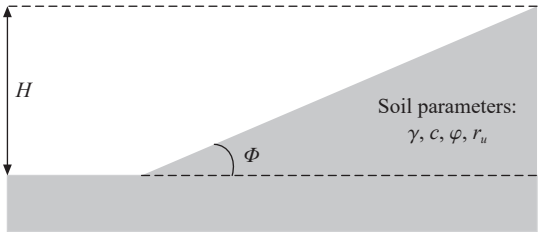


图 1　边坡特征参数

Fig. 1　Slope parameters

2.2　数据收集与分析

数据收集是建立机器学习分类模型的首要步骤，数据集的质量直接影响模型的预测性能。对于边坡稳定性这类工程问题，选择文献中具有代表性的数据构建模型至关重要。本研究通过文献检索和数据整理，收集了 168 组边坡案例[18−22]，并汇编成数据集，其中，稳定状态（标签为 1）和失稳状态（标签为 0）的边坡案例数量均为 84。这些边坡样本来源于中国、美国、英国和印度等地区，已广泛应用于边坡稳定性预测。

在建立模型之前，为了解数据集的分布情况，采用 Origin 绘图软件对该数据集进行描述性统计和可视化分析。统计结果见表 1。

表 1　数据集描述性统计

Table 1　Data set descriptive statistics

Feature	γ/(kN·m^{-3})	c/kPa	φ/(°)	Φ/(°)	H/m	r_{u}/kPa	S
Max	31.30	300.00	45.00	59.00	511.00	0.50	1.0
Min	12.00	0	0	16.00	3.60	0	0
Mean	21.76	34.12	28.73	36.10	104.19	0.22	0.5
Median	20.96	19.96	30.24	35.00	50.00	0.25	
Standard deviation	4.15	45.96	10.61	10.25	133.08	0.16	

图 2 为数据集的可视化分析结果，其中，对角线上的直方图反映了变量的分布情况，上三角位置的散点图显示了变量之间的相关性和趋势。从直方图可以看出，数据集的分布较为分散，大多数特征不服从正态分布。从散点图可以看出：对于稳定状态的边坡，其容重、内聚力和内摩擦角相对较大；孔隙

压力比越大，边坡越倾向于失稳。同时，对特征进行斯皮尔曼相关性分析，得到相关系数，如图 2 中下三角所示。相关系数均小于 0.5，即特征之间呈现弱相关性或中等相关性，不存在强相关性。此外，结合特征之间的散点分布情况，可以看出不同特征之间及特征与边坡状态之间存在较为复杂的非线性关系，这也表明边坡为复杂的非线性系统。

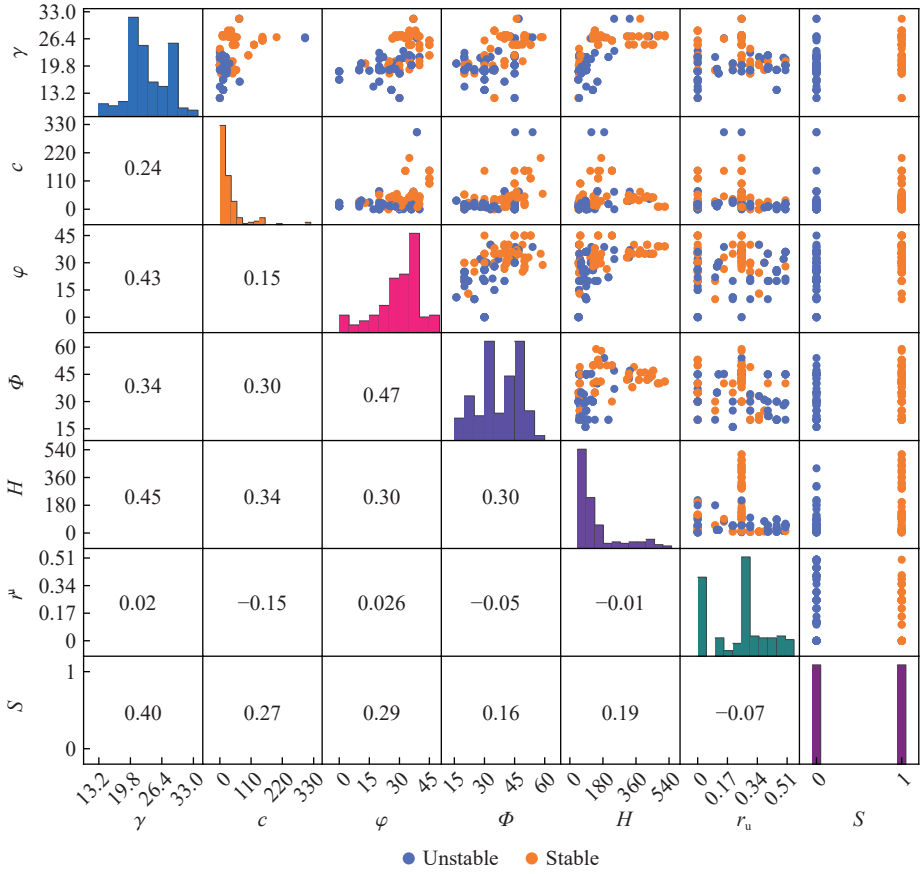


图 2　特征散点分布和相关系数

Fig. 2　Scatter distribution and correlation coefficient of dataset

考虑到数据集各个特征的单位和量纲不同，为避免某些特征量级的差异对模型产生不利影响，采用最大-最小值归一化处理方法对数据集进行预处理，处理方式见式 (6)。数据归一化处理不仅可以加快模型收敛速度，提高模型训练效率，还可以提高模型的精度和稳定性。

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

式中：x为原始特征数据，$x_{\max}$ 、$x_{\min}$分别为特征最大值和特征最小值，x_{norm}为归一化处理后的数据。

归一化处理后，原始数据将线性映射到 [0, 1] 区间内。随后，将数据集按照 7∶3 的比例随机划分为训练集（包括稳定样本 59 例、失稳样本 58 例）和测试集（包括稳定样本 25 例、失稳样本 26 例），分别用于构建模型和评估其预测性能。

2.3　模型分类性能指标

混淆矩阵（confusion matrix）是衡量分类模型性能的常用工具，可实现真实值与模型预测值之间差异的可视化，图 3 为一个二分类模型混淆矩阵。图 3 中，T_{p} 代表预测稳定且实际稳定的样本数量；F_{n} 代表预测失稳但实际稳定的样本数量；F_{p} 代表实际失稳但预测稳定的样本数量；T_{n} 代表实际失稳且预测失稳的样本数量。

基于混淆矩阵，可计算出多个分类性能指标。本研究选取 5 个指标评估模型的分类性能，即准确

率(accuracy, A_{cc})、精确率(precision, P_{re})、召回率(recall, R_e)、精确率与召回率的加权平均值 $F_{1\text{-score}}$ 以及受试者工作特性曲线(receiver operating characteristic, ROC)下的面积(area under the ROC curve, AUC)。前四者的表达式为[13]

$$A_{cc}=\frac{T_p+T_n}{T_p+F_n+F_p+T_n} \tag{7}$$

$$P_{re}=\frac{T_p}{T_p+F_p} \tag{8}$$

$$R_e=\frac{T_p}{T_p+F_n} \tag{9}$$

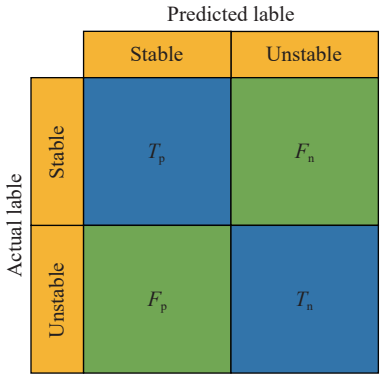


图 3　混淆矩阵

Fig. 3　Confusion matrix

$$F_{1\text{-score}}=\frac{2P_{re}R_e}{P_{re}+R_e} \tag{10}$$

准确率是分类模型最基本的性能指标之一，为模型正确预测的比例；精确率为模型预测结果为正例中真正例的比例；召回率为模型正确预测正例的比例；$F_{1\text{-score}}$ 是一种评估模型分类性能的综合指标。

假正率(false positive rate, F_{pr})和真正率(true positive rate, T_{pr})的表达式为

$$F_{pr}=\frac{F_p}{F_p+T_n} \tag{11}$$

$$T_{pr}=\frac{T_p}{T_p+F_n} \tag{12}$$

ROC 曲线是以假正率为横坐标、真正率为纵坐标的曲线。依据预测概率和实际标签，将不同阈值下的 F_{pr} 和 T_{pr} 绘制在坐标系中，连接各点形成 ROC 曲线。ROC 曲线越接近左上角，则模型性能越好。AUC 值是衡量模型分类性能的一个综合指标，取值范围在 0～1 之间，AUC 值越接近 1，则模型性能越好。

2.4　超参数调整

超参数调整是机器学习模型优化中至关重要的一个环节。RF 模型中有众多可调参数，在工程实践中，需要根据实际情况进行参数调优，寻找最佳的参数组合，使模型性能最优。常用的超参数优化方法有随机搜索、网格搜索和智能优化算法。

在进行超参数调整时常结合交叉验证[23]。交叉验证是一种评估模型性能并防止模型过拟合的技术。5 折交叉验证的原理如图 4 所示，将训练集随机分为 5 份，得到 5 个子集，其中 4 个子集用于训练模型，1 个子集用于测试模型，得到 RF 模型在测试子集上的准确率；重复以上步骤，每次选取不同的测试子集，直至每个子集都充当过一次测试集为止；最后，得到模型准确率的平均值，并以此作为优化过程中的适应度函数 F_{fit}。

2.5　WOA-RF 模型

本研究依托 Anaconda 3 软件中的 Jupyter Notebook 板块，结合 Scikit-learn 工具包，使用 WOA 优化算法对 RF 模型中的决策树个数 $n_{estimators}$、最大分离特征数 $N_{max_features}$ 以及决策树最大深度 D_{max_depth} 3 个超参数进行优化，以 5 折交叉验证后模型的平均准确率作为适应度函数，在训练集上多次迭代寻求最优超参数组合，最终构建最优 WOA-RF 混合分类模型。WOA-RF 模型的构建流程见图 5。

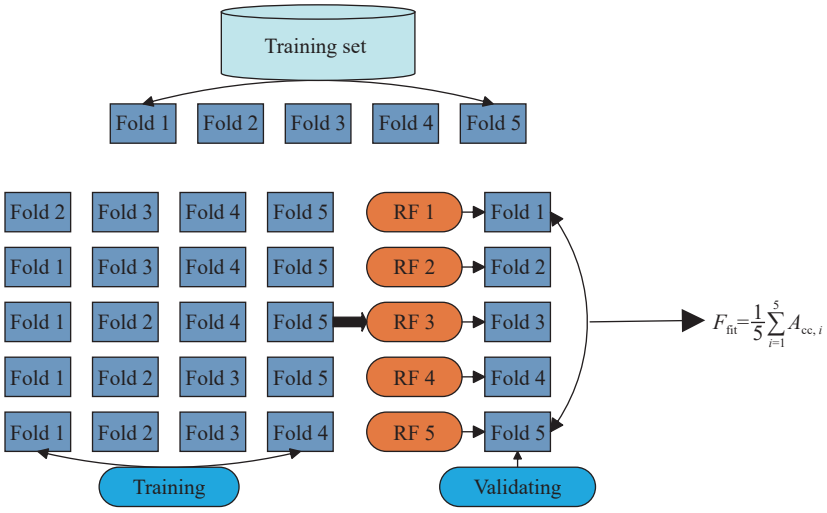


图 4　交叉验证和适应度函数计算步骤

Fig. 4　Cross validation and fitness function calculation process

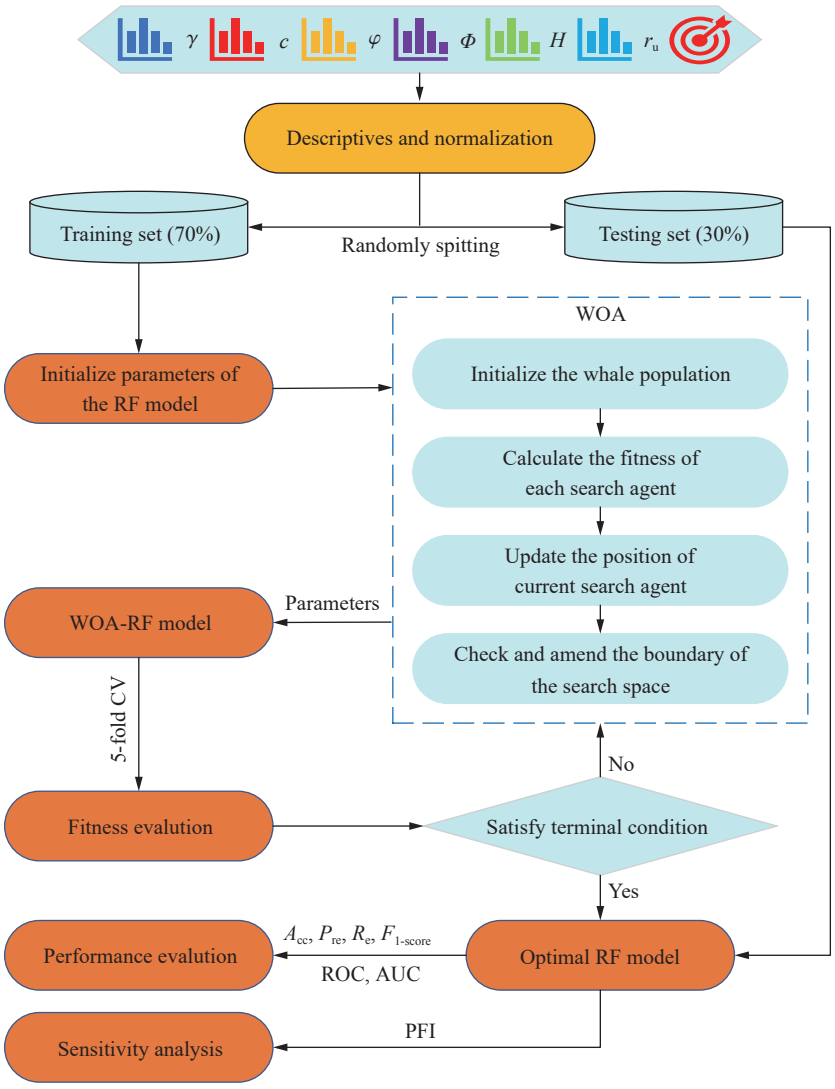


图 5　WOA-RF 模型构建流程

Fig. 5　WOA-RF model construction process

具体步骤总结如下：

(1) 边坡数据集的描述性统计分析和归一化处理，按照 7∶3 随机划分为训练集和测试集；

(2) RF 参数初始化，设置待优化参数的搜索上限和下限；

(3) WOA 参数的初始化，设置鲸鱼种群数量；

(4) 确认最佳搜索代理，并按照式 (1)～式 (5) 迭代更新位置；

(5) 以 5 折交叉验证后的模型平均准确率作为适应度函数，记录每一代最佳搜索代理的适应度函数值；

(6) 满足终止迭代条件时，输出最优超参数组合，并以此构建最优 WOA-RF 模型；

(7) 评估最优 WOA-RF 模型的性能，并使用 PFI(permutation feature importance)算法分析特征重要性。

3　结果与讨论

3.1　超参数优化结果

构建混合模型时，种群数量会显著影响模型的性能和优化过程[24]。为了使混合模型具有最佳性能，将种群数量分别设置为 20、40、60、80 和 100。图 6 显示了不同种群数量下模型适应度随迭代次数的变化情况。当迭代结束时，得到最优超参数组合，以此构建最优 WOA-RF 模型。

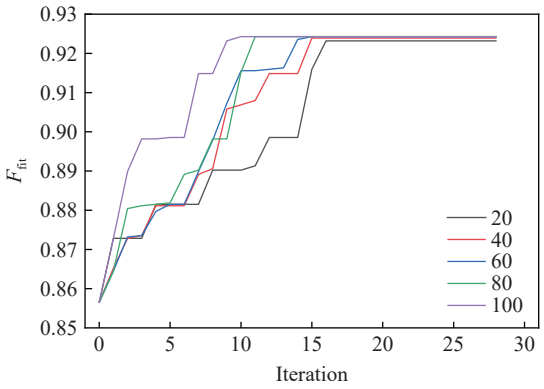


图 6　不同种群数量下优化过程中适应度的变化曲线

Fig. 6　Fitness variation for different population size during optimization process

从图 6 可以看出：各种群数量下模型经交叉验证后的适应度在迭代初期的增幅较大，收敛时均大于 0.92，说明 WOA 可以有效地优化 RF 模型的性能；当迭代次数达到 17 时，适应度均收敛，模型性能趋于稳定，即 WOA 优化算法表现出较好的搜索能力；随着种群数量的增加，模型收敛速度加快，特别是种群数量为 100 时，迭代 9 次后模型便开始收敛，适应度略优于其他种群，为最大值 0.924。RF 模型的超参数设置范围、优化结果以及其他基础参数设置列于表 2，其中，$N_{min_samples_leaf}$ 和 $N_{min_samples_split}$ 分别为叶子节点所需的最小样本数和节点分裂所需的最小样本数。

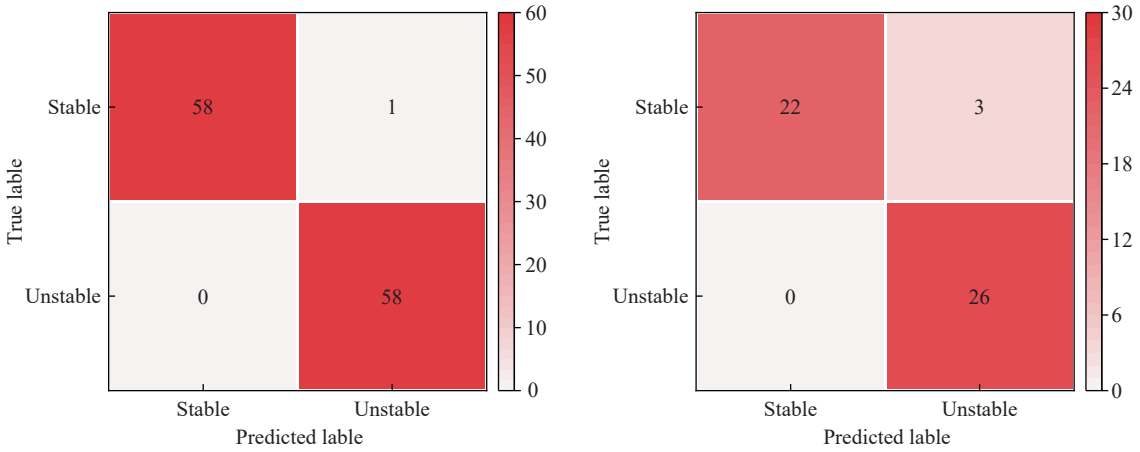
表 2　WOA-RF 模型参数设置

Table 2　Parameter settings in WOA-RF model

Title	$n_{estimators}$	$N_{max_features}$	D_{max_depth}	$N_{min_samples_leaf}$	$N_{min_samples_split}$	Criterion
Range	[1, 500]	[1, 6]	[1, 30]			
Optimized value	167	3	11	1	2	Gini

3.2　WOA-RF 模型预测结果

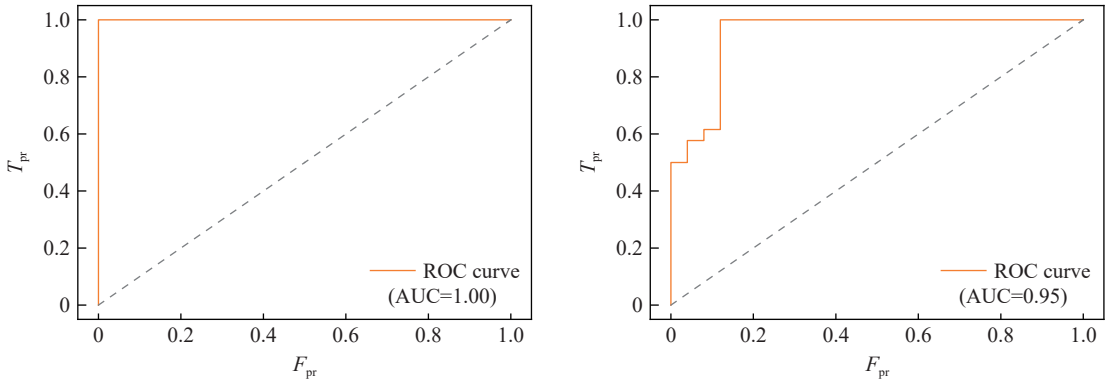
分别将训练集和测试集输入最优 WOA-RF 模型中，训练和测试结果的混淆矩阵和分类性能指标如图 7 所示，ROC 曲线如图 8 所示。从混淆矩阵可以看出，在训练集和测试集上，分别有 1 个和 3 个稳定样本被错误预测为失稳，但所有的失稳样本均被正确预测，因此，所建立的混合模型在识别边坡失稳上具有良好的性能，符合工程实际需求。计算各分类性能指标可得：训练集上，模型的 A_{cc} 为 0.99，P_{re} 为 1.00，R_e 为 0.98，AUC 值为 1.00，说明模型的训练效果较好；测试集上，模型的 A_{cc} 为 0.94，P_{re} 为 1.00，R_e 为 0.88，AUC 值为 0.95，说明模型的泛化性能较好。



(a) Training set, A_{cc}=0.99, P_{re}=1, R_e=0.98, $F_{1\text{-score}}$=0.99　　(b) Test set, A_{cc}=0.94, P_{re}=1, R_e=0.88, $F_{1\text{-score}}$=0.94

图 7　WOA-RF 模型在训练集和测试集上的混淆矩阵和分类性能指标

Fig. 7　Confusion matrix and classification performance of WOA-RF model for training and test sets



(a) Training set　　(b) Test set

图 8　WOA-RF 模型在训练集和测试集上的 ROC 曲线

Fig. 8　ROC curves of WOA-RF model for training and test sets

3.3　不同分类模型对比分析

为了进一步探索 WOA 优化超参数的可行性和 WOA-RF 模型的预测性能，使用 WOA 优化算法对 4 种广泛应用的机器学习模型（K 近邻算法（K-nearest neighbors, KNN）、SVM、ANN、DT）进行优化，优化后的模型将与 WOA-RF 模型进行对比分析。这些模型都将基于训练集建立，并通过测试集评估其性能。这些模型的训练集和测试集与 WOA-RF 模型的训练集和测试集保持一致，以确保结果的可靠性。

表 3 为各模型优化前后测试集上的 A_{cc} 和 $F_{1\text{-score}}$。可以看出，WOA 优化后，各模型的 A_{cc} 和 $F_{1\text{-score}}$ 均得到了不同程度的提升，说明 WOA 可实现对 4 种机器学习模型的超参数优化和模型性能的提升。其中，ANN 模型性能的提升最高，A_{cc} 和 $F_{1\text{-score}}$ 分别提高 50.9% 和 51.9%。WOA-RF 模型的 A_{cc} 和 $F_{1\text{-score}}$ 均为 0.94，相较于 RF 模型，这 2 个指标均提升了 11.9%。优化前后 RF 模型的性能均优于优化前后其余 4 种模型的性能。

表 4 为各模型优化后在测试集上的分类性能指标，根据不同模型在每个指标中的性能进行排名，性能越好则排名数值越大，将 4 个指标的排名数值相加，得到总得分。由表 4 可以看出，优化后的 5 个模型的总得分从高到低依次为 WOA-RF（20）、WOA-SVM（16）、WOA-KNN（11）、WOA-ANN（9）、WOA-DT（4）。由于 WOA-RF 和 WOA-SVM 的总得分较高，WOA-RF 和 WOA-SVM 可较好地进行边坡稳定性预测。此外，WOA-RF 的总得分显著高于 WOA-DT，说明由 DT 集成后的 RF 可大幅提升模型的预测性能。

表 3　各模型优化前后的A_{cc}和 $F_{1\text{-score}}$

Table 3　A_{cc} and $F_{1\text{-score}}$ before and after optimization

Model	A_{cc}	$F_{1\text{-score}}$	Model	A_{cc}	$F_{1\text{-score}}$
KNN	0.76	0.76	WOA-KNN	0.82	0.82
SVM	0.80	0.80	WOA-SVM	0.85	0.84
ANN	0.53	0.52	WOA-ANN	0.80	0.79
DT	0.71	0.71	WOA-DT	0.75	0.75
RF	0.84	0.84	WOA-RF	0.94	0.94

表 4　各模型优化后分类性能指标及排名

Table 4　Classification performance and ranking of models after optimization

Model	A_{cc}	A_{cc} rank	P_{re}	P_{re} rank	R_e	R_e rank	$F_{1\text{-score}}$	$F_{1\text{-score}}$ rank	Total score
WOA-KNN	0.82	3	0.82	2	0.82	3	0.82	3	11
WOA-SVM	0.84	4	0.85	4	0.84	4	0.84	4	16
WOA-ANN	0.80	2	0.83	3	0.76	2	0.79	2	9
WOA-DT	0.75	1	0.75	1	0.72	1	0.73	1	4
WOA-RF	0.94	5	1.00	5	0.88	5	0.94	5	20

图 9 为各模型在测试集上的 ROC 曲线和 AUC 值。可以看出：没有任何一个模型的 ROC 曲线可完全包络其他模型；WOA-RF、WOA-KNN 和 WOA-SVM 这 3 个模型的 ROC 曲线更接近左上角，AUC 值分别为 0.950、0.864 和 0.846；WOA-RF 的 AUC 值最大，大于 0.9。因此，WOA-RF 的预测性能最好。

此外，为综合评估各模型的分类性能，基于以上 5 个指标，绘制了雷达图，如图 10 所示。结果表明，构建的 WOA-RF 模型的各个指标均优于其他模型，表现出极好的性能。因此，选择 WOA-RF 模型作为预测边坡稳定性的最佳模型。

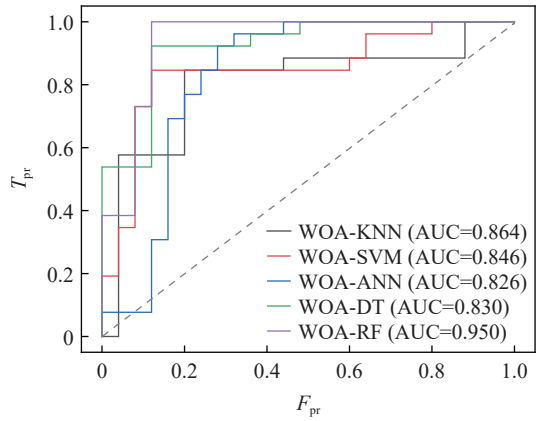


图 9　各模型在测试集上的 ROC 曲线

Fig. 9　ROC curves of different models for the test set

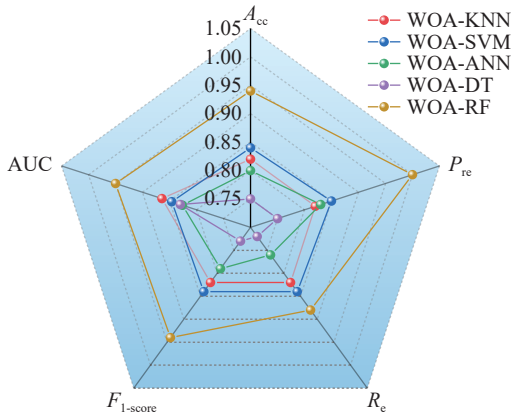


图 10　各模型分类性能指标雷达图

Fig. 10　Radar chart of performance indicators for different models

3.4　特征重要性分析

探索特征重要性并确定其对边坡稳定性的影响至关重要，分析结果可为边坡防治提供依据。采用 PFI 算法，通过随机打乱特征值序列和计算比较模型性能的衰减程度，评估特征重要性。对训练好的 WOA-RF 模型进行多次特征重要性分析，可获得特征的权重系数均值和标准差，如图 11所示。

图 11 中，横坐标为特征权重系数均值，误差棒为特征权重系数标准差，纵坐标为各个特征。显然，

γ(权重系数均值为 0.16)是影响边坡稳定性的最敏感因素，H、Φ、c、r_u和 φ 的权重系数均值相差不大，分别为 0.103、0.091、0.082、0.076 和 0.064，因此，特征重要性由高到低依次为：γ、H、Φ、c、r_u、φ。这些特征对边坡稳定性都有不可忽视的作用，作为模型的输入是合理的。

此外，需要注意的是，结合武梦婷等[9]、Qi 等[13]以及 Guyon 等[25]的研究结果可以发现，使用不同的数据集、分类模型和特征重要性分析方法，特征权重系数均值及排序会有所不同，在实际应用过程中需结合边坡工程情况选择相对重要的几个特征(γ、H、Φ、c)为着手点进行分析与防治。未来，随着更多边坡案例的出现和对特征更加综合全面的考虑，将获得更具代表性的结果。

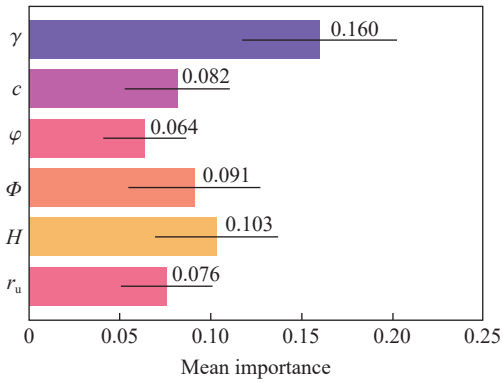


图 11　特征重要性分值(重复次数：100)

Fig. 11　Feature importance score (iteration: 100)

4　结　论

(1) 边坡是一个复杂系统，特征之间以及特征与边坡稳定性之间存在较复杂的非线性关系。

(2) WOA 可有效地实现超参数调整以及对 RF 模型性能的提升。优化后的 WOA-RF 模型在训练集上的准确率为 0.99，测试集上的准确率为 0.94，测试集上的准确率、精确率、召回率和 $F_{1\text{-score}}$ 分别提升了 11.9%、19.0%、4.8% 和 11.9%。

(3) 对比构建的 WOA-KNN、WOA-SVM、WOA-ANN、WOA-DT 和 WOA-RF 这 5 种混合集成模型，WOA-RF 模型的各个分类性能指标均为最优，综合性能最好。

(4) γ 的特征权重系数最大，表明其对边坡稳定性影响最大。整体来看，其余 5 个特征权重系数差别不大，但对边坡稳定性都有不可忽视的作用。因此，在进行边坡稳定性预测时，γ、H、Φ、c、r_u和 φ 这 6 个特征都应考虑在内。

(5) 所构建的模型将边坡稳定性视为二分类问题(稳定与失稳)。然而，不同类型边坡的安全等级划分标准有所不同，因此，下一步将基于机器学习对边坡稳定性进行多元分类分析和回归分析。

参考文献：

[1] 陈昌富, 李伟, 张嘉睿, 等. 山区公路边坡工程智能分析与设计研究进展 [J]. 湖南大学学报 (自然科学版), 2022, 49(7): 15–31.
CHEN C F, LI W, ZHANG J R, et al. State-of-the-art of intelligent analysis and design in slope engineering of highways in mountainous areas [J]. Journal of Hunan University (Natural Sciences), 2022, 49(7): 15–31.

[2] 何丽平, 汪晓俊, 郭剑雄, 等. 砂泥岩互层岩质边坡爆破振动衰减规律现场试验研究 [J]. 高压物理学报, 2023, 37(5): 055301.
HE L P, WANG X J, GUO J X, et al. Field experimental research on blasting vibration attenuation law of sand-mudstone interbedded rock slope [J]. Chinese Journal of High Pressure Physics, 2023, 37(5): 055301.

[3] LIN Y, ZHOU K P, LI J L. Prediction of slope stability using four supervised learning methods [J]. IEEE Access, 2018, 6: 31169–31179.

[4] KURTZ W, LAPIN A, SCHILLING O S, et al. Integrating hydrological modelling, data assimilation and cloud computing for real-time management of water resources [J]. Environmental Modelling & Software, 2017, 93: 418–435.

[5] 杨杰, 马春辉, 程琳, 等. 高陡边坡变形及其对坝体安全稳定影响研究进展 [J]. 岩土力学, 2019, 40(6): 2341–2353, 2368.
YANG J, MA C H, CHENG L, et al. Research advances in the deformation of high-steep slopes and its influence on dam safety [J]. Rock and Soil Mechanics, 2019, 40(6): 2341–2353, 2368.

[6] DONG L J, WESSELOO J, POTVIN Y, et al. Discrimination of mine seismic events and blasts using the Fisher classifier, naive Bayesian classifier and logistic regression [J]. Rock Mechanics and Rock Engineering, 2016, 49(1): 183–211.

[7] QI C C, TANG X L. A hybrid ensemble method for improved prediction of slope stability [J]. International Journal for Numerical and Analytical Methods in Geomechanics, 2018, 42(15): 1823–1839.

[8] 张凯, 张科. 基于 LightGBM 算法的边坡稳定性预测研究 [J]. 中国安全科学学报, 2022, 32(7): 113–120.
ZHANG K, ZHANG K. Prediction study on slope stability based on LightGBM algorithm [J]. China Safety Science Journal, 2022, 32(7): 113–120.

[9] 武梦婷, 陈秋松, 齐冲冲. 基于机器学习的边坡安全稳定性评价及防护措施 [J]. 工程科学学报, 2022, 44(2): 180–188.
WU M T, CHEN Q S, QI C C. Slope safety, stability evaluation, and protective measures based on machine learning [J]. Chinese Journal of Engineering, 2022, 44(2): 180–188.

[10] LIN S, ZHENG H, HAN B, et al. Comparative performance of eight ensemble learning approaches for the development of models of slope stability prediction [J]. Acta Geotechnica, 2022, 17(4): 1477–1502.

[11] KARIR D, RAY A, KUMAR BHARATI A, et al. Stability prediction of a natural and man-made slope using various machine learning algorithms [J]. Transportation Geotechnics, 2022, 34: 100745.

[12] 陈光耀, 汪明武, 金菊良. 基于 CMFOA-SVM 的边坡稳定性评价模型 [J]. 长江科学院院报, 2023, 40(2): 95–101.
CHEN G Y, WANG M W, JIN J L. CMFOA-SVM model for evaluating slope stability [J]. Journal of Changjiang River Scientific Research Institute, 2023, 40(2): 95–101.

[13] QI C C, TANG X L. Slope stability prediction using integrated metaheuristic and machine learning approaches: a comparative study [J]. Computers & Industrial Engineering, 2018, 118: 112–122.

[14] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1): 5–32.

[15] MIRJALILI S, LEWIS A. The whale optimization algorithm [J]. Advances in Engineering Software, 2016, 95: 51–67.

[16] LIN S, ZHENG H, HAN C, et al. Evaluation and prediction of slope stability using machine learning approaches [J]. Frontiers of Structural and Civil Engineering, 2021, 15(4): 821–833.

[17] KHAJEHZADEH M, KEAWSAWASVONG S. Predicting slope safety using an optimized machine learning model [J]. Heliyon, 2023, 9(12): e23012.

[18] LU P, ROSENBAUM M S. Artificial neural networks and grey systems for the prediction of slope stability [J]. Natural Hazards, 2003, 30(3): 383–398.

[19] SUN J L, WU S C, ZHANG H J, et al. Based on multi-algorithm hybrid method to predict the slope safety factor-stacking ensemble learning with Bayesian optimization [J]. Journal of Computational Science, 2022, 59: 101587.

[20] YAN X M, LI X B. Bayes discriminant analysis method for predicting the stability of open pit slope [C]//2011 International Conference on Electric Technology and Civil Engineering (ICETCE). Lushan: IEEE, 2011: 147–150.

[21] HOANG N D, PHAM A D. Hybrid artificial intelligence approach based on metaheuristic and machine learning for slope stability assessment: a multinational data analysis [J]. Expert Systems with Applications, 2016, 46: 60–68.

[22] SAH N K, SHEOREY P R, UPADHYAYA L N. Maximum likelihood estimation of slope stability [J]. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 1994, 31(1): 47–53.

[23] STONE M. Cross-validatory choice and assessment of statistical predictions [J]. Journal of the Royal Statistical Society: Series B (Methodological), 1974, 36(2): 111–133.

[24] LIU Z D, LI D Y. Intelligent hybrid model to classify failure modes of overstressed rock masses in deep engineering [J]. Journal of Central South University, 2023, 30(1): 156–174.

[25] GUYON I, GUNN S, NIKRAVESH M, et al. Feature extraction: foundations and applications (studies in fuzziness & soft computing) [M]. Berlin, Heidelberg: Springer-Verlag, 2006.

Slope Stability Prediction Based on WOA-RF Hybrid Model

ZHANG Jiantao, LIU Zhixiang, ZHANG Shuangxia, GUO Tengfei, YUAN Congxiang

（*School of Resources and Safety Engineering*, *Central South University*, *Changsha* 410083, *Hunan*, *China*）

Abstract: To effectively predict slope stability and prevent slope instability occurrence, a hybrid model WOA-RF, combining whale optimization algorithm (WOA) and random forest (RF) was proposed. Based on the collected slope cases, the classification and generalization performance of the model was evaluated according to the classification performance indicators given by the confusion matrix and the area under the receiver operating characteristic curve. Additionally, WOA was used to optimize four widely used machine learning models, and the optimized machine learning models were compared with WOA-RF. The results demonstrate that WOA is effective in optimizing hyperparameters and improving model performance. The optimal WOA-RF model achieves an accuracy of 0.99 on training set and of 0.94 on test set. After optimization, the accuracy, the precision, the recall, and the hamonic mean of the precision and recall are increased by 11.9%, 19.0%, 4.8%, and 11.9%, respectively. Comparative analysis reveals that the WOA-RF model is superior to the others in all indicators. Furthermore, the feature importance ranking was determined. Analysis of the feature importance indicates that unit weight is the most sensitive feature affecting slope stability. The established WOA-RF model is proved effective in predicting slope stability and facilitating the development of appropriate protective measures based on the predicted results.

Keywords: slope stability prediction; machine learning; whale optimization algorithm; random forest; feature importance