

基于 Attention-Stacking 的边坡稳定性预测

李小双^{1,2}, 缪鑫¹, 蒋宏伟^{1,3*}, 王孟来⁴, 庞建涛⁴

1. 常州大学城市建设学院, 江苏 常州 213164;
2. 齐鲁理工学院土木工程学院, 山东 济南 250200;
3. 河南理工大学能源科学与工程学院, 河南 焦作 454003;
4. 云南磷化集团有限公司, 云南 昆明 650607



摘要:传统的边坡稳定性预测方法依赖精确地质参数且难以处理非线性问题,而单一机器学习模型泛化能力有限。为了提高预测精度和稳健性,提出了一种基于注意力机制的多模型融合方法。该方法以随机森林(RF)、支持向量机(SVM)和XGBoost为基学习器,并通过贝叶斯优化基学习器的超参数,构建 Attention-Stacking 集成框架,引入双重注意力机制的元学习器,分别对输入特征和基模型输出进行动态权重分配。该方法的准确率达0.8604,精确率、召回率和F1分数分别为0.9206、0.8169和0.8657,均优于单一模型和传统Stacking方法。特征重要性分析显示,模型能够自动聚焦于内摩擦角和黏聚力等关键参数。研究表明,注意力机制能够有效提升集成模型的适应性和可解释性,为边坡稳定性预测提供了一种可靠的数据驱动解决方案。

关键词:边坡稳定性预测;机器学习;注意力机制;Stacking;贝叶斯优化

中图分类号: TU457 文献标志码: A 文章编号: 1005-2518(2026)03-0661-13 DOI: 10.11872/j.issn.1005-2518.2026.03.329
引用格式: 李小双, 缪鑫, 蒋宏伟, 等. 基于 Attention-Stacking 的边坡稳定性预测[J]. 黄金科学技术, 2026, 34(3): 661-673. LI Xiaoshuang, MIAO Xin, JIANG Hongwei, et al. Slope Stability Prediction Based on Attention-Stacking[J]. Gold Science and Technology, 2026, 34(3): 661-673.

边坡稳定性是岩土工程和地质灾害防治领域的核心课题(胡平等, 2023),直接关系到基础设施安全、人民生命财产安全和生态环境可持续发展。随着我国基础设施建设的不断推进,尤其是在山区高速公路、铁路、水利枢纽和矿山开采等重大工程中,高陡边坡广泛分布,边坡失稳引发的滑坡和崩塌等地质灾害频发,造成严重的经济损失和人员伤亡(易健翔等, 2024)。

传统的稳定性分析多依赖于极限平衡法和有限元法(Azarafza et al, 2021; 苏振宁等, 2022; 刘亮等, 2024; 祝经中等, 2024; 卢林娟等, 2025)。虽然传统边坡稳定性分析方法正在从经验性极限平衡法(郭志边等, 2025)向更严格的力学解析体系演进(常仕奇等, 2024; 周璐等, 2025),但这些方法对地质参数的精确性要求高,且难以应对复杂非线性(兰恒星等, 2022; 徐孟超等, 2025)和不确定性因素

收稿日期: 2025-11-11; 修订日期: 2026-01-18

基金项目: 云南省中央引导地方科技发展项目“重要生态敏感区磷矿山固废充填保水绿色开采技术应用研究”(编号: 202407AC110019)、国家自然科学基金项目“干湿—冷热—循环载荷耦合下高陡岩质边坡裂隙岩体的疲劳失稳机理研究”(编号: 42477142)、“干湿交替—爆破振动耦合下含空区高陡岩体力学响应特征及其失稳机理研究”(编号: 42277154)、“降雨—采动耦合作用下高陡边坡软弱夹层弱化与多场响应机制研究”(编号: 52404113)、山东省自然科学基金项目“露井复合采动下岩体响应的演化特征及其破坏机制研究”(编号: ZR2022ME188)、江苏省中青年学术带头人项目(SZC2509200010)和常州大学高层次人才引进项目“降雨入渗与露井复合采动耦合作用下高陡岩质边坡岩体响应的演化特征及其致灾机制研究”(编号: ZMF24020037)联合资助

作者简介: 李小双(1983—), 男, 湖北随州人, 教授, 博士生导师, 从事矿山露天开采与边坡工程方面的研究工作。E-mail: xsli2011@cczu.edu.cn

***通信作者:** 蒋宏伟(1992—), 男, 江苏泰兴人, 博士, 讲师, 从事地质灾害监测预警研究工作。E-mail: jianghongwei@cczu.edu.cn

© Editorial Department of Gold Science and Technology (CC BY-NC-ND)

的影响,限制了其在实际工程中的广泛应用。为了克服传统方法的局限性,数据驱动的机器学习技术已被广泛应用于边坡稳定性预测(Wang et al, 2024; 杜中亚等, 2025), 现有研究将其转化为二元分类问题, 并采用单一模型开展预测。相关研究成果表明, SVM、RF 和梯度提升等模型在边坡稳定性评估中均表现出良好效果(张梦涵等, 2022; Yang et al, 2023; 石峻峰等, 2025; 徐智超等, 2025)。然而, 单一模型受自身结构的约束, 泛化能力有限, 难以充分表征边坡系统复杂、非线性和多源异质的演化规律, 且性能易随数据分布波动而不稳定。其中, RF 易受样本不平衡影响(Maxwell et al, 2018), SVM 对噪声和异常值敏感(Cervantes et al, 2020), XGBoost 在复杂工况下存在过拟合风险(Zhang et al, 2020)。

为了克服单一模型在边坡稳定性预测中泛化能力不足和稳健性差等问题, 集成学习已成为提升预测性能的有效途径, 其通过融合多个基学习器输出, 可降低模型方差和偏差, 增强整体泛化能力。目前, 主流集成策略包括 Bagging、Boosting 和 Stacking 这 3 类(Hong et al, 2005; 谭勇等, 2024; 孟凡修, 2025)。其中, Stacking 作为高阶集成框架, 通过引入元学习器对异构基模型输出再学习, 充分利用各基模型优势, 提升预测精度和稳健性, 在复杂工程预测中表现出良好的潜力(孔坤锋等, 2025)。然而, 传统的 Stacking 仍存在一定的局限: 一是元学习器多采用静态加权机制, 忽视了不同基模型在不同输入区域的局部预测能力差异; 二是基模型输出缺乏置信度建模和上下文感知, 可能导致元学习器过度依赖不稳定或过拟合模型; 三是基学习器类型差异显著时, 固定权重融合易引发信息冗余或决策冲突, 限制集成系统优化潜力。

针对上述问题, 本研究提出了一种基于注意力机制的动态加权 Stacking 模型, 用于边坡稳定性预测。该方法在传统 2 层 Stacking 架构基础上, 构建了一个基于全连接神经网络的双重注意力元学习器, 实现对输入特征与基模型预测的联合建模和动态融合。基学习器选择 RF、SVM 和 XGBoost。在元学习阶段, 模型通过特征注意力模块自动学习各地质参数在不同样本中的相对重要性; 然后, 通过模型注意力模块, 基于当前样本的特征表示与各基模

型的预测输出, 生成基于每一个样本的权重向量, 实现“因样本而异”的模型集成策略。

1 数据预处理

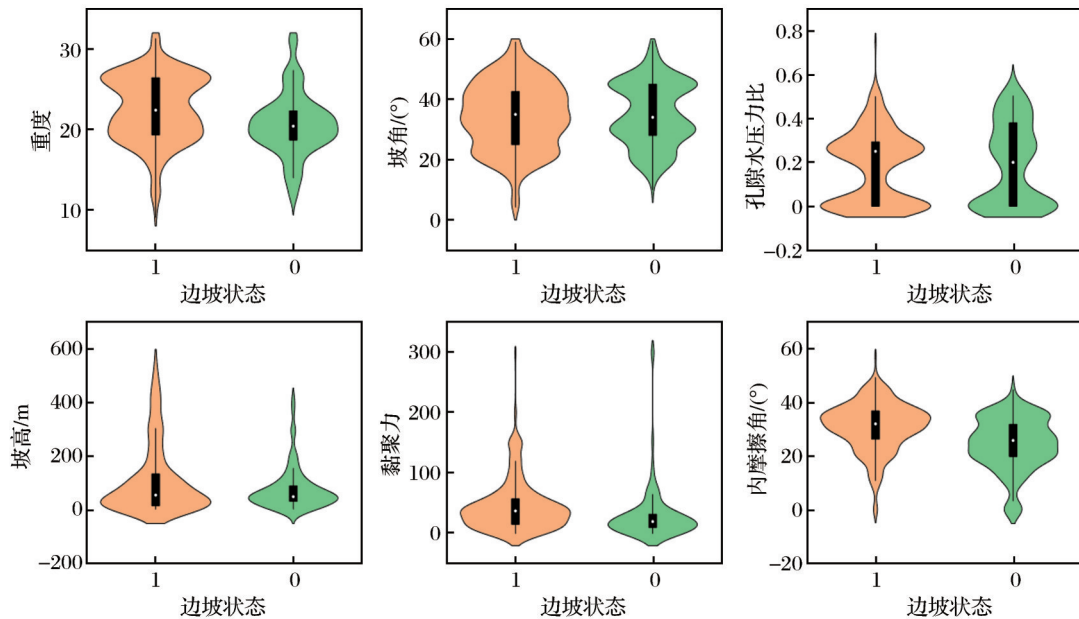
1.1 数据集的建立

影响岩土边坡稳定性的因素错综复杂, 但总体可划分为岩土体力学性质、边坡几何特征和外部环境因素这 3 类。因此, 引入特征坡高(H)、坡角(β)、重度(γ)、黏聚力(c)、内摩擦角(φ)和孔隙水压力比(γ_u)这 6 个主要参数, 对边坡稳定性进行预测, 并将边坡状态划分为 2 类: 1 表示稳定状态, 0 表示失稳状态。这 6 个参数是极限平衡法等经典理论中控制边坡安全系数的基本变量, 其中, 黏聚力和内摩擦角决定了土体的抗剪强度; 重度直接影响边坡的自重, 自重较大时会增加边坡失稳的风险; 坡高和坡角影响边坡的几何形态和应力分布; 水是重要的外部因素, 能够通过有效应力来削减抗滑力。

为了确保数据的真实性和可靠性, 本研究采用的数据均来自现有文献(Zhou et al, 2019; 张伟龙, 2020; Bai et al, 2022; Zhu et al, 2023; Huang et al, 2024), 文献数据涉及国内外多个地区的露天矿和水库库区等工程场景, 共有 643 组样本, 其中失稳状态数据为 288 组, 稳定状态数据为 355 组。为便于更直观地理解数据集, 通过绘制小提琴图来描述数据的分布情况。如图 1 所示, 与稳定状态相比, 失稳状态下的内摩擦角和黏聚力更集中于较低值。稳定边坡的重度分布整体高于不稳定边坡, 根据箱型图, 坡角均集中在 $20^\circ \sim 40^\circ$, 不稳定状态下的孔隙水压力明显高于稳定状态下的边坡。坡高集中在 $0 \sim 100$ m。在不同特征参数中, 稳定与失稳状态下的小提琴图形态较为相似, 离群值较少, 表明数据集构建合理, 数据质量可靠。

1.2 数据可视化

特征之间若存在强烈的多重共线性, 不仅会增加模型复杂度, 还会削弱其泛化能力。因此, 为了更好地理解数据集的特征分布及其相互关系, 对所收集的边坡稳定性数据进行可视化分析。绘制了各特征之间相关性和分布情况的矩阵散点图, 如图 2 所示, 根据边坡状态的不同, 划分为 2 种不同颜色, 其中蓝色代表稳定, 红色代表失稳。图中下三



注:小提琴图外部轮廓的宽度代表数据点分布的密度;小提琴图中心的矩形为箱型图,其上、下边缘分别为上四分位数和下四分位数,中间为中位数

图1 各特征参数分布小提琴图
Fig. 1 Violin diagrams of each characteristic parameter distribution

角是不同状态下的散点分布图,所绘制的趋势线代表2个特征之间的线性回归关系,斜率越大代表相关性越强,反之则越小,可以看出这些趋势线的整体斜率较小,说明特征之间的相关性较低。对角线是单个特征在不同状态下的分布图,可见特征之间没有明显的分布规律。上三角是特征之间的皮尔逊相关系数,其中坡高与重度之间的相关系数最高,为0.502,孔隙水压力与黏聚力之间的相关系数最低,为-0.048,其他大部分相关系数均为0.2~0.4,认为这些特征之间相对独立。

2 算法模型

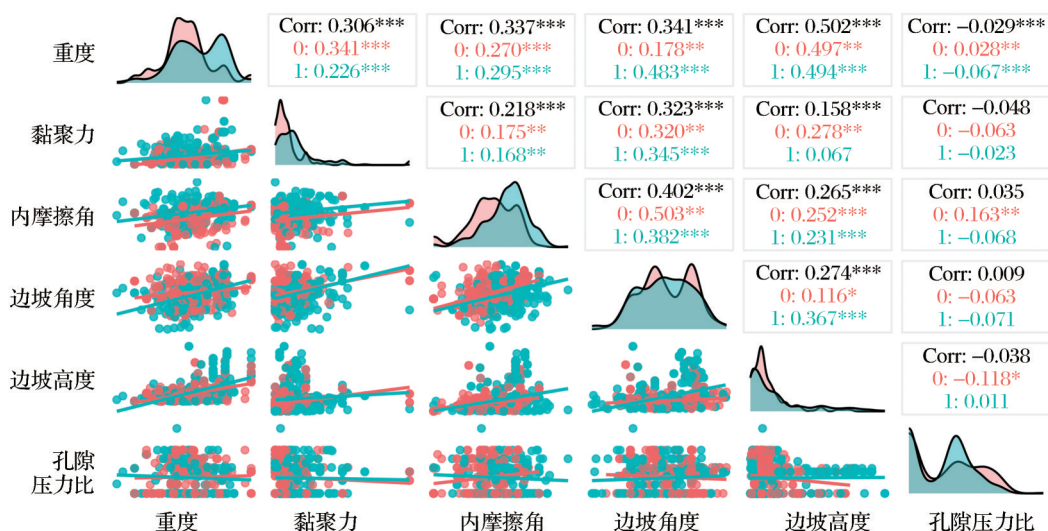
2.1 基学习器选择

(1)RF算法模型。随机森林是Breiman(2001)提出的一种基于集成学习思想的机器学习算法,通过构建并整合多棵决策树的预测结果,显著提升了模型的稳定性和预测精度。该算法的核心机制在于双重随机性,训练流程如图3所示。首先,在训练阶段,随机森林采用自助采样法从原始数据集中有放回地抽取多个不同的子样本集,每个子样本集用于独立训练一棵决策树,从而构建出一个多样化

的“森林”。其次,为进一步增加树与树之间的差异性,并降低过拟合风险,随机森林在每次节点分裂时,并不是考察全部特征,而是随机选取一个特征子集,且仅在该子集中寻找最优的分裂方式。

(2)SVM算法模型。支持向量机是一种基于统计学习理论的机器学习方法(Dietrich et al, 1999),其核心思想是通过寻找一个最优超平面来实现对数据的有效划分。该算法在处理线性不可分问题时,引入核技巧,将原始数据映射到高维特征空间,使数据更容易被线性分离,本研究采用径向基函数作为核函数。同时,SVM通过最大化类别之间的间隔来提升模型的泛化能力,减少过拟合风险。图4所示为支持向量机算法示意图。

(3)XGBoost算法模型。XGBoost是一种基于梯度提升框架的集成学习算法(Mitchell et al, 2017)。该算法通过构建多个弱学习器并以加法方式逐步优化模型预测结果,具有较高的预测精度和泛化能力。XGBoost的核心思想是通过迭代训练多个决策树,每棵树拟合当前模型残差,并将所有树的结果相加,从而获得最终预测值。其目标函数包括损失函数和正则化项2个部分,具体形式如下:



注:0表示2个特征参数之间在边坡失稳状态下的相关系数;1表示2个特征参数之间在边坡稳定状态下的相关系数;*,**和***分别表示在0.05、0.01和0.001水平下显著相关,图中红色和蓝色分别为失稳和稳定状态下的结果

图2 数据库相关矩阵散点分布图

Fig. 2 Scatter plot of database correlation matrix

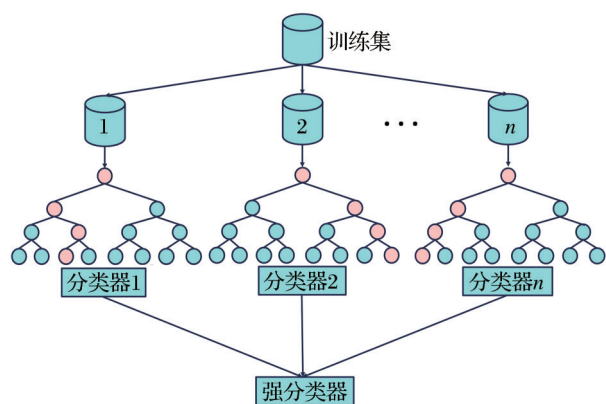


图3 RF算法示意图

Fig. 3 Schematic diagram of RF algorithm

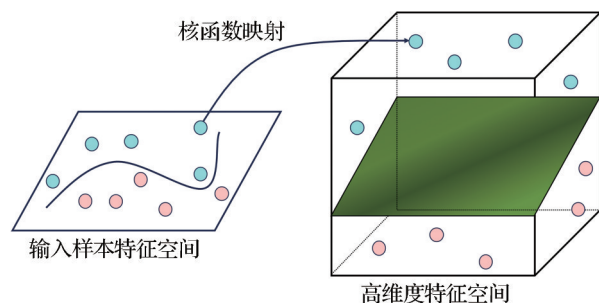


图4 SVM算法示意图

Fig. 4 Schematic diagram of SVM algorithm

$$Obj(f) = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^k \Omega(f_k) \quad (1)$$

式中: $L(y_i, \hat{y}_i)$ 为损失函数, 衡量模型预测值 $\hat{y}_i$ 与真实值 y_i 之间的误差; $\Omega(f_k)$ 为第 k 棵树的正则化项。 $\Omega(f_k)$ 定义为

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (2)$$

式中: T 为该树的叶子节点个数; ω_j 为第 j 个叶子节点的权重; γ 为控制叶子数量的 L1 正则化系数; λ 为控制叶子权重大小的 L2 正则化系数。最终预测结果由所有决策树的输出累加而成, 表示为

$$\hat{y}_i = \sum_{k=1}^k f_k(x_i) \quad (3)$$

2.2 贝叶斯优化算法

贝叶斯优化算法是一种基于概率模型的全局优化方法(Kaedi et al, 2017), 适用于复杂黑箱函数的超参数调优。其核心思想是通过构建目标函数的代理模型, 并结合采集函数指导下一步采样, 从而在有限评估次数内高效逼近最优解。贝叶斯优化算法的训练过程如图5所示。

贝叶斯优化的具体过程如下:

(1) 在超参数搜索空间中随机选取若干初始点进行评估, 形成初始观测数据集, 其中每个点对应一组超参数及其对应的性能指标, 本研究以 AUC 值作为优化目标。随后, 利用高斯过程回归作为代理模型, 基于已有观测数据对目标函数进行建模。

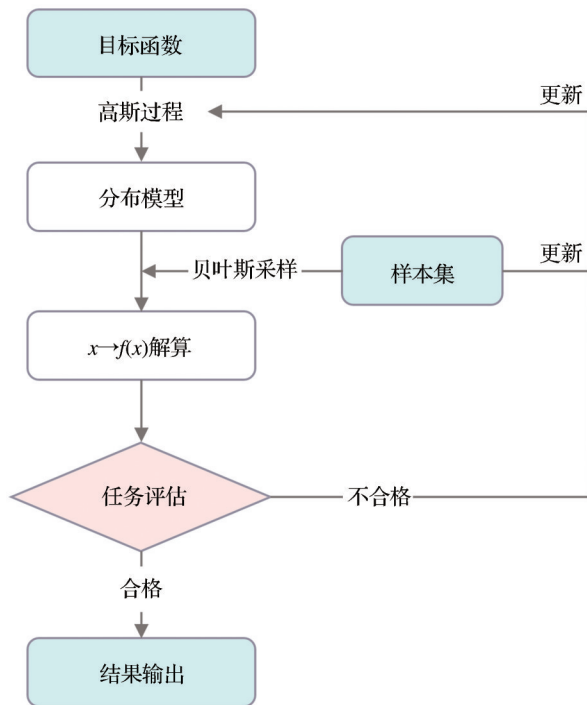


图5 贝叶斯优化算法流程

Fig. 5 Flow chart of Bayesian optimization algorithm

代理模型不仅能够预测未知点的性能,还能够提供预测不确定性,为平衡探索和开发提供依据。

(2)采用期望改善作为采集函数,权衡当前已知最优解与候选点的潜在提升空间,选择最具潜力的超参数组合进行下一轮评估。为鼓励适度探索,设置采集函数的kappa参数为2.5,xi参数为0,仅关注优于当前最优值的改进。

(3)每次评估新点之后,将其结果更新至观测数据集,并重新拟合代理模型,迭代优化过程。本文设定最大迭代次数为500次,前5次为随机初始化阶段,后495次基于采集函数主动搜索,直至达到最大迭代次数后终止,最终输出最优超参数组合。表1所示为基学习器的超参数。

2.3 Attention-Stacking模型

传统Stacking集成算法是一个2层的模型结构。第1层由多个不同的算法构成,如决策树、支持向量机和K近邻等,这些算法统称为基学习器。第2层由一种算法构成,通常为逻辑回归或线性回归,该算法称为元学习器。本文使用逻辑回归作为第2层的元学习器,与Attention-Stacking算法进行比较。在训练阶段,通常运用K折交叉验证的思想

表1 超参数选取

Table 1 Hyperparameter selection

算法	调参	取值范围	最优取值
RF	max_depth	(3,15)	13
	max_features	(1,6)	3
	min_samples_leaf	(1,102)	7
	min_samples_split	(2,256)	78
	n_estimators	(10,200)	58
SVM	C	(0.01,10)	46.87
	gamma	(0.01,10)	1.21
	Coef0	(0,10)	6.26
XGB	num_boost_round	(10,200)	68
	eta	(0.01,0.5)	0.5
	max_depth	(3,15)	5
	subsample	(0.1,1)	0.93
	colsample_bytree	(0.1,1)	0.31
	colsample_bylevel	(0.1,1)	0.87
	colsample_bynode	(0.1,1)	0.41

将训练集划分为多个子数据集,将各基学习器K折交叉验证的预测结果输入第2层进行元学习器的训练。在测试阶段,各基学习器在K折交叉验证时建立的模型对测试集进行预测,将同一基学习器不同模型的预测结果取均值输入元学习器,获得最终的测试结果。

与Stacking元学习器使用静态、固定的权重来融合基学习器不同,本研究提出了一种利用全连接神经网络实现双重注意力机制——特征注意力机制和模型注意力机制的元学习器。模型第1层选用随机森林、极限梯度提升和支持向量机作为基学习器,由于本研究包含643个边坡样本,因此采用5折交叉验证生成元学习器的训练特征,每次训练使用训练集80%的样本,能使基学习器充分学习,避免因训练数据不足导致模型欠拟合,剩余20%的验证样本有助于获得稳定的性能估计。同时,使用全部训练数据重新训练每个基学习器,获得用于最终测试集预测的最终模型。

第2层元学习器主要为特征注意力和模型注意力模块,特征注意力模块将原始特征作为输入,通过一个浅层前馈神经网络生成权重,具体计算过程如下:

$$z = \text{ReLU}(\mathbf{W}_1 x + \mathbf{b}_1) \quad (4)$$

式中: $\mathbf{W}_1$ 为第1层权重矩阵; x 为输入特征; $\mathbf{b}_1$ 为偏

置向量; $ReLU$ 为激活函数, 引入非线性, 能够让神经网络学习复杂模式。

$$a_f = \text{Sigmoid}(\mathbf{W}_2 z + \mathbf{b}_2) \quad (5)$$

式中: $\mathbf{W}_2$ 为第 2 层权重矩阵; z 为经过第 1 层变换后的特征; $\mathbf{b}_2$ 为偏置向量; Sigmoid 为激活函数, 将输出压缩到 $(0, 1)$, 作为权重。最后, 将特征注意力权重 a_f 与原始特征逐元素相乘, 获得最终结果, 表示为

$$\tilde{x} = x \odot a_f \quad (6)$$

模型注意力模块将加权后的特征和 3 个基学习器的预测结果作为输入, 再通过 1 个神经网络结构, 生成各基学习器的融合权重。具体计算过程如下:

$$h = \text{ReLU}(\mathbf{V}_1 \tilde{x} + \mathbf{c}_1) \quad (7)$$

式中: $\mathbf{V}_1$ 为第 1 层权重矩阵; $\mathbf{c}_1$ 为偏置向量。

$$s = \mathbf{V}_2 h + \mathbf{c}_2 \quad (8)$$

式中: $\mathbf{V}_2$ 为第 2 层权重矩阵; $\mathbf{c}_2$ 为偏置向量; $s = [s_1, s_2, s_3]$, 表示每个基模型在当前样本上的得分。最后归一化为权重, 表示为

$$a_m = \text{Softmax}(s) \quad (9)$$

在训练阶段, 元学习器通过最小化二元交叉熵损失函数 (Binary Cross-Entropy, BCE) 来优化其性能, 对于每个训练批次模型计算预测概率与真实标签之间的 BCE 损失, 并以此作为优化目标。模型采用 Adam 优化器, 利用反向传播算法逐层计算损失函数对元学习器中所有可学习参数的梯度, 并以此

更新参数。训练过程批大小 (batch_size) 分别设置为 16、32 和 64, 训练轮次 (epochs) 分别设置为 100、500、900 和 1 000。最终经过测试 batch_size 为 32, epochs 为 900, 能使模型在充分学习的同时, 逐步收敛至较优的集成策略。

经过模型注意力模块后得到的权重与基学习器的预测结果输入至元分类器中, 获得最终预测结果。Attention-Stacking 算法预测流程如图 6 所示。

3 结果与分析

3.1 模型指标评估

本文预测边坡稳定性为二分类问题, 采用混淆矩阵 (表 2) 和受试者工作特征 (ROC) 曲线, 混淆矩阵的二级分类指标包括准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall) 和三级指标 F1 分数 (F1-Score)。混淆矩阵可对分类结果进行可视化, 其所反映的内容如图 7 所示。通过混淆矩阵, 计算各项指标: 准确率 (Accuracy) 是指模型正确预测的样本数占总样本数的比例, 反映模型整体的判别能力; 精确率 (Precision) 表示在模型预测为不稳定的样本中, 实际确实发生失稳的比例; 召回率 (Recall) 表示在所有实际不稳定的样本中, 被模型成功识别出来的比例; F1 分数 (F1-Score) 是精确率与召回率的调和平均数, 能够综合反映模型在二者之间的平衡能力。上述指标的计算公式为

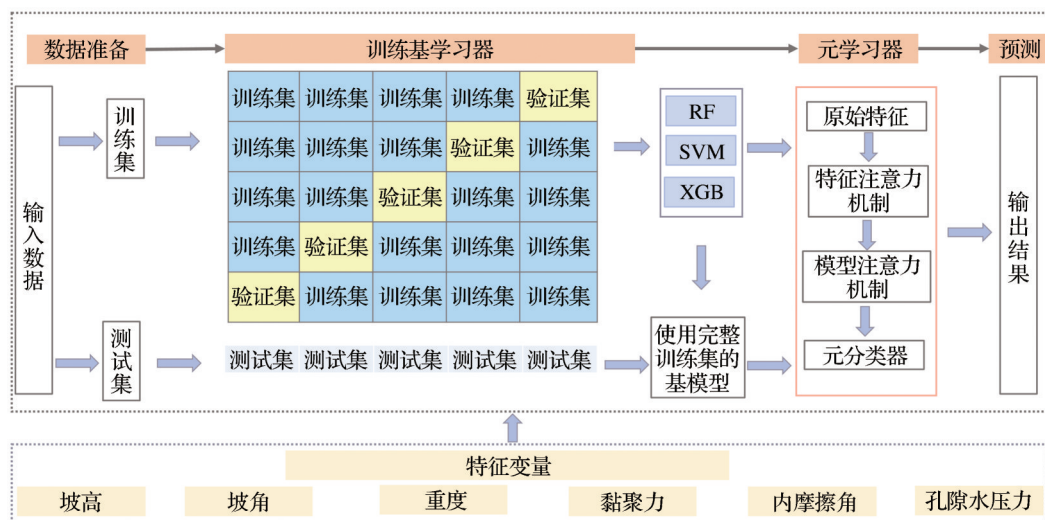


图 6 Attention-Stacking 算法预测流程

Fig. 6 Process of Attention-Stacking algorithm prediction

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F1 = \frac{2Precision \times Recall}{Precision + Recall} \quad (13)$$

式中: TN 表示真阴性; FP 表示假阳性; FN 表示假阴性; TP 表示真阳性。

表 2 混淆矩阵
Table 2 Confusion matrix

真实结果	预测结果	
	稳定	失稳
稳定	真阳性(TP)	假阴性(FN)
失稳	假阳性(FP)	真阴性(TN)

为了更好地展示模型性能,将 Attention-Stacking 算法与单一模型和 Stacking 算法的结果进行对比。如图 7 所示,分别为 RF、SVM、XGBoost、Stacking 和 Attention-Stacking 的混淆矩阵。由图 7 可以看出,这 5 种算法在判别边坡处于稳定状态的正确率高于判别边坡处于失稳状态,其中基于注意力机制改进后的堆叠模型和随机森林仅有 5 个被误判为失稳。 TP 、 FN 和 TN 这 3 项指标改进后的堆叠模型明显优于其他模型。图 8 所示为这 5 个模型的受试者工作特征曲线,可以发现 AUC 值均超过 0.85,说明这 5 个模型均具有良好的分类能力,而改进后的 Attention-Stacking 模型的 AUC 值为 0.93,超过其他 4 种模型,且 Attention-Stacking 算法的 ROC 曲线更靠近左上角,表明其在不同阈值下均能保持较高的真阳性率和较低的假阳性率。

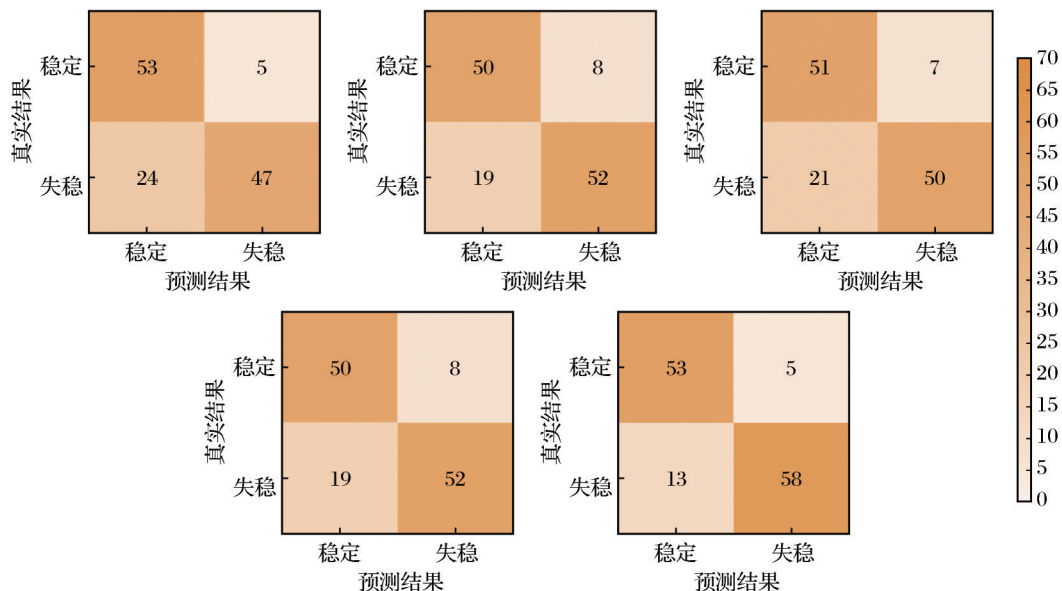


图 7 模型混淆矩阵

Fig. 7 Model confusion matrix

根据混淆矩阵计算出的准确率、召回率、精确率和 $F1$ 值,如表 3 所示。Attention-Stacking 模型达到最高准确率,数值为 0.8604,比其他模型高 7%~8%。这表明该模型在整体分类任务中的表现最为出色。

由表 3 也可以看出,Attention-Stacking 模型的精确率和召回率也远高于其他模型。由于其高精确率(0.9206)与高召回率(0.8169)的协同优化,最

终推动 $F1$ 分数达到 0.8657,显著优于其他方法。上述结果充分证明,引入的注意力机制通过自适应加权融合策略,使元学习器能够动态地聚焦于不同基模型的最可靠预测输出,并有效挖掘关键判别性特征。最终,通过引入注意力机制,Stacking 框架在复杂且不平衡的边坡稳定性预测任务中实现了性能突破,其高精度和高稳健性的特点,展现出巨大的工程应用潜力。

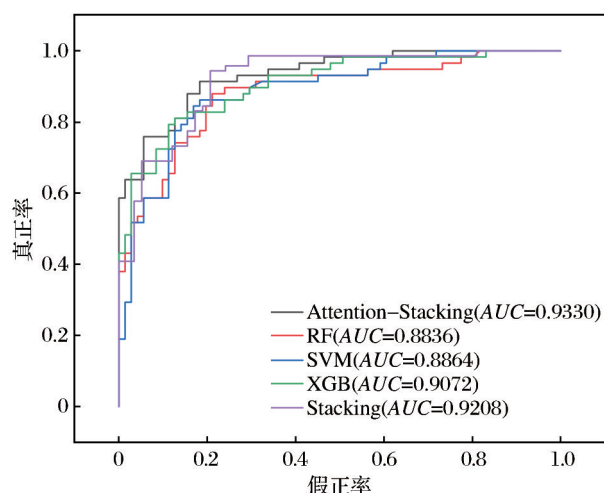


图8 模型 ROC 曲线

Fig. 8 ROC curves of model

表3 模型性能评估指标结果

Table 3 Results of model performance evaluation metrics

模型	准确率	精确率	召回率	F1
RF	0.7752	0.6883	0.9138	0.7852
SVM	0.7907	0.7246	0.8621	0.7874
XGB	0.7829	0.7083	0.8793	0.7846
Stacking	0.7907	0.7323	0.8666	0.7938
Attention-Stacking	0.8604	0.8169	0.9206	0.8657

3.2 特征重要性分析

基于元学习器中的特征注意力模型对6个特征的权重进行分析,如图9所示,各特征所获得的权重具有明显的差异性分布。内摩擦角、黏聚力和重度位列前3位,权重分别为0.6623、0.6377和0.4613,且这3个特征的权重远远高于其他特征,该现象与3个特征在岩土工程稳定性分析中的核心物理意义高度吻合。内摩擦角和黏聚力作为决定岩土抗剪强度的直接参数,被模型赋予最高权重,表明模型能够准确地识别并聚焦于这2个最关键的控制性因素。重度作为影响滑体自重应力的重要参数,其权重居第3位,进一步证明了模型决策逻辑的合理性。

相比之下,坡角、坡高和孔隙水压力所获得的注意力权重相对较低,分别为0.4183、0.3891和0.3626。这一分布特征同样符合岩土工程的理论认识:坡角和坡高主要定义了边坡的几何形态和临空条件,是边坡稳定性问题的外部诱因;而孔隙水

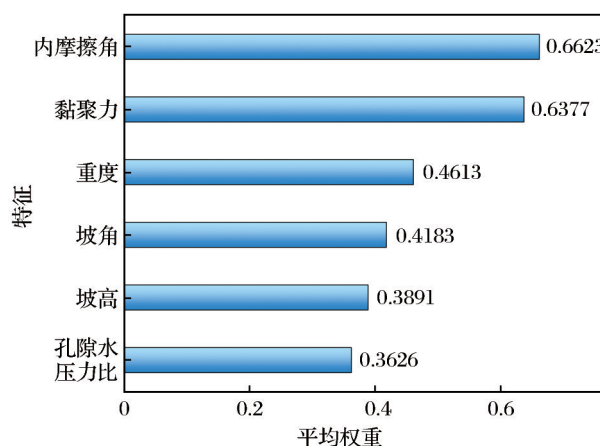


图9 特征平均权重

Fig. 9 Average weight of characteristics

压力比则通过影响有效应力来削弱抗滑力。这三者虽然对边坡稳定性具有重要影响,但其作用通常通过与岩土体强度参数的耦合来体现。模型能够自动学习到这种内在的物理层级关系。这种将数据驱动与物理机理相结合的能力,极大地增强了模型在边坡稳定性评估这类高风险决策任务中的可信度和可用性。

3.3 基模型权重分析

为了进一步探究本研究所提出的基于注意力机制改进的Stacking模型的内部决策机制,对模型注意力模块输出的基模型权重进行分析,如图10所示,展示了XGB、SVM和RF这3种基模型在测试集上获得的平均注意力权重分布。注意力权重呈

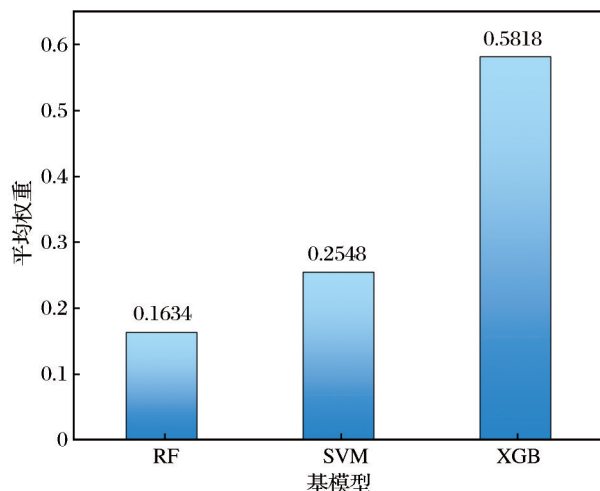


图10 基模型平均权重

Fig. 10 Average weights of the base model

现出高度非均匀分布,三者权重之和为1。XGBoost 获得了显著高于其他模型的权重,其权重值接近0.6,这表明元学习器在整体上最信赖 XGBoost 所提供的预测结果,将其视为基模型中的主导模型。SVM 所获得的权重约为 XGBoost 的一半,权重值为0.2548,位居第2,表明其预测提供了重要且独特的补充信息。随机森林的权重最低,说明其输出与 XGB 存在较高的冗余性,元学习器抑制了其贡献,从而优化了整体模型的决策效率。

这种权重等级表明,元学习器并非进行简单的模型平均或多数投票,而是学会了一种精准且有根据的信任分配策略。该结果暗示在此研究涉及的数据集和任务上,XGBoost 所学习的模式与目标变量之间具有最高的一致性、稳定性和可靠性。SVM 因其独特的核函数与决策边界,提供了 XGBoost 无法捕捉的某些判别模式,因而获得了相当的权重。相比之下,尽管随机森林本身是强大的模型,但其基于树的集成方式与 XGBoost 存在一定的相通性,导致其提供的增量信息有限,所以被分配了较低的权重。更重要的是,平均权重仅揭示了全局趋势,而该模型的核心优势在于其动态加权能力。对于每一个独立的样本,元学习器都会实时生成一套独一无二的权重。例如:对于预测置信度高的一些样

本,权重分配可能高度集中于主导模型;对于特征模糊、难以判断的一些样本,元学习器则会采取更均衡的权重分配,综合考量所有模型的意见,通过集体决策来降低误判风险。这种“因样本而异”的微观动态调整能力,是模型实现高性能和高稳健性的关键,极大地增强了模型决策过程的透明度和可解释性。

4 工程验证

为综合评价所提出的 Attention-Stacking 模型的泛化能力和工程实用性,本研究从一个独立的外部数据源中随机选取 10 个边坡案例进行验证。如表 4 所示,模型成功预测了其中 9 个案例的稳定状态,验证准确率达到 90%。唯一出现判别不一致的案例,其模型预测概率为 0.53,很接近决策阈值(0.5)。深入分析表明,该案例的参数组合存在显著的矛盾,黏聚力和内摩擦角均处于较高水平,提供了较强的内在抗剪强度。同时,极大的坡高和坡角以及存在的孔隙水压力具有较大的不稳定因素,这导致了模型的误判。总体而言,本次外部验证证实了模型具有良好的泛化能力,同时也明确了在未来的模型训练中要注意在极端工况下的鲁棒性。

表 4 案例验证结果
Table 4 Case validation results

序号	$\gamma/(\text{kN}\cdot\text{m}^{-3})$	$c/(\text{kPa})$	$\varphi/(^{\circ})$	$\beta/(^{\circ})$	$H/(\text{m})$	γ_w	实际状态	预测概率	预测状态
1	19.06	11.75	28	35	21	0.11	失稳	0.31	失稳
2	14	11.97	26	30	88	0	失稳	0.03	失稳
3	27.3	14	31	41	110	0.25	稳定	0.97	稳定
4	27	14	31	41	110	0.25	稳定	0.97	稳定
5	21.98	19.96	36	45	50	0	失稳	0.06	失稳
6	19.97	19.96	36	45	50	0.5	失稳	0.05	失稳
7	18.77	30.01	9.99	25.02	50	0.1	稳定	0.88	稳定
8	20.96	30.01	35.01	40.02	12	0.4	稳定	0.97	稳定
9	27.3	31.5	29.7	41	135	0.25	稳定	0.96	稳定
10	27	32	33	42.6	301	0.25	失稳	0.53	稳定

5 结论

构建了一个包含 643 组样本,涵盖坡高、坡角、重度、黏聚力、内摩擦角及孔隙水压力比 6 个关键

特征的边坡稳定性数据库。采用 Attention-Stacking 模型对边坡稳定性进行预测,通过引入 RF、SVM 和 XGBoost 作为基学习器,使用贝叶斯优化基模型超参数,以充分挖掘输入特征与稳定性状

态之间的非线性映射关系。

(1)相较于RF、SVM和XGBoost等单一模型以及传统Stacking集成方法,本文模型在准确率、精确率、召回率和F1分数上均获得显著提升,证明了引入注意力机制进行动态融合的有效性。

(2)通过特征重要性分析发现,模型自动赋予内摩擦角、黏聚力和重度等关键岩土力学参数更高的权重,这与边坡稳定性分析的经典理论高度吻合,表明模型不仅具备数据驱动特性,还能够捕捉潜在的物理机理,显著提升了结果的可信度。

(3)权重分析表明,元学习器能够根据具体样本特征,动态调整对各基模型预测结果的信任程度,而非简单平均。这种“因样本而异”的策略能够有效利用不同模型的优势,降低了单一模型误判的风险,从而提高了整体模型的泛化能力和稳健性。

(4)尽管本文所提方法取得了良好效果,但仍有进一步优化的空间。未来工作将从以下几个方面展开:首先,引入更丰富的数据特征,如地质构造、地下水动态监测数据和遥感影像等,从而更全面地描述边坡状态;其次,可探索更先进的注意力机制或序列模型,以捕捉特征间更复杂的依赖关系;最后,将模型应用于在线监测与实时预警系统,并在更多样化的工程场景中进行验证,是推动其走向实际工程应用的关键步骤。

参考文献 (References):

Azarafza M, Akgün H, Ghazifard A, et al, 2021. Discontinuous rock slope stability analysis by limit equilibrium approaches—a review [J]. *International Journal of Digital Earth*, 14 (12): 1918-1941.

Bai G X, Hou Y L, Wan B F, et al, 2022. Performance evaluation and engineering verification of machine learning based prediction models for slope stability [J]. *Applied Sciences*, 12 (15): 7890.

Breiman L, 2001. Random forests [J]. *Machine Learning*, 45 (1): 5-32.

Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, et al, 2020. A comprehensive survey on support vector machine classification: applications, challenges and trends [J]. *Neurocomputing*, 408: 189-215.

Dietrich R, Oppen M, Sompolsky H, 1999. Statistical mechanics of support vector networks [J]. *Physical Review Letters*, 82 (14): 2975-2978.

Hong W P, Song Y S, Lim S G, 2005. Stability evaluation of the cut slope using artificial neural network model [J]. *Journal of the Korean Society of Civil Engineers*, 25 (4C): 275-283.

Huang J X, Lu D, Lin W, et al, 2024. Enhancing slope stability prediction through integrated PCA-SSA-SVM modeling: a case study of Longlian expressway [J]. *Frontiers in Earth Science*, 12: 1429601.

Kaedi M, Ahn C W, 2017. Robust optimization using Bayesian optimization algorithm: early detection of non-robust solutions [J]. *Applied Soft Computing*, 61: 1125-1138.

Maxwell A E, Warner T A, Fang F, 2018. Implementation of machine-learning classification in remote sensing: an applied review [J]. *International Journal of Remote Sensing*, 39 (9): 2784-2817.

Mitchell R, Frank E, 2017. Accelerating the XGBoost algorithm using GPU computing [J]. *PeerJ Computer Science*, 3: e127.

Wang Y H, Wang L Q, Zhang W G, et al, 2024. A physics-informed machine learning solution for landslide susceptibility mapping based on three-dimensional slope stability evaluation [J]. *Journal of Central South University*, 31 (11): 3838-3853.

Yang Y K, Zhou W, Jiskani I M, et al, 2023. Slope stability prediction method based on intelligent optimization and machine learning algorithms [J]. *Sustainability*, 15 (2): 1169.

Zhang W G, Zhang R H, Wu C Z, et al, 2020. State-of-the-art review of soft computing applications in underground excavations [J]. *Geoscience Frontiers*, 11 (4): 1095-1106.

Zhou J, Li E M, Yang S, et al, 2019. Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories [J]. *Safety Science*, 118: 505-518.

Zhu W Q, Zhang S H, Li Y H, et al, 2023. Efficient slope reliability analysis based on representative slip surfaces: a comparative study [J]. *Frontiers in Earth Science*, 11: 1100104.

常仕奇,董晓强,刘晓凤,等, 2024. 基于极限平衡的边坡稳定性严格计算方法 [J/OL]. *太原理工大学学报*, 1-11 [2025-11-30]. <https://link.cnki.net/urlid/14.1220.N.20241203.1833.008>.

Chang Shiqi, Dong Xiaoliang, Liu Xiaofeng, et al, 2024. Strict calculation method for slope stability based on limit equilibrium [J/OL]. *Journal of Taiyuan University of Technology*, 1-11 [2025-11-30]. <https://link.cnki.net/urlid/14.1220.N.20241203.1833.008>.

- 杜中亚,刘翠,2025.数据驱动下边坡位移预测 SSA-BP 模型及应用研究[J].自动化应用,66(10):168-170. Du Zhongya, Liu Cui, 2025. Research on SSA-BP model and its application in slope displacement prediction under data driven[J]. Automation Application, 66(10):168-170.
- 郭志边,闫振国,刘志勇,2025.基于极限平衡法的黏性土边坡稳定性分析[J].路基工程,(4):76-79. Guo Zhibian, Yan Zhenguo, Liu Zhiyong, 2025. Stability analysis of cohesive soil slope based on limit equilibrium method [J]. Subgrade Engineering, (4):76-79.
- 胡平,熊治文,廖小平,等,2023.边坡安全风险评估研究综述[J].铁道建筑,63(2):108-115. Hu Ping, Xiong Zhiwen, Liao Xiaoping, et al, 2023. Review of slope safety risk assessment[J]. Railway Engineering, 63(2):108-115.
- 孔坤峰,陈宜玲,陈锋,等,2025.多模型决策级融合土质边坡稳定性预测模型及可解释性分析[J/OL].铁道科学与工程学报,1-13 [2025-11-30]. <https://doi.org/10.19713/j.cnki.43-1423/u.T20250822>. Kong Kunfeng, Chen Yiling, Chen Feng, et al, 2025. Multi-model decision level fusion of soil slope stability predictive model and interpretability analysis[J/OL]. Journal of Railway Science and Engineering, 1-13 [2025-11-30]. <https://doi.org/10.19713/j.cnki.43-1423/u.T20250822>.
- 兰恒星,彭建兵,祝艳波,等,2022.黄河流域地质地表过程与重大灾害效应研究与展望[J].中国科学(地球科学),52(2):199-221. Lan Hengxing, Peng Jianbing, Zhu Yanbo, et al, 2022. Research and prospect of geological surface processes and major disaster effects in the Yellow River Basin[J]. Scientia Sinica (Terrae), 52(2):199-221.
- 刘亮,徐森民,王迪,等,2024.基于最优边坡角的露天矿山边坡稳定性分析与安全评价研究[J].中国矿业,33(增刊2):98-104. Liu Liang, Xu Senmin, Wang Di, et al, 2024. Research on slope stability analysis and safety evaluation of open-pit mine based on optimal slope angle [J]. China Mining Magazine, 33(Suppl.2):98-104.
- 卢林娟,陈俊智,任春芳,2025.Fuzzy-AHP 法在顺层岩质边坡稳定性分析中的应用[J].黄金科学技术,33(3):615-625. Lu Linjuan, Chen Junzhi, Ren Chunfang, 2025. Application of Fuzzy-AHP method in the stability analysis of bedding rocky slopes[J]. Gold Science and Technology, 33(3):615-625.
- 孟凡修,2025.多特征 Adaboost 算法在多波束点云滤波中的应用[J].海洋测绘,45(2):19-23. Meng Fanxiu, 2025. Application of multi-feature Adaboost in multi-beam point cloud filtering[J]. Hydrographic Surveying and Charting, 45(2):19-23.
- 石峻峰,周琳,任宇联,等,2025.基于 INGO-RF 的边坡稳定性预测模型[J].安全与环境学报,25(4):1380-1390. Shi Junfeng, Zhou Lin, Ren Yulian, et al, 2025. Slope stability prediction model based on INGO-RF [J]. Journal of Safety and Environment, 25(4):1380-1390.
- 苏振宁,邵龙潭,2022.基于有限元极限平衡法的三维边坡稳定性[J].工程科学学报,44(12):2048-2056. Su Zhenning, Shao Longtan, 2022. Three-dimensional slope stability based on the finite element limit equilibrium method[J]. Chinese Journal of Engineering, 44(12):2048-2056.
- 谭勇,陈记,杨忠民,等,2024.基于 CatBoost 集成学习的边坡稳定性预测方法[J].科学技术与工程,24(30):13153-13160. Tan Yong, Chen Ji, Yang Zhongmin, et al, 2024. Slope stability prediction method based on CatBoost ensemble learning [J]. Science Technology and Engineering, 24(30):13153-13160.
- 徐孟超,王运敏,李小双,2025.基于 MatDEM 的露天转地下充填法开采中采场与边坡稳定性的影响研究[J].黄金科学技术,33(6):1139-1154. Xu Mengchao, Wang Yunmin, Li Xiaoshuang, 2025. Study on the impact of open-pit to underground filling mining on stope and slope stability based on MatDEM [J]. Gold Science and Technology, 33(6):1139-1154.
- 徐智超,陈匀杉,邓超,等,2025.基于 PSO-SVM 模型的边坡稳定性预测研究[J].安全与环境学报,25(9):3531-3537. Xu Zhichao, Chen Yunshan, Deng Chao, et al, 2025. Predicting slope stability using a PSO-SVM model [J]. Journal of Safety and Environment, 25(9):3531-3537.
- 易健翔,王宸琛,邹金锋,2024.基于修正拟动力法的岩质边坡三维稳定性分析方法及应用[J].铁道科学与工程学报,21(6):2295-2304. Yi Jianxiang, Wang Chenchen, Zou Jinfeng, 2024. Three-dimensional stability analysis approach for the rock slope based on the modified pseudo-dynamic method [J]. Journal of Railway Science and Engineering, 21(6):2295-2304.
- 张梦涵,魏进,卞海丁,2022.基于机器学习的边坡稳定性分析方法:以国内 618 个边坡为例[J].地球科学与环境学报,44(6):1083-1095. Zhang Menghan, Wei Jin, Bian Haiding, 2022. Slope stability analysis method based on machine learning: taking 618 slopes in China as examples [J]. Journal of Earth Sciences and Environment, 44(6):1083-1095.
- 张伟龙,2020.基于机器学习的边坡稳定性分析研究[D].太原:太原理工大学. Zhang Weilong, 2020. The Research of

- slope stability analysis based on machine learning[D]. Taiyuan: Taiyuan University of Technology.
- 周璐,陶志鹏,谢玉堂,等,2025.基于非线性 Mohr-Coulomb 破坏准则的边坡稳定性严格上限变分解析[J/OL].华东交通大学学报,1-13[2025-11-30].<https://doi.org/10.16749/j.cnki.jecjtu.20250908.002>.Zhou Lu, Tao Zhipeng, Xie Yutang, et al, 2025. Analysis of strict upper limit variation of slope stability based on nonlinear Mohr-Coulomb failure criterion[J/OL]. Journal of East China Jiaotong University, 1-13[2025-11-30]. <https://doi.org/10.16749/j.cnki.jecjtu.20250908.002>.
- 祝经中,朱飞,2024.某露天矿山高边坡稳定性分析研究[J].中国矿业,33(增刊1):124-127.Zhu Jingzhong, Zhu Fei, 2024. Analysis and research on high slope stability of an open-pit mine[J]. China Mining Magazine, 33(Suppl.1): 124-127.

Slope Stability Prediction Based on Attention-Stacking

LI Xiaoshuang^{1,2}, MIAO Xin¹, JIANG Hongwei^{1,3}, WANG Menglai⁴, PANG Jiantao⁴

1.School of Urban Construction, Changzhou University, Changzhou 213164, Jiangsu, China;

2.School of Civil Engineering, Qilu University of Technology, Jinan 250200, Shandong, China;

3.School of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo 454003, Henan, China;

4.Yunnan Phosphate Group Co., Ltd., Kunming 650607, Yunnan, China

Abstract: The prediction of slope stability is a vital aspect of ensuring safety in geotechnical engineering, as it significantly influences the reliability of infrastructure design and disaster mitigation strategies, particularly in mountainous areas. Traditional prediction methods, such as limit equilibrium approaches and numerical modeling, are often limited by their dependence on simplified assumptions and precise geological parameters, which are notoriously challenging to measure accurately in complex field conditions. Furthermore, these deterministic methods frequently fail to account for the highly nonlinear, heterogeneous, and stochastic nature of slope failure mechanisms. Although machine learning (ML) has emerged as a promising alternative, individual ML models often exhibit limited generalization capabilities, high variance, or bias when applied to diverse geological datasets, resulting in suboptimal predictive performance. To address these ongoing challenges, this study introduces a novel, robust data-driven framework: an Attention-Enhanced Multi-Model Fusion method based on an advanced Stacking ensemble architecture. The proposed system integrates three distinct and complementary base learners—Random Forest (RF), Support Vector Machine (SVM), and XGBoost—selected for their unique capabilities in managing high-dimensional, non-linear geotechnical data. To ensure optimal performance, the hyperparameters of each base learner are meticulously tuned using Bayesian optimization, which effectively navigates the complex parameter space to mitigate the risk of overfitting. The core innovation of this research resides in the development of the meta-learner, which incorporates an advanced dual attention mechanism. In contrast to traditional stacking ensembles that employ static or linear weighting schemes, our attention-based meta-learner dynamically allocates weights to both input features and the outputs of base models. This mechanism enables the model to autonomously identify and focus on critical geotechnical indicators, such as internal friction angle and cohesion, while adaptively prioritizing the most reliable base predictor for specific geological scenarios. Comprehensive experimental evaluations underscore the superior efficacy of the proposed Attention-Stacking framework. The model achieved an impressive accuracy of 0.8604, significantly surpassing standalone base models and traditional stacking methods. A detailed performance analysis reveals a precision of 0.9206, a recall of 0.8169, and an F1-score of 0.8657, indicating an exceptional

balance between minimizing false positives and accurately identifying unstable slopes. Furthermore, feature importance analysis highlights the model's enhanced interpretability; the attention mechanism effectively emphasizes key physical parameters that align with established geotechnical theories, thereby bridging the gap between black-box AI and engineering physics. By concurrently enhancing predictive accuracy, robustness, and transparency, this study presents a reliable, adaptive solution for slope stability assessment, offering substantial potential for advancing risk management protocols and decision-making processes in contemporary geotechnical engineering practices.

Key words: slope stability prediction; machine learning; attention mechanism; Stacking; Bayesian optimization