

基于 EO-LightGBM 融合模型的边坡稳定性预测方法研究

白龙发, 李保珠, 时鹏程

(昆明理工大学国土资源工程学院, 昆明 650093)

摘要: 边坡稳定性预测是确保边坡工程安全、经济和环保的关键环节。通过一系列数据对边坡的稳定性状态进行预测, 不仅避免了进行复杂的计算, 还能更好地了解边坡的稳定情况。基于此, 本文提出了一种基于 EO-LightGBM 模型的边坡稳定性预测方法, 并通过学习边坡的重度、内聚力、摩擦角、边坡角、边坡高度、孔隙压力比这 6 个特征和稳定性状态之间的关系, 实现对边坡稳定性的预测。此外, 通过与其他模型的对比研究, 验证了本文所提出的 EO-LightGBM 模型的可靠性。研究结果表明: 使用 EO 优化后的 LightGBM 模型误判率更低, 具有更高的准确性, 可以更好地预测边坡的稳定性状态; 同时也将为建立基于数据驱动的边坡稳定性智能决策控制平台提供参考。

关键词: 边坡稳定性; 融合模型; 机器学习; 智能预测; 数据驱动

中图分类号: TU457 **文献标识码:** A

STUDY ON SLOPE STABILITY PREDICTION METHOD BASED ON EO-LIGHTGBM FUSION MODEL

BAI Long-fa, LI Bao-zhu, SHI Peng-cheng

(Faculty of Land Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China)

Abstract: Slope stability prediction is a key link to ensure the safety, economy and environmental protection of slope engineering. Through a series of data, the stability state of the slope is predicted, which not only avoids complex calculations, but also better understands the stability of the slope. Based on this, this paper proposes a slope stability prediction method based on EO-LightGBM model, and realizes the prediction of slope stability by learning the relationship between the six characteristics of slope gravity, cohesion, friction angle, slope angle, slope height, pore pressure ratio and stability state. In addition, the reliability of the EO-LightGBM model proposed in this paper is verified by comparing with other models. On the surface of the research results, the LightGBM model optimized by EO has lower misjudgment rate and higher accuracy, which can better predict the stability state of the slope. At the same time, it will also provide a reference for the establishment of an intelligent decision-making control platform for slope stability based on data-driven.

Key words: slope stability; fusion model; machine learning; intelligent prediction; data-driven

0 引言

边坡是一个受多种因素共同影响的复杂地质体,随着我国大规模工程建设的快速发展,岩土边坡工程数量急剧增加,边坡失稳灾害的规模和频率也呈增强态势^[1-3]。而边坡稳定性是土木工程、公路、铁路和水利工程中的一个长期关注的关键问题^[4-7]。由于各种自然和人为因素,边坡失稳可能导致严重的经济损失、环境破坏和人员伤亡。因此,对边坡稳定性的研究和预测具有重要的工程实践和社会意义^[8-9]。

随着技术的快速发展,许多传统的边坡稳定性分析方法已被机器学习和数据驱动的方法所取代。冯夏庭等^[10]在 1995 年提出利用神经网络估计边坡稳定性;Qian 等^[11]基于极限分析方法使用人工神经网络(ANN)对非均质土质边坡进行了稳定性预测,并获得了较高的准确性。除此之外,Zhou 等^[12]、Liu 等^[13]、张凌凡等^[14],还将梯度提升(gradient boosting machine)、极限学习机(extreme learning machine)、决策树等算法模型应用到边坡稳定性预测中。然而,大多数学者对于安全系数预测的研究,往往是通过单一算法或者改进的单一算法进行研究。尽管模型精度上有所提升,但由于模型单一,加之数据量比较小,因此,在预测上存在一定的不稳定性,不具有很强的说服力。

与此同时,基于决策树的机器学习方法,如 LightGBM,因其出色的预测能力和计算效率在多个领域得到了广泛应用。平衡优化器(EO)方法也在多个领域,特别是在复杂优化问题中,显示了其独特的优势。基于此,本文提出了一种基于 EO-LightGBM 模型的边坡稳定性预测方法,根据收集到的 210 组边坡稳定性状态数据,结合 LightGBM 的预测能力和 EO 的全局搜索优势,为边坡稳定性预测提供了一个更为准确和鲁棒的方法,还同时克服了单个模型预测效果不稳定的缺点。通过对比验证,证明了 EO-LightGBM 融合模型在边坡稳定性预测方面优于其他现有方法,将为工程实践提供一个新的、高效的工具,从而更好地预测可能的边坡失稳事件。

1 数据来源及分析

影响边坡稳定性的因素主要有岩土体重度、黏聚力、内摩擦角、坡角、坡高及孔隙水压力比等,并且数据集的大小与质量直接影响模型构建的可靠性。为了准确可靠地反映各模型预测效果,通过大量文

献查询与数据汇编,从文献^[15]收集到包含以上特征的 210 组边坡数据样本,同时将数据集中安全系数 >1 的定性为“稳定”,将安全系数 <1 的定性为“破坏”,用以对边坡稳定性状态的定性评价预测。数据集各影响因素的分布情况如图 1 所示。

此外,为减小数据差异以及精准表示算法的精度,考虑到数据集采集受地质条件和稳定性状态评价指标差异的影响,在划分数据集之前,将数据汇总后重新进行数据随机排列,其中选取前 180 组数据作为训练集,余下 30 组用作测试集,以确保模型训练和测试的有效性。

2 边坡稳定性预测方法

2.1 平衡优化器算法(EO)

平衡优化器(equilibrium optimizer, EO)是 Afshin Faramarzi 等于 2020 年提出的一种较新的优化算法^[16-17]。该算法基于控制容积的质量平衡现象而设计,展示了优越的搜索能力和高效的收敛特性。通常可通过一阶微分方程加以描述:

$$V \frac{dC}{dt} = QC_{eq} - QC + G \quad (1)$$

式中, V 代表控制容积; C 代表控制容积内的浓度值; Q 代表流进或流出控制容积的容量流量; C_{eq} 是在控制体积中无质量生成(即达到平衡)时的浓度; G 代表控制容积内部的质量生成速率。

求解公式(1)可得:

$$C = C_{eq} + (C_0 - C_{eq})F + G(1 - F)/\lambda V \quad (2)$$

$$F = e^{-\lambda(t-t_0)} \quad (3)$$

式中, F 表示指数项系数; λ 表示流动率; C_0 表示控制容积在时间 t_0 的初始浓度。

EO 算法根据公式(2)进行迭代,搜索全局最优解。在整个寻优的过程中, C 为当前的新解, C_0 为上一次迭代得到的解, C_{eq} 为当前搜索到的最优解。解的迭代涵盖了对目前最佳解附近的细致搜索以及整体寻优范围的广泛探查,如图 2 所示。

平衡优化器具体实现步骤如下:

第一,算法对每个优化变量在其上下界之间执行随机赋初值:

$$C_i^0 = C_{\min} + r_i(C_{\max} - C_{\min}), i = 1, 2, \dots, n \quad (4)$$

式中, C_{\min} 和 C_{\max} 分别为优化变量的上下界; r_i 表示序号为 i 个体的随机数向量,为 $0 \sim 1$ 范围内的随机数。

第二,为防止算法陷入局部最优值,式(2)中的平衡状态将在当前 5 个最佳候选解中进行选取(图 2),此举旨在增强算法的全局寻优效能。由这些候

图 1 数据集各影响因素的分布情况

选解构成的平衡状态集合如下所示:

$$C_{eq, pool} = \{C_{eq,1}, C_{eq,2}, C_{eq,3}, C_{eq,4}, C_{eq,ave}\} \quad (5)$$

式中, $C_{eq,1}, C_{eq,2}, C_{eq,3}, C_{eq,4}$ 分别为当前迭代次数内搜索到的最优的 4 个解, $C_{eq,ave}$ 表示这 4 个解的平均状态。上述 5 个候选解均有 20% 的概率被选择。

第三,为了平衡算法的全局搜索和局部搜索,将公式(3)进行如下改进:

$$F = a_1 * \text{sign}(r - 0.5) [e^{-\lambda} - 1] \quad (6)$$

式中, a_1 为全局搜索的权重系数; r 和 λ 表示随机数向量,它们的每个元素值均为 0~1 范围内的随机数。

第四,设计生成速率,该过程可加强算法的局部寻优能力。

$$G = G_{CP} (C_{eq} - \lambda C) \quad (7)$$

$$G_{CP} = \begin{cases} 0.5r_1, r_2 \geq 0.5 \\ 0, \text{ otherwise} \end{cases} \quad (8)$$

式中, G_{CP} 为生成速率控制参数向量; r_1 为随机数向量,其每个元素值均为 0~1 之间的随机数; r_2 为 0~1 范围内的随机数。

最后,按照公式(9)对个体的解进行更新。

$$C = C_{eq} + (C - C_{eq})F + G(1 - F)/\lambda V \quad (9)$$

2.2 LightGBM 集成学习

LightGBM(light gradient boosting machine)算法是由微软公司于 2017 年提出的基于梯度提升

图 2 平衡优化器寻解过程

迭代决策树 (GBDT) 的一种新型集成学习算法^[18-19]。基于 GBDT 的框架,LightGBM 整合了独立特征打包方法(exclusive feature bunding, EFB)与基于梯度的单边取样策略(gradient-based one-side sampling, GOSS)。EFB 技术与 GOSS 技术不仅可充分利用样本数据,同时也可提高模型的训练速度。

LightGBM 集成了多个 CART 回归树:

$$f_T(X) = \sum_{t=1}^T f_t(X), f_t \in F \quad (10)$$

式中: f_i 对应第 k 棵决策树; F 为所有树的集合空间。回归树可以用另一种形式表示为 $\omega_{q(x)}$, $q \in \{1, 2, \dots, J\}$, J 为叶子数, q 表示树的决策规则, ω 表示叶节点的样本权重。

以牛顿法为基础采用累加方式进行模型训练:

$$G_t \cong \sum_{i=1}^n \left(g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right) = \sum_{j=1}^J \left(\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} (h_i + \lambda) \omega_j^2 \right) \right) \quad (11)$$

式中, g_i 、 h_i 分别为损失函数的一阶和二阶梯度统计量; I_j 为叶子 J 样本集。

在给定的树结构 $q(x)$, 每个叶子节点上的最佳叶子重量分数 ω_j^* 和极值 G_T^* 如下所示:

$$\omega_j^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}, G_T^* = - \frac{1}{2} \sum_{j=1}^J \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} \quad (12)$$

最后, 将拆分后的目标函数进行融合添加:

$$G = \frac{1}{2} \left[\frac{\left(\sum_{i \in I_L} g_i \right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_R} g_i \right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i \right)^2}{\sum_{i \in I} h_i + \lambda} \right] \quad (13)$$

式中, I_L 、 I_R 分别表示左右分支的样本集; λ 为惩罚模型的系数。

并且, 在 GBDT 的框架中, LightGBM 技术引入了优化过的直方图方法和带深度约束的叶子增长机制(leaf-wise)(图 3)。不仅显著缩短了模型的计

算时长, 同时在确保效率的前提下避免了过度拟合。

图 3 LightGBM 算法中树的生长策略

3 边坡稳定性预测与分析

3.1 EO-LightGBM 融合模型的建立与训练

采用 EO-LightGBM 边坡稳定性预测模型对边坡稳定性预测的流程如图 4 所示。本文根据第 2 节所描述的 210 组边坡岩性数据与边坡的稳定性状态一一对应, 构建预测模型所需数据集。其中, 特征值为岩土体重度、黏聚力、内摩擦角、坡角、坡高及孔隙水压力等, 目标值为边坡稳定性状态(稳定“1”/破坏“0”)。数据集构建完成后, 将其按 8:2 的比例划分为训练集(包含 180 组数据)与测试集(包含 30 组数据)。将训练集的 180 组数据输入至 LightGBM 模型进行训练。同时, 由于 LightGBM 算法包含多个

图 4 边坡稳定性预测流程

重要的超参数,本文采用平衡优化器(EO)对 LightGBM 的超参数进行调整。研究定义了一个目标函数来评估给定参数集下模型的性能,并让 EO 算法搜索参数空间,以找到使目标函数最小化的最优参数集,使模型的预测效果达到最优。因此,训练过程中通过 EO 对 LightGBM 的最优超参数进行自动搜索,得到最优的预测模型 EO-LightGBM。此外,为了评估模型的泛化能力,我们使用了 5 折交叉验证方法。训练完毕后,所得的最优预测模型被保存。同时,可通过提前划分好的测试集数据对上述最优模型进行检验,以检验该模型对边坡稳定性预测的准确度。除了 EO-LightGBM 模型,本文还训练了传统的 LightGBM、Xgboost、Adaboost 和 Random Forest 模型进行比较。每个模型的性能指标都被计算并对比,以对比 EO-LightGBM 模型的优势。

3.2 预测结果及分析

将 30 组测试集数据输入 4.1 节已经训练完成并保存的最优预测模型进行边坡的稳定性运算,以验证 EO-LightGBM 模型预测边坡稳定性状态的准确度。同时,将 EO-LightGBM 模型的预测结果,与仅使用 LightGBM 算法的预测模型得到的边坡稳定性状态结果及实际边坡稳定性状态进行对比,并将预测有误的在图中以(×)的形式标注,如图 5 和图 6 所示。

图 5 LightGBM 模型预测结果

由图 5 和图 6 可知,使用 LightGBM 模型对边坡稳定性状态的测试集数据进行预测,误判了 4 个,误判率为 13.33%。而经过平衡优化器算法对 LightGBM 算法的超参数进行优化后,预测效果得以改进,误判了 3 个,误判率降低为 10%。因此,在针对大样本的预测过程中,经过平衡优化器优化后的 LightGBM 算法的准确度更加可信。说明本文

所提出的 EO-LightGBM 预测模型对边坡稳定性状态的预测更具有说服力。

图 6 EO-LightGBM 模型预测结果

3.3 模型的优越性比较

为了进一步验证 EO-LightGBM 模型对边坡稳定性状态预测的可靠性和稳定性,本文选取了 Xgboost、Adaboost 和 Random Forest 这 3 种被广泛使用的算法作为对比,同时进行了边坡稳定性状态的预测研究。各模型的预测结果如表 1 所示。其中误判的样本在其预测结果后以“×”表示。

根据表 1 可知,在使用本文所提出的 EO-LightGBM 预测模型时,边坡稳定状态预测效果最好,在 30 组测试集数据中误判了 3 个,其余几个被广泛使用的算法的误判情况均多余 3 个。其中: Xgboost 模型误判了 4 个,误判率为 13.33%; Adaboost 模型误判了 5 个,误判率为 16.67%; Random Forest 模型效果相对最差,误判了 7 个,误判率高达 23.33%。EO-LightGBM 模型与 LightGBM、Xgboost 模型相比较,虽然误判减少数量不大,但误判的减少也代表了重要的改进,特别是在小样本数据集上。同时每次误判的减少在边坡稳定性预测中都非常关键,尤其是在防止灾害和减少损失方面,即使是轻微的准确度提升,也可能对实际操作和决策产生显著影响。

因此,可以进一步说明本文所提出的 EO-LightGBM 模型具有更强的学习能力,对边坡稳定性状态预测的准确度高,能有效地应用于边坡稳定状态的预测任务。

综上所述,本研究所提出的 EO-LightGBM 融合模型通过结合 LightGBM 的高效预测能力与 EO 算法的全局搜索优势,显著提高了模型的预测准确性和鲁棒性。通过 EO 算法对 LightGBM 的超参数进行自动优化,模型在边坡稳定性预测中展现

表 1 各模型对测试集的预测结果对比

测试集样本	真实值	LightGBM	EO-LightGBM	Xgboost	Adaboost	Random Forest
1	1	1	1	1	1	1
2	1	1	1	0(×)	1	1
3	1	0(×)	1	1	1	0(×)
4	1	1	0(×)	1	1	1
5	1	1	1	1	1	0(×)
6	1	1	1	0(×)	1	1
7	1	1	1	1	1	1
8	1	1	1	1	0(×)	1
9	1	0(×)	1	1	1	1
10	1	1	1	1	1	0(×)
11	1	1	1	1	1	1
12	1	1	1	1	1	1
13	1	1	1	1	1	1
14	1	1	1	1	1	1
15	0	1(×)	0	1(×)	1(×)	0
16	0	0	1(×)	0	0	1(×)
17	1	1	1	1	1	1
18	0	0	0	0	1(×)	0
19	1	1	1	1	1	0(×)
20	1	1	1	0(×)	1	1
21	1	1	1	1	1	1
22	1	0(×)	1	1	1	0(×)
23	1	1	1	1	1	1
24	1	1	0(×)	1	0(×)	1
25	1	1	1	1	1	1
26	0	0	0	0	0	0
27	1	1	1	1	1	1
28	1	1	1	1	0(×)	0(×)
29	1	1	1	1	1	1
30	1	1	1	1	1	1

出比传统算法更低的误判率,从而证明了其对于不同数据分布的强大适应能力。此外,与 Xgboost、Adaboost 和 Random Forest 等其他模型比较进一步验证了 EO-LightGBM 模型在准确度和鲁棒性方面的显著优势。

4 结论

准确地获取边坡的稳定状态是保障边坡安全建设的关键。因此,本文以 LightGBM 机器学习模型为基础,利用 EO 算法对其超参数进行优化,构建了 EO-LightGBM 预测模型对边坡稳定性状态进行预测。得出如下结论:

(1) 本文利用 EO 对 LightGBM 算法进行超参数优化,优化后的模型对边坡稳定状态的预测准确度有了明显提升,由误判 4 个降低至误判 3 个,验证了 EO-LightGBM 模型对边坡稳定性预测的可行性。

(2) 通过与其他几种被广泛应用的算法预测结果的对比,可以发现本文所提出的 EO-LightGBM 模型的准确度(90%)均高于 Xgboost 等其他几个模

型。证实了 EO-LightGBM 预测模型的可靠性,可以更好地运用到边坡稳定性状态预测中。

(3) 由于本文仅采用 210 组含有 6 个参数的边坡数据,受限于样本数量以及参数数量。如若能增加边坡其他参数以及大样本进行训练,必定能进一步提升算法的预测精度。

参考文献

- [1] 徐永福,程岩,唐宏华. 膨胀土边坡失稳特征及其防治技术标准[J]. 中南大学学报(自然科学版),2022,53(1):1-20.
- [2] 涂义亮,刘新荣,钟祖良,等. 三类边坡失稳判据的统一性[J]. 岩土力学,2018,39(1):173-180,190.
- [3] 王浩,王晓东,泮俊. 超高速路堑边坡治理工程案例研究 I:边坡失稳机制模拟分析[J]. 岩石力学与工程学报,2017,36(4):899-909.
- [4] 吴静,郭德宇,张乐文,等. 金沙江溪洛渡水电站拱坝及边坡稳定性分析研究[J]. 岩石力学与工程学报,2023,42(10):1-18.
- [5] 乔木,胡斌,李京,等. 考虑矿山边坡稳定性和经济效益的边坡控制参数敏感性[J]. 中国有色金属学报,2023,33(8):1-14.
- [6] 石玉玲,常洲,安宁,等. 基于干湿循环试验的黄土路堑浅层边坡长期稳定性分析[J]. 交通运输工程学报,2023,23(4):104-115.

- [7] 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.
- [8] 张凌凡,陈忠辉,周天白,等.基于梯度提升决策树的露天矿边坡多源信息融合与稳定性预测[J].煤炭学报,2020,45(S1):173-180.
- [9] 杨天鸿,王赫,董鑫,等.露天矿边坡稳定性智能评价研究现状、存在问题及对策[J].煤炭学报,2020,45(6):2277-2295.
- [10] 冯夏庭,王泳嘉,丁恩保.智能化的边坡稳定性分析方法[J].东北大学学报,1995(05):453-457.
- [11] Z. G. Qian, et al. An artificial neural network approach to inhomogeneous soil slope stability predictions based on limit analysis methods[J]. Soils and Foundations, 2019, 59(2): 556-569.
- [12] Jian Zhou, et al. Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories[J]. Safety Science, 2019, 118:505-518.
- [13] Liu Z, Shao J, Xu W, et al. An extreme learning machine approach for slope stability evaluation and prediction[J]. Nat Hazards, 2014, 73(2):787-804.
- [14] 张凌凡,陈忠辉,周天白,等.基于梯度提升决策树的露天矿边坡多源信息融合与稳定性预测[J].煤炭学报,2020,45(S1):173-180.
- [15] Junlong S, Shunchuan W, Huajin Z, 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.
- [16] Faramarzi A, Heidarinejad M, Stephens B, et al. Equilibrium optimizer: A novel optimization algorithm [J]. Knowledge-Based Syst. 2020, 191:105190.
- [17] Kardani N, Bardhan A, Gupta S, de al. Predicting permeability of tight carbonates using a hybrid machine learning approach of modified equilibrium optimizer and extreme learning machine [J]. Acta Geotech, 2022, 17(4):1239-1255.
- [18] Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree [J]. Advances in neural information processing systems, 2017, 30:3147-3157.
- [19] Liang W, Luo S, Zhao G, et al. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms [J]. Mathematics, 2020, 8(5):235-247.

作者简介: 白龙发(1997—),男,硕士研究生,研究方向为水文地质与工程地质。E-mail: 1249709468@qq.com

通讯作者: 李保珠(1973—),男,博士,副教授,从事水文地质教学、研究工作。E-mail: 147035644@qq.com

《地质灾害与环境保护》征稿启事

本刊主管单位:四川省教育厅;主办单位:成都理工大学、地质灾害防治与地质环境保护国家重点实验室。学术性期刊,季刊。1990年9月创刊,国内外公开发行人。本刊已加入《中国学术期刊(光盘版)》和“中国期刊网”,“万方数据网络系统(ChinaInfo)数字化期刊群”,“中文科技期刊数据库”。欢迎投稿,欢迎订阅。

征稿内容: 自然灾害(特别是地质灾害)与人为灾害防治,生态环境、资源环境和地质环境保护,水文地质与工程地质,岩土工程与工程勘察,与本学科有关的边缘和交叉学科等方面的优秀论文,最新科技成果,新理论、新方法、新技术研讨,及工程技术经验总结,新成就、新动向方面的综述、述评,新书刊、新设备评介等。

来稿注意事项:

(1) 文章正文应主题突出,数据翔实可靠,论点正确,图件清晰。

(2) 每篇稿件必须有:中英文篇名、摘要(摘要以第三人称书写)、图表名,所在单位中英文名称;第一作者简介,主要包括:姓名、性别、出生年、职务、职称,最高学历、专业、研究方向,主要科技成果及其获奖情况,以及出版专著、发表论文等简况。

(3) 来稿可电子邮件投稿,地址:dzzh@cdu.edu.cn,也可在线投稿:https://cont.cdu.edu.cn/jx_dzzh。

《地质灾害与环境保护》编辑部

2024年12月25日