

DOI:10.11918/202404059

基于 GA-LightGBM 算法的 TBM 掘进参数与岩体等级关系

曹贞洋¹, 龚敏¹, 吴昊骏¹, 龚潇雨², 吴晓东¹, 胡广风¹, 王思杰¹

(1. 北京科技大学 资源与安全工程学院, 北京 100083; 2. 北京科技大学 国家材料服役安全科学中心, 北京 100083)

摘要: 为建立反映 TBM 掘进参数与岩体等级关系的岩体等级识别模型, 并提高模型的构建效率和识别率, 以某隧道工程为背景开展研究。现场勘测岩体特征并基于 BQ 法和 RQD 值划分岩体等级, 采集 TBM 工作数据并筛选出与岩体特性变化相关的主要掘进参数; 基于轻量梯度提升器 (LightGBM) 算法, 拟合 TBM 掘进参数与岩体等级的关系, 并利用遗传算法 (GA), 优化 LightGBM 的超参数, 最终建立 GA-LightGBM 岩体等级识别模型。结果表明: GA-LightGBM 模型的识别准确率达到 93.5%, 高于支持向量机模型和随机森林模型的准确率, 且模型训练速度比梯度提升决策树算法提高了 8 倍; 5 种 TBM 掘进参数与岩石强度和岩体完整性等特性存在相关关系, 其中总推进力可作为感知岩体特征的主要判据。研究提供了一种高效分析 TBM 掘进参数并准确识别岩体等级的方法, 为现场快速感知岩体等级并实时调整作业参数提供支撑。

关键词: 隧道掘进机; 岩体等级识别模型; 遗传算法; LightGBM 算法

中图分类号: TU452

文献标志码: A

文章编号: 0367-6234(2025)07-0022-11

Relationship between TBM tunneling parameters and rock mass grades based on GA-LightGBM algorithm

CAO Zhenyang¹, GONG Min¹, WU Haojun¹, GONG Xiaoyu², WU Xiaodong¹, HU Guangfeng¹, WANG Sijie¹

(1. School of Resources and Safety Engineering, University of Science and Technology Beijing, Beijing 100083, China;

2. National Center for Materials Service Safety, University of Science and Technology Beijing, Beijing 100083, China)

Abstract: In order to establish an identification model of rock mass grade reflecting the relationship between TBM tunneling parameters and rock mass categories, improve model building efficiency and recognition rate, a research was carried out on background of a tunnel project. The rock mass characteristics were surveyed and graded based on BQ method and RQD, the TBM working data was collected and the main excavation parameters related to the change of rock mass characteristics were screened. The relationship between TBM tunneling parameters and rock mass grade was fitted based on Light Gradient Boosting Machine (LightGBM) algorithm, and the hyperparameters of LightGBM were optimized using genetic algorithm (GA), then a GA-LightGBM model of rock mass grades identification was established. Results: The accuracy of the GA-LightGBM recognition model reached 93.5%, which was higher than that of the support vector machine model and the random forest model. The model training speed is 8 times faster than the gradient boosting decision tree algorithm. Five TBM tunneling parameters were related to rock strength and rock mass integrity, and the total propulsion force could be used as the main criterion for sensing rock mass characteristics. The study provides an efficient method for analyzing TBM excavation parameters and accurately identifying rock mass grades, providing support for rapid on-site perception of rock mass grades and real-time adjustment of operating parameters.

Keywords: TBM; identification model of rock mass grade; genetic algorithm; LightGBM algorithm

在隧道、竖井等地下工程建设中, 与钻爆法相比, 机械破岩技术更加安全、环保, 近年来受到高度重视, 以 TBM 为代表的机械开挖方法在地下工程中得到广泛应用。

掘进机破岩是刀盘与岩体间相互作用过程^[1],

掘进机破岩参数与所破岩体的强度、完整性之间 (通常用岩体等级表示) 存在复杂对应关系, 如能建立这种关系模型将有助于现场根据掘进数据快速识别岩体等级^[2], 及时准确获取地质特征, 并据此实时调整掘进机参数, 达到提高掘进效率、降低成本和

收稿日期: 2024-04-23; 录用日期: 2024-08-03; 网络首发日期: 2025-07-08

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20250708.0857.002>

基金项目: 国家重点研发计划专项 (2021YFB3401501)

作者简介: 曹贞洋 (1996—), 男, 博士研究生; 龚敏 (1963—), 男, 教授, 博士生导师

通信作者: 龚敏, gongmustb@163.com

实现安全施工的目的^[3]。因此确立掘进参数与岩体参数的关系对 TBM 破岩高效化、智能化均具有较为重要的意义。

利用人工智能算法构建掘进参数与岩体参数的关系模型并进行岩体识别是目前的研究热点^[4-5]。Shi 等^[6]利用模糊 C 均值算法辨识了现场黏性土、砾质砂、岩石等地质条件,Liu 等^[7]基于 BP 神经网络和模拟退火算法利用 TBM 掘进参数预测了抗压强度等岩体参数,朱梦琦等^[8]、Liu 等^[9]基于集成 CART 算法结合 TBM 作业参数预测了围岩等级。

但目前研究尚存在有待改进的两个方面:首先是对岩体等级的划分较为粗糙,同等级岩体的特性也存在较大差异;另一方面,面对现场海量数据,模型快速识别岩体等级、实时指导现场作业尤为重要^[10],上述研究中样本量较少,尚未面临如何提高模型构建速度这一难题。

为此,本文提出采用 BQ 法和 RQD 值两步分级的方法划分岩体等级,利用遗传算法(GA)^[11-12]和 LightGBM 算法^[13-15]处理现场采集的数据,构建反映掘进参数与岩体等级关系的岩体等级识别模型,并在湖北某隧道 TBM 开挖工程中开展了应用。采用上述方法可提高岩体等级识别模型的构建效率和泛化性能,实现对掘进工作面岩体条件的快速感知,为隧道 TBM 实时调整掘进参数、提高掘进效率提供支撑。

1 研究方法及原理

1.1 总体思路

为高效分析现场大规模数据、快速构建掘进参数与岩体等级的关系模型,并根据掘进参数准确识别岩体等级,本文基于 GA-LightGBM 算法,开展掘进参数与岩体等级关系研究,构建岩体等级识别模型方法,研究路线如图 1 所示,主要包括数据采集、数据预处理、模型构建和结果评估共 4 个方面。

数据采集:在现场实验区段勘测岩石抗压强度、岩体完整性(含 RQD)等特性参数,并采集 TBM 在不同岩体中掘进状态时的工作数据。

数据预处理:在岩体基本质量分级(BQ法)的基础上,采用 RQD 指标来细分岩体等级,同时分析数据之间的相关性和掘进数据的分布特征,摒除冗余特征,筛选出与岩体特性强相关的主要掘进参数。

模型构建:将 TBM 主要掘进参数作为输入特征、岩体等级作为输出变量构建模型,基于训练集数据应用遗传算法优化 LightGBM 的超参数,将最优组合代入 LightGBM 中训练岩体等级识别模型。

结果评估:利用测试集数据计算所得模型的识

别准确率,并将其识别准确率、训练速度与其他算法模型进行对比,评估模型的优越性;根据特征重要性,分析各掘进参数对模型识别岩体的贡献程度。

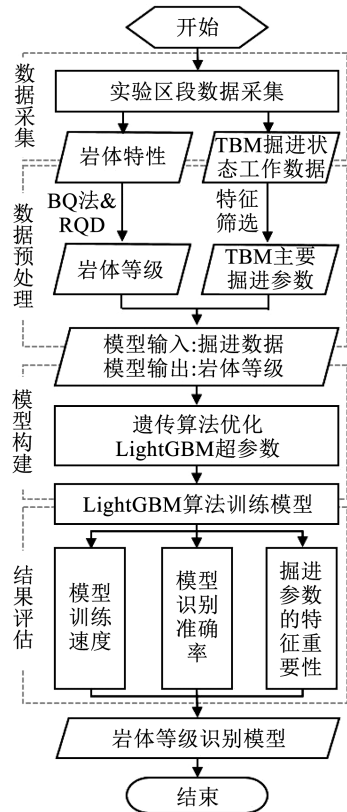


图 1 研究路线

Fig. 1 Research route

经过上述 4 个环节,最终得到准确率高、训练速度快的岩体等级识别模型,以下具体介绍各环节实现方法。

1.2 数据采集及预处理方法

通过掘进参数识别岩体等级,首先需要获取现场不同地段的岩体特征和相应的掘进参数两类数据。对于前者,由于现场勘探条件有限,掘进区段的岩体特征数据无法随时随地获取,只能选取部分区段进行勘测,根据勘测结果确定岩体等级;对于后者,TBM 传感系统实时采集了足够丰富的数据量,但参数种类较多,可能存在不相关数据和冗余数据,需在模型训练前进行参数筛选。

为更精细地划分岩体等级,本文采用 BQ 法和 RQD 值两步分级的方法。工程上主要按照《工程岩体分级标准》(GB/T 50218—2014)中岩体基本分级方法(BQ法)来划分岩体等级,BQ 值计算公式为

$$BQ = 100 + 3R_c + 250K_v \quad (1)$$

式中:当 $R_c > 90K_v + 30$ 时,应以 $R_c = 90K_v + 30$ 和 K_v 代入计算 BQ 值;当 $K_v > 0.04R_c + 0.4$ 时,应以 $K_v = 0.04R_c + 0.4$ 和 R_c 代入计算 BQ 值。

岩体完整性的变化对 TBM 掘进效率影响较大^[16],然而同等级岩体的 K_v 值差别可能很大。以工程建设常见的 III 级岩体 ($BQ = 351 \sim 450$) 为例,III 级岩体的完整性可为较破碎 ($K_v = 0.35 \sim 0.55$)、较完整 ($K_v = >0.55 \sim 0.75$) 和完整 ($K_v > 0.75$),所对应 K_v 的取值范围较大。

此外,现场定量测定 K_v 的过程容易受到外界因素干扰^[17]。RQD 也反映了岩体结构面的发育程度^[18],其变化对 TBM 掘进效率有较大的影响^[19],在现场可根据岩芯较为直观地测定 RQD。

本文在划分岩体等级时,为提高岩体完整程度的权重,并兼顾 RQD 值对 TBM 掘进的影响,先按照 BQ 法初步划分岩体等级;当处于同一等级岩体的 RQD 值相差大于 25 时,则对该等级岩体进一步细分,按现场 RQD 值从大到小依次定义为 a、b、c 等多个等级。

在筛选 TBM 掘进参数时,通过分析参数间的相关性和样本的分布特征,确定构建岩体等级识别模型所需的主要掘进参数。首先剔除与所有岩体特征均为极弱相关(相关性系数 < 0.2) 的掘进参数,再利用掘进参数之间的相关性,从相关性极强(相关性系数 > 0.8) 的参数组中剔除冗余参数。

采用上述方法,采集并预处理现场数据,确定 TBM 主要掘进参数和相应的岩体等级,这些数据是构建岩体等级识别模型所需的。

1.3 模型构建方法

由于 TBM 运行参数与岩体特征之间的复杂关系难以用单个函数关系式表达,故采用人工智能算法,以掘进参数为输入特征、岩体等级为输出变量,拟合两者的关系,构建岩体等级识别模型,进而实现对现场岩体等级的准确、快速感知。

在处理大量数据多分类问题上,LightGBM 算法具有准确率高、运行速度快的优点,但该算法需要调节的超参数较多。超参数寻优是构建高性能模型的关键环节,网格搜索算法在搜索空间过大时,计算量急剧增大、寻优速度缓慢;随机搜索算法搜索速度快,但所得最优值具有极大的偶然性;遗传算法基于种群进化理论,从多个方向同时搜索最优值,具有良好的随机性、全局性和并行性,提高了超参数的搜索效率。

因此,本文结合 LightGBM 算法和遗传算法的优势,构建岩体识别模型。由于 LightGBM 是以梯度提升决策树算法(GBDT)为基础,故依次对 GBDT 算法、LightGBM 以及遗传算法进行简要介绍。

1.3.1 梯度提升决策树算法(GBDT)

GBDT 算法是一种以分类回归树(CART 树)作

为基学习器的集成算法,其通过多次迭代产生 CART 树,每次迭代都旨在减小模型拟合值与真实值的差距,以准确预测样本类别。

以三分类问题为例,训练样本数据 $\{X, Y\}$ 中 X 为输入特征, y 为样本类别,利用 one-hot 编码将 y 转换为概率数组,若某样本实际类别 y 为第 2 类,则其实际概率数组 Y 可表示为 $[0, 1, 0]$ 。GBDT 算法每次迭代均会基于损失误差最小化,训练出三棵独立的分类回归树,计算各样本属于每种类别的拟合概率值,进而得到拟合概率数组为

$$T_i(X) = [p_i^1(X), p_i^2(X), p_i^3(X)] \quad (2)$$

式中: $T_i(X)$ 表示第 i 次迭代的拟合概率数组; $p_i^1(X)$ 、 $p_i^2(X)$ 、 $p_i^3(X)$ 表示样本属于各类别的拟合概率。

如式(3)所示,在进行 n 次迭代后,根据实际概率数组 Y 与前 n 次迭代的拟合值计算出负梯度 $r_i(X)$ 。若 $r_i(X)$ 在误差允许范围内,说明模型拟合效果很好,可根据式(4)计算 n 次迭代的拟合概率数组之和 $F(X)$,作为最终结果输出。

$$r_i(X) = Y - \sum_{i=1}^n T_i(X) \quad (3)$$

$$F(X) = \sum_{i=1}^n T_i(X) \quad (4)$$

按照上述流程训练出识别模型后,向模型输入任意一组新数据的特征值 X_0 ,均会输出一个概率数组 $F(X_0) = [p^1(X_0), p^2(X_0), p^3(X_0)]$,将数组中概率最大的类别作为数据的分类结果。

1.3.2 LightGBM 算法

基于 GBDT 算法,LightGBM 采用了基于梯度的单边采样算法,保留对信息增益贡献值更大的大梯度样本,并随机抛弃部分小梯度样本以精简数据集;基于模型特征的空间稀疏性,将很少同时取零的特征作为互斥特征捆绑合并,从而降低特征维度,提升训练速度。

另外,在训练学习过程中,LightGBM 采用直方图算法^[20]降低了节点分裂过程的运行消耗;并采用按叶子生长策略,减少了极小增益节点的不必要分裂,加快了模型计算速度。

1.3.3 遗传算法

遗传算法(genetic algorithm)作为一种随机全局搜索优化算法,使用概率机制展开迭代而不受目标函数梯度的约束,只要设置好种群大小、繁殖代数、交叉概率和变异概率等参数,即可全局自动寻优,具有较高搜索效率。

图 2 为遗传算法流程图,算法在参数初始化后形成含有 m 条染色体(即 m 组参数)的种群,通过

计算每条染色体的适应度大小, 评估每组参数值的优劣。通过选择、交叉、变异等操作产生新一代种群, 循环 K 次后将种群中适应度最高的个体作为最优解。

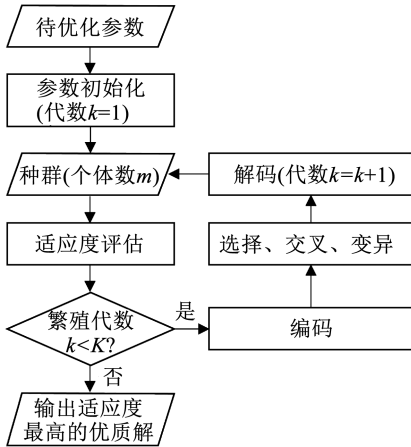


图 2 遗传算法流程图

Fig. 2 Flowchart of Genetic Algorithm

该组参数越好。对于本文所求解的分类模型超参数寻优问题, 可直接将模型在验证集上的分类准确率 (Acc) 作为适应度 f 。

1.3.4 五折交叉验证法

为提高上述 Acc 数值的可信度, 在求解这一指标时可采用交叉验证法, 本文采用的是常用的五折交叉验证。将数据集分为 5 个大小相似的互斥子集, 且每个子集类别占比相似, 依次选择 1 个子集作为验证集、其余 4 个子集作为训练集, 进行 5 次训练和验证。计算 5 次验证集准确率的均值, 作为当前个体的适应度, 即

$$f = \text{Acc} = \frac{1}{5} \sum_{r=1}^5 \frac{N_{r1}}{N_r} \quad (5)$$

式中: N_{r1} 为在第 r 折验证时验证集中得到正确分类的样本数量, N_r 为第 r 折时验证集样本总数。

1.3.5 GA-LightGBM 模型

结合上述相关算法及方法, 提出 GA-lightGBM 模型的构建方法 (图 3), 主要包括以下 8 个步骤:

适应度是种群更新的重要指标, 值越大代表着

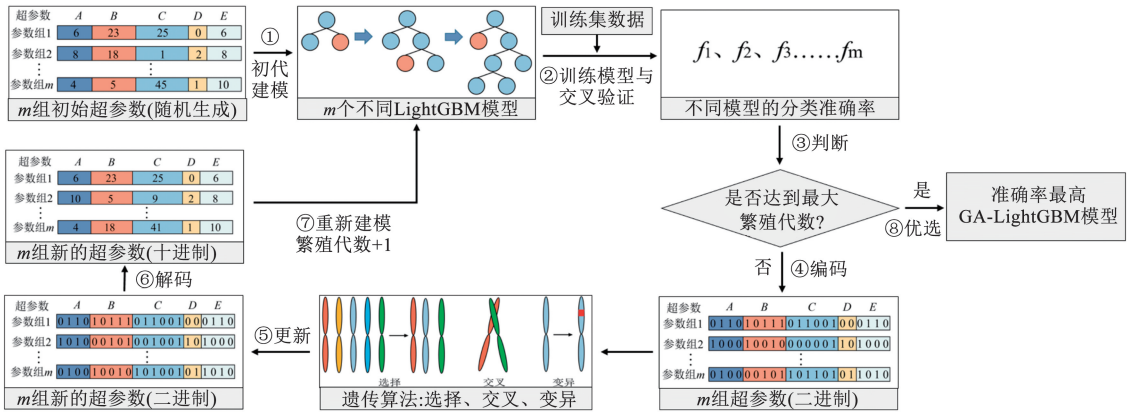


图 3 GA-LightGBM 模型构建方法及流程

Fig. 3 Construction method and flow of GA-LightGBM model

1) 初代建模。利用随机数函数, 随机生成 m 组 LightGBM 模型所需要的超参数 (十进制); 超参数类型包括学习率、迭代次数、树模型的最大深度等, 据此初步构建 m 个不同的 LightGBM 模型。此时的模型繁殖代数为 1。

2) 训练模型与交叉验证。将训练集数据代入到上述 m 个 LightGBM 模型中, 进行分别训练, 并利用五折交叉验证法, 获得各模型分类准确率。

3) 判断。步骤 2 的循环次数等于遗传算法的繁殖代数, 判断繁殖代数是否达到设定的最大值, 若为是, 进行步骤 8; 若为否, 进行步骤 2~7。

4) 编码。对 m 组超参数进行二进制编码, 得到 m 组链状数据, 用于后续交叉、变异操作。

5) 更新超参数。基于模型准确率大小, 保留准确率高的超参数组, 剔除准确率低的超参数组; 利用

遗传算法对各组超参数进行交叉、变异操作, 得到 m 组新的超参数 (二进制)。

6) 解码。将 m 组新的二进制超参数解码为十进制形式。

7) 重新建模。利用 m 组新的超参数重新构建 LightGBM 模型, 并再次进行步骤 2 的模型训练和步骤 3 的判断流程。

8) 优选。当繁殖代数达到预设最大值时, 从当前的 m 个模型中优选出分类准确率最高的模型, 作为最终的 GA-LightGBM 模型。

1.4 结果评估指标

将模型对各级岩体的分类准确率和模型的训练速度作为指标, 评估基于 GA-LightGBM 模型能否准确、快速地反映掘进参数与岩体等级之间的关系。

特征重要性表示了输入特征与模型的相关程

度,本文在模型构建成功后,根据特征重要性评估各输入特征(掘进参数)对模型预测结果的贡献大小。LightGBM 算法利用各个特征在所有基决策树上的信息增益(gain)之和表示该特征的重要性评分^[21],为便于统计和对比,本文对所有特征的重要性评分进行归一化处理。

2 工程应用

2.1 工程背景

本研究依托于湖北省十堰市水资源配置隧道工程四标段项目,该工程项目总线长 26 km,岩石以中硬岩和坚硬岩为主,部分区域存在软岩。图 4 为四标段项目各桩号段的主要岩性,以及地勘阶段的岩体分级情况。坚硬岩区段岩体完整,地勘结果为 II 级岩体;中硬岩区段岩体较完整,地勘结果为 III 级和 IV 级岩体;软岩区域为断层破碎带,地勘结果为 V 级岩体。现场对 IV 级与 V 级岩体采用钻爆法开挖,对 II、III 级岩体采用 TBM 机械开挖。

在 TBM 掘进施工段,岩石成分包括石英片岩、花岗岩、花岗斑岩、闪长岩。为保障施工安全高效、最大程度发挥 TBM 的掘进优势,需根据工作面岩体

情况不断调整作业参数。

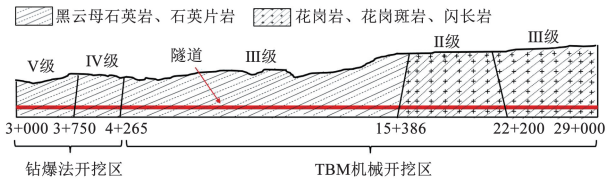


图 4 隧道各区段岩性及分级情况

Fig. 4 Lithology and classification of each section of the tunnel

实现上述目标的关键,在于识别动态工作面的岩体特征。利用长期地质勘探获取所有区段的岩体特征,在经济成本和时间效益上均难以接受。因此,提出基于现场 TBM 掘进数据感知岩体等级的方法。

2.2 现场数据采集与预处理

现场选取 4 个区段,利用地质钻探勘测了岩体特征(表 1),同时通过机载传感器采集到掘进参数的大量数据。采用 1.2 节所述 BQ 和 RQD 值两步分级法,将 4 个区段岩体划分为 3 个等级:II 级岩体(桩号 16+000—16+200 段、23+000—23+200 段)、III a 级岩体(桩号 10+000—10+200 段)、III b 级岩体(桩号 6+000—6+200 段)。

表 1 隧道掘进区段岩体特征及等级

Tab. 1 Rock mass characteristics and category in tunnelling sections

区段桩号	岩石成分	单轴抗压强度	单轴抗拉强度	完整性系数	岩石质量	岩体等级
		R_c /MPa	R_m /MPa	K_v	指标 RQD	
6+000—6+200	黑云母石英片岩、二云石英片岩	60	2.35	0.73	24.0	III b 级
10+000—10+200	奥陶系黑云母、石英片岩	50	2.64	0.81	53.1	III a 级
16+000—16+200	变质花岗岩、花岗斑岩	75	3.37	0.80	78.9	II 级
23+000—23+200	变质斜长花岗岩、花岗斑岩	75	3.37	0.80	78.9	II 级

TBM 机载传感器的采样间隔为 1 s,数据包括停机、掘进和回退 3 种状态。删除 TBM 停机状态(推进速度为 0)、回退状态(总推进力为 0)的数据,保留正常掘进的数据。为避免数据偶然波动的影响,采用平均值采样法,将每分钟 60 组数据的平均值作为该时段的有效值。按上述预处理数据后,得到有效掘进数据共 44 131 组,II 级、III a 级、III b 级围岩掘进数据分别有 23 898 组、9 126 组和 11 107 组。图 5 为 TBM 在 3 种不同岩体中掘进时,部分参数数据变化的示例曲线。由图 5 初步可知,岩体条件不同时,掘进数据的变化存在着一定差别。

在现场采集的 TBM 数据中,参与机-岩作用的掘进参数有 9 种,分别为推进速度、刀盘转速、总推进力、刀盘转矩、贯入度、推进位移、刀盘平均电流、推进缸无杆腔压力、推进泵压力。为降低数据冗余,更加简明、清晰地反映岩体特征与掘进参数的关系,

在模型训练前先对掘进参数、岩体参数进行 Spearman 相关性分析,根据相关性结果筛选参数。

参与相关性分析的岩体参数包括岩石强度、完整性系数 K_v 、RQD、岩体渗水率和岩石软化系数。其中,岩体渗水率和岩石软化系数是和项目场地条件密切相关的参数。基于本文依托的水资源配置隧道工程的数据,计算岩体参数、TBM 掘进参数的相关性情况如图 6~8 所示。

图 6 为该隧道工程中各岩体参数的相关性情况。岩体渗水率与岩石抗压强度、RQD 值成明显的负相关关系,抗压强度越大、RQD 值越大,意味着岩体越坚固、完整,渗水率越小。岩石软化系数是岩石在饱和状态和干燥状态的单轴抗压强度之比,因此其与岩石抗压强度有极强的相关性。完整性系数 K_v 和 RQD 值都表征着岩体的完整程度,两者同样具有相关性。

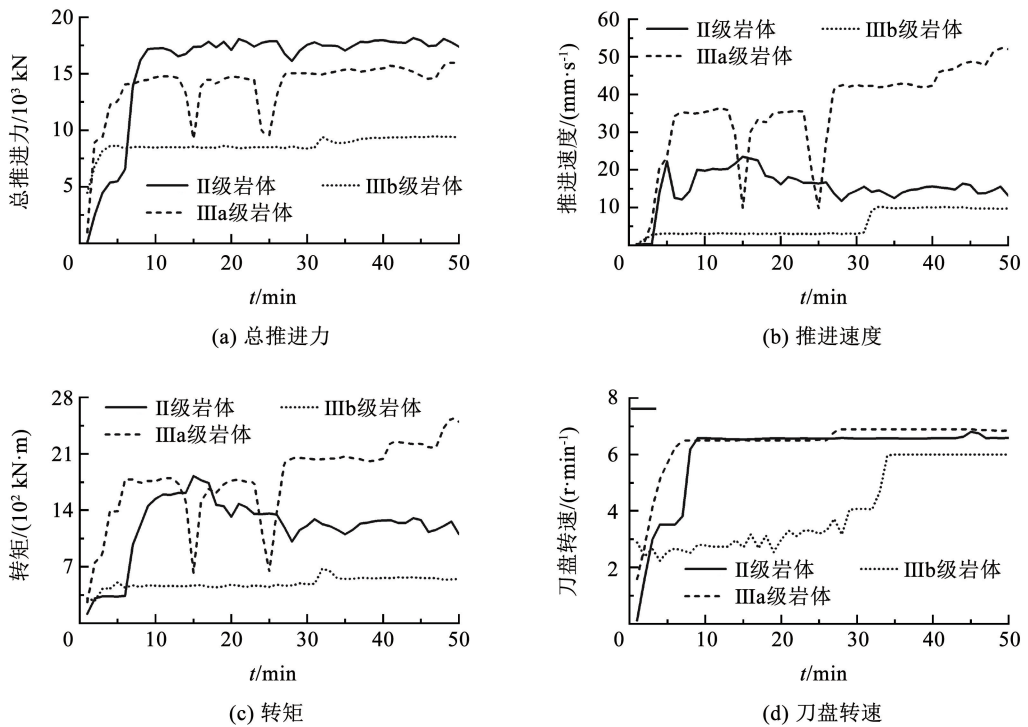


图5 TBM 正常掘进时部分参数数据变化曲线

Fig. 5 Change curve of some parameters of TBM during normal driving

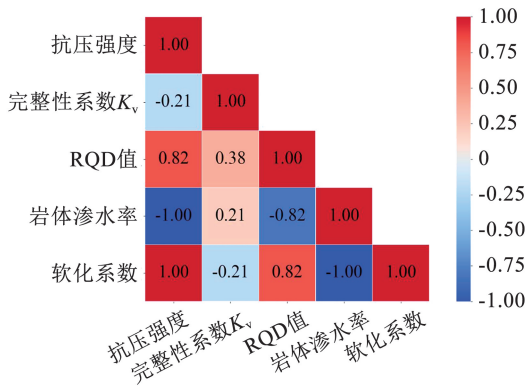


图6 岩体参数间的相关性系数

Fig. 6 Correlation coefficients of rock mass parameters

图7为TBM掘进参数与岩体参数的相关性情况。其中,抗压强度与刀盘转速、总推进力等多个掘进参数的相关性较好;与 K_v 的相关性系数大于0.4的掘进参数仅有刀盘转矩;总推进力、推进缸无杆腔压力与RQD值成极强相关(相关性系数 >0.8);岩体渗水率、岩石软化系数与各掘进参数的相关性情况,和抗压强度一致。大部分掘进参数均和岩体参数具有较好相关性,但刀盘平均电流仅与 K_v 相关性系数大于0.2,与其他岩体参数的相关性系数为0.1左右。

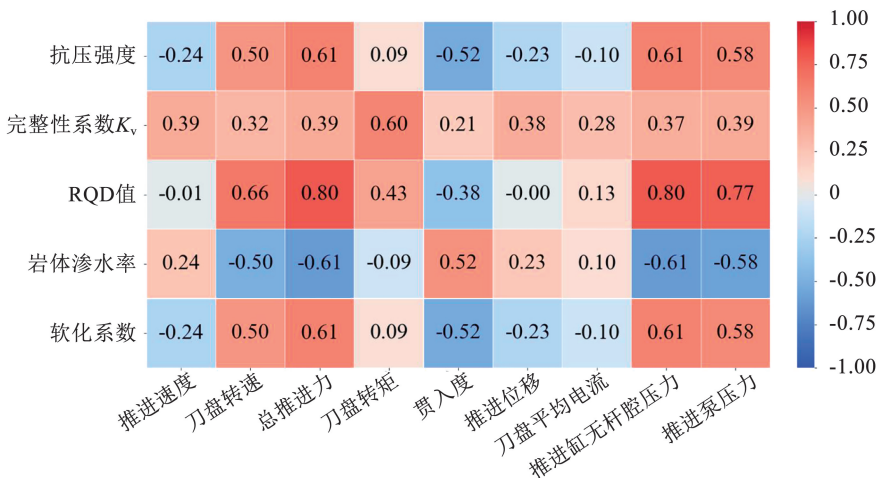


图7 掘进参数与岩体参数的相关性系数

Fig. 7 Correlation coefficients between tunneling parameters and rock mass parameters

图 8 为各掘进参数之间的相关性情况,部分掘进参数之间存在极强相关性(相关性系数 > 0.8),具体有:推进速度与推进位移;总推进力、推进缸的无杆腔压力与推进泵压力;刀盘扭矩与刀盘平均电流。这表明掘进参数存在着冗余。

结合图 7、8 各参数相关性情况,剔除与岩体特征极弱相关的参数和冗余参数,从 9 种掘进参数中筛选出总推进力、刀盘扭矩、刀盘转速、推进速度、贯入度作为 TBM 主要掘进参数,图 9 ~ 13 为 5 种参数在不同岩体条件下的样本分布情况。

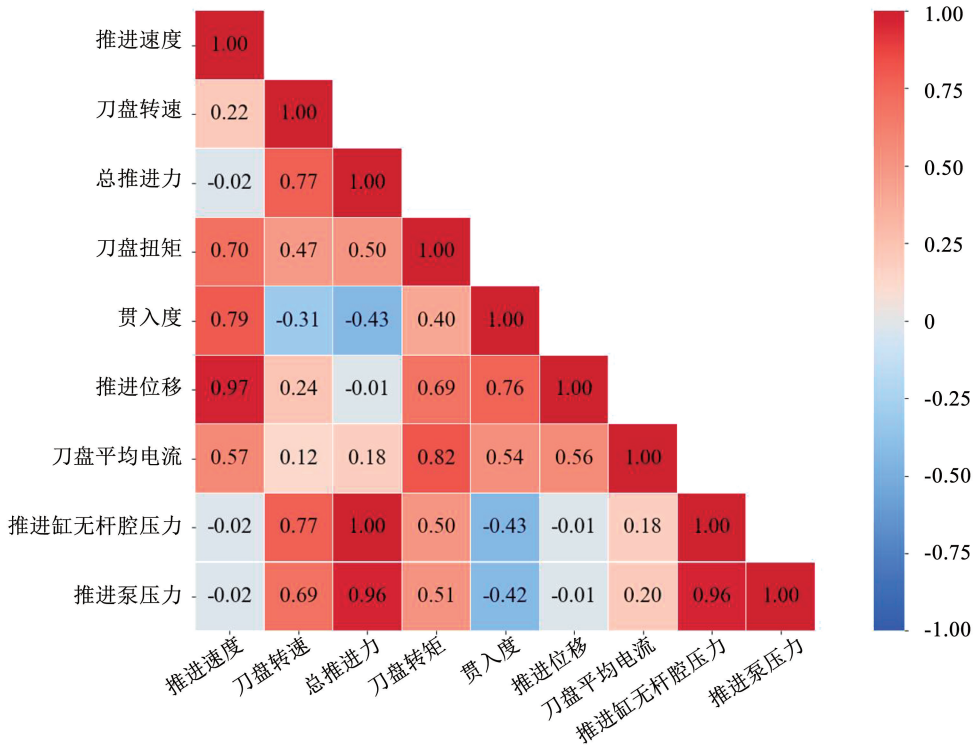


图 8 掘进参数间的相关性系数

Fig. 8 Correlation coefficients of tunneling parameters

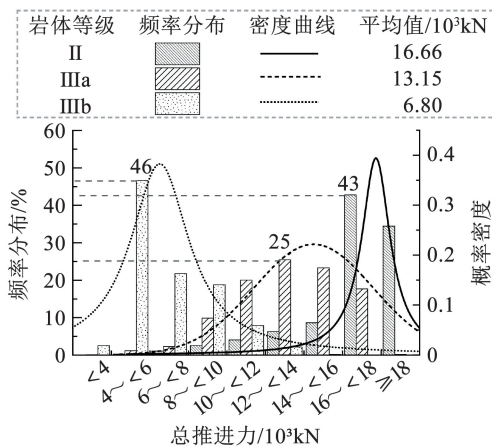


图 9 总推进力分布及其概率密度曲线

Fig. 9 Total propulsive force distribution and probability density curve

由图 9 可知,总推进力随岩体质量改善有明显的增大趋势。刀盘扭矩和刀盘转速作为滚刀作用于岩体的重要参数,结合图 7 的相关性系数分析图 10、图 11 发现,刀盘扭矩的分布受 K_v 值的影响较大,而刀盘转速则主要受 RQD 值的影响;刀盘扭矩和刀盘转速在 IIIb 级岩体 ($K_v = 0.73$, RQD = 24.0) 的主要

取值区间和平均值要远小于 II 级岩体 ($K_v = 0.80$, RQD = 78.9)、IIIa 级岩体 ($K_v = 0.81$, RQD = 53.1)。

推进速度和贯入度表征着 TBM 的掘进效率,由图 12 和图 13 可知,推进速度在 K_v 较小的 IIIb 级岩体中平均值较大;贯入度在岩石强度较低的 IIIa 级岩体 ($R_c = 50$ MPa) 中平均值较大。

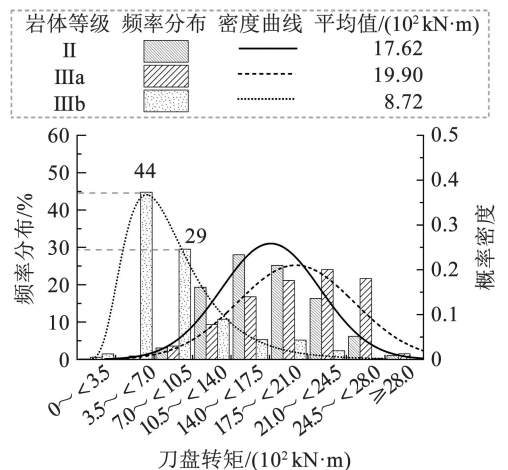


图 10 刀盘扭矩分布及其概率密度曲线

Fig. 10 Cutter head torque distribution and probability density curve

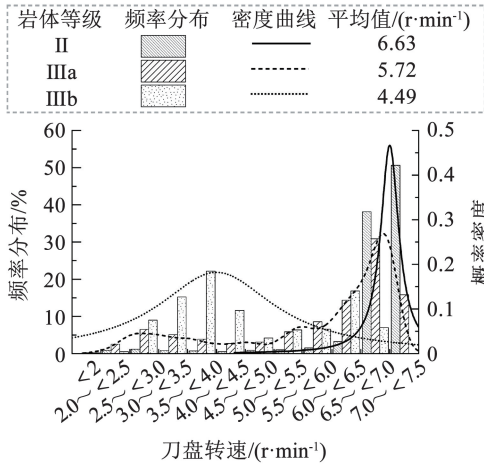


图 11 刀盘转速分布及其概率密度曲线

Fig. 11 Cutter head speed distribution and probability density curve

合上述掘进参数与岩体等级的关系,本文采用了人工智能算法构建岩体等级识别模型。

2.3 岩体等级识别模型的构建

2.3.1 数据集的划分

为满足岩体等级识别模型构建和评估的需求,本文将各类岩体数据按 8:2 的比例划分(表 2)。其中,80% 数据在模型构建时会按照五折交叉验证法再次划分为训练集和验证集,用于训练模型和超参数寻优;20% 数据作为测试集,不参与模型构建和训练过程,只用于评估最终所得模型的识别性能。

表 2 训练集与测试集的数据分配

Tab. 2 Data allocation of training set and test set

岩体等级	数据总量	训练集 + 验证集	测试集
II	23 898	19 118	4 780
III a	9 126	7 300	1 826
III b	11 107	8 885	2 222

2.3.2 GA-LightGBM 模型训练与超参数寻优

将总推进力、刀盘转矩、刀盘转速、掘进速率和贯入度作为模型训练中的输入特征,岩体等级作为模型训练的输出结果,根据图 3 的方法与流程构建岩体等级识别模型。

在模型训练过程中选取合适的超参数是提高模型性能的重要途径,如果超参数值设置不当,模型很难有较好的识别效果。LightGBM 模型需调节的超参数及其搜索范围见表 3。

表 3 模型的超参数及搜索范围

Tab. 3 The hyperparameters of model and the search scope

超参数名称	含义	搜索范围	采样间隔
learning_rate	学习率	[0.02,0.6]	0.02
n_estimators	迭代次数	[5,300]	5
max_depth	树模型的最大深度	[2,9]	1
num_leaves	单棵树上叶子节点数	[2,500]	1
min_child_sample	每个叶子节点最小样本量	[0,1 000]	10
bagging_fraction	每次迭代选择样本的占比	[0.7,1]	0.1
feature_fraction	每次迭代选择特征的占比	[0.7,1]	0.1
reg_alpha	L1 正则化参数	[0,1]	0.1
reg_lambda	L2 正则化参数	[0,10]	0.5
cat_smooth	降低噪声影响的系数	[1,30]	1

按照表 3 中的搜索范围和采样间隔,若采用网格搜索算法进行超参数寻优,大约需进行 8×10^{13} 次搜索,计算量巨大。为提高搜索效率,采用遗传算法处理 LightGBM 的超参数设置问题。设置遗传算法的种群大小为 100,并且随机产生初始种群;繁殖代数为 80 代,每代繁殖采用轮盘赌法选择出新种群;

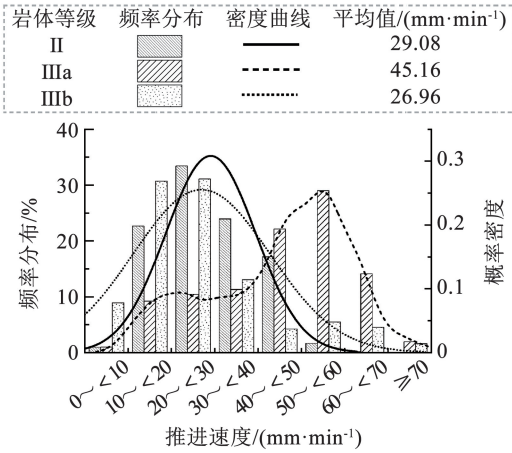


图 12 推进速度分布及其概率密度曲线

Fig. 12 Propulsion speed distribution and probability density curve

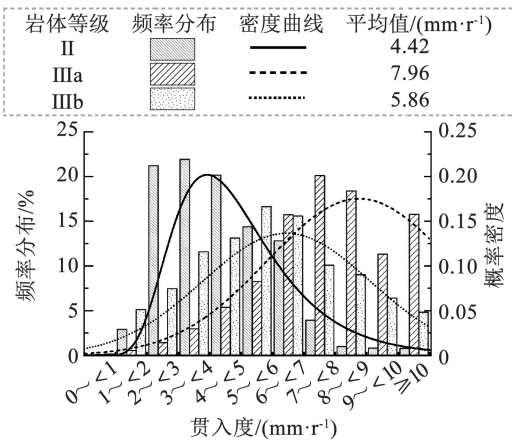


图 13 贯入度分布及其概率密度曲线

Fig. 13 Penetration distribution and probability density curve

综上分析,总推进力、刀盘转矩、刀盘转速、推进速度和贯入度等掘进参数与岩石强度、岩体完整性系数 K_v 、岩石基本质量指标 RQD 值等存在相关关系,不同岩体的掘进参数数据分布存在差异。为拟

种群中个体交叉概率为 0.4、变异概率为 0.05。

值得注意的是,超参数 max_depth 与 num_leaves 需满足式(6)的关系,因此在每次随机生成 max_depth 值后,需根据 max_depth 的取值,限定 num_leaves 的随机取值范围。同时,为防止交叉、变异操作对 max_depth 和 num_leaves 的影响,在解码过程中,若 num_leaves 超过式(6)中最大值,则将最大值直接赋给 num_leaves,以保证程序正常运行。

$$D_{\max} \leq 2^{N_1} - 1 \quad (6)$$

式中: D_{\max} 为 LightGBM 模型中决策树的最大深度(max_depth), N_1 为 LightGBM 中决策树叶子节点数。

图 14 为模型在验证集上的识别准确率随遗传算法繁殖代数变化情况,随着繁殖代数增加,识别准确率不断提高,意味着模型超参数得到了不断优化。经过 80 代繁殖,验证集的识别准确率达到 93.69%,对应各超参数的最优组合如表 4 所示。按照表 4 设置 LightGBM 的各超参数数值,训练并构建出最终的岩体等级识别模型(GA-LightGBM 模型)。

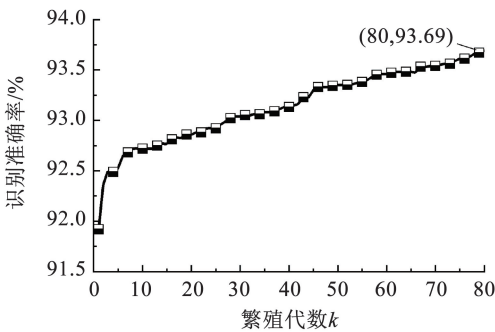


图 14 验证集识别准确率变化曲线

Fig. 14 Recognition accuracy change curve of validation set

表 4 模型超参数的最优组合

Tab. 4 The optimal combination of model hyperparameters

超参数	最优值	超参数	最优值
learning_rate	0.56	bagging_fraction	0.9
n_estimators	270	feature_fraction	0.7
max_depth	9	reg_alpha	0.9
num_leaves	464	reg_lambda	0
min_child_sample	10	cat_smooth	25

2.4 GA-LightGBM 模型性能评估及特征重要性

图 15 为测试集数据在 GA-LightGBM 模型中预测结果的混淆矩阵,各类别大部分样本均能得到准确预测,模型识别出的岩体等级与地勘结果吻合性较好。如图 15 所示,本文在地勘结果为 II 级的岩体

中,抽取了 4 761 组掘进参数样本,将其代入 GA-LightGBM 模型中,识别出岩体等级为 II 级的样本有 4 510 组,准确率达到 94.7%;对于地勘结果为 III 级的岩体,在利用 BQ 法和 RQD 值两步法、细分为 III a 级和 III b 级后,GA-LightGBM 模型同样可以较好的区分出 III a 级和 III b 级岩体;III a 级、III b 级的测试样本数共有 4 066 组,被误识别为 II 级的样本仅有 250 组。

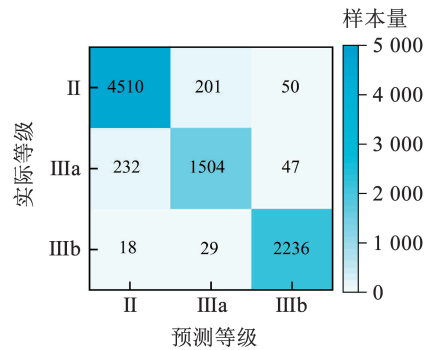


图 15 GA-LightGBM 模型预测结果混淆矩阵

Fig. 15 Confusion matrix of GA-LightGBM model's prediction

统计本文模型与支持向量机、随机森林、GBDT 模型在各等级岩体中的识别效果,并计算 4 种模型在 CPU 为 i5-10210U、内存 8GB、GPU 为 MX250 的计算机上的训练耗时(表 5)。

表 5 不同模型的识别率

Tab. 5 The recognition accuracy of different models

模型	识别率/%				训练耗时/s
	II 级	III a 级	III b 级	整体	
SVM	91.4	74.3	93.8	88.6	34
RF	92.7	80.2	95.8	90.9	28
GBDT	94.4	83.6	98.0	93.2	175
GA-LightGBM	94.7	84.4	97.9	93.5	19

由表 5 可知,GA-LightGBM 模型的整体识别准确率达到 93.5%,高于 SVM 模型和 RF 模型,与 GBDT 模型相近;其训练速度比 GBDT 模型提高了 8 倍。GA-LightGBM 模型在识别准确率、运行速度上均有更好的性能,得到了现场设计及施工单位的肯定。

图 16 为 5 种掘进参数在 GA-LightGBM 模型构建中的特征重要性。对比各参数的重要性评分,总推进力的重要程度远大于其他 4 种参数,这表明在利用掘进参数感知岩体等级时,总推进力可作为主要判据,应重点关注该参数的数据变化。

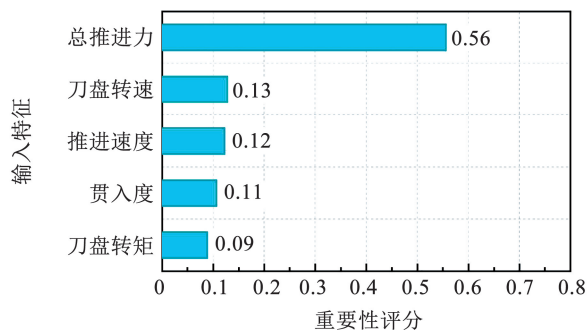


图 16 GA-LightGBM 模型各输入特征重要性评分

Fig. 16 Importance scoring of input characteristics in GA-LightGBM model

3 讨论

随着隧道建设向深地进军, TBM 在掘进作业中会遇到高地应力或大裂隙区等复杂的地质环境。开展复杂地质条件对 TBM 掘进参数的影响研究, 对 TBM 进行深地掘进具有重要意义。

根据工程地勘报告显示, 本文研究中数据采集所涉及到的 4 个区段(6 + 000—6 + 200、10 + 000—10 + 200、16 + 000—16 + 200、23 + 000—23 + 200), 其初始应力场的应力值分别为 1.03 ~ 4.87 MPa、2.4 ~ 5.0 MPa、10.5 MPa 和 19.0 MPa, 岩石强度应力比分别为 9.65、10.4、6.4 和 3.1, 应力量级属于低 ~ 高地应力。因此, 所构建的 GA-LightGBM 模型适用于高地应力地区。

需要指出的是, 由于本次研究的数据样本有限, 未能深入分析复杂环境(包括高地应力、大断层破碎带)对 TBM 掘进参数的影响, 高地应力、大断层等复杂地质条件与 TBM 掘进效率的具体关系有待进一步的研究。

4 结论

本文依托于湖北省十堰市水资源配置隧道工程项目, 开展了 TBM 掘进参数与岩体等级关系研究, 得到以下结论:

1) 利用 GA-LightGBM 算法研究掘进参数与岩体特征关系, 构建了岩体等级识别模型, 实现了对掘进工作面岩体等级的准确、快速感知。GA-LightGBM 模型的整体识别率达到了 93.5%, 高于 SVM、RF 模型, 训练速度比 GBDT 模型提高了 8 倍。

2) 通过分析不同岩体条件下掘进参数分布特征和构建岩体等级识别模型, 确定了掘进机的总推进力、刀盘转矩、刀盘转速、推进速度和贯入度等掘进参数与岩石强度、岩体完整程度等特性存在相关

关系, 证明了可通过现场掘进参数实时感知岩体特征。

3) 对比各掘进参数在 GA-LightGBM 模型构建中的重要性评分, 确定了总推进力是基于掘进参数感知岩体等级过程中的主要判据。

4) 本文提出的基于 GA-LightGBM 算法高效分析 TBM 掘进参数并准确识别岩体等级的方法, 为实时感知岩体特征进而调整掘进参数、提高掘进效率提供支撑, 对于 TBM 破岩高效化、智能化具有积极意义。

参考文献

- [1] 许宇, 李兴高, 杨益, 等. 盾构切刀切削混凝土过程中的动态响应试验[J]. 哈尔滨工业大学学报, 2021, 53(5): 182
XU Yu, LI Xinggao, YANG Yi, et al. Dynamic response mechanism of shield cutter in concrete cutting[J]. Journal of Harbin Institute of Technology, 2021, 53(5): 182
- [2] 吴鑫林, 张晓平, 刘泉声, 等. TBM 岩体可掘性预测及其分级研究[J]. 岩土力学, 2020, 41(5): 1721
WU Xinlin, ZHANG Xiaoping, LIU Quansheng, et al. Prediction and classification of rock mass boreability in TBM tunnel[J]. Rock and Soil Mechanics, 2020, 41(5): 1721
- [3] 张海龙, 汤杨, 任汀, 等. 复合地层隧道围岩强度实时估算研究[J]. 岩土力学, 2022, 43(7): 1877
ZHANG Hailong, TANG Yang, REN Ting, et al. Real-time estimation of surrounding rock strength in composite stratatunnels[J]. Rock and Soil Mechanics, 2022, 43(7): 1877
- [4] LIU Bin, WANG Ruitui, GUAN Zengda, et al. Improved support vector regression models for predicting rock mass parameters using tunnel boring machine driving data[J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2019, 91: 102958
- [5] LIU Zaobao, LI Long, FANG Xingli, et al. Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network[J]. Automation in Construction, 2021, 125: 103647
- [6] SHI Maolin, ZHANG Liyong, SUN Wei, et al. A fuzzy c-means algorithm guided by attribute correlations and its application in the big data analysis of tunnel boring machine[J]. Knowledge-Based Systems, 2019, 182(C): 104859
- [7] LIU B, WANG R, ZHAO G, et al. Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm[J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2020, 95(C): 103103
- [8] 朱梦琦, 朱合华, 王昕, 等. 基于集成 CART 算法的 TBM 掘进参数与围岩等级预测[J]. 岩石力学与工程学报, 2020, 39(9): 1860
ZHU Mengqi, ZHU Hehua, WANG Xin, et al. Study on CART-based ensemble learning algorithms for predicting TBM tunneling

- parameters and classing surrounding rockmasses[J]. Chinese Journal of Rock Mechanics and Engineering, 2020, 39(9): 1860
- [9] LIU Quansheng, WANG Xinyu, HUANG Xing, et al. Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data [J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2020, 106: 103595
- [10] 袁亮, 张平松. TBM 施工岩巷掘探一体化技术研究进展与思考 [J]. 煤田地质与勘探, 2023, 51(1): 21
YUAN Liang, ZHANG Pinsong. Research progress and thinking on integrated tunneling and detection technology of rock roadway with TBM[J]. Coal Geology and Exploration, 2023, 51(1): 21
- [11] 崔学杰, 晏鄂川, 陈武. 基于改进遗传算法的岩体结构面产状聚类分析[J]. 岩土力学, 2019, 40(增刊 1): 374
CUI Xuejie, YAN Echuan, CHEN Wu. Cluster analysis of discontinuity occurrence of rock mass based on improved genetic algorithm[J]. Rock and Soil Mechanics, 2019, 40(Sup.1): 374
- [12] 朱云伟, 黄海松, 魏建安. 基于 GA-LightGBM 的刀具磨损状态在线识别[J]. 组合机床与自动化加工技术, 2021(10): 83
ZHU Yunwei, HUANG Haisong, WEI Jian'an. Online tool wear status identification based on GA-LightGBM[J]. Modular Machine Tool and Automatic Manufacturing Technique, 2021(10): 83
- [13] 王璞, 刘洋, 黄智仁. 一种轻量级梯度提升机的交通模式识别 [J]. 哈尔滨工业大学学报, 2019, 51(9): 96
WANG Pu, LIU Yang, HUANG Zhiren. Transportation modes recognition using a light gradient boosting machine[J]. Journal of Harbin Institute of Technology, 2019, 51(9): 96
- [14] 刘尧, 陈改革, 刘振国, 等. 关联规则挖掘驱动的盾构机刀盘健康评估方法[J]. 中国机械工程, 2023, 34(11): 1326
LIU Yao, CHEN Gaige, LIU Zhenguo, et al. Health assessment method of shield machine cutterheads driven by association rule mining[J]. China Mechanical Engineering, 2023, 34(11): 1326
- [15] 叶飞, 冯浩岚, 梁兴, 等. 基于 WOA-LightGBM 的盾构切削锚杆参数预测[J]. 同济大学学报(自然科学版), 2022, 50(12): 1761
YE Fei, FENG Haolan, LIANG Xing, et al. Prediction of shield performance in cutting anchor bolts based on WOA-LightGBM[J]. Journal of Tongji University (Natural Science), 2022, 50(12): 1761
- [16] 何发亮, 谷明成, 王石春. TBM 施工隧道围岩分级方法研究 [J]. 岩石力学与工程学报, 2002(9): 1350
HE Faliang, GU Mincheng, WANG Shichun. Study on classification method of surrounding rock of TBM construction tunnel [J]. Chinese Journal of Rock Mechanics and Engineering, 2002(9): 1350
- [17] 殷明伦, 张晋勋, 江玉生, 等. 节理数表征岩体完整系数的结构面类别修正研究[J]. 岩土力学, 2021, 42(4): 1133
YIN Minglun, ZHANG Jinxun, JIANG Yusheng, et al. Study of correction of the structural plane category based on the rock mass integrity coefficient characterized by the volumetric joint count[J]. Rock and Soil Mechanics, 2021, 42(4): 1133
- [18] 殷欣, 高峰, 刘泉声, 等. 面向隧道掘进机可掘性评价的多算法融合优化模型及其工程应用[J]. 岩石力学与工程学报, 2022, 41(增刊 1): 2757
YIN Xin, GAO Feng, LIU Quansheng, et al. Multi-algorithm fusion-optimization model and its engineering application for boreability evaluation of tunnel boring machine[J]. Chinese Journal of Rock Mechanics and Engineering, 2022, 41(Sup.1): 2757
- [19] 吴志军, 方立群, 翁磊, 等. 基于 TBM 掘进性能的岩体分级及可掘性等级感知识别方法[J]. 岩石力学与工程学报, 2022, 41(增刊 1): 2684
WU Zhijun, FANG Liqun, WENG Lei, et al. A classification and boreability perception and recognition method for rock mass based on TBM tunneling performance [J]. Chinese Journal of Rock Mechanics and Engineering, 2022, 41(Sup.1): 2684
- [20] KE Guolin, MENG Qi, FINLEY T, et al. LightGBM: a highly efficient gradient boosting decision tree [C]//31st International Conference on Neural Information Processing Systems (NIPS 2017). Red Hook: Curran Associates Inc., 2017: 3149
- [21] 董学超, 郭明伟, 王水林. 基于 LightGBM 的超大沉井下沉状态预测及传感器优化布置[J]. 岩土力学, 2023, 44(6): 1789
DONG Xuechao, GUO Mingwei, WANG Shuilin. Sinking state prediction and optimal sensor placement of super large open caissons based on LightGBM[J]. Rock and Soil Mechanics, 2023, 44(6): 1789