

doi:10.3969/j.issn.1007-7545.2024.02.013

基于机器学习方法预测地浸过程中 铀浸出金属量的变化

余东原¹, 罗跃^{1,2}, 梁大业³, 李立尧^{1,2}

1. 东华理工大学 水资源与环境工程学院, 南昌 330013;
2. 东华理工大学 核资源与环境重点实验室, 南昌 330013;
3. 中核内蒙古矿业有限公司, 呼和浩特 010010)

摘要:在地浸采铀过程中,准确预测铀浸出金属量具有重要意义。使用多元线性回归、多层感知机(MLP)和随机森林(RF)多种机器学习方法分别建立预测模型。结果表明:1)相比于传统的多元线性回归算法,MLP和RF两种方法能够得到预测性能更好的模型。2)多层感知机(MLP)模型在预测铀浸出金属量变化的上性能表现最佳($R^2=0.91$)。3)在相同的预测精度下,RF模型比MLP模型耗时更短,超参数设置更加简单。4)在以总流量和每平方米铀量为铀浸出金属量的关键因素进行分析时,总流量对铀浸出金属量的权重占比为81.6%。

关键词:地浸采铀;线性回归模型;机器学习预测

中图分类号:TL212.1⁺2 文献标志码:A 文章编号:1007-7545(2024)03-0092-07

Predicting the Variation of Uranium Leaching Metal Content in Ground-leaching Process Based on Machine Learning Methods

YU Dongyuan¹, LUO Yue^{1,2}, LIANG Daye³, LI Liyao^{1,2}

1. School of Water Resource and Environmental Engineering, East China University of Technology, Nanchang 330013, China;
2. Key Laboratory of Nuclear Resource and Environment, East China University of Technology, Nanchang 330013, China;
3. Inner Mongolia Mining Co., Ltd., CNNC, Huhhot 010010, China)

Abstract: In the process of in-situ leaching of uranium, it is of great significance to accurately predict the amount of uranium metal extracted. Multiple machine learning methods such as multiple linear regression, Multi-Layer Perceptron (MLP) and Random Forest (RF) were used to build prediction models. The results indicate that: 1) Compared with the traditional multiple linear regression algorithms, MLP and RF methods can obtain models with better prediction performance. 2) The MLP model has the best performance in predicting the change of uranium leaching metal content ($R^2=0.91$). 3) With the same prediction accuracy, the RF model takes less time than that of the MLP model, and the hyperparameters setting is simpler. 4) The weight ratio of the total flow rate to the amount of uranium metal leached is 81.6% when the total flow rate and the amount of uranium leaching metal content per square meter are taken as the key factors.

Key words: ground-leaching uranium mining; linear regression model; machine learning prediction

收稿日期:2023-10-23

基金项目:国家自然科学基金资助项目(42062017,42372295);江西省自然科学基金资助项目(20224BAB203038,20212BAB213006)

作者简介:余东原(1998-),男,硕士研究生;通信作者:罗跃(1985-),男,博士,副教授

铀是一种重要的放射性元素,可用于发电和其他核能应用,具有广泛的应用和意义。目前我国主要开发砂岩型铀矿,采用原地浸出方法开采(简称为地浸采铀)。该方法利用注液井向含矿含水层中注入浸出剂,创造有利于铀矿溶解和迁移的水化学环境,再通过抽液井抽取富含铀的地下水,实现原位回收铀。该过程涉及复杂的水文地球过程,铀的浸出金属量与铀品位、抽水量相关,但它们之间的量化关系较复杂。为了更好地规划和优化铀资源开采,减少采铀成本和对环境所造成的影响,提高铀浸出量,需要准确预测地浸采铀中铀浸出金属量的变化。

近年来,机器学习在各领域表现出良好的应用前景。机器学习结合了分类和回归的主要功能,提供了从杂乱、不规则和高维数据中挖掘隐藏模式或推断相关性的算法^[1-2]。通常,机器学习由三个关键过程组成,即训练数据(特征值)的输入、主题模型的构建和目标结果的输出。一旦训练模型通过验证,就可以在接收新的输入后快速预测相应的输出。相比于传统的数据处理方式,机器学习在效率、适应性、灵活性和拓展性上都具有极大的优势,成为了许多领域的关键技术。雷林等通过将神经网络与改进型遗传算法相结合,构建神经网络预测模型,开展反应溶质运移模拟(RTM)理论和应用研究,为地浸采铀工艺参数提供了优化建议^[3-4]。此外,PEETS等^[5]和JIA等^[6]结合机器学习算法对水和土壤中物质的毒性进行了预测。JIA^[7]等和YANG等^[8]则识别土壤重金属污染的潜在来源、污染的驱动因素和热点区域。REN等^[9]使用贝叶斯集成机器学习框架用于预测局部臭氧浓度。ZHANG等^[10]以黄钾铁矾的浸出过程为例,使用机器学习技术对矿物浸出工艺进行了优化。

因此,本文引入机器学习方法,通过将铀品位数据、抽液流量以及铀产出数据训练模型,得到一个可预测铀浸出金属量变化的模型^[11-23]。

1 研究方法

1.1 数据获取与处理

研究区域为内蒙某铀矿的一个采区,以五点式抽注单元布置钻孔(图1)。该采区共有39个抽液孔,抽液数据包括每天的流量数据和铀浸出数据,自2021年7月起,记录时长为一年。为消除数据特征之间的尺度差异,更好地展现特征之间的关系,提高数据的可解释性以及模型的稳定性,在建模之前对数据进行归一化处理,将数据转为无量纲数。

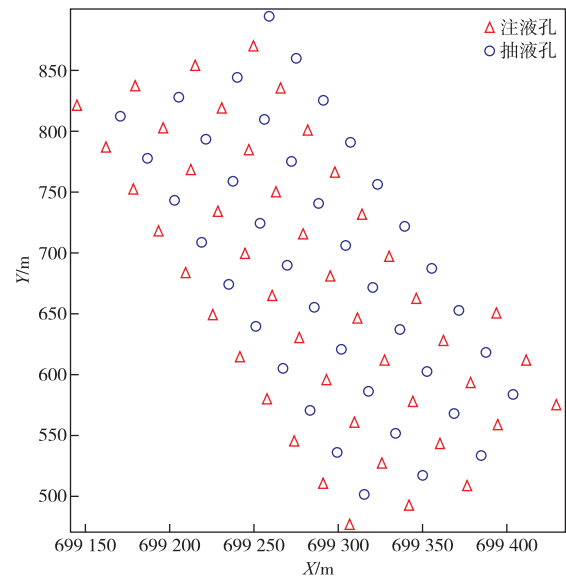


图1 抽注分布图

Fig. 1 Map of pumping and injection distribution

1.2 多元线性回归

使用多元线性回归预测分析时,使用普通最小二乘法和逐步回归获得最佳预测,并假设因变量和自变量之间存在线性关系。建立合适的模型来描述数据的相关性。

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (1)$$

式中, \hat{y} 表示因变量(要预测的值); x_1, x_2, \dots, x_n 表示自变量(用来预测 \hat{y} 的变量); $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ 表示回归系数(用来衡量自变量对因变量的影响)。

1.3 多层感知机(MLP)

多层感知机是一种基于前馈神经网络的监督学习模型通过多层神经元的组合和激活函数的非线性变换,可以学习到更复杂的特征表示和非线性关系。在训练过程中,神经网络通过计算损失函数对模型参数的梯度,然后沿着梯度的反方向更新参数,不断迭代优化模型。逐步调整权重和偏置,使得预测值与实际值之间的误差不断减小,最终得到较好的拟合效果。

预测表达式:

$$Z^{(1)} = \mathbf{W}^{(1)} \mathbf{X} + \mathbf{b}^{(1)}$$

$$a^{(1)} = f(Z^{(1)})$$

$$Z^{(2)} = \mathbf{W}^{(2)} a^{(1)} + \mathbf{b}^{(2)}$$

⋮

$$Z^{(L-1)} = \mathbf{W}^{(L-1)} a^{(L-2)} + \mathbf{b}^{(L-1)}$$

$$a^{(L-1)} = f(Z^{(L-1)})$$

$$Z^{(L)} = \mathbf{W}^{(L)} a^{(L-1)} + \mathbf{b}^{(L)}$$

$$\hat{y} = f(Z^{(L)}) \quad (2)$$

式中, L 为多层感知机的总层数; \mathbf{X} 为输入向

量; $W^{(i)}$ 和 $b^{(i)}$ 为第 i 层的权重矩阵和偏置向量; $f(\cdot)$ 为激活函数; $a^{(i)}$ 为第 i 层激活输出; $Z^{(i)}$ 为第 i 层加权输入; \hat{y} 是多层感知机的输出结果。

1.4 随机森林(RF)

随机森林是一种集成学习方法,由多个决策树组成。每个决策树都是通过对训练数据的随机采样和特征随机选择来构建的,每个决策树的预测结果被综合考虑,最终的预测结果是基于所有决策树的平均值或投票结果。

预测表达式:

$$\hat{y} = \frac{1}{T} \sum_{i=1}^T f_i(X) \quad (3)$$

式中, \hat{y} 为随机森林模型的输出结果; T 为决策树的数量; $f_i(X)$ 为第 i 棵决策树对输入 X 的预测结果。

1.5 模型预测精度评价

采用浸出金属量的实际值和预测值之间的均方根误差和确定系数两个常见的评价指标来衡量模型性能。均方根误差(RMSE)也是衡量实际值与预测值之间的平均差异的指标。对预测误差进行了平方,使得较大的误差具有更高的惩罚,反映了预测值与真实值之间的差异程度。因此, RMSE 越小,表示预测模型的准确性越高。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

确定系数(R^2)可以衡量模型对因变量变异的解释程度。 R^2 的取值范围为 $[0, 1]$,越接近 1 表示模型的预测精度越高。

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \quad (5)$$

式中, y_i 为实际值; \hat{y}_i 为预测值; \bar{y}_i 为实际值的平均值。

本文使用机器学习中常用的 Sklearn 库开展模型训练与评估, Sklearn 能够实现高效的算法应用,涵盖了本文所需的机器学习算法:多元线性回归,多层感知机(MLP)和随机森林(RF)。

2 特征值筛选

为了提高模型的预测能力和训练速度,减少冗余信息的干扰,需找到对目标值预测贡献度最大的特征变量,在进行特征值筛选时,使用皮尔逊相关系数(r)衡量各个变量之间线性相关程度,选取合适的

特征值作为输入变量。在此基础上对总流量、平米铀量、浸出金属和浸出率进行分析,各个参数之间的皮尔逊相关系数见图 2。结果表明:浸出液金属量与总流量和平米铀量之间存在着不同程度的相关性:浸出金属量与总流量呈现正相关($r=0.4$),与平米铀量呈现正相关($r=0.41$);浸出率与总流量呈现正相关($r=0.42$),与平米铀量相关性较差($r=-0.15$)。考虑到铀浸出的环境是复杂的,影响浸出金属量的因素之间可能存在复杂的相互作用,进一步使用多元线性回归和机器学习来分析这些因素对浸出金属量造成的影响。

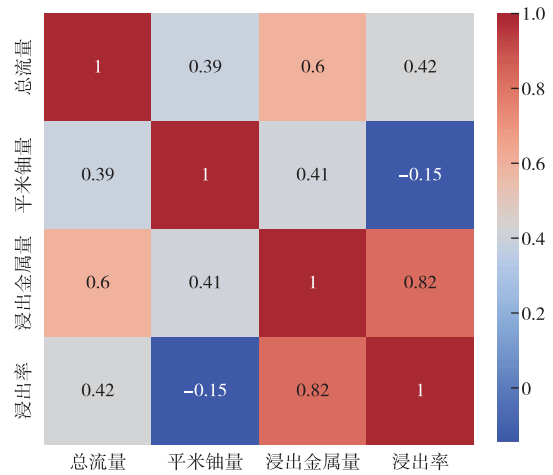


图 2 皮尔逊相关系数热力矩阵图
Fig. 2 Pearson correlation coefficient thermodynamic matrix diagram

在此次数据分析中,将浸出金属量作为目标变量,总流量和平米铀量数据作为特征变量,将数据输入到机器学习模型中,训练出一个能够准确预测铀浸出金属量的模型为后续的分析 and 决策提供支持。

3 结果与讨论

由于各个采区相互独立且采区抽液钻孔数较少、数据之间的差异较大,不考虑模型的泛化性能。基于以上条件,将此采区的数据同时作为训练集和测试集,通过训练好的模型预测数据的变化,并将其与真实值进行比较来评估模型的预测能力。在模型训练完成后对比分析,选取预测能力最佳的模型进行情景分析,验证结果的准确性。

3.1 预测铀浸出金属量的机器学习模型

使用线性回归预测模型对铀浸出金属量进行预测时,拟合效果较差,确定系数仅为 0.31,均方根误

差为 0.14(图 3)。这表明多元线性回归模型对于铀浸出金属量的整体解释能力较弱。由于总流量、平米铀量和浸出金属量之间并不是简单的线性关系,因此使用 MLP 和 RF 来地拟合浸出金属量的变化。

在使用机器学习模型时,参数设置会对模型的性能和行为产生影响。因此,需要调节合适的参数组合,以使模型的性能达到预期效果。两个机器学习的超参数设置如表 1 所示。

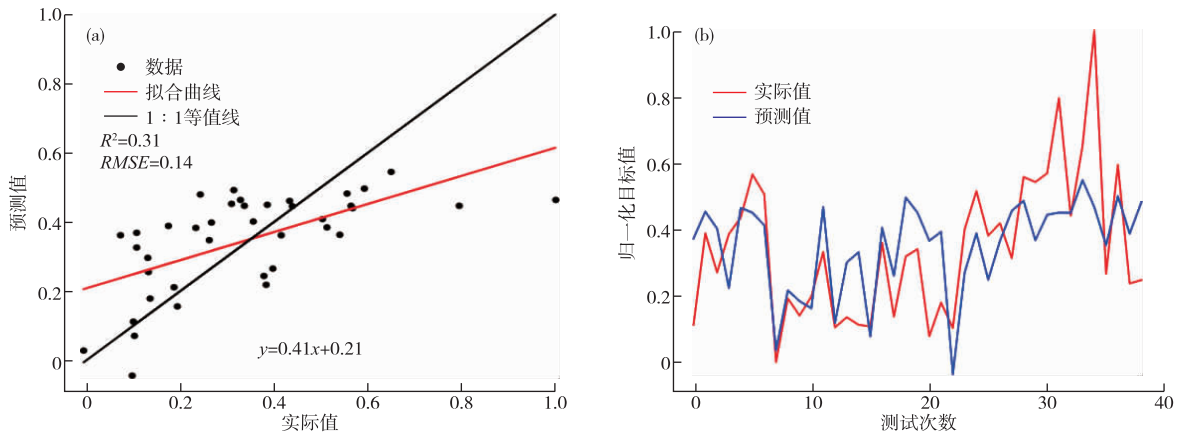


图 3 线性回归预测误差图(a)和线性回归预测折线图(b)

Fig. 3 Linear regression prediction error chart (a) and linear regression prediction line chart (b)

表 1 机器学习超参数设置

Table 1 Machine learning hyperparameters settings

方法	功能	优化参数
多层感知机	MLP	hidden_layer_size:(5 000,2 000,200,20)
		activation:relu
		solver:lbgfs
随机森林	Random Forest	max_iter:5000
		n_estimators:100
		random_state:42
		min_samples_split:2
		min_samples_leaf:1

如图 4 所示,MLP、RF 两种机器学习的均方根误差分别为 0.06、0.07,确定系数分别是 0.91、0.89。MLP 和 RF 模型的拟合效果均比多元线性回归模型更佳,其中 MLR 模型拟合效果最好,但耗时相对于 RF 模型会更久、超参数设置会更复杂。

3.2 特征变量重要性评价

为比较总流量和平米铀量对浸出金属量的影响,进行特征重要性评估。在此次评估中,由于 MLP 模型在隐藏层之间使用了非线性激活函数模型的权重来计算,各变量的重要性可能受到非线性关系的影响,使特征重要性的解释变得复杂,导致评

估结果可能不够准确。因此,特征重要性评价使用 RF 模型,评估结果显示,总流量和平米铀量对结果的权重占比分别为 81.6%和 18.4%。这表明总流量是影响浸出金属量的主要因素。

3.3 情景分析

使用预测精度较高的 MLP 模型,以总流量 50%对应的抽液结果来衡量模型预测铀浸出金属量变化的能力,绘制实际值和预测值对应的浸出金属等值线图。如图 5 所示,实际值和预测值之间的差异较小,均方根误差为 0.06,说明该模型能够准确预测浸出金属量,达到了预期效果。

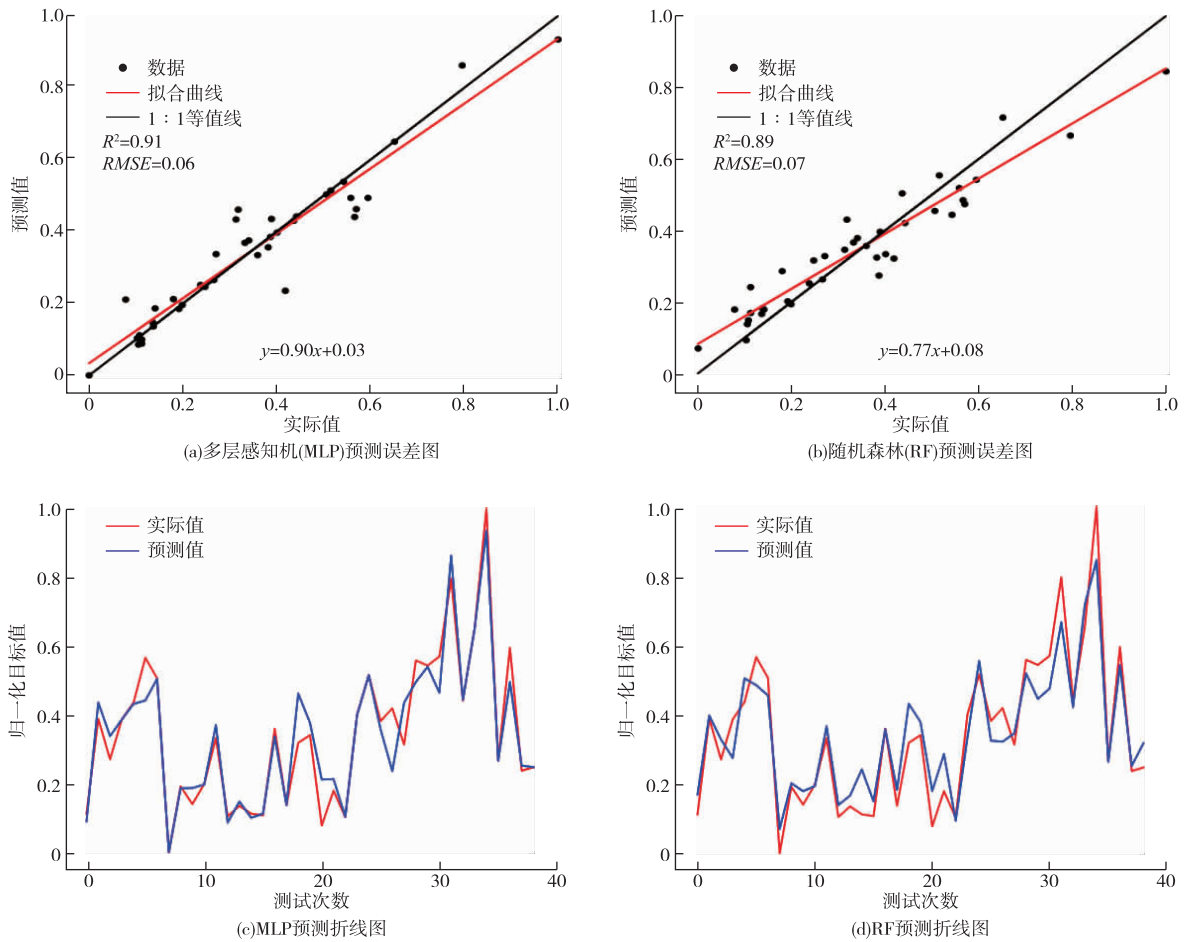


图 4 MLP 及 RF 预测误差图和折线图

Fig. 4 Prediction error chart and prediction line chart of MLP and RF models

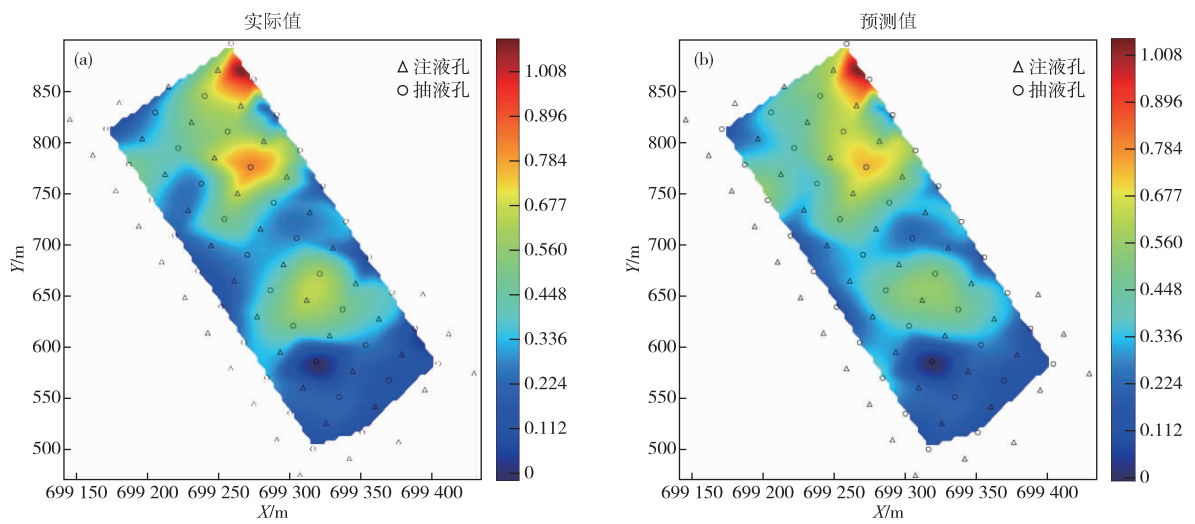


图 5 实际值(a)和预测值(b)等值线图

Fig. 5 Contour plots of actual values (a) and predicted value (b)

4 结论

1)通过相关性分析可知,浸出金属量与总流量和平米铀量之间存在不同程度的相关性。

2)使用多元线性模型建立浸出金属量与总流量和平米铀量的预测模型拟合效果较差,表明浸出金属量与总流量和平米铀量的关系较复杂,不是简单的线性关系。

3)多层感知机(MLP)和随机森林(RF)两种机器学习方法均可以建立较好的预测模型。其中MLP模型的确定系数较高,达到了0.91。MLP和RF方法能够较好地处理影响铀浸出金属量相关数据,自动地学习和适应新的数据,发现隐藏的模式和关系,处理也更为便捷。

参考文献

- [1] LIU X, LU D W, ZHANG Q, et al. Data-driven machine learning in environmental pollution: gains and problems[J]. *Environmental Science & Technology*, 2022,56(4):2124-2133.
- [2] ZHONG S F, ZHANG K, BAGHERI M, et al. Machine learning: new ideas and tools in environmental science and engineering[J]. *Environmental Science & Technology*, 2021,55(19):12741-12754.
- [3] 雷林,雷泽勇. 酸法地浸浸出液铀浓度神经网络预测模型的研究[J]. *矿冶工程*, 2007,27(4):17-20.
LEI L, LEI Z Y. Neural network predicting model for uranium concentration of acid in-situ leach liquor[J]. *Mining and Metallurgical Engineering*, 2007,27(4):17-20.
- [4] 易卫平,周泉,余芸珍. 浸出液铀浓度预测系统 SPUC 的研制[J]. *南华大学学报(自然科学版)*, 2004,18(4):105-108.
YI W P, ZHOU Q, YU Y Z. Development of system on predicting uranium concentration from pregnant solution[J]. *Journal of Nanhua University (Natural Sciences Edition)*, 2004,18(4):105-108.
- [5] PEETS P, WANG W C, MACLEOD M. MS2Tox machine learning tool for predicting the ecotoxicity of unidentified chemicals in water by nontarget LC-HRMS[J]. *Environmental Science & Technology*, 2022,56(22):15508-15517.
- [6] JIA X Y, CAO Y N, O'CONNOR D, et al. Mapping soil pollution by using drone image recognition and machine learning at an arsenic-contaminated agricultural field[J]. *Environmental Pollution*, 2021,270:116281. 1-116281. 10.
- [7] JIA X L, HU B F, MARCHANT B P, et al. A methodological framework for identifying potential sources of soil heavy metal pollution based on machine learning: a case study in the Yangtze Delta, China[J]. *Environmental Pollution*, 2019,250:601-609.
- [8] YANG S Y, DAVID T, YANG D, et al. A synthesis framework using machine learning and spatial bivariate analysis to identify drivers and hotspots of heavy metal pollution of agricultural soils[J]. *Environmental Pollution*, 2021,287:117611. 1-117611. 10.
- [9] REN X, MI Z Y, CAIT, et al. Flexible bayesian ensemble machine learning framework for predicting local ozone concentrations[J]. *Environmental Science & Technology*, 2022,56(7):3871-3883.
- [10] ZHANG Z, X M Z, ZHANG D, et al. Application of machine learning in a mineral leaching process: taking pyrolusite leaching as an example [J]. *Robotics & Machine Learning Daily News*, 2023:68-69.
- [11] 刘石玉,刘金辉,周义朋,等. 地浸采铀过程中含矿层堵塞特征研究[J]. *有色金属(冶炼部分)*, 2022(8):65-75.
LIU S Y, LIU J H, ZHOU Y P, et al. Study on characteristics of ore containing layer blockage during in-situ leaching of uranium mining [J]. *Nonferrous Metals (Extractive Metallurgy)*, 2022(8):65-75.
- [12] 陈万利. 巴彦乌拉铀矿地浸过程中含矿层化学堵塞机理研究[D]. 南昌:东华理工大学, 2017.
CHEN W L. Study on the chemical jamming mechanism of in-situ leaching process in ledge of Bayan Ula uranium ore deposit [D]. Nanchang: East China University of Technology, 2017.
- [13] 郑和秋野. 巴彦乌拉铀矿现场地浸条件试验与地浸场水动力模拟[D]. 南昌:东华理工大学, 2017.
ZHENG-HE Q Y. Site conditions experiments for field test on in-situ leaching and hydrodynamic simulation of uranium in Bayan-Ula [D]. Nanchang: East China University of Technology, 2017.
- [14] 邢拥国,刘金辉,唐华章,等. 巴彦乌拉矿床铀矿体分布与铀矿物赋存特征[J]. *东华理工大学学报(自然科学版)*, 2017,40(1):1-6.
XING Y G, LIU J H, TANG H Z, et al. Characteristic of uranium body distribution and uranium mineral storage in Bayanwula uranium deposit[J]. *Journal of East China University of Technology (Natural Science)*, 2017,40(1):1-6.
- [15] 邓锦勋,许影,赵利信,等. 不同氧化剂在酸法地浸铀矿山难浸出矿石中的应用研究[J]. *中国矿业*, 2018,27(11):116-120.
DENG J X, XU Y, ZHAO L X, et al. Application of different oxidants for hard leaching ores in an acid in-situ leaching uranium mine[J]. *China Mining Magazine*,

- 2018,27(11):116-120.
- [16] 刘迎九,谢水波,周泉,等.某铀矿床酸法地浸现场试验及化学堵塞成因分析[J].南华大学学报(自然科学版),2007,21(1):10-13.
LIU Y J, XIE S B, ZHOU Q, et al. The field test of acid in-situ leaching in the uranium deposit and study on the chemical jamming[J]. Journal of University of South China (Social Science Edition), 2007, 21(1): 10-13.
- [17] 汪润超,李寻,罗跃,等.基于数值模拟分析赤铁矿对地浸采铀的影响[J].有色金属(冶炼部分),2021(8):87-93.
WANG R C, LI X, LUO Y, et al. Influence of hematite on in-situ leaching of uranium based on numerical simulation[J]. Nonferrous Metals (Extractive Metallurgy), 2021(8):87-93.
- [18] 王兵,罗跃,李寻,等.基于地球化学模拟分析杂质矿物对酸法地浸采铀的影响[J].原子能科学技术,2021,55(6):976-983.
WANG B, LUO Y, LI X, et al. Influence of impurity mineral on acid leaching uranium based on geochemical simulation[J]. Atomic Energy Science and Technology, 2021, 55(6):976-983.
- [19] 蒋小辉,叶善东,陈国贞,等.新疆某高耗酸矿床碱法地浸采铀试验研究[J].中国矿业,2012,21(增刊1):301-305.
JIANG X H, YE S D, CHEN G Z, et al. Alkaline in-situ leaching test of a high-acid-consumption uranium mineral deposit in Xinjiang[J]. China Mining Magazine, 2012, 21(Suppl. 1):301-305.
- [20] 孙娇.广西大新铀矿床铀的浸出特征及成矿物质来源研究[D].成都:成都理工大学,2014.
SUN J. The leaching characteristics of uranium and source material research in Daxin uranium deposit of Guangxi[D]. Chengdu: Chengdu University of Technology, 2014.
- [21] 陶峰,张传飞,冯国平,等.某砂岩型铀矿 $\text{CO}_2 + \text{O}_2$ 地浸采铀试验[J].有色金属(冶炼部分),2022(6):56-61.
TAO F, ZHANG C F, FENG G P, et al. $\text{CO}_2 + \text{O}_2$ in-situ leaching of uranium from a sandstone type uranium deposit [J]. Nonferrous Metals (Extractive Metallurgy), 2022(6):56-61.
- [22] 刘亚洲,罗跃,李寻,等.巴彦乌拉铀矿床酸法浸铀的水岩反应堵塞机理[J].有色金属(矿山部分),2022,74(5):44-51.
LIU Y Z, LUO Y, LI X, et al. Water rock reaction plugging mechanism of acid leaching uranium in Bayanwula uranium deposit[J]. Nonferrous Metals (Mining Section), 2022, 74(5):44-51.
- [23] 丰章哲,刘金辉,阳奕汉,等.地浸采铀过程中的矿物溶解沉淀特征[J].有色金属(冶炼部分),2022(10):72-80.
FENG Z Z, LIU J H, YANG Y H, et al. Characterization of mineral dissolution precipitation during uranium in-situ leaching [J]. Nonferrous Metals (Extractive Metallurgy), 2022(10):72-80.