



引文格式: 孔垂瑾, 袁辉, 来志庆, 等. 运用“指针矿物”黑云母识别斑岩铜矿成矿岩体: 机器学习视角[J]. 西北地质, 2026, 59(4): 48–62. DOI: 10.12401/j.nwg.2025189

Citation: KONG Chuijin, YUAN Hui, LAI Zhiqing, et al. Identifying Porphyry Copper Ore-Forming Intrusions Using Biotite as An Indicator Mineral: A Machine Learning Perspective[J]. Northwestern Geology, 2026, 59(4): 48–62. DOI: 10.12401/j.nwg.2025189

运用“指针矿物”黑云母识别斑岩铜矿成矿岩体: 机器学习视角

孔垂瑾^{1,2,3}, 袁辉^{1,2,3}, 来志庆^{1,2,3}, 李三忠^{1,2,3}, 钟世华^{1,2,3,*}

(1. 中国海洋大学海洋地球科学学院, 山东 青岛 266100; 2. 海底科学与探测技术教育部重点实验室, 山东 青岛 266100; 3. 青岛海洋科技中心, 海洋矿产资源评价与探测技术功能实验室, 山东 青岛 266237)

摘要: 准确识别成矿岩体是斑岩铜矿找矿勘查的关键。前人开展了大量基于锆石、磷灰石等岩浆副矿物微量元素识别岩浆成矿潜力的研究, 然而副矿物的微量元素组成受控因素较为复杂, 这可能在一定程度影响对成矿岩浆信息的指示能力。黑云母是中酸性岩浆岩中常见的造岩矿物, 其主量元素组成主要受控于原始岩浆, 因此可以作为“指针矿物”指示岩浆成矿潜力。笔者在系统汇编全球斑岩铜矿成矿岩体与全球非成矿中酸性岩体黑云母数据的基础上, 选取 9 种黑云母主量元素成分, 选取支持向量机、随机森林和极限梯度提升 3 种机器学习模型, 用于识别斑岩铜矿成矿岩体。结果表明, 训练得到的 3 种模型均表现良好, 准确率分别为 0.956 1、0.961 4 和 0.958 7, 说明该方法能够有效识别斑岩铜矿成矿岩体。此外, 笔者借助机器学习解释性算法 SHAP 对 3 种模型进行特征贡献度分析。结果显示, 对输出结果影响最大的 3 个训练特征依次为 FeO^T 、 Na_2O 和 MnO , 且成矿岩体的黑云母整体表现为低 FeO^T 、低 MnO 和高 Na_2O 的特征组合。研究表明, 黑云母主量元素能够有效指示岩体成矿潜力, 这为斑岩铜矿成矿岩体识别及类似矿床找矿方法研究提供了新的思路。

关键词: 斑岩铜矿; 黑云母; 主量元素; 机器学习; 成矿岩体识别

中图分类号: P585; P612

文献标志码: A

文章编号: 1009-6248(2026)04-0048-15

Identifying Porphyry Copper Ore-Forming Intrusions Using Biotite as An Indicator Mineral: A Machine Learning Perspective

KONG Chuijin^{1,2,3}, YUAN Hui^{1,2,3}, LAI Zhiqing^{1,2,3}, LI Sanzhong^{1,2,3}, ZHONG Shihua^{1,2,3,*}

(1. College of Marine Geosciences, Ocean University of China, Qingdao 266100, Shandong, China; 2. Key Laboratory of Submarine Geosciences and Prospecting Techniques, Ministry of Education, Qingdao 266100, Shandong, China; 3. Laboratory for Marine Mineral Resources, Qingdao Marine Science and Technology Center, Qingdao 266237, Shandong, China)

收稿日期: 2026-03-13; 修回日期: 2026-04-10; 责任编辑: 吕鹏瑞

基金项目: 国家自然科学基金项目(42472104), 山东省自然科学基金青年科学基金(B类)项目(ZR2025QB19), 山东省泰山学者青年专家项目(tsqn202507105), 深地国家科技重大专项(2025ZD1009710)联合资助。

作者简介: 孔垂瑾(2004–), 女, 本科生, 地质学专业。E-mail: chuijinkong@163.com。

* 通讯作者: 钟世华(1989–), 男, 博士, 副教授, 从事地质大数据与成矿研究。E-mail: zhongshihua@ouc.edu.cn。

© Editorial Office of *Northwestern Geology*. OA under CC BY-NC-ND

Abstract: Accurate identification of ore-forming intrusions is critical for porphyry copper exploration. Previous studies have extensively used trace-element compositions of magmatic accessory minerals, such as zircon and apatite, to evaluate the metallogenic potential of magmas. However, the trace-element compositions of accessory minerals are controlled by multiple factors, which may limit their effectiveness in recording ore-forming magma signatures. Biotite is a common rock-forming mineral in intermediate to felsic igneous rocks, and its major-element compositions are mainly controlled by the parental magma, making it a potential indicator mineral for magmatic metallogenic potential. Based on a systematic compilation of biotite data from porphyry copper ore-forming intrusions and barren intermediate to felsic intrusions worldwide, this study selected nine major elements of biotite to construct three machine-learning models, namely support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGB), for identifying porphyry copper ore-forming intrusions. The results show that all three models perform well, with accuracies of 0.956 1, 0.961 4, and 0.958 7, respectively, indicating that this approach can effectively identify ore-forming intrusions related to porphyry copper mineralization. SHAP analysis further reveals that FeO^T , Na_2O , and MnO are the three most important features controlling model outputs, and that biotite from ore-forming intrusions is generally characterized by low FeO^T , low MnO , and high Na_2O . These results demonstrate that the major-element compositions of biotite can effectively indicate the metallogenic potential of intrusions, providing a new approach for identifying ore-forming intrusions in porphyry copper systems and for guiding exploration of similar deposit types.

Keywords: porphyry copper deposit; biotite; major elements; machine learning; ore-forming intrusion identification

斑岩铜矿提供了世界上约 75% 的铜、50% 以上的钼以及 25% 以上的金资源 (Sillitoe, 2010), 在全球铜矿资源供应中具有重要意义, 因而长期受到学术界和工业界的广泛关注 (Richards, 2011)。大量研究表明, 斑岩铜矿的形成与中酸性岩浆活动密切相关 (Heinrich, 2024), 中酸性岩浆的成分直接决定了斑岩铜矿化能否形成以及规模大小 (Ballard et al., 2002; Sun et al., 2015; Huang et al., 2024)。因此, 建立准确识别中酸性侵入体成矿潜力的方法, 对于深化斑岩铜矿成矿机制认识及指导找矿勘查均具有重要意义。

近年来, 随着矿物微区分析技术的发展, 越来越多研究尝试通过挖掘矿物地球化学成分变化规律, 建立能够指示成矿信息的地球化学方法 (Qiu et al., 2024), 以提高找矿勘查效率、降低成矿预测成本。这类能够记录和指示成矿特征的矿物通常被称为“指针矿物” (Indicator Mineral; Plouffe et al., 2024)。指针矿物既可以是岩浆岩中的造岩矿物 (如橄榄石、云母等) 和副矿物 (如石榴子石、锆石和磷灰石等), 也包括岩浆-热液矿床热液阶段形成的硅酸盐矿物和硫化物 (Lehtonen et al., 2015)。目前, 指针矿物已经被广泛应用于岩浆矿床、火山块状硫化物矿床、密西西比河谷型矿床以及斑岩型矿床等多种矿床类型的找矿预测

与成因研究中 (王玉峰等, 2020; Wilkinson et al., 2020; 李华伟等, 2023; Lin et al., 2023)。

在斑岩铜矿成矿岩体识别研究中, 应用最为广泛的指针矿物是锆石、磷灰石等常见副矿物 (Huang et al., 2023; 郭广慧等, 2023)。这是由于副矿物在中酸性岩中广泛发育, 封闭温度高、抗蚀变能力强, 因此被认为是记录原始岩浆成矿信息的良好载体 (Hoskin et al., 2003; 王海然等, 2013; Lee et al., 2021; Carrasco-Godoy et al., 2024)。然而, 越来越多的研究指出, 包括锆石在内的岩浆副矿物成分不仅受原始岩浆成分控制, 还受到温度、结晶顺序及后期流体改造等多种因素影响 (Bouzari et al., 2016; Tatnell et al., 2023), 这可能会导致对中酸性岩体成矿潜力的误判 (Barber et al., 2021)。因此, 探索新的指针矿物, 对于提高成矿岩体识别的可靠性具有重要意义。

相比岩浆副矿物, 造岩矿物在岩浆体系中含量更高、分布更广, 其晶格中的主量元素种类更为丰富, 化学组成对母岩浆性质及后续岩浆演化过程也更为敏感 (Yu et al., 2022; Pan et al., 2024)。因此, 作为指针矿物的造岩矿物在指示成矿岩体信息时具有优势 (Huang et al., 2023)。黑云母是中酸性岩浆岩中常见的铁镁质造岩矿物, 其晶体结构中可容纳多种主量元

素,包括 Ti、Mg、Fe、Mn 等。已有研究表明,黑云母主量元素及挥发分相关组分可在一定程度上记录岩浆结晶条件及岩浆-热液演化信息(Henry et al., 2005; Yu et al., 2022; Liu et al., 2023; Baidya et al., 2024)。例如,黑云母的 X_{Mg} [即 $Mg/(Mg+Fe)$] 或 $Fe/(Fe+Mg)$ 可在一定程度上指示体系氧逸度变化(Ishihara, 1977; Pan et al., 2024),而氧逸度正是区分斑岩铜矿成矿岩体和非成矿岩体的重要因素之一(Heinrich, 2024)。因此,利用黑云母主量元素组成识别斑岩铜矿成矿岩体在理论上是可行的。

近年来,随着全球地球化学数据库(如 GEOROC 等)的持续建设与更新,大规模岩石-矿物地球化学数据的获取成为可能。相较于传统依赖少量经验性指标和单一判别图解的方法,机器学习能够综合利用多维地球化学信息,识别复杂的非线性关系,因此更适用于多元素耦合控制的成矿问题。近年的研究表明,机器学习方法已在副矿物成因分类、成矿潜力识别及构造背景判别等方面取得良好应用效果(Luo et al., 2024; Ge et al., 2025)。例如,支持向量机(Support Vector Machine, 简称 SVM)、随机森林(Random Forest, 简称 RF)和极限梯度提升(eXtreme Gradient Boosting, 简称 XGB)等算法已成功应用于全岩地球化学成矿肥力分类(Nathwani et al., 2022)、磷灰石微量元素矿床类型识别(Qiu et al., 2024)以及锆石微量元素构造背景判别(Luo et al., 2024)等研究中,显示出对高维、多源地球化学数据的良好适应性和较高判别能力。已有研究还表明,基于磷灰石微量元素数据构建的支持向量机与随机森林模型能够较高精度地区分不同花岗岩成因类型;基于全岩地球化学数据的随机森林模型亦可有效识别成矿岩体与非成矿岩体(刘嘉情等, 2023, 2025)。

上述研究进展说明,机器学习不仅能够提升地质大数据分析效率,还能够从复杂变量关系中提取对成矿体系识别最有效的信息,从而为成矿岩体与非成矿岩体判别提供可靠的方法学基础。对于中酸性岩浆岩而言,黑云母作为常见的富铁镁造岩矿物,其主量元素组成能够在一定程度上记录岩浆演化过程及相关成矿信息,因此有望成为识别斑岩铜矿成矿岩体的重要载体。将机器学习方法引入黑云母成分数据分析,有助于进一步揭示黑云母主量元素组成与斑岩铜矿成矿潜力之间的内在联系,并为成矿岩体识别提供新的技术路径。

基于此,笔者系统汇编了全球斑岩铜矿成矿岩体与全球非成矿中酸性岩体中的黑云母主量元素数据,采用支持向量机、随机森林和极限梯度提升 3 种机器学习算法开展模型训练,构建能够自动识别中酸性岩体斑岩铜矿成矿潜力的智能分类模型。此外,进一步结合了机器学习解释性方法,对训练模型进行分析,定量评估各主量元素特征对分类结果的贡献,这对深入揭示斑岩铜矿控矿因素、丰富斑岩铜矿成矿理论具有重要意义。

1 数据汇编及预处理

1.1 数据汇编

笔者在系统梳理公开发表文献和数据库资料(GEOROC)的基础上,汇编建立了一个用于中酸性岩浆岩成矿潜力判别的黑云母主量元素数据库。该数据库中的黑云母均为岩浆成因黑云母,包括来自全球各地的 473 条斑岩铜矿成矿岩体黑云母主量元素数据和 3 273 条非成矿中酸性岩体黑云母主量元素数据(表 1)。针对每条黑云母数据,选取 9 种主量元素(SiO_2 、 TiO_2 、 Al_2O_3 、 FeO^T 、 MnO 、 MgO 、 CaO 、 Na_2O 、 K_2O)作为机器学习模型输入特征。选择上述 9 种主量元素作为训练特征,主要基于以下两方考虑:①大部分黑云母电子探针主量元素分析均包含这些元素,因此能够确保汇编的数据库涵盖目前大多数黑云母分析结果;②这些元素被证明能够有效反映岩浆信息及相关成矿信息,且已有研究表明,其组成特征对区分成矿与非成矿岩体具有较好的指示意义(Moshefi et al., 2018)。

笔者对来自两种岩体类型的黑云母主量元素地球化学特征进行了统计分析(表 2),并绘制了黑云母主量元素箱线图(图 1)。成矿岩体黑云母 SiO_2 含量为 30.17%~51.23%,平均值为 37.62%,中值为 37.63%;非成矿岩体黑云母 SiO_2 含量为 19.48%~56.36%,平均值为 36.64%,中值为 36.50%。成矿岩体黑云母 TiO_2 含量为 0.24%~18.13%,平均值为 3.85%,中值为 3.52%;非成矿岩体黑云母 TiO_2 含量为 0.01%~19.16%,平均值为 3.30%,中值为 3.11%。成矿岩体黑云母 Al_2O_3 含量为 2.75%~35.59%,平均值为 14.43%,中值为 14.01%;非成矿岩体黑云母 Al_2O_3 含量为 1.11%~34.92%,平均值为 14.68%,中值为 14.52%。成矿岩体黑云母 FeO^T 含量为 2.09%~30.31%,平均值为 16.32%,中值为 16.00%;

表 1 本研究使用的来自成矿岩体与非成矿岩体的黑云母数据来源

Tab. 1 Data sources of biotite samples used in this study

编号	地区	矿床名称	矿床类型	黑云母类型	数据量	数据来源
1	中国安徽宣城	茶亭	斑岩铜金矿床	成矿黑云母	14	肖庆玲等, 2018
2	中国安徽铜陵	舒家店	斑岩铜矿床	成矿黑云母	6	李万滢, 2024
3	中国东北	多处Mo-Cu	斑岩钼铜矿床	成矿黑云母	16	Kan et al., 2023
4	中国东昆仑	多处侵入体	斑岩铜相关岩浆系统	成矿黑云母	8	Zhang et al., 2024
5	中国大兴安岭北部	小柯勒河	斑岩铜矿床	成矿黑云母	3	邓昌州, 2019
6	中国吉林-黑龙江东部陆缘	洋灰洞子	斑岩铜矿床	成矿黑云母	13	褚小磊, 2022
		小西南岔	斑岩铜矿床	成矿黑云母	7	
7	中国青海南部	纳日贡玛	斑岩钼(铜)矿床	成矿黑云母	19	郝金华等, 2010
		甲玛	斑岩铜矿床	成矿黑云母	10	
		拉抗俄	斑岩铜钼矿床	成矿黑云母	11	
8	中国西藏冈底斯	邦铺	斑岩铜钼矿床	成矿黑云母	4	刘诗文, 2015; 吴松, 2016; 唐攀等, 2017; 唐攀, 2018; Tang et al., 2019
		驱龙	斑岩铜钼矿床	成矿黑云母	31	
		朱诺	斑岩铜矿床	成矿黑云母	11	
9	中国西藏	玉龙	斑岩铜矿床	成矿黑云母	4	刘诗文, 2015
	中国新疆	包古图	斑岩铜矿床	成矿黑云母	16	
10	中国云南	普朗	斑岩铜金矿床	成矿黑云母	14	刘学龙等, 2013; 曹康, 2018; 张少颖, 2022
		铜厂沟	斑岩钼(铜)矿床	成矿黑云母	13	
11	中国云南西北部	红牛-红山	斑岩-矽卡岩铜矿床	成矿黑云母	40	李守奎, 2023; 周杰虎等, 2024
		Pulang (PL)	斑岩铜矿床	成矿黑云母	2	
		Songnuo (SN)	斑岩铜矿床	成矿黑云母	2	
12	中国云南西北部(中甸弧区)	Langdu (LD)	斑岩-矽卡岩铜矿床	成矿黑云母	2	Pan et al., 2024
13	伊朗	Deh-Bala	斑岩铜矿床	成矿黑云母	15	Kazemi et al., 2022
14	伊朗西部	Darezerezhk et al.	斑岩铜矿床	成矿黑云母	7	Taghavi et al., 2022
15	伊朗西北部	Sungun	斑岩铜钼矿床	成矿黑云母	23	Moshefi et al., 2018
16	伊朗克尔曼带	Chahfiruzeh et al.	斑岩铜矿床	成矿黑云母	75	Zarasvandi et al., 2019; Heidari et al., 2019
17	印度尼西亚	Batu Hijau	斑岩铜金矿床	成矿黑云母	3	Idrus, 2018
18	希腊东北部	Maronia	斑岩铜钼(金)矿床	成矿黑云母	42	Schaarschmidt et al., 2021
19	美国缅因州	Deboullie	斑岩铜钼矿床	成矿黑云母	3	Loferski et al., 1995
20	加拿大	多岩体对比	斑岩铜钼系统	成矿黑云母	20	Zhang et al., 2016
21	加拿大	Sisson Brook	斑岩铜钼矿床	成矿黑云母	28	Zhang, 2015
22	秘鲁北部	Rio Blanco(白河)	斑岩铜钼矿床	成矿黑云母	9	陈念, 2022
23	阿根廷	Quebrada de la Mina、Altar North	斑岩铜金矿床	成矿黑云母	2	Maydagán et al., 2017
24	全球	—	—	非成矿黑云母	3 273	GEOROC

表 2 成矿与非成矿岩体黑云母主量元素(%)统计特征

Tab. 2 Statistical characteristics of major elements (%) in biotite from ore-forming and barren intrusions

元素特征	成矿岩体黑云母			非成矿岩体黑云母		
	含量	平均值	中值	含量	平均值	中值
SiO ₂	30.17~51.23	37.62	37.63	19.48~56.36	36.64	36.50
TiO ₂	0.24~18.13	3.85	3.52	0.01~19.16	3.30	3.11
Al ₂ O ₃	2.75~35.59	14.43	14.01	1.11~34.92	14.68	14.52
FeO ^T	2.09~30.31	16.32	16.00	3.15~38.94	20.61	20.03
MnO	0.01~2.29	0.23	0.22	0.01~10.63	0.42	0.32
MgO	0.67~21.09	13.31	13.57	0.002~23.961	9.998	10.26
CaO	0.01~3.28	0.06	0.03	0.01~13.28	0.28	0.04
Na ₂ O	0.02~2.16	0.18	0.17	0.008~2.6	0.24	0.124
K ₂ O	4.7~10.69	9.05	9.236	0.016~11.76	9.01	9.27

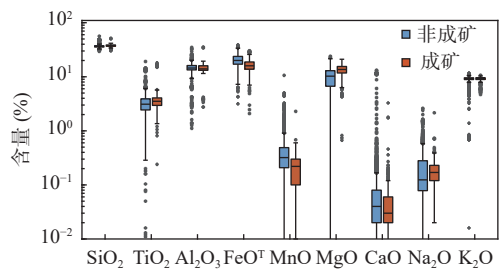


图1 成矿岩体与非成矿岩体黑云母9种主量元素箱线图
Fig. 1 Box plots of nine major elements in biotite from ore-forming and barren intrusions

非成矿岩体黑云母 FeO^T 含量为 3.15%~38.94%，平均值为 20.61%，中值为 20.03%。成矿岩体黑云母 MnO 含量为 0.01%~2.29%，平均值为 0.23%，中值为 0.22%；非成矿岩体黑云母 MnO 含量为 0.01%~10.63%，平均值为 0.42%，中值为 0.32%。成矿岩体黑云母 MgO 含量为 0.67%~21.09%，平均值为 13.31%，中值为 13.57%；非成矿岩体黑云母 MgO 含量为 0.002%~23.961%，平均值为 9.998%，中值为 10.26%。成矿岩体黑云母 CaO 含量为 0.01%~3.28%，平均值为 0.06%，中值为 0.03%；非成矿岩体黑云母 CaO 含量为 0.01%~13.28%，平均值为 0.28%，中值为 0.04%。成矿岩体黑云母 Na₂O 含量为 0.02%~2.16%，平均值为 0.18%，中值为 0.17%；非成矿岩体黑云母 Na₂O 含量为 0.008%~2.6%，平均值为 0.24%，中值为 0.124%。成矿岩体黑云母 K₂O 含量为 4.7%~10.69%，平均值为 9.05%，中值为 9.24%；非成矿岩体黑云母 K₂O 含量为 0.016%~11.76%，平均值为 9.01%，中值为 9.27%。

整体来看，两类黑云母的许多主量元素含量均表

现出一定差异。成矿岩体黑云母相较于非成矿岩体黑云母，在 MgO 含量上整体偏高，而 FeO^T、MnO 及 CaO 含量相对偏低，反映出两类黑云母记录了不同的原始岩浆信息。

1.2 数据预处理

为保证模型训练的可靠性与稳定性，在开展机器学习建模之前，有必要对黑云母地球化学数据库进行系统预处理，主要包括缺失值处理、数据集划分、类别不平衡处理和数据标准化处理。

(1)缺失值处理：为保证输入数据的完整性并尽可能减少人为干预带来的误差，本研究仅保留同时包含 9 种主量元素(SiO₂、TiO₂、Al₂O₃、FeO^T、MnO、MgO、CaO、Na₂O、K₂O)的黑云母样本，缺失任一指标的数据均予以剔除，以避免插值填补可能引入的人为偏差。筛选后数据集仍保持 3 746 条有效样本，能够满足模型训练需求。

(2)数据集划分：为避免后续类别不平衡处理造成数据泄漏，笔者对原始数据集进行训练集与测试集划分。具体而言，按照 8 : 2 的比例将样本划分为训练集和测试集，并保持成矿岩体黑云母与非成矿岩体黑云母在两个子集中的类别比例基本一致。测试集在划分完成后即保持原始数据分布，不再参与任何过采样、参数优化或交叉验证过程，仅用于最终模型泛化性能评估。该处理流程有助于保证测试结果能够更真实地反映模型在未知样本上的判别能力。

(3)类别不平衡处理：在机器学习分类任务中，训练样本类别分布不均衡是较为常见的问题(Japkowicz et al., 2002; He et al., 2009)。本研究汇编的黑云母数

据同样存在明显类别不平衡现象。在3746条样本中,来自成矿岩体的黑云母数据为473条,而来自非成矿岩体的黑云母数据为3273条,比例约为1:7,数量明显不平衡。若不加处理,模型往往偏向于样本数量占优的类别,从而削弱对成矿样本的识别能力。为降低这一影响,通常可采用过采样、欠采样或二者结合的方式对数据分布进行调整(Chawla et al., 2002; Kotsiantis et al., 2006)。综合考虑样本规模以及信息保留程度,本研究采用过采样技术(Synthetic Minority Over-sampling Technique, SMOTE)对少数类样本进行扩增处理。需要指出的是,SMOTE仅在训练集内部实施,用于构建平衡训练数据集,测试集始终保持原始分布,不参与过采样处理。这样既能够缓解训练阶段的类别不平衡问题,又可以避免因在全数据集上过采样而导致的信息泄漏。

(4)标准化处理:由于不同主量元素含量在数值尺度和分布范围上存在差异,若直接用于模型训练,可能导致特征权重受量纲影响而产生偏差。为消除不同变量量纲差异带来的干扰,笔者采用Z-score标准化方法对特征进行标准化处理,将各特征转换为均值为0、标准差为1的无量纲变量,使不同尺度的数据处于可比状态。标准化过程同样仅基于训练集参数进行,并应用于测试集,以保证训练和预测过程的一致性。该处理有助于提升模型参数求解的稳定性、减少模型运算时长,并增强不同特征在模型中的协同作用,从而提高整体训练效率和预测性能。

2 机器学习训练

2.1 机器学习模型选取

为实现成矿岩体与非成矿岩体黑云母的有效判别,本研究选取支持向量机、随机森林和极限梯度提升三种机器学习算法进行对比分析(图2)。

支持向量机是一种以统计学习理论为基础的监督学习方法(Cortes et al., 1995),其核心思想是在高维特征空间中构建最优分类超平面,使两类样本之间的间隔(Margin)最大化,从而获得良好的泛化能力。针对非线性分类问题,支持向量机可通过核函数将原始特征空间映射至高维空间,使线性不可分问题转化为高维空间中的线性可分问题(Vapnik, 1999)。在地球化学数据分析中,由于元素之间常存在复杂的非线性耦合关系(如Fe-Mg替代、Ti与氧逸度条件的响应关

系等),传统线性方法难以准确刻画其边界特征。因此,本研究采用径向基核函数构建支持向量机模型,以增强其对复杂非线性结构的刻画能力。此外,支持向量机在小样本条件下具有较好的稳定性和泛化性能,适用于成矿样本数量相对较少的情形。

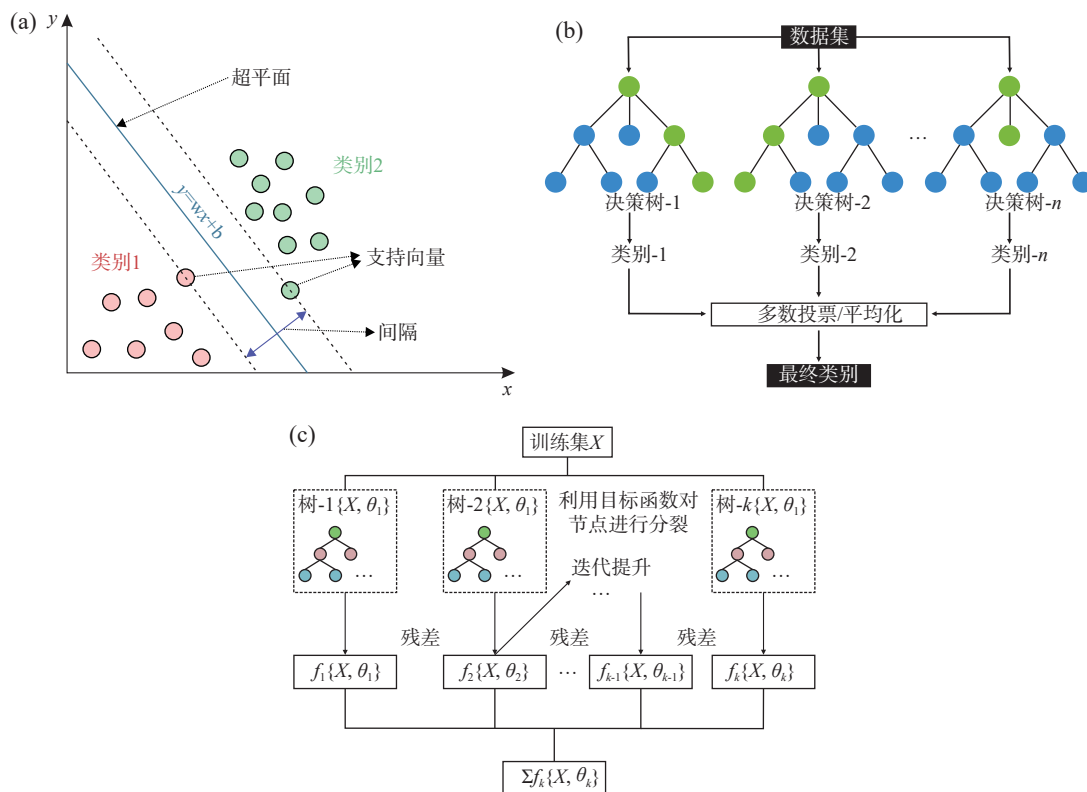
随机森林是一种基于Bagging思想构建的集成学习算法(Breiman, 2001)。该方法通过对原始样本进行自助抽样(Bootstrap Sampling),构建多棵相互独立的决策树,并通过投票机制获得最终分类结果,从而降低单棵决策树易过拟合的风险。随机森林模型能够处理高维数据和变量之间的复杂相互作用关系,对异常值和噪声具有较强的鲁棒性。该方法可输出特征重要性排序,为识别对成矿判别具有关键意义的地球化学指标提供依据。在矿床预测与地球化学分类研究中,随机森林模型已被广泛应用并取得较好效果(Rodriguez-Galiano et al., 2015)。

极限梯度提升是一种基于梯度提升决策树(Gradient Boosting Decision Tree, GBDT)框架发展的高效集成算法(Chen et al., 2016)。与传统GBDT相比,极限梯度提升在目标函数中引入正则化项以控制模型复杂度,并通过二阶梯度优化提升训练效率和收敛速度,同时支持并行计算,具有较强的泛化能力和计算性能。在地球化学分类问题中,元素之间往往存在高度非线性关系和交互效应,极限梯度提升通过逐步拟合残差的方式迭代优化模型,能够有效捕捉复杂特征结构,在类别不平衡条件下仍能保持较高稳定性。

2.2 模型训练和参数优化

对于构建的黑云母主量元素数据库,笔者按照8:2的比例将样本划分为训练集与测试集,并保持成矿岩体黑云母与非成矿岩体黑云母在两个子集中的类别比例基本一致(图3),以降低类别分布偏移对模型评估结果的影响。

在模型训练阶段,考虑到原始训练数据中成矿样本数量明显少于非成矿样本,文中仅在训练集内部采用合成少数类过采样技术对少数类样本进行扩增,以构建类别相对平衡的训练数据集。需要强调的是,SMOTE处理严格限定在训练集范围内,测试集始终保持原始样本结构,以避免类别不平衡处理导致数据泄漏,从而保证测试结果能够更真实地反映模型在未知样本上的判别能力。在此基础上,为降低单次数据划分所带来的偶然性偏差,并提高模型评估结果的可靠性,笔者采用五折交叉验证对平衡后的训练集进行



a. 支持向量机模型结构示意图; b. 随机森林模型结构示意图; c. 极限梯度提升模型结构示意图

图2 本研究采用的3种机器学习模型结构示意图

Fig. 2 Schematic diagrams of the three machine learning models used in this study

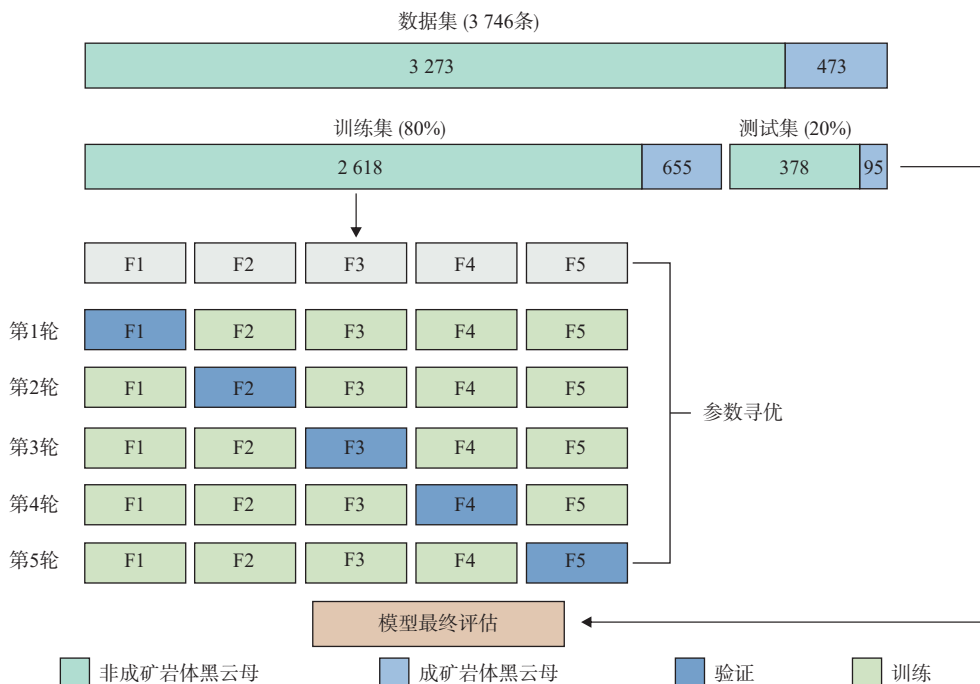


图3 数据集划分与五折交叉验证流程示意图

Fig. 3 Schematic diagram of dataset splitting, training-set oversampling, and fivefold cross-validation workflow

内部评估(Kohavi, 1995)。具体而言,将训练集划分为5个互不重叠的子集,每次选取一个子集作为验证

集,其余4个子集用于模型训练,循环进行5次,使每个子集均作为验证集参与一次评估。最终,以5次验

证结果的平均值作为模型内部性能评价依据, 减弱随机划分带来的波动, 在一定程度上缓解过拟合问题。

在超参数优化方面, 本研究采用贝叶斯优化方法对模型关键参数进行自动搜索。相比传统网格搜索或随机搜索方法, 贝叶斯优化通过构建目标函数的概率代理模型(通常为高斯过程), 并基于贝叶斯推断不断更新对参数空间的后验分布, 在较少的迭代次数下高效逼近最优解(Snoek et al., 2012)。该方法通过在探索与利用之间取得平衡, 能够显著提高参数寻优效率, 尤其适用于参数空间较大且模型训练代价较高的情形。笔者以交叉验证结果作为目标函数, 在训练集内部迭代搜索最优超参数配置。

对于支持向量机模型, 优化参数包括惩罚系数 C 与径向基核函数参数 γ 。参数 C 控制模型对误分类样本的惩罚强度, 而 γ 决定核函数的影响范围, 两者共同影响决策边界的复杂程度(Vapnik, 1999)。当 C 值为 14.067 8, γ 值为 1.0 时, 支持向量机模型获得最优分类效果。

对于随机森林模型, 优化参数包括决策树数量 ($n_estimators$)、最大树深度 (max_depth) 及最大特征数 ($max_features$)。这些参数分别影响模型稳定性、表达能力及随机性程度, 从而决定模型泛化性能(Breiman, 2001)。当 $n_estimators$ 值为 100, $max_features$ 值为 2, max_depth 值为 23 时, 随机森林模型获得最优分类效果。

对于极限梯度提升模型, 优化参数包括学习率 ($learning_rate$)、最大树深度 (max_depth)、子样本比例 ($subsample$) 及正则化参数等。上述参数直接影响模型复杂度与迭代更新幅度, 是控制模型性能的关键因素(Chen et al., 2016)。 $learning_rate$ 值为 0.296 0, max_depth 值为 6, $subsample$ 值为 0.890 9 时, 极限梯度提升模型获得最优分类效果。

在贝叶斯优化框架下, 通过交叉验证评价指标作为目标函数, 不断更新参数组合的后验分布, 最终确定最优超参数配置。优化完成后, 采用独立测试集对最终模型进行性能检验, 以评估其在未知数据上的泛化能力(图 3)。

2.3 模型评价方法

为系统评估不同模型对成矿与非成矿黑云母样本的判别能力, 本研究采用混淆矩阵(Confusion Matrix)与受试者工作特征曲线(Receiver Operating Characteristic Curve, 简称 ROC 曲线)综合评价支持向量机模型、

随机森林模型与极限梯度提升模型的性能。

混淆矩阵用于刻画预测结果与真实类别之间的对应关系, 相关评价指标包括准确率(Accuracy)、精确率(Precision)、召回率(Recall)以及特异度(Specificity)值, 以多维度评估机器学习模型性能(Powers, 2020)。评价指标计算如公式(1)所示。

$$\begin{aligned} \text{准确率} &= \frac{TP+TN}{TP+TN+FP+FN} \\ \text{精确率} &= \frac{TP}{TP+FP} \\ \text{召回率} &= \frac{TP}{TP+FN} \\ \text{特异度} &= \frac{TN}{TN+FP} \end{aligned} \quad (1)$$

式中: TP (True Positive)表示实际为成矿样本且被正确识别的数量; TN (True Negative)表示实际为非成矿样本且被正确识别的数量; FP (False Positive)表示实际为非成矿样本但被误判为成矿样本; FN (False Negative)表示实际为成矿样本但被误判为非成矿样本。

ROC 曲线用于刻画分类模型在不同判别阈值条件下的性能变化情况。该曲线以假阳性率(False Positive Rate, FPR)为横轴、真阳性率(True Positive Rate, TPR)为纵轴, 在 [0,1] 范围内绘制而成, 其计算过程如公式(2)所示。

$$\begin{aligned} TPR &= \frac{TP}{TP+FN} \\ FPR &= \frac{FP}{TN+FP} \end{aligned} \quad (2)$$

式中: FPR 表示在所有真实负样本中被误判为正类的比例; TPR 则表示在所有真实正样本中被正确识别的比例。通常情况下, FPR 越低、 TPR 越高, 说明模型的分类能力越强(Fawcett, 2006)。

在 ROC 曲线的基础上, 可通过计算曲线下的面积(Area Under the Curve, AUC)对模型整体判别能力进行量化评价。AUC 的取值范围为 0~1, 其数值越接近 1, 表明模型对正负样本的区分能力越强(Hanley et al., 1982; 潘北斗等, 2022)。一般而言, 当 AUC 为 0.5~0.7 时, 模型判别能力较弱; 当 AUC 为 0.7~0.9 时, 模型具有较好的区分效果; 当 AUC 大于 0.9 时, 则说明模型具有较高的分类可靠性(Fawcett, 2006)。

3 结果

3.1 模型评价结果

基于独立测试集预测结果, 笔者获得了支持向量

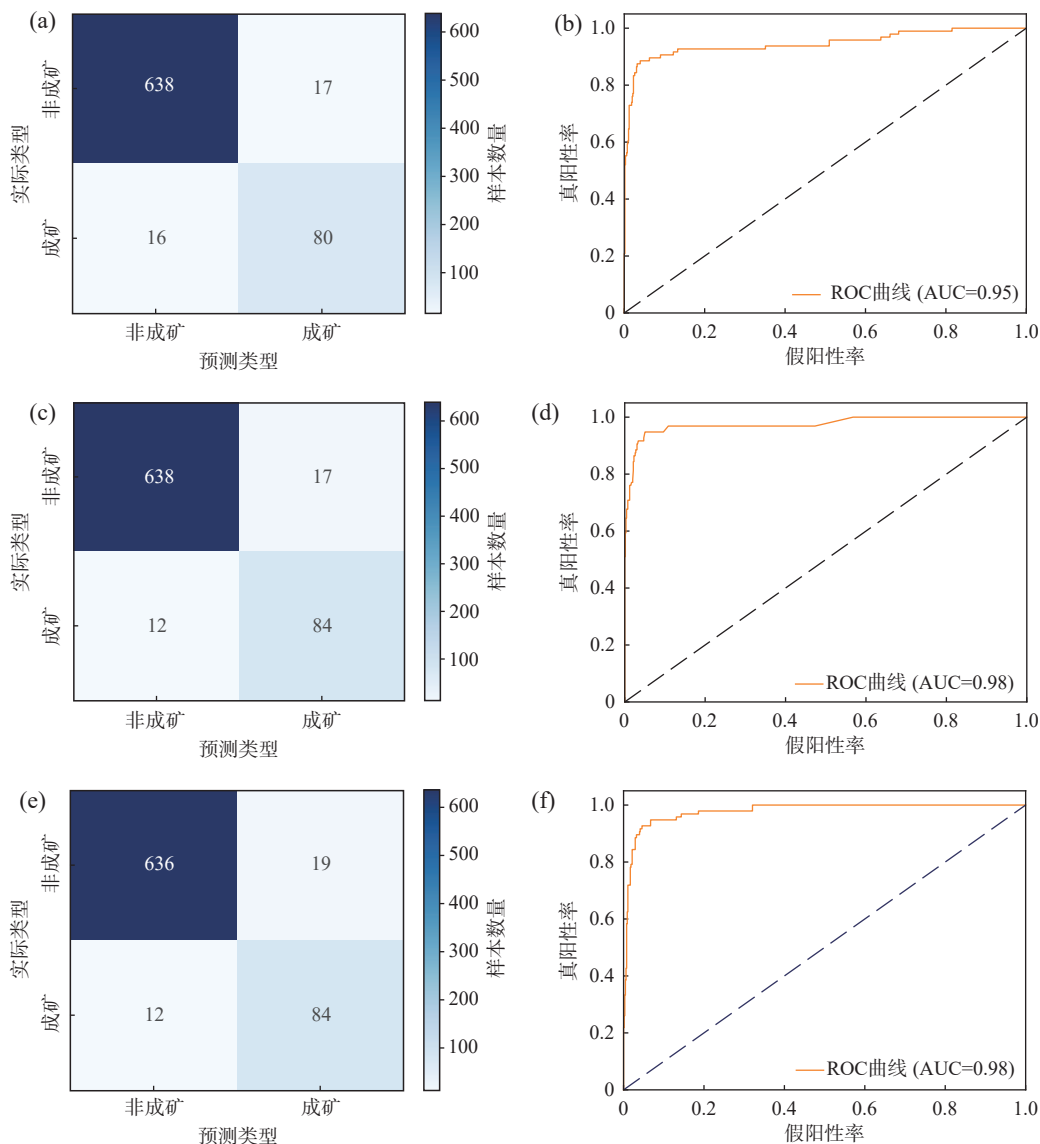
机、随机森林和极限梯度提升 3 种模型对应的混淆矩阵及 ROC 曲线(图 4), 并进一步计算了各项评价指标(表 3)。总体来看, 3 种模型均表现出较高的分类性能, 表明黑云母主量元素在区分成矿岩体与非成矿岩体方面具有较强的判别能力。

其中, 支持向量机模型的整体准确率为 0.956 1, AUC 值为 0.95; 随机森林模型的整体准确率为 0.961 4, AUC 值为 0.98; 极限梯度提升模型的整体准确率为 0.958 7, AUC 值同样为 0.98。3 种模型的准确率均大于 0.95, AUC 值均达到 0.95 及以上, 说明基于黑云母主量元素构建的机器学习模型具有较好的整体判别

效能。

从成矿样本识别效果来看, 支持向量机模型的成矿样本精确率和召回率分别为 0.824 7 和 0.833 3; 随机森林模型分别为 0.831 7 和 0.875 0; 极限梯度提升模型分别为 0.815 5 和 0.875 0。可见, 随机森林和极限梯度提升模型对成矿样本的召回率均高于支持向量机模型, 说明二者在减少成矿样本漏判方面表现更优。

从非成矿样本识别效果来看, 3 种模型的非成矿样本精确率均较高, 变化范围为 0.975 5~0.981 5, 召回率为 0.971 0~0.974 0, 表明 3 种模型对多数类样本均具有稳定的识别能力。其中, 随机森林模型在保持



a. 支持向量机混淆矩阵图; b. 支持向量机 ROC 曲线; c. 随机森林混淆矩阵图; d. 随机森林 ROC 曲线;
e. 极限梯度提升混淆矩阵图; f. 极限梯度提升 ROC 曲线

图4 独立测试集上 3 种模型的混淆矩阵与 ROC 曲线

Fig. 4 Confusion matrices and ROC curves of the three models on the independent test set

表3 三种机器学习模型在测试集上的评价指标结果

Tab. 3 Evaluation metrics results of the three machine learning models on the test set

模型	准确率	成矿样本精确率	成矿样本召回率	非成矿样本精确率	非成矿样本召回率	特异度	AUC值
支持向量机	0.956 1	0.824 7	0.833 3	0.975 5	0.974 0	0.974 0	0.95
随机森林	0.961 4	0.831 7	0.875 0	0.981 5	0.974 0	0.974 0	0.98
极限梯度提升	0.958 7	0.815 5	0.875 0	0.981 5	0.971 0	0.971 0	0.98

较高非成矿样本识别能力的同时,对成矿样本也取得了较高召回率,综合表现最佳。

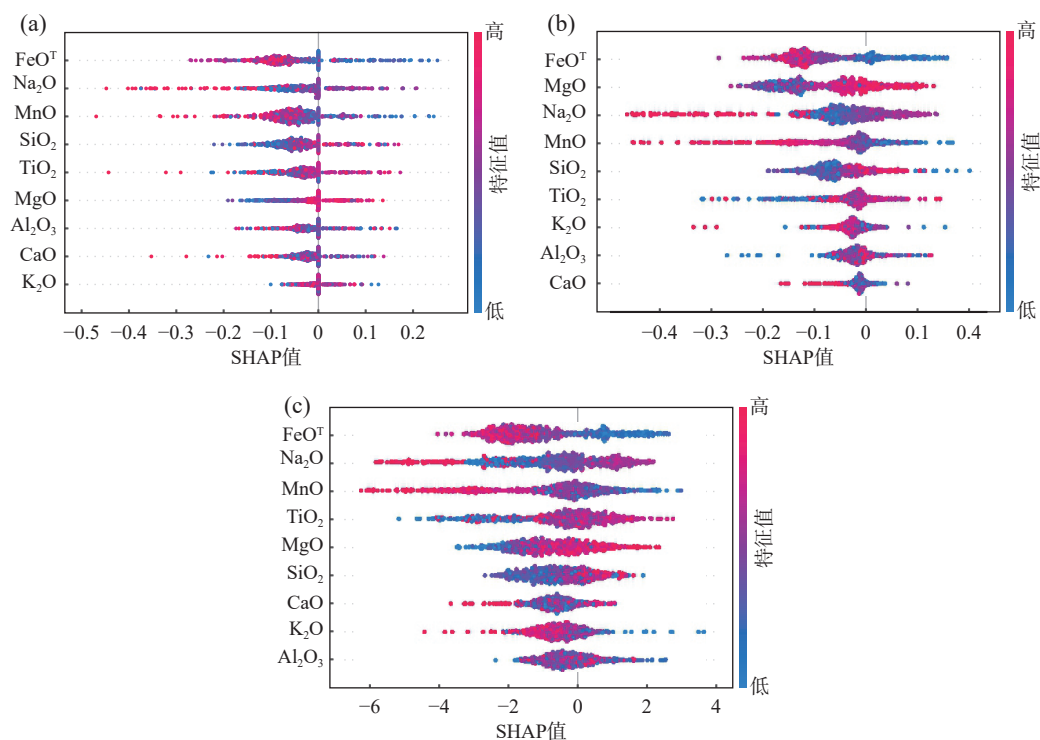
综合比较显示,3种模型均能够有效识别成矿岩体与非成矿岩体黑云母,但随机森林模型在准确率、成矿样本召回率和AUC值等指标上总体表现最优;极限梯度提升模型次之;支持向量机模型整体表现稳定,但对成矿样本的识别能力相对稍弱。因此,从整体性能和找矿应用意义来看,随机森林模型是本研究数据集上的最优分类模型。

3.2 SHAP 特征贡献分析

为进一步揭示各主量元素特征对模型输出结果的贡献大小及作用方向,并验证其地球化学意义,本文采用机器学习解释性算法 SHAP (SHapley Additive ex Planations) 对支持向量机、随机森林和极限梯度提

升3种模型进行特征贡献解释分析(Lundberg et al., 2017)。SHAP 值可量化单个特征对模型预测结果的边际贡献,揭示特征对分类结果的正负影响方向,同时验证各主量元素特征的地球化学意义。

3种模型中成矿岩体黑云母的 SHAP 蜂群图(图5)显示,各主量元素在不同模型中的重要性排序整体表现出较高一致性,说明模型对成矿判别信息的提取较为稳定。其中,FeO^T在3种模型中均位居重要性排序首位,表现出最强的特征贡献度,表明其为区分成矿与非成矿黑云母的核心特征;Na₂O和MnO也始终位于贡献度较高的位置,是区分成矿岩体和非成矿岩体黑云母的重要指标。此外,MgO与TiO₂在随机森林和极限梯度提升模型中贡献度相对更高,而在支持向量机模型中排序略低,反映出不同模型对非线性结构



a. 支持向量机模型成矿岩体黑云母的 SHAP 图解; b. 随机森林模型成矿岩体黑云母的 SHAP 图解; c. 极限梯度提升模型成矿岩体黑云母的 SHAP 图解; 图中点表示单条黑云母数据的 SHAP 值, 颜色表示对应特征取值的高低

图5 3种模型成矿岩体黑云母的 SHAP 特征贡献分析结果

Fig. 5 SHAP feature importance results for ore-forming intrusion biotite in the three models

的捕捉能力存在差异。相比之下, SiO_2 、 K_2O 、 Al_2O_3 与 CaO 的 SHAP 值整体集中于零附近, 贡献度相对较弱, 表明其对模型整体判别能力的贡献相对有限。

从特征作用方向来看, 3 种模型同样表现出较高一致性。 FeO^T 整体表现为负向控制特征, 即 FeO^T 含量越低, 对样本被判定为成矿岩体黑云母的正向贡献越大; 反之, 较高 FeO^T 值则更有利于样本被判定为非成矿岩体黑云母。 MnO 的变化趋势与 FeO^T 基本一致, 即低 MnO 含量更可能为成矿类别, 高 MnO 含量则倾向于非成矿类别。与此相反, Na_2O 表现为正向控制特征, 即较高 Na_2O 值多对应为成矿类别, 而低 Na_2O 值则多对应非成矿类别。 MgO 和 TiO_2 整体上也多表现为正向贡献, 但其贡献强度在不同模型中存在差异。

综合 3 种模型的 SHAP 分析结果, 可以归纳出成矿岩体黑云母的主要化学特征组合为低 FeO^T 、低 MnO 和高 Na_2O , 并伴随相对较高的 MgO 与 TiO_2 。其中, FeO^T 为最稳定且贡献度最高的判别特征, Na_2O 与 MnO 构成重要辅助判别指标, 而 MgO 与 TiO_2 在支持向量机和随机森林模型中表现出更为明显的区间分段特征。总体而言, 3 种模型在关键变量排序及作用方向上的高度一致性, 不仅增强了模型解释结果的可信度, 也说明黑云母主量元素中确实蕴含着能够有效反映岩体成矿潜力的重要信息。

4 讨论

4.1 模型可靠性

本研究基于全球斑岩铜矿成矿岩体和非成矿中酸性岩体黑云母主量元素数据库, 构建了基于支持向量机、随机森林和极限梯度提升 3 种机器学习模型的中酸性岩体成矿潜力判别方法。结合多种模型评价指标和机器学习解释性算法结果, 文中建立的分类方法具有较好的可靠性和稳定性。

模型评价指标显示, 3 种模型在测试集上均表现出较高的分类性能, 其准确率为 0.956 1~0.961 4, 成矿岩体黑云母的精确率为 0.815 5~0.831 7, 召回率为 0.833 3~0.875 0; 非成矿岩体黑云母的精确率为 0.975 5~0.981 5, 召回率为 0.971 0~0.974 0, AUC 值为 0.95~0.98。这些结果共同表明 3 种模型均能够有效区分成矿与非成矿岩体黑云母, 其中随机森林模型在准确率、召回率及 AUC 值等指标上总体表现最优, 说明其在本研究数据集上具有更好的综合判别能力和泛化能力。

值得注意的是, 在原始数据集存在明显类别不平衡的情况下, 笔者并未仅依赖准确率评价模型性能, 而是结合召回率与 AUC 值综合判断模型对成矿样本的识别能力, 因此所得结果更具找矿应用意义。结果表明, 随机森林模型无论是召回率还是 AUC 值, 均很高, 进一步证明随机森林模型在识别来自成矿岩体和非成矿岩体黑云母方面具有较高的有效性。

SHAP 解释结果显示, 种模型在主要特征变量的重要性排序及其作用方向上表现出较高一致性。 FeO^T 在 3 个模型中均为最重要的输入特征, Na_2O 和 MnO 对模型输出结果也十分重要。与此同时, 特征作用方向也较为稳定: 在 3 个模型中, FeO^T 含量越低, 样本越倾向于被判定为成矿岩体黑云母; MnO 表现的规律与 FeO^T 类似; 而 Na_2O 表现出相反特征, 即 Na_2O 含量越高, 样本越倾向于被判定为成矿岩体黑云母。这种跨模型解释一致性表明, 模型判别结果并非偶然统计拟合, 而是反映了数据内部较为稳定的成矿岩体与非成矿岩体黑云母之间真实存在的地球化学差异, 从解释层面进一步增强了模型结果的可信度。此外, 基于 SHAP 的解释结果也与前人认识一致。例如, Azadbakht et al. (2020) 对来自加拿大新不伦瑞克省志留纪—泥盆纪花岗岩的黑云母开展了系统研究, 发现来自非成矿岩体的黑云母 FeO^T 含量(平均值为 0.84)明显高于来自斑岩铜钼矿床成矿岩体的黑云母(0.57)。该研究也发现成矿岩体黑云母的 MnO 含量(0.16)显著低于非成矿岩体(0.71)。这些研究进一步证明文中构建的机器学习模型具有可靠性。

更为重要的是, 机器学习识别出的成矿岩体信息也可被已知的斑岩铜矿成矿岩浆形成过程所解释。大量的岩浆表明, 斑岩铜矿成矿岩浆通常形成于较高氧逸度条件下, 而较高氧逸度环境有利于 S、Cu 等成矿元素在岩浆—流体体系中的迁移、富集与沉淀 (Heinrich, 2024)。黑云母的 FeO^T 和 MnO 含量可能就与岩浆的氧化还原程度有关。Fe 和 Mn 均属于变价元素, 在熔体中分别以 +2 价和 +3 价以及 +2 价和 +4 价的形式存在。对于黑云母, Fe 主要以 +2 价的形式存在于黑云母矿物晶格中, 与 Mg^{2+} 形成类质同象替代; 而 Mn 则主要以 +2 的形式进入黑云母晶格。相比于 Fe^{2+} 和 Mn^{2+} , Fe^{3+} 和 Mn^{4+} 则较难进入到黑云母中。相比于非成矿岩浆, 成矿岩浆由于氧逸度高, Fe 和 Mn 主要以 Fe^{3+} 和 Mn^{4+} 的形式存在, 这就导致从成矿岩浆中结晶出的黑云母具有较低的 FeO^T 和 MnO 含量, 而非

成矿岩浆由于氧逸度低,Fe和Mn主要以 Fe^{2+} 和 Mn^{2+} 的形式存在,更容易进入黑云母晶格,从而导致黑云母富集 FeO^{T} 和 MnO 。

综上所述,文中所建立的基于黑云母成分的成矿岩体智能识别方法不仅准确率高,而且能够被已知地质过程所解释,可信度高,应用前景大。

4.2 存在问题和展望

本研究构建了基于黑云母成分的斑岩铜矿成矿岩体智能识别方法,为斑岩铜矿找矿勘查提供了一种绿色、高效且经济的新思路,也为深化斑岩铜矿成矿理论认识提供了新的研究视角。但目前的研究还存在一些问题,尚需在未来的工作中进一步完善提升。

本研究所采用的成矿岩体黑云母数据相对有限,仅占全部数据的约12.6%。采用SMOTE过采样技术对少数类样本进行了扩增,在一定程度上缓解了类别不平衡现象。然而,已有研究表明,极端不平衡的样本库即便经过类不平衡处理,仍会对训练获得的机器学习模型稳定性与泛化能力产生影响。未来,随着来自成矿岩体黑云母数据报道的增多,将能够在一定程度上降低数据库类不平衡现象,从而提升模型预测精度与可靠性。

本研究仅选取9种主量元素作为模型训练特征。尽管主量元素主要受控于原始岩浆成分,受其他因素干扰小,在理论上能够更好的指示岩浆原始信息,但其组成仍可能受后期热液蚀变影响,从而在一定程度上削弱原始岩浆信号。因此,未来研究中可进一步结合黑云母微量元素(如Li、V、Co、Ni等)及卤素元素含量,构建更加丰富的多维特征空间,以提高模型预测结果的准确性。

需要指出的是不同类型矿物在岩浆演化过程中记录的信息存在差异,单一矿物可能难以全面反映复杂成矿系统的多阶段演化过程。未来可探索黑云母与锆石、磷灰石等多种指针矿物联合使用的方法,通过多矿物特征融合进一步提升成矿岩体识别精度,为深部找矿提供更加可靠的技术支撑。

5 结论

(1)系统汇编了全球斑岩铜矿成矿岩体与全球非成矿中酸性岩体的黑云母主量元素数据,并基于3种机器学习模型,构建了识别斑岩铜矿成矿潜力的智能分类方法。机器学习训练结果显示,3种模型均具有

较高的准确率(0.9561~0.9614),并且随机森林模型在所有评价指标上表现最优。

(2)SHAP分析表明,黑云母 FeO^{T} 含量在3个模型中均对输出结果发挥最为关键作用, MnO 、 Na_2O 也对输出结果影响大,来自成矿岩体的黑云母整体表现为低 FeO^{T} 、低 MnO 和高 Na_2O 的特征,其中 FeO^{T} 和 MnO 的特征主要受岩浆氧逸度控制。

(3)结合机器学习方法,黑云母主量元素特征不仅能够有效记录原始岩浆信息,还能够用来准确约束岩浆成矿潜力。将该方法应用于其他指针矿物,将进一步提高岩浆成矿潜力识别的准确率和效率。

利益冲突声明(Conflict of Interests):

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

参考文献(References):

- 曹康. 云南普朗超大型斑岩铜金矿床岩浆作用与矿床成因[D]. 北京: 中国地质大学(北京), 2018.
- CAO Kang. Magmatism and genesis of the giant Pulang porphyry Cu-Au deposit, Yunnan, SW China[D]. Beijing: China University of Geosciences (Beijing), 2018.
- 陈念. 秘鲁北部中新世成矿作用[D]. 北京: 中国地质大学(北京), 2022.
- CHEN Nian. Miocene metallogenesis in northern Peru: A case study of the Rio Blanco porphyry Cu-Mo deposit[D]. Beijing: China University of Geosciences (Beijing), 2022.
- 褚小磊. 吉黑东部陆缘中生代斑岩型铜成矿作用研究[D]. 长春: 吉林大学, 2022.
- CHU Xiaolei. Mesozoic porphyry copper mineralization in the eastern continental margin of Jilin and Heilongjiang Province[D]. Changchun: Jilin University, 2022.
- 邓昌州. 大兴安岭北部中生代斑岩铜矿: 成岩与成矿[D]. 长春: 吉林大学, 2019.
- DENG Changzhou. Petrology and metallogenesis of the porphyry Cu deposits in the northern Great Xing'an Range[D]. Changchun: Jilin University, 2019.
- 郭广慧, 钟世华, 李三忠, 等. 运用机器学习和锆石微量元素构建花岗岩成矿潜力判别图解: 以东昆仑祁漫塔格为例[J]. 西北地质, 2023, 56(6): 57-70.
- GUO Guanghui, ZHONG Shihua, LI Sanzhong, et al. Constructing Discrimination Diagrams for Granite Mineralization Potential by Using Machine Learning and Zircon Trace Elements: Example from the Qimantagh, East Kunlun[J]. Northwestern Geology, 2023, 56(6): 57-70.
- 郝金华, 陈建平, 田永革, 等. 青海纳日贡玛斑岩铜(钼)矿含矿

- 斑岩矿物学特征及成岩成矿意义[J]. 地质与勘探, 2010, 46(3): 367-376.
- HAO Jinhua, CHEN Jianping, TIAN Yongge, et al. Mineralogical features of porphyries in the Narigongma Mo (-Cu) deposit, southern Qinghai and their implications for petrogenesis and mineralization[J]. *Geology and Exploration*, 2010, 46(3): 367-376.
- 李华伟, 杨志明. 岩浆锆石和磷灰石矿物化学及在斑岩矿床领域的应用[J]. 地质学报, 2023, 97(4): 973-1001.
- LI Huawei, YANG Zhiming. Geochemistry of magmatic zircon and apatite and its applications in porphyry deposits[J]. *Acta Geologica Sinica*, 2023, 97(4): 973-1001.
- 李守奎. 滇西北铜厂沟斑岩钼(铜)矿床岩浆-流体作用与成矿机制[D]. 昆明: 昆明理工大学, 2023.
- LI Shoukui. Magmatism, Hydrothermalism and Metallogenic Mechanism of the Tongchanggou Porphyry Mo (Cu) Deposit in Northwestern Yunnan[D]. Kunming: Kunming University of Science and Technology, 2023.
- 李万滢. 安徽铜陵舒家店斑岩铜矿床流体包裹体与蚀变矿物学研究[D]. 合肥: 中国科学技术大学, 2024.
- LI Wanying. Geochemistry of fluid inclusions and altered minerals in Shujiadian porphyry copper deposit, Tongling, Anhui Province[D]. Hefei: University of Science and Technology of China, 2024.
- 刘嘉情, 钟世华, 李三忠, 等. 基于机器学习的磷灰石源岩识别方法[J]. 大地构造与成矿学, 2025, 49(2): 467-480.
- LIU Jiaqing, ZHONG Shihua, LI Sanzhong, et al. A Machine Learning-based Method for Identifying Apatite Source Rocks[J]. *Geotectonica et Metallogenia*, 2025, 49(2): 467-480.
- 刘嘉情, 钟世华, 李三忠, 等. 基于机器学习和全岩成分识别东昆仑祁漫塔格斑岩-矽卡岩矿床成矿岩体和贫矿岩体[J]. 西北地质, 2023, 56(6): 41-56.
- LIU Jiaqing, ZHONG Shihua, LI Sanzhong, et al. Identification of mineralized and barren magmatic rocks for the porphyry-skarn deposits from the Qimantagh, East Kunlun: based on machine learning and whole-rock compositions[J]. *Northwestern Geology*, 2023, 56(6): 41-56.
- 刘诗文. 岩浆成因黑云母对成矿作用的标识[D]. 北京: 中国地质大学(北京), 2015.
- LIU Shiwen. The Identification of Magmatic Origin Biotite to Mineralization[D]. Beijing: China University of Geosciences (Beijing), 2015.
- 刘学龙, 李文昌, 尹光侯, 等. 云南格咱岛弧普朗斑岩型铜矿年代学、岩石矿物学及地球化学研究[J]. 岩石学报, 2013, 29(9): 3049-3064.
- LIU Xuelong, LI Wenchang, YIN Guanghou, et al. The geochronology, mineralogy and geochemistry study of the Pulang porphyry copper deposits in Geza arc of Yunnan Province[J]. *Acta Petrologica Sinica*, 2013, 29(9): 3049-3064.
- 潘北斗, 李宝, 王福臣, 等. 不同类型矿床(点)空间关联关系定量研究——以中南半岛及临近地区为例[J]. 地质与勘探, 2022, 58(3): 532-544.
- PAN Beidou, LI Bao, WANG Fuchen, et al. Quantitative study on spatial correlation of different types of deposits (occurrences) — A case study of Indochina Peninsula and adjacent areas[J]. *Geology and Exploration*, 2022, 58(3): 532-544.
- 唐攀, 唐菊兴, 郑文宝, 等. 西藏拉抗俄斑岩铜钼矿床黑云母矿物化学特征[J]. 地学前缘, 2017, 24(5): 265-282.
- TANG Pan, TANG Juxing, ZHENG Wenbao, et al. Mineral chemistry of hydrothermal biotites from the Lakang'e porphyry Cu-Mo deposit, Tibet[J]. *Earth Science Frontiers*, 2017, 24(5): 265-282.
- 唐攀. 西藏冈底斯成矿带斑岩铜多金属矿床黑云母矿物学特征——对成矿环境的指示[D]. 成都: 成都理工大学, 2018.
- TANG Pan. Mineralogical characteristics of biotite from the porphyry copper-polymetallic deposits in Gangdese metallogenic belt, Tibet: Implication for metallogenetic environment[D]. Chengdu: Chengdu University of Technology, 2018.
- 王海然, 赵红格, 乔建新, 等. 锆石 U-Pb 同位素测年原理及应用[J]. 地质与资源, 2013, 22(3): 229-232+242.
- WANG Hairan, ZHAO Hongge, QIAO Jianxin, et al. Theory and application of zircon U-Pb isotope dating technique[J]. *Geology and Resources*, 2013, 22(3): 229-232+242.
- 王玉峰, 肖丙建, 宋明春, 等. 鲁西大井头地区金刚石找矿指示矿物特征及其研究意义[J]. 地质学报, 2020, 94(9): 2712-2727.
- WANG Yufeng, XIAO Bingjian, SONG Mingchun, et al. Characteristics and research significance of diamond prospecting indicator minerals in the Dajingtou area, western Shandong Province[J]. *Acta Geologica Sinica*, 2020, 94(9): 2712-2727.
- 吴松. 西藏冈底斯朱诺超大型斑岩铜矿床: 岩浆与成矿[D]. 北京: 中国地质大学(北京), 2016.
- WU Song. The super-large Zhunuo porphyry Cu deposit in the Gangdese belt, southern Tibet: magmatism and mineralization[D]. Beijing: China University of Geosciences (Beijing), 2016.
- 肖庆玲, 周涛发, 袁峰, 等. 安徽茶亭铜金矿床赋矿石英闪长玢岩黑云母成分特征及成岩成矿指示意义[J]. 矿床地质, 2018, 37(6): 1179-1194.
- XIAO Qingling, ZHOU Taofa, YUAN Feng, et al. Compositional characteristics, petrogenesis and metallogenetic significance of biotite from the ore-bearing quartz diorite porphyry in Chating Cu-Au deposit, Anhui Province[J]. *Mineral Deposit*, 2018, 37(6): 1179-1194.
- 张少颖. 普朗斑岩型铜金矿床岩浆-热液演化与成矿模式[D]. 北京: 中国地质大学(北京), 2022.
- ZHANG Shaoying. Magmatic-hydrothermal evolution and mineralization model of the Pulang porphyry copper-gold deposit[D]. Beijing: China University of Geosciences (Beijing), 2022.
- 周杰虎, 刘学龙, 周云满, 等. 滇西北红牛-红山铜矿床复式岩体黑云母特征及成岩成矿指示意义[J]. 矿床地质, 2024, 43(3): 503-526.
- ZHOU Jiehu, LIU Xuelong, ZHOU Yunman, et al. Geochemical characteristics of biotite from composite pluton and its implications for diagenesis and mineralization in Hongniu-Hongshan copper deposit, northwestern Yunnan[J]. *Mineral Deposit*, 2024, 43(3): 503-526.

- Azadbakht Z, Lentz D R, McFarlane C R, et al. Using magmatic biotite chemistry to differentiate barren and mineralized Silurian–Devonian granitoids of New Brunswick, Canada[J]. *Contributions to Mineralogy and Petrology*, 2020, 175(7): 69.
- Baidya A S, Maiti G, Mondal S, et al. Biotite chemistry as an indicator of hydrothermal deposit types and fluid sources: Insights from big data compilation, multivariate statistical analysis, and machine learning[J]. *Journal of Geochemical Exploration*, 2024, 259: 107442.
- Ballard J R, Palin J M, Campbell I H. Relative oxidation states of magmas inferred from Ce (IV)/Ce (III) in zircon: application to porphyry copper deposits of northern Chile[J]. *Contributions to Mineralogy and Petrology*, 2002, 144(3): 347–364.
- Barber N D, Edmonds M, Jenner F, et al. Amphibole control on copper systematics in arcs: Insights from the analysis of global datasets[J]. *Geochimica et Cosmochimica Acta*, 2021, 307: 192–211.
- Bouzari F, Hart C J R, Bissig T, et al. Hydrothermal alteration revealed by apatite luminescence and chemistry: A potential indicator mineral for exploring covered porphyry copper deposits[J]. *Economic Geology*, 2016, 111(6): 1397–1410.
- Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- Carrasco-Godoy C, Campbell I H, Cajal Y. Quantifying the criteria used to identify zircons from ore-bearing and barren systems in porphyry copper exploration[J]. *Economic Geology*, 2024, 119(5): 1035–1058.
- Chawla N V, Bowyer K W, Hall L O, et al. SMOTE: synthetic minority over-sampling technique[J]. *Journal of artificial intelligence research*, 2002, 16: 321–357.
- Chen T, Guestrin C. Xgboost: A scalable tree boosting system[A]. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*[C]. 2016, 785–794.
- Cortes C, Vapnik V. Support-vector networks[J]. *Machine Learning*, 1995, 20(3): 273–297.
- Fawcett T. An introduction to ROC analysis[J]. *Pattern Recognition Letters*, 2006, 27(8): 861–874.
- Ge Y Z, Zhang Z J, Zhou Y Z, et al. Explainable machine learning reveals apatite fertility and porphyry copper mineralization processes in the syn-and post-subduction settings[J]. *Ore Geology Reviews*, 2025, 183: 106679.
- Hanley J A, McNeil B J. The meaning and use of the area under a receiver operating characteristic (ROC) curve[J]. *Radiology*, 1982, 143(1): 29–36.
- He H, Garcia E A. Learning from imbalanced data[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2009, 21(9): 1263–1284.
- Heidari M, Zarasvandi A, Rezaei M, et al. Physicochemical attributes of parental magma in collisional porphyry copper systems; using biotite chemistry, case study: Chahfiruzeh porphyry copper deposit[J]. *Journal of Economic Geology*, 2019, 10(2): 561–586.
- Heinrich C A. The chain of processes forming porphyry copper deposits—An invited paper[J]. *Economic Geology*, 2024, 119(4): 741–769.
- Henry D J, Guidotti C V, Thomson J A. The Ti-saturation surface for low-to-medium pressure metapelitic biotites: Implications for geothermometry and Ti-substitution mechanisms[J]. *American Mineralogist*, 2005, 90(2–3): 316–328.
- Hoskin P W O, Schaltegger U. The composition of zircon and igneous and metamorphic petrogenesis[J]. *Reviews in Mineralogy and Geochemistry*, 2003, 53: 27–62.
- Huang M L, Bi X W, Hu R Z, et al. Linking porphyry Cu formation to tectonic change in postsubduction settings: a case study from the Giant Yulong Belt, Eastern Tibet[J]. *Economic Geology*, 2024, 119(2): 279–304.
- Huang M L, Zhu J J, Chiaradia M, et al. Apatite volatile contents of porphyry Cu deposits controlled by depth-related fluid exsolution processes[J]. *Economic Geology*, 2023, 118(5): 1201–1217.
- Idrus A. Petrography and Mineral Chemistry of Magmatic and Hydrothermal Biotite in Porphyry Copper-Gold Deposits: A Tool for Understanding Mineralizing Fluid Compositional Changes During Alteration Processes[J]. *Indonesian Journal on Geoscience*, 2018, 5(1): 47–64.
- Ishihara S. The magnetite-series and ilmenite-series granitic rocks[J]. *Mining Geology*, 1977, 27(145): 293–305.
- Japkowicz N, Stephen S. The class imbalance problem: A systematic study[J]. *Intelligent Data Analysis*, 2002, 6(5): 429–449.
- Kan J, Qin K, Wang L, et al. The influence of fluid-exsolving depth on mineralization quality: Evidence from biotite and zircon mineralogy and fluid inclusions from the 460 Gaodi porphyry Mo-Cu deposit, NE China[J]. *Minerals*, 2023, 13(5): 699.
- Kazemi K, Modabberi S, Xiao Y, et al. Geochronology, whole-rock geochemistry, SrNd isotopes, and biotite chemistry of the Deh-Bala intrusive rocks, Central Urumieh-Dokhtar Magmatic Arc (Iran): Implications for magmatic processes and copper mineralization[J]. *Lithos*, 2022, 408: 106544.
- Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]. *IJCAI*, 1995, 14(2): 1137–1145.
- Kotsiantis S, Kanellopoulos D, Pintelas P. Handling imbalanced datasets: A review[J]. *GESTS International Transactions on Computer Science and Engineering*, 2006, 30(1): 25–36.
- Lee R G, Plouffe A, Ferbey T, et al. Recognizing porphyry copper potential from till zircon composition: a case study from the Highland Valley porphyry district, south-central British Columbia[J]. *Economic Geology*, 2021, 116(4): 1035–1045.
- Lehtonen M, Lahaye Y, O'Brien H, et al. Novel technologies for indicator mineral-based exploration[J]. *Geological Survey of Finland, Special Paper*, 2015, 57: 23–62.
- Lin S, Hu K, Cao J, et al. Geochemistry and origin of hydrothermal apatite in Carlin-type Au deposits, southwestern China (Gao-long deposit)[J]. *Ore Geology Reviews*, 2023, 154: 105312.
- Liu Y, Gao J F, Qi L, et al. Textural and compositional variation of mica from the Dexing porphyry Cu deposit: constraints on the behavior of halogens in porphyry systems[J]. *Acta Geochimica*,

- 2023, 42(2): 221–240.
- Loferski P J, Ayuso R A. Petrography and mineral chemistry of the composite Deboullie pluton, northern Maine, USA: Implications for the genesis of Cu Mo mineralization[J]. Chemical Geology, 1995, 123(1–4): 89–105.
- Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. Advances in Neural Information Processing Systems, 2017, 30.
- Luo L, Chen G, Li Z. Identifying tectonic settings of porphyry copper deposits using zircon trace elements—A semi-supervised machine learning method[J]. Ore Geology Reviews, 2024, 171: 106170.
- Maydagán L, Franchini M, Chiaradia M, et al. Petrogenesis of Quebrada de la Mina and Altar North porphyries (Cordillera of San Juan, Argentina): Crustal assimilation and metallogenic implications[J]. Geoscience frontiers, 2017, 8(5): 1135–1159.
- Moshefi P, Hosseinzadeh M R, Moayyed M, et al. Comparative study of mineral chemistry of four biotite types as geochemical indicators of mineralized and barren intrusions in the Sungun Porphyry Cu-Mo deposit, northwestern Iran[J]. Ore Geology Reviews, 2018, 97: 1–20.
- Nathwani C L, Wilkinson J J, Fry G, et al. Machine learning for geochemical exploration: classifying metallogenic fertility in arc magmas and insights into porphyry copper deposit formation[J]. Mineralium Deposita, 2022, 57(7): 1143–1166.
- Pan Y, Dong G, Tsunogae T, et al. Contrasting mineralized and barren porphyries in the Zhongdian Arc, insights from biotite and apatite compositions and halogen fugacity[J]. Scientific Reports, 2024, 14(1): 12110.
- Plouffe A, Lee R G, Byrne K, et al. Tracing Detrital Epidote Derived from Alteration Halos to Porphyry Cu Deposits in Glaciated Terrains: The Search for Covered Mineralization[J]. Economic Geology, 2024, 119(2): 305–329.
- Powers D M W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation[J]. arXiv preprint arXiv: 2010.16061, 2020.
- Qiu K F, Zhou T, Chew D, et al. Apatite trace element composition as an indicator of ore deposit types: A machine learning approach[J]. American Mineralogist, 2024, 109(2): 303–314.
- Richards J P. Magmatic to hydrothermal metal fluxes in convergent and collided margins[J]. Ore Geology Reviews, 2011, 40(1): 1–26.
- Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M, et al. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines[J]. Ore geology reviews, 2015, 71: 804–818.
- Schaarschmidt A, Klemm R, Regelous M, et al. The formation of shoshonitic magma and its relationship to porphyry-type mineralisation: the Maronia pluton in NE Greece[J]. Lithos, 2021, 380: 105911.
- Sillitoe R H. Porphyry copper systems[J]. Economic Geology, 2010, 105(1): 3–41.
- Snoek J, Larochelle H, Adams R P. Practical bayesian optimization of machine learning algorithms[J]. Advances in Neural Information Processing Systems, 2012, 25.
- Sun W, Huang R, Li H, et al. Porphyry deposits and oxidized magmas[J]. Ore Geology Reviews, 2015, 65: 97–131.
- Taghavi A, Maanijou M, Lentz D R, et al. Biotite compositions and geochemistry of porphyry-related systems from the central Urumieh Dokhtar Magmatic Belt, western Yazd, Iran: Insights into mineralization potential[J]. Lithos, 2022, 412: 106593.
- Tang P, Tang J, Lin B, et al. Mineral chemistry of magmatic and hydrothermal biotites from the Bangpu porphyry Mo (Cu) deposit, Tibet[J]. Ore Geology Reviews, 2019, 115: 103122.
- Tatnell L, Anenburg M, Loucks R. Porphyry copper deposit formation: identifying garnet and amphibole fractionation with REE pattern curvature modeling[J]. Geophysical Research Letters, 2023, 50(14): e2023GL103525.
- Vapnik V N. An overview of statistical learning theory[J]. IEEE Transactions on Neural Networks, 1999, 10(5): 988–999.
- Wilkinson J J, Baker M J, Cooke D R, et al. Exploration targeting in porphyry Cu systems using propylitic mineral chemistry: A case study of the El Teniente deposit, Chile[J]. Economic Geology, 2020, 115(4): 771–791.
- Yu K, Li G, Zhao J, et al. Biotite composition as a tracer of fluid evolution and mineralization center: A case study at the Qulong porphyry Cu-Mo deposit, Tibet[J]. Mineralium Deposita, 2022, 57(6): 1047–1069.
- Zarasvandi A, Heidari M, Raith J, et al. Geochemical characteristics of collisional and pre-collisional porphyry copper systems in Kerman Cenozoic Magmatic Arc, Iran: Using plagioclase, biotite and amphibole chemistry[J]. Lithos, 2019, 326: 279–297.
- Zhang J, Chai J, Zhang L, et al. Lower crustal copper enrichment and mobilization indicated by Late Permian amphibole gabbros and granodiorites in the eastern Kunlun Orogen, northern Tibet[J]. Lithos, 2024, 482: 107740.
- Zhang W. Petrological and Metallogenic studies of the Nashwaak Granite and felsic dykes associated with the Sisson Brook W-Mo-(Cu) deposit, west-central New Brunswick, Canada[J]. Geology, 2015: 134203054.
- Zhang W, Lentz D R, Thorne K G, et al. Geochemical characteristics of biotite from felsic intrusive rocks around the Sisson Brook W-Mo-Cu deposit, west-central New Brunswick: An indicator of halogen and oxygen fugacity of magmatic systems[J]. Ore Geology Reviews, 2016, 77: 82–96.