



DOI:10.12404/j.issn.1671-1815.2402216

引用格式:肖昆,焦常伟,杨亚新,等.基于机器学习的铀矿层岩性测井评价研究进展[J].科学技术与工程,2025,25(12):4827-4839.

Xiao Kun, Jiao Changwei, Yang Yaxin, et al. Research progress on lithologic logging evaluation of uranium ore layers based on machine learning[J]. Science Technology and Engineering, 2025, 25(12): 4827-4839.

天文学、地球科学

基于机器学习的铀矿层岩性测井评价研究进展

肖昆¹, 焦常伟^{1*}, 杨亚新¹, 黄笑², 王殿学², 段忠义¹, 徐艺宸¹

(1. 东华理工大学核资源与环境国家重点实验室, 南昌 330013; 2. 核工业二四三大队, 赤峰 024000)

摘要 近年来,人工智能在各个领域展现出了强大的模式识别和分类能力,为岩性识别提供了新的思路。从支持向量机、神经网络、集成学习这3种方法出发,综述这些机器学习算法的基本原理、优缺点及其在铀矿层岩性识别领域的研究进展和应用情况。结果表明:机器学习通过训练模型可以有效识别出测井数据与不同岩性之间的关联,将岩性识别过程转化为机器学习的过程,可以极大地提高岩性识别自动化程度和识别准确率,具有重要的现实意义和广阔的发展前景。

关键词 铀矿层; 岩性识别; 机器学习; 分类问题; 测井评价

中图分类号 P631.8; 文献标志码 A

Research Progress on Lithologic Logging Evaluation of Uranium Ore Layers Based on Machine Learning

XIAO Kun¹, JIAO Chang-wei^{1*}, YANG Ya-xin¹, HUANG Xiao², WANG Dian-xue²,
DUAN Zhong-yi¹, XU Yi-chen¹

(1. Key Laboratory of Nuclear Resources and Environment, East China University of Technology, Nanchang 330013, China;
2. Nuclear Industry Group No. 243, Chifeng 024000, China)

[Abstract] In recent years, artificial intelligence has demonstrated strong pattern recognition and classification capabilities across various fields, providing new insights for lithology identification. Starting from three methods: support vector machines, neural networks, and ensemble learning, the basic principles, advantages and disadvantages of these machine learning algorithms were reviewed, as well as their research progress and application in the field of uranium ore bed lithology identification. The results show that machine learning can effectively identify the correlation between logging data and different lithologies through model training, transforming the process of lithology identification into a machine learning process. This can greatly improve the automation level and accuracy of lithology identification, holding significant practical importance and a broad development prospect.

[Keywords] uranium deposit; lithology identification; machine learning; classification problem; logging evaluation

岩性是岩石命名的基础,是颜色、结构、胶结物和矿物成分的总体反映,地层内岩性的准确识别对储层划分和矿藏评价起着关键作用^[1]。地球物理测井具有垂直分辨率高、连续性好、数据采集方便等优点,是当下常用的获取地层内信息的手段^[2]。利用获取的测井资料对地层岩性识别是一项基础工作,传统岩性识别的方法主要以交会图技术为主。然而,随着地下资源岩性矿藏地质结构日益复杂化,对岩性识别也提出了更高的要求^[3]。在复杂

地层条件下,有时一条或几条测井曲线无法准确、快速地识别出岩性,因为大多数的原生岩相有着相似的测井响应特征,难以在交会图内进行有效的聚类和识别。

随着人工智能的迅速发展,机器学习为岩性识别提供了新的研究方向^[4-5]。反向传播(back propagation, BP)算法的提出为神经网络在计算机科学和机器学习领域的广泛应用奠定了基础。Samuel等^[6]通过对比试验证实了神经网络在岩性分类方

收稿日期:2024-03-28; 修订日期:2025-01-21

基金项目:江西省主要学科学术和技术带头人培养计划(20204BCJ23027);核资源与环境国家重点实验室联合创新基金(2022NRE-LH-18);江西省自然科学基金(20232BAB203072)

第一作者:肖昆(1987—),男,汉族,江西抚州人,博士,副教授。研究方向:地球物理测井理论与方法。E-mail:xiaokun0626@163.com。

*通信作者:焦常伟(1998—),男,汉族,河南周口人,硕士研究生。研究方向:铀矿资源测井评价。E-mail:1821155424@qq.com。

投稿网址:www.stae.com.cn

面的有效性。基于神经网络方法,刘明军^[7]开发了一款岩性识别软件系统,并对其进行了检验,该系统展现出了较高的识别速度和准确率。Cortes 等^[8]提出了支持向量机(support vector machine, SVM),其学习策略就是间隔最大化,对于非线性的分类同样具有较高适用性。Smirnoff 等^[9]基于测井和地表地质等数据,利用支持向量机方法实现了区域的三维地质建模。牟丹等^[10]整合了研究区内的岩心及岩矿鉴定资料与测井数据,应用测井数据建立支持向量机两分类和多分类岩性识别模型,通过对 4 个测试井中的 800 个岩性数据进行识别,达到 82.3% 的准确率。2001 年,随机森林(random forest, RF)等集成算法的出现,为地质勘探和资源开发提供了更有效的工具^[11]。为解决碳酸盐岩储层复杂岩性识别方法存在的精度低、泛化能力差、预测不稳定等问题,周雪晴等^[12]提出了一种基于粗糙集-随机森林算法的复杂岩性识别方法,显著提升了预测准确率和稳定性。Ao 等^[13]采用修剪随机森林(pruning random forest, PRF)算法在中国渤海西部地震资料上的实验中展现出卓越的岩性预测准确度和鲁棒性。上述研究证明,机器学习已成为岩性识别和自动分层等测井解释中的有力工具,分析研究测井数据岩性识别的机器学习算法是当前测井解释工作的重要课题^[14]。鉴于此,对目前基于机器学习的铀矿层岩性识别方法进行总结,以期能够掌握铀矿层岩性识别的研究方法和未来的发展趋势。

1 机器学习算法与铀矿层岩性测井评价研究进展

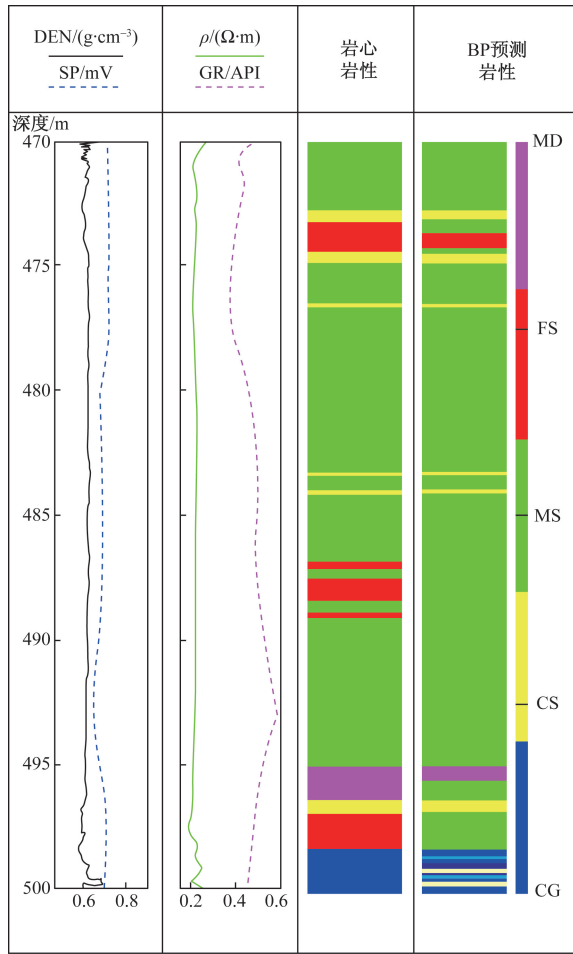
机器学习作为实现人工智能的方法,通过用算法来解析数据,从中总结与学习,然后对真实世界中的事件做出决策和预测^[15]。近年来,随着人工智能的深入研究,机器学习技术在孔隙度预测、划分油气层、岩性识别、渗透率预测等方面的应用越来越广泛。解决岩性分类任务主要分为 3 个阶段:预处理、分类、后处理。预处理对识别质量有着重要影响,因此必须选择适当的预处理方法。分类器应该选取非线性的,并能够识别模糊、重叠的类别。为了优化后处理阶段的岩性识别质量,必须考虑与邻近井的相关性。

1.1 神经网络

神经网络也被称为人工神经网络,通过模仿生物神经元之间相互传递信号的方式,达到学习经验的目的。神经网络最早是由 McCulloch 等^[16]于 20 世纪 40 年代提出,具有自学习、自组织、自适应性的特点,使得网络可以处理未知的系统。

与传统的机器学习方法相比,以人工神经元为基础的神经网络方法能够对更多的数据进行更深层次的信息挖掘。

在铀矿领域,神经网络目前多用于对勘探过程中的岩性与铀异常层进行预测,并取得了良好的应用效果^[17]。许建华^[18]应用神经网络模型自动识别岩性,取得了良好的解释效果。为解决传统神经网络算法难以收敛的问题,祖秀兰等^[19]引入学习参数调整算法和隐节点调整算法,提出了一个改进的神经网络模型,证明了提高神经网络预测能力的关键在于提高样本的“质量”,为神经网络岩性识别研究提供了新的思路。李继安^[20]建立了基于 BP 算法改进的神经网络模型,通过统计学的方法选择视电阻率、密度和井径 3 个参数作为模型的输入特征,结果显示,岩性识别的平均符合率达到 84.60%,应用效果符合预期,但该方法在其他地区的适用性尚待进一步验证。Muhamediyev 等^[21]使用人工神经网络等自学习系统进行铀矿测井数据解释,初步研究结果显示,利用人工神经网络可以实现 52%~73% 的可解释数据与实验结果的符合率,通过一个统一的系统同时使用多种分类算法,岩识别的质量得以进一步提高。Muhamediyev 等^[22]以哈萨克斯坦的 Inkai 铀矿为研究对象,分析比较神经网络、K 近邻和朴素贝叶斯等算法在铀矿层岩性识别上的应用效果,实验结果表明,几乎在所有情况下,神经网络方法的精度都比其他两种方法高 5%~15%,而且神经网络算法也表现出了更强的可训练性和可改进性。为了快速获取铀矿矿集区中铀矿异常分布信息,康乾坤等^[23]采用神经网络模型对松辽盆地的特定铀矿区进行异常识别,研究结果显示,模型在识别铀矿的异常层和矿化层方面取得了 86.55% 的准确率,该方法在快速获取未知钻孔信息方面展现出良好的发展潜力。易敏等^[24]基于深度学习框架 Tensorflow 构建一维卷积神经网络模型,以新疆白杨河铀矿床 8 个钻孔测井数据为研究对象进行蚀变矿物识别,模型测试精确率达 87%,而且模型具有较好的稳定性,但网络本身的局限性使得该模型的可推广性受到限制。Sun 等^[25]提取密度(DEN)、自然伽马(GR)、电阻率(ρ)、自然电位(SP)、井径(CAL)等测井曲线作为模型输入特征,采用神经网络模型对内蒙古塔然高勒铀矿床进行岩性智能识别,识别结果与岩心对比发现(图 1^[25]),该方法准确度较高且比岩心更加精细。通过混淆矩阵对模型的可靠性进行分析(表 1^[25]),综合岩性分类准确率为 88.31%,其中砾岩的预测精度达到 100%。由于在训练模型时,泥岩和细砂岩的岩性样本相对较少,因



CS 为粗粒砂岩;MS 为中粒砂岩;FS 为细粒砂岩;
CG 为钙质砂岩;MD 为泥岩

图 1 基于神经网络模型岩性识别成果图^[25]

Fig. 1 Based on neural network model rock recognition results map^[25]

表 1 塔然高勒铀矿岩性识别混淆矩阵^[25]

Table 1 Tanlangole uranium lithology identification confusion matrix^[25]

预测	真实					总计
	砾岩	粗砂岩	中砂岩	细砂岩	泥岩	
砾岩	26	0	0	0	0	26
粗砂岩	10	40	2	0	0	52
中砂岩	0	4	391	25	4	424
细砂岩	0	0	20	55	5	80
泥岩	0	0	1	6	10	17
灵敏度/%	72.22	90.91	94.44	63.95	52.63	74.83
特异性/%	100.00	99.63	81.72	79.81	100.00	92.23
精度/%	100.00	76.92	92.22	68.75	58.82	79.34
准确率/%			88.31			

此导致这两种岩性的预测准确率较低,这说明模型的数据样本均匀性对于预测结果具有重要影响。陈维政^[26]提出了一种基于网格搜索算法优化的人工神经网络对内蒙古某铀矿钻井岩性数据进行分类,使用网格搜索算法优选出最优隐藏层神经元个

数及学习率,实验结果表明,该模型能很好的对砂岩进行分类,总体分类准确率达到了 80%,但是对于不可渗透的粉砂岩等岩性识别效果较差,适用性有待进一步验证。

传统神经网络在建模和预测过程中过于依赖已知的地层先验信息,具有一定的局限性,因此有必要引入自组织神经网络(self-organizing neural networks, SOM)用以识别岩性^[27-28]。SOM 的结构使其具有自适应性,能够适应不同数据分布和模式,基于 SOM 的可视化特性,模型的输出可以直观地呈现在二维或三维映射中,有助于地质工作者的理解和解释。此外,通过这种无监督学习的神经网络算法,可以有效地解决新开铀矿区内样本数量较少,获取标准样本困难等问题^[29-30]。徐建国等^[31]提出 SOM 模型对获取的铀矿测井数据样本进行岩性自动分类,其运行结果与岩心编录岩性结果相吻合,不仅验证了该方法的可靠性,也为无监督学习在铀矿层岩性识别方面提供了重要参考。张平等^[32]将松辽盆地南部的实际测井资料输入 SOM 模型,岩性的识别准确率达到了 86.67%。蔡中超^[33]采用 SOM 岩性识别模型对测井数据展开了综合岩性识别,研究发现, SOM 方法在岩性的分类方面表现出较快的收敛速度和较高的识别准确率,然而,进一步研究显示,在加入 GR 数据的情况下,模型的性能反而会降低。对于薄层岩性的识别, SOM 的可靠性和适用性仍有待进一步验证。

循环神经网络(recurrent neural network, RNN)是一种递归神经网络,专用于处理序列信息。常规 RNN 容易丢失先前的信息,导致梯度爆炸和梯度消失等问题^[34-35],为了解决这一问题, Hochreiter 等^[36]首次提出了长短期记忆网络(long short-term memory, LSTM), LSTM 是一种改进的 RNN 模型,相比 RNN 具有更简单的循环单元,已成为常用的 RNN 改进网络。

实际上,测井数据的采集是一个按时间顺序进行的动态过程,最终产生随深度变化的离散数据点。这些数据点实际上代表了测井区域相邻岩体测井响应的加权和,因此,测井数据在空间上具有局部和全局的关联性。传统的机器学习方法通常假设测井数据在空间和时间上是互相独立的离散点,忽略了测井数据本身的特性。LSTM 神经网络结构的计算过程中采用的特殊门控结构能够考虑到局部和全局关联性,使其非常适用于处理测井数据。为了确定机器学习算法在铀矿层岩性分类任务中的精度上限, Kuchin 等^[37]开发出一种数字钻孔模型,使得生成一套完整的测井数据成为可能,不

仅可以避免主观的专家评估,同时通过这些生成的数据验证了 LSTM 在岩性分类方面的有效性。Wu 等^[38]将卷积神经网络和 LSTM 相结合,在经过优化算法调参后,该算法取得的效果明显优于其他传统机器学习算法。周渊凯等^[39]重点研究了深层网络在铀矿层岩性识别中的应用,利用深层网络提取复杂的岩性模式,建立 LSTM 和 8 层全连接神经网络模型,结果显示,深层网络可以在一定程度上缓解岩性样本不均衡对分类效果的影响,但这种提升是有限的,而且深层网络需要的计算资源较大,容易产生过拟合现象。

陈炫沂^[40]以大庆地区铀矿资源潜力区为研究对象,将 W2 井段测井数据作为样本,同时采用神经网络、RF、XGBoost 等机器学习算法与 LSTM 方法进行岩性识别效果对比,结果显示, LSTM、XGBoost 和 RF 算法在岩性预测方面表现良好,其结果与实际岩性基本一致。相反, BP 和 KNN 算法的预测结果存在较大偏差,主要集中在将粉砂岩和细砂岩误判为泥岩。BP 和 KNN 算法出现较多预测错误的薄岩层, XGBoost 和 RF 算法对一些薄层岩性未能做出预测, LSTM 对于薄层岩性的预测更为全面和准确。高精度的 LSTM 岩性识别模型为薄层段岩性识别提供了数据基础。

神经网络在铀矿层识别上的优势:①神经网络具有强大的非线性建模能力和抗干扰能力,在拟合数据的复杂关系方面,效果良好;②利用自组织神经网络模型,可以在缺少矿井内先验信息的情况下进行岩性识别;③ LSTM 模型在岩性识别性能上表现稳定。难点:①对于深层网络或大规模数据集,神经网络模型通常需要较长的训练时间和大量的计算资源;②神经网络对数据质量要求较高,特别是对于输入特征的准确性和完整性。

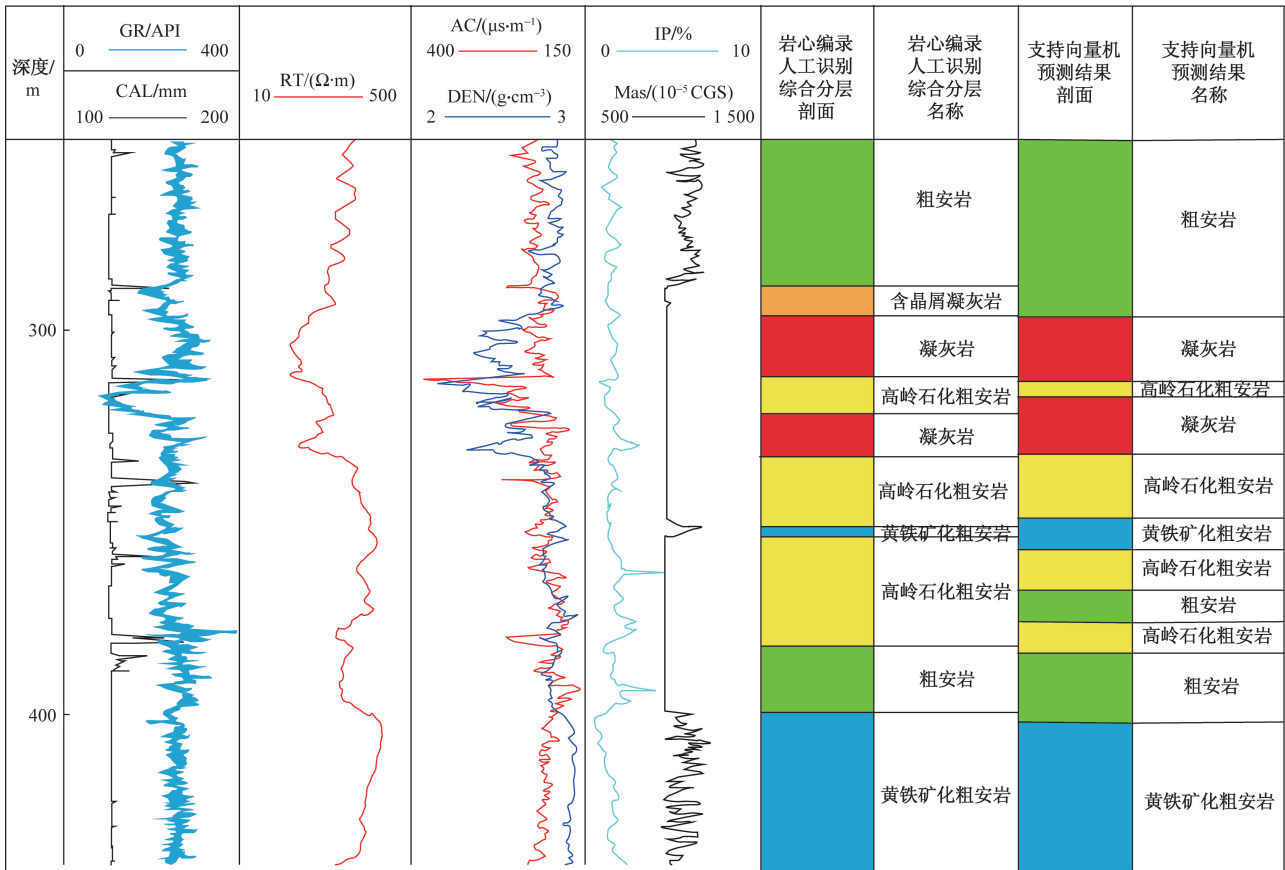
1.2 支持向量机

支持向量机是以最小化错误率理论界限为思想,以统计学习理论为基础,能较好地解决小样本学习问题的一种新的机器学习方法,简单来讲,就是从训练样本中学习构造函数来识别未知样本的类别。因其结构简单、学习速率快、适应能力强、推广能力强,而且内存占用很小等特点,在岩性识别领域深受欢迎^[41-43]。

岩性识别方法的选择主要依赖于相应数据的类型,支持向量机作为近年来发展较为迅速的一种智能方法,有力地促进了岩性识别技术的发展^[43]。数据解释的方法涵盖了多种学习系统,如支持向量分类、人工神经网络、 k -最近邻(k -nearest neighbor, KNN)等。Amirgaliev 等^[44]在上述几种方法的基础

上,提出邻接立方体法,通过对少数解释算法结果进行积分,使得铀矿测井岩性识别的准确性提高了 2%~3%。Kuchin 等^[45]对机器学习方法在哈萨克斯坦铀矿床的应用进行了研究,验证了非线性分类器 SVM 在铀矿床岩性分类方面的有效性,进一步研究发现,通过利用与邻近井的相关性,可以有效提高铀矿床岩性分类的质量。对铀矿勘探来说,岩性识别可以帮助确定潜在的矿化带和有利的岩石类型。高文利等^[46]借助对地球物理测井和钻孔岩心编录等数据的系统研究,完成了岩性的人工识别与支持向量机判别(图 2^[46]),建立了钻孔测井解释岩性剖面,总体上两种方法的判别结果相近,为庐枞盆地深部寻找铀提供了重要的依据。

支持向量机被认为是当前分辨能力较高的岩性识别方法之一^[47],通过对模型进行优化,可提高其识别准确率,如表 2^[48-52]所示。为解决岩性识别中岩层界面模糊定位问题,张翔等^[48]引入岩层样本对岩性类别的隶属度,提出了一种基于模糊支持向量机算法的岩性分类方法,结果表明,采用模糊支持向量机方法识别岩性的正确率为 94.8%,明显高于传统支持向量机方法的 89.5%,为沉积环境下的岩性识别提供了重要的参考价值。钟仪华等^[49]提出一种基于主成分分析的最小二乘支持向量机的岩性识别预测模型,其岩性识别的准确率达到 92.5%,这一方法显著简化了网络结构,提高岩性识别准确率的同时,运算速度也得到了显著提升。SVM 通过引入核函数,如多项式核或径向基函数(RBF)核,能够更灵活地捕捉岩性数据中的复杂模式,从而提高了模型的表达能力^[53]。为了消除人工选择参数时的随机性和不确定性,提升算法的自动化水平,研究人员开展了多项工作。郭瑞华等^[54]将粒子群优化算法(particle swarm optimization, PSO)与最小二乘支持向量机(least squares support vector machine, LSSVM)相结合,提出 PSO-LSSVM 模型,研究结果显示,该方法能有效描述测井数据与岩性类别之间的非线性映射关系,并具备较高的识别精度。陈钢花等^[50]采用粒子群算法优化支持向量机模型参数,将 30 多口井的数据输入支持向量机岩性识别模型,测试样本准确率为 83.56%,达到了预期效果。张昭杰等^[51]基于遗传算法构建了支持向量机岩性识别模型,该模型对样本预测总体符合率达 81.60%,对泥岩和中砂岩的识别准确率分别为 88.6%和 88.3%,明显优于传统测井岩性识别方法。Mou 等^[52]提出了一种基于变分不等式的支持向量机算法,通过 IPPA(interchangeable proximal point algorithm)算法求解 SVM 的最优解,将其应用



GR 为自然伽马测井曲线;DEN 为密度测井曲线;CAL 为井径测井曲线;RT 为电阻率测井曲线;AC 为声波时差测井曲线;IP 为极化率测井曲线;Mas 为磁化率测井曲线,其单位中的 CGS 是厘米-克-秒 (centimeter-gram-second) 单位制的缩写,Mas 测井曲线反映的是岩石的磁化率,在 CGS 单位制中,磁化率是一个无量纲的量,它表示物质被磁化的难易程度

图 2 支持向量机的判别结果^[46]

Fig. 2 Discriminant results of support vector machines^[46]

表 2 SVM 模型优化方法^[48-52]

Table 2 SVM model optimization methods^[48-52]

优化方法	方法描述	参考文献
引入隶属度	根据样本重要性赋予隶属度	张翔等 ^[48]
主成分分析法	提取测井数据中影响岩性识别因素	钟仪华等 ^[49]
粒子群算法	优化核函数参数 γ 和惩罚因子 C	陈钢花等 ^[50]
遗传算法	优选核函数参数 σ 和惩罚因子 C	张昭杰等 ^[51]
变分不等式算法	优化问题转变为求解不等式	Mou 等 ^[52]

于岩性分类。结果显示,基于 IPPA 的 SVM 模型的岩性预测准确率达 96%,远高于 LSSVM 模型的 75%,同时也提升了运算速度。虽然结合支持向量机与其他算法能有效克服其难以同时处理多个岩性的问题,但 SVM 模型对于特征的选择和提取非常敏感。如果选择的特征不能很好地捕捉薄层岩性的关键特征,模型的性能可能会下降^[55]。

支持向量机算法在铀矿层识别上的优势:①测井数据通常包含多个测井曲线,每个曲线代表一个特征,构成高维数据,而 SVM 在高维数据空间中表现出色;②SVM 在处理小样本数据集时通常

表现良好,它倾向于具有良好的泛化性能,可以处理样本较少的岩性识别问题。缺陷:①支持向量机方法在交叉验证和准确性方面表现不佳,所研究的地质数据缺乏明显的规律,该方法受到许多外部和自然因素的影响;②利用 SVM 方法进行机器学习,需要拥有大量的数据,并且针对特定数据集,每次训练数据的分布都应该是平衡的;③SVM 算法直接作用于特征向量和支撑向量,利用适当的核函数(如高斯核、多项式核或 sigmoid 核)在二维或三维平面上执行计算,需要兼容的硬件和较长的处理时间。

1.3 集成学习算法

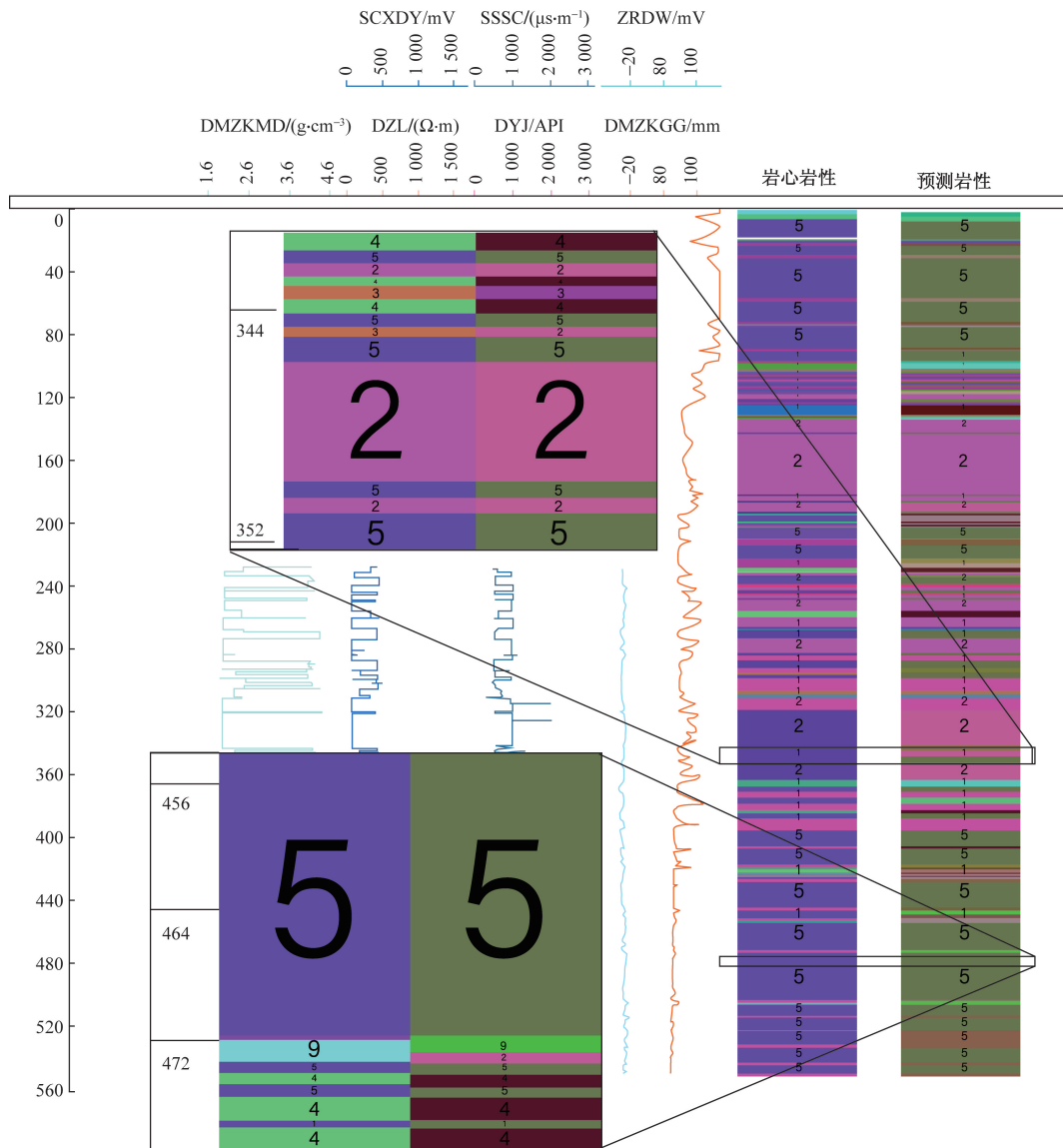
集成(integrated)学习算法将多个分类器组合在一起,综合多个分类器的优点进行决策,它采用一系列弱学习器进行学习,并按照一定的规则对各种学习结果进行整合,从而获得比单一的弱学习器学习方法更好的学习效果^[56-57]。目前常用于铀矿层识别的集成学习算法主要有:基于 Bagging 思想的算法和基于 Boosting 思想的算法^[58-59]。

随机森林是一种基于 Bagging 思想的集成学习算法,使用决策树作为基学习器,通过软投票或硬投票方式整合多个决策树的预测结果^[60-61]。机器学习主要用于数据的回归和分类,而随机森林可以同时胜任这两种任务,已逐渐成为当前数据挖掘、生物信息学等领域的研究热点^[62]。为了寻找在岩性识别中实用性较高的模型,康乾坤等^[23]研究分析了随机森林算法、神经网络和支持向量机算法在岩性分类中的应用,通过对比验证,随机森林模型取得了最高的分类效率,而且在通过变量的优选后,随机森林模型的预测准确率还可以得到进一步提高。王志宏等^[63]将因子分析浓缩数据的优势引入到储层岩性判别指标的分析中,从而降低因素间的关联度,并建立了随机森林分类模型,结果显示,所得到的判别模型泛化误差满足要求,这种方法为结合测井数据对储层岩性的判别提供了一条新的研究途径。Chen 等^[64]的研究发现,主成分分析能够识别与铀相关的元素组合,从而有助于揭示深埋铀矿床的储层特征。采用元素组合的线性判别分析(LDA)和 RF 模型,可以有效地识别受含铀热液活动影响的砂岩下伏区域,为在全球元古代盆地中勘探深埋铀矿床提供了方法。马东来等^[65]采用随机森林模型对直罗组铀矿目的层中的 5 种岩性(泥岩、细砂岩、中砂岩、粗砂岩和砾岩)进行自动分类,总体识别准确率达到 82.85%,然而,对于粗砂岩的识别准确率仅为 48%,还需要进一步提高。Kong 等^[66]利用随机森林模型对二连盆地马尼特拗陷进行矿产远景定位,结果显示,模型在研究区铀矿化预测中准确率达到 92%,为进一步的铀矿勘探提供了有价值的地质见解。

基于 Boosting 的算法包括梯度提升决策树(gradient boosting decision tree, GBDT)、自适应增强(adaptive boosting, AdaBoost)等。GBDT 的基本原理是通过迭代的方式构建一系列决策树模型,每棵树都试图纠正前一棵树的预测误差。因此,选择合适的决策树模型在一定程度上影响着 GBDT 模型最终的泛化能力^[67-68]。韩启迪等^[69]将 GBDT、KNN、SVM 和决策树 4 种机器学习算法分别用于岩性识别领域,最终发现 GBDT 模型展现出了集成学习模型的优势,其识别准确率明显优于其他 3 种单一分类模型。为了验证 GBDT 在识别致密砂岩储层岩性方面相较于传统算法的优越性,Gu 等^[70]选用 GBDT 算法进行储层岩性预测,并比较了 GBDT 与其他机器学习算法在岩性识别上的性能差异,通过实验验证了 GBDT 在岩性识别中的优越性。马陇飞等^[71]比

较了 GBDT、随机森林、支持向量机和神经网络 4 种机器学习算法模型,最终 4 种机器学习模型都取得了较高的识别准确率,其中 GBDT 在所有评价指标中取得了最高分。鉴于交会图法对于研究区内的岩性识别效果不明显,段忠义等^[72]建立 GBDT 模型,通过网格搜索法对 GBDT 模型的参数进行优化调整,最终岩性识别准确率达到 98.52%,为砂岩型铀矿地层测井识别与评价提供了重要的参考价值。GBDT 模型采用了贪婪算法,虽然简单高效,但是难以全局优化,而且容易产生过拟合。

XGBoost 算法的应用能够提高模型的泛化能力并减少过拟合的风险,是一种基于 GBDT 框架的改进算法或者说工程实现,由于该算法采用多线程和分布式计算的方法,因此运算时间也得到了大幅缩短。Merembayev 等^[73]以哈萨克斯坦铀矿为研究对象,基于神经网络、XGBoost 等 5 种机器学习算法对矿床内部地层标记和岩性识别,结果显示,所有算法对测试地层数据精度较高,但在岩性分类方面的精度较低,这一方法有助于自动定义地层水平,从而改善铀矿的开采流程。Sun 等^[74]研究发现,经过贝叶斯算法优化后的 XGBoost 模型,在各项评价指标上均优于 GBDT 模型,Zhang 等^[75]构建 3 种机器学习模型,用于地层岩性识别,结果显示,集成学习中的 XGBoost 方法在地层识别中表现出明显优势,其岩性识别准确率最高。为了解决传统机器学习模型泛化能力低的问题,Zhang 等^[76]利用集成学习中的 AdaBoost 算法对研究区域的测井资料进行模型试验,结果显示,模型在测试集上的分类准确率约为 98.42%,为了更直观的了解模型的效果,绘制原始岩性与预测岩性结果对比图(图 3^[76])。可以看出,该模型对较厚地层岩性的识别准确率较高,但对薄层岩性的识别能力较低,表明该模型需要进一步优化。为了提高模型的参数寻优效率,邹学钢^[77]使用麻雀搜索算法对 XGBoost 算法进行优化,进而构建 XGBoost 岩性识别模型,3 组试验结果表明,优化后的 XGBoost 模型岩性识别准确率均高于 95%,尽管相较于模型优化之前提高了约 20%,但该模型对于少数类岩性召回率偏低仍是一个亟待解决的问题。Mukhamediev 等^[78]选择经典机器学习算法和集成算法建立岩性识别模型,利用浮动数据窗口对铀矿测井数据进行转换,以达到增强岩性分类效果的目的,结果表明,增加浮动数据窗口的大小可以将岩性识别准确率提高 6%~12%;XGBoost 模型在岩性分类中取得了最好的效果(F1_score 为 0.705, F1_score 用于评估分类模型的性能,是精确率和召回率的调和平均数)。上述研究证明



SCXDY 为三侧向电压测井曲线, mV; SSSC 为双采集时差测井曲线, $\mu\text{s}/\text{m}$; DMZKMD 为密度测井曲线, g/cm^3 ; ZRDW 为自然电位测井曲线, mV; DZL 为钻井液电阻率测井曲线, $\Omega \cdot \text{m}$; DYJ 为短源距放射性测井曲线, API; DMZKGG 为井径测井曲线, mm; 1 为泥岩; 2 为粉砂岩; 3 为细砂岩; 4 为中砂岩; 5 为粗砂岩

图3 岩性识别对比图^[76]

Fig. 3 Comparison of lithology identification^[76]

了将集成学习算法应用于铀矿测井数据分类中的有效性和优越性。

集成学习方法在铀矿层识别的优势: ①通过将多个分类器组合起来, 集成学习可以综合利用每个分类器的优势, 弥补单个分类器的不足, 从而提高整体分类准确性; ②通过集成学习可以降低过拟合的风险, 因为不同的分类器可能会在不同的数据子集上进行训练, 从而减少对某些数据的过度拟合。难点: ①集成学习通常需要训练多个分类器, 训练时间和计算复杂性可能显著增加; ②无法有效的解释每个弱分类器对于结果的影响大小。

2 机器学习与铀矿层岩性识别分析

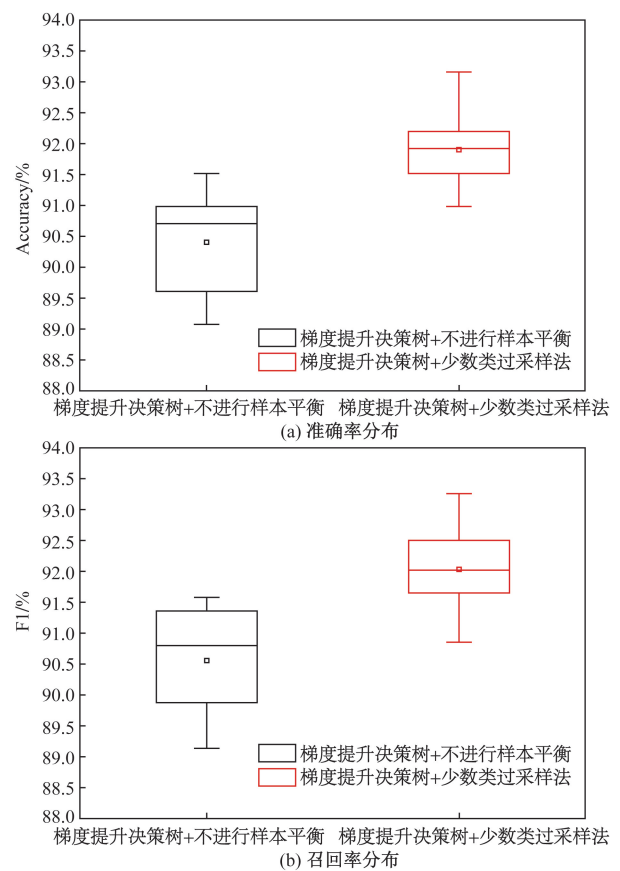
综上所述, 集成学习通过将多个基本模型组合起来, 能够有效地降低模型的偏差和方差, 并从不同的角度捕捉和利用岩性数据的多样性, 在铀矿层岩性识别方面展现出巨大的潜力。通过进一步对不同机器学习方法在铀矿层识别的效果进行分析, 如表3^[32,40,51,65,72]所示, 结果表明, 机器学习算法在处理数据样本不平衡或岩层较薄的情况下存在一定的分类困难问题, 因为机器学习模型需要使用足够数量的数据样本去训练, 才能有效地学习和泛化。针对以上问题, 研究出一种对非均衡数据进行

表3 铀矿层识别的不同机器学习方法对比^[32,40,51,65,72]Table 3 Comparison of different machine learning methods for uranium seam identification^[32,40,51,65,72]

机器学习方法	输出类别	准确率/%	主要优缺点	参考文献
神经网络(ANN)	矿化层、异常层非矿化层	86.55	抗干扰能力强,能快速获取未知孔的异常信息	张平等 ^[32]
长短期记忆神经网络(LSTM)	泥岩、粉砂岩、细砂岩、粗砂岩、中砂岩	85.00	模型的循环单元简单,性能稳定,少量样本岩性识别效果较差	陈炫沂 ^[40]
支持向量机(SVM)	泥岩、泥质粉砂岩、中细砂岩、砂砾岩	81.60	适用于小样本,中砂岩和泥岩识别易混淆	张昭杰等 ^[51]
随机森林(RF)	泥岩、细砂岩、中砂岩、粗砂岩、砾岩	82.85	有较高鲁棒性和泛化能力,薄层岩性识别效果较差	马东来等 ^[65]
梯度提升决策树(GBDT)	黏土、泥岩、粉砂岩、细砂岩、中砂岩、粗砂岩、砂砾岩	98.52	简单高效,可以提供特征重要性评估,但薄层岩性识别准确率较低	段忠义等 ^[72]

高效、高精度分类的算法尤为重要。

研究表明,引入人工少数类过采样法(synthetic minority over-sampling technique, SOMTE)可以在一定程度上改善样本不均衡问题。SOMTE 算法作为典型的过采样算法,通过生成合成样本来增加少数类样本比例,从而实现训练集数据的平衡,提高少数类别的识别准确率^[79-80]。段忠义等^[81]将 XGBoost、SOMTE 随机森林、GBDT、KNN 模型应用于松辽盆地砂岩型铀矿层的识别,对比这些模型的识别结果,发现 SOMTE 随机森林模型的预测精度达到 95.02%,与 XGBoost 模型的预测精度相近,验证了 SOMTE 算法的可行性。为解决地层中不同类型岩性比例失衡的问题,陈炫沂^[40]通过 SMOTE 算法对测井数据进行平衡处理,并以此为基础建立了 SMOTE-LSTM 模型。结果显示,与 SOMTE 算法处理数据前的 LSTM 模型相比,样本数量较少的粉、细砂岩和中、粗砂岩的预测准确率提高了约 20%,而且总体岩性识别准确率也略有上升,进一步验证了该算法的可行性。为处理不平衡的数据,Ibrahim 等^[82]提出一种基于合成少数过采样技术和极限梯度增强(SMOTE-XGB)的客观混合方法,用于 Tarkwaian 古砂矿地层的岩性分类,结果显示,当 SMOTE 算法应用于不平衡数据集时,少数类(尤其是矿化岩性)的分类性能显著提高。针对传统方法的局限性和测井数据的数据不平衡问题,Zhou 等^[83]建立了基于集成算法的 GBDT 学习模型,通过 SMOTE 算法平衡岩性数据的类别数量,如图 4^[83]所示,GBDT 模型的岩性识别准确率(Accuracy)和召回率(F1)都明显提升,进一步对比其他平衡算法发现,SMOTE 算法对于模型性能的提升效果最为显著,表明了该方法在提高岩性识别性能方面的有效性。为了解决研究区内测井资料不均衡、岩性小比例难以预测等问题,Shi^[84]基于 SMOTE 建立 logic-CNN 岩性识别模型,结果表明,使用 SMOTE 后,泥岩、泥质粉砂岩、



GBDT + NoBalancing 为梯度提升决策树 + 不进行样本平衡;
GBDT + SMOTE 为梯度提升决策树 + 少数类过采样法
图4 GBDT + NoBalancing 方法与 GBDT + SMOTE 方法的比较^[83]

Fig. 4 Comparison of GBDT + NoBalancing method and GBDT + SMOTE method^[83]

粉砂岩、火山岩和凝灰岩的识别率分别提高了 1.30%、3.03%、1.90%、1.03% 和 12.09%。

综上所述,证实 SMOTE 算法在处理不平衡数据时的有效性和准确性,然而,SMOTE 算法对于不同机器学习模型的性能影响程度存在差异。为了更好地将该方法应用于不同的地质勘探领域,还有待深入研究。

3 展望

岩性识别对储层预测和储层评价都有着极其重要的参考价值,利用机器学习算法与测井资料相结合进行岩性识别,对于矿藏的智能化开发也具有重大意义。总结神经网络、支持向量机、集成学习等几种机器学习算法在铀矿层识别方面的研究现状,分析了将机器学习算法应用于铀矿层的岩性识别、多样本和多数据信息准确分类中的可行性,也指出了当前机器学习在岩性识别中存在的问题。没有任何一种机器学习模型是完美的,如何为现有的问题寻找最优解是一个永恒不变的话题。

虽然机器学习为铀矿层识别提供了一个更加方便快捷的方法,使得岩性识别的效率大大提高,但在实践过程中,还是存在一些不足,对于未来机器学习在铀矿层识别方面还需要重点关注以下几个方面:①探索能够提高岩性分类质量的深度学习模型,比如卷积神经网络,它能够考虑到测井曲线的形状特征,因此可以更好地模仿研究人员进行岩性分类的过程;②结合多种数据源进行铀矿层岩性智能识别将成为未来的发展趋势。除了传统的测井数据外,铀矿层岩性智能识别还可以融合地质图像、地球物理数据、地球化学数据等多模态信息;③机器学习具有一定的“黑盒”性质,未来可以着重研究模型在铀矿层岩性识别中的可解释性;例如,可以结合局部解释方法和全局解释方法,同时考虑模型的整体结构和局部决策过程,这样不仅可以增强研究人员对模型的信任和理解,也可以促进模型的改进和优化;④以随机森林等算法为代表的监督学习方法在岩性识别领域中高度依赖大量准确的岩性标签数据,然而,岩性标签一旦存在不准确的情况,极易导致模型预测产生偏差;同时,高质量岩性标签的获取难度较大,并且可能引发过拟合或欠拟合等问题,这在很大程度上限制了监督学习方法在岩性识别中的应用;未来,可深入探索无监督学习方法,以降低对岩性标签的依赖程度,进而提高岩性识别的准确性与可靠性;⑤探索多个模型相融合在岩性识别中具有重要意义和广阔应用前景。单一模型难以有效的处理岩性识别中涉及多种复杂因素,可能会导致模型泛化能力有限。多种模型相融合,可以从多个角度综合考虑问题,充分发挥各自的优势。

随着人工智能和世界范围内核能的迅猛发展,全球对于铀矿的需求将不断增加,铀矿层岩性智能识别将走上快速发展的高速通道。

4 结论

(1)归纳不同机器学习算法在铀矿层岩性识别中的应用研究现状,总结了不同机器学习算法在岩性识别的优势与不足,并提出针对性的建议。

(2)将机器学习算法应用于铀矿层的岩性识别中,运用机器学习方法较强的非线性拟合能力和泛化能力,可以有效地弥补传统线性数学方法的不足,有监督学习的特点可以解决诸多需要人工经验判别的问题,大幅度提升铀矿层识别的效率与精度。

(3)岩性智能识别方法不仅可以挖掘数据背后蕴藏的巨大潜力,也使得数据处理更加准确和便捷,已经逐渐成为铀矿层识别的重要发展方向。

参 考 文 献

- [1] 李增华. 一种基于机器学习的砂岩型铀矿层类型识别方法及装置. CN202211303813.9[P]. 2023-01-10.
Li Zenghua. A method and device for identifying lithological types of sandstone-type uranium deposits based on machine learning: CN202211303813.9[P]. 2023-01-10.
- [2] 操应长,姜在兴,夏斌,等. 利用测井资料识别层序地层界面的几种方法[J]. 石油大学学报(自然科学版), 2003(2): 23-26.
Cao Yingchang, Jiang Zaixing, Xia Bin, et al. Several methods for identifying sequence stratigraphic interfaces using well logging data [J]. Journal of Petroleum University (Natural Science Edition), 2003(2): 23-26.
- [3] 周正龙,王贵文,冉冶,等. 致密油储集层岩性岩相测井识别方法:以鄂尔多斯盆地合水地区三叠系延长组7段为例[J]. 石油勘探与开发, 2016, 43(1): 61-68.
Zhou Zhenglong, Wang Guiwen, Ran Ye, et al. Logging identification method for lithology and lithofacies of tight oil reservoirs: a case study of the Triassic Yanchang Formation Member 7 in Heshui area, Ordos Basin [J]. Petroleum Exploration and Development, 2016, 43(1): 61-68.
- [4] 袁彪. 基于机器学习的岩性识别模型研究[D]. 北京:中国地质大学, 2021.
Yuan Biao. Research on rock identification models based on machine learning [D]. Beijing: China University of Geosciences, 2021.
- [5] Xie Y X, Zhu C Y, Wen Z, et al. Evaluation of machine learning methods for formation lithology identification: a comparison of tuning processes and model performances [J]. Journal of Petroleum Science and Engineering, 2018, 160: 182-193.
- [6] Samuel J. Rogers, 钱志良. 利用神经网络由测井数据确定岩性[J]. 国外油气勘探, 1993(1): 76-83.
Samuel J. Rogers, Qian Zhiliang. Determination of lithology from well logging data using neural networks [J]. Foreign Oil and Gas Exploration, 1993(1): 76-83.
- [7] 刘明军. 测井数据岩性识别软件系统的开发与应用[J]. 地球物理学进展, 2013, 28(2): 908-913.
Liu Mingjun. Development and application of well logging data lithology identification software systems [J]. Progress in Geophysics,

- 2013, 28(2): 908-913.
- [8] Cortes C, Vapnik V. Support-vector networks[J]. *Machine Learning*, 1995, 20: 273-297.
- [9] Smirnoff A, Bolsvert E, Paradis S J. Support vector machine for 3D modelling from sparse geological information of various origins[J]. *Computers & Geosciences*, 2008, 34(2): 127-143.
- [10] 牟丹, 王祝文, 黄玉龙, 等. 基于 SVM 测井数据的火山岩岩性识别——以辽河盆地东部坳陷为例[J]. *地球物理学报*, 2015, 58(5): 1785-1793.
Mou Dan, Wang Zhuwen, Huang Yulong, et al. Lithology identification of volcanic rocks based on SVM well logging data: a case study of the eastern sag of the Liaohe Basin[J]. *Chinese Journal of Geophysics*, 2015, 58(5): 1785-1793.
- [11] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45: 5-32.
- [12] 周雪晴, 张占松, 张超谟, 等. 基于粗糙集——随机森林算法的复杂岩性识别[J]. *大庆石油地质与开发*, 2017, 36(6): 127-133.
Zhou Xueqing, Zhang Zhansong, Zhang Chaomo, et al. Identification of complex lithologies based on rough set-random forest algorithm[J]. *Daqing Petroleum Geology and Development*, 2017, 36(6): 127-133.
- [13] Ao Y, Li H, Zhu L, et al. Identifying channel sand-body from multiple seismic attributes with an improved random forest algorithm[J]. *Journal of Petroleum Science and Engineering*, 2019, 173: 781-792.
- [14] 朱枫帆. 基于机器学习的测井数据岩性识别分析研究[D]. 荆州: 长江大学, 2021.
Zhu Fengfan. Research on lithology identification analysis of well logging data based on machine learning[D]. Jingzhou: Yangtze University, 2021.
- [15] Wang H, Lei Z, Zhang X, et al. Machine learning basics[J]. *Deep Learning*, 2016, 1: 98-164.
- [16] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. *The Bulletin of Mathematical Biophysics*, 1943, 5: 115-133.
- [17] Wu Q, Li Z, Wang Y, et al. Combination of seismic attributes using clustering and neural networks to identify environments with sandstone-type uranium mineralization [J]. *Acta Geophysica*, 2023, 71(6): 2715-2731.
- [18] 许建华. 反向传播神经网络模型及其在测井资料岩性自动识别中的应用[J]. *石油物探*, 1993, 32(3): 53-59.
Xu Jianhua. Application of backpropagation neural network model in automatic lithology recognition of well logging data[J]. *Petroleum Exploration and Development*, 1993, 32(3): 53-59.
- [19] 祖秀兰, 周蓉生, 马英杰. 人工神经网络在铀矿测井解释中的应用[J]. *铀矿地质*, 2001, 17(4): 239-244.
Zu Xiulan, Zhou Rongsheng, Ma Yingjie. Application of artificial neural networks in uranium mine logging interpretation[J]. *Uranium Geology*, 2001, 17(4): 239-244.
- [20] 李继安. 人工智能神经网络在岩性识别、孔隙度和渗透率预测中的应用——以十红滩铀矿床为例[J]. *西北地质*, 2010, 43(2): 32-37.
Li Ji'an. Application of artificial intelligence neural networks in lithology identification, porosity and permeability prediction: a case study of the Shihongtan uranium deposit[J]. *Northwestern Geology*, 2010, 43(2): 32-37.
- [21] Muhamediyev R I, Amirgaliyev Y, Iskakov S K, et al. Integration of results from recognition algorithms applied to the uranium deposits[J]. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2014, 18(3): 347-352.
- [22] Muhamediyev R, Yakunin K, Iskakov S, et al. Comparative analysis of classification algorithms[C]//2015 9th International Conference on Application of Information and Communication Technologies (AICT). Beijing: IEEE, 2015: 96-101.
- [23] 康乾坤, 路来君, 尚殷民. 基于测井数据的砂岩型铀矿异常识别 BP 神经网络方法应用[J]. *科学技术与工程*, 2020, 20(9): 3476-3484.
Kang Qiankun, Lu Laijun, Shang Yinmin. Application of BP neural network method in sandstone-type uranium anomaly identification based on well logging data[J]. *Science Technology and Engineering*, 2020, 20(9): 3476-3484.
- [24] 易敏, 叶发旺, 张川, 等. 基于深度学习的蚀变矿物识别——以新疆白杨河铀矿床为例[J]. *铀矿地质*, 2021, 37(4): 673-682.
Yi Min, Ye Fawang, Zhang Chuan, et al. Identification of altered minerals based on deep learning: a case study of the Baiyanghe uranium deposit in Xinjiang [J]. *Uranium Geology*, 2021, 37(4): 673-682.
- [25] Sun Y Q, Chen J, Yan P, et al. Lithology identification of uranium-bearing sand bodies using logging data based on a BP neural network[J]. *Minerals*, 2022, 12(5): 248-256.
- [26] 陈维政. 神经网络在砂岩型铀矿测井中的岩性识别应用[J]. *黑龙江科学*, 2024, 15(12): 8-12.
Chen Weizheng. Application of neural networks in lithology identification for sandstone-type uranium ore logging[J]. *Heilongjiang Science*, 2024, 15(12): 8-12.
- [27] 陈晖, 胡泽根, 李云鹏, 等. 自组织神经网络技术在新油田快速评价中的应用[J]. *重庆科技学院学报(自然科学版)*, 2017, 19(2): 15-17.
Chen Hui, Hu Zegen, Li Yunpeng, et al. Application of self-organizing neural network technology in rapid evaluation of new oil-fields[J]. *Journal of Chongqing University of Science and Technology (Natural Science Edition)*, 2017, 19(2): 15-17.
- [28] 侯俊胜, 尉中良. 自组织神经网络在测井资料解释中的应用[J]. *测井技术*, 1996, 20(3): 197-200.
Hou Junsheng, Wei Zhongliang. Application of self-organizing neural network in interpretation of well logging data[J]. *Well Logging Technology*, 1996, 20(3): 197-200.
- [29] 王宗俊, 董洪超, 范廷恩, 等. 基于无监督学习的测井岩相分析技术及其应用[J]. *石油物探*, 2021, 60(3): 403-413.
Wang Zongjun, Dong Hongchao, Fan Ting'en, et al. Analysis technology and application of well logging lithofacies based on unsupervised learning[J]. *Petroleum Exploration*, 2021, 60(3): 403-413.
- [30] Cunningham P, Cord M, Delany S J. Machine learning techniques for multimedia: case studies on organization and retrieval[M]. Berlin: Springer, 2008: 21-49.
- [31] 徐建国, 吴显礼, 王卫国. 自组织竞争人工神经网络在砂岩型铀矿测井数据解释中的应用[J]. *世界核地质科学*, 2008, 25(2): 114-118.
Xu Jianguo, Wu Xianli, Wang Weiguo. Application of self-organizing

- zing competitive artificial neural network in interpretation of sandstone-type uranium mine logging data[J]. *World Nuclear Geoscience*, 2008, 25(2): 114-118.
- [32] 张平, 潘保芝, 张莹, 等. 自组织神经网络在火成岩岩性识别中的应用[J]. *石油物探*, 2009, 48(1): 53-56.
Zhang Ping, Pan Baozhi, Zhang Ying, et al. Application of self-organizing neural network in lithology identification of igneous rocks[J]. *Petroleum Exploration*, 2009, 48(1): 53-56.
- [33] 蔡中超. 自组织竞争神经网络在砂岩型铀矿测井数据解释中的应用研究[D]. 南昌: 东华理工大学, 2015.
Cai Zhongchao. Research on the application of self-organizing competitive neural network in interpretation of sandstone-type uranium mine logging data[D]. Nanchang: East China University of Technology, 2015.
- [34] 刘礼文, 俞弦. 循环神经网络(RNN)及应用研究[J]. *科技视界*, 2019(32): 54-55.
Liu Liwen, Yu Xian. Research on recurrent neural network (RNN) and its applications[J]. *Science and Technology Vision*, 2019(32): 54-55.
- [35] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Trans Neural Network*, 2002, 5(2): 157-166.
- [36] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [37] Kuchin Y I, Mukhamediev R I, Yakunin K O. One method of generating synthetic data to assess the upper limit of machine learning algorithms performance[J]. *Cogent Engineering*, 2020, 7(1): 1718821.
- [38] Wu L, Dong Z, Li W, et al. Well-logging prediction based on hybrid neural network model[J]. *Energies*, 2021, 14(24): 8583.
- [39] 周渊凯, 刘祐. 基于深度学习方法的测井岩性识别研究[J]. *铀矿地质*, 2024, 40(2): 336-345.
Zhou Yuankai, Liu Hu. Study on logging lithology identification based on deep learning methods[J]. *Uranium Geology*, 2024, 40(2): 336-345.
- [40] 陈炫沂. 砂岩型铀矿潜力区岩性特征智能识别方法研究[D]. 长春: 吉林大学, 2023.
Chen Xuanyi. Study on intelligent recognition method of lithological characteristics in potential areas of sandstone-type uranium deposits[D]. Changchun: Jilin University, 2023.
- [41] Deng C, Pan H, Fang S, et al. Support vector machine as an alternative method for lithology classification of crystalline rocks[J]. *Journal of Geophysics and Engineering*, 2017, 14(2): 341-349.
- [42] 柳成志, 滕立惠. 利用支持向量机识别松辽盆地火山岩岩性[J]. *地质与资源*, 2014, 23(3): 288-291.
Liu Chengzhi, Teng Lihui. Utilizing support vector machines to identify lithology of volcanic rocks in the Songliao Basin[J]. *Geology and Resources*, 2014, 23(3): 288-291.
- [43] 巩维嘉. 基于集成学习的岩性识别研究与应用[D]. 大庆: 东北石油大学, 2023.
Gong Weijia. Research and application of lithology identification based on ensemble learning[D]. Daqing: Northeast Petroleum University, 2023.
- [44] Amirgaliev E, Isabaev Z, Iskakov S, et al. Recognition of rocks at uranium deposits by using a few methods of machine learning [C]//*Soft Computing in Machine Learning*. New York: Springer International Publishing, 2014: 33-40.
- [45] Kuchin Y, Grundspenķis J. Machine learning methods for identifying composition of uranium deposits in Kazakhstan[J]. *Applied Computer Systems*, 2017, 22(1): 21-27.
- [46] 高文利, 孔广胜, 潘和平, 等. 庐枞盆地科学钻探地球物理测井及深部铀异常的发现[J]. *地球物理学报*, 2015, 58(12): 4522-4533.
Gao Wenli, Kong Guangsheng, Pan Heping, et al. Discovery of deep uranium anomalies through scientific drilling geophysical logging in the Lufeng Basin[J]. *Chinese Journal of Geophysics*, 2015, 58(12): 4522-4533.
- [47] 吴施楷, 曹俊兴. 基于连续限制玻尔兹曼机的支持向量机岩性识别方法[J]. *地球物理学进展*, 2016, 31(2): 821-828.
Wu Shikai, Cao Junxing. Lithology identification method based on continuous restricted Boltzmann machine support vector machine [J]. *Progress in Geophysics*, 2016, 31(2): 821-828.
- [48] 张翔, 肖小玲, 严良俊, 等. 基于模糊支持向量机方法的岩性识别[J]. *石油天然气学报*, 2009, 31(6): 115-118.
Zhang Xiang, Xiao Xiaoling, Yan Liangjun, et al. Lithology identification based on fuzzy support vector machine method[J]. *Journal of Oil and Gas Technology*, 2009, 31(6): 115-118.
- [49] 钟仪华, 李榕. 基于主成分分析的最小二乘支持向量机岩性识别方法[J]. *测井技术*, 2009, 33(5): 425-429.
Zhong Yihua, Li Rong. Lithology identification method based on principal component analysis and least squares support vector machine[J]. *Well Logging Technology*, 2009, 33(5): 425-429.
- [50] 陈钢花, 王军. 粒子群算法在砂砾岩体岩性识别中的应用[J]. *测井技术*, 2015, 39(2): 171-174.
Chen Ganghua, Wang Jun. Application of particle swarm algorithm in lithology identification of conglomerate bodies[J]. *Well Logging Technology*, 2015, 39(2): 171-174.
- [51] 张昭杰, 方石. 基于遗传算法优化的支持向量机在岩性识别中的应用[J]. *世界地质*, 2019, 38(2): 486-491.
Zhang Zhaojie, Fang Shi. Application of support vector machine in lithology identification based on genetic algorithm optimization [J]. *World Geology*, 2019, 38(2): 486-491.
- [52] Mou D, Wang Z, Tan X. A variational inequality approach with SVM optimization algorithm for identifying mineral lithology [J]. *Journal of Applied Geophysics*, 2022, 204: 104747.
- [53] Liu Y, Zhu R, Zhai S, et al. Lithofacies identification of shale formation based on mineral content regression using LightGBM algorithm: a case study in the Luzhou block, South Sichuan Basin, China[J]. *Energy Science & Engineering*, 2023, 11(11): 4256-4272.
- [54] 郭瑞华, 曹庆年. 基于粒子群优化的最小二乘支持向量机在岩性识别中的应用[C]//第十一届中国青年信息与管理学者大会论文集. 西安: 西安石油大学, 2009: 62-66.
Guo Ruihua, Cao Qingnian. Application of particle swarm optimization-based least squares support vector machine in lithology identification [C]//*Proceedings of the 11th China Youth Scholars' Conference on Information and Management*. Xi'an: Xi'an Shiyou University, 2009: 62-66.
- [55] 全容超. 机器学习在岩性智能识别中的应用研究进展[J]. *化工矿物与加工*, 2022, 51(8): 43-47.
Tong Rongchao. Research progress on the application of machine learning in intelligent lithology recognition [J]. *Chemical Engi-*

- neering & Mining, 2022, 51(8): 43-47.
- [56] 王日升. 基于 Spark 的一种改进的随机森林算法研究[D]. 太原: 太原理工大学, 2018.
Wang Risheng. Research on an improved random forest algorithm based on Spark[D]. Taiyuan: Taiyuan University of Technology, 2018.
- [57] Gong K, Ye Z, Chen D, et al. Investigation on automatic recognition of stratigraphic lithology based on well logging data using ensemble learning algorithm[C]//SPE Asia Pacific Oil and Gas Conference and Exhibition. Dallas: Society of Petroleum Engineers, 2018; D021S016R003.
- [58] 连克强. 基于 Boosting 的集成树算法研究与分析[D]. 北京: 中国地质大学, 2018.
Lian Keqiang. Research and analysis of ensemble tree algorithm based on Boosting[D]. Beijing: China University of Geosciences, 2018.
- [59] 任涛, 林梦楠, 陈宏峰, 等. 基于 Bagging 集成学习算法的地震事件性质识别分类[J]. 地球物理学报, 2019, 62(1): 383-392.
Ren Tao, Lin Mengnan, Chen Hongfeng, et al. Classification of seismic event properties based on Bagging ensemble learning algorithm[J]. Chinese Journal of Geophysics, 2019, 62(1): 383-392.
- [60] 李菊花, 秦顺利, 王洁, 等. 随机森林算法在吉木萨尔页岩油藏中的应用[J]. 长江大学学报(自然科学版), 2023, 20(2): 69-76.
Li Juhua, Qin Shunli, Wang Jie, et al. Application of random forest algorithm in Jimsar shale reservoirs[J]. Journal of Yangtze University (Natural Science Edition), 2023, 20(2): 69-76.
- [61] Schonlau M, Zou R Y. The random forest algorithm for statistical learning[J]. The Stata Journal, 2020, 20(1): 3-29.
- [62] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. 统计与信息论坛, 2011(3): 32-38.
Fang Kuangnan, Wu Jianbin, Zhu Jianping, et al. A review of random forest methods[J]. Statistical and Information Forum, 2011(3): 32-38.
- [63] 王志宏, 韩璐, 戚磊. 随机森林分类方法在储层岩性识别中的应用[J]. 辽宁工程技术大学学报(自然科学版), 2015, 34(9): 1083-1088.
Wang Zhihong, Han Lu, Qi Lei. Application of random forest classification method in reservoir lithology identification[J]. Journal of Liaoning University of Technology (Natural Science Edition), 2015, 34(9): 1083-1088.
- [64] Chen S, Hattori K, Grunsky E C. Identification of sandstones above blind uranium deposits using multivariate statistical assessment of compositional data, Athabasca Basin, Canada[J]. Journal of Geochemical Exploration, 2018, 188: 229-239.
- [65] 马东来, 蔡煜琦, 孙远强. 基于随机森林算法的塔然高勒地区测井数据岩性识别[J]. 世界核地质科学, 2023, 40(1): 43-50, 67.
Ma Donglai, Cai Yuqi, Sun Yuanqiang. Lithology identification of logging data in the Tarahn Gol area based on random forest algorithm[J]. World Nuclear Geoscience, 2023, 40(1): 43-50, 67.
- [66] Kong W, Chen J, Zhu P. Machine learning-based uranium prospectivity mapping and model explainability research[J]. Minerals 2024, 14: 128.
- [67] Li L J, Yu Y, Bai S S, et al. Towards effective network intrusion detection: a hybrid model integrating GINI index and GBDT with PSO[J]. Journal of Sensors, 2018, 6: 1-9.
- [68] Yu Z, Wang Z, Zeng F, et al. Volcanic lithology identification based on parameter-optimized GBDT algorithm: a case study in the Jilin Oilfield, Songliao Basin, NE China[J]. Journal of Applied Geophysics, 2021, 194: 104443.
- [69] 韩启迪, 张小桐, 申维. 基于梯度提升决策树(GBDT)算法的岩性识别技术[J]. 矿物岩石地球化学通报, 2018, 37(6): 1173-1180.
Han Qidi, Zhang Xiaotong, Shen Wei. Rock identification technology based on gradient boosting decision tree (GBDT) algorithm[J]. Bulletin of Mineralogy, Petrology and Geochemistry, 2018, 37(6): 1173-1180.
- [70] Gu Y F, Zhang D Y. Lithology prediction of tight sandstone reservoirs using GBDT[J]. Progress in Geophysics, 2021, 36(5): 1956-1965.
- [71] 马陇飞, 萧汉敏, 陶敬伟, 等. 基于梯度提升决策树算法的岩性智能分类方法[J]. 油气地质与采收率, 2022, 29(1): 21-29.
Ma Longfei, Xiao Hanmin, Tao Jingwei, et al. Intelligent classification method of lithology based on gradient boosting decision tree algorithm[J]. Oil and Gas Geology and Recovery Efficiency, 2022, 29(1): 21-29.
- [72] 段忠义, 肖昆, 杨亚新, 等. 松辽盆地砂岩型铀矿钻孔岩性的测井识别[J]. 地球物理学进展, 2023, 38(6): 2490-2501.
Duan Zhongyi, Xiao Kun, Yang Yaxin, et al. Logging identification of sandstone-type uranium deposits in the Songliao Basin[J]. Progress in Geophysics, 2023, 38(6): 2490-2501.
- [73] Merembayev T, Yunussov R. Machine learning algorithms for classification geology data from well logging[C]//14th International Conference on Electronics Computer and Computation (ICECCO). Beijing: IEEE, 2018; 206-212.
- [74] Sun Z, Jiang B, Li X, et al. A data-driven approach for lithology identification based on parameter-optimized ensemble learning[J]. Energies, 2020, 13(15): DOI: 10.3390/en13153903.
- [75] Zhang J, He Y, Zhang Y, et al. Well-logging-based lithology classification using machine learning methods for high-quality reservoir identification: a case study of Baikouquan Formation in Mahu Area of Junggar Basin, NW China[J]. Energies, 2022, 15(10): DOI: 10.3390/en15103675.
- [76] Zhang X L, Sun Q, He K. et al. Lithology identification of logging data based on improved neighborhood rough set and AdaBoost[J]. Earth Science Informatics, 2022, 15: 1201-1213.
- [77] 邹学钢. 基于 ADASYN-SSA-XGBoost 的铀矿测井岩性识别研究[D]. 南昌: 东华理工大学, 2023.
Zou Xuegang. Research on uranium ore logging lithology identification based on ADASYN-SSA-XGBoost[D]. Nanchang: Donghua University of Science and Technology, 2023.
- [78] Mukhamediev R I, Kuchin Y, Popova Y, et al. Determination of reservoir oxidation zone formation in uranium wells using ensemble machine learning methods[J]. Mathematics, 2023, 11(22): 46-87.
- [79] Chawla N, Hall L O, Bowyer K H, et al. SMOTE: synthetic minority over-sampling technique[J]. Journal of Artificial Intelligence Research, 2002, 16(1): 321-357.

- [80] Gui Z, Chen J G, Wang C B. Classification of geological imbalance data based on PSA-SMOTE random forest: a case study of geochemical data in eastern Tianshan Mountains[J]. Journal of Guilin University of Technology, 2017, 37(4): 587-593.
- [81] 段忠义, 肖昆, 杨亚新, 等. 基于集成学习的松辽盆地砂岩型铀矿地层岩性自动识别研究[J]. 原子能科学技术, 2023, 57(12): 2443-2454.
Duan Zhongyi, Xiao Kun, Yang Yaxin, et al. Research on automatic identification of lithology in sandstone-type uranium deposits in the Songliao Basin based on ensemble learning[J]. Atomic Energy Science and Technology, 2023, 57(12): 2443-2454.
- [82] Ibrahim B, Ahenkorah I, Ewusi A, et al. A novel XRF-based lithological classification in the Tarkwaian paleo placer formation using SMOTE-XGBoost[J]. Journal of Geochemical Exploration, 2023, 245: 107147.
- [83] Zhou K B, Zhang J, Ren Y, et al. A gradient boosting decision tree algorithm combining synthetic minority oversampling technique for lithology identification[J]. Geophysics, 2020, 85(4): 147-158.
- [84] Shi S. Lithology identification of uranium drilling based on SMOTE algorithm logic-CNN[J]. Journal of Biotech Research, 2024, 18: 53-64.