

文章编号:1009-9603(2022)01-0021-09

DOI:10.13673/j.cnki.cn37-1359/te.2022.01.003

基于梯度提升决策树算法的岩性智能分类方法

马陇飞^{1,2}, 萧汉敏^{1,2,3}, 陶敬伟⁴, 苏致新⁵

(1.中国科学院大学 工程科学学院,北京 100049; 2.中国科学院渗流流体力学研究所,河北 廊坊 065007;
3.中国石油勘探开发研究院,北京 100083; 4.上海帕科信息科技有限公司,上海 200235;
5.中国石油大庆油田有限责任公司 勘探事业部,黑龙江 大庆 163453)

摘要:岩性识别是油气勘探开发领域一项重要的基础工作。针对致密砂岩储层岩石成分复杂、岩性多样和岩性常规测井识别受限等问题,利用机器学习算法在数据分析上的强大功能,采用泛化能力出众的梯度提升决策树(GBDT)算法解决岩性识别中人力和物力耗费大的问题。以鄂尔多斯盆地三叠系延长组长7段致密砂岩储层为研究对象,通过敏感分析选取声波时差、自然伽马、电阻率、泥质含量、自然电位、有效孔隙度、含水饱和度和密度8个测井参数,构建基于GBDT算法的岩性识别模型,结合实际数据进行验证和应用效果分析。与朴素贝叶斯、随机森林、支持向量机和人工神经网络算法岩性识别相比,GBDT算法岩性识别准确率达到92%,高精度的GBDT算法岩性识别模型为致密砂岩储层岩性精确识别提供了新的解决途径。

关键词:致密砂岩储层;岩性识别;机器学习;GBDT模型;鄂尔多斯盆地

中图分类号:TE319

文献标识码:A

Intelligent lithology classification method based on GBDT algorithm

MA Longfei^{1,2}, XIAO Hanmin^{1,2,3}, TAO Jingwei⁴, SU Zhixin⁵

(1.College of Engineering Science, University of Chinese Academy of Sciences, Beijing City, 100049, China; 2.Institute of Porous Flow & Fluid Mechanics, Chinese Academy of Sciences, Langfang City, Hebei Province, 065007, China; 3.PetroChina Research Institute of Petroleum Exploration & Development, Beijing City, 100083, China; 4.Shanghai Pukka Information Tech Ltd., Shanghai City, 200235, China; 5.Exploration Department, Daqing Oilfield Co., Ltd., PetroChina, Daqing City, Heilongjiang Province, 163453, China)

Abstract: Lithology identification is a vital basic work in the field of oil and gas exploration and development. Tight sandstone reservoirs suffer from complex rock composition, diverse lithology, and limited lithology identification by conventional logging. As machine learning is powerful in data analysis, this paper proposed a gradient boosting decision tree (GBDT) algorithm with strong generalization ability to cut down large manpower and material resource consumption in lithology identification. Taking Chang7 Member of Yanchang Formation in Ordos Basin as the research object, it selects eight logging parameters including acoustic time difference (AC), natural gamma ray (GR), resistivity (RT), clay content (SH), natural potential (SP), effective porosity (POR), water saturation (S_w), and density (DEN) through sensitive analysis to build a lithology identification model based on GBDT algorithm. Actual data were applied to verify and analyze the application effect. The accuracy of the GBDT algorithm can reach 92%, compared with that of other methods such as naive Bayes, random forest, support vector machine, and artificial neural network for lithology identification. The high-precision lithology identification model based on the GBDT algorithm provides a new solution for tight sandstone reservoir evaluation.

Key words: tight sandstone reservoir; lithology identification; machine learning; GBDT model; Ordos Basin

岩性识别是油气勘探开发领域一项重要的基础工作,地层对比、沉积模拟和有利区带预测等地

收稿日期:2021-07-23。

作者简介:马陇飞(1995—),男,陕西宝鸡人,在读硕士研究生,从事油气田开发、机器学习和深度学习方面的研究。E-mail:1092797696@qq.com。

基金项目:中国石油基础性前瞻性研究专项“致密储层渗流通道表征技术及渗流机理研究”(2021DJ2201)。

质工作离不开岩性资料。因此,如何从测井数据中获取准确的岩性信息逐渐成为地质和地球物理领域的研究热点^[1-8]。砂岩和泥岩是典型的储层与非储层岩性,它们有各自特定的测井响应,因此砂、泥岩储层岩性预测一般采用2条或3条测井曲线组成的二维或三维交会图完成。随着常规油气藏的枯竭,油气公司开始转向非常规油气藏的勘探开发,如页岩油气藏和低渗透油藏。非常规油气藏复杂的沉积环境给测井解释带来了更多的挑战,需要综合考虑时间、设施预算和人工成本。因此,碳酸盐岩、致密砂岩和页岩等非常规储层的有效开发是当前勘探工作的新目标^[9-15]。非常规油气藏的岩性成分复杂,测井响应相似,很难进行岩性预测^[2-3,16-23]。例如,鄂尔多斯盆地三叠系延长组长7段致密砂岩储层岩性主要为细砂岩、泥岩、碳质泥岩和泥质粉砂岩等,这些岩性的声波时差、电性和放射性测井曲线通常呈现相似值,因此,用交会图方法无法正确显示每种岩性的特定测井响应^[16-17,24]。为了解决复杂岩性预测问题,在对交会图研究的基础上,创造性提出了更多有价值的方法,目前主要分为统计学方法和人工神经网络方法。

早期岩性解释技术不需要计算,通过直观地识别多条测井曲线之间的测井信号分离或独特趋势,采用定性方法完成。在实践中,这种方法能够快速评估,特别是在一致的间隔深度上,但该方法需要大量数据集分析,并依赖于该地区的地质历史^[5,25]。现代计算机的进步促进了岩性定量识别方法的发展,提高了识别的速度和准确性。用作岩性识别工具的数学技术多种多样,比如聚类分析、模糊逻辑和人工神经网络等算法。其中机器学习算法作为人工神经网络算法的一个分支,在世界范围内蓬勃发展和应用,将该算法应用于储层岩性识别评价变得尤为重要。在机器学习算法中应用监督学习算法,该算法可以学习、识别各种数据集之间的复杂关联,提取必要的信息,提高储层岩性评价精度,降低勘探风险,确定盈亏平衡点^[26-27]。梯度提升决策树(GBDT, gradient boosting decision tree)是解决数据不均衡问题的机器学习算法,其通过回归树与目标值之间的残差进行归类分析,再利用逐步提升算法不断减小残差,以使计算值逐渐逼近目标值^[28-32]。由于不同的残差在回归树下会得到不同的处理,所以即使样本中有错误点,训练的结果也不会受影响,表现出GBDT算法不仅有合理的训练过程,还有良好的鲁棒性^[33]。为此,基于GBDT算法在模式识

别上具有多种优势,提出利用GBDT算法来识别致密砂岩储层岩性,并对模型的计算原理和实验验证进行阐述,生成研究区储层岩性分类图,以期鄂尔多斯盆地三叠系延长组长7段致密砂岩储层的岩性精确识别提供新的解决途径。

1 GBDT算法原理及步骤

1.1 GBDT算法原理

GBDT算法是为了确保通过不断减少训练过程中产生的学习误差率来进行分类或回归,并面对实际生产数据类型不均衡而开发的,从一开始就被普遍使用。GBDT算法具有高预测精度、处理一致和离散形式的数据能力、对某些鲁棒损失函数的利用以及对异常值的极端鲁棒性等属性^[34]。GBDT算法如图1所示,在GBDT算法的每次迭代中使用一种称为分类回归树(CART tree)的基学习器。CART tree适用于高偏差、低方差和足够的深度。在回归问题中,执行迭代过程,每个基学习器基于前一个基学习器的学习误差率进行训练,采用梯度下降技

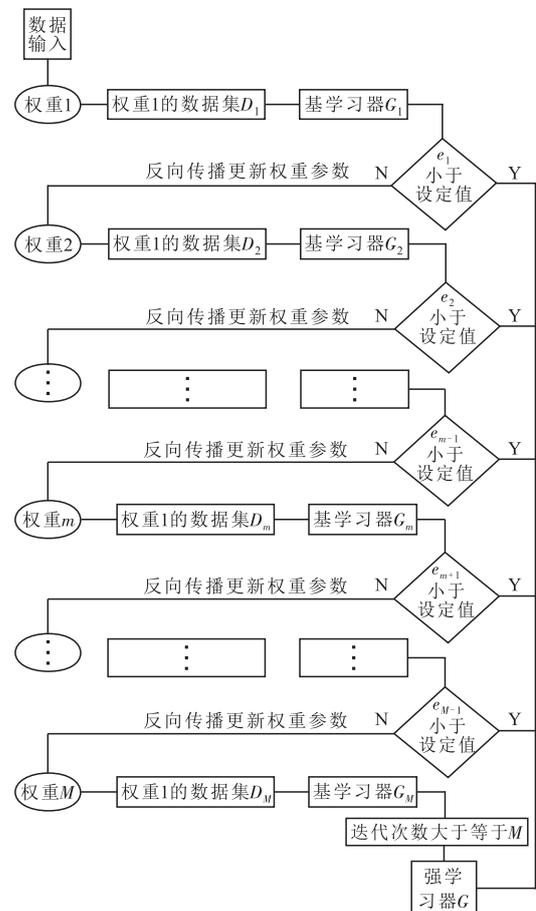


图1 GBDT算法示意(据文献[35]修改)

Fig.1 Training program of GBDT algorithm (Modified according to Reference[35])

术对损失函数进行负梯度拟合回归树,以保证决策模型的持续改进^[36]。

GBDT算法的主要思路是:在梯度方向上训练一个新的学习器来降低前一个学习器的学习误差率,并且新的学习器就是在前一个学习器的基础上迭代生成,其计算公式^[35]为:

$$F_m(x) = F_{m-1}(x) + \rho h(x) \quad 1 \leq m \leq M \quad (1)$$

GBDT算法的具体步骤主要分为以下4步。

第1步,首先对一个学习器进行初始化,初始学习器计算公式为:

$$F_0(x) = \operatorname{argmin}_{\rho} \sum_{i=1}^N L(y_i, \rho) \quad (2)$$

第2步,计算当前损失函数的负梯度值,迭代次数为M,即计算此次迭代中回归树的拟合目标,其计算公式为:

$$r_{m,i} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x_i) = F_{m-1}(x_i)} \quad i = 1, 2, 3, \dots, N \quad (3)$$

第3步,通过第m次迭代,获得最优的基分类器,其计算公式为:

$$\alpha_m = \operatorname{argmin}_{\alpha} \sum_{i=1}^N [r_{m,i} - \beta h(x_i; \alpha_m)] \quad (4)$$

基于线性寻优等方式计算最优的学习率 ρ_m ,更新下一个学习器,其计算公式为:

$$F_m(x) = F_{m-1}(x) + \rho_m h(x; \alpha_m) \quad (5)$$

若 $m = M$ 迭代结束,否则重复步骤2—3。

第4步,生成最终的强学习器G。

1.2 GBDT算法岩性识别模型构建

根据GBDT算法原理,提出解决岩性识别的GBDT算法计算流程如图2所示。由于岩性识别为多分类问题,故损失函数宜选择交叉熵损失函数;优化器选择在迭代过程中自适应调整学习率的Adam优化器进行网络学习,以提升网络学习速度。为防止网络发生过拟合,使用dropout正则化提升模型的泛化能力。

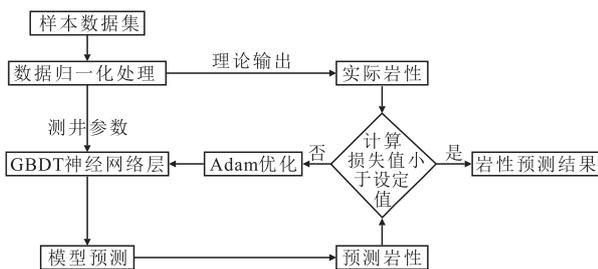


图2 岩性识别的GBDT算法计算流程
Fig.2 Calculation flow of GBDT algorithm for lithology identification

2 应用实例

本次研究数据为鄂尔多斯盆地三叠系延长组长7段致密砂岩储层岩性数据,该储层为半深湖-深湖重力流沉积,因受后期成岩等地质影响,岩性类型复杂多样^[37]。现场钻井取心和岩屑录井资料显示,主要有泥岩、碳质泥岩、泥质粉砂岩、粉砂质泥岩、细砂岩和油页岩等6种岩石类型,其中细砂岩储层为主要的储油层。

2.1 测井参数敏感性分析

通过对测井参数和岩性的分析^[38-40],选取声波时差(AC)、自然伽马(GR)、电阻率(RT)、泥质含量(SH)、自然电位(SP)、有效孔隙度(POR)、含水饱和度(S_w)和密度(DEN)8个测井参数,以A12井为例,对碳质泥岩、泥岩、泥质粉砂岩、粉砂质泥岩和细砂岩等5种岩性做盒形图进行分析。由图3可知:不同岩性对应的测井响应特征存在明显差异。主要体现在:①细砂岩和泥质粉砂岩的声波时差最低(图3a),而其他岩性的声波时差存在重叠部分,因此不能通过声波时差有效划分岩性。细砂岩的自然伽马和泥质含量低(图3b,3d),而泥岩的泥质含量表现出较高值,则可以通过声波时差和泥质含量来划分泥岩与细砂岩。②有效孔隙度和含水饱和度几乎无法有效地区分岩性(图3f,3g)。③自然电位能够有效地区分岩性(图3e),电阻率和密度均存在重叠部分,无法有效地区分岩性(图3c,3h)。④数据夹杂噪音,导致个别岩性测井参数取值区间较大,甚至超出正常值,使得盒形图的极值范围变化较大。因此,为进一步统计有实际意义的不同岩性的测井参数响应特征值,将各测井参数值的25%~75%作为实际岩性响应范围的上下限(表1)。

结合A12井不同岩性测井参数盒形图(图3)和岩性测井参数响应特征值上、下限统计结果(表1),可以得到以下结论:①不同岩性对应测井参数响应特征虽各不相同,但存在一定的重叠部分,即测井参数对岩性响应敏感性差异较大。②各测井参数对细砂岩响应特征的变化范围较小,其中有效孔隙度取值范围最大,泥质含量取值范围最小;碳质泥岩、泥质粉砂岩、粉砂质泥岩的有效孔隙度、含水饱和度取值范围基本不变;碳质泥岩的声波时差、电阻率、自然伽马和密度取值范围变化大;泥岩的泥质含量和声波时差取值范围大,其他测井参数取值范围变化小。

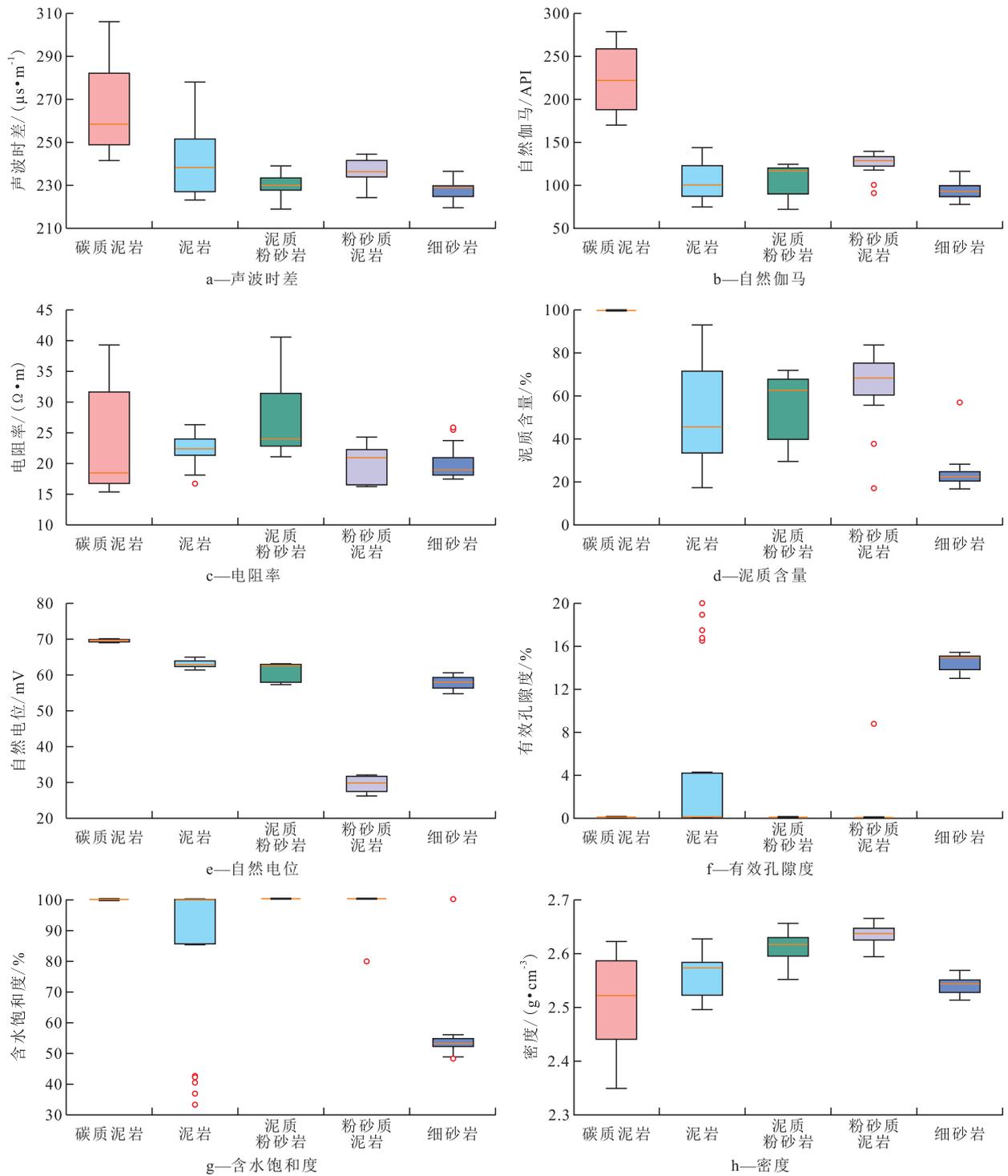


图3 鄂尔多斯盆地三叠系延长组长7段A12井不同岩性测井参数盒形图

Fig.3 Box diagram of different lithologic logging parameters in Well A12 of Chang7 Member of Triassic Yanchang Formation, Ordos Basin

2.2 模型训练结果对比

选取鄂尔多斯盆地三叠系延长组长7段A12, A13, A15, A19和A20井测井曲线共计3 375个训练样本数据点和1 020个验证数据点,其中前4口井的碳质泥岩、泥岩、泥质粉砂岩、粉砂质泥岩和细砂岩的样品个数分别为501,680,515,812和847个。利

用Python软件编程,分别编写机器学习算法(朴素贝叶斯算法、随机森林算法和支持向量机算法)、GBDT算法和神经网络算法程序,将8条测井曲线与对应致密砂岩储层岩性分类数据点进行训练学习,并在测试集上进行分类预测,得到其对应岩性的混淆矩阵图,从而评估算法的岩性识别效果。

表1 鄂尔多斯盆地三叠系延长组长7段不同岩性测井参数响应特征值(25%~75%)
Table1 Response characteristic values of different lithologic logging parameters in Chang7 Member of Triassic Yanchang Formation, Ordos Basin (25%~75%)

岩性	AC/($\mu\text{s}\cdot\text{m}^{-1}$)	GR/API	RT/($\Omega\cdot\text{m}$)	SH/%	SP/mV	POR/%	S _w /%	DEN/($\text{g}\cdot\text{cm}^{-3}$)
碳质泥岩	248.98~282.30	188.06~258.61	16.80~31.68	100.00	69.70~70.22	0.10	100	2.44~2.59
泥岩	227.19~251.49	87.86~123.43	21.35~23.95	34.09~71.29	62.20~64.2	0.05~3.40	85.47~100	2.52~2.58
泥质粉砂岩	227.69~233.45	89.64~121.66	23.01~31.53	39.82~68.43	58.23~63.31	0.10	99.86	2.60~2.63
粉砂质泥岩	233.95~242.22	122.25~134.11	16.58~22.29	60.80~75.82	27.75~31.56	0.09	99.99	2.64~2.65
细砂岩	224.68~226.69	86.68~100.31	18.10~20.84	20.26~24.79	56.42~59.63	11.11~12.11	52.04~55.49	2.53~2.55

朴素贝叶斯、随机森林和支持向量机算法分别通过 sklearn.tree, sklearn.ensemble 和 sklearn.svm 程序包直接导入运行,算法参数调节直接通过调节相关函数的参数取值范围,人工神经网络算法由 Keras 程序包导入,网络层设置4层,训练次数设置200次,学习率取值为0.002。GBDT算法通过 sklearn.ensemble 程序包导入,训练次数设置1600次,学习率取值为0.001。

2.2.1 机器学习算法识别效果

朴素贝叶斯算法 通过朴素贝叶斯算法判别分析岩性得到的混淆矩阵图(图4)可以看出,岩性识别准确率仅为57%。其中对碳质泥岩、泥岩、泥质粉砂岩、粉砂质泥岩和细砂岩的识别准确率分别为79%,39%,0%,52%和70%,说明朴素贝叶斯算法判别分析岩性的识别效果较差,不适合基于测井资料的致密砂岩储层岩性评价。

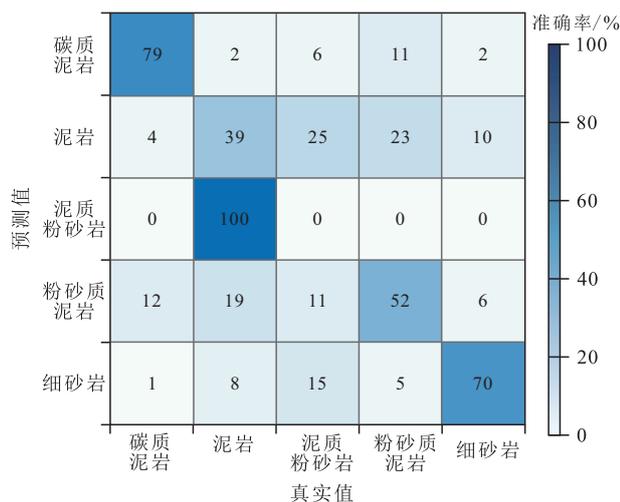


图4 朴素贝叶斯算法判别分析岩性混淆矩阵图

Fig.4 Lithology confusion matrix of discriminant analysis by Naive Bayes algorithm

随机森林算法 随机森林算法是机器学习的一种监督算法,是利用多棵树对样本岩性数据进行训练并预测的一种分类器算法。通过随机森林算法判别分析岩性的混淆矩阵图(图5)可以看出,随机森林算法岩性识别准确率为74%。其中对细砂

岩的识别准确率高达94%,但对其他岩性识别的准确率较低,因此,随机森林算法也不适用于基于测井资料的致密砂岩储层岩性评价。

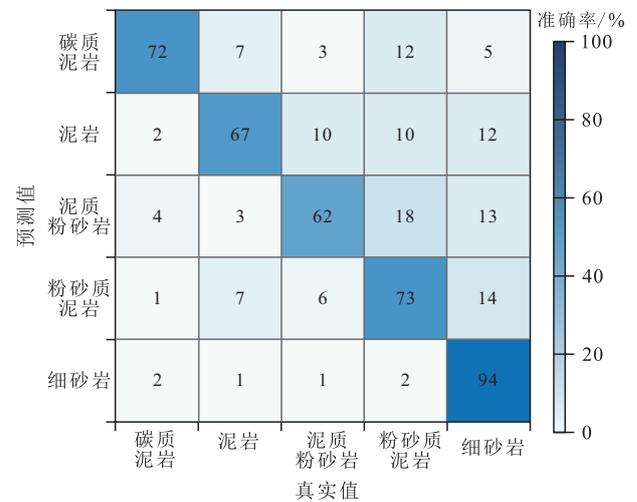


图5 随机森林算法判别分析岩性混淆矩阵图

Fig.5 Lithology confusion matrix of discriminant analysis by random forest algorithm

支持向量机算法 支持向量机算法是当前较为流行的机器学习算法,其为一种二分类算法,当处理多分类问题时,需要构造合适的多分类器。由支持向量机算法判别分析岩性的混淆矩阵图(图6)可以看出,该算法的岩性识别准确率为63%,对碳质泥岩、泥质粉砂岩和细砂岩识别准确率较高,但对泥岩和粉砂质泥岩识别准确率较低。因此,支持向量机算法也不适合基于测井资料的致密砂岩储层岩性评价。

2.2.2 人工神经网络算法识别效果

人工神经网络算法是当前人工智能领域非常重要的算法。由人工神经网络判别分析岩性的混淆矩阵图(图7)可以看出,人工神经网络算法的岩性识别准确率为83%。对泥质粉砂岩的识别准确率为55%,对细砂岩识别准确率达86%。可见人工神经网络算法对致密砂岩储层岩性识别较差,因此人工神经网络算法也不适用于基于测井资料的致密砂岩储层岩性评价。

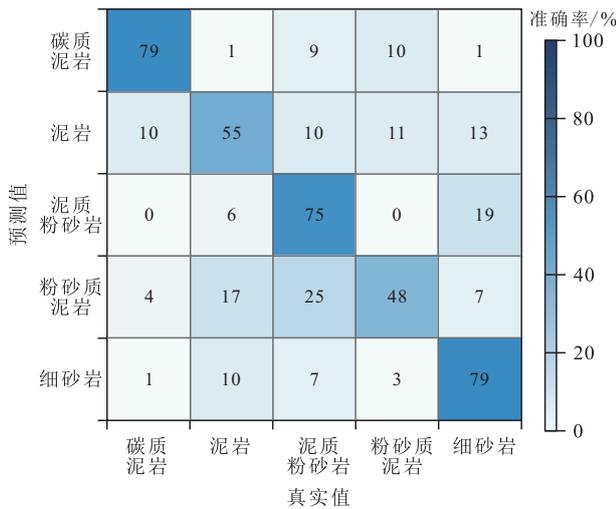


图6 支持向量机算法判别分析岩性的混淆矩阵图

Fig.6 Lithology confusion matrix of discriminant analysis by support vector machine algorithm

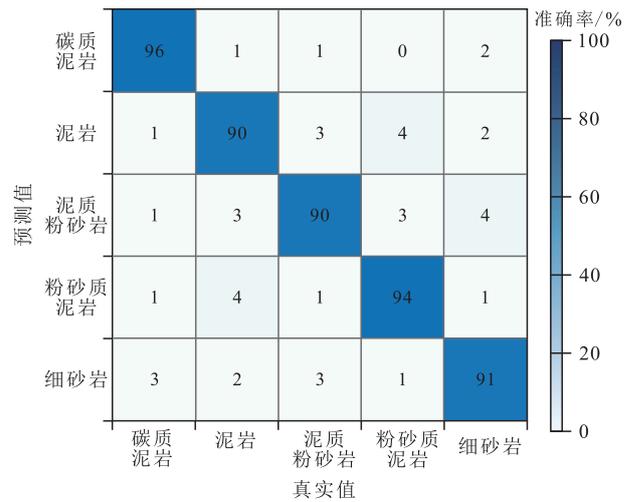


图8 GBDT算法判别分析岩性的混淆矩阵图

Fig.8 Lithology confusion matrix of discriminant analysis by GBDT algorithm

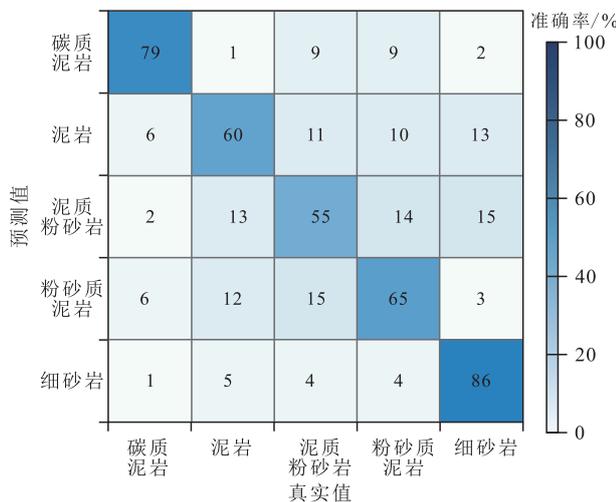


图7 人工神经网络算法判别分析岩性的混淆矩阵图

Fig.7 Lithology confusion matrix of discriminant analysis by artificial neural network algorithm

2.2.3 GBDT 算法识别效果

由 GBDT 算法判别分析岩性的混淆矩阵图(图 8)可以看出,GBDT 算法的岩性识别准确率为 92%,对碳质泥岩、泥岩、泥质粉砂岩、粉砂质泥岩和细砂岩识别准确率分别为 96%、90%、90%、94% 和 91%,识别效果较好。由此可见,GBDT 算法非常适合致密砂岩储层岩性识别,其精度足以提供可靠的预测结果。

利用训练好的 GBDT 算法对研究区 A23 井储层岩性进行评价。结果(图 9)表明,GBDT 算法的岩性识别结果与该井钻井取心、岩屑录井结论接近。因此,可以推断 GBDT 算法适合对鄂尔多斯盆地三叠系延长组长 7 段致密砂岩储层岩性类型进行识别,且精度更高,在致密砂岩储层岩性评价中具有更显著的优势。

鄂尔多斯盆地三叠系延长组长 7 段储层为典型的低-特低孔隙度、低-特低渗透率的致密砂岩储层^[27]。朴素贝叶斯、随机森林、支持向量机和人工神经网络算法需要岩性类型数据量基本一致,才可以保证模型的预测准确率较高,而实际生产中每一种岩性类型的数据量是不均衡的,当处理提供的岩性类型数据量不均衡时,朴素贝叶斯、随机森林、支持向量机和人工神经网络算法岩性识别的准确率普遍较低,而笔者提出的 GBDT 算法是针对数据量不均衡而设计的,对于处理数据量不均衡问题有着独特的优势。GBDT 算法对于致密砂岩储层岩性识别的准确率较高,能够在实际生产中得到很好的应用。

3 结论

GBDT 算法在训练过程中对数据量不均衡的不同类型训练样本做出不同的处理,使得建立的模型在识别相似预测样本时能够给出准确的判定,避免了数据强化的操作,有效捕捉到了岩性与测井参数之间的非线性关系,实现了对岩性的有效识别。

与朴素贝叶斯、随机森林、支持向量机和人工神经网络等算法相比,基于 GBDT 算法的岩性识别方法更加精确,准确率达到 92%,表明该算法具有良好的泛化能力,即鲁棒性很好,能够解决实际生产需求。

符号解释

- D_1 ——训练基学习器 G_1 的权重 1 的数据集;
- D_2 ——训练基学习器 G_2 的权重 1 的数据集;

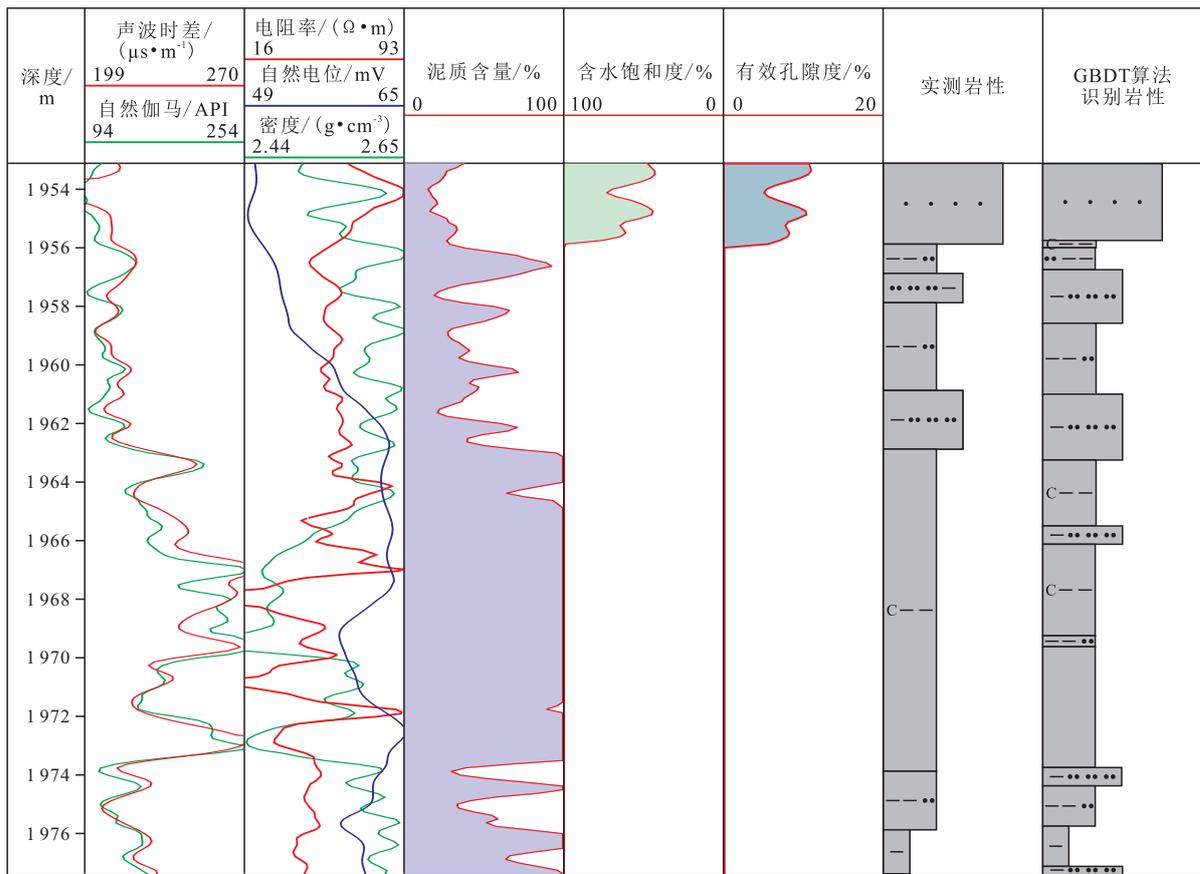


图9 鄂尔多斯盆地三叠系延长组长7段 A23 井 GBDT 算法岩性识别结果

Fig.9 Lithologic identification results of Well A23 of Chang7 Member of Triassic Yanchang Formation in Ordos Basin based on GBDT algorithm

D_m ——训练基学习器 G_m 的权重 1 的数据集;
 D_M ——训练基学习器 G_M 的权重 1 的数据集;
 e_m ——迭代第 m 次的基学习器 G_m 的学习误差率;
 $F_m(x)$ ——当前学习器;
 $F_{m+1}(x)$ ——新学习器;
 $F_{m-1}(x)$ ——前一个学习器;
 $F_{m-1}(x_i)$ ——第 i 次迭代的前一个学习器;
 $F_0(x)$ ——初始化学习器;
 $F(x_i)$ ——前一个学习器的预测结果;
 G ——强学习器;
 G_1 ——权重 1 的数据集训后得到权重 1 的基学习器;
 G_2 ——权重 1 的数据集训后得到权重 2 的基学习器;
 G_m ——权重 1 的数据集训后得到权重 m 的基学习器;
 G_M ——权重 1 的数据集训后得到权重 M 的基学习器;
 $h(x)$ ——当前损失函数负梯度方向上拟合得到的基学习器;
 $h(x_i; \alpha_m)$ ——有最优基分类器 α_m 的损失函数负梯度方向上拟合得到的基学习器;
 i ——样本数据编号, $i = 1, 2, 3, \dots, N$;
 $L(y_i, \rho)$ ——损失函数;

$\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}$ ——当前迭代中回归树的拟合目标;
 m ——迭代次数;
 M ——迭代总次数;
 N ——总数;
 r ——损失函数的负梯度值;
 $r_{m,i}$ ——当前损失函数的负梯度值;
 x ——样本数据;
 x_i ——第 i 个样本数据;
 y_i ——第 i 个预测目标;
 α_m ——最优的基分类器;
 β ——系数;
 ρ ——学习率;
 ρ_m ——最优的学习率。

参考文献

[1] JIANG J, RUI Z, HAZLETT R, et al. An integrated technical-economic model for evaluation CO₂ enhanced oil recovery development[J]. Global Warming Focus, 2019, 247: 191-211.
 [2] CHEN W, XIE X, WANG J, et al. A comparative study of logistic model tree, random forest, and classification and regression tree

- models for spatial prediction of landslide susceptibility [J]. *Catena*, 2017, 151: 147-160.
- [3] ULVMOEN M, HAMMER H. Bayesian lithology/fluid inversion-comparison of two algorithms [J]. *Computational Geosciences*, 2010, 14(2): 357-367.
- [4] HSIEH B Z, LEWIS C, LIN Z S. Lithology identification of aquifers from geophysical well logs and fuzzy logic analysis: Shui-Lin Area, Taiwan [J]. *Computational Geosciences*, 2005, 31(3): 263-275.
- [5] LV A, CHENG L, AGHIGHI M A, et al. A novel workflow based on physics-informed machine learning to determine the permeability profile of fractured coal seams using downhole geophysical logs [J]. *Marine and Petroleum Geology*, 2021, 131: 1-14.
- [6] 王志新. 油田开发规划多目标不确定性优化模型及其应用 [J]. *大庆石油地质与开发*, 2020, 39(5): 86-90.
WANG Zhixin. Multi-objective uncertainty optimization model of oilfield development planning and its application [J]. *Petroleum Geology & Oilfield Development in Daqing*, 2020, 39(5): 86-90.
- [7] 闫健, 秦大鹏, 王平平, 等. 鄂尔多斯盆地致密砂岩储层可动流体赋存特征及其影响因素 [J]. *油气地质与采收率*, 2020, 27(6): 47-56.
YAN Jian, QIN Dapeng, WANG Pingping, et al. Occurrence characteristics and main controlling factors of movable fluid in tight sandstone reservoirs in Ordos Basin [J]. *Petroleum Geology and Recovery Efficiency*, 2020, 27(6): 47-56.
- [8] 陈晓智, 陈桂华, 李祺鑫, 等. 致密砂岩气储量不确定性计算方法及意义 [J]. *断块油气田*, 2019, 26(1): 17-20, 41.
CHEN Xiaozhi, CHEN Guihua, LI Qixin, et al. Uncertainty reserves calculation for tight sandstone gas and its significance [J]. *Fault-Block Oil and Gas Field*, 2019, 26(1): 17-20, 41.
- [9] CUI G D, REN S R, EZEKIEL J, et al. The influence of complicated fluid-rock interactions on the geothermal exploitation in the CO₂ plume geothermal system [J]. *Applied Energy*, 2018, 227: 49-63.
- [10] 谷志宇, 刘恩涛, 王香增, 等. 鄂尔多斯盆地东南部延长组七段页岩发育特征及勘探潜力 [J]. *油气地质与采收率*, 2021, 28(1): 95-105.
GU Zhiyu, LIU Entao, Wang Xiangzeng, et al. Development characteristics and exploration potential of shale in Chang7 Member in Southeast of Ordos Basin [J]. *Petroleum Geology and Recovery Efficiency*, 2021, 28(1): 95-105.
- [11] RUI Z H, CUI K H, WANG X Q, et al. A comprehensive investigation on performance of oil and gas development in Nigeria: technical and non-technical analyses [J]. *Energy*, 2018, 158: 666-680.
- [12] 彭金宁, 邱岐, 王东燕, 等. 苏北盆地古近系阜宁组致密油赋存状态与可动性 [J]. *石油实验地质*, 2020, 42(1): 53-59.
PENG Jinning, QIU Qi, WANG Dongyan, et al. Occurrence and recoverability of tight oil in Paleogene Funing Formation, Subei Basin [J]. *Petroleum Geology & Experiment*, 2020, 42(1): 53-59.
- [13] WEIJERMARS R. Natural resource wealth optimization: a review of fiscal regimes and equitable agreements for petroleum and mineral extraction projects [J]. *Natural Resources Research*, 2015, 24: 385-441.
- [14] MCGLADE C, SPEIRS J, SORRELL S. Unconventional gas—a review of regional and global resource estimates [J]. *Energy*, 2013, 55(1): 571-584.
- [15] THOMAS S Ahlbrandt. Future petroleum energy resources of the world [J]. *International Geology Review*, 2002, 44(12): 1 092-1 104.
- [16] GU Y F, BAO Z D, RUI Z H. Complex lithofacies identification using improved probabilistic neural networks [J]. *Petrophysics*, 2018, 59(2): 245-267.
- [17] SALEM H S. Computer modeling of porosity and lithology for complex reservoirs using well-log measurements [J]. *Energy Sources*, 2000, 22(6): 515-524.
- [18] 韩载华, 赵靖舟, 孟选刚, 等. 鄂尔多斯盆地三叠纪湖盆东部“边缘”长7段烃源岩的发现及其地球化学特征 [J]. *石油实验地质*, 2020, 42(6): 991-1 000.
HAN Zaihua, ZHAO Jingzhou, MENG Xuangang, et al. Discovery and geochemical characteristics of Chang 7 source rocks from the eastern margin of a Triassic lacustrine basin in the Ordos Basin [J]. *Petroleum Geology & Experiment*, 2020, 42(6): 991-1 000.
- [19] 姚泾利, 周新平, 惠潇, 等. 鄂尔多斯盆地西缘古峰庄地区低级序断层封闭性及其控藏作用 [J]. *中国石油勘探*, 2019, 24(1): 72-81.
YAO Jingli, ZHOU Xinping, HUI Xiao, et al. Sealing capabilities and reservoir controlling effects of low-grade faults in Gufengzhuang area, western margin of the Ordos Basin [J]. *China Petroleum Exploration*, 2019, 24(1): 72-81.
- [20] 张哲豪, 魏新善, 弓虎军, 等. 鄂尔多斯盆地定边油田长7致密砂岩储层成岩作用及孔隙演化规律 [J]. *油气地质与采收率*, 2020, 27(2): 43-52.
ZHANG Zhehao, WEI Xinshan, GONG Hujun, et al. Diagenesis characteristics and evolution of porosity of Chang7 tight sandstone reservoir in Dingbian Oilfield, Ordos Basin [J]. *Petroleum Geology and Recovery Efficiency*, 2020, 27(2): 43-52.
- [21] 赖生华, 白璞, 叶超, 等. 鄂尔多斯盆地高家河地区山西组二段下部地震沉积学研究 [J]. *石油实验地质*, 2020, 42(1): 79-87.
LAI Shenghua, BAI Pu, YE Chao, et al. Seismic sedimentology of lower part of second member of Shanxi Formation in Gaojiahe area, Ordos Basin [J]. *Petroleum Geology & Experiment*, 2020, 42(1): 79-87.
- [22] 刘成川, 杨建, 卜淘, 等. 川西窄河道致密砂岩气藏井型井网设计 [J]. *特种油气藏*, 2020, 27(1): 82-88.
LIU Chengchuan, YANG Jian, BU Tao, et al. Well-type and well-pattern design for the narrow channel tight sandstone gas reservoir in Western Sichuan [J]. *Special Oil & Gas Reservoirs*, 2020, 27(1): 82-88.
- [23] 高玉巧, 蔡潇, 何希鹏, 等. 页岩气勘探开发实验测试技术挑战与发展方向 [J]. *油气藏评价与开发*, 2021, 11(2): 164-175.
GAO Yuqiao, CAI Xiao, HE Xipeng, et al. Challenges and development direction of experimental testing technology for shale gas exploration and development [J]. *Reservoir Evaluation and Development*, 2021, 11(2): 164-175.
- [24] HUANG B, PAN B. Characteristics of log responses and lithology determination of igneous rock reservoirs [J]. *Journal of Geophysics*

- and Engineering, 2004, 1(1): 51-55.
- [25] 李国欣, 王峰, 皮学军, 等. 非常规油气藏地质工程一体化数据优化应用的思考与建议[J]. 中国石油勘探, 2019, 24(2): 147-152.
LI Guoxin, WANG Feng, PI Xuejun, et al. Optimized application of geology-engineering integration data of unconventional oil and gas reservoirs [J]. China Petroleum Exploration, 2019, 24(2): 147-152.
- [26] DEV V A, EDEN M R. Formation lithology classification using scalable gradient boosted decision trees[J]. Computers and Chemical Engineering, 2019, 128: 392-404.
- [27] 刘巍, 刘威, 谷建伟, 等. 利用卡尔曼滤波和人工神经网络相结合的油藏井间连通性研究[J]. 油气地质与采收率, 2020, 27(2): 118-124.
LIU Wei, LIU Wei, GU Jianwei, et al. Research on interwell connectivity of oil reservoirs based on Kalman filter and artificial neural network [J]. Petroleum Geology and Recovery Efficiency, 2020, 27(2): 118-124.
- [28] 梁宇, 任战利, 王彦龙, 等. 鄂尔多斯盆地子长地区延长组流体包裹体特征与油气成藏期次[J]. 石油与天然气地质, 2011, 32(2): 182-191.
LIANG Yu, REN Zhanli, WANG Yanlong, et al. Characteristics of fluid inclusions and reservoiring phases in the Yanchang Formation of Zichang area, the Ordos Basin [J]. Oil & Gas Geology, 2011, 32(2): 182-191.
- [29] YANG S Y, WU J P, DU Y M, et al. Ensemble learning for short-term traffic prediction based on gradient boosting machine [J]. Journal of Sensors, 2017: 1-15.
- [30] 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: 1-9.
- [31] LIAO Z, WAN S, HE Y, et al. Classification of small GTPases with hybrid protein features and advanced machine learning techniques [J]. Current Bioinformatics, 2018, 13(5): 492-500.
- [32] ZHANG C S, ZHANG Y, SHI X J, et al. On incremental learning for gradient boosting decision trees[J]. Neural Processing Letters, 2019, 50(1): 957-987.
- [33] 谷宇峰, 张道勇, 鲍志东, 等. GBDT识别致密砂岩储层岩性[J]. 地球物理学进展, 2021, 36(5): 1956-1965.
GU Yufeng, ZHANG Daoyong, BAO Zhidong, et al. Lithology prediction of tight sandstone reservoirs using GBDT[J]. Progress in Geophysics, 2021, 36(5): 1956-1965.
- [34] MA X, DING C, LUAN S, et al. Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method[J]. Journal of Robotics & Machine Learning, 2017, 18: 2303-2310.
- [35] 王瑞妮. 基于流形学习和GBDT的异常流量检测方法[D]. 哈尔滨: 哈尔滨工程大学, 2020.
WANG Ruini. Anomaly traffic detection based on manifold learning and GBDT[D]. Harbin: Harbin Engineering University, 2020.
- [36] XIA Y, LIU C, LI Y, et al. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring [J]. Expert Systems with Applications, 2017, 78: 225-241.
- [37] JEROME H Friedman. Greedy function approximation: a gradient boosting machine [J]. The Annals of Statistics, 2001, 29(5): 1189-1232.
- [38] 罗群, 吴安彬, 王井伶, 等. 中国北方页岩气成因类型、成气模式与勘探方向[J]. 岩性油气藏, 2019, 31(1): 1-11.
LUO Qun, WU Anbin, WANG Jingling, et al. Genetic types, generation models, and exploration direction of shale gas in northern China[J]. Lithologic Reservoirs, 2019, 31(1): 1-11.
- [39] 靳军, 王剑, 杨召, 等. 准噶尔盆地克一百断裂带石炭系内幕储层测井岩性识别[J]. 岩性油气藏, 2018, 30(2): 85-92.
JIN Jun, WANG Jian, YANG Zhao, et al. Well logging identification of Carboniferous volcanic inner buried-hill reservoirs in Ke-Bai fault zone in Junggar Basin[J]. Lithologic Reservoirs, 2018, 30(2): 85-92.
- [40] 崔景伟, 朱如凯. 致密砂岩层内强钙质胶结物成因机制及其意义——以鄂尔多斯盆地三叠系延长组长7油层组为例[J]. 吉林大学学报: 地球科学版, 2020, 50(4): 957-967.
CUI Jingwei, ZHU Rukai. Mechanism of strong calcium cementation in tight sandstone and its significance: A case study on Triassic Chang7 oil formation of Yanchang formation in Ordos basin [J]. Journal of Jilin University: Earth Science Edition, 2020, 50(4): 957-967.

编辑 常迎梅