

基于成像测井灰度值的GA-SVR在白云岩储层孔隙度预测中的应用

——以川西地区L组为例

熊晨皓¹, 李响²

¹中石化勘探分公司物探研究院, 四川 成都

²成都理工大学能源学院, 四川 成都

收稿日期: 2022年3月25日; 录用日期: 2022年4月18日; 发布日期: 2022年4月25日

摘要

孔隙度的准确性是优质储层识别和建模的重要基础。白云岩储层孔隙结构复杂, 发育类型多, 其孔隙度难以精确预测。通过采集川西地区L组白云岩储层182个取心样本、常规测井和成像测井资料, 将6种机器学习模型应用于不同测井参数的训练与测试, 并评价不同模型的预测性能。结果表明, 引入不敏感损失函数度量的遗传算法支持向量回归模型的预测效果最好, 决定系数高达0.87。采用常规测井与成像测井相结合的机器学习综合算法, 能有效降低白云岩储层非均质性的影响, 提高孔隙度的解释精度, 为优质白云岩储层的识别奠定了良好的基础。

关键词

孔隙度, 白云岩, 机器学习, 成像测井, 常规测井

Application of GA-SVR Based on Gray Value of Imaging Logging in Porosity Prediction of Dolomite Reservoir

—A Case Study of L Formation in Western Sichuan

Chenhao Xiong¹, Xiang Li²

¹Geophysical Research Institute, Sinopec Exploration Company, Chengdu Sichuan

²College of Energy, Chengdu University of Technology, Chengdu Sichuan

Received: Mar. 25th, 2022; accepted: Apr. 18th, 2022; published: Apr. 25th, 2022

文章引用: 熊晨皓, 李响. 基于成像测井灰度值的 GA-SVR 在白云岩储层孔隙度预测中的应用[J]. 地球科学前沿, 2022, 12(4): 515-524. DOI: 10.12677/ag.2022.124052

Abstract

The accuracy of porosity is an important basis for the identification and modeling of high-quality reservoir. The pore structure of heterogeneous dolomite reservoir is complex and there are many types of development, so it is difficult to accurately predict. By collecting 182 core samples, conventional logging data and imaging logging data of dolomite reservoir in L Formation in Western Sichuan. Six machine learning models are applied to the training and testing of different logging parameters, and the prediction performance of different models is evaluated. The results show that the genetic algorithm support vector regression model with epsilon has the best prediction effect, and the coefficient of determination is 0.87. The integrated method of combining conventional logging and imaging logging with machine learning can effectively reduce the influence of dolomite reservoir heterogeneity, improve the interpretation accuracy of porosity, lay a good foundation for the identification of high-quality dolomite reservoir, and explore a new way of porosity interpretation.

Keywords

Porosity, Dolomite, Machine Learning, Imaging Logging, Well Logging

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

对储层评价而言,物性参数至关重要,其中孔隙度是一个很重要的参数。对于白云岩储层,岩体结构中孔缝洞系统的发育程度和分布范围难以表征[1],对优质储层的准确评价带来了巨大挑战。取心和地球物理测井是油气工业中确定孔隙度的主要方法,通常取心成本高、耗时长、样本少,且仅适用于部分储层。通过寻找常规测井电性参数与孔隙度的相关性往往不能得到满意的效果,很大程度上依赖于解释人员的经验,这一问题在致密碳酸盐岩储层中更为突出[2][3]。

为了解决这些问题,大量机器学习方法被引入到孔隙度预测中,并进行了多方面的研究。常用的机器学习模型主要为神经网络和支持向量机。其中神经网络细分为卷积神经网络(CNN)[4],反向传播神经网络(BP)[5][6],数据分组处理神经网络(GMDH)[7]等。支持向量机[8]按其应用方式分为支持向量分类机(Support Vector Machine, SVM)和支持向量回归机(Support Vector Regression, SVR),其代表优化算法为遗传算法(Genetic Algorithm, GA),粒子群(Particle Swarm Optimization, PSO)及网格搜索(Grid Search, GS)等,这些方法多数具有可靠的预测性能。在成像测井应用方面一些学者进行了深入的探讨,陈义祥、李晓辉等通过电成像测井参数定义算法对孔隙度值进行计算[9][10],平海涛等提出了基于电成像孔隙度谱分析的复杂储层渗透率模型,实现了核磁共振测井与电成像测井的有机结合[11]。不少学者采用图像处理技术对成像测井图像上缝洞面积进行提取,并建立相关性间接的进行孔隙度预测及评价[12][13]。在实际应用中,各种方法均存在不同程度的缺陷,利用常规测井存在相关性较差,解释结果偏差较大;成像测井在采集时易受噪声及异物的影响,出现局部异常值。

为此,本文采用机器学习模型结合上述两种测井作进一步研究,通过对6种机器学习模型的优选,将常规测井与成像测井结合,建立一种综合方法。利用M1井非试气段实测岩心样本,建立预测模型,再用M1井试气段实测岩心样本对模型进行测试校正。利用校正后的模型对M2井L组白云岩储层进行

了可靠的孔隙度解释, 并以 M2 井 L 组实测岩心样本和实际测试段单井日产气量数据作为验证。

2. 方法原理

2.1. 支持向量回归机

支持向量回归机是支持向量在函数回归领域的应用。寻求最优超平面使所有样本点离超平面的“总偏差”最小, 对于非线性回归问题, 一般用线性函数拟合:

$$f(x) = \omega \cdot x_i + b, i = 1, 2, \dots, n \quad (1)$$

式中: ω, b 为模型系数, 其中 $x_i \in R^n$ 为输入量。在利用 SVR 进行训练时, 可引入参数 ε , 即不敏感损失函数度量(epsilon) [14], 以确保对偶变量的稀疏性, 保证全局最小解及提高泛化能力:

$$\begin{cases} y_i - f(x_i) \leq \varepsilon + \xi_i \\ f(x_i) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad i = 1, 2, \dots, n \quad (2)$$

式中: ξ_i, ξ_i^* 为松弛因子。核函数作为支持向量回归机的重要组成部分, 经常用于将非线性可分问题从原始特征空间映射到高维空间, 从而将非线性可分问题转化为线性可分问题, 不同的核函数会导致不同的。核函数作为支持向量回归机的重要组成部分, 经常用于将非线性可分问题从原始特征空间映射到高维空间, 从而将非线性可分问题转化为线性可分问题, 不同的核函数会导致不同的 SVR 性能, 常规的核函数有线性核(Linear Kernel)、多项式核(Polynomial Kernel)及径向基核函数(Radial Basis Function, RBF)等。针对碳酸盐岩非均质性较强、线性不可分等特点, RBF 核函数具有明显的优势。为找到最优的模型参数, 引入网格搜索(GS)、遗传算法(GA)及粒子群(PSO)三种寻优算法, 图 1 为三种寻优算法的工作流程。网格搜索作为最常用方法虽然在小范围搜索时能够得到全局最优解, 但在范围较大时往往会很费时间。而采用遗传和粒子群算法就可以不必遍历网格内的所有的参数点, 极大的节省了计算时间, 但缺点是比网格搜索法更容易陷入局部最优解。

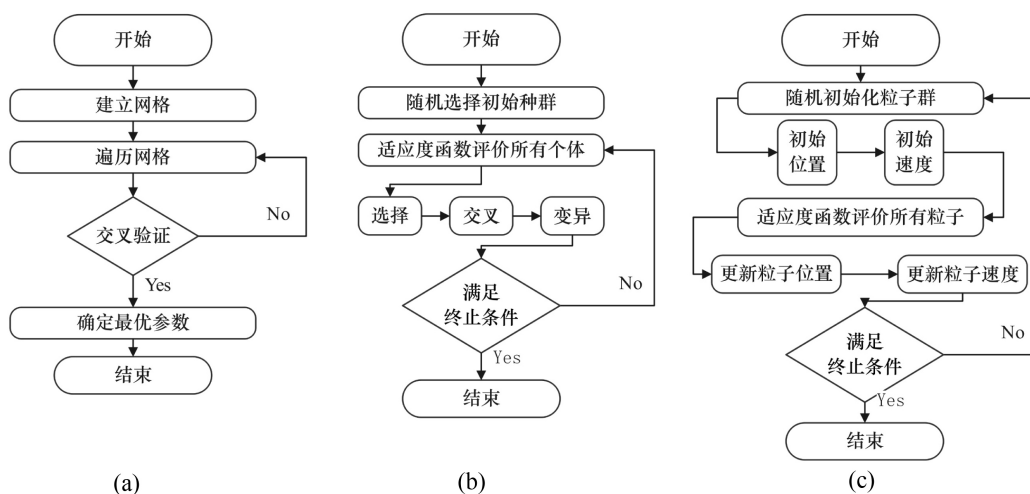


Figure 1. Work flow of parameter optimization: (a) GS; (b) GA; and (c) PSO

图 1. 寻优算法工作流程图: (a) 网格搜索; (b) 遗传算法; (c) 粒子群算法

2.2. 模型评价指标

为更科学的评价各机器学习模型的预测准确度, 采用决定系数。(Coefficient of determination, R^2)、

均方根误差(Root mean square error, RMSE)、平均绝对误差(Average absolute error, AAE)和最大绝对误差(Maximum absolute error, MAE)等四种指标。其中 R^2 、RMSE 和 AAE 反映了预测值与实测值的总体偏差, R^2 越大, RMSE、AAE 越小, 模型性能越好, MAE 反应了预测值与实测值的局部偏差, MAE 越小, 预测准确度越高。计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

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

$$AAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

$$MAE = \max |y_i - \hat{y}_i|, i = 1, \dots, n \quad (6)$$

式(3)~(6)中: y_i 为样本数值, \hat{y}_i 为该样本预测值, \bar{y} 为样本平均值。

2.3. 新方法预测孔隙度的流程

目前利用机器学习算法对孔隙度进行预测多数情况基于常规测井电性参数。相对于常规测井, 电成像测井具有其独特的自身优势, 采集数据时可实现井周 360 度高分辨率成像。成像测井图像真实的体现了沿井深度各地层的表面情况, 其明暗程度一定程度上反映了该区域孔隙结构的发育情况, 由于在非均质碳酸盐岩中, 泥质含量相对较小, 对孔隙度值的影响甚小, 成像测井图像的灰度值可作为确定孔隙度值的一种参数。

本文首先根据实际取心样本孔隙度值、常规测井电性参数及成像测井图像等包含孔隙度信息的高分辨率数据源, 通过岩心归位、岩心标定等常规方法进行深度校正, 采用 Python 语言编程对成像测井图像进行批量处理, 实现了沿深度方向的灰度值提取, 再将样本孔隙度值与测井参数进行关联。在实际操作过程中, 由于成像测井图像通常一口井分为数段, 故在编写算法的时候将相邻图片的原始灰度值矩阵进行拼接操作, 并在深度重叠的部分采用算术平均值法进行拼接, 最后得到完整的单井灰度值曲线。通过对 6 种不同的机器学习模型进行对比分析评价, 确定最佳孔隙度预测模型, 最后对全井孔隙度值进行解释。图 2 为基于常规测井和电成像测井结合的机器学习孔隙度预测流程。

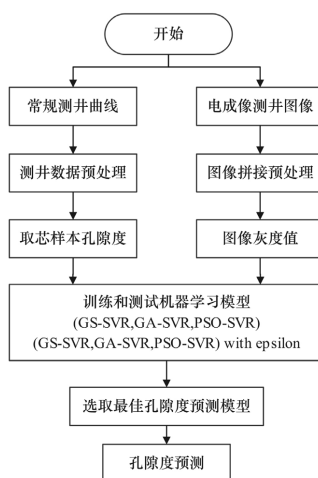


Figure 2. Porosity prediction technology flow

图 2. 孔隙度预测流程

3. 实验数据及模型

3.1. 灰度值与样本分析

川西地区 L 组的白云岩沉积厚度较大, 分布范围较广。由于储层非均质性较强, 储层规模、质量特征缺乏深入认识, 严重制约了勘探突破。研究区中 M1 井、M2 井均为重要的工业气井。

将 M1 井白云岩地层 182 个岩心样本进行总孔隙度和灰度值的频率分布分析, 如图 3 所示。孔隙度值分布在 0.3%~12.25% 区间, 超过 70% 的孔隙度值小于 4%, 灰度值主要分布在 140~180, 仅有少数灰度值小于 140。本文将常规测井曲线自然伽马(GR)、声波时差(DT)、补偿中子(CNL)、密度(DEN)、成像测井灰度值及其求导斜率等作为输入参数, 这些参数包涵了岩石孔隙度的绝大部分原始信息。

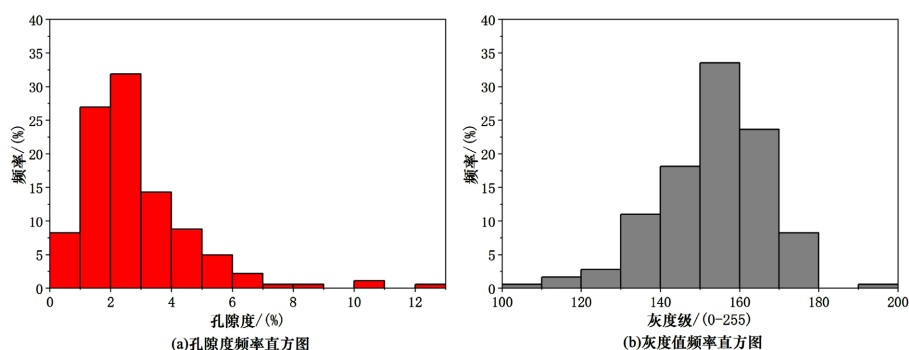


Figure 3. Frequency distribution of the core porosity and gray value
图 3. 岩心孔隙度和灰度值频率分布图

图 4 表明孔隙度与测井响应参数没有明显的线性关系, 进一步证实了白云岩储层的非均质性。为寻找一个可靠的机器学习模型, 将 M1 井所有可用的取心样本数据分为两组: 为保证训练样本的丰富性, 根据孔隙度频率直方图, 抽取不同频率段约 20% 的样本作为测试数据集(40 个样本), 其余样本作为训练数据集(142 个样本)。

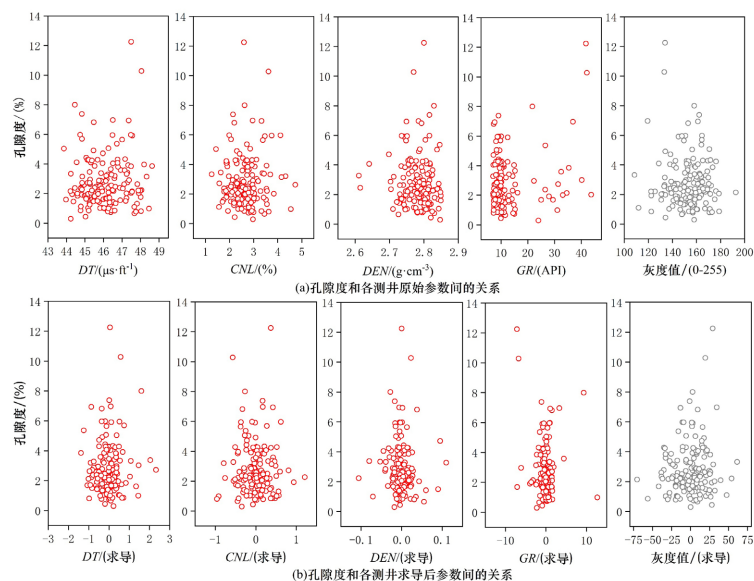


Figure 4. Cross-plot of porosity and various logging parameters
图 4. 孔隙度和各测井参数值交会图

3.2. 基于常规测井的机器学习模型优选

将 8 种测井参数作为输入参数, 孔隙度值作为输出参数, 样本测井参数响应范围见表 1。

Table 1. Log response range of core samples

表 1. 样本测井响应范围

参数	$DT/\mu\text{s}\cdot\text{ft}^{-1}$	slope of DT	$CNL/\%$	slope of CNL	$DEN/\text{g}\cdot\text{cm}^{-3}$	slope of DEN	GR/API	slope of GR
平均值	46.08	-0.01	2.66	<0.01	2.78	<0.01	12.19	-0.16
分布区间	43.9~48.6	-1.4~2.3	1.3~4.7	-1.0~1.2	2.60~2.90	-0.1~0.1	7.0~43.7	-9.3~12.7

4 种评价指标(R^2 , RMSE, AAE, MAE)用于评价训练后 6 种机器学习模型的各项性能。图 5 为不同机器学习模型孔隙度实测值与预测值的交会图, 数据点越靠近预期拟合线 $y = x$ 表明预测准确率越高。不难看出, 在该区域仅依靠常规测井电性参数对白云岩孔隙度进行预测难以取得较好的效果, 不仅归咎于白云岩的非均质性, 还存在于岩心归位、取心深度及实验数据等多方面。各机器学习模型之间的主要区别反映在对 2.5%~7.5%孔隙度区间预测的能力。

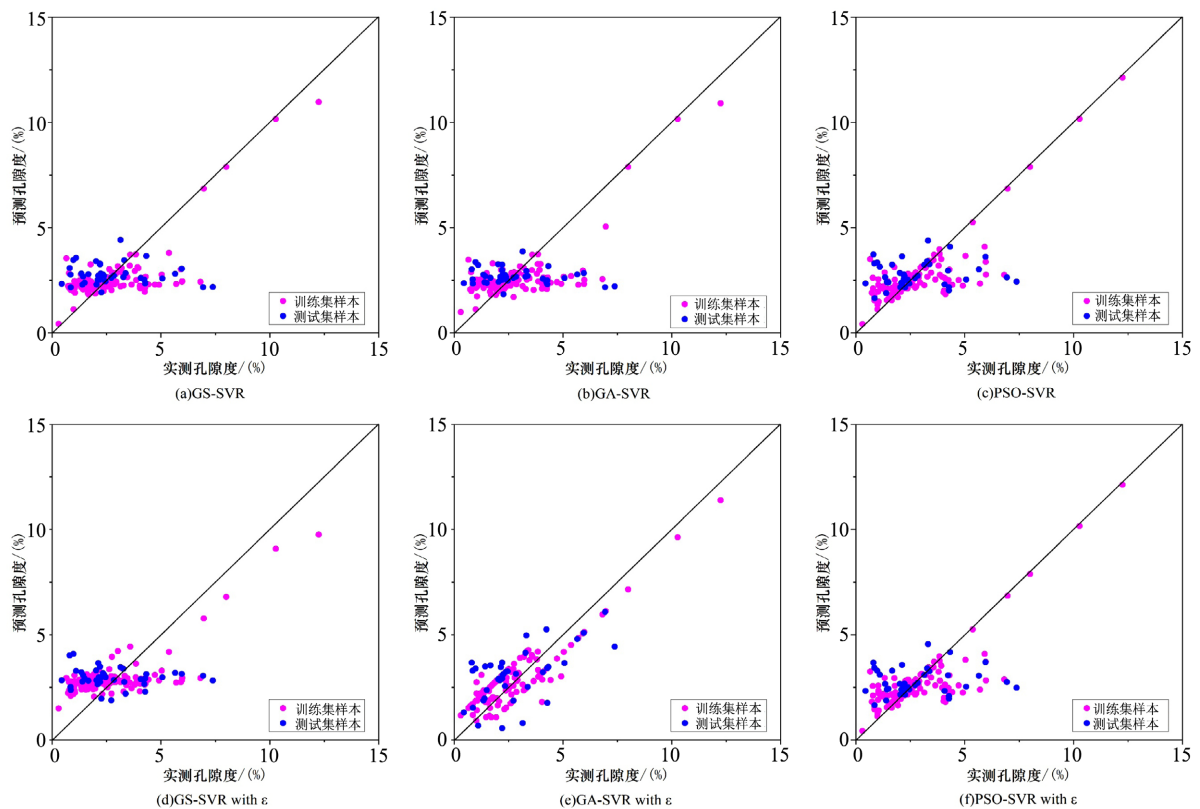


Figure 5. Cross-plot of the core porosity and porosity predicted by the various machine learning models

图 5. 不同机器学习模型实测孔隙度与预测孔隙度交会图

图 6 可以看出, 在引入不敏感损失函数度量 ε 下的 GA-SVR 模型的预测效果最好, R^2 最大, RMSE、AAE、MAE 最小, 在训练集和测试集中分别为 0.79、0.86、107.51、2.95 和 0.35、1.38、47.38、3.95。该研究表明采用 GA-SVR with ε 模型优于其他模型, 能获得全局最小解, 具有最佳的泛化能力。

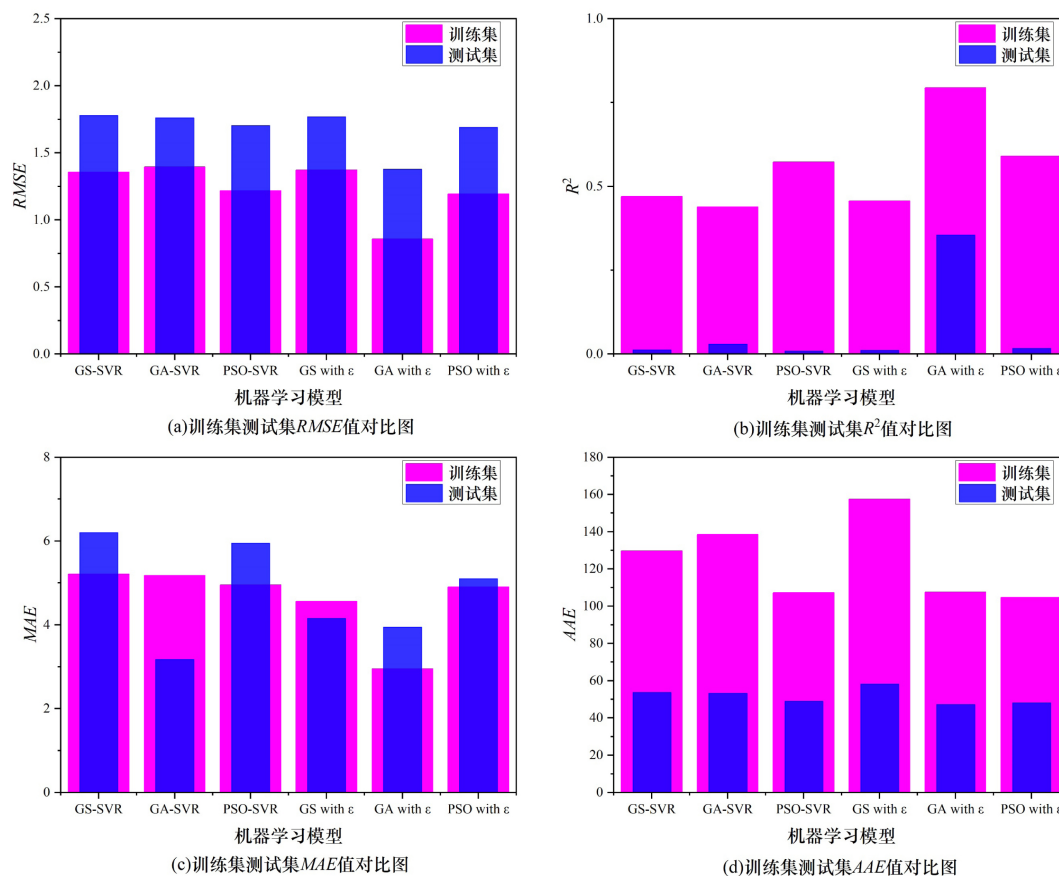


Figure 6. Performance comparison of different machine learning models based on conventional log
图 6. 基于常规测井建模的不同机器学习模型性能对比图

3.3. 基于常规测井与成像测井结合的模式建立

通过对全部测井参数进行不同的组合的寻优, 确定将 5 种测井参数作为输入参数, 孔隙度值作为输出参数, 训练集样本测井响应参数见表 2。图 7 为 6 种机器学习模型孔隙度实测值与预测值的交会图, 从图中可清楚的看到随着成像测井灰度值的引入预测效果有显著的提升, 不足的地方主要表现在 2%~3% 孔隙度区间, 表明训练模型对该区域的泛化能力较弱, 该区间对应的储层孔隙结构较复杂。图 8 可看出, GA-SVR with ϵ 模型性能最佳, R^2 最大, RMSE、AAE、MAE 最小, 在训练集和测试集中分别为 0.90、0.46、71.78、3.83 和 0.87、0.46、17.97、2.79。

Table 2. Log response parameter of training set samples
表 2. 训练集样本测井响应参数

样本号	深度/m	孔隙度/%	$DT/\mu\text{s}\cdot\text{ft}^{-1}$	slope of DT	GR/API	slope of GR	灰度值/0~256
1	5421.375	1.0	46.1	1.4	31.0	12.7	160.7
2	5421.75	2.0	47.5	<0.1	43.7	-2.0	131.9
3	5421.875	12.3	47.5	<0.1	41.7	-7.2	133.8
...							
141	5424.25	3.0	46.2	0.4	22.3	-6.1	151.1
142	5424.5	2.2	46.5	<0.1	16.2	-1.3	167.3

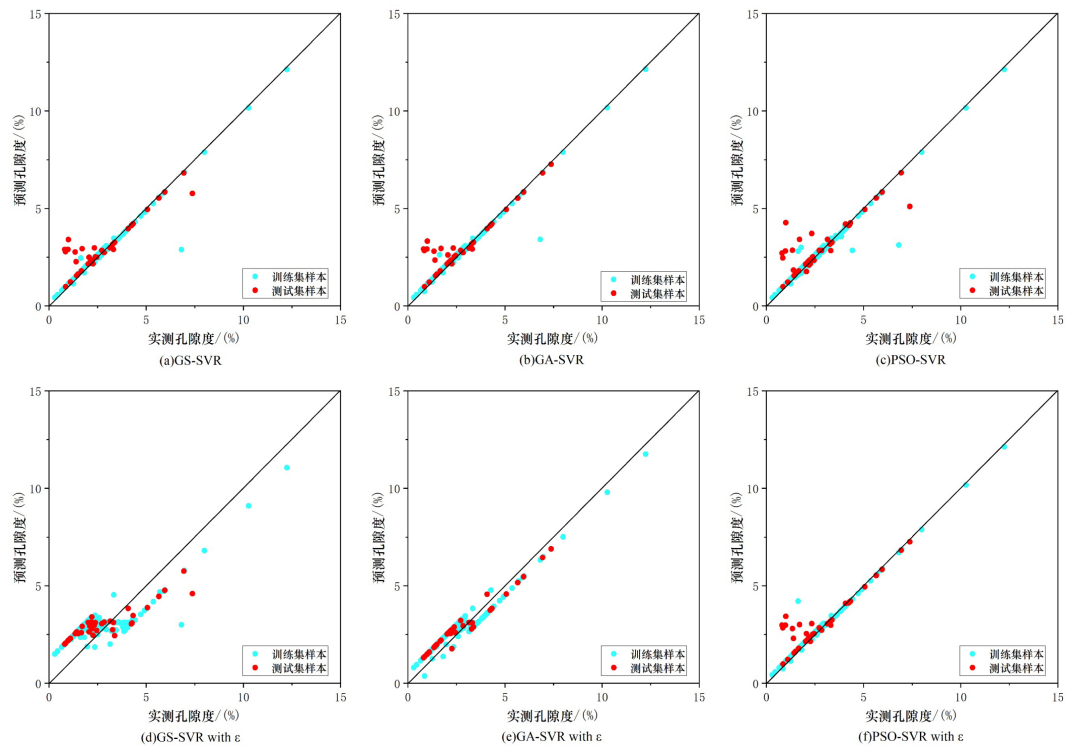


Figure 7. Cross-plot of the core porosity and porosity predicted by the various machine learning models
图 7. 不同机器学习模型实测孔隙度与预测孔隙度交会图

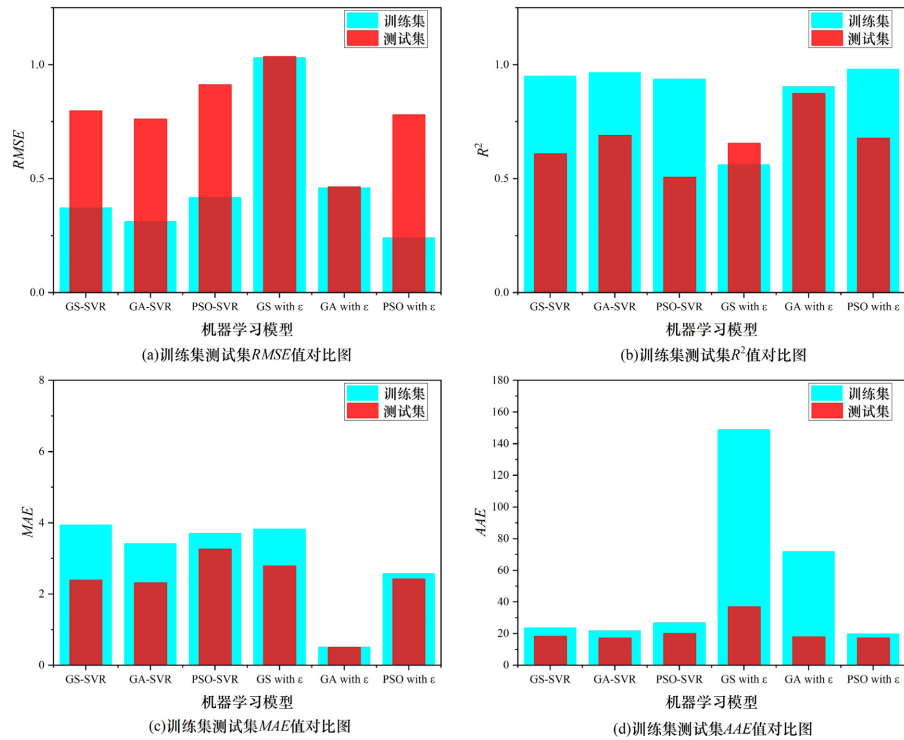


Figure 8. Performance comparison of different machine learning models with conventional logging and imaging logging combined
图 8. 常规测井与成像测井结合不同机器学习模型性能对比图

实验结果表明遗传算法 GA-SVR with ε 不仅在正确率方面优于其他算法, 实际操作过程中, 计算速度方面较网格搜索算法也有明显的优势。研究表明, 通过 6 种机器学习模型的对比, 在支持向量机模型算法中引入不敏感损失函数度量 ε 能提高其泛化能力, 针对样本较少的情况预测结果更佳精准, 模型更加稳定。研究表明通过引入成像测井图像灰度值作为训练参数, 能减轻白云岩储层孔隙度非均质性的不利影响, 也从侧面说明了在该区域中常规测井参数自然伽马(GR)与声波时差(DT)蕴含着白云岩储层孔隙度的绝大部分信息。将 GA-SVR with ε 机器学习模型与常规测井、成像测井相结合的方法比常规孔隙度线性或非线性预测方法更准确、可靠, 有利于非均质白云岩的优质储层的识别与评价。

3.4. 数据准备

由于不同地区白云岩原生孔隙、次生孔隙度发育情况不一, 沉积环境各异, 通过单口测井数据建模达到较高的精度并不足以证明该方法的可靠性。为此, 选取 M2 井 L 组实测样本数据验证本方法的应用效果。

采用 GA-SVR with ε 机器学习模型对 M2 井测井参数进行训练, 并对孔隙度进行预测。将数据分为两组用于对比本方法与常规测井预测方法的应用效果。

1) 第一组训练数据集由 M2 井自然伽马(GR)及其斜率、声波时差(DT)及其斜率, 成像测井灰度值等组成, 测试集为孔隙度值。

2) 第二组训练数据集由 M2 井常规测井参数自然伽马(GR)及其斜率、声波时差(DT)及其斜率等组成, 测试集为孔隙度值, 用于验证。

3.5. 孔隙度解释实例与效果分析

图 9 为 M2 井 L 组储层两种不同孔隙度预测方法对比。其中图(a)左 6 道为实测孔隙度, 左 7 道为常规测井方法预测孔隙度(蓝色离散曲线)与实测孔隙度(黄色散点)的组合图道, 其最大偏差达超过 50%。

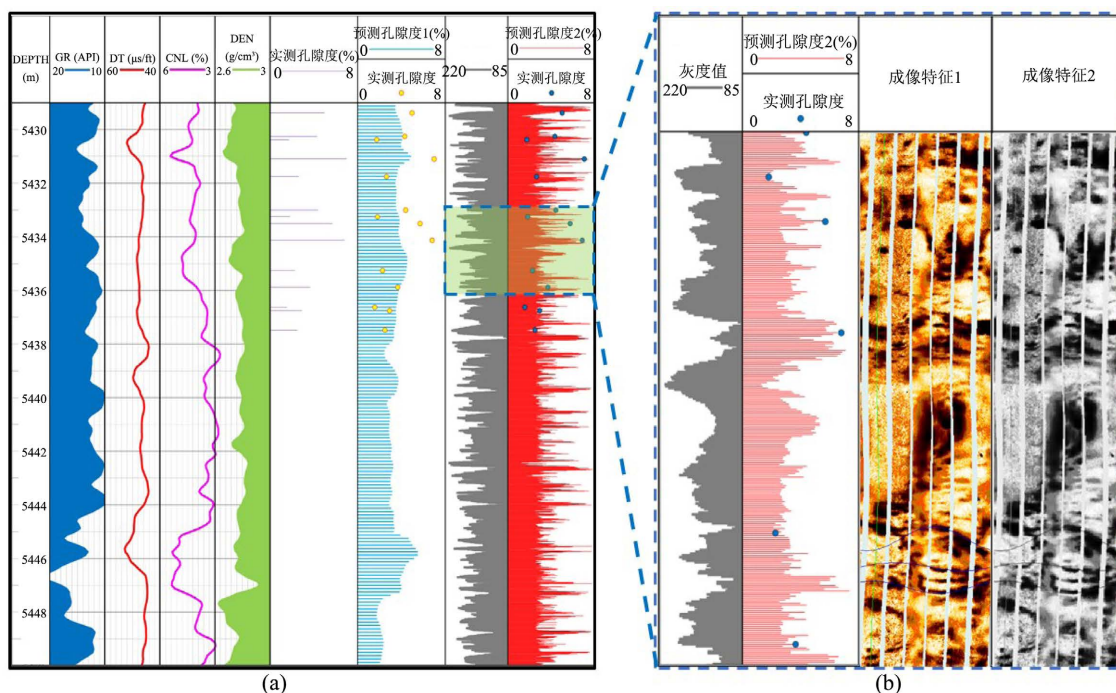


Figure 9. Porosity interpretation result map of L formation in well M2. (a) Results of single well; (b) Partial enlarged drawing
图 9. M2 井 L 组测井孔隙度解释成果图。(a) 单井成果图; (b) 局部放大图

左 8 道为通过 Python 程序提取的成像测井的灰度值, 右 1 道为采用本方法预测的孔隙度(红色离散曲线)与实测孔隙度(蓝色散点)的组合图道, 最大偏差小于 4%。图(b)中成像特征 2 道为成像测井灰度图像, 成像特征 1 道为成像测井原始图像。通过对比可以看出, 预测结果与实际取心资料、单井产量成像特征等钻探信息一致, 说明本方法针对该研究区孔隙度预测具有较高的可信度。在 5435 m 左右深度, 灰度值异常下降、孔隙度分布普遍偏高, 这是由于成像测井图像分辨率较低, 采集过程中存在一定噪声, 出现局部异常值, 不影响对白云岩优质储层的识别。

4. 结论

1) 川西地区 L 组白云岩储层的非均质性给储层评价和预测带来了很大的困难。本文结合常规测井和成像测井, 运用了 6 种机器学习模型, 确定了一种可靠的孔隙度解释方法。

2) 针对研究区白云岩的非均质性, 在支持向量回归算法中引入不敏感损失函数度量 ϵ 能有效地提高模型的泛化能力, 其中 GA-SVR with ϵ 模型性能最佳, R^2 最大, RMSE、AAE、MAE 最小, R^2 在训练集和测试集中分别为 0.90 和 0.87, 用该模型对孔隙度进行预测较可靠, 较精准。

3) 将机器学习模型应用于常规测井与成像测井相结合的储层参数预测方法, 明显优于常规方法, 有利于研究区白云岩储层孔隙度的精细解释与识别。

参考文献

- [1] 司马立强, 疏壮志. 碳酸盐岩储层测井评价方法及应用[M]. 北京: 石油工业出版社, 2009.
- [2] Malik, A.S., Boyko, O., Atkar, N. and Young, W.F. (2001) A Comparative Study of MR Imaging Profile of Titanium Pedicle Screws. *Acta Radiologica*, **42**, 291-293. <https://doi.org/10.1080/028418501127346846>
- [3] Zhong, Z. and Carr, T.R. (2019) Application of a New Hybrid Particle Swarm Optimization-Mixed Kernels Function-Based Support Vector Machine Model for Reservoir Porosity Prediction: A Case Study in Jacksonburg-Stringtown Oil Field, West Virginia, USA. *Interpretation*, **7**, T97-T112. <https://doi.org/10.1190/INT-2018-0093.1>
- [4] Li, Y., Kang, Z.J., Xue, Z.J. and Zheng, S. (2018) Theories and Practices of Carbonate Reservoirs Development in China. *Petroleum Exploration and Development Online*, **45**, 712-722. [https://doi.org/10.1016/S1876-3804\(18\)30074-0](https://doi.org/10.1016/S1876-3804(18)30074-0)
- [5] 杨柳青, 陈伟, 查蓓. 利用卷积神经网络对储层孔隙度的预测研究与应用[J]. 地球物理学进展, 2019, 34(4): 1548-1555.
- [6] 连承波, 李汉林, 渠芳, 蔡福龙, 张军. 基于测井资料的 BP 神经网络模型在孔隙度定量预测中的应用[J]. 天然气地球科学, 2006(3): 382-384.
- [7] 冯帅, 王连君, 邢雅文, 王娅妮, 周斌, 杨洲鹏. 二连盆地乌兰花凹陷安山岩有利储层预测[J]. 石油地球物理勘探, 2017, 52(S1): 117-122.
- [8] Mathew Nkurlu, B., Shen, C., Asante-Okyere, S., Mulashani, A.K., Chungu, J. and Wang, L. (2020) Prediction of Permeability Using Group Method of Data Handling (GMDH) Neural Network from Well Log Data. *Energies (Basel)*, **13**, Article No. 551. <https://doi.org/10.3390/en13030551>
- [9] 乐友喜, 袁全社. 支持向量机方法在储层预测中的应用[J]. 石油物探, 2005, 44(4): 388-392.
- [10] 陈义祥, 任小锋, 牟瑜, 陈惠, 俞保财, 姚海林, 等. 基于电成像测井的致密碳酸盐岩储层有效性评价方法[J]. 测井技术, 2020, 44(1): 49-54.
- [11] 李晓辉, 周彦球, 缙艳红, 王玉华, 姜宝彦. 电成像测井孔隙度分析技术及其在碳酸盐岩储层产能预测中的应用[J]. 吉林大学学报(地球科学版), 2012, 42(4): 928-934.
- [12] 平海涛, 秦瑞宝, 孙建孟, 刘小梅, 黄涛. 电成像测井孔隙度谱分析及其在复杂储层渗透率评价中的应用研究[J/OL]. 地球物理学进展, 2020: 1-9. <http://kns.cnki.net/kcms/detail/11.2982.P.20201109.1408.122.html>, 2020-11-09.
- [13] 刘瑞林, 谢芳, 肖承文, 张丽莉. 基于小波变换图像分割技术的电成像测井资料裂缝、孔洞面孔率提取方法[J]. 地球物理学报, 2017, 60(12): 4945-4955.
- [14] 景建恩, 梅忠武, 李舟波. 塔河油田碳酸盐岩缝洞型储层的测井识别与评价方法研究[J]. 地球物理学进展, 2003, 18(2): 336-341.