

DOI: 10.3969/j.issn.1672-7703.2021.02.012

基于大数据的测井智能解释系统开发与应用

石玉江¹ 刘国强² 钟吉彬³ 王娟³ 张文静³

(1 中国石油长庆油田公司数字化与信息管理部; 2 中国石油勘探与生产分公司; 3 中国石油长庆油田公司勘探开发研究院)

摘要: 人工智能技术在测井解释中的应用由来已久, 但单方法应用研究多, 系统集成应用少。当今云计算、大数据和人工智能技术的规模应用, 再次推动了测井解释智能化发展。以多学科数据融合的数据湖为基础, 在专业软件的基础上进行智能解释模块开发, 将测井智能解释引入到传统的解释流程中, 辅助测井分析人员快速挖掘隐藏的高价值信息。利用大数据治理工具将数据湖连接到智能模型, 实现大数据与智能算法融合。通过系统开发实现基于大数据的智能解释模型和传统专业软件的集成, 搭建基于地质图件导航的测井智能解释环境, 形成地质、油藏背景下的智能化测井解释工作模式, 增强测井评价复杂储层的能力, 提高工作效率。

关键词: 大数据; 智能算法; 测井解释; 智能解释环境

中图分类号: P631.9

文献标识码: A

Development and application of intelligent logging interpretation system based on big data

Shi Yujiang¹, Liu Guoqiang², Zhong Jibin³, Wang Juan³, Zhang Wenjing³

(1 Digitalization and Information Management Department, PetroChina Changqing Oilfield Company; 2 PetroChina Exploration & Production Company; 3 Research Institute of Exploration and Development, PetroChina Changqing Oilfield Company)

Abstract: Artificial intelligence technology has been used in well logging interpretation for a long time, but there are more researches on single interpretation method and less on application of system integration. Nowadays, the large-scale application of cloud computing, big data and artificial intelligence technology has promoted the intelligent development of logging interpretation once again. Based on the data lake of multidisciplinary data fusion, the intelligent interpretation module is developed in the logging interpretation software. That is, the intelligent logging interpretation method is introduced into the traditional interpretation process to assist logging analysts to quickly mine the hidden high-value information. Big data governance tools are used to connect the data lake to the intelligent model, so that the big data and intelligent algorithm are integrated. The system integrates intelligent interpretation model based on big data with traditional professional software and builds an intelligent logging interpretation environment based on geological map navigation to realize the intelligent logging interpretation working mode under the background of geology and reservoir integration, strengthen the ability of complex reservoir evaluation, and improve the working efficiency.

Key words: big data, intelligent algorithm, logging interpretation, intelligent logging interpretation environment

引用: 石玉江, 刘国强, 钟吉彬, 等. 基于大数据的测井智能解释系统开发与应用[J]. 中国石油勘探, 2021,26(2):113-126.
Shi Yujiang, Liu Guoqiang, Zhong Jibin, et al. Development and application of intelligent logging interpretation system based on big data[J]. China Petroleum Exploration, 2021,26(2):113-126.

基金项目: 中国石油天然气股份有限公司投资信息化重点项目“勘探开发一体化协同研究及应用平台(一期)建设”(PetroChina-IT-2017-N104)。

第一作者简介: 石玉江(1971-), 男, 甘肃平凉人, 博士, 2012年毕业于西北大学, 教授级高级工程师, 主要从事测井技术应用、地质综合研究和数字化与信息管理工作。地址: 陕西省西安市长庆兴隆园小区长庆油田公司数字化与信息管理部, 邮政编码: 710021。E-mail: syj_cq@petrochina.com.cn

收稿日期: 2020-09-08; 修改日期: 2021-02-08

0 引言

人工智能技术作为第四次工业革命的重要推动力,是大数据、算法和算力的重要结合,给各行各业带来了革命性的变革,也为测井技术的发展提供了强大的技术支持。测井技术的应用深度与广度主要取决于资料解释的信息丰度,其中的重要环节依赖于解释方法与软件的开发^[1-3]。目前,国内外测井解释软件功能齐全、种类繁多,促进了解释技术进步,提高了测井解释水平,解决了一批测井解释评价的难题。但是随着勘探开发目标的日益复杂化和隐蔽化,现有的解释方法和软件难以准确描述储层的复杂状态^[4]。测井对储层的非线性响应,导致处理解释质量不高,要准确研究非均匀地层测井信息所蕴含的地质信息,需要研发新的测井解释方法^[5]。人工智能作为一种改进计算机求解问题的方法^[1],可以弥补传统解释方法的不足,使测井分析人员摆脱大量低知识层次的分析工作。将人工的知识和经验系统化、条理化形成“智能助手”,帮助测井分析人员解决更深层次的地质问题。

为此,国内外进行了大量的实践。斯伦贝谢公司推出的 DELFI 勘探开发认知环境,将复杂的数值模拟、数据分析等变得智能和快捷。国内也在积极探索适合中国特色的测井智能解释道路,这给传统测井解释技术发展带来了新动能和良好的发展前景^[6-8]。20 世纪,中国测井科技工作者在人工智能单方法应用和专家系统开发等方面都取得了重要的研究成果^[9-22]。然而,这些成果并没有及时地转化为生产力。究其原因,一是缺少足量的专业大数据作为智能计算的基础;二是缺少符合测井智能解释业务的集成应用环境。

前人取得的科研成果为测井智能解释系统的开发和应用奠定了基础。按照测井解释软件走集成化道路^[3]的思路,长庆油田通过系统开发和应用环境搭建,实现了大数据、人工智能和测井解释软件的集成应用。将“数据”“算法”“场景”3 个关键要素有机结合,形成了基于大数据的测井智能解释环境,提高了测井参数计算精度和流体识别准确率,提升了测井分析效率,促进了测井解释由单井解释向多井评价、油藏分析转变。

1 总体思路和系统架构

1.1 总体思路

测井解释软件是测井资料处理与解释的重要工

具和载体,是运用各种现代技术综合解决地质问题的重要手段,是测井解释技术能力的综合体现,也是衡量测井技术水平的一个重要指标^[8]。斯伦贝谢公司在 Techlog 解释软件中开发了主成分分析、回归、聚类、决策树、神经网络等智能功能模块。这些模型的训练样本主要是软件的工区数据,一方面数据量不足,另一方面针对算法的数据预处理较少,导致应用受限。DELFI 系统建立了跨专业协同工作环境,采用开放式的软件开发模式,通过 Studio 集中管理 Techlog 与 Petrel、ECLIPS 等地质、油藏软件的数据,实现了地质、工程业务智能化与协同化。然而,涉及国家地质数据的安全问题,中国不能像引进专业软件一样引进 DELFI 系统,只能走自主研发的道路。

长庆油田经过 10 余年的数字化油田建设,建成了数字化油气藏研究系统(RDMS)^[6],盘活了 15 万余口井的地质、钻井、测井、录井、分析化验、试油(气)等多学科数据。RDMS 实现了专业数据自由推送、实时关联、在线可视等基础功能;开发的基于地质图件的数据导航系统(CQGIS)实现了多学科数据自动收集、整理,以及软件的无缝衔接^[23-25],这些建设成果逐渐成为大数据技术应用的基础。针对测井软件开发的数据接口打通了数据湖的通道,通过以专业软件为基础开发大数据治理工具,实现样本数据快速提取^[26-30]。在专业软件中开发智能解释模块,将智能应用引入到传统测井解释流程中,实现专业软件、智能应用和数据湖的耦合集成。在 RDMS、CQGIS 和解释软件之间,搭建起以地质、油藏为约束的测井智能解释评价环境。

1.2 系统架构

系统自下而上由数据层、功能层和应用层组成,从功能上分为样本数据治理、数据综合分析、模型训练与优选、模型评估与管理,应用上分为建模应用和极简应用(图 1)。在数据层开发数据解编模块处理测井数据体,实现测井数据与试油、岩心分析、录井等资料的数据模型和数据结构统一。数据治理工具在深度域上进行数据治理,完成连续数据和离散数据的匹配。解释人员通过专业软件选择储层特征数据,并依据试油结论给特征数据打上标签,将这些数据按层位存入样本数据库。进入样本库的数据可直接用于模型训练与优化,形成用于测井参数计算和流体识别的智能模型,通过模型优选,组合成针对性推荐智能模型组,供用户调用。

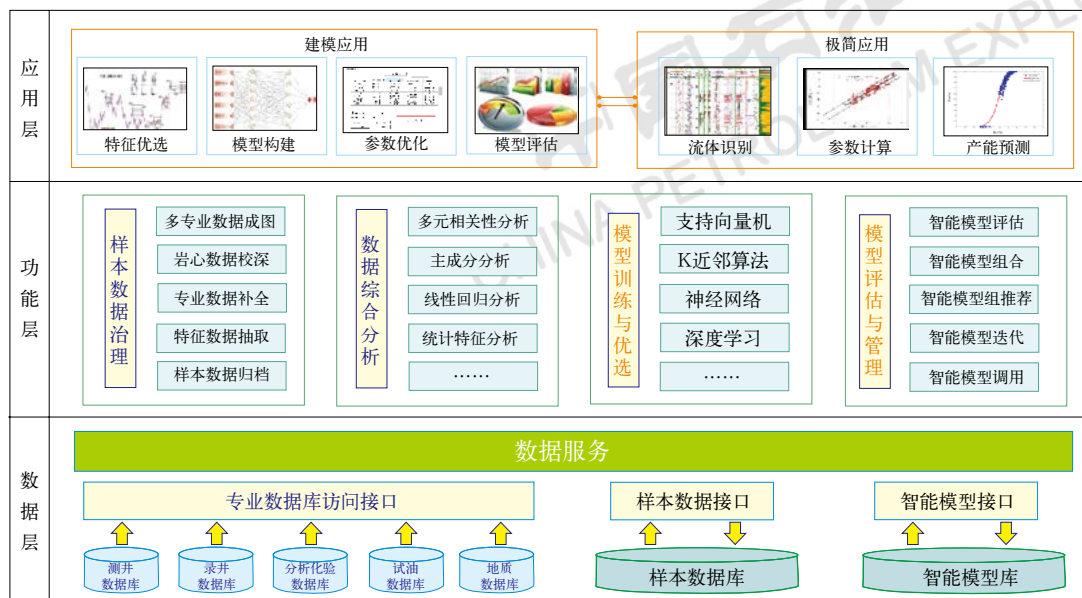


图1 基于大数据的测井智能解释系统架构图

Fig.1 Architecture of intelligent logging interpretation system based on big data

系统应用分为极简应用和建模应用两种模式。极简应用模式面向普通用户，普通用户不用关心模型训练和参数优化等工作，根据系统推荐的最优模型组合，开展单井或多井的专业解释工作，在智能系统的辅助下专注于参数计算、流体识别及产能预测工作。建模应用模式面向专家用户，主要负责特征优选、模型构建、参数优化、模型评估工作，是系统正常运行的基础。

在此系统架构下，智能系统运行的业务流程大致分为：①依托 RDMS 数据流进行样本数据治理，建设机器学习样本库；②根据智能算法特点，通过数据综合分析工具在样本库中进行特征优选，并进行预处理，形成模型训练的特征矩阵；③通过超参数设置构建模型，并进行模型迭代训练，优化参数，对模型进行评估；④组合单一功能的智能模型，形成智能模型组；⑤用户通过极简应用模式调用智能模型组进行测井解释评价。其中①~④在建模应用模式下进行，⑤在极简应用模式下进行。

2 测井大数据治理

在专业解释软件的基础上开发大数据治理工具，实时绘制测井专业图件，对多专业数据进行快速集成和直观展示，对数据进行补充、融合、抽取等治理。让专家用户专注于储层特征的分析、选取和打标签工作。

通过开发数据接口，在数据湖中自动收集并基本信息（如坐标、钻井液、钻头等信息）、地质分层、试油、录井、分析物性等数据，并自动转换为数据治理工具的数据格式，批量推送到治理工具；专业软件调用测井绘图模板直接成图，并可以自由补充其他数据，完成特征匹配和数据检查后，选择储层特征值，打上相应标签（如油层、油水同层、差油层、含油水层、干层、泥岩、砂岩、煤、隔夹层等）。最后，一键式自动提取目标井基本信息、单点数据、特征值、试油数据、岩心分析数据等，并批量提交入库。以上过程构成了测井大数据治理流程（图2）。

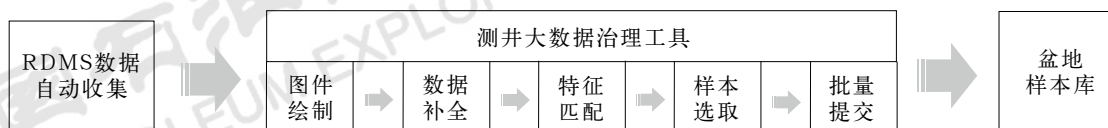


图2 大数据治理流程图

Fig.2 Diagram of big data governance process

选取样本数据应当遵循以下原则：①尽可能选择具有单层试油数据的储层提取样本数据；②以一个压裂试油单元为最小单位选取特征值，一个单元可以包含若干特征值，每个特征值的标签可以不同；③以储量计算标准为依据确定标签属性，尤其要仔细鉴定油水同层和含水层；④在一个试油单元中尽可能多地选取不同标签的数据，除了选取有效储层外，还应附带上下盖层及隔夹层的数据；⑤样本库数据要尽量保证各标签样本数据相对平衡。

按照地质分层，样本库进行分层管理数据，利用“分区分层，分而治之”的数据治理策略，逐步建立盆地级的机器学习样本库。数据治理和样本库建设是一个长期的过程，贯穿于智能系统应用的整个生命周期。后期或可采取以智能促智能的思路加快样本数据库建设。

3 智能模型训练和管理

经过长期的研究和实践，前人^[9-22]将有监督学习、无监督学习、强化学习等算法纷纷引入到测井解释中，形成了各具特色的应用技术。由于样本数据量较少，数据缺乏精细的预处理，导致模型泛化能力较低，距工业化应用存在一定距离。本次系统开发将这些成果进行了优化集成，开发了“模型参数自动寻

优”“模型组合优化推荐”“专业图件在线绘制”“应用效果实时展示”等功能模块，实现了智能算法集成应用。

3.1 智能模型训练

结合测井解释业务特点，对支持向量机、K-近邻、神经网络等算法进行优化，使智能算法与传统解释理念融合。通过集成TensorFlow深度神经网络算法^[31-32]，保证系统功能可扩展。以上算法的核心原理及训练方法成熟可靠，在系统开发过程中主要在数据预处理和模型自动训练方面做了大量工作，下面以支持向量机识别复杂油水层为例进行简要介绍。

3.1.1 智能算法与测井解释方法的结合

研究认为，支持向量机分类算法思想与测井常规油水识别图版基本思想是一致的，目标是寻找不同类别数据的边界。支持向量机学习不同类别数据之间的决定边界，通常只有位于边界的部分数据点起到重要作用，这些点被称为“支持向量”（图3a）^[33]。在传统的交会图解释方法中，利用声波时差与电阻率交会图（图3b）确定油层类的电性下限值，然后把各下限值用直线连接起来作为油水界限（图3b）。二者的基本思想都是通过统计分析，确定出一个合理的决策边界^[34]，用于数据分类。

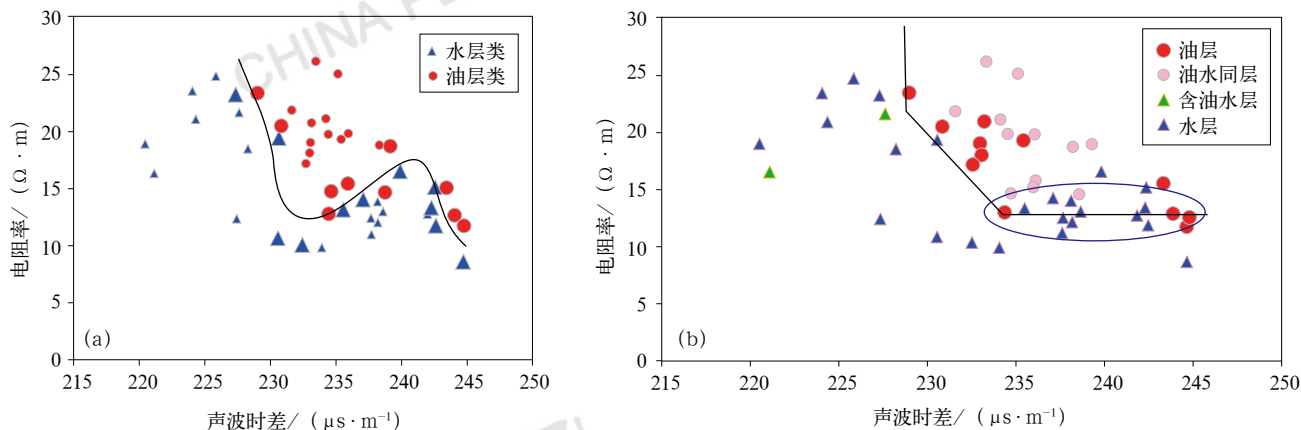


图3 支持向量机分类原理与常规油水识别图

Fig.3 Comparison of oil-water identification between support vector machine classification principle and conventional cross plot method

(a) 支持向量机原理示意图，图中尺寸较大的数据点是支持向量，黑色曲线是决策边界；(b) 声波时差与电阻率交会图。
(a) 中油层类数据点对应 (b) 中油层和油水同层数据点，(a) 中水层类数据点对应 (b) 中水层和含水层数据点

在复杂储层的测井解释中，常规油水识别图版适应性较差，在图版上表现为油水界限不清晰，如图3b 蓝圈数据点所示。油层类和水层类混杂在一起，常规解释方法无法确定有效的油水界限，导致图版解释符合率较低，主要原因是常规二维交会图仅使用了

电阻率和声波时差两个测井信息，未使用其他测井资料信息。为了充分挖掘这些信息，测井分析家通过近似条件假设和严格的数学模型推导^[1]，构建储层敏感参数，制作各种解释图。通过这种类似降维的方法，将更多的测井信息引入到二维图中，但仍然没有使用

全部测井信息。

在二维图中增加一个维度的测井信息, 带来的相关信息量成倍增加, 通常表示为三维流体识别图(图4a), 这种交会图在实际使用时非常不方便。即使在单个二维交会图中的油水界限都很不规则(图4b中黑色曲线), 如果再增加图4c、图4d所示的两个交会图的不规则油水边界, 油水决策边界将

变为三维空间的一个超平面。这个超平面无法直观展示, 应用难度大。支持向量机利用核函数可以将多维数据映射到非线性的高维空间, 通过超参数控制决策超平面的复杂程度^[33], 可以准确逼近这个超平面。支持向量机利用灵活的决策边界寻优技术, 在非线性高维空间中开展测井解释, 这给正确识别非常规复杂储层流体性质提供了有效的解决方案。

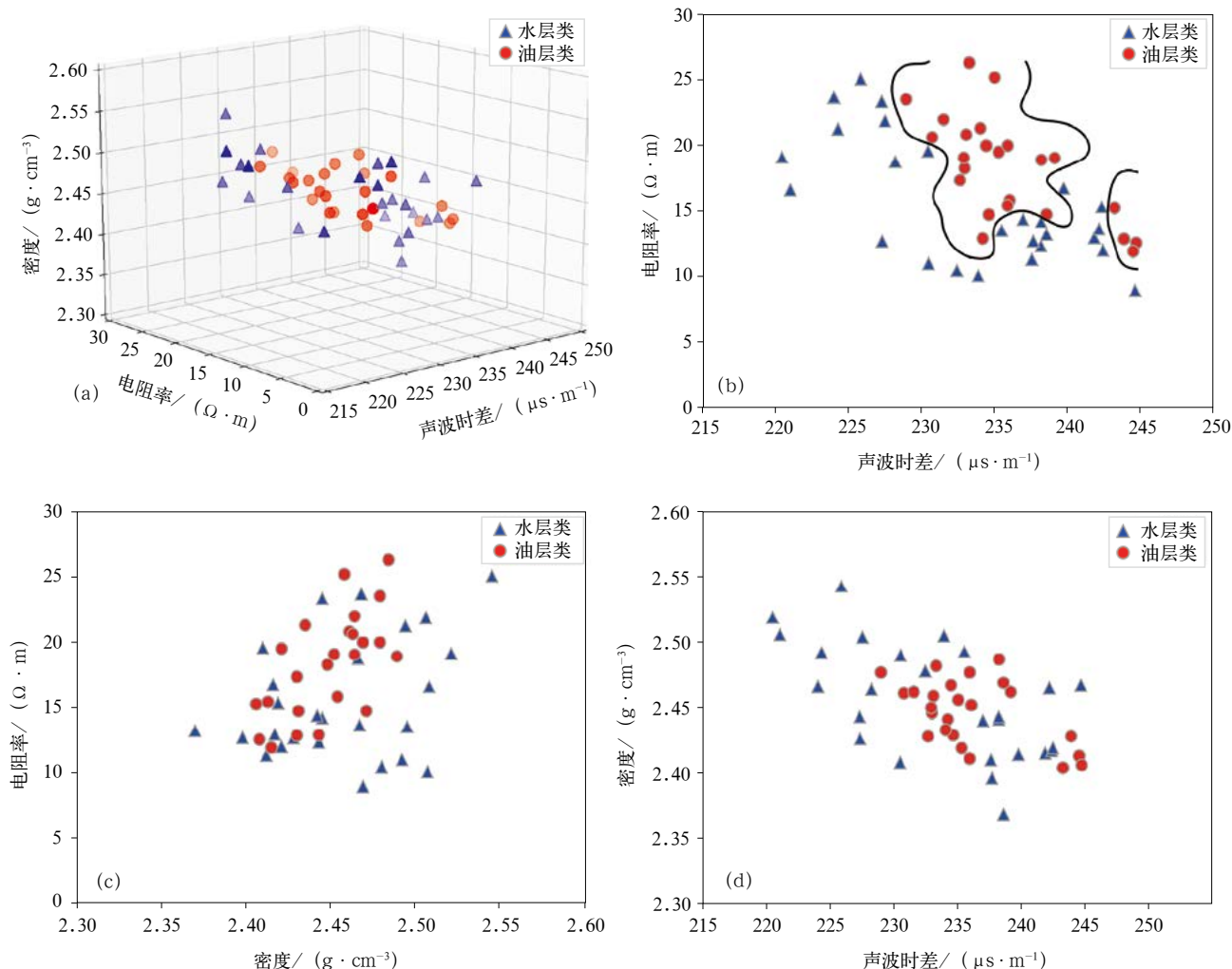


图4 流体识别图及其决策边界

Fig.4 Fluid identification chart and decision boundary

(a) 三维流体识别图; (b) 声波时差与电阻率交会图; (c) 密度与电阻率交会图; (d) 声波时差与密度交会图

3.1.2 样本数据预处理

目前, 大多数机器学习模型都采用梯度下降算法进行模型训练, 其中心思想是迭代调整参数使成本函数最小化^[34], 当成本函数达到全局最小, 得到的参数组合就是最佳拟合参数; 否则, 模型要么过拟合, 要么欠拟合, 泛化应用效果均不会太好。大多数机器学习算法对数据尺度非常敏感, 当输入模型的多维数

据尺度差异很大时, 会使梯度下降算法寻找成本函数的全局最小值变得异常困难。

用支持向量机进行分类时, 支持向量对决策边界起到了决定性作用。如图5a, 因为密度与电阻率数据数值差异较大, 支持向量之间的“距离”小, 分类效果非常差; 经过预处理后二者尺度变得接近(图5b), 分类效果明显改善。支持向量之间的“距离”被放大, 决策边界更容易求取^[33]。

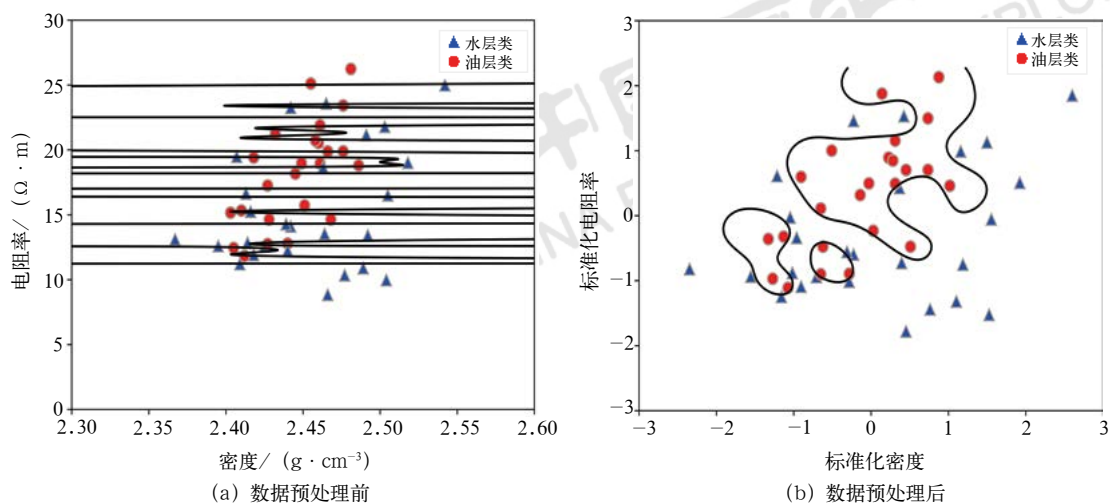


图5 数据预处理前后密度—电阻率交会图及决策边界

Fig.5 Comparison of density-resistivity cross plot and decision boundary before and after data preprocessing

众所周知,各测井曲线之间数据尺度差异大,数据分布形式也有较大不同。如图6中声波时差和自然伽马曲线数据尺度和分布形态都比较相近,类似于正态分布特征;而电阻率和密度曲线不但数据尺度相差大,分布形态也不相同。数据的这些特征都会影响最优决策边界的计算。依据机器学习特征工程要求,结合测井数据特点,对电阻率曲线进行对数变换后,与其他曲线同时采用Z分数标准化方法^[33] [公式(1)]进行

预处理,使数据尺度和分布形态一致(图7),这对于大多数机器学习算法是非常重要的。公式(1)中需要用到区域样本数据的平均值和标准差。由统计学大数定律^[35]可知,对于某条测井曲线,在同一地区或地层范围内,当样本数据足够多时,数据的标准差和平均值趋于稳定。因此,对于样本数据的预处理可以直接使用大样本数据的标准差和平均值,利用公式(1)直接计算,保证了数据预处理标准的一致性。

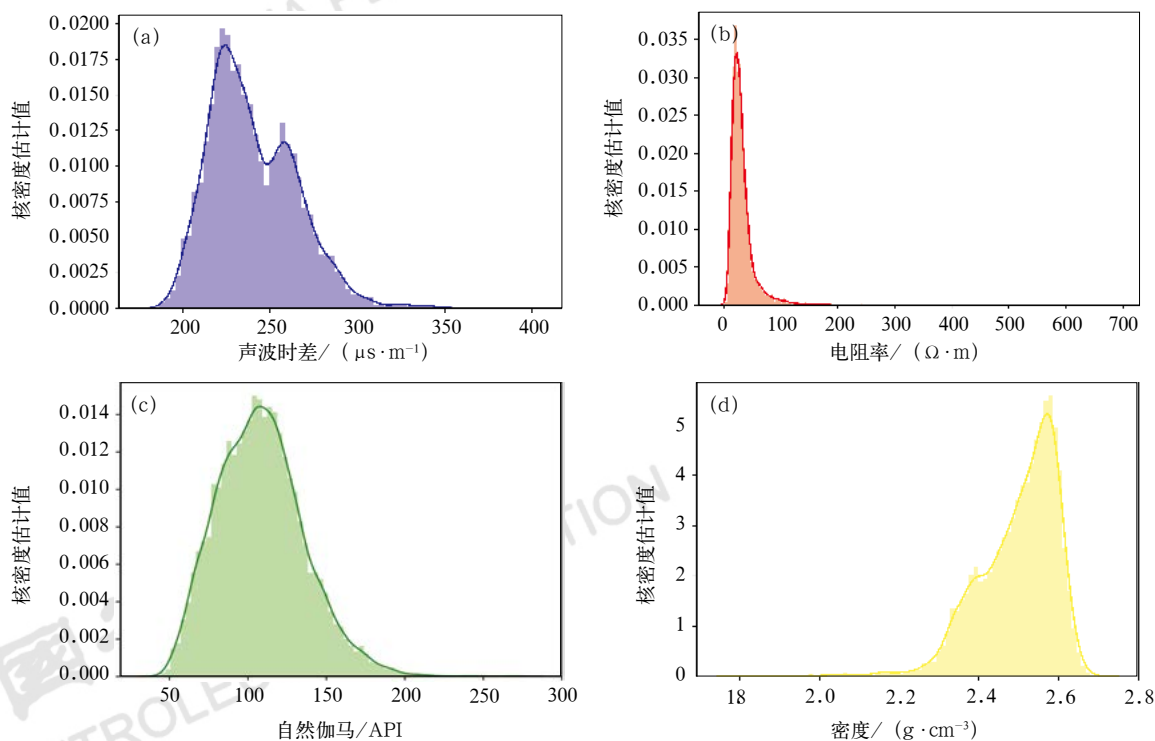


图6 预处理前测井数据概率密度图

Fig.6 Distribution of logging data before preprocessing

(a) 未预处理的声波时差; (b) 未预处理的电阻率; (c) 未预处理的自然伽马; (d) 未预处理的密度

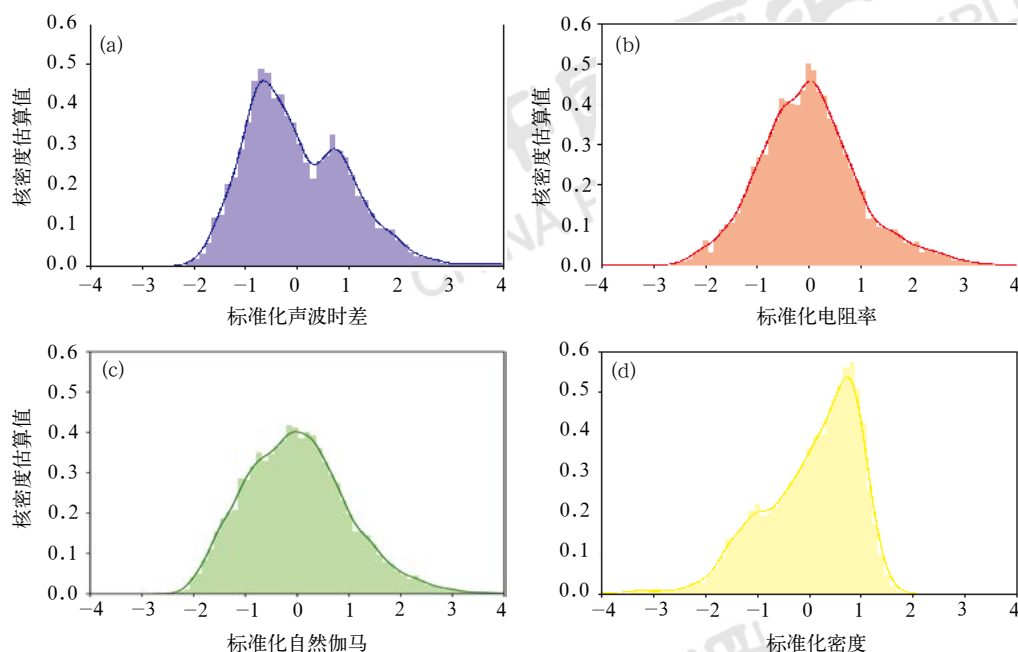


图7 预处理后测井数据概率密度图

Fig.7 Distribution of logging data after preprocessing

(a) 预处理后声波时差；(b) 预处理后电阻率；(c) 预处理后自然伽马；(d) 预处理后密度

$$z = (x - \mu) / \sigma \quad (1)$$

式中 z ——特征缩放后的特征值；

x ——特征缩放前的原始测井数据值；

μ ——数据的平均值；

σ ——数据的标准差。

通过以上数据预处理，既消除了数据分布形态的差异，又对数据进行了统一的标准化处理。通过专业软件成图与原始测井曲线对比可知，标准化处理之后的曲线保留了原始测井曲线的绝大部分信息（图8）。在保留地层特征信息的同时，又将适合人类分析的数据，转换到了适合机器学习的特征空间。

实践证明，数据预处理可以大幅提升支持向量机对流体识别的准确度。对10884个样本数据进行试验分析，预处理前回判符合率为83%，预处理后回判符合率为97%，回判符合率提高了14%。与支持向量机算法一样，K-近邻、K均值、逻辑回归、神经网络、主成分分析等算法都对数据尺度非常敏感^[33]，因此，输入机器学习模型的测井数据都需要进行预处理。

3.1.3 模型训练

智能模型训练的关键是超参数寻优。超参数的数量和性质由智能算法本身及训练算法、激活函数、目标函数等共同决定。通过人工调参的方式训练模型，

难度非常大。随着机器学习技术进步，带交叉验证的网格搜索法是目前智能模型超参寻优常用的一种方法。该方法降低了模型训练难度，使得智能模型训练逐渐实现了半自动化或自动化。

支持向量机用于分类时，常用的核函数为线性核、多项式核、拉普拉斯核、高斯径向基核等^[34]，以高斯径向基核函数为例，有正则化参数 C 和高斯核宽度 γ 。决策边界是基于其与支持向量之间的距离及支持向量重要性来计算^[31]。数据点之间的距离[公式(2)]由高斯核函数计算得到，其中正则化参数 C ，决定模型受限程度。当 C 值小说明模型非常受限，每个数据点的影响范围较小（图9a）；当 C 值增大时，数据点对模型的影响变大（图9b）。利用网格搜索法优选适当的超参数，将数据点分为不同的类型，决策边界将根据超参数不同而变得复杂（图9c、d）。模型训练就是设计适当的模型评估算法，优选出最优的超参数组合（图9c），避免机器学习模型欠拟合（图9a、b）或过拟合（图9d）。

$$k(x_1, x_2) = e^{-\gamma \|x_1 - x_2\|^2} \quad (2)$$

式中 x_1, x_2 ——数据点；

$\|x_1 - x_2\|$ ——欧氏距离；

γ ——控制高斯核宽度的超参数。

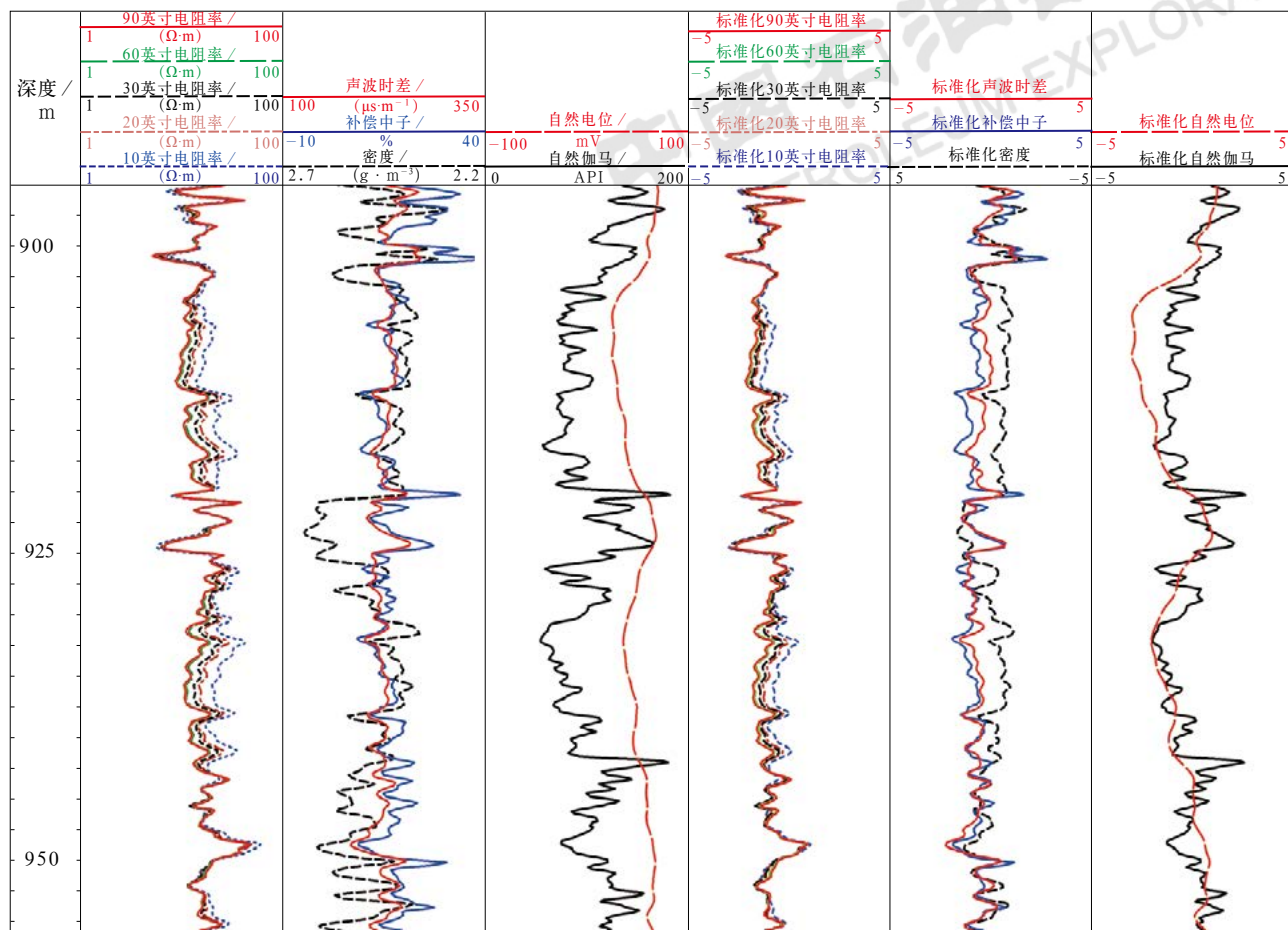


图8 测井曲线预处理效果图

Fig.8 Comparison diagram of logging curve before and after data preprocessing

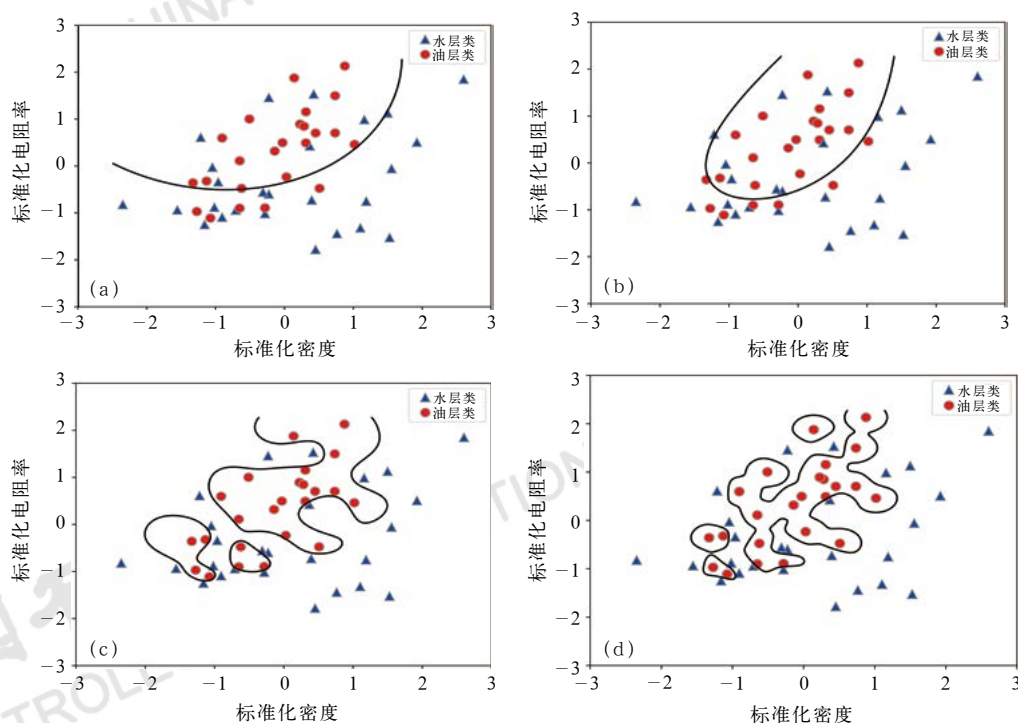


图9 支持向量机不同超参数对同一数据计算的决策边界对比图

Fig.9 Comparison of decision boundary of support vector machines with different super parameters for the same logging data

(a) 超参数设置为 $C=100$, $\gamma=0.01$; (b) 超参数设置为 $C=10000$, $\gamma=0.01$; (c) 超参数设置为 $C=100$, $\gamma=4$; (d) 超参数设置为 $C=10000$, $\gamma=8$

混淆矩阵通常用于判断模型精度和计算召回率等指标,但是这些参数对于测井分析人员来说并不容易理解。在上述评价的基础上,利用专业图件的形式直接展示计算成果,是符合专业习惯的。如图10中智能计算孔隙度采用神经网络计算得到,人工计算孔隙度利用密度曲线依据岩石物理模型计算得到。这两条

曲线与岩心分析孔隙度对比可知,神经网络计算结果与岩石物理模型计算结果相当,在细微处神经网络计算更加准确。同样,智能计算渗透率曲线采用K-近邻模型计算得到,与人工计算渗透率曲线相比更接近岩心分析渗透率。智能解释结论采用支持向量机识别,其油气结论与人工解释结论基本一致。

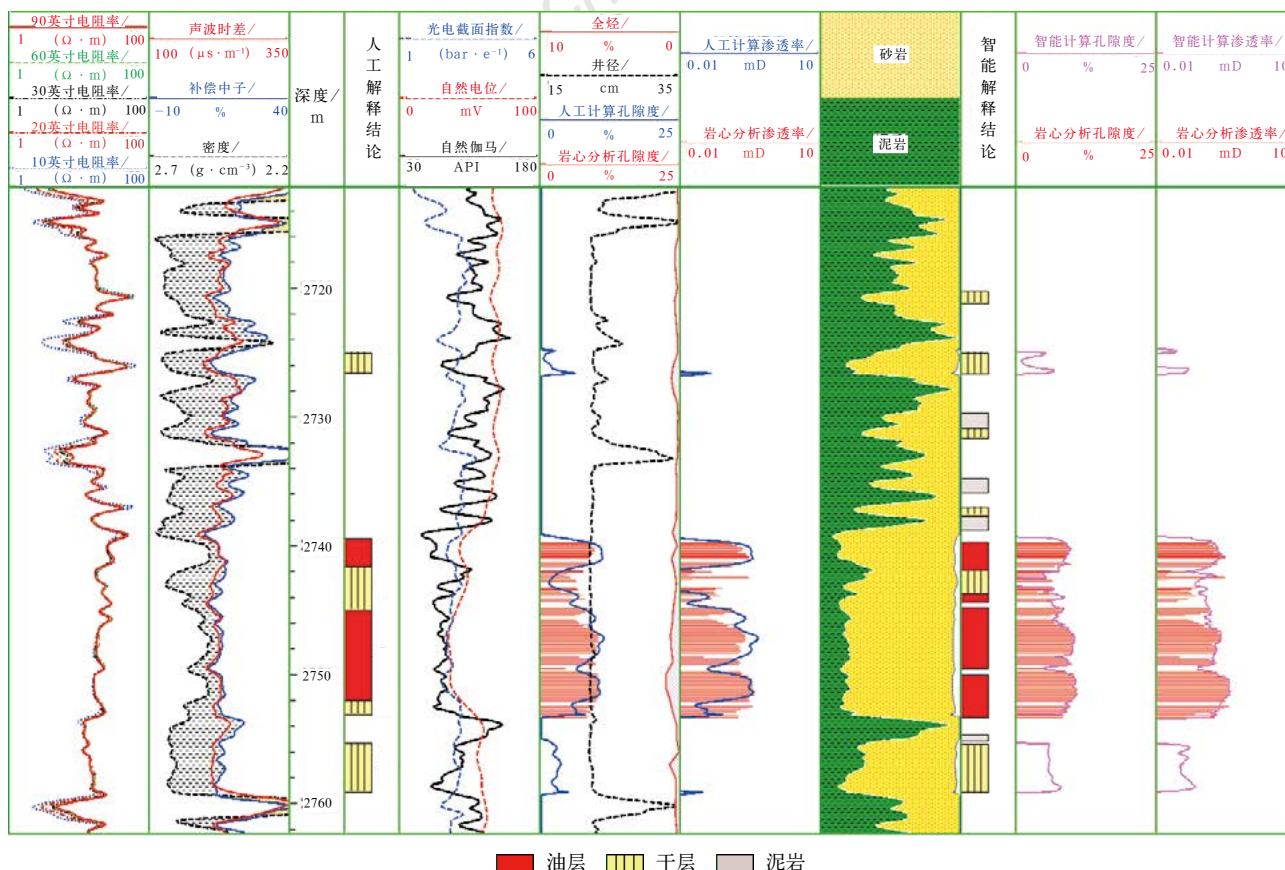


图10 智能解释结果与人工解释结果对比图

Fig.10 Comparison of intelligent and manual logging interpretation results

利用参数自动寻优算法,实现了支持向量机模型半自动训练,K-近邻全自动化训练。前馈神经网络实现了部分超参数自动优化,用户需要设定网络结构,循环次数由系统自动确定。深度神经网络实现了全自动化训练。智能模型对油水层判识准确率达到85%以上,与人工解释精度相当,孔隙度和渗透率计算精度高于人工计算精度,能够满足大多数测井解释评价工作的要求。

3.2 模型组合与自动推荐

系统集成的机器学习算法根据不同的目的,可以训练成不同的智能模型。比如支持向量机既可用于流体分类,也可用于回归计算。同样,神经网络既可

用于计算孔隙度,也可以计算渗透率或饱和度,甚至重构其他特征曲线。所以,在系统应用过程中必定会形成大量功能单一的模型,这给模型应用造成一定的困难。

模型以一定的规则命名并保存在智能模型库中,实现统一管理。如孔隙度计算模型“2019_石油_姬塬_长₆_孔隙度_BPNN_GPADCA9-2_物_(9-10-10-1)_1.45”表示的意义是:2019年建立的针对姬塬地区长₆石油储层的孔隙度计算模型,采用的算法是BP神经网络,输入数据为岩心分析的物性数据和常规测井曲线GR、PE、AC、DEN、CNL,以及AT90、AT60、AT30、AT20,神经网络结构是输入层9个特征向量,输出层1个目标向量,中间包含

两个隐藏层,每个隐藏层具有10个神经元,模型的计算误差是1.45%。在模型库中还保存了训练模型时使用的样本数据索引及超参数、神经元阈值及权重等其他参数。这些数据保证了该模型可以随时调用,并为后期实现模型迁移学习提供基础。当样本库数据更新后,系统可以调出智能模型重新训练、评估,实现模型循环迭代更新。

以上训练完成的智能模型功能单一,只能完成孔隙度、渗透率、饱和度计算和流体识别中的一个。要实现智能模型极简应用,就要对单个模型进行优化、组合。需要专家用户在构建智能解释模型后,针对区域地质特征优选性能最佳的模型进行组合,按一定规则命名,推荐给用户使用。如“2019_石油_姬塬_长6_孔渗饱流体组合模型_V1”表示的意义是:该模型组合为2019年建立的用于姬塬长6地层石油孔隙度、渗透率、饱和度计算和油水识别的模型组合,版本号为V1。

以上功能模块与长庆油田CQGIS地质图件导航系统结合,在极简应用模式下,系统获取到目标井所属区块和目标层位,自动推荐最佳模型组,实现一键式在线智能解释;通过专业软件将智能解释成果和原始测井曲线绘制成专业图件,并且在同一环境下,用户可以继续开展后续深入的测井、地质研究工作。

4 测井智能解释应用场景

智能系统要在传统石油行业中发挥作用,必须与专业应用场景结合。通过数据接口软件将数据治理工具与区域数据湖连接起来,可以自由调取海量的专业数据。开发应用接口将CQGIS地质图件导航系统与测井智能解释软件连接起来,支撑起极简应用和建模应用,形成具有长庆特色的“RDMS+智能计算”的测井智能解释环境。

4.1 大数据治理场景

RDMS是长庆油田数字化建设的重要成果,利用数据接口连接数据湖、测井大数据治理工具和样本库,形成了基于大数据治理的数据流,解决了样本库中数据来源问题。

在这个环境下,测井分析人员不仅可以开展数据治理工作,还能够对样本数据进行质量评估。建立样本库需要对样本井进行筛选,对专业数据进行甄别,获取样本“金数据”。用户只需要批量导入井名,系

统将自动收集数据并完成格式转换,替用户完成数据加载等外围工作。在多井对比环境下,依据试油数据对储层流体性质进行重新认识,更正错误的解释结论;确认岩心分析的孔隙度、渗透率、饱和度等数据与测井曲线特征匹配;校正录井资料与测井、试油数据的响应特征,标识出每口井目的层位的测井特征值,分层位批量提交样本数据。同时,这些数据保存在本地计算机工区中,可以随时调整并同步更新到样本库。通过样本数据质量分析,形成各类统计图表(图6)进行主成分分析、多参数相关性分析等。

大数据治理场景将数据治理从智能模型训练中分离出来,可以大幅提升数据治理速度,快速形成盆地级样本库,为智能应用提供高质量的学习数据。

4.2 测井智能解释应用

智能解释应用场景是系统应用的核心。在极简应用模式下,用户利用地质图件导航,确定目标井和目的层位,系统根据这些信息自动推荐智能模型组合,并启动智能解释软件模块,同步完成孔隙度、渗透率、饱和度参数计算、流体性质自动识别和产能预测,实时在线成图,生成储量参数图表。

在建模应用模式下,系统根据地质图件的层位信息,将样本库中对应层位的样本井点显示在地质图件上。用户根据目标区域地质、油藏概况,在图件上选择一组样本井作为学习样本(这一组样本井可视为盆地样本数据库的一个子集),并将这些信息推送到建模工具,选择相应算法进行模型构建、训练、评估和参数优化等工作。当性能满足应用要求后,按照模型管理规则存入模型库。当目标区域内的孔隙度、渗透率、饱和度、流体性质、产能预测等计算模型都训练完成后,将这些模型组合成智能模型组推荐并发布应用。此后,极简应用用户就可以直接使用这个模型组合开展测井解释工作。

将模型建模和应用分离,降低了智能模型应用难度。专家用户需要具有较强的测井解释评价能力,熟悉智能模型训练和评价方法,主要任务是建立和维护智能模型,包括模型应用效果评价,如果泛化能力降低,则需要将模型进行循环迭代训练,直到满足新的应用需求。极简用户只需熟悉应用环境和使用流程即可。他们的主要任务是使用“智能助手”辅助他们完成相应研究工作,如果发现模型精度降低,可将问题通过系统记录到模型库中,也可将某些井作为样本井

添加到样本库中。专家用户根据这些信息完善样本数据，进行模型迭代优化，使整个智能系统得到循环学习、迭代提升。

5 应用效果

测井智能解释系统的成功开发和落地应用，强化了地质、油藏对测井评价的指导作用，基本实现了测井解释的自动化、智能化。该系统在参数计算和流体识别等方面的精度都达到或超过了人工解释，智能评价结果与试油、投产等数据更加匹配，初步具备了辅助人工进行流体识别和油藏评价的能力，实现了多井批量处理，大幅提高了测井解释的工作效率。

H369 井是鄂尔多斯盆地西部的一口老井，其地理位置远离样本井分布范围，且在 H369 井周围仅有

3 口邻井。该井 2631.5~2634.6m 人工解释结论是含油水区（如图 11）。岩屑录井描述为 1.3m 油迹和 0.5m 油斑。对该含油水区进行射孔，试油日产 21.5t 油和 56.1m³ 水，试油结论为油水同层，人工解释出现了较大失误。

在极简应用模式下，自动处理得到的智能解释结果为油水同层（图 11），与人工解释结论截然相反，但是与试油结果一致。智能计算孔隙度、渗透率和饱和度与人工计算结果接近，与岩心分析物性数据匹配较好，符合岩石物理规律。该系统利用支持向量机识别流体性质，神经网络计算孔隙度，K-近邻模型计算渗透率和饱和度，这些模型组合成一个智能模型组，同时完成 4 个方面的工作，不但提高了测井解释效率，计算结果也满足测井评价需要。

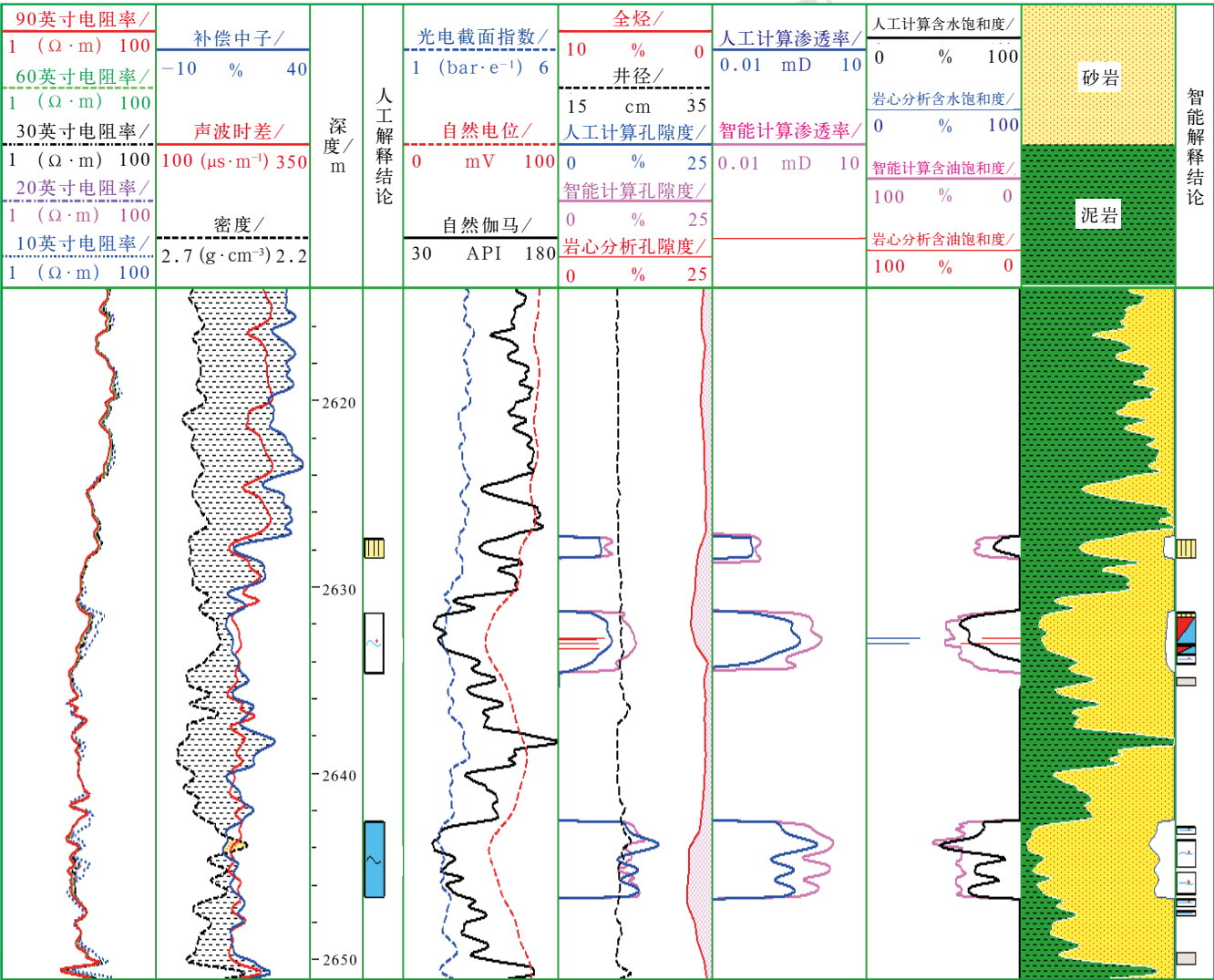


图 11 H369 井单井智能解释成果图

Fig.11 Single well intelligent interpretation results of well H369



图 12 H369 井及邻井批量智能解释成果图

Fig.12 Batch processing results of well H369 and its adjacent wells

为进一步认识该油藏,对 H369 井和周围 3 口邻井进行批量智能解释,在线绘制智能解释成果图(图 12),多井对比图自动按照海拔深度对齐,有利于多井对比分析。由图 12 可知,除 H369 井能够正确识别油水层外,邻井 H262 井和 H418 井智能解释结论均与试油结论吻合,H388 井智能解释结论为水层。在智能解释成果的基础上,进行精细小层对比,认为该区地质构造对油气成藏具有一定的控制作用,H262 井海拔稍高,试油产纯油(日产油 6.0t),H418 井海拔稍低,试油为油水同层(日产油 4.9t、日产水 9.3m³)。H369 井和 H388 井的油顶海拔比 H418 井低了 49m。H388 井距 H369 井仅 300m,海拔相当,但 H388 井人工解释结论和智能解释结论都为水层,分析认为 H369 井是孤立的构造—岩性小油藏,油藏面积较小,与 H262 井油藏并没有连通,分别属于两个小型构造—岩性油藏,这一结论与地质分析结论一致。

后期区域开发结果也证实了以上分析结果。区域开发时,在 H369 井周围部署了 4 口开发井,开发效果不好。投产的 3 口井(H231—43 井、H230—42 井、H231—42 井)试采期间产水率达到了 80%~100%。H230—42 井投产初期产油效果稍好,但在投产的 18 个月内,10 个月产水率为 100%,8 个月产水率为 80%~85%,累计产油仅 12t。从油藏开发效果来看,依据智能解释系统分析得到的结果可靠,证明系统可以较好地辅助人工认识油藏。

6 结语

随着大数据和计算机技术的飞跃发展,机器学习、深度学习等人工智能技术在测井解释中的应用,要快速走出前人的研究和试验阶段,进入高速发展阶段,当前正是进行测井解释智能化升级的最佳时期。中国测井解释软件走的是集成化发展道路,取得了较大的成功。智能化应用作为测井解释评价的一种手段,一方面要兼顾与测井解释软件的融合,另一方面还要具备较强的开放性,要将近年来测井专业信息化建设和智能化应用成果集成起来,形成开放性的测井智能应用环境。通过“数据”“算法”“场景”三者有机融合,构建“大数据+智能计算+专业软件”的测井智能解释应用环境。利用智能模型通过大数据分析提高测井参数计算和流体识别精度;通过信息化系统打通测井评价各专业之间的数据壁垒,实现数据的自动收集、转换和推送,完成外围数据准备工作,提升测井评价效率;在专业软件里实现智能解释与传统岩石物理

解释结合,互为补充,提升测井评价整体效果。测井智能解释在长庆油田的应用实践表明,这条智能化发展道路是正确的,有利于测井智能应用生态进一步发展。

参考文献

- [1] 曾文冲. 现代测井技术发展趋势的分析[J]. 测井技术, 1989,13(1):1-11.
Zeng Wenchong. An analysis of the development trend of modern logging technique[J]. Well Logging Technology, 1989,13(1):1-11.
- [2] 李乐中, 张立. 斯伦贝谢公司解释系统及软件介绍[J]. 国外测井技术, 1992,7(1):111-113.
Li Lezhong, Zhang Li. Introduction to interpretation system and software of schlumberger[J]. World Well Logging Technology, 1992,7(1):111-113.
- [3] 陆大卫, 张世刚. 测井软件走集成化道路[J]. 测井技术, 1996,20(1):5-11.
Lu Dawei, Zhang Shigang. On integration of well logging software[J]. Well Logging Technology, 1996,20(1):5-11.
- [4] 李浩, 刘双莲, 魏修平. 浅析我国测井解释技术面临的问题与对策[J]. 地球物理学进展, 2010,25(6):2084-2090.
Li Hao, Liu Shuanglian, Wei Xiuping. Analysis of the problem and strategies in logging interpretation techniques in our country[J]. Progress in Geophysics, 2010,25(6):2084-2090.
- [5] 谭廷栋. 中国石油测井学科的发展及展望[J]. 地球物理学报, 1997(增刊 1):344-350.
Tan Tingdong. Prospect and development for petroleum well logging science in China[J]. Chinese Journal of Geophysics, 1997(S1):344-350.
- [6] 石玉江, 王娟, 程启贵, 等. 数字化油藏研究理念与实践[M]. 北京: 石油工业出版社, 2020.
Shi Yujiang, Wang Juan, Cheng Qigui, et al. Concept and practice of digital reservoir research[M]. Beijing: Petroleum Industry Press, 2020.
- [7] 赵丽莎, 史永彬, 金玮, 等. 基于梦想云的测井智能化解释应用研究[J]. 中国石油勘探, 2020,25(5):97-103.
Zhao Lisha, Shi Yongbin, Jin Wei, et al. Application research on intelligent logging interpretation based on E&P Dream Cloud[J]. China Petroleum Exploration, 2020,25(5):97-103.
- [8] 邹德江, 谢关宝. 大数据时代的现代测井解释技术探讨[J]. 石油地质与工程, 2016,30(6):51-54.
Zou Dejiang, Xie Guanbao. Exploration of modern logging interpretation technology in the era of big data[J]. Petroleum Geology and Engineering, 2016,30(6):51-54.
- [9] 白云程, 何川. 人工智能专家系统在测井综合解释中的应用[J]. 测井技术, 1984,8(4):16-32.
Bai Yuncheng, He Chuan. Application of artificial intelligence expert system in logging comprehensive interpretation[J]. Well Logging Technology, 1984,8(4):16-32.
- [10] 伊广林, 吕健儒, 吴付东, 等. 石油储集层评价专家系统[J]. 测井技术, 1986,10(1):1-8.
Yi Guanglin, Lü Jianru, Wu Fudong, et al. Expert system for oil reservoir evaluation[J]. Well Logging Technology, 1986,10(1):1-8.
- [11] 袁祖贵. 砂泥岩地层测井解释专家系统在孤岛油田的应用[J]. 测井技术, 1988,12(3):74-79.
Yuan Zugui. Application of sand shale formation logging interpretation expert system in Gudao Oilfield[J]. Well Logging

- Technology, 1988,12(3):74-79.
- [12] 胡克珍, 王继贤, 詹晓君, 等. 利用测井曲线进行小层对比的专家系统[J]. 测井技术, 1989,13(4):71-76.
Hu Kezhen, Wang Jixian, Zhan Xiaojun, *et al.* An expert system for detailed correlation of well logs[J]. Well Logging Technology, 1989,13(4):71-76.
- [13] 夏宏泉, 刘红岐. BP 神经网络在测井资料标准化中的应用[J]. 测井技术, 1996,20(3):201-206.
Xia Hongquan, Liu Hongqi. Application of BP neural network to normalization of logging data[J]. Well Logging Technology, 1996,20(3):201-206.
- [14] 陈遵德. 测井数据模式识别中的信息优化方法[J]. 测井技术, 1998, 22(6):43-46.
Chen Zunde. A method for information optimization in pattern recognition with log data[J]. Well Logging Technology, 1998,22(6):43-46.
- [15] 宋子齐, 谭成仟, 吴少波, 等. 灰色系统与神经网络技术分析软件在油田勘探开发中的应用[J]. 测井技术, 2000,24(6):423-427.
Song Ziqi, Tan Chengqian, Wu Shaobo, *et al.* Application of grey system & neural network analysis software in oilfield exploration and exploitation[J]. Well Logging Technology, 2000,24(6):423-427.
- [16] 李洪奇, 李雄炎, 谭锋奇, 等. 基于数据挖掘技术的测井评价方法[J]. 测井技术, 2009,33(1):16-21.
Li Hongqi, Li Xiongyan, Tan Fengqi, *et al.* A log evaluation method based on the data mining technique[J]. Well Logging Technology, 2009,33(1):16-21.
- [17] 李洪奇, 谭锋奇, 徐长福, 等. 基于决策树方法的砾岩油藏岩性识别[J]. 测井技术, 2010,34(1):16-21.
Li Hongqi, Tan Fengqi, Xu Changfu, *et al.* Lithology identification of conglomerate reservoir based on decision tree method[J]. Well Logging Technology, 2010,34(1):16-21.
- [18] 杨莉娜, 杨斌, 鲁洪江, 等. 五百梯气田长兴组储层裂缝发育程度的常规测井神经网络预测[J]. 中国石油勘探, 2011, 16(1):10,63-69.
Yang Lina, Yang Bin, Lu Hongjiang, *et al.* Prediction of fracture development in Changxing Formation reservoir of Wubaiti Gas Field from conventional logging data using neural Networks[J]. China Petroleum Exploration, 2011,16(1):10,63-69.
- [19] 李虎, 范宜仁, 丛云海, 等. 基于改进 SADE 算法的神经网络预测储层物性[J]. 测井技术, 2012,36(6):585-589.
Li Hu, Fan Yiren, Cong Yunhai, *et al.* A new method predicting reservoir properties with neural network based on SADE algorithm[J]. Well Logging Technology, 2012,36(6):585-589.
- [20] 李艳华, 王红涛, 王鸣川, 等. 基于 PCA 和 KNN 的碳酸盐岩沉积相测井自动识别[J]. 测井技术, 2017,41(1):57-63.
Li Yanhua, Wang Hongtao, Wang Mingchuan, *et al.* Automatic identification of carbonate sedimentary facies based on PCA and KNN using logs[J]. Well Logging Technology, 2017,41(1):57-63.
- [21] 石玉江, 周金昱, 钟吉彬, 等. 重构电阻率曲线识别水淹层的方法及应用[J]. 测井技术, 2018,42(1):42-48.
Shi Yujiang, Zhou Jinyu, Zhong Jibin, *et al.* Method and application for identifying waterflooded layer with reconstructed resistivity curves[J]. Well Logging Technology, 2018,42(1):42-48.
- [22] 谭茂金, 白洋, 王谦, 等. 当非常规油气遇到人工智能: 多源数据驱动下非常规油气测井智能解释方法研究进展[C]. 2019 年油气地球物理学学术年会, 2019.
Tan Maojin, Bai Yang, Wang Qian, *et al.* When unconventional oil and gas meet artificial intelligence: progress of intelligent interpretation methods for unconventional oil and gas logging driven by multi-source data[C]. Proceedings of 2019 Annual Meeting of Oil and Gas Geophysics, 2019.
- [23] 杨华, 石玉江, 王娟, 等. 油气藏研究与决策一体化信息平台的构建与应用[J]. 中国石油勘探, 2015,20(5):1-8.
Yang Hua, Shi Yujiang, Wang Juan, *et al.* Construction and application of reservoir research and decision-making integrated information platform[J]. China Petroleum Exploration, 2015, 20(5):1-8.
- [24] 石玉江, 王娟, 魏红芳, 等. 基于梦想云的油气藏协同研究环境构建与应用[J]. 中国石油勘探, 2020,25(5):15-22.
Shi Yujiang, Wang Juan, Wei Hongfang, *et al.* Construction and application of oil and gas reservoir collaborative research environment based on E&P Dream Cloud[J]. China Petroleum Exploration, 2020,25(5):15-22.
- [25] 付金华, 石玉江, 王娟, 等. 长庆油田勘探开发服务型共享数据中心构建研究[J]. 中国石油勘探, 2017,22(6):1-8.
Fu Jinhua, Shi Yujiang, Wang Juan, *et al.* Construction of service-oriented shared data center for exploration and development of Changqing Oilfield[J]. China Petroleum Exploration, 2017,22(6):1-8.
- [26] 杨勇, 黄文俊, 王铁成, 等. 梦想云数据连环湖建设研究[J]. 中国石油勘探, 2020,25(5):82-88.
Yang Yong, Huang Wenjun, Wang Tiecheng, *et al.* Research on construction of Data Interlinked Lakes of E&P Dream Cloud[J]. China Petroleum Exploration, 2020,25(5):82-88.
- [27] 杜金虎, 时付更, 杨剑锋, 等. 中国石油上游业务信息化建设总体蓝图[J]. 中国石油勘探, 2020,25(5):1-8.
Du Jinhu, Shi Fugeng, Yang Jianfeng, *et al.* Overall blueprint of information construction of PetroChina upstream business[J]. China Petroleum Exploration, 2020,25(5):1-8.
- [28] 王同良. 油气行业数字化转型实践与思考[J]. 石油科技论坛, 2020, 39(1):29-33.
Wang Tongliang. Practice and thinking of oil and gas industrial digitalization transformation[J]. Petroleum Science and Technology Forum, 2020,39(1):29-33.
- [29] 付锁堂, 石玉江, 丑世龙, 等. 长庆油田数字化转型智能化发展成效与认识[J]. 石油科技论坛, 2020,39(5):9-15.
Fu Suotang, Shi Yujiang, Chou Shilong, *et al.* Achievements from digital transformation and intelligent development of Changqing Oilfield[J]. Petroleum Science and Technology Forum, 2020,39(5):9-15.
- [30] 徐鹏, 高健伟, 陈溯, 等. 勘探开发数据资产化管理实践与思考[J]. 石油科技论坛, 2020,39(5):34-40.
Xu Peng, Gao Jianyi, Chen Su, *et al.* Research and practice of asset management of exploration and development data[J]. Petroleum Science and Technology Forum, 2020,39(5):34-40.
- [31] Andreas Müller, Sarah Guido. Introduction to machine learning with Python—a guide for data scientists[M]. California: O'Reilly Media, 2016.
- [32] Aurélien Geron. Hands-on machine learning with scikit-learn and tensor flow—concepts, tools, and techniques for building intelligent systems[M]. California: O'Reilly Media, 2019.
- [33] Sinan Ozdemir, Divya Susarla. Feature engineering made easy[M]. Birmingham: Packt Publishing, 2018.
- [34] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
Zhou Zhihua. Machine learning[M]. Beijing: Tsinghua University Press, 2016.
- [35] 贾俊平, 何晓群, 金勇进. 统计学[M]. 北京: 中国人民大学出版社, 2014.
Jia Junping, He Xiaogun, Jin Yongjin. Statistics[M]. Beijing: China Renmin University Press, 2014.