

DOI: 10.3969/j.issn.1672-7703.2020.05.013

基于梦想云的测井智能化解释应用研究

赵丽莎¹ 史永彬^{2,3} 金玮^{2,3} 李华¹ 塔斯肯¹

(1 中国石油勘探开发研究院; 2 中国石油集团东方地球物理勘探有限责任公司; 3 北京中油瑞飞信息技术有限责任公司)

摘要: 迄今为止, 地球物理测井仍然是石油工业中用于识别地下岩层物性以及是否含油气性的最重要的技术手段。利用大数据和人工智能方法, 基于梦想云智能数据分析平台开展测井智能化解释, 以数据收集和拼接为基础, 进行特征抽取, 结合专家知识和经验构建模型, 开展机器学习和模型训练, 完成模型测试与迭代, 并分别对砂岩、泥岩的判别(阈值为0.4, 60口井)和油层、水层判别(阈值为0.5, 47口井)进行了智能化解释技术验证。解释结果显示: 砂岩、泥岩判别用时约2~4h, 泥岩和砂岩判别准确率分别为94%和90%, 综合判别准确率为92%; 油层、水层判别用时约4~6h, 油层和水层判别准确率分别为93%和70%, 综合判别准确率为86%, 有效提高了多口井测井解释自动化程度和结果符合率, 大数据分析技术应用为传统的石油测井解释技术发展带来了新的动能和良好的发展前景。

关键词: 梦想云; 测井; 大数据; 人工智能; 测井解释

中图分类号: TE19

文献标识码: A

Application research on intelligent logging interpretation based on E&P Dream Cloud

Zhao Lisha¹, Shi Yongbin^{2,3}, Jin Wei^{2,3}, Li Hua¹, Ta Siken¹

(1 PetroChina Research Institute of Petroleum Exploration & Development; 2 BGP Inc., CNPC; 3 Richfit Information Technology Co., Ltd.)

Abstract: Up to now, geophysical logging is still the most important technical means to identify the physical properties of subsurface strata and whether they are hydrocarbon-bearing. Using the technology of big data and artificial intelligence, based on Dream Cloud intelligent data analysis platform, this paper develops the intelligent well logging interpretation. On the basis of data collection and splicing, features of log data are extracted. Combined with expert knowledges and experiences, models are established, machine learning and model training are carried out, and model testing and iteration are completed. The intelligent logging interpretation technology is verified by sandstone and mudstone identification (cut-off value is 0.4 based on 60 wells) and oil and water layer identification (cut-off value is 0.5 based on 47 wells). The interpretation results show that, it takes about 2-4 hours to distinguish sandstone and mudstone. The accuracy rate of mudstone and sandstone identification is 94% and 90% respectively, and the comprehensive accuracy rate is 92%. And It takes about 4-6 hours to distinguish oil layer and water layer. The accuracy rate of oil layer and water layer identification is 93% and 70% respectively, and the comprehensive accuracy rate is 86%. This technology can effectively improve the automation degree and result coincidence rate of multi-well logging interpretation. The application of big data analysis technology brings new motive force and good prospects for the development of traditional well logging interpretation technology.

Key words: Dream Cloud, well logging, big data, artificial intelligence, well log interpretation

0 引言

测井解释技术是石油上游业务中用于油气层识别、老区挖潜、老井复查等关键业务的核心技术^[1-4]。传统

的测井解释是一项综合性极强的工作, 其中除需要使用数十种测井资料外, 还需要参考钻井、录井、试油、地质、油藏等各专业的海量资料数据; 这些数据种类多、数据分析处理难度大, 且专家水平、经验参差不齐,

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: 赵丽莎(1990-), 女, 北京人, 博士, 2017年毕业于北京大学, 工程师, 现主要从事油气田勘探开发、人工智能及大数据应用等方面工作。地址: 北京市海淀区学院路20号中国石油勘探开发研究院, 邮政编码: 100083。E-mail: zhaolisha@petrochina.com.cn

收稿日期: 2020-07-10; 修改日期: 2020-07-17

存在测井解释结果与实际情况符合率低的情况,例如,在油田气藏储层解释研究中,部分被解释判别为“差油层”的储层,经专家重新研究评价和试采后,变成了油气生产潜力层。因此,油气储层研究中迫切需要引入新的技术来解决传统研究和认识上的局限问题^[5-8]。

测井及其关联数据具备体量大、种类多、变化快等大数据特征。大数据算法主要包括分类、回归、聚类、关联、神经网络方法、数据挖掘等^[9-13]。基于机器学习、模式学习、统计学等基础,大数据算法力求从海量的、不完全的、有噪声的、模糊的、随机的数据中发现隐含在其中有价值的、潜在的有用信息和知识。目前大数据的各种算法已经被广泛应用于手写识别、语音识别、人脸识别、面部表情分类、车牌识别、文本提取、最优路径分析、数据挖掘等各个方面,并且取得了不错的效果^[14-17]。

随着石油勘探开发难度加大,应用大数据、人工智能等新方法、新技术来解决勘探开发业务痛点问题的研究探索工作呈现加速趋势^[18-20]。国内外均涌现出一些将大数据算法应用于测井解释的研究探索性案例。例如,运用支持向量机技术进行岩性识别,使用线性回归或深度学习算法预测泥质含量和储层物性参数等,均取得了一定的效果。相比于传统的测井解释研究,大数据算法在测井解释领域具有较大的优越性,

如通过大数据算法,将输入样本映射到高维空间,将一个非线性问题转换成线性问题等,可以很好地解决测井解释过程中存在的数据类别多、数据量大、定性问题多等需要依靠专家经验的难题。然而,由于实际地质情况的复杂性,大数据算法的多样性也带来了新的问题,即处于不同地质环境的含油气储层,使用相同的大数据算法模型可能会得到不同的评价结果,因而有必要研究在相似地质环境中,改善大数据算法模型的适应性问题。探索大数据算法在测井解释中的适用性应用,对于提高测井解释成果准确度和降低成本具有重大意义。

1 技术方案

1.1 方案设计

基于梦想云智能数据分析平台开展测井智能化解解释不同于传统测井解释方法,它是以梦想云丰富的勘探开发数据湖为基础,以PaaS云平台为支撑,通过数据收集与拼接,结合专家知识和经验,进行特征构建和模型构建,开展机器学习和模型训练,指导单井纵向到层、多井横向对比自动解释,为新井解释、老井复查降低劳动强度,实现测井解释新技术的探索与突破。其总体架构设计如图1所示。

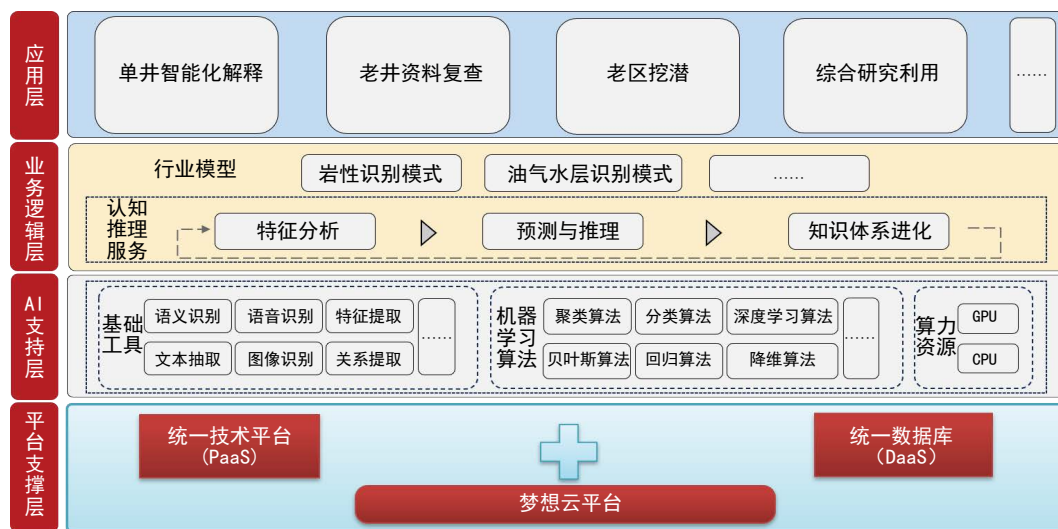


图1 基于梦想云平台的测井智能化解解释应用架构

Fig.1 Application architecture of intelligent well logging interpretation based on Dream Cloud platform

1.2 技术路线

测井智能化解解释技术路线共分为4个部分,包括数据收集、特征构建、模型构建,以及模型测试与

迭代(图2)。

(1) 数据收集包括对结构化数据,如钻井、邻井、测井等数据,以及非结构化数据,如地质资料、解释结论、井史资料等收集整理工作。

(2) 特征构建是在基于大数据分析的专业数据基础之上进行的。研究并分析测井解释数据要求、业务流程和专家经验等内容,采用知识图谱、大数据存储等技术,建立包含专家经验、业务规则、分析成果、数据标准在内的领域知识库,提供更快捷的数据查询、钻取和关联,以支撑大数据分析技术对数据的需求。

(3) 模型构建指构建测井解释的大数据分析算法、模型。基于测井数据及其他数据提供的特征值,应用深度学习算法对特征值解释结论进行训练(数据样本训练和专家训练),产出智能测井结果。通过示范区

块数据迭代得到初步的训练算法和模型,通过扩大机器学习的数据范围,逐步迭代,进一步获得结果符合度更高的测井智能化解释模型和算法。

(4) 模型测试与迭代过程结束后,可以形成完整的测井智能解释系统,建立流程体系。在大数据算法和模型的基础上,建立成熟的基于大数据分析的测井智能化解释流程体系,并根据应用场景形成人机交互的业务人员可以应用的智能化解释系统,能够实现特征值提取、模型算法选择、模型训练、可视化展示等功能。



图2 测井智能化解释技术路线

Fig.2 Technical route of intelligent well logging interpretation

1.3 实施方案

1.3.1 数据收集与拼接

本次研究共收集了测井解释相关的测井曲线、录井、取心、地层水、小层、井口地理位置等数据对象,

上述数据对象的含义及形式见表1;上述数据除井口地理位置仅利用井号外,其他数据对象均采用由测点、区段、小层号等组合形成的唯一标识;上述数据均为直井数据。

表1 测井解释相关的各类数据的含义及特征

Table 1 Meaning and characteristics of various data related to well logging interpretation

数据名称	含义	原始数据列	数据特征
测井曲线	测井仪记录的当前深度的各类物理数据,形成曲线	自然电位(SP)、自然伽马(GR)、深/浅双侧向电阻率(RLLD/RLLS)、微球形聚焦电阻率(MSFL)、声波时差(AC)、补偿中子(CNL)、密度(DEN)	连续测点,采样间隔为0.125m
录井	某个深度区段的岩石的物理性状	岩石性质、颜色、含油级别	离散区段
取心	某个特定测点的岩石的物理性状	岩石性质、颜色、含油级别	离散测点
地层水	当前层位地层水的水型及离子浓度	水型、pH值、总矿化度、各类离子浓度	小层号
小层	各井内指示自上一个测点至本测点对应的小层号	测深点、对应层位	小层号
井口地理位置	井口的大地坐标	x坐标和y坐标	井号

对数据实际中存在的问题,进行如下预处理:

(1) 数据缺失值、异常值或重复值。当某些井缺失部分数据时,为保证数据的连续性,需要对数据充实处理;对异常值或重复值,需要做针对性处理。

(2) 数据缺乏统一标志或维度,测井数据以连续

测点、离散测点表示,而录井数据是以区段、小层号等形式表示,这些数据无法直接整合在一起,需要将录井数据处理成测井数据形式;此外,由于基于机器学习的建模需要使用统一维度的数据,因此,数据维度(即数据类型)不同则不能用于数据建模。

为了较好地整合数据,需首先为每类数据进行异常值处理及数据拼接,数据拼接后,所有类型的数据将对齐到测井曲线的最细粒度或采样,即将钻井、录井、地质与岩性分层等信息对应到测井的每个采样深度。

同一类数据的异常值处理和拼接应一并完成。在拼接测井曲线、录井、取心数据时,以测井曲线为基础,首先将录井数据拼接给测井曲线,形成一份新数据,之后将取心数据拼接给这套新数据,逐步追加形成完整数据。

1.3.2 特征构建思路

对于岩性识别和油层、水层识别特征的构建主要参照以下基本思路进行。对于自然电位(SP)、自然伽马(GR)、深/浅双侧向电阻率(RLLD/RLLS)、微球形聚焦电阻率(MSFL)、声波时差(AC)、补偿中子(CNL)、密度(DEN),利用其英文缩写标识并进行描述。

1.3.2.1 单井测井曲线数值及物理特征

对于GR值,砂岩较小、泥岩较大;在砂岩中,油层较大、水层较小。

对于SP值,砂岩较大、泥岩较小;在砂岩中,油层较小、水层较大。

对于三电阻率数据RLLD、RLLS和MSFL,泥岩低于砂岩;在砂岩中,油层较大、水层较小。

对于薄层,因为测量结果会受到上下邻近区段的干扰,因此其特征值要取尖峰,乃至对尖峰进行放大;厚层则应取中段数据计算。

受密度和浮力的影响,在同一小层内的每一个连续渗透层(上下的泥岩或围岩超过2m)中,油层应位于水层上方,气层处于油层上方。

1.3.2.2 辅助信息

地层水会提高或降低当前层的整体电阻率,砾岩会提高当前层电阻率,砂岩会降低当前层电阻率。上述现象导致出现高阻水层和低阻油层。

对于取心和录井数据,含油级别分为富含油、饱含油、油浸、油斑、油迹、荧光。因为油相对于水密度较小,对于研究区段的邻井,在垂深上,邻井的一些区段若与该井区段同属一个小层,同时邻井区段处于相对较浅的位置,则可推测邻井的同区段为相对有利层(为油层可能性较大);若邻井区段处于相对较深的位置,则可推测邻井的同区段为相对不利层(为水层可能性较大)。

1.3.3 特征构建过程

特征构建过程包括业务参数构建、异常值剔除与

标准化、整体特征构建、补充其他特征等部分。

基于本项研究所提出的特征构建思想,形成了较多的特征。由于岩性识别与油层、水层识别构建的特征有一定重叠,在此按组将所有构建的特征分别进行叙述,并给出每组特征的应用范围。

1.3.3.1 业务参数构建

业务参数是基于曲线数据计算各个测点的当前各业务参数值,与区段的统计量等无关。在计算业务参数后,再进行统一标准化及后续面向区段的数据特征构建。

计算各类孔隙度,包括声波孔隙度、密度孔隙度和中子孔隙度:计算纵波速度和泥质含量;基于声波孔隙度、密度孔隙度、中子孔隙度计算横波速度、剪切模量、泊松比和流体因子。

1.3.3.2 异常值剔除与标准化

异常值剔除与标准化过程包括:①业务参数异常值剔除;②进行测井曲线的标准化;③进行岩石名称、颜色、含油级别的映射;④进行地理位置坐标和地层水数据标准化;⑤整体特征构建——岩性识别模块。

本部分利用曲线数据和部分业务参数进行实现,对于曲线数据,利用基于原始数据及此前进行标准化的结果数据进行特征构建,即对于所有曲线数据,利用了原值和标准化值两套值。

对于RLLD、RLLS、MSFL以及对应的标准化后的值,利用最小值作为基线;而对于GR、SP、AC、CNL、DEN以及对应的标准化后的值,利用最大值和最小值的平均值作为基线,考察其偏离基线的情况。在各井中计算各自的基线值,并标记这些基线值为base。

针对岩性特征识别,在一个程序指示的顶—底深度的区段内进行构建。由于此前根据深度将数据进行过拼接,因而可以便捷地根据深度找到所有数据。区段内的特征构建过程为:

(1) 分别取GR、SP、AC、CNL、DEN、RLLD、RLLS、MSFL特征在本区段内的平均值作为原值。因为薄层会受上下区段干扰,对原值特征进行如下调整:待判别区段厚度在2m以上,不予变动;待判别区段厚度在2m以下,记区段长为 L ,将原始的平均特征值 f 变为 $\text{base} + (f - \text{base}) [1 + \lg(2/L)]$,即将 f 向基线base的反方向移动,调整其偏离的大小为原来的 $[1 + \lg(2/L)]$ 倍,保证过渡在 $L=2$ 点连续且适当对薄层扩展。当 L 小于1时, f 变为 $[1 + \lg(2/1)]$,防止调整倍数不合理地过大。

(2) 计算GR、SP、AC、CNL、DEN、RLLD、

RLLS、MSFL 在本区段中与各自基线 base 差值的平均值。

(3) 提取深/浅侧向电阻率的幅度差 RLLD-RLLS, 以及在参数计算中密度—补偿中子运算的孔隙度的差值。

(4) 计算其他所有构建业务参数的区段内平均值, 包括基于声波孔隙度、中子孔隙度和密度孔隙度计算的结果, 共 3 套。

1.3.3.3 整体特征构建——油层、水层识别

本部分利用曲线数据、部分业务参数以及各类辅助信息实现。在此, 对曲线数据, 仅利用原始数据进行特征构建。

对于任一待判别区段, 其基线值仅取研究区段上下 50m (合计 100m) 区段的曲线数据进行计算, 而非整个井的曲线数据。这是因为单井数据随着深度的变化使一些维度的数据发生整体偏移, 对较长的井段必须考虑这种变化。

为了更清楚地描述数据, 相比岩性识别, 这里基于不同的数据基准构建了两套基线, 以研究区段上下 50m 的所有区段计算出的基线, 记为 base_c; 仅以研究区段上下 50m 的所有砂岩区段计算出的基线, 记为 base_f; 岩性数据可参考录井数据进行计算, 在建模过程中则是将所有解释结论包含的区段作为该井内所有的砂岩数据。在这种基线背景下, 对于两套基线基于的数据系统 base_c 和 base_f, 对 RLLD、RLLS、MSFL 和 GR、SP、AC、CNL、DEN 分别计算利用 1% 分位数和 99% 分位数作为一套基线的最小值和最大值。

类似于岩性识别, 油层、水层区段特征在一个程序指示的顶—底深度的区段内进行构建。这里, 在区段内构建特征如下:

(1) 取 GR、SP、AC、CNL、DEN、RLLD、RLLS、MSFL 特征在本区段内的平均值作为原值。类似于岩性识别, 因为薄层会受上下区段干扰, 对原值特征进行如下调整: 待判别区段厚度在 2m 以上, 不予变动; 待判别区段厚度在 2m 以下, 计算基线 GR、SP、AC、CNL、DEN 为基线 base_c 的最大值和最小值的平均值, RLLD、RLLS、MSFL 为基线 base_c 的最小值。类比岩性特征识别, 在此记这些值为 base; 记区段长为 L , 将原始的平均特征值 f 变为 $\text{base} + (f - \text{base}) [1 + \lg(2/L)]$, 即将 f 向基线 base 的反方向移动, 调整其偏离的大小为原来的 $[1 + \lg(2/L)]$ 倍, 当 L 小于 1 时, f 变为 $[1 + \lg(2/1)]$ 。

(2) 计算所有原始特征 GR、SP、AC、CNL、DEN、

RLLD、RLLS、MSFL 与 base_c、base_f 的最小值、最大值的差, 合计得到 4 套移动参数, 与 base_c 的差值描述了当前数据特征相比泥岩基线的变动、与 base_f 的差值描述了当前数据特征相比当前其他砂岩区段的相对位置。

(3) 提取深/浅侧向电阻率的幅度差 RLLD-RLLS, 以及在参数计算中密度—补偿中子运算的孔隙度的差值。

(4) 计算其他所有构建业务参数的区段内平均值, 包括基于声波时差孔隙度、补偿中子孔隙度和密度孔隙度计算的结果, 共 3 套。

1.4 模型训练与实际特征构建

对于岩性识别, 训练数据主要以原始曲线数据为数据基础构建特征, 以录井解释结论为标签。在构建训练数据的特征时, 以滑动窗口的形式构建每段泥岩和砂岩的所有样本。在本次研究中, 设置的区段长度即窗口为 2m、步长为 0.5m。在进行测试及预测时, 将自动处理整段井数据, 以 2m 为窗口、以 0.5m 为步长进行待预测样本的构建及建模判别。这表示当提供所有曲线数据时, 能自动地对砂岩、泥岩进行判别。

对于油层、水层识别, 训练数据结合了原始曲线数据以及各类辅助数据, 以综合解释结论为标签, 将油层、可能油气层、油水同层均看作油层。在构建训练数据特征时, 以测井解释结论所有记录条目作为油层、水层的所有样本构建范围。在进行测试时, 同样利用测井解释结论提到的所有样本条目; 而进行预测时, 用户需手工指定所有待预测的井号及对应区段, 由系统提供解答。这表示本系统可以半自动地对一个区段的油层、水层进行判别。

上述特征构建完成后产生了大量特征值, 但由于已有各类数据的缺失, 特别是各井、各样本数据的缺失情况不同, 在实际过程上会产生大量的空值。对此, 本模型构建多套数据基础, 在每套数据集上进行模型构建。

1.4.1 构建训练数据集

为了尽可能地利用各个样本所含有的所有数据, 本模型通过如下方式构建多套数据集, 保证每套数据集上全部数据非空。

对于每口单井, 首先考察各井部分缺失的数据列及全部缺失的数据列。之后将其集合进行去重复, 得到的数据在各井间存在多种不同的缺失类型。这种处理方式较为简明, 得到的数据集数目可控。

当收集了所有数据缺失的类型后, 即可将不同类型的数据缺失列分别从特征集合中排除, 形成保证在本数据缺失类型下必须保有的全部数据列, 此后对该

套数据集仅用这些数据列进行建模,该套数据集将用于对这些数据列上全部非空的样本进行建模。

对于每一条训练数据及测试数据的样本,逐一与各种数据缺失类型下所需的数据列进行比较,观察是否符合其所需数据列要求,即在这些列上全部非空。若是全部非空,则表示该样本可在该数据集下进行建模训练或预测。如此,便可得到每个样本能在哪些数据模型下用于训练,或是被预测。

最终,对于岩性识别无须构建多套样本,对于油层、水层识别则构建了7套数据集。每套数据集内的样本在这些数据列上全部非空。一般而言,一套数据集数据完整性要求越高,这套数据集所含的样本数目越少。

1.4.2 模型构建

对于岩性分类,采用XGBoost单模型进行分类判别。

对于油层、水层分类,模型构建采用多模型分类+模型集成的方式。在训练阶段,以Lasso、SVM、GBDT、KNN和XGBoost为基模型。同时,由于大多数模型水层相比油层明显较多,本模型对数据集进行一次平衡,将水层随机抽取使数据集内油层、水层条数相等,形成另一套数据集,使上述5个基模型同时对原始数据集和平衡后的样本数据集进行建模,即构建10个基模型。

基模型构建完毕并产生基模型预测结果后,以XGBoost模型为集成模型进行训练。最终形成在某一套数据集下的集成模型。

在训练过程中,以80%的数据作为基模型的训练数据,对于剩下的20%的数据,以训练好的基模型对其进行预测,并将预测结果作为训练数据、油解释结论作为标签,训练集成模型。在预测阶段,将一条样本利用10个基模型进行预测,预测结果交给集成模型进行预测,形成该样本在当前数据集下的预测结果。

1.4.3 模型结果产出

对于岩性分类,判别结果为一个区段为砂岩的概率,判断的阈值为0.5。

对于油层、水层分类,由于根据数据完整性构建了多套数据集,模型评价过程实际上是在多套数据集上分别产生了一个样本的预测结果,返回结果为一个区段为油层的概率。在本项目中,对于多套结果,取其中的最大值,当其判别概率大于阈值0.4时,则认为这个区段为油层。调高阈值可提升油层的准确率、降低油层的召回率,反之同理。

2 研究成果

2.1 阈值为0.4时砂岩、泥岩的判别准确率

该模型共对60口井完成上述的判别过程,共耗时约2~4h,此运行时间会随数据集的不同和模型迭代次数的不同发生变化。判别结果如表2及图3所示,该结果显示,此模型的判别准确率基本满足了砂岩、泥岩的自动划分要求。

表2 砂岩、泥岩判别准确率结果

Table 2 Accuracy of sandstone and mudstone identification

类型	准确率/%	召回率/%	样本数/个
泥岩	94	85	39825
砂岩	90	96	55732
总体	92	91	95557

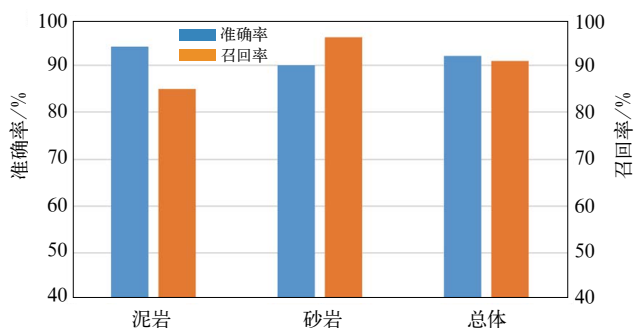


图3 砂岩、泥岩判别准确率统计图

Fig.3 Statistical chart of accuracy of sandstone and mudstone identification

2.2 阈值为0.5时油层、水层判别准确率

该模型对47口井完成上述判别过程,共耗时约4~6h,此运行时间会随数据集的不同和模型迭代次数的不同发生变化。判别结果如表3及图4所示,结果显示,有些油层特征不明显的油层未被识别出来,部分高阻水层、低阻油层难以被模型识别。但总体而言,该模型判别的准确率与专家人工解释结果的准确率相当。

表3 油层、水层判别准确率结果

Table 3 Accuracy of oil layer and water layer identification

类型	准确率/%	召回率/%	样本数/个
水层	93	85	1774
油层	70	85	726
总体	86	85	2500

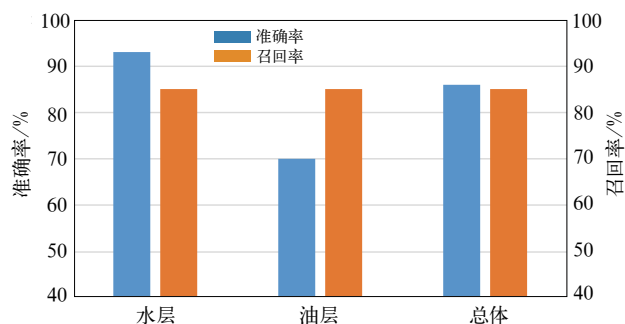


图4 油层、水层判别准确率统计图

Fig.4 Statistical chart of accuracy of oil layer and water layer identification

3 结论

梦想云平台为中国石油勘探开发业务应用提供了强有力的技术支撑,也为测井大数据、人工智能技术应用和发展奠定了良好基础,利用梦想云数据湖提供的丰富数据资源、专家知识和成果,以及云平台所具备的机器深度学习的算例与算法,实现了传统测井解释工作的智能化应用新突破,提高了测井解释的自动化程度和结果符合率,同时,也丰富了梦想云地质研究业务应用场景,通过进一步完善,可实现在油田测井解释工作中的工业化推广。

大数据时代到来给测井解释技术进步和发展带来了新契机,同时也给测井解释工程师们带来了新挑战。在大数据应用背景下,测井解释工程师们不仅要传统的测井数据理解释理念进行转变,还需要将现代大数据分析及人工智能技术应用于测井解释中,在提升解释效率的同时确保解释精度。大数据分析技术应用为传统的石油测井解释技术发展带来了新动能和良好的发展前景。

参考文献

- [1] 张龙海, 周灿灿, 刘国强, 等. 孔隙结构对低孔低渗储集层电性及测井解释评价的影响[J]. 石油勘探与开发, 2006, 33(6): 671-676.
Zhang Longhai, Zhou Cancan, Liu Guoqiang, *et al.* Influence of pore structures on electric properties and well logging evaluation in low porosity and permeability reservoirs[J]. Petroleum Exploration and Development, 2006, 33(6): 671-676.
- [2] 周灿灿, 王昌学. 水平测井解释技术综述[J]. 地球物理学进展, 2006, 21(1): 152-160.
Zhou Cancan, Wang Changxue. Technology review on the log interpretation of horizontal well[J]. Progress in Geophysics, 2006, 21(1): 152-160.
- [3] Wyllie M R J. Petroleum engineering. (Book reviews: the fundamentals of well log interpretation)[J]. Science, 1964, 143(3606): 560.
- [4] Moran J H, Gianzero S. Effects of formation anisotropy on

- resistivity-logging measurements[J]. Geophysics, 1979, 44(7): 1266-1286.
- [5] Bartetzko A, Pezard P, Goldberg D, *et al.* Volcanic stratigraphy of DSDP/ODP hole 395A: an interpretation using well-logging data[J]. Marine Geophysical Researches, 2001, 22(2): 111-127.
- [6] 杨小兵, 杨争发, 谢冰, 等. 页岩气储层测井解释评价技术[J]. 天然气工业, 2012, 32(9): 33-36.
Yang Xiaobing, Yang Zhengfa, Xie Bing, *et al.* Log interpretation and evaluation of shale gas reservoir[J]. Natural Gas Industry, 2012, 32(9): 33-36.
- [7] Wang Y, Xia H, Chen K. Optimized log interpretation for carbonate reservoirs in Tarim Basin[J]. Journal of Southwest Petroleum Institute, 2000, 22(4): 18-21.
- [8] Shen M Y, Chen J K, Yao W. Development of the well logging interpretation of thin layers[J]. World Well Logging Technology, 2005, 35(3): 385-398.
- [9] Min C, Mao S, Liu Y. Big data: a survey[J]. Mobile Networks & Applications, 2014, 19(2): 171-209.
- [10] Vlahogianni E I. Computational intelligence and optimization for transportation big data: challenges and opportunities [M]. Engineering and Applied Sciences Optimization, Springer International Publishing, 2015.
- [11] Agrawal D, Bernstein P, Bertino E, *et al.* Challenges and opportunities with big data 2011-1[J]. Proceedings of the VLDB Endowment, 2012, 5(12): 2032-2033.
- [12] Philip Chen C L, Zhang C Y. Data-intensive applications, challenges, techniques and technologies: a survey on big data[J]. Information Sciences, 2014, 275(11): 314-347.
- [13] Nguyen T T T, Armitage G. A survey of techniques for internet traffic classification using machine learning[J]. IEEE Communications Surveys & Tutorials, 2009, 10(4): 56-76.
- [14] 王栋. 大数据时代的现代测井解释技术[J]. 化工设计通讯, 2018, 44(11): 238.
Wang Dong. Analysis of modern logging interpretation technology in the age of big data[J]. Chemical Engineering Design Communications, 2018, 44(11): 238.
- [15] 邹德江, 谢家宝. 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. Discussion on the technology of modern logging interpretation of big data era[J]. Petroleum Geology and Engineering. 2016, 30(6): 51-54.
- [16] 杜金虎, 时付更, 张仲宏, 等. 中国石油勘探开发梦想云研究与实践[J]. 中国石油勘探, 2020, 25(1): 58-66.
Du Jinhu, Shi Fugeng, Zhang Zhonghong, *et al.* Research and practice of Dream Cloud for exploration and development of PetroChina[J]. China Petroleum Exploration, 2020, 25(1): 58-66.
- [17] Baaziz A, Quoniam L. How to use big data technologies to optimize operations in upstream petroleum industry[J]. Social Science Electronic Publishing, 2014, 1(1): 19-25.
- [18] Staff P. Application of big data analytics to optimize the operations in the upstream petroleum industry[C]. International Conference on Computing for Sustainable Global Development, IEEE, 2015.
- [19] Li D W, Shi G R. Optimization of common data mining algorithms for petroleum exploration and development[J]. Acta Petrolei Sinica, 2018, 39(2): 240-246.
- [20] Anifowose F, Labadin J, Abdulraheem A. Ensemble learning model for petroleum reservoir characterization: a case of feed-forward back-propagation neural networks[C]. Revised Selected Papers of Pakdd International Workshops on Trends & Applications in Knowledge Discovery & Data Mining, 2013.