



石油钻探技术
Petroleum Drilling Techniques
ISSN 1001-0890, CN 11-1763/TE

《石油钻探技术》网络首发论文

题目: 人工智能钻井技术研究方法及其实践
作者: 杨传书, 李昌盛, 孙旭东, 黄历铭, 张好林
收稿日期: 2020-10-15
网络首发日期: 2021-06-04
引用格式: 杨传书, 李昌盛, 孙旭东, 黄历铭, 张好林. 人工智能钻井技术研究方法及其实践. 石油钻探技术.
<https://kns.cnki.net/kcms/detail/11.1763.TE.20210603.1530.002.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

钻井完井

doi:10.11911/syztjs. 2020136

人工智能钻井技术研究方法及其实践

杨传书, 李昌盛, 孙旭东, 黄历铭, 张好林

(中国石化石油工程技术研究院, 北京 102206)

摘要: 人工智能技术飞速发展, 部分行业已取得明显效果, 但人工智能技术在钻井领域的应用尚处于探索阶段。为推动人工智能技术在钻井领域的应用, 在简述了钻井行业人工智能应用研究情况的基础上, 提出了将人工智能技术具体应用到钻井领域的“三轮驱动”方法论, 分析了钻井领域适合开展人工智能研究的业务场景以及人工智能技术工具, 提出了基于方法论进行项目优选的评价方法, 给出了评价实例, 并以井下故障复杂实时诊断为例简述了钻井工程人工智能应用研究的过程。指出了钻井领域开展人工智能应用研究存在的不足, 提出了钻井人工智能技术的发展建议, 以期推动钻井人工智能技术的发展。

关键词: 钻井; 人工智能; 大数据; “三轮驱动”方法论; 井下故障

中图分类号: TP18; TE938 **文献标志码:** A

Research Method and Practice of Artificial Intelligence Drilling Technology

YANG Chuanshu, LI Changsheng, SUN Xudong, HUANG Liming, ZHANG Haolin

(Sinopec Research Institute of Petroleum Engineering, Beijing, 102206, China)

Abstract: With the rapid development of artificial intelligence (AI) technology, it has made remarkable breakthroughs in many fields. However, the application of AI in drilling engineering is still in the primary stage. In order to promote the application of AI technology in drilling, based on a brief description of the research situation of its application in drilling engineering, a “three-wheel drive” methodology for the specific application of AI technology in drilling area was proposed. Then, business application scenarios and AI technology tools suitable for the research of AI in drilling engineering were analyzed. After putting forward a method of evaluating and optimizing projects based on the methodology with examples, the research process of AI application in drilling was illustrated by the real-time diagnosis of complex downhole failures. Finally, the shortcomings were identified and suggestions were given for the application of AI in drilling engineering, so as to promote the development of AI drilling technology.

Key words: drilling; artificial intelligence; big data; the “three drives” methodology; downhole failure

近年来, 新一代人工智能技术快速发展, 在智慧安防、智慧金融、智慧零售等领域取得了良好的应用效果。因此, 全社会掀起了“人工智能+”的研究热潮, 石油行业也不例外, 从与日陡增的文献可见一斑。尽管一些报道称在某些场景取得了很好的试验结果, 但总体来说人工智能技术在石油行业的应用研究仍处在探索阶段, 在钻井工程领域尚未取得工业应用的实质进展。从林林总总的文献报道来看, 人工智能的应用场景很多, 人工智能算法也很多, 那么如何能让更多的研究者拨云见日, 采用一种科学的方法加快人工智能应用技术研发呢? 笔者从方法论入手, 阐述一种将人工智能

收稿日期: 2020-10-15; 改回日期: 2021-03-10。

作者简介: 杨传书(1972—), 男, 湖北黄冈人, 1996年毕业于石油大学(华东)机械制造工艺与设备、计算机技术及应用专业, 2009年获中国石油大学(华东)石油与天然气工程专业工程硕士学位, 高级工程师, 主要从事石油工程信息技术、石油工程软件方面的研究与管理工作。E-mail: yangcs.sripe@sinopec.com。

基金项目: 中国石化新领域培育科技攻关项目“基于大数据的司钻智脑系统研制(一期)”(编号: XLY19001)和中国石化科技攻关项目“基于远程支持平台的工程数据实时分析技术研究”(编号: P18001)部分研究内容。

技术应用到钻井领域的技术路径。

1 钻井领域人工智能应用概况

人工智能（简称 AI）是计算机科学的一个分支，是一门认知学科，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器。AI 包括 6 大领域^[1]：机器人、计算机视觉、自然语言处理、认知与推理、机器学习、博弈与伦理，广泛应用于社交、医疗、商业等领域。支撑 AI 崛起的 3 大支柱条件：数据、算力和算法。

BP 公司与 IBM 公司联合研发了下套管卡管智能预测系统，预测精度达到 85%。沙特阿美公司利用决策树和 k -邻近算法预测井涌，预测精度达到 90% 以上^[2]。Oceanit 和 Shell 公司合作研发了智能定向钻井系统，利用 14 口井的定向钻井数据进行强化学习训练，基于当前施工参数预测未来的压差和扭矩，预测误差分别为 0.21% 和 2.72%，利用该系统可实现高效定向钻进^[3]。Aramco 公司利用人工智能方法计算基于时间的钻井液安全密度窗口，与传统的物理解析算法相比，计算速度大幅提高^[4]。国内管志川等人^[5]采用神经网络方法研究钻井过程中的井下复杂情况，对漏涌卡等风险进行识别和预测。殷志明等人^[6]采用机器学习的方式研究了钻速预测过程中的离群点问题，提出了一种融合的离群点检测算法。Yin Qishuai 等人^[7]通过应用大数据挖掘技术研究实时录井数据识别钻井状态，分析钻井非生产时间，形成提高钻井效率的方案。

近年来，国外油公司或油服公司纷纷与人工智能研究机构进行强强联合，加快油气领域人工智能应用研究，比如壳牌、斯伦贝谢均与微软合作，道达尔与谷歌合作，威德福与英特尔合作，KBR 与 IBM 合作。同时，各大公司认识到数据共享平台及开放性人工智能创新环境的重要性，斯伦贝谢推出的 DELFI 云平台和哈里伯顿推出的 DecisionSpace® 365 云平台，都是具备全业务链数据管理、大数据及支撑人工智能研究的开放式云平台。国内中国石油集团发布了“梦想云”，其包括云平台、数据湖、数据中台、应用商店、人工智能应用等关键技术，为加快包括钻井在内的油气人工智能技术研究奠定了基础^[8]。中国石化也着手打造石化智云平台，并开展智能油气田试点建设，但支撑钻井大数据及人工智能研究的基础仍然薄弱。

2 钻井领域 AI 技术方法论

工业界 AI 不同于人机博弈和商业营销，仅仅靠简单的数据驱动是行不通的，而必须从实际问题出发，用领域知识来约束人工智能建模^[9]。钻井是一项复杂的工程，必须找到合适的场景、合适的的数据，并将行业知识与 AI 算法相融合，才能将 AI 技术真正应用与钻井领域。

2.1 AI 建模基础

大数据是 AI 的基石，进行 AI 建模，必须确保有可用的数据。可用的数据应包含三大特征：数据量要大，只有原始数据量大，才有可能构建足够的样本用于训练；数据要准确，利用错误数据和假数据无法构建正确的关联关系；数据维度全面，即数据项必须涵盖应用场景所需的主要特征量。

钻井领域 AI 技术不仅依赖各类工程数据，还依赖于井下环境数据，即岩石与地层特征数据。工程类数据包括井眼轨迹、井身结构、钻具组合、钻头、钻井液体系、钻井液性能、实时综合录井（包

括工程参数、钻井液参数、气测参数)、岩屑录井、测井、故障复杂情况记录、钻时等数据,上述部分数据为时间序列,部分为深度序列,部分兼而有之,不同场合都可以用。此外,还可包括一些计算数据,如环空压力、摩阻扭矩、机械比能等。岩石与地层特征数据包括岩石力学参数、岩石产状、岩石物理性质、地层三压力、地层岩性等,这些参数大多是基于测录井数据计算或解释,进而通过试验或实测校验得到的,而测录井数据的解释可以利用 AI 技术。

2.2 AI 模型应用场景

美国的李杰等人提出智能制造系统需要从以下 3 个方向开展研究^[10]:第 1 个方向是在解决可见问题的过程中积累经验和知识,从而规避这些问题;第 2 个方向则需要依靠数据去分析问题产生的隐性线索、关联性和根本原因,利用预测分析将不可见问题显性化,进而解决不可见问题;第 3 个方向是通过对知识的深度挖掘,建立知识和问题之间的相关性,从旧知识中产生新知识,并利用新知识精确建模,从设计层面就能够避免可见和不可见问题的发生。参考这一思路,结合钻井工程业务特点,大致可以梳理出钻井领域 AI 应用场景,如表 1 所示

表 1 钻井领域 AI 应用场景及解决的问题

Table 1 AI application scenarios and problems to be solved in drilling engineering

序号	业务点	应用场景	解决的具体问题
1	机械钻速预测	工程设计	结合大量邻井实钻数据、新井地质与工程设计方案,预测每个井段的机械钻速,进而预测钻井周期、钻井进度及成本
2	机械钻速预测	施工(实时)	结合大量邻井实钻数据以及当前正钻井具体参数,预测当前或下一井段的机械钻速,为优化钻压、转速等工程参数提供依据
3	地层三压力预测	工程设计	根据邻井测录井数据、压力测试数据及研究成果,进行目标井地层三压力预测
4	岩石力学参数预测	工程设计	结合邻井测录井数据、岩石力学试验数据及研究成果,预测目标井的各岩层的岩石可钻性、弹性模量、黏聚力、泊松比等
5	钻头优选	工程设计	结合当前井的地质特征、区域钻井大数据,优选最适合当前地层的钻头
6	钻井液优选	工程设计	结合当前地质环境、区域钻井大数据推荐最适合当前地层的钻井液体系及性能参数
7	井下工况识别	施工(实时)	从大量历史实时数据中,分析不同工况下的参数变化规律,进而实时判别井下工况(钻进、起钻、下钻、循环等),进而实时分析钻井时效,为故障预警奠定基础
8	卡钻预警	施工(实时)	结合卡钻历史案例数据、当前工程参数变化趋势进行压差卡钻、坍塌卡钻、缩径卡钻的征兆预警
9	井涌溢流预警	施工(实时)	结合井涌溢流历史案例数据、当前工程参数变化趋势进行井涌溢流的征兆预警
10	井漏预警	施工(实时)	结合井漏历史案例数据、当前工程参数变化趋势进行井漏的征兆预警
11	故障处理方案推荐	施工	当井下出现复杂情况或发生故障时,能根据大量历史事故处理案例和专家知识库,推荐最合适的解决方案
12	钻井工程参数推荐	施工(实时)	结合基于历史数据的钻速预测及风险预警模型,推荐低风险、高钻速的参数组合
13	钻头磨损监测	施工(实时)	结合历史数据中的工程参数及取出钻头磨损情况记录,对当前钻头进行磨损状况监测及预测,便于确定合理的起钻时间
14	岩屑浓度预测	施工(实时)	基于钻井工况和地质录井相关参数预测当前井筒中的岩屑分布
15	起下钻速度控制	施工(实时)	匹配最相似邻井历史案例,提取邻井无风险起下钻的最优速度,为司钻推荐合理的起下钻速度区间,实现安全提效
16	水力计算模型校正	施工(实时)	利用邻井数据学习获得修正系数,实时校准水力模型计算,提高计算实时响应速度(不进行稳压影响下的复杂迭代计算)
17	摩阻系数监测	施工(实时)	通过大量实时数据与大钩载荷之间的内在关系,预测不同井段的摩阻系数,解决摩阻影响因素复杂,难以建立物理计算模型的难题
18	地层岩性识别	施工(实时)	从历史数据中学习规律,利用随钻测录井数据,实时识别岩性,为现场钻井参数优化提供直接依据
19	设备故障预测	全周期	分析不同设备的历史运行数据,捕捉设备损坏前的状态特征,预测设

			备维护周期,以缩短停工时间
20	井场视频异常识别	施工(实时)	利用计算机视觉技术,结合钻井井场的人员活动特征,实现井场人员异常行为(如人员倒地、坠落、人员徘徊、敏感区域人员进入、井场周边异常人员等)的实时识别,为钻井 HSSE 管理提供高效到手段

2.3 AI 建模的支撑技术

2.3.1 大数据支撑平台

机器学习是基于数据的科学方法,因此首先要把各类数据、海量数据组织起来,并转换成 AI 算法可直接调用的数据集,常规数据管理系统无法完成这一工作。大数据支撑平台主要实现的功能包括:异构数据存储、数据预处理、数据聚合、数据过滤、数据变换、特征提取、特征衍生、特征降维、样本标记、数据服务等。由于涉及到异构数据、数据体量大,该平台不可避免要用到 ETL 数据提取工具、Hadoop 分布式存储与计算工具、Spark 大数据计算引擎等流行工具。

2.3.2 AI 建模算法及工具库

成熟的 AI 算法数量庞大,且仍在被不断研发中。表 2 列举了钻井工程领域可能用到的算法及其适用场景^[11-12]。

表 2 常用 AI 算法及其适用场景

Table 2 Common AI algorithms and the suitable scenarios

算法大类	算法名称	特点及适用场景
分类	支持向量机	数据特征较多、线性模型
	平均感知器	训练时间短、线性模型
	逻辑回归	训练时间短、线性模型
	贝叶斯点机	训练时间短、线性模型,允许缺失值,更适用于增量训练
	决策森林	训练时间短、精确度高
	决策丛林	训练时间短、精确度高、内存占用量小
	神经网络	精确度高、数据量大、训练时间较长
聚类	XGBoost	精确度高、训练时间短、允许存在缺失值
	层次聚类	训练时间短、数据量大
	K-means	精确度高、训练时间短,仅适用于数值型数据训练
	FCM	精确度高、训练时间短
回归	SOM 神经网络	大数据量、运行时间较长
	排序回归	对数据进行分类排序
	泊松回归	预测事件次数
	森林分位数回归	预测分布
	线性回归	训练时间短、线性模型
	贝叶斯线性回归	线性模型,训练数据量较少
	神经网络回归	精确度高、训练时间较长
	决策森林回归	精确度高、训练时间短
时序序列	XGBoost 回归	精确度高、训练时间短、内存占用较大
	自回归 (AR)	描述当前值与历史值之间的关系,用变量自身的历史时间数据对自身进行预测,必须满足平稳性的要求,适用于自相关系数拖尾、偏相关系数截尾的情形
	移动平均 (MA)	关注 AR 中的误差项的累加,消除预测中的随机波动,适用于自相关系数截尾、偏相关系数拖尾的情形
	自回归移动平均 (ARMA)	以上二者的结合,适用于自相关系数和偏相关系数均为拖尾的情形

我国在 AI 基础研究方面比较薄弱,目前尚未形成可用的算法平台。国内研究人员目前主要使用国外开源的算法库(已经把一些算法进行了程序封装,可直接调用),比较受欢迎的有 TensorFlow、Apache SystemML、Caffe、Apache Mahout、OpenNN、PyTorch、Neuroph、Deeplearning4j、Mycroft、

OpenCog 等, 每个算法库都有其优缺点, 需要甄别选用或组合使用^[13]。

2.4 “三轮驱动”方法论

钻井工程 AI 项目成功的关键在于 3 个条件: 有实际需求的业务场景、有满足条件的数据、有适用的技术工具 (即 AI 工具)。业务、数据、技术三者任何一项欠缺都不可能理想的结果, 好比三轮车的 3 个轮子, 任何 1 个轮子有缺陷都影响行车速度和平稳性。在 AI 项目确立时应遵循“三轮驱动”方法 (见图 1), 优先选择三者交集中心区的项目。

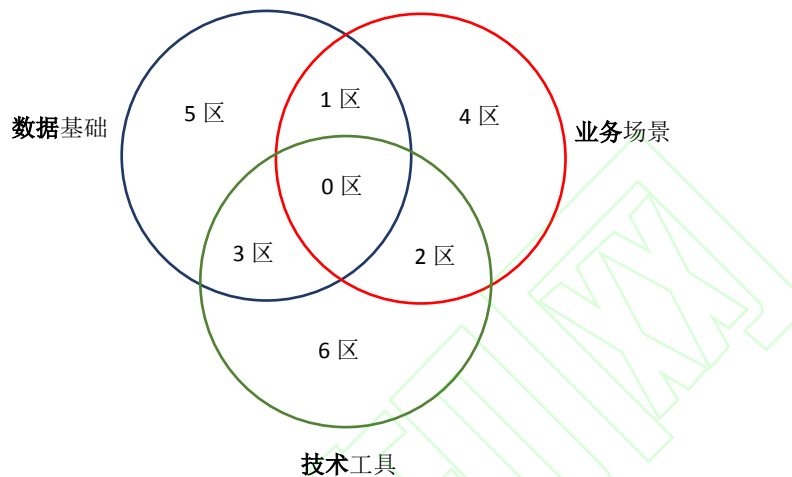


图1 “三轮驱动”方法论
Fig.1 Three-wheel Drive Methodology

实际研究工作中, 分别从业务、数据、技术三方面进行现状梳理, 针对每种组合情形, 分别归入图 1 的 7 个区: 0 区代表三者兼备, 具备开展 AI 建模的条件, 可启动应用研究工作; 1 区、2 区、3 区代表具备了其中 2 项条件, 可进一步探索另一条件的可行性, 比如搜集必须的数据、挖掘业务需求, 或者寻找适合的 AI 方法, 这些情形可通过努力逐渐向 0 区过度; 4 区、5 区、6 区是极不成熟的情形, 建议暂时放弃。

对于 0 区的情形, 还可对每个条件进行评估: 业务场景方面, 可从需求迫切程度、预期应用广度、潜在社会经济效益等方面量化评价; 数据基础方面, 可从数据完整性、数据量、准确率、可标记率等维度量化评价; 技术工具方面, 可从算法丰富程度、该方法在其他行业应用成熟度、样本需求量、可迁移性、算法工具获取难易程度等角度量化评价。最终从中选出三者均为高分的项目开展研究, 快速应用。

3 钻井领域 AI 技术应用研究实践

3.1 项目优选

以表 1 所列的业务场景为纲, 以“三轮驱动”方法为指导, 根据笔者所在团队的技术现状, 分别从 3 个方面进行评估, 开展了项目优选, 限于篇幅所限, 表 3 中列出了部分项目的评估情况。

表 3 钻井领域 AI 应用项目评价 (示例)
Table 3 Project evaluation (example) of AI application in drilling engineering

序号	项目名称	数据基础	AI 工具	综合评价
----	------	------	-------	------

1	机械钻速预测（设计）	历史井的地层岩性数据、岩石力学数据、地层压力数据、钻井实时参数、钻井液性能、钻具组合及钻头数据齐全，完全满足建模所需	神经网络、决策森林等回归工具	模型训练的3个条件均满足，模型应用时所需的数据也能提供，可为区域钻井时效分析及方案优化、新井周期预估提供新途径，预计效果优于传统手段。推荐！
2	参数优化与钻头状态评价（实时）	历史数据同上，满足建模需求；模型应用时需要的钻井参数、钻时数据可实时获取	同上	模型训练所需的历史数据满足，应用所需的参数可实时采集，随钻过程中可以最优钻速为目标进行钻井参数优化分析，满足应用条件。推荐！
3	井下故障复杂诊断预警	综合录井实时数据容易获取，存量也大	支持向量机、神经网络、决策森林、XGBoost等分类工具	故障复杂发生前一般都有相关参数异常变化征兆，随钻实时数据可实时获取，技术可行性相对较弱，历史案例样本数量偏少，但对于预防与控制井下故障有着重要意义，现场急需
4	地层岩性识别	传统工作中，通过测井解释得到地层岩性，而测井数据及历史井岩性数据都比较齐全，实钻过程中也可以获取测井数据	神经网络、支持向量机等分类或聚类工具	3个条件均能较好的满足，且行业知识对于测井数据与岩性的关联有较好的认知，条件成熟
5	设备故障预测	大部分井场设备状态尚未实现实时监测，历史案例数据积累不足		方法可行但数据基础薄弱，难以完成模型训练。建议尽快规模化升级井场关键设备，实现数据采集，待条件成熟再启动AI项目

由表3可以看出，由于模型训练主要使用历史数据，而模型应用必须使用实际数据作为入口参数，模型训练和实际应用场景下的数据满足程度可能有差异，某些情况下即便模型可以训练出来，也很难在实际工作场景应用，比如随钻机械钻速预测。

3.2 实践案例

钻井工程领域亟待解决且适合应用人工智能技术的问题，包括地层特征识别、钻井参数优化、井眼轨道优化、井下故障复杂诊断预警等。下面以钻井井下风险实时识别为例，简要阐述人工智能建模及试验验证情况。

3.2.1 数据提取与预处理

从实时分析数据集和复杂情况案例知识库中自动提取钻井复杂情况样本，包括复杂情况类型、发生时间、井深、录井参数等。由于录井实时数据存在空值、异常、重复、噪声等情况，如不处理将对模型产生影响，因此对数据进行预处理，选取多种数据处理方法，开发数据提取和预处理模块。以钻井实时数据为例，最常见的问题是数据项值缺失和数据项值异常（如超出正常值范围）、数据记录重复，对于这些异常，一般采取数据项整列删除、数据记录整条删除、利用本列均值替换等方式进行处理，通过开发数据预处理软件模块（如图2所示），预先设置对每种异常的处理方式，软件即可自动批量进行数据清洗，得到正常可用的数据，亦可在随钻过程中对动态采集的数据进行实时清洗，以满足实时智能预测的应用需求。

数据清洗界面																																																																																																								
请输入井号: 顺南蓬1井		开始清洗		停止清洗																																																																																																				
字段名	缺失值处理	异常值处理	归一化处理	该字段的上限值	该字段的下限值																																																																																																			
JH	整例删除	平均值修复																																																																																																						
SJ	整例删除	平均值修复																																																																																																						
DGFH	整例删除	平均值修复		2000	0																																																																																																			
ZY	整例删除	平均值修复		100	0																																																																																																			
ZPZS	整例删除	平均值修复		80	0																																																																																																			
ZFNI	整例删除	平均值修复		100	0																																																																																																			
LGYL	整例删除	平均值修复		50	0																																																																																																			
TY	整例删除	平均值修复		50	0																																																																																																			
<table border="1"> <thead> <tr> <th>JH</th><th>SJ</th><th>JS</th><th>CZJS</th><th>CDJS</th><th>ZTWZ</th><th>DGWZ</th><th>DGFH</th><th>ZY</th><th>ZPZS</th><th>ZFNI</th><th>LGYL</th><th>TY</th><th>ZS</th><th>CTJ1</th> </tr> </thead> <tbody> <tr> <td>顺南蓬1井</td><td>2018/10/23 9:09:41</td><td>6639.81</td><td>6639.81</td><td>6639.81</td><td>6508.87</td><td>24.95</td><td>1000</td><td>0</td><td>0</td><td>0</td><td>15.98</td><td>0</td><td>14.73</td><td>8.71</td> </tr> <tr> <td>顺南蓬1井</td><td>2018/10/23 9:09:51</td><td>6639.81</td><td>6639.81</td><td>6639.81</td><td>6508.87</td><td>24.95</td><td>1000</td><td>0</td><td>0</td><td>0</td><td>15.87</td><td>0</td><td>14.73</td><td>8.71</td> </tr> <tr> <td>顺南蓬1井</td><td>2018/10/23 9:09:51</td><td>6639.81</td><td>6639.81</td><td>6639.81</td><td>6508.87</td><td>24.95</td><td>1000</td><td>0</td><td>0</td><td>0</td><td>15.87</td><td>0</td><td>14.73</td><td>8.71</td> </tr> <tr> <td>顺南蓬1井</td><td>2018/10/23 9:09:51</td><td>6639.81</td><td>6639.81</td><td>6639.81</td><td>6508.87</td><td>24.95</td><td>1000</td><td>0</td><td>0</td><td>0</td><td>15.87</td><td>0</td><td>14.73</td><td>8.71</td> </tr> <tr> <td>顺南蓬1井</td><td>2018/10/23 9:10:16</td><td>6639.81</td><td>6639.81</td><td>6639.81</td><td>6508.87</td><td>24.95</td><td>1000</td><td>0</td><td>0</td><td>0</td><td>15.70</td><td>0</td><td>14.73</td><td>8.72</td> </tr> </tbody> </table>															JH	SJ	JS	CZJS	CDJS	ZTWZ	DGWZ	DGFH	ZY	ZPZS	ZFNI	LGYL	TY	ZS	CTJ1	顺南蓬1井	2018/10/23 9:09:41	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.98	0	14.73	8.71	顺南蓬1井	2018/10/23 9:09:51	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.87	0	14.73	8.71	顺南蓬1井	2018/10/23 9:09:51	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.87	0	14.73	8.71	顺南蓬1井	2018/10/23 9:09:51	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.87	0	14.73	8.71	顺南蓬1井	2018/10/23 9:10:16	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.70	0	14.73	8.72
JH	SJ	JS	CZJS	CDJS	ZTWZ	DGWZ	DGFH	ZY	ZPZS	ZFNI	LGYL	TY	ZS	CTJ1																																																																																										
顺南蓬1井	2018/10/23 9:09:41	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.98	0	14.73	8.71																																																																																										
顺南蓬1井	2018/10/23 9:09:51	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.87	0	14.73	8.71																																																																																										
顺南蓬1井	2018/10/23 9:09:51	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.87	0	14.73	8.71																																																																																										
顺南蓬1井	2018/10/23 9:09:51	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.87	0	14.73	8.71																																																																																										
顺南蓬1井	2018/10/23 9:10:16	6639.81	6639.81	6639.81	6508.87	24.95	1000	0	0	0	15.70	0	14.73	8.72																																																																																										
查看清洗结果																																																																																																								

图2 数据预处理模块

Fig.2 Module of Data Preprocessing

3.2.2 业务标签

基于业务专家理论及油田历史井实钻经验,按照正常作业、故障征兆、故障发生等作业状态,对实钻录井时间域原始数据划分时间段,如以3 min为时间窗格,通过专家分析对井漏前、发生井漏和井漏后的录井数据进行打标签,分别定义为非井漏样本、井漏样本和非井漏样本(见图3)。

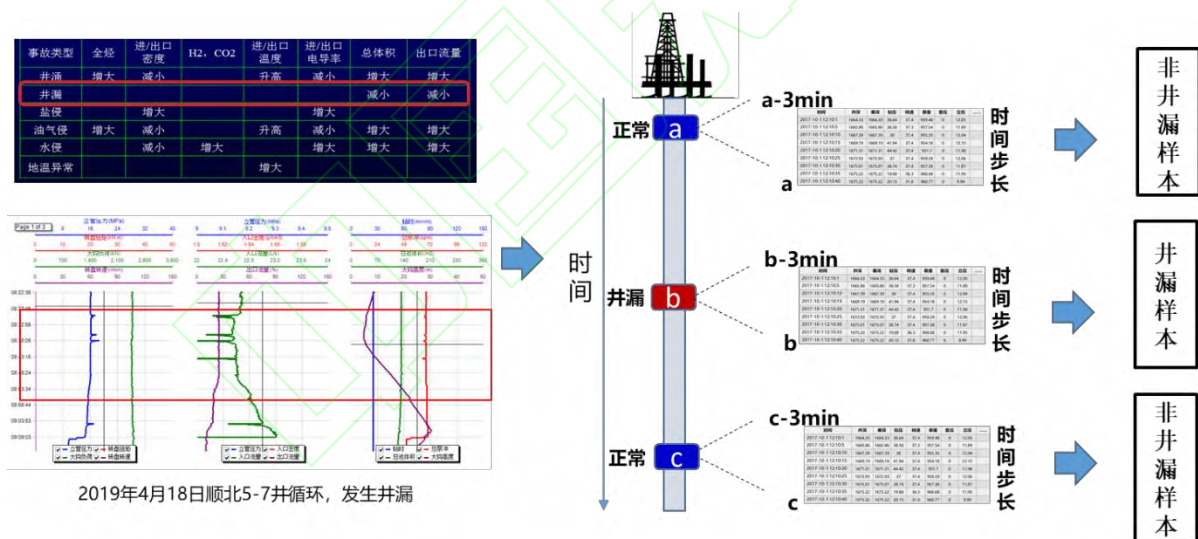


图3 专家业务标签逻辑

Fig.3 Expert Business Tag Logic

3.2.3 特征选取

根据数据相关性分析结果,确定不同钻井复杂情况的特征参数。以井漏为例,对入口流量、出口流量、钻井液池内钻井液体积、立管压力等参数的原始数据,提取绝对值、变化率、偏度、峰度、相关性等特征,找到井漏与非井漏样本之间数据变化的差异性,通过斯皮尔曼相关性分析方法,分析特征与井漏的相关系数,形成井漏识别模型训练的特征选取逻辑(见图4),最终确定9个特征参数,即:工况、入口流量、出口流量、计量罐内钻井液体积变化、钻井液总体积变化、立管压力、

立管压力变化率、套管压力、套管压力变化率。

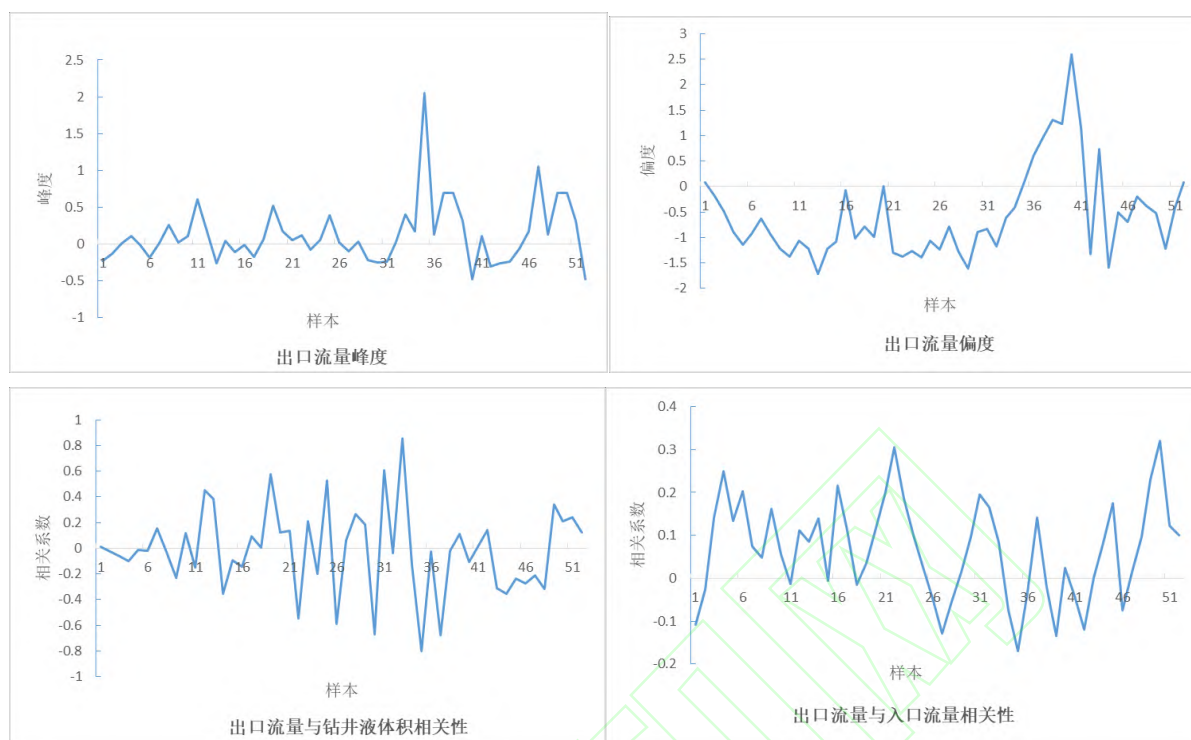


图4 井漏特征选取逻辑
Fig.4 Well Leakage Feature Selection Logic

再以卡钻为例,卡钻发生前的参数变化征兆为大钩载荷、立压、扭矩及转速异常波动,经过对大量案例分析发现,卡钻发生过程的参数变化呈现随时间变化的规律,因此引入了平均值和峰值作为新的特征参数。此外,由于每个参数都存在一些正常波动,每相邻两组数据(一般相隔5~10 s)的变化不能准确体现参数真正的变化趋势,而将数据按连续时间窗划分为处理单元可很好地解决这一问题。最终确定了14个特征参数(每3 min一个数据窗格作为计算单元):工况、大钩载荷、过去30 min大钩载荷的平均值、大钩载荷变化率峰值、立压、过去30 min立压的平均值、立压变化率峰值、扭矩、过去30 min扭矩的平均值、扭矩变化率峰值、钻压、入口流量、转速。

3.2.4 算法选取与模型构建

根据复杂情况的特征参数,将钻井复杂情况识别问题处理为二分类问题,即井漏和非井漏、溢流和非溢流、卡钻和非卡钻。大数据分类算法丰富,典型常用算法有决策树、随机森林、支持向量机等,目前应用较多的主要是以决策树为基础的集成算法,比如XGBoost就是其中效果最好的算法之一。该算法由GBDT(梯度提升迭代决策树)改进而来,是由多个基学习器共同构成的整体模型,既可以用于分类也可以用于回归问题中。利用XGBoost算法,依据混淆矩阵进行推算,验证模型准确率(1-误报)与召回率(1-漏报),通过反复调整模型框架参数(比如调整最大迭代次数(n),迭代步长(D)等),逐步得到最优模型。

在钻井井下故障复杂智能识别建模过程中,首先提取80%的案例数据集加载到算法中进行训练,然后用剩余的20%案例数据进行验证,然后局部更换数据集中的案例,重复上述过程,通过验证识别结果与实际案例的吻合度,促进模型自学习自提升。利用某地区57口井的114个复杂案例作为基

本数据样本集, 附加 100 个正常钻进的数据样本, 共同构建了模型训练样本集。研究发现, 基于 XGBoost 训练的模型, 井漏识别准确率 80%, 溢流识别准确率 82.7%, 卡钻识别准确率 71.2%, 与经典 GBDT 相比, 识别准确率和计算速度上明显提高。此外, 该结果也比笔者此前基于传统方法建模的识别准确率有所提高, 相信随着案例样本的不断积累和增加, 该模型的识别能力会进一步提高。

4 存在问题与发展建议

4.1 存在的问题

1) 钻井工程的 AI 应用场景很多, 但国内石油企业“数据孤岛”问题较严重, 而且钻井历史数据的标准化、完整性、正确性欠缺, 数据的自动化采集程度较低(数据频度、准确性和时效性不足), 较难构建出针对每个应用场景的高质量训练样本。

2) 开展大数据及人工智能应用研究工作离不开大数据相关技术的支撑, 传统手段难以支撑海量数据的抽取、处理、转换、服务, 而针对钻井工程大数据的支撑平台刚刚起步。

4.2 发展建议

1) 建立集团级数据共享中心, 加快数据治理, 打破油田分公司之间、专业板块之间的信息壁垒, 提高数据质量, 将科研团体的研发能力同生产企业的数据库优势结合起来, 共同促成企业数字转型和钻井的智能化。

2) 借鉴国外经验, 同知名的 AI 专业研究机构合作, 首先构建一体化大数据及人工智能研究支撑平台, 针对钻速预测及参数优化、井眼轨道优化、地层特征识别等场景, 优选或自研人工智能算法, 选择复杂疑难的钻井工区, 加强特征工程研究及模型验证, 注重现场可实施性。

3) 进行井场设备的数字化升级换代, 大幅提高现场数据的自动化采集程度, 为钻井 AI 技术发展积累高质量的“原材料”, 促进钻井自动化智能化技术快速发展。

5 结束语

随着人工智能算法的不断丰富和石油行业数据积累, 人工智能方法为钻井技术提供了新的有效途径和重要补充, 特别是在利用传统试验方法难以构建准确物理模型的场景, 人工智能方法将大有作为。在大数据及人工智能技术的推动下, 钻井行业正在经历一场数字化转型革命。钻井人工智能应用场景很多, 每个场景所需的基础条件和算法各不相同, “三轮驱动”的方法论可在一定程度上帮助研究者走出“无从下手”的困境, 能够快速评估优选出开展实践的智能应用场景, 并按照规范的步骤开展智能模型构建与应用试验, 从整体上促进钻井技术转型升级。

参考文献

Reference

- [1] 李开复, 王咏刚. 人工智能[M]. 北京: 文化发展出版社, 2017: 188-290.
LEE Kai-Fu, WANG Yonggang. Artificial intelligence[M]. Beijing: Cultural Development Press, 2017
- [2] ALOUHALI R, ALJUBRAN M, GHARBI S, et al. drilling through data: automated kick detection using data mining[R]. SPE 193687, 2018.
- [3] POLLOCK J, STOECKER-SYLVIA Z, VEEDU V, et al. Machine learning for improved directional drilling[R]. OTC 28633, 2018.
- [4] PHAN D T, LIU Chao, ALTAMMAR M J, et al. Application of artificial intelligence to predict time-Ydependent safe mud weight windows for

inclined wellbores[R].IPTC 19900, 2020.

- [5] 管志川, 胜亚楠, 许玉强, 等. 基于 PSO 优化 BP 神经网络的钻井动态风险评估方法[J]. 中国安全生产科学技术, 2017, 13 (8): 5-11.
GUAN Zhichuan, SHENG Ya'nan, XU Yuqiang, et al. Dynamic risk assessment method of drilling based on PSO optimized BP neural network[J]. Journal of Safety Science and Technology, 2017, 13(8):5-11.
- [6] 殷志明, 刘书杰, 谭扬, 等. 基于机器学习的深水钻井大数据处理方法研究[J]. 海洋工程装备与技术, 2019, 6 (11): 446-453.
YIN Zhiming, LIU Shujie, TAN Yang, et al. Research on outlier marking method of deepwater drilling big data in machine learning[J]. Ocean Engineering Equipment and Technology, 2019, 6(11): 446-453.
- [7] YIN Qishuai, YANG Jin, ZHOU Bo. et al. Improve the drilling operations efficiency by the big data mining of real-time logging[R]. SPE/IADC 189330, 2018.
- [8] 林伯韬, 郭建成. 人工智能在石油工业中的应用现状探讨[J]. 石油科学通报, 2019, 4 (4): 403-413.
LIN Botao, GUO Jiancheng. Discussion on current application of artificial intelligence in petroleum industry[J]. Petroleum Science Bulletin, 2019, 4(4): 403-413.
- [9] 李杰. 工业人工智能[M]. 刘宗长, 高虹安, 贾晓东, 译. 上海: 上海交通大学出版社, 2019: 29-113.
LEE Jay. Industrial artificial intelligence[M]. LIU Zongchang, GAO Hongan, JIA Xiaodong, Translated. Shanghai: Shanghai Jiao Tong University Press, 2019.
- [10] 李杰, 倪军, 王安正. 从大数据到智能制造[M]. 上海: 上海交通大学出版社, 2016: .
LEE Jay, NI Jun, WANG Anzheng. From BigData to Intelligent Manufacturing[M]. Shanghai: Shanghai Jiao Tong University Press, 2016.
- [11] CSDN. 人工智能算法总结[EB/OL]. [2019-12-26] <https://blog.csdn.net/JimBraddock/article/details/100376290>.
- [12] CSDN. 机器学习 (三) ——时间序列法[EB/OL]. [2019-12-26] <https://blog.csdn.net/livan1234/article/details/80863021>.
- [13] CSDN. 10 大热门人工智能开源工具 (框架) [EB/OL]. [2019-12-26] <https://blog.csdn.net/zhangbijun1230/article/details/80714032>.

[编辑 刘文臣]