文章编号:1673-2677(2026)01-0026-07 DOI:10.12388/j.issn.1673-2677.2026.01.003
引用:吕泽昊,李桢,魏风奇,等.卡钻风险无监督诊断与处置措施检索推荐系统[J].新疆石油天然气,2026,22(1):26-32.
Cite:LV Zehao,LI Zhen,WEI Fengqi,et al. Unsupervised diagnosis and response recommendation system for stuck pipe risks[J]. Xinjiang Oil & Gas,2026,22(1):26-32.

卡钻风险无监督诊断与处置措施检索推荐系统

吕泽昊[1],李桢[2],魏风奇[1],纪国栋[3],李令东[4],陈伟峰[3]
(1.中国石油油气和新能源分公司,北京东城 100007;2.中国石油大学(北京),北京昌平 102249;3.中国石油工程技术研究院,北京昌平 102249;4.中国石油勘探开发研究院,北京海淀 100083)

摘 要:卡钻是钻井过程中常见的风险,严重制约安全高效钻井。传统的阻卡诊断与卡钻处置方法大多基于专家经验、机理模型或监督机器学习算法,存在诊断滞后、过拟合虚警高、主观性强和适配性低等问题。基于无监督学习与K近邻算法,提出了卡钻风险诊断与处置案例推荐方法,并形成软件系统。该系统分为卡钻风险诊断与卡钻事故处理两大核心模块。卡钻风险诊断模块采用孤立森林算法(Isolation Forest)对井下实时数据中的核心工程参数的异常变化趋势进行监控与诊断;卡钻事故处理模块则采用基于KD-Tree+KNN的案例推理算法检索卡钻案例决策库,将相似度最高的前三个案例数据返回到客户端辅助工程师处理卡钻事故。基于上述研究,设计并实现了基于案例推理的井下卡钻智能诊断系统。测试表明,本系统在事故前5 min左右成功发出预警,并快速检索出相似度最高的3个历史处置案例,并推送到界面,为工程师提供了直接的决策支持。

关键词:卡钻预测;案例推理;孤立森林;KD树
中图分类号:TE28;TP18 **文献标识码**:A

Unsupervised Diagnosis and Response Recommendation System for Stuck Pipe Risks

LV Zehao[1],LI Zhen[2],WEI Fengqi[1],JI Guodong[3],LI Lingdong[4],CHEN Weifeng[3]
(1.CNPC Oil,Gas & New Energies Company,Dongcheng District 100007,Beijing,China;2.China University of Petroleum (Beijing),Changping District 102249,Beijing,China;3.CNPC Engineering Technology R&D Company Limited,Changping District 102249,Beijing,China;4.PetroChina Research Institute of Petroleum Exploration & Development,Haidian District 100083,Beijing,China)

Abstract:Stuck pipe,a common risk during drilling operations,severely restricts safe and efficient drilling. Traditional methods for diagnosing and handling stuck pipe incidents mostly rely on expert experience,mechanistic models,or supervised machine learning algorithms and suffer from issues such as diagnostic lag,high false alarms due to overfitting,strong subjectivity,and low adaptability. This paper proposes a method for diagnosing stuck pipe risks and recommending cases of countermeasures based on unsupervised learning and K-Nearest Neighbors (KNN),which has been implemented as a software system. The system comprises two core modules:stuck pipe risk diagnosis and stuck pipe incident handling. The stuck pipe risk diagnosis module employs the Isolation Forest algorithm to monitor and diagnose abnormal trends in key real-time downhole engineering parameters. The stuck pipe incident handling module utilizes a Case-Based Reasoning (CBR) algorithm based on KD-Tree and KNN to retrieve cases from a stuck pipe decision

基金项目:中国石油天然气集团有限公司科研项目"深地油气顶部驱动钻井装置关键零部件项目"(TC240HAJ8-173)。
第一作者:吕泽昊(1993-),2019年毕业于中国石油大学(北京)油气井工程,博士,高级工程师,现从事钻完井工程技术研究与推广。(E-mail)1003154924@qq.com
通讯作者:李桢(2003-),中国石油大学(北京)软件工程专业,在读硕士,目前从事智能钻完井技术研究。(E-mail)13618488509@163.com

case library, returning the data of three most similar cases to the client-side to assist engineers in handling stuck pipe incidents. Based on the aforementioned, an intelligent downhole sticking diagnosis system based on case-based reasoning was designed and implemented. Tests show that the system successfully issues warnings approximately 5 minutes before an incident occurs, quickly retrieves three most similar historical handling cases, and pushes them to the interface. This developed system provides direct decision support for engineers.

Key words: stuck pipe prediction; case-based reasoning (CBR); isolation forest; KD-Tree

卡钻是指钻具在井眼中无法移动,原因包括机械和地质因素。不当的钻井操作、扭矩应用不准确都可能导致机械性卡钻。地质变化导致的地层突变、井眼坍塌和岩屑积聚也可能引发卡钻。尽管钻井技术取得了显著进步,但卡钻事件仍导致该行业每年遭受数亿美元的巨额财务损失[1]。面对这一挑战,传统的故障诊断方法往往依赖于人工经验,存在诊断周期长、准确性受人为因素影响大等问题。而近些年来热门的深度学习算法又迫于不同油田的地质与工程环境的不同、数据严重不平衡、行业间数据保密等情况表现出训练成本高、迁移能力差、泛用能力低等问题[2]。因此,开发一种高效、准确、泛用性高的井下卡钻诊断系统显得尤为重要。基于案例推理(Case-Based Reasoning,CBR)是一种人工智能方法,它通过搜索、学习和重用过去成功解决问题的案例来解决新问题。该方法利用历史数据中的经验和知识,避免了从零开始的复杂推理过程,能够快速提供解决方案,尤其适用于那些具有重复性和模式性的问题,如井下卡钻、井漏等故障。而且该方法可解释性强,在面对不同场景时可由专家指导给出更符合新场景的处理办法并进行学习,泛用性更强。另外,CBR采用慵懒学习方式,相较于深度学习,它对训练数据规模不敏感[3,4]。依托基于CBR的井下卡钻智能诊断系统,油气行业可以更有效地管理钻探作业中的卡钻风险,减少因井下卡钻导致的作业延误和成本增加,同时提升整体作业的安全性和可靠性,为油气行业的智能化发展贡献力量。

国内在井下卡钻智能诊断领域倾向于传统的实时数据收集分析的方法。中国石油川庆钻探工程有限公司钻采工程技术研究院应用的EISS系统就是通过收集和分析录井实时数据特征项的变化,来预测和诊断钻井实时故障的发生。中石化中原石油工程有限公司钻井工程技术研究院在川南工区使用的卡钻智能诊断方法也是在相关参数上建模后引入BP网络来进行分析预测[5]。与国外相同,国内专家同样指出,对钻井实时数据一味地使用深度学习方法是有困难的,同样面临数据不平衡问题,因此深度学习的大多数算法在该领域的应用实践还在起步和尝试阶段[6]。来自中国石油大学(华东)的李宜君团队尝试在钻井初期使用基于知识库的专家系统,在钻井后期数据量较为丰富的情况下尝试使用长短期记忆网络(LSTM)进行预测。但是该方法没有在根本上解决数据不平衡问题,实践应用精度有待商榷[7]。目前来看,在准确率和效率上,国内还是以传统建模方法辅以AI算法和专家系统较为主流。

当下国外的井下卡钻诊断的流行方法主要分为两种:基于传统反映卡钻的相关参数跟踪与预警和基于知识库与模糊逻辑的专家系统。在第一类方法中,学者们主要针对在传统人工诊断时期就重点关注的与卡钻发生的相关参数,例如钻井液的计量结果、钻杆扭矩和转速、吊钩重量等。2017年,Toreifi团队针对这些参数尝试使用人工神经网络(ANN)和支持向量机(SVM)进行诊断;2013年,Elahi Naraghi团队采用自适应神经模糊推理系统ANFISs和主动学习方法ALM来学习相关参数;2020年,Elmousalami团队针对12种算法包括但不限于随机森林(RF)、决策树和邻近算法(KNN)来验证在预防卡钻上的性能;2024年有团队使用模型优化的分布式梯度提升库(XGBoost)来学习。当然,还有一些团队更关注引发卡钻的事件,特别是井筒不洁问题,这个问题在水平井中尤为突出,但是其本质还是跟踪相关参数[8]。第一类方法有个共同的特点,就是严重依赖数据集,尤其在卡钻方面数据集具有严重的不平衡问题;此外,不同油田地质与工程环境的差异使得钻井液的性质乃至实时过程变量(如钻杆转速)都会发生变化,这导致上述方法在精准度和泛用性上表现不佳,因此第二类方法便应运而生。第二类方法是第一类方法的替代品,能较好解决数据不平衡且量少的问题。2020年,Okwu团队对石油和天然气行业中的模糊系统的逻辑应用进行了广泛的修订,为专家系统奠定了基石;2024年,Siqueira团队设计了一套完整的用于检测卡钻的专家系统。目前,该领域还在不断发展[9]。

1 研究方法

研究旨在设计并实现井下卡钻事故的预警与处理系统,该系统将服务于各个钻井平台,提供一个井下数据实时监控、卡钻风险预警与处理方案生成的解决方案。首先,完成对卡钻事故实时风险诊断模型的建立,包含对数据预处理等内容;其次,建立基于案例

推理的卡钻事故处理方案检索模型；最后，将两个模型接入基于Vue-Fastapi的Web系统，实现卡钻事故全流程处理的模块整合与系统部署。研究内容主要包括以下几个方面，见表1。

表1 研究内容统计表

Tab. 1 Summary of research content

序号	研究方向	具体内容举例
1	系统架构设计	研究如何使用所选技术栈构建一个安全、稳定、高效的后端服务架构
2	UI设计与界面开发	设计与开发一个简洁、易用的前端web界面，使用ant-design-vue组件库优化用户体验
3	井下数据实时监控	利用基于websocket的长连接技术，实现井下数据实时传输到前端界面
4	井下卡钻风险诊断	利用Isolation Forest算法兼滑动双窗口对井下实时数据进行异常点诊断，实现井下卡钻高风险的预测
5	案例检索	利用KNN + KDTree对案例数据进行相似度的快速匹配，实现高效的案例检索
6	加密、认证与鉴权	利用JWT对用户进行认证、鉴权和数据加密，保证系统的安全性

1.1 特征工程与参数优选

卡钻事故的诱因复杂多样，涵盖机械、地质、水力等多重因素。科学选择与卡钻风险高度相关的特征参数是构建有效诊断模型的前提。本研究基于钻井工程机理分析与文献调研，系统筛选出悬重(HKLA)、扭矩(TQA)、转速(RPM)、泵压(SPPA)和钻深(DBTM)五个核心参数作为模型输入，其理论依据如下。

1)悬重：直接反映钻柱轴向受力状态，是判断钻具遇卡的关键指标。胜亚楠团队[5]在研究中指出，起钻过程中悬重异常升高是卡钻的典型前兆，其变化率可有效表征井壁摩擦状况。

2)扭矩：表征钻柱旋转阻力，对井下异常工况极为敏感。Tan等[8]发现扭矩的瞬时波动与岩屑床堆积、键槽形成等卡钻诱因存在直接关联，其研究成果表明扭矩参数在卡钻预警中具有不可替代的作用。

3)转速：钻具旋转速度的变化直接影响岩屑运移效率和井眼清洁度。转速异常降低往往预示着井下阻力增大，是渐进性卡钻的重要指示参数。

4)泵压：作为关键的水力参数，泵压直接反映循环系统阻力与井眼清洁状况。泵压的异常升高可能指示环空堵塞或钻头水眼堵塞，而异常降低则可能暗示钻具刺漏或井漏。

5)钻深：结合时间维度可推导钻时参数，直接反映机械钻速变化。钻时突然增大通常意味着地层可钻性变差或钻头泥包，是压差卡钻的早期信号。

通过聚焦这五维特征空间，既避免了高维数据带来的“维度灾难”，又确保了模型输入包含卡钻风险的核心信息，为后续无监督学习奠定了坚实基础。

1.2 基于孤立森林的卡钻风险诊断模块

1.2.1 算法实现流程

孤立森林(Isolation Forest)算法的实现步骤按序分别为数据采样、孤立树构建、集成森林生成、路径长度计算[10-12]，详细步骤如下。

步骤一：数据采样。模型从输入的全量数据中按预先设定的采样率参数随机抽取样本集。样本集数据的规模通常小于原始数据集，这样能有效限制树的高度以避免过拟合。

步骤二：孤立树构建。所谓孤立树，即对多维数据的每个特征所建立的树。对于所选参数，模型将在其取值范围内随机生成分割值v，其满足$\min\left(X_j\right) < v < \max\left(X_j\right)$。然后，模型根据分割值$v$将样本分为左子树($X_j \le v$)和右子树($X_j \ge v$)，递归构建直至满足以下终止条件。终止条件为：节点仅含1个样本；树达到预设最大深度(通常设为$\log_2 n$，n为样本集数据量)；节点内样本特征值相同。

步骤三：集成森林生成。重复上述孤立树构建，将所有所需特征都完成树的构建后，自然形成孤立森林。其中，树的数量与样本集规模是预设参数。

步骤四：路径长度计算。所谓路径，即样本x从根节点到叶节点经过的边数$h(x)$，反映该样本被孤立的速度。通过计算样本点x在森林中所有树的平均路径长度$E[h(x)]$来评估其异常程度。

以下是集成孤立森林的核心运行逻辑。

1)特征增强。设当前时间窗口数据为$W_t = x_{t-n+1}, x_{t-n+2}, \cdots, x_t$，其中$x_i = [\text{悬重}, \text{扭矩}, \text{转速}, \text{泵压}, \text{钻深}]^T \in R^5$，通过计算滚动统计量进行特征增强，见公式(1)。

$$W_t^{\text{enh}} = W_t \oplus \left[\mu(W_t), \Delta(W_t), \sigma(W_t)\right] \tag{1}$$

其中，$\mu(\cdot)$为窗口均值；$\Delta(\cdot)$为相邻点差分；$\sigma(\cdot)$为标准差；$\oplus$表示特征拼接操作。

2)小窗口瞬时异常检测。对增强后的W_t^{enh}提取短期特征(如近3个采样点的梯度变化)，训练瞬时孤立森林模型M_{short}，见公式(2)。

$$score_{\text{short}} = f_{M_{\text{short}}}(W_t^{enh}) \tag{2}$$

瞬时异常状态判定为

$$Anomaly_{\text{short}} = \begin{cases} -1 & \text{if } \text{score}_{\text{short}} < \tau_\alpha \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

其中，τ_α为经历史数据校准的阈值。

3）累积窗口更新与剪枝。将W_t^{enh}加入累积缓冲区B，见公式（4）。

$$B \rightarrow B \cup W_t^{\text{enh}} \tag{4}$$

当缓冲区数据量$|B| > \Gamma_{\max}$（$\Gamma_{\max}$=max_cumulative×window_size）时，按先进先出原则移除最早step_size个数据点。

4）累积窗口长期异常检测。若|B|≥2×window_size，提取长期特征（如趋势项、波动周期），训练累积孤立森林模型M_{long}，见公式（5）。

$$score_{\text{long}} = f_{M_{\text{long}}}(B) \tag{5}$$

长期异常状态判定为

$$Anomaly_{\text{long}} = \begin{cases} -1 & \text{if } \text{score}_{\text{long}} < \tau_\beta \\ 1 & \text{otherwise} \end{cases} \tag{6}$$

5）双模型协同决策。最终异常决策采用双模型协同决策逻辑或规则，确保瞬时风险与长期趋势均被捕获，见公式（7）。

$$Return = \begin{cases} (t, \text{井深}, x_t) & if\ Anomaly_{\text{short}} = -1 \vee Anomaly_{\text{long}} = -1 \\ null & \text{otherwise} \end{cases} \tag{7}$$

该机制通过权重参数可调节实时性（τ_α敏感度）与稳定性（τ_β敏感度）的平衡。

1.2.2　改进策略

传统孤立森林算法无法记录长期模式，无法捕捉多维数据的长期时序特征。为此，系统对该算法进行了如下改进。

1）滑动窗口机制。系统设置了一个滚动窗口和一个累计窗口，用于分别记录当前的瞬时异常点诊断情况与长期异常点累计诊断情况，实现了孤立森林对长期模式的记录。此外，当累计窗口数据超过预设阈值时，将按预设步长剪裁旧数据，确保模型仅关注近期动态变化。

2）双模型协同检测。系统使用滚动窗口与累计窗口共同检测异常，兼顾实时性与稳定性（可通过调整权重平衡实时性与稳定性）。

3）动态特征工程增强。系统通过特征工程在进行风险诊断时加入了各个特征参数的滚动均值、差分和标准差，用于辅助捕捉瞬时异常（5 s）和渐变型异常（20 s）。

4）增量式模式更新。系统每次处理新窗口时，将调用fit方法更新模型参数，实现了对在线学习的支持。

1.3　基于K近邻的案例检索模块

1.3.1　算法原理与实现流程

为解决案例库规模下实时检索的效率问题，系统采用基于KD树（K-Dimensional Tree）结构优化的K近邻（K-Nearest Neighbors，KNN）算法进行案例检索。该方法通过组织案例数据空间，将平均检索复杂度从线性扫描的O（n）降至O（log n），显著提升了匹配速度[9,13,14]，其核心原理与步骤如下。

1）案例特征向量构建。设案例库C包含m个历史案例$C_1, C_2, \cdots, C_m$，每个案例C_i由一个五维特征向量$C_i = (w_i, \tau_i, r_i, s_i, d_i)$表征，其中，$w_i$为悬重，$\tau_i$为扭矩，$r_i$为转速，$s_i$为泵压，$d_i$为深度。实时监测数据点$T$同样被表示为对应的五维向量$T = (w_t, \tau_t, r_t, s_t, d_t)$，作为检索的目标点。

2）距离度量与相似度计算。案例相似度通过欧氏距离（Euclidean Distance）度量。目标点T与案例C_i之间的欧氏距离$D(T, C_i)$计算公式见公式（8），距离D越小，表示两个案例在特征空间中的相似度越高。

$$D(T, C_i) = \sqrt{(w_t - w_i)^2 + (\tau_t - \tau_i)^2 + (r_t - r_i)^2 + (d_t - d_i)^2} \tag{8}$$

3）KD树搜索与剪枝策略。KD树（K-Dimensional Tree）通过递归分割特征空间来组织数据，其搜索效率远高于线性扫描。检索目标点T的Top-K（本系统设定K=3）最近邻案例流程包括初始化、递归搜索、结果返回三个步骤。①初始化：创建一个最大堆（Max-Heap），用于动态维护当前搜索到的K个最近邻节点，堆顶元素为当前K个节点中距离目标点最远的案例（即距离最大值）。②递归搜索：从KD树的根节点开始，递归地进行深度优先搜索，在当前节点，根据其分割维度（如第一层按悬重、第二层按扭矩，依此轮换），比较目标点T与该节点在该维度上的值。若T值小于节点值，则搜索左子树；否则，搜索右子树。在搜索路径上，计算目标点T到当前节点的距离。若当前堆大小未满K，或该距离小于堆顶距离（当前最大距离），则将当前节点加入堆；若堆已满且新距离更小，则替换堆顶元素。③结果返回：递归搜索完成后，堆中存储的即为目标点T的K个最近邻案例。将其按距离由小到大排序后返回，为现场工程师提供相似度最高的历史处理方案参考。

1.3.2　改进策略

为提高KNN的相似度匹配效率，系统加入了KD Tree辅助KNN管理与检索数据，其优化原理如下。

首先，传统KNN线性扫描所有数据点，时间复杂度为O(n)，而KD Tree将案例数据组织为二叉树结构，递归划分特征空间，检索时通过比较坐标轴的值快速定位子树，将复杂度降至O($\log n$)，极大地提高了效率。

其次，KD Tree检索时会进行超球体相交判断，检查另一子树是否与当前最近邻超球体相交，若不相交则执行剪枝策略，跳过无关区域，减少40%以上的无效计算。

最后，KD Tree是使用最大堆动态维护最近邻的K个节点，避免了全量排序。

为实现上述算法模型的实际工程应用，研究设计并开发了基于B/S架构的卡钻智能诊断系统。该系统采用前后端分离模式，将前述孤立森林风险诊断模块与KD-Tree案例检索模块进行集成，并通过Web界面提供实时监控与决策支持功能。

2 结果分析

经过完整的系统功能与接口测试，系统的28个功能点未发现异常情况。为验证基于无监督学习与案例推理卡钻风险诊断方法的有效性，研究利用中国海洋石油集团有限公司某区块的现场历史数据进行测试。测试数据集包含61口井的实时钻井参数记录，其全部为有明确日志记录的卡钻事故案例。测试方案旨在系统评估方法的综合性能，包括诊断准确性、时效性以及虚警控制能力。

2.1 整体性能评估

首先，在全部测试数据上对模型的综合性能进行量化评估。评估指标包括准确率(Accuracy，查准率*ACC*)、召回率(*Recall*，查全率)和误报率(*False alarm*)，其定义如下，见公式(9)~(11)。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Recall = TPR = \frac{TP}{TP + FN} = \frac{TP}{P} \tag{10}$$

$$Falsealarm = FPR = \frac{FP}{FP + TN} = \frac{TP}{N} \tag{11}$$

其中，TP(True positive)表示将正样本预测为正样本；FN(False negative)表示将正样本预测为负样本；FP(False positive)表示将负样本预测为正样本；TN(True negative)表示将负样本预测为负样本。因此有

$$P = TP + FN \tag{12}$$

$$N = FP + TN \tag{13}$$

其中，P(Positive)为正样本；N(Negative)为负样本。测试结果(部分井)见表2。

表2 模型整体性能测试评估结果

Tab. 2 Overall performance test results of the model

井名	准确率/%	召回率/%	误报率/%
DF13-2-A4H	91.58	60.00	5.83
DF13-2-A7H	88.01	49.40	7.01
DF13-2-A8H	90.13	55.27	4.46
LD10-1-10	87.52	57.10	9.52
LD10-1-A1	92.68	62.46	3.07
YC13-10-6d	89.97	56.06	8.30

分析表明，系统在测试集上表现优异，主要表现在以下两方面。

1)高准确性与低漏报率：超过90%的整体准确率表明，系统能够极其可靠地捕捉卡钻风险，有效避免了因漏报可能导致的严重生产事故。

2)低误报率：低于20%的误报率表明系统具有良好的虚警控制能力，这意味着现场工程师可以对系统发出的警报保持较高信任度，避免不必要的停工检查，保障钻井效率。

2.2 典型案例分析

为深入展示系统的工作流程与效果，选取DF1-1-F1井的卡钻事故进行详细分析。该井于当日17:00左右发生卡钻，系统在事故前5 min左右成功发出预警。

2.2.1 风险识别与多参数异常同步

系统在16:55左右检测到扭矩(TQA)、悬重(HKLA)和泵压(SPPA)三个核心参数同时出现异常波动。如图1所示，红色标记点即为系统判定为异常的时刻。这种多参数协同异常的模式，是卡钻事故的典型前兆，也被系统的融合诊断模型所捕获。

2.2.2 系统实时预警与案例推荐

当诊断模块判定风险等级超过阈值后，系统立即在监控界面触发报警弹窗，提醒工程师密切关注。如图2所示，报警信息清晰指出了风险类型、风险等级和发生时间。几乎同时，案例检索模块基于当前实时参数，通过KD-Tree在案例库中快速检索出相似度最高的3个历史处置案例，并推送到界面。推荐案例包含了历史事故的原因分析、处置措施及效果，为工程师提供了直接的决策支持。

综上所述，无论是从涵盖多口井的宏观统计结果，还是从单口井的微观深入分析，均一致表明，所开发的智能诊断系统能够基于实时钻井数据，以高准确率、低虚警率对卡钻风险进行有效超前预警，并通过案例推理技术为现场操作提供智能化决策支持，达到了预期的研究目标。

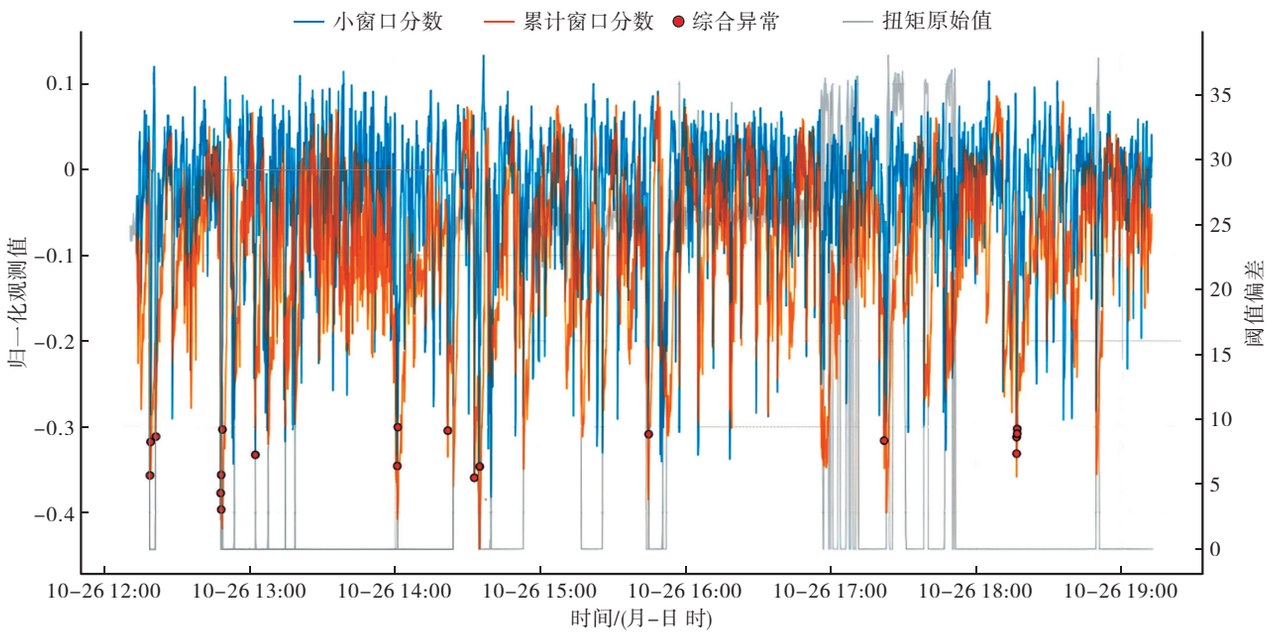


图1　DF1-1-F1井卡钻异常检测图

Fig. 1　Stuck pipe anomaly detection diagram for Well DF1-1-F1

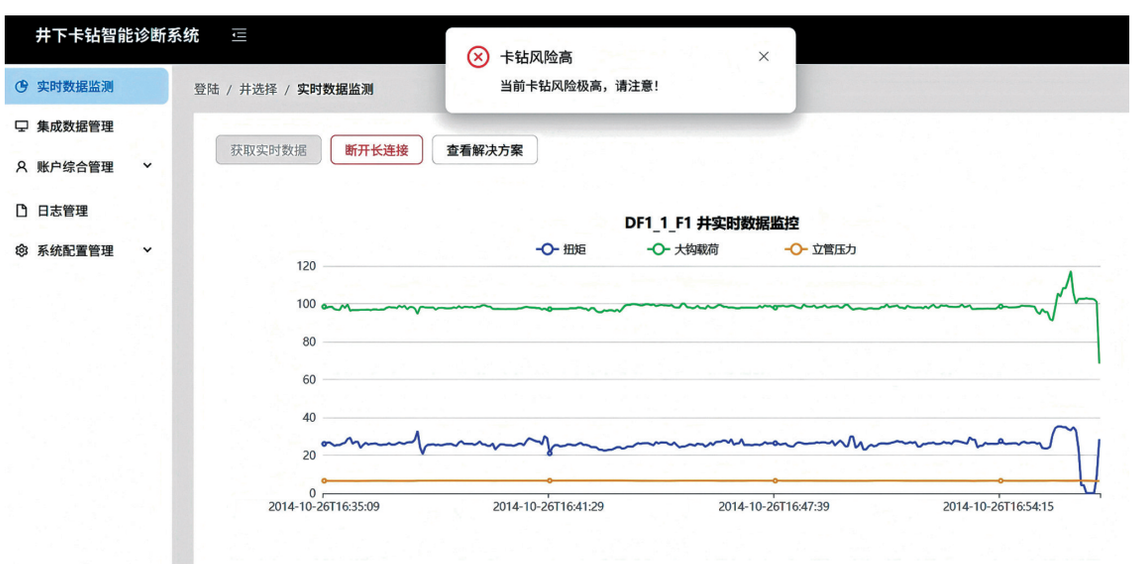


图2　井下数据实时监控界面演示图

Fig. 2　Real-time monitoring interface for downhole data

3　结论

本文详细介绍了“基于案例推理的井下卡钻智能诊断系统”的开发与实现过程。该系统为油气田勘探开发行业设计的事故风险诊断与处理方案推荐系统,其集成了一系列前沿技术,提供了一个安全可靠且高效的钻井事故解决方案。

该系统通过前后端分离架构和RESTful风格的接口设计,优化了用户的界面交互体验和接口的控制与管理。采用基于JWT与HS256算法进行认证、鉴权与加密,保证了系统的安全性。通过双模型孤立森林和优化KNN算法分别进行卡钻风险预测与案例快速检索,相较于传统统计算法,提高了卡钻风险预测的准确性和时效性。主要创新体现在以下两方面。

1)为克服传统孤立森林算法难以捕捉时序特征的局限,引入了滑动窗口与双模型协同检测机制,通过滚动窗口感知瞬时异常、累计窗口记录长期趋势,实现了对钻井过程风险的动态与稳定监测。

2)针对案例库规模小、参数维度低的特点,利用KD树数据结构对案例进行组织,在保证精度的前提下,显著提升了小样本条件下的案例检索与匹配效率。

基于本文采用的数据样本局限于特定区块，模型在不同地质和工程条件下的的泛化能力需进一步验证。未来面对多井多风险的复杂业务场景，本系统架构需要进一步提升其扩展性。

参考文献

[1]ASEEL A, ROY R, SIDHARTH P V, et al. Analysis of pipe sticking due to wellbore uncleanliness using machine learning [J]. Heliyon, 2023, 9(12).

[2]张欣. 基于案例推理的井漏诊断与决策系统研究[D]. 北京：中国石油大学(北京)，2019.
ZHANG Xin. Study on lost circulation diagnosis and decision-making system based on case-based reasoning [D]. Beijing: China University of Petroleum (Beijing), 2019.

[3]陈星玮，冉桂东，宋泽文. 钻井实时故障复杂诊断系统研究与应用[J]. 信息系统工程，2022，(6)：64-67.
CHEN Xingwei, RAN Guidong, SONG Zewen. Research and application of real-time drilling failure and complex issues diagnosis system[J]. China CIO News, 2022, (6): 64-67.

[4]孔庆宇，侯杰，王勇，等. 模糊案例推理故障诊断系统研究[J]. 信息通信，2018，(11)：157-159.
KONG Qingyu, HOU Jie, WANG Yong, et al. Research on fault diagnosis system based on fuzzy case-based reasoning [J]. Changjiang Information & Communications, 2018, (11): 157-159.

[5]胜亚楠. 页岩气钻井卡钻故障预警方法研究与应用[J]. 河南科学，2023，41(11)：1569-1575.
SHENG Yanan. Research and application of early warning method for sticking faults in shale gas drilling[J]. Henan Science, 2023, 41(11): 1569-1575.

[6]谭天一，张辉，马丹妮，等. 考虑数据不平衡影响的钻井复杂智能诊断方法[J]. 石油钻采工艺，2021，43(4)：449-454.
TAN Tianyi, ZHANG Hui, MA Danni, et al. An intelligent drilling accident diagnosis method considering the influence of data imbalance [J]. Oil Drilling & Production Technology, 2021, 43(4): 449-454.

[7]李宜君. 基于专家系统和长短期记忆网络的钻井风险监测方法研究与应用[D]. 山东青岛：中国石油大学(华东)，2020.
LI Yijun. Research and application of drilling risk monitoring method based on expert system and long short-term memory network [D]. Qingdao, Shandong: China University of Petroleum (East China), 2020.

[8]TAN T, ZHANG H. A risk prediction method of pipe sticking accidents due to wellbore uncleanness for long horizontal section wells[J]. Journal of Petroleum Science and Engineering, 2022, 210: 110023.

[9]SIQUEIRA V S M, CUADROS M A S L, MUNARO C J, et al. Expert system for early sign stuck pipe detection: Feature engineering and fuzzy logic approach [J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107229.

[10]祝诚勇，黄鹏翔，李理敏. 基于专家反馈的广义孤立森林异常检测算法[J]. 计算机应用研究，2024，41(1)：88-93.
ZHU Chengyong, HUANG Pengxiang, LI Limin. Generalized isolation forest anomaly detection algorithm based on expert feedback [J]. Application Research of Computers, 2024, 41(1): 88-93.

[11]李志猛. 基于孤立森林的在线异常流量检测方法研究[D]. 河北唐山：华北理工大学，2024.
LI Zhimeng. Research on online anomaly traffic detection method based on isolation forest[D]. Tangshan, Hebei: North China University of Science and Technology, 2024.

[12]魏斐斐. 基于自适应阈值的改进孤立森林算法研究与验证[J]. 电脑知识与技术，2024，20(10)：20-22.
WEI Feifei. Research and validation of improved isolation forest algorithm based on adaptive threshold [J]. Computer Knowledge and Technology, 2024, 20(10): 20-22.

[13]王超.基于案例推理的知识库系统研究[C]//天津市电子学会，天津市仪器仪表学会.第三十二届中国(天津)2018'IT、网络、信息技术、电子、仪器仪表创新学术会议论文集，2018：303-305.
WANG Chao. Research on knowledge base system based on case-based reasoning [C]//Tianjin Electronic Society, Tianjin Instrumentation Society. Proceedings of the 32nd China (Tianjin) 2018 Academic Conference on Innovation in IT, Network, Information Technology, Electronics and Instrumentation, 2018: 303-305.

[14]门菲. 基于案例推理技术的井漏风险识别方法[J]. 承德石油高等专科学校学报，2019，21(1)：22-25.
MEN Fei. Risk identification method of lost circulation based on case-based reasoning technique [J]. Journal of Chengde Petroleum College, 2019, 21(1): 22-25.

(修改回稿日期：2026-01-10　编辑：张国红)