中文引用格式:武胜男,张来斌,胡一鸣,等. 面向溢流风险的多分类智能识别模型[J]. 中国安全科学学报,2026,36(1):72-80.
英文引用格式:WU Shengnan, ZHANG Laibin, HU Yiming, et al. A multi-class intelligent identification model for kick risk [J]. China Safety Science Journal,2026,36(1):72-80.

面向溢流风险的多分类智能识别模型*

武胜男[1,2]教授, 张来斌[1,2]教授, 胡一鸣[1,2,3], 崔 蓉[1,2], 刘书杰[4], 殷志明[5]
(1 中国石油大学(北京) 安全与海洋工程学院,北京 102249;2 应急管理部 油气生产安全与应急技术重点实验室,北京 102249;3 中海油安全技术服务有限公司,天津 300450;4 中海油海南能源有限公司,海南 海口 570100; 5 中海油研究总院有限责任公司,北京 100028)

中图分类号:X937 **文献标志码**:A **DOI**: 10.16265/j.cnki.issn1003-3033.2026.01.0911
资助项目:国家重点研发计划项目(2022YFC2806504)。

【摘 要】 为提升钻井过程已发溢流风险的识别准确率,融合特征工程与机器学习方法,建立一种多类别溢流风险智能识别模型。首先,基于现场实测溢流数据,采用小波变换实现噪声抑制;其次,结合光滑样条函数提取关键参数的动态变化趋势,据此分析溢流特征参数的异常波动行为,建立低、中、高3类风险等级的判定准则,根据钻井数据的变化特征标记溢流风险;然后,引入麻雀搜索算法(SSA)优化极限学习机(ELM),构建基于改进型ELM(IELM)的多分类溢流风险智能识别模型;最后,所构建的风险数据集进行模型训练、调优与测试,验证模型的识别性能。结果表明:IELM模型在分类精度与判别稳定性方面均优于原始ELM及反向传播(BP)神经网络模型,能够更准确和高效地识别不同等级的溢流风险。

【关键词】 溢流风险; 识别模型; 改进极限学习机(IELM); 神经网络; 表征参数

A multi-class intelligent identification model for kick risk

WU Shengnan[1,2], **ZHANG Laibin**[1,2], **HU Yiming**[1,2,3], **CUI Rong**[1,2], **LIU Shujie**[4], **YIN Zhiming**[5]
(1 College of Safety and Ocean Engineering, China University of Petroleum (Beijing), Beijing 102249, China; 2 Key Laboratory of Oil and Gas Safety and Emergency Technology, Ministry of Emergency Management, Beijing 102249, China; 3 CNOOC Safety & Technology Services Co., Ltd., Tianjin 300450, China; 4 CNOOC Hainan Energy Co., Ltd., Haikou Hainan 570100, China; 5 CNOOC research Institute Co., Ltd., Beijing 100028, China)

Abstract: In order to improve the identification accuracy of kick risk during drilling, a multi-category kick risk intelligent identification model was proposed by integrating feature engineering and machine learning techniques. Firstly, a wavelet transform was employed to achieve noise suppression based on field-measured kick data. Secondly, the dynamic variation trends of key parameters were extracted using smooth spline functions, and the abnormal fluctuation behaviors of kick-related characteristic parameters were analyzed. Based on this, a three-level risk classification criterion (low, medium, and high) was established, and kick risks were labeled according to the variation features of drilling data. Then, the

* 文章编号:1003-3033(2026)01-0072-09; 收稿日期:2025-08-14; 修稿日期:2025-10-20

sparrow search algorithm (SSA) was introduced to optimize the extreme learning machine (ELM), and a multi-classification kick risk intelligent identification model based on IELM was constructed. Finally, the performance of the model was validated through training, tuning, and testing on the constructed risk dataset. The results show that the IELM model outperforms the original ELM and back-propagation (BP) neural network model in terms of classification accuracy and discrimination stability, and is capable of identifying different levels of kick risks more accurately and efficiently.

Keywords: overflow risk; identification model; improved extreme learning machine (IELM); neural network; characteristic parameters

0 引　言

目前,石油开采逐渐从浅层向深层、超深层[1]拓展。在石油开采中,井喷是最为严重的一种井控事故,溢流是井喷发展的早期标志。早期溢流监测方法主要包括地面监测和井下监测 2 大类;从监测手段来看,分为钻井液微流量监测、液面监测、综合录井参数监测、随钻压力监测等[2]。这些方法操作手段单一,容易受到钻井现场环境的干扰,导致溢流发现具有一定的滞后性[3],且一定程度上依赖于专家经验,识别结果具有很大的主观性。近年来,随着计算机技术的快速发展,基于录井数据与机器学习的溢流识别方法已成为该领域的热点。

当前研究主要围绕数据处理、模型结构优化和核心算法改进等方面展开。在数据处理方面,针对钻井数据不平衡、噪声与高维特性,学者们采用多种预处理策略。如张禾等[4]利用边界合成少数类过采样算法能够生成支持向量机(Support Vector Machine,SVM)分类边界样本的优点平衡样本数据集,构建了基于立管压力、全烃值和出口流量的识别模型;YANG Jin 等[5]采用主成分分析降维多源录井参数,并利用人工神经网络识别气侵溢流。在模型结构优化方面,通过构建更贴合数据特性的网络架构,以提升特征提取与模式识别的能力,如陈青等[6]综合运用归一化、高通滤波和局部加权线性回归进行数据预处理,并采用图卷积网络与双向循环神经网络提升识别性能;ZHANG Dezhi 等[7]通过改进门控循环单元,提出级联门控循环单元网络,实现了溢流早期识别,准确率提升 5.88%。在核心算法改进方面,部分工作侧重于针对性的调整与增强现有识别模型,以提升其性能与稳健性,如史肖燕等[8]选用随机森林集成学习模型,通过评估 14 种录井参数的特征重要性与分类决策,完成溢流与井漏的异常判断。

综上,钻井事件智能识别领域仍面临诸多挑战:①多数研究采用现场真实数据,但由于溢流发生时常伴随人为干预与控制,导致有效异常样本稀缺,数据质量参差不齐[9-10]。②钻井日志记录往往较为宽泛,直接根据现场日志判断溢流发生情况会导致判断不准确,所以数据预处理与高质量标注是溢流识别工作的关键步骤。③有研究指出[7],传统二元分类难以反映溢流风险状态,亟须探索能够输出溢流等级的标记与识别方法。

鉴于此,笔者拟提出一种基于特征工程的多元分类溢流风险标记方法,并将麻雀搜索法(Sparrow Search Algorithm, SSA)思维融入极限学习机(Extreme Learning Machine,ELM),建立改进的 ELM (Improved ELM,IELM)模型,智能识别溢流风险。

1 溢流识别机器学习算法与原理

1.1 溢流风险识别原理

溢流是指在钻井过程中,当井底压力 P_b 低于地层孔隙压力 P_p 时,即满足条件:

$$P_b < P_p \tag{1}$$

在正压差 $\Delta P = P_p - P_b > 0$ 的驱动下,地层流体开始侵入井筒,如果溢流不及时控制,可能会发展成井涌甚至井喷[11]。

而溢流风险识别可看作一个基于多变量时间序列的异常检测问题,多变量时序数据定义为:

$$\boldsymbol{X} = [x_1, x_2, \cdots, x_t] \in \mathbb{R}^{D\times C} \tag{2}$$

式中:$\boldsymbol{X}$ 为输入数据集;$\mathbb{R}$ 为实数集;D 为总时间步数;D 为特征数;x_t 为第 t 个观测时刻的数据。

溢流识别是通过机器学习算法训练一个函数 f: $\mathbb{R}^{D\times C} \to \mathbb{R}$,根据每个时刻输出异常判断溢流风险等级,标签向量 $\boldsymbol{y}_t \in \{1,2,3\}$,其中,1、2、3 分别表示当前时刻处于低、中、高溢流风险状态[12]。

1.2 ELM 算法

ELM 是一种单隐层前馈神经网络,适用于各种回归和分类问题,其优化目标是通过计算各节点权

值、偏置,使得输出数据 O_j 尽可能地接近数据标签 $\boldsymbol{T}_j$,如下式:

$$O_j = \sum_{i=1}^{L} \boldsymbol{\beta}_i f(\boldsymbol{W}_i X_j + b_i) = \boldsymbol{T}_j \tag{3}$$

式中:$\boldsymbol{W}_i$ 为输入层各节点到隐藏层第 i 节点的权值向量,$\boldsymbol{W}_i = [\omega_{1i}, \omega_{2i}, \cdots, \omega_{Ni}]$;$b_i$ 为隐藏层第 i 节点的偏置;$\boldsymbol{\beta}_i$ 为隐藏层第 i 节点到输出层各节点的权值向量,$\boldsymbol{\beta}_i = [\beta_{i1}, \beta_{i2}, \cdots, \beta_{iM}]$;$X_j$ 为输入数据;$f(x)$ 为隐藏层的激活函数,常见的激活函数有 Sigmoid 函数、tanh 函数等[13]。

将式(3)改写成矩阵的形式,见下式:

$$\boldsymbol{T} = \boldsymbol{H}\boldsymbol{\beta} \tag{4}$$

$$\boldsymbol{H} = \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1n} \\ H_{21} & H_{22} & \cdots & H_{2n} \\ \vdots & \vdots & & \vdots \\ H_{n1} & H_{n2} & \cdots & H_{nn} \end{bmatrix} \tag{5}$$

式中:$\boldsymbol{T}$ 为由所有输入数据(K 组)所对应的标签构成的矩阵,$\boldsymbol{T} = [\boldsymbol{T}_1^{\mathrm{T}}, \boldsymbol{T}_2^{\mathrm{T}}, \cdots, \boldsymbol{T}_K^{\mathrm{T}}]^{\mathrm{T}}$;$\boldsymbol{\beta}$ 为由所有隐藏层节点到输出层的权值向量所构成的矩阵,$\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_{\mathrm{L}}]^{\mathrm{T}}$;$\boldsymbol{H}$ 为由所有隐藏层节点经激活函数运算后输出的值所构成的矩阵,$H_{ij} = f(\boldsymbol{W}_{\mathrm{i}} \mathrm{X}_{\mathrm{j}} + \mathrm{b}_{\mathrm{i}})$。

ELM 算法先"随机生成"输入层到隐藏层的权值向量 $\boldsymbol{W}_i$ 和节点偏置 b_i,再通过矩阵运算和广义逆求解隐藏层到输出层的权值矩阵 $\boldsymbol{\beta}$,求解如下式:

$$\boldsymbol{\beta} = \boldsymbol{H}^{+} \boldsymbol{T} \tag{6}$$

式中 $\boldsymbol{H}^{+}$ 为 $\boldsymbol{H}$ 的广义逆矩阵,可通过满秩分解方法求出。

1.3 SSA

SSA 是一种受麻雀觅食和反捕食行为启发的群体智能优化算法。SSA 算法流程如图 1 所示。

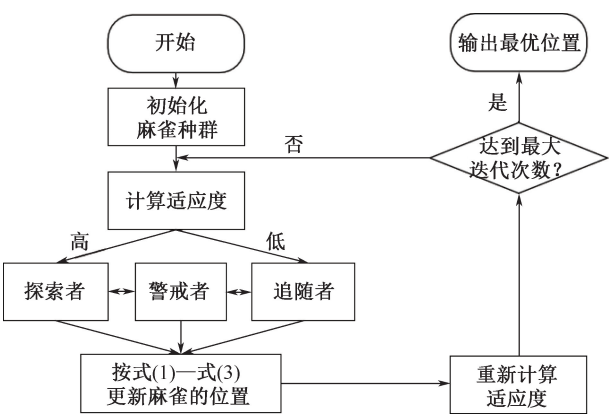


图 1 SSA 的流程

Fig. 1 Process of SSA

图 1 中,"探索者"位置变更如下式所示:

$$x_{i,j}^{q+1} = \begin{cases} x_{i,j}^{q} \exp(-i/\alpha \cdot q_{\max}) & \text{if} \quad R_2 < S \\ x_{i,j}^{q} + Q \cdot L & \text{if} \quad R_2 \geqslant S \end{cases} \tag{7}$$

式中:$x_{i,j}^{q}$ 为空间中第 i 只麻雀在第 j 维度上第 q 次迭代的位置;$q_{\max}$ 为所设置的最大迭代次数;α 为[0,1]的一个随机值;R_2 为随机的预警值,取值范围为[0,1];S 为所设置的危险预警阈值,取值范围为[0.5,1];Q 为满足正态分布的随机数;$\boldsymbol{L}$ 为大小为 d 的单位向量;d 为麻雀的维度大小。

"追随者"位置变更如下式:

$$x_{i,j}^{q+1} = \begin{cases} Q \exp(x_w^q - x_{i,j}^q / i^2) & \text{if} \quad i > \dfrac{n}{2} \\ x_b^{q+1} + | x_{i,j}^q - x_b^{q+1} | B^{+} L \end{cases} \tag{8}$$

式中:x_b^{q+1} 为适应度值最大的位置;x_w^q 为适应度值最小的位置;$\boldsymbol{B}$ 为全是元素 1 或-1 组成的大小为 d 的向量;$\boldsymbol{B}^{+}$满足关系式 $\boldsymbol{B}^{+} = \boldsymbol{B}^{\mathrm{T}}(\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}})^{-1}$。

"感知者"位置变更如下式:

$$x_{i,j}^{q+1} = \begin{cases} x_b^q + l | x_{i,j}^q - x_b^q | & \text{if} \quad f_i > f_g \\ x_{i,j}^q + k\left(\dfrac{| x_{i,j}^q - x_w^q |}{(f_i - f_w) + e}\right) & \text{if} \quad f_i = f_g \end{cases} \tag{9}$$

式中:l 为步长调节因子,服从于标准正态分布;k 为取值区间为[-1,1]随机值,表示麻雀个体移动的方向;e 为接近 0 的常数,防止分母出现 0;f_i 为当前迭代下第 i 只麻雀个体的适应度值;f_g 为当前迭代下麻雀种群整体的适应度最大值;f_w 为当前迭代下麻雀种群整体的适应度最小值。

2 溢流风险多分类智能识别模型

为提高溢流风险识别模型的分类精度与稳定性,建立基于 SSA 算法的 IELM 分类识别模型,并提出一种溢流数据分级与标记方法,如图 2 所示。

2.1 溢流数据分级与标记

根据现场日志,选择一段外界干扰少,并包含正常、异常情况的数据,2 种类型的数据量上尽可能均衡。选取立管压力、活动泥浆池体积、钻井液相对出口流量和气测值 4 种参数作为溢流关键表征参数,输入模型进行训练[14],使用 db4 小波将原始数据分解 4 层,并采用软阈值的方式进行降噪处理。

对于高压天然气井,溢流类型通常以气侵溢流为主,气侵发生时,气测值对气体侵入响应极为灵敏,数值在溢流初期迅速上升并持续维持于较高水平;活动泥浆池体积的无故增大也被工程实践广泛

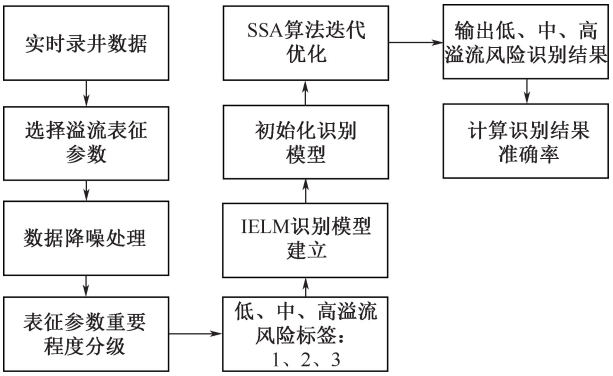


图 2　溢流识别模型建立流程

Fig. 2　Overflow recognition model development process

视作溢流的典型特征,部分研究[15]进一步将其累计增量作为判断溢流严重程度的依据;若钻井液相对入口流量保持稳定而钻井液相对出口流量上升,同样指示溢流发生[16],但仅在溢流量较大时才具备良好识别性[17];立管压力可以间接反映井筒流体状态,在溢流和井漏工况下均可能呈现下降趋势[18-19],因此,通常作为辅助判据。在对象为高压天然气井时,优先考虑参数对溢流的响应速度与敏感性,兼顾参数对溢流表征的确定性,并综合现有研究[2,5],总结 4 种表征参数的异常形式,并对参数的重要性进行排序,见表 1。

表 1　参数重要性排序

Table 1　Sorting of parameter importance

表征参数	参数异常形式	重要级别
气测值	数据呈现快速上升趋势、数据大于一定阈值、离群异常数据	Ⅰ-核心参数
活动泥浆池体积	数据呈现上升趋势、离群异常数据	Ⅱ-典型参数
钻井液相对出口流量	数据呈现上升趋势、离群异常数据	Ⅲ-重要参数
立管压力	数据呈现下降趋势、离群异常数据	Ⅳ-辅助参数

根据各参数是否表现溢流异常特征,以及各参数对溢流判断的重要级别,将正常数据与异常数据进一步划分为低、中、高溢流风险 3 种标签,划分规则见表 2。

立管压力与钻井液相对出口流量的变化同时受人为调控钻井泵的影响,为剔除钻井泵改变对立管压力所带来的线性影响,利用立管压力真实值减去钻井泵线性预测的立管压力值,同理对钻井液相对出口流量做同样处理,并采用光滑样条拟合表征参数。

最后构建数据集 $\boldsymbol{A}$,$\boldsymbol{A}$ 为一个 $m\times n$ 的矩阵,其中,m 为数据量,n 为特征参数和标签。

表 2　溢流分级方式

Table 2　Well kick classification method

溢流分级	划分规则	风险标签
低溢流风险	不多于 1 个非核心参数出现异常	1
中溢流风险	核心参数出现异常	2
	核心参数与典型参数不同时异常时 2 个参数出现异常	
高溢流风险	核心参数与典型参数同时出现异常	3
	3 个及以上参数出现异常	

2.2　基于 IELM 的溢流智能识别模型

由于在权重和偏置选择方面算法随机性的固有特性,ELM 算法的可控性和稳定性较差,不足以处理复杂数据。采用 IELM[20],通过将新型优化算法 SSA 与 ELM 相结合,并采用适用于分类任务的适应度函数,利用 SSA 的稳定性和快速性优势搜索 ELM 的最优权值及偏置,从而弥补其不稳定性和缺乏可控性的缺陷。IELM 算法流程如图 3 所示。

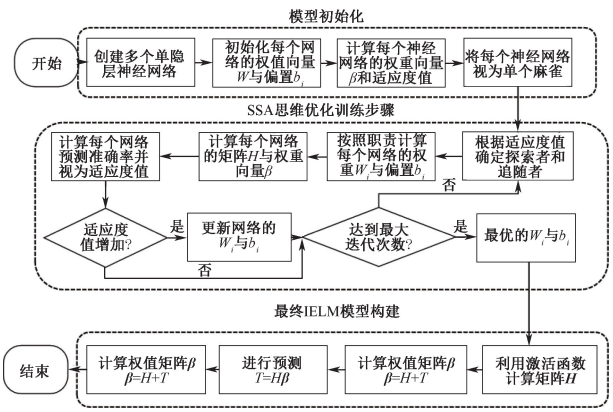


图 3　IELM 算法流程

Fig. 3　Process of IELM algorithm

具体流程如下:

1) 初始化。①随机生成多个单隐藏层神经网络的权重向量 $\boldsymbol{W}_i$ 和偏置 b_i;②计算每个网络的初始适应度(如预测误差);③根据适应度排序,选择性能较好的网络作为探索者,其余网络作为追随者。构建适用于分类任务的适应度函数,如下式[20]:

$$g(\boldsymbol{W},b)=(1-\omega)\frac{Z_t}{Z_t+F_t}+\omega\frac{Z_t}{Z_v+F_v}$$
$$\omega\in[0,1] \tag{10}$$

式中:ω 为对适应度函数所设置的权重参数;Z_t 和 F_t 分别为训练集中正确预测和错误预测的实例数量;Z_v 和 F_v 分别为验证集中正确预测和错误预测的实例数量。

2) 迭代优化(SSA 优化阶段)。重复以下步骤

直到达到最大迭代次数。①探索者更新,按式(7)调整 $\boldsymbol{W}_i$ 和 b_i;②追随者更新,向最佳网络的方向调整 $\boldsymbol{W}_i$ 和 b_i;③警戒者机制(随机选择部分网络),如果性能较差,则向最优解靠近;否则进行随机扰动;④适应度计算,更新每个网络的适应度;⑤保留最优解,仅保留优化后性能提升的网络。

3) 构建最终 IELM 模型。①选择适应度最高的 $\boldsymbol{W}_i$ 和 b_i;②用最优参数构建 ELM 网络,用于预测或分类。

3 高压天然气井溢流风险智能识别实例

3.1 研究对象介绍与数据处理标记

根据现场日志,对象井于持续钻进至 4 023 m 时机械钻速加快,后续在进行划眼操作的过程中气测值快速升高,井口返出由 19%上升至 24%,活动池泥浆增加 0.6 m^3,判断发生气侵溢流现象。截取正常钻进到发生溢流后的数据,同时剔除现场工人开始调整钻井液进行压井操作后的数据,最终获取现场有效数据 5 839 组,数据整体信息见表 3。

绘制钻井泵数据曲线如图 4 所示,由图 4 可知:现场曾多次调整钻井泵的数值,分析立管压力、钻井液相对出口流量这 2 种数据时,应尽可能降低人为调整钻井泵的影响。

以立管压力为例,绘制钻井泵-立管压力散点图,如图 5 所示,2 条直线分别对应划眼-钻进与划眼-循环工况(序号 1—2 678、2 679—5 839)。针对 2 工况分别以钻井泵为自变量、立管压力为因变量建立线性回归模型,决定系数 R^2 分别为 0.981 4 和 0.995 4,反映了钻井泵对立管压力的线性影响。

表 3 所选数据的信息

Table 3 Information of the selected data

数据序号	时间范围	井筒深度/m	钻头深度/m	工况
1—714	8:54:15—9:53:40	3 995.29~4 002.43	3 995.29~4 002.43	纯钻进
715—868	9:53:45—10:06:30	4 002.43	3 992.93~4 002.43	划眼
869—976	10:06:35—10:15:30	4 002.43~4 003.45	4 002.43~4 003.45	纯钻进
977—1 003	10:15:35—10:17:45	4 003.45	4 002.78~4 003.45	划眼
1 004—2 678	10:17:50—12:37:20	4 003.45~4 023.55	4 003.45~4 023.55	纯钻进
2 679—3 825	12:37:25—14:12:55	4 023.55	3 994.31~4 023.55	划眼(其间溢流)
3 826—5 839	14:13:00—17:00:55	4 023.55	4 020.93~4 021.41	循环(其间溢流)

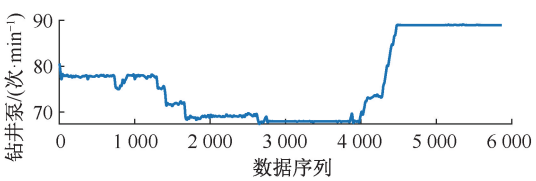


图 4 钻井泵数据曲线

Fig. 4 Mud pump data curve during drilling

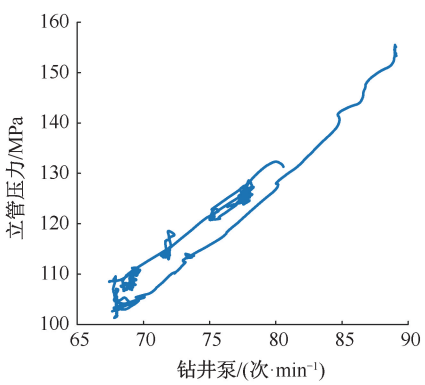


图 5 钻井泵-立管压力数据曲线

Fig. 5 Drilling pump-Standpipe pressure correlation curve

分别绘制钻进工况和循环工况的钻井泵-立管压力回归曲线,如图 6 所示。

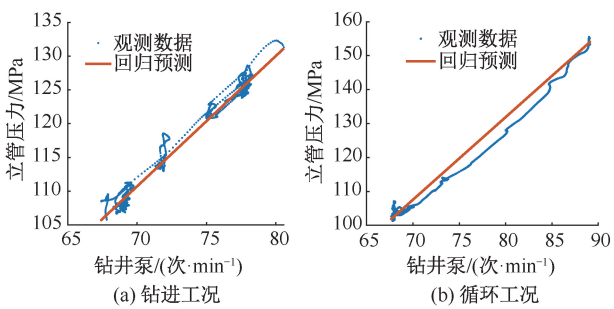


图 6 钻井泵-立管压力回归曲线

Fig. 6 Drilling pump-Standpipe pressure regression fitcurve

绘制残差曲线如图 7 所示。钻井液相对出口流量数据剔除钻井泵影响后的残差如图 8 所示,由图 8 可知:在钻井泵稳定阶段(图 4 中序列 2 971—3 471),残差与原始数据趋势一致,能保留立管压力下降及钻井液相对出口流量的波动特征;在异常增大(如序列 1 671—1 841)时也能有效反映。相比原始数据在钻井泵变化时的联动(如 1 271—1 671),残差数据减少了钻井泵变化带来的干扰。由此可

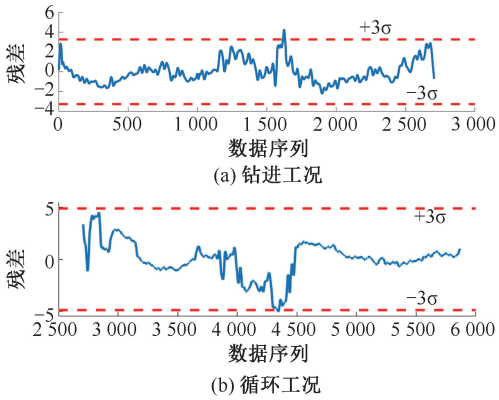


图 7　立管压力数据剔除钻井泵影响后的残差

Fig. 7　Residuals of standpipe pressure data after removing the influence of drilling pump

见:残差有效降低了钻井泵影响,保留了关键趋势,适用于后续立管压力、钻井液相对出口流量的异常分析。

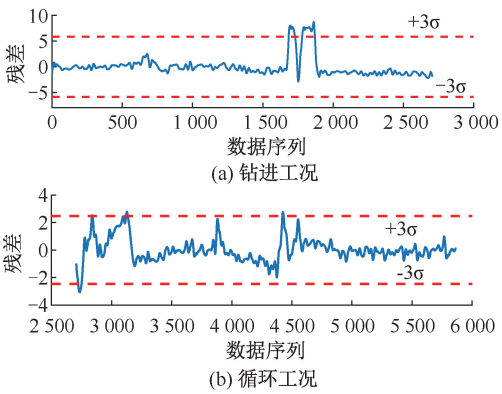


图 8　钻井液相对出口流量数据剔除钻井泵影响后的残差

Fig. 8　Residuals of drilling fluid relative outlet flow rate data after removing the influence of drilling pump

绘制钻井液相对出口流量残差如图 9 所示。

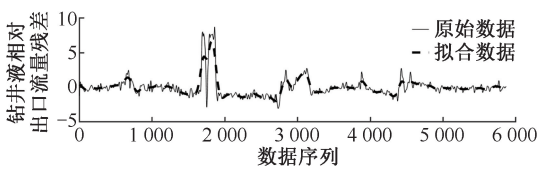


图 9　钻井液相对出口流量残差

Fig. 9　Residual of drilling fluid relative outlet flow rate

拟合钻井液相对出口流量残差数据并绘制其拟合曲线,如图 10 所示,光滑样条拟合所得到的曲线很好地保留了原始数据复杂的趋势,且曲线完全可导。

针对气测值数值过大的异常状况,结合日志记录,将阈值设置为 10%,并采用 3σ 原则辨别这部分离群异常数据。

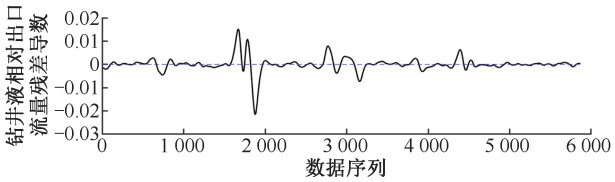


图 10　对钻井液相对出口流量残差数据拟合求导

Fig. 10　Derive by fitting the drilling fluid relative outlet flow rate residual data

根据表 3 分析数据异常并赋予标签,构建数据集 $\boldsymbol{A}$,$\boldsymbol{A}$ 为一个 5 839×5 的矩阵,前 4 列分别为立管压力、活动泥浆池体积、气测值、钻井液相对出口流量降噪后的数据,最后一列为数据所对应的标签。数据溢流风险标签如图 11 所示。

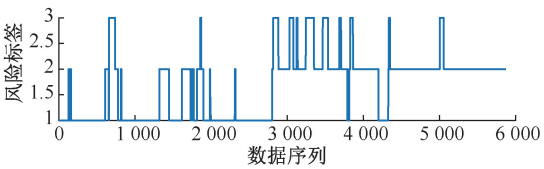


图 11　数据溢流风险标签

Fig. 11　Well kick risk label

3.2　模型训练与调试

3.2.1　ELM 模型的训练与调试

通过 Macro-average 方法计算 F_1 值,作为模型的评估指标[21]。以 880 个隐层节点作为最优超参数,ELM 模型在数据训练集 $\boldsymbol{A}_1$ 上重新训练,并识别 $\boldsymbol{A}_1$ 与测试集 $\boldsymbol{A}_2$ 的溢流风险,绘制混淆矩阵,如图 12 所示,计算模型准确率,即训练集为 79.81%,测试集为 80.14%。

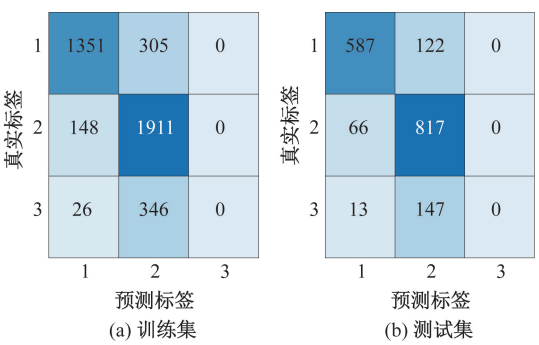


图 12　最优超参数下 ELM 模型溢流识别结果的混淆矩阵

Fig. 12　Confusion matrix of ELM model well kick identification results under the optimal hyper-parameters

3.2.2　IELM 模型的训练与调试

根据 ELM 参数调试结果,将 IELM 隐层节点数

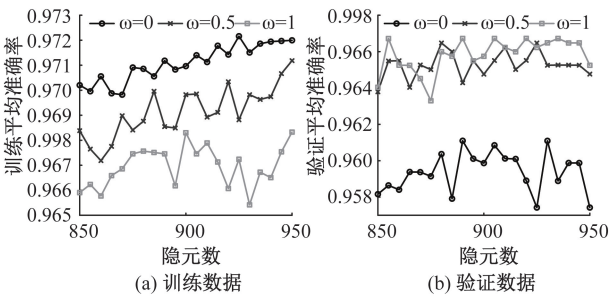
调至 850~950,步长为 5 进行网格搜索。适应度函数采用式(8),分别设置 ω 为 0、0.5、1 进行模型调试,超参数设置见表 4。

表 4 麻雀搜索的超参数设置

Table 4 Hyper-parameter settings in sparrow search

种群数量	迭代次数	发现者比例	感知者比例	预警阈值
100	20	0.5	0.2	0.7

使用训练集 $\boldsymbol{A}_1$ 在不同隐层节点数和权重 ω 下训练 IELM 模型,如图 13 所示,当 ω 为 0.5 时,验证准确率显著提升,综合考虑,最终选取 $\omega=0.5$,隐层节点数为 925,此时平均验证准确率达 96.88%,训练准确率为 96.65%。

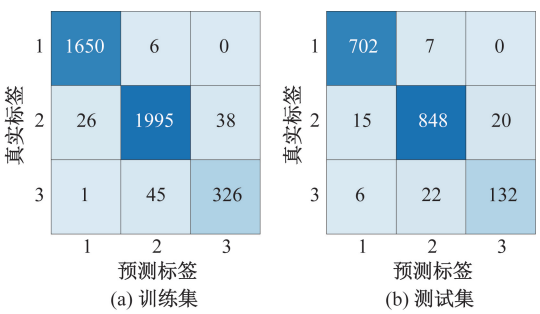


(a) 训练数据 (b) 验证数据

图 13 不同隐藏节点数量下的 IELM 模型平均准确率

Fig. 13 Average accuracy of the IELM model with different numbers of hidden nodes

使用数据集 $\boldsymbol{A}_1$、$\boldsymbol{A}_2$ 进行溢流风险识别,如图 14 所示,IELM 在训练集 $\boldsymbol{A}_1$ 和测试集 $\boldsymbol{A}_2$ 上的准确率分别达 97.16%和 96.01%,F_1 值 93.01%。



(a) 训练集 (b) 测试集

图 14 最优超参数下 IELM 模型溢流识别结果的混淆矩阵

Fig. 14 Confusion matrix of IELM model well kick identification results under optimal hyper-parameters

3.2.3 不同识别模型对比分析

为充分体现 IELM 的溢流风险识别效果,选用反向传播(Back Propagation,BP)神经网络、SVM 模型进行对比。

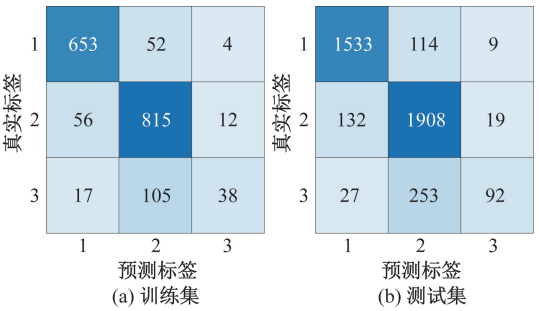
由于 ELM 为单隐层网络,选用单隐层 BP 神经网络进行模型对比,BP 网络迭代次数设为 2 000 次(对应 IELM 中 SSA 的 100 个种群迭代 20 次),超参数设置见表 5。

表 5 BP 神经网络的超参数设置

Table 5 Hyper-parameter settings of BP neural network

迭代次数	学习率	隐层激活函数	输出层激活函数	训练方式
2 000	0.1	正切 S 型	线性型	梯度下降

以 182 节点为最优参数训练 BP 网络,绘制识别结果混淆矩阵如图 15 所示,识别准确率分别为训练集 86.44%、测试集 85.96%。

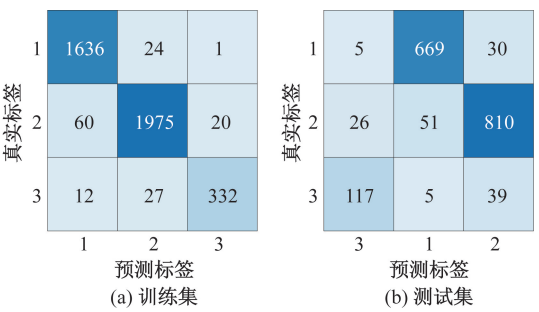


(a) 训练集 (b) 测试集

图 15 最优超参数下 BP 模型溢流识别结果的混淆矩阵

Fig. 15 Confusion matrix of BP model well kick identification results under the optimal hyper-parameters

选取 SVM 模型进行训练与测试,并通过网格搜索法寻找最优超参数。结果表明,当惩罚函数 $c=1\,000$,核参数 $\gamma=1$ 时,模型取得最佳识别性能。此时如图 16 所示,训练集识别准确率为 96.48%,测试集识别准确率为 91.10%。



(a) 训练集 (b) 测试集

图 16 最优超参数下 SVM 模型溢流识别结果的混淆矩阵

Fig. 16 Confusion matrix of SVM model well kick identification results under the optimal hyper-parameters

计算 IELM 算法的总体溢流风险识别准确率及各级准确率,并对比 IELM 与 BP、SVM 识别模型,IELM 溢流风险识别准确率高于 BP、SVM 模型,且低风险识别准确率达 99%,降低了误报警率,但高风险识别相对较低,需特别关注漏警问题。识别结

果汇总见表 6。

表 6　识别结果汇总

Table 6　Summary of identification results　%

数据集	准确率	IELM	ELM	BP	SVM
数据训练集 A_1	低风险	99.64	81.58	92.57	98.49
	中风险	96.89	92.81	92.67	96.11
	高风险	87.63	0	24.73	89.49
	整体	97.16	79.81	86.44	96.48
数据测试集 A_2	低风险	99.01	82.79	92.10	95.03
	中风险	96.03	92.53	92.30	91.32
	高风险	82.5	0	23.75	72.67
	整体	96.01	80.14	85.96	91.10
	F_1 分数	93.01	—	71.47	87.03

4　结　论

1）通过拟合残差分析与特征联动标注溢流风险,减弱了人为经验对溢流识别的干扰影响。通过分析钻井过程中各关键参数的联动变化特征,构建低、中、高 3 类溢流风险标签体系,形成高质量溢流风险数据集,为模型训练与风险识别提供可靠支撑,实现对溢流及潜在风险的准确判别。

2）引入 SSA 优化 ELM 网络结构,构建 IELM 多分类模型,通过模拟麻雀觅食与反捕食行为实现对网络权值与偏置的智能调节,显著提升算法在多类别识别任务中的鲁棒性与稳定性。

3）对比分析 IELM 多分类模型与原始 ELM 模型及经典 BP 神经网络、SVM 识别模型,结果表明:IELM 模型在溢流风险的整体分类准确率、稳定性方面均优于对比模型,展现出良好的实用性与工程推广价值,具备应用于深水钻井智能预警系统的潜力。

4）目前研究仅针对钻井过程溢流事件展开识别,对复杂工况下钻井事件的识别仍有不足。未来可围绕多源钻井数据融合、深度学习与时序建模引入、模型可解释性增强以及工程应用集成等方向展开,进一步提升溢流风险识别模型的泛化能力与实用价值。

参 考 文 献

[1] XU Chengyuan, YAN Xiaopeng, KANG Yili, et al. Structural failure mechanism and strengthening method of fracture plugging zone for lost circulation control in deep naturally fractured reservoirs [J]. Petroleum Exploration and Development, 2020, 47(2): 430-440.

[2] 李玉飞, 张博, 孙伟峰. 基于 SVM 和 D-S 证据理论的早期溢流智能识别方法[J]. 钻采工艺,2020,43(5):27-30,6.
LI Yufei, ZHANG Bo, SUN Weifeng. Research on intelligent early kick identification method based on SVM and D-S evidence theory [J]. Drilling & Production Technology, 2020, 43(5): 27-30,6.

[3] 邴磊. 基于 Attention-GRU 算法早期溢流识别预警方法[J]. 海洋石油,2025,45(2):76-82.
BING Lei. Early overflow identification based on Attention-GRU algorithm early warning method[J]. Offshore Oil, 2025, 45(2):76-82.

[4] 张禾, 池紫欣. 基于 BSMOTE-SVM 算法的溢流风险评价 [J]. 控制工程, 2023, 30(12): 2 173-2 178.
ZHANG He, CHI Zixin. Overflow risk evaluation based on BSMOTE-SVM algorithm [J]. Control Engineering of China, 2023, 30(12): 2 173-2 178.

[5] YANG Jin, SUN Ting, ZHAO Ying, et al. Advanced real-time gas kick detection using machine learning technology[C]. Proceedings of the 29th International Ocean and Polar Engineering Conference, 2019:1 404-1 409.

[6] 陈青, 黄志强, 孔祥伟, 等. 基于多录井参数特征同步的溢流事故监测研究 [J]. 应用数学和力学, 2025, 46(2): 241-53.
CHEN Qing, HUANG Zhiqiang, KONG Xiangwei, et al. Study on overflow accident monitoring based onsynchronous features of multiple well logging parameters [J]. Applied Mathematics and Mechanics, 2025, 46(2): 241-53.

[7] ZHANG Dezhi, SUN Weifeng, DAI Yongshou, et al. A hierarchical early kick detection method using a cascaded GRU network [J]. Geoenergy Science and Engineering, 2023, 222: DOI:10. 1016/j. geoen. 2022. 211390.

[8] 史肖燕, 周英操, 赵莉萍, 等. 基于随机森林的溢漏实时判断方法研究 [J]. 钻采工艺, 2020, 43(1): 9-12,7.
SHI Xiaoyan, ZHOU Yingcao, ZHAO Liping, et al. Losses in real time on basis of random forest method [J]. Drilling & Production Technology, 2020, 43(1): 9-12,7.

[9] WANG Chao, LIU Gonghui, YANG Zhirong, et al. Downhole working conditions analysis and drilling complications detection method based on deep learning [J]. Journal of Natural Gas Science and Engineering, 2020, 81: DOI: 10. 1016/j. jngse. 2020. 103485.

[10] ZHANG Zhi, SUN Baojiang, WANG Zzhiyuan, et al. Early monitoring method of downhole accident driven by physics based model and data driven methods coupling [J]. Geoenergy Science and Engineering, 2023, 221:DOI:10. 1016/j.

jngse. 2020. 103485.

[11] LIU Wei, FU Jiasheng, DENG Song, et al. Overflow identification and early warning of managed pressure drilling based on series fusion data-driven mode [J]. Processes, 2024, 12: DOI: 10. 3390/pr12071436.

[12] 陈贵波,刘东升,李加卫,等. 基于时间和通道融合的多变量时间序列异常检测[J/OL]. 计算机应用与软件:1-13. [2025-09-01]. https://link. cnki. net/urlid/31. 1260. TP. 20250919. 1739. 004.
CHEN Guibo, LIU Dongsheng, LI Jiawei, et al. Multivariate time series anomaly detection based on time and channel fusion[J/OL]. Computer Applications and Software: 1-13. [2025-09-01]. https://link. cnki. net/urlid/31. 1260. TP. 20250919. 1739. 004.

[13] WANG Shuhua, CHEN Zan, CHEN Shengnan. Applicability of deep neural networks on production forecasting in Bakken shale reservoirs [J]. Journal of Petroleum Science and Engineering, 2019, 179: 112-125.

[14] 杨向前, 张苹茹, 武胜男, 等. 基于数据模型协作的海上钻井溢流早期预测预警 [J]. 中国安全科学学报, 2024, 34(4): 93-100.
YANG Xiangqian, ZHANG Pingru, WU Shengnan, et al. Early prediction and waring of ofshore drilling overlow based ondata model collaboration [J]. China Safety Science Joural, 2024, 34(4): 93-100.

[15] 葛亮, 滕怡, 肖国清, 等. 基于井下环空参数的溢流智能预警技术研究 [J]. 西南石油大学学报:自然科学版, 2023, 45(2): 126-134.
GE Liang, TENG Yi, XIAO Guoqing, et al. Research on overlow intelligent warning teclmology based on downhole annulus parameters [J]. Joumnal of Southwest Petroleum university: Science & Technology Edition, 2023, 45(2): 126-134.

[16] 李开荣, 陈俊男, 段丽娟, 等. 钻井液池体积精准监测装置的研制与应用 [J]. 录井工程, 2022, 33(4): 92-96.
LI Kairong, CHEN Junnan, DUAN Lijuan, et al. Development and application of accurate monitoring device for drilling fluid pit volume [J]. Mud Logging Engineering, 2022, 33(4): 92-96.

[17] 张继德, 张永刚, 韩国生, 等. 钻井液录井参数在钻井工程异常预报中的应用 [J]. 录井工程, 2010, 21(3): 39-44, 76-77.
ZHANG Jide, ZHANG Yonggang, HAN Guosheng, et al. The application of drilling fluid logging parameter in the abnormal prediction of drilling engineering [J]. Mud Logging Engineering, 2010, 21(3): 39-44, 76-77.

[18] OSAROGIAGBON A, MUOJEKE S, VENKATESAN R, et al. A new methodology for kick detection during petroleum drilling using long short-term memory recurrent neural network [J]. Process Safety and Environmental Protection, 2020, 142: 126-137.

[19] SUN Weifeng, LI Weihua, ZHANG Dezhi, et al. Lost circulation monitoring using bi-directional LSTM and data augmentation [J]. Geoenergy Science and Engineering, 2023, 225: DOI: 10. 1016/j. geoen. 2023. 211660.

[20] WU Shengnan, HU Yiming, ZHANG Laibin, et al. Intelligent risk identification for drilling lost circulation incidents using data-driven machine learning [J]. Reliability Engineering & System Safety, 2024, 252: DOI: 10. 1016/j. ress. 2024. 110407.

[21] POWERS D. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation [J]. Journal Machine Learning Technologies, 2011, 2: 37-63.

作者简介： 武胜男 (1986—),女,山西大同人,博士,教授,主要从事深层、深水油气开采及关键安全装备风险评估与预警、可靠性与测试性维护方面的研究。E-mail: wushengnan@ cup. edu. cn。