



结合LSTM自编码器与集成学习的井漏智能识别方法

孙伟峰 冯剑寒 张德志 李威桦 刘凯 戴永寿

An Intelligent Lost Circulation Recognition Method Using LSTM-Autoencoder and Ensemble Learning

SUN Weifeng, FENG Jianhan, ZHANG Dezhi, LI Weihua, LIU Kai, DAI Yongshou

在线阅读 View online: <http://doi.org/10.11911/syztjs.2024006>

您可能感兴趣的其他文章

Articles you may be interested in

基于支持向量机的钻井工况实时智能识别方法

Real-Time Intelligent Identification Method under Drilling Conditions Based on Support Vector Machine

石油钻探技术. 2019, 47(5): 28–33 <http://doi.org/10.11911/syztjs.2019033>

基于人工智能的抽油机井结蜡预警方法

An Early Warning Method Based on Artificial Intelligence for Wax Deposition in Rod Pumping Wells

石油钻探技术. 2019, 47(4): 97–103 <http://doi.org/10.11911/syztjs.2019093>

基于可信度分析的钻井液脉冲信号识别方法

A New Drilling Fluid Pulse Signal Identification Method Based on Credibility Analysis

石油钻探技术. 2018, 46(3): 120–126 <http://doi.org/10.11911/syztjs.2018047>

延长油田西部地区低压易漏地层固井技术

Cementation Technology for Low-Pressure Formations Susceptible to Lost Circulation in Western Area of the Yanchang Oilfield

石油钻探技术. 2017, 45(4): 53–58 <http://doi.org/10.11911/syztjs.201704009>

基于测录井资料的环西—彭阳地区延安组储层流体性质识别方法研究

A Study on the Logging-Based Identification Method for Reservoir Fluid Properties of the Yan'an Formation in the HuanxiPengyang

石油钻探技术. 2020, 48(5): 111–119 <http://doi.org/10.11911/syztjs.2020079>

钻井液漏失位置测量仪的研制及试验

The Development and Testing of Lost Circulation Position Detector

石油钻探技术. 2017, 45(6): 55–59 <http://doi.org/10.11911/syztjs.201706010>



扫码关注公众号，获取更多信息！

◀ 钻井完井 ▶

doi:10.11911/syztjs.2024006

引用格式: 孙伟峰, 冯剑寒, 张德志, 等. 结合 LSTM 自编码器与集成学习的井漏智能识别方法 [J]. 石油钻探技术, 2024, 52(3): 61-67.

SUN Weifeng, FENG Jianhan, ZHANG Dezhi, et al. An intelligent lost circulation recognition method using LSTM-autoencoder and ensemble learning [J]. Petroleum Drilling Techniques, 2024, 52(3): 61-67.

结合 LSTM 自编码器与集成学习的井漏智能识别方法

孙伟峰¹, 冯剑寒¹, 张德志², 李威桦¹, 刘 凯², 戴永寿¹

(1. 中国石油大学(华东)海洋与空间信息学院, 山东青岛 266580; 2. 中国石油大学(华东)控制科学与工程学院, 山东青岛 266580)

摘 要: 为了解决传统的井漏智能识别模型因井漏样本数量受限导致其识别准确率低的问题, 提出了一种长短期记忆 (long short-term memory, LSTM) 网络与自编码器 (auto-encoder, AE) 相结合、集成 LSTM-AE 的井漏智能识别方法。首先, 采用正常样本训练多个包含不同隐藏层神经元数目的 LSTM-AE 模型, 利用重构得分筛选出识别效果较好的几个模型作为基识别器; 然后, 采用集成学习对多个基识别器的识别结果进行加权融合, 解决单一模型因对样本局部特征过度学习导致的误报与漏报问题, 提高模型的识别准确率。从某油田 18 口井的钻井数据中选取了 6 000 组正常钻进状态下的立压、出口流量、池体积数据, 对集成 LSTM-AE 模型进行训练和测试, 结果表明, 提出方法的识别准确率达到 94.7%, 优于其他常用的智能模型的识别结果, 为井漏识别提供了一种新的技术途径。

关键词: 井漏识别; 长短期记忆网络; 自编码器; 集成学习

中图分类号: TE28 文献标志码: A 文章编号: 1001-0890(2024)03-0061-07

An Intelligent Lost Circulation Recognition Method Using LSTM-Autoencoder and Ensemble Learning

SUN Weifeng¹, FENG Jianhan¹, ZHANG Dezhi², LI Weihua¹, LIU Kai², DAI Yongshou¹

(1. College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao, Shandong, 266580, China;
2. College of Control Science and Engineering, China University of Petroleum (East China), Qingdao, Shandong, 266580, China)

Abstract: To enhance the low recognition accuracy of traditional intelligent lost circulation models, which suffer from limited samples, this study combined the long short-term memory (LSTM) network and auto-encoder (AE) to create an integrated LSTM-AE-based intelligent lost circulation recognition model. Initially, multiple LSTM-AE models with varying numbers of hidden neurons were trained using normal samples. Several models with better recognition performance were selected as base recognizers based on their reconstruction scores. Subsequently, the recognition results from these base recognizers were fused using ensemble learning. This approach addresses the tendency of a single model to produce false alarms and missed alarms due to overlearning of local sample characteristics, thereby improving the recognition accuracy of the model. The integrated LSTM-AE model was trained and tested using 6 000 sets of stand pipe pressure, outlet flow, and mud pit volume data from 18 wells under normal drilling conditions in an oilfield. The results show that the proposed method achieves a recognition accuracy of 94.7%, surpassing the recognition results of other commonly used intelligent models. This approach offers a novel method for lost circulation recognition.

Key words: lost circulation recognition; long short-term memory networks; auto-encoder; ensemble learning

井漏是钻井过程中经常遇到的井下复杂情况^[1]。钻井液频繁漏失不但会延长钻井周期, 甚至会引起井喷等更严重的钻井事故。传统的井漏监测方法依靠人工监测录井参数实现, 但录井参数的变化具有缓慢、细微的特点, 导致监测结果及时性差、准确率低。近年来, 随着大数据与人工智能在油气田开发

收稿日期: 2023-11-26; 改回日期: 2024-02-26。

作者简介: 孙伟峰 (1982—), 男, 山东东营人, 2005 年毕业于山东大学通信工程专业, 2010 年获山东大学信号与信息处理专业博士学位, 教授, 博士生导师, 主要从事钻井风险智能预警、高频雷达海洋遥感等方面的研究工作。E-mail: sunwf@upc.edu.cn。

基金项目: 国家自然科学基金项目“基于深度学习的深地叠前时空域地震子波提取方法研究”(编号: 42274159)资助。

领域中的应用^[2], 基于机器学习的钻井风险识别方法^[3-7] 受到越来越多的关注。由于钻井风险识别所用的录井参数皆为时序数据, 具备时序信息处理能力的长短期记忆(long short-term memory, LSTM)网络^[8-10] 被应用于风险识别。然而, 井漏样本数量较少, 且与正常工况下的数据量不平衡, 使模型倾向于将识别结果划分到多数样本的类别^[11], 识别准确率不高。与上述方法相比, 利用正常样本训练的自编码器(auto-encoder, AE)成为解决钻井风险识别问题的新途径。传统自编码器在网络异常检测^[12]、轴承故障诊断^[13]、轨迹异常检测^[14] 等方面取得了良好的应用效果, 但其处理时间序列的能力不足, 无法有效提取录井参数的变化特征。集成学习(ensemble learning)通过融合多个单一学习器的识别结果, 降低了单一学习器因过度学习局部特征导致的误报、漏报, 能够获得更好的泛化能力和更优的识别结果。目前, 集成学习已经与卷积神经网络^[15]、支持向量机^[16]、循环神经网络^[17] 相结合, 用来提升模型的分类与预测性能。

基于上述考虑, 为了在井漏样本数量不足条件下准确识别井漏, 笔者提出了集成 LSTM-AE 的井漏智能识别方法。利用 LSTM 网络作为自编码器的隐含层, 构建包含不同隐藏层神经元数目的多个 LSTM-AE 模型, 通过重构得分筛选出具有较好性能的 LSTM-AE 模型作为基识别器, 对多个基识别器的重构误差与阈值进行加权集成, 提升井漏识别的准确率, 实验结果验证了提出方法的有效性。

1 时间序列重构原理

1.1 LSTM 网络时间序列处理原理

LSTM 网络作为循环神经网络(recurrent neural network, RNN)的变种, 通过改变门控状态来控制传输状态, 解决了 RNN 网络存在的梯度消失的问题, 目前已成为处理时间序列数据的标准方法^[18]。LSTM 网络由遗忘门、输入门和输出门组成(见图 1)。

遗忘门的任务是接收上一时刻的记忆并决定所要保留和遗忘的部分, 输入门对输入进行选择性记忆, 输出门决定将要输出的信息, 遗忘门、输入门、输出门的计算公式分别为:

$$f_t = \sigma(\mathbf{W}_{fh}h_{t-1} + \mathbf{W}_{fx}x_t + \mathbf{W}_{fc}c_{t-1} + \mathbf{b}_f) \quad (1)$$

$$i_t = \sigma(\mathbf{W}_{ih}h_{t-1} + \mathbf{W}_{ix}x_t + \mathbf{W}_{ic}c_{t-1} + \mathbf{b}_i) \quad (2)$$

$$o_t = \sigma(\mathbf{W}_{oh}h_{t-1} + \mathbf{W}_{ox}x_t + \mathbf{W}_{oc}c_t + \mathbf{b}_o) \quad (3)$$

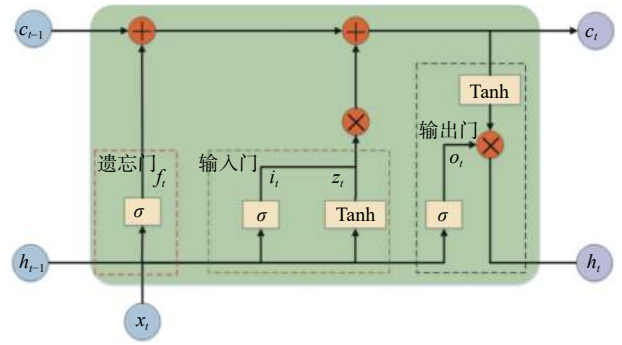


图 1 LSTM 网络结构

Fig.1 Structure of LSTM network

式中: f_t , i_t 和 o_t 分别为遗忘门、输入门和输出门; h_{t-1} 为上一时刻的输出; x_t 为当前时刻的输入; c_t , c_{t-1} 分别为当前时刻与上一时刻记忆单元的存储信息; \mathbf{W}_{fh} , \mathbf{W}_{fx} 和 \mathbf{W}_{fc} 为遗忘门的权重参数矩阵; \mathbf{W}_{ih} , \mathbf{W}_{ix} 和 \mathbf{W}_{ic} 为输入门的权重参数矩阵; \mathbf{W}_{oh} , \mathbf{W}_{ox} 和 \mathbf{W}_{oc} 为输出门的权重参数矩阵; \mathbf{b}_f , \mathbf{b}_i 和 \mathbf{b}_o 为网络偏置参数矩阵; σ 表示 sigmoid 函数。

输入门的值接近于 1 时, 即接收全部当前候选记忆单元的值; 遗忘门的值接近于 1 时, 即拒绝接收当前候选记忆单元的值, 最终的输出则是由输出门和当前单元状态共同决定。LSTM 层选择性地接收当前信息, 有效获取数据序列内部的长时相关性, 具备处理时间序列的能力。当前记忆单元、候选记忆单元、当前时刻输出的公式分别为:

$$c_t = c_{t-1}f_t + z_t i_t \quad (4)$$

$$z_t = \tanh(\mathbf{W}_{zy}h_{t-1} + \mathbf{W}_{zx}x_t + \mathbf{b}_z) \quad (5)$$

$$h_t = o_t \tanh(c_t) \quad (6)$$

式中: \mathbf{W}_{zy} 和 \mathbf{W}_{zx} 为记忆单元的权重参数矩阵; \mathbf{b}_z 为网络偏置参数矩阵; h_t 为当前时刻的输出; z_t 为当前候选记忆单元。

1.2 LSTM-AE 时间序列重构原理

传统自编码器(auto-encoder, AE)由输入层、隐含层、输出层构成。在训练阶段, 通过对正常样本进行重构训练, 使自编码器熟悉正常样本的数据分布规律, 实现对正常样本特征的提取。为了提升传统自编码器的时序处理能力, 使用 LSTM 层代替传统自编码器中的全连接层, 构建 LSTM-AE 网络, 其结构如图 2 所示。LSTM-AE 对时间序列重构的过程由编码器(encoder)与解码器(decoder)实现。位于输入层与隐含层之间的编码器在高维空间对输入时间序列的特征进行提取, 并编码为低维空间特征, 编码阶段实际是一个保留部分特征的数据压缩

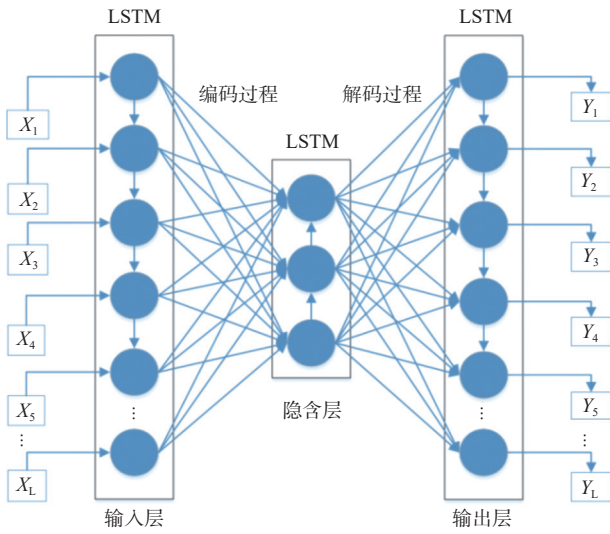


图 2 LSTM-AE 结构

Fig.2 Structure of LSTM-AE

过程, 可以描述为:

$$h = \eta(W_1 X + b_1) \quad (7)$$

式中: W_1 和 b_1 分别为编码器中 LSTM 层的权重与偏置矩阵; h 为编码后处于隐含层的低维特征; η 为神经元间的激活函数; X 为输入的时间序列。

解码器位于隐含层与输出层之间, 将根据隐含层的低维空间特征, 使其重新映射回高维空间, 输出与输入具有相同维度的时间序列, 从而完成了时间序列的重构, 过程可以描述为:

$$Y = \eta(W_2 h + b_2) \quad (8)$$

式中: W_2 和 b_2 分别为解码器中 LSTM 层的权重与偏置矩阵; Y 为编码后得到的时间序列。

训练 LSTM-AE 的目的是通过不断更新编码器与解码器的权重与偏置, 最小化损失, 使最终输出的时间序列与输入的时间序列相同, 在实现对样本序列无损重构的同时, 使模型学习到正常样本的特征。重构损失 L 的计算公式为:

$$L = \|Y - X\|^2 \quad (9)$$

2 集成 LSTM-AE 的井漏识别方法

集成 LSTM-AE 模型对正常样本中的监测参数变化特征进行学习后, 当测试样本中的监测参数特征与正常样本中的监测参数特征差别较大时, 测试样本的重构误差较高, 模型将测试样本判别为井漏样本。因此, 需先选取变化规律能反映井漏特征的监测参数, 再对数据进行预处理后输入模型。

2.1 监测参数选取

井漏发生的机理较为复杂, 单一参数的变化并不能唯一确定井漏的发生, 因此需要结合多个监测参数的变化对井漏进行识别^[19]。当井漏发生时, 随着钻井液从井筒沿着漏失通道进入地层, 环空中钻井液的上返速度降低, 造成环空摩阻减小, 导致立管压力降低^[20], 并且在入口流量没有显著变化的情况下, 出口流量减少; 最后, 部分钻井液流入地层导致上返至地面的钻井液减少, 钻井液池中钻井液的体积减小。因此, 选择钻井液池中钻井液体积、出口流量百分比、立管压力作为集成 LSTM-AE 模型的输入参数。

2.2 数据预处理

为了降低噪声与不同参数间的数值范围对数据的影响, 采用滤波与归一化的方法对钻井数据进行预处理。

2.2.1 钻井数据滤波

在钻井数据的采集过程中, 数据不可避免地存在许多噪声及急剧变化的跳变点, 这些干扰对于 LSTM-AE 模型提取正常样本中监测参数的全局特征有较大的影响, 因此需要使用滤波的方法减少这类因素的干扰。

SG 滤波是平滑时间序列数据滑动窗口的一种加权平均算法, 其加权系数是通过一个滤波窗口内给定高阶多项式的最小二乘拟合得到, SG 滤波的公式为:

$$y_j^* = \frac{\sum_{i=-n}^n a_i y_{j+i}}{m} \quad (10)$$

式中: y_j^* 为时间序列拟合后的值; y 为原始的时间序列; a_i 为滤波系数; m 为卷积数目; n 为滤波窗口大小, 与多项式的阶数共同决定最终的滤波效果。

2.2.2 钻井数据归一化

为了降低监测参数之间特征尺度及量纲不同对自编码器提取监测参数特征的影响, 通常需要进行归一化处理, 常用的数据归一化方法为最大-最小值归一化, 其计算公式为:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (11)$$

式中: x 为当前值; x' 为归一化后得到的值; $\min(x)$ 和 $\max(x)$ 分别表示一组数据中的最小值、最大值。

2.3 LSTM-AE 的集成过程

在训练阶段, LSTM-AE 模型学习到正常样本的特征, 根据测试样本重构误差与阈值的大小关系判

断测试样本是否为井漏样本;但 LSTM-AE 模型会过度学习训练样本中的局部特征而遗漏掉对其他特征的学习,使识别准确率降低。针对这一问题,提出了基于重构得分的模型筛选方法,从多个 LSTM-AE 模型中筛选出识别性能较好的基识别器,使用集成的方法集成多个基识别器对井漏进行组合判别。由于不同基识别器对样本特征的学习存在差异,被单一模型遗漏掉的特征可被其他模型学习到,弥补了单一模型特征学习能力的不足,提升井漏识别的准确性。集成 LSTM-AE 模型的建立分为以下3步:

1) LSTM-AE 模型训练。将正常样本集划分为训练集 D_1 与测试集 D_2 , 设置 k 个结构不同的 LSTM-AE 模型。训练阶段,使用训练集 D_1 对 k 个 LSTM-AE 模型分别进行训练,得到每个样本的重构误差,并将第 k 个 LSTM-AE 模型的误差阈值设定为:

$$T_k = Q_k \quad (12)$$

式中: Q_k 为第 k 个 LSTM-AE 模型对 D_1 的所有重构误差的 95% 分位数; T_k 为第 k 个 LSTM-AE 模型的误差阈值。

2) 基识别器筛选。根据每一个自编码器的误差阈值 Q_n , 其重构得分计算公式为:

$$S_k = \frac{1}{2} \ln \left[\frac{1 - Q_k}{Q_k} \right] \quad (13)$$

式中: S_k 为其重构得分。重构得分越高,代表模型在训练时的损失越小,一定程度上表示模型的重构效果较好。

取得分最高的 p 个 LSTM-AE 模型作为基识别器,对 S_p 进行归一化,得到每个子 LSTM-AE 模型的最终权重矩阵 S 为:

$$S = [S_1, S_2, \dots, S_p] \quad (14)$$

式中: S_p 为第 p 个基识别器所占的权重。

3) 井漏判别。将正常样本测试集 D_2 与井漏样本测试集 D_3 输入基识别器中,分别得到一组重构误差,最终测试集中第 i 个样本集成后的重构误差 e_i 为每个基识别器的重构误差的加权和:

$$e_i = S_1 e_{i1} + S_2 e_{i2} + \dots + S_p e_{ip} \quad (15)$$

式中: e_i 为第 i 个序列集成后的重构误差; e_{pi} 为第 p 个基识别器对第 i 个序列的重构误差。

最终的误差阈值 T 为每个基识别器的误差阈值的加权和:

$$T = S_1 T_1 + S_2 T_2 + \dots + S_p T_p \quad (16)$$

式中: T_p 为第 p 个基识别器的误差阈值; T 为最终的误差阈值。若 $e_i \leq T$, 则被模型判断为正常钻进状态;若 $e_i > T$, 则模型给出井漏报警。

3 井漏识别试验结果与分析

3.1 LSTM-AE 井漏识别模型的搭建

3.1.1 数据来源与样本集划分

本次试验中所使用的数据集来源于国内具有井漏记录的 18 口井,对数据集中的时间序列使用长度为 20、步长为 1 的滑动窗口共生成 6 000 个数据样本,滑动窗口所生成的样本包含了时间序列的趋势变化信息,满足了井漏识别的实时性,其中正常样本用于模型的训练与测试,而井漏样本只用于模型的测试,每一个样本包含钻井液池中钻井液体积、出口流量百分比和立管压力各 20 个数据点,滑动窗口过程如图 3 所示。

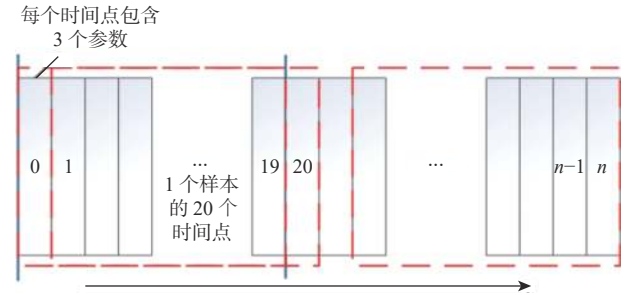


图 3 滑动窗口示意

Fig.3 Sliding window

按照 9:1, 8:2, 7:3 和 6:4 等 4 种常用比例划分训练集与测试集,对模型进行训练与测试,对比不同测试集模型的识别准确率。测试结果表明,当训练集与测试集比例为 8:2 时,识别准确率最高,达到了 84.5%。因此,按照 8:2 的比例对样本进行划分,用于训练的正常样本为 4 000 个,用于测试的正常样本与井漏样本各 1 000 个。

3.1.2 模型训练

为了训练出多个性能不同的 LSTM-AE 模型,根据传统自编码器结构对称的特点,在 4 层 LSTM 层组成的模型中,使用第二层 LSTM 层将输入样本特征降维至低维空间,前 2 层 LSTM 层的作用是逐步将数据量压缩,避免直接降维导致特征大量丢失,后 2 层 LSTM 层通过将数据扩充实现重构。为了保证在降维的同时实现数据量的压缩,第二层 LSTM 层神经元数量需小于前一层 LSTM 层的神经元数,在此条件下,通过改变第二层 LSTM 层的神经

经元数,特征降维后所保留的特征产生了变化,因此具有构建多个不同识别效果的 LSTM-AE 模型的可能。构建第二层 LSTM 层神经元数为 2~15 个、其他参数均相同的 14 个 LSTM-AE 模型,对所有 LSTM-AE 模型进行训练,并根据重构得分筛选出基识别器,重构得分高的前 6 个 LSTM-AE 模型的重构结果见表 1。从表 1 可以看出,当神经元数量为 12 与 11 时,模型性能最好。因此,优先选择神经元数量为 12 与 11 的 LSTM-AE 模型进行集成。

表 1 不同神经元数量 LSTM-AE 模型重构结果的对比
Table 1 Comparison of reconstruction performance of LSTM-AE with different numbers of neurons

序号	LSTM层神经元数	误差阈值	重构得分
1	12	0.019	3.95
2	11	0.045	3.05
3	4	0.060	2.74
4	3	0.071	2.57
5	6	0.074	2.51
6	15	0.076	2.49

3.2 集成结果与对比分析

3.2.1 评价指标

采用准确率、误报率和漏报率对模型性能进行评估^[21]。此外,采用查准率、查全率比较本文方法与其他智能方法的性能,计算公式分别为:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (17)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (18)$$

式中: P 为查准率; R 为查全率; N_{TP} 为实际为井漏且判别为井漏的样本数; N_{TN} 为实际为正常且判别为正常的样本数; N_{FP} 为实际为正常但判别为井漏的样本数; N_{FN} 为实际为井漏但判别为正常的样本数。

3.2.2 集成结果与对比试验

根据重构得分的高低,优先选取重构得分较高的 LSTM-AE 模型作为基识别器,基识别器数量过多时,集成后的模型会包含重构效果不好的基识别器,最终识别的准确率不会有提升,所以选取 2~6 个基识别器进行集成测试。对效果最好的单一基识别器模型与不同个数基识别器集成的 LSTM-AE 模型的漏报率、误报率、准确率进行对比,结果如图 4 所示。基识别器数量过多时,漏报率较高,导致识别准确率下降;基识别器数量为 3 时,集成效果最

好,准确率达到 94.7%,与识别效果最好的单一基识别器相比,准确率提高了 4.6 个百分点。

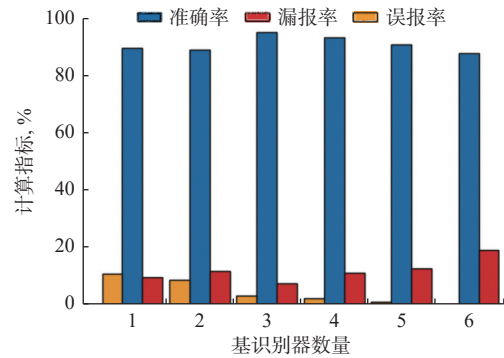


图 4 不同数量基识别器集成效果的对比

Fig.4 Comparison of integrated effects for different numbers of base recognizers

为了进一步评估本模型的性能,选取 4 种智能识别方法与集成 LSTM-AE 模型进行对比,结果见表 2。由表 2 可知,在样本数量不平衡的条件下,SVM 等井漏识别方法趋向于把测试样本划分为正常样本,尽管误报率略低,但查全率、准确率均低于集成后 LSTM-AE 模型,验证了井漏样本不足条件下本文方法的有效性。

表 2 不同模型井漏识别结果的对比

Table 2 Comparison of lost circulation recognition results of different models

方法	误报率, %	查全率, %	查准率, %	准确率, %
集成LSTM-AE	2.6	97.1	92.3	94.7
LSTM	3.9	95.8	90.3	93.4
SVM	1.0	98.7	79.0	89.0
随机森林	1.4	98.8	85.1	92.0
决策树	7.8	91.0	82.6	87.4

3.2.3 案例分析

根据某井的记录日志,该井发生井漏前后监测参数的变化如图 5 所示,现场于 78 min 即监测到此井钻井液池中钻井液的体积减少了 1.0 m³,给出了井漏报警。

选择得分最高的 3 个 LSTM-AE 模型作为基识别器,采用各基识别器与集成后的模型对图 5 中井漏样本进行识别,结果如图 6 所示(图 6 中,绿线与红线分别为现场标定的正常钻进状态与发生井漏的时间段),现场报警时间在第 58 min。

从图 6 可以看出,基识别器 1 的重构误差于 46 min 超出阈值,给出风险报警,比现场记录的报警时间

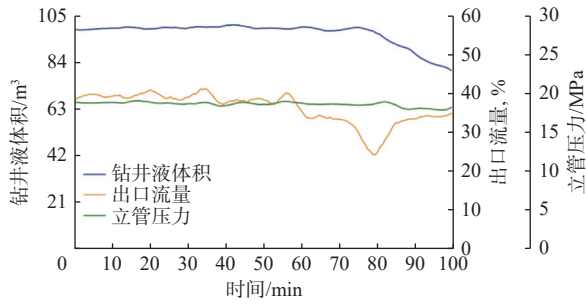


图5 某井井漏前后参数的变化情况
Fig.5 Parameter variations before and after lost circulation in a well

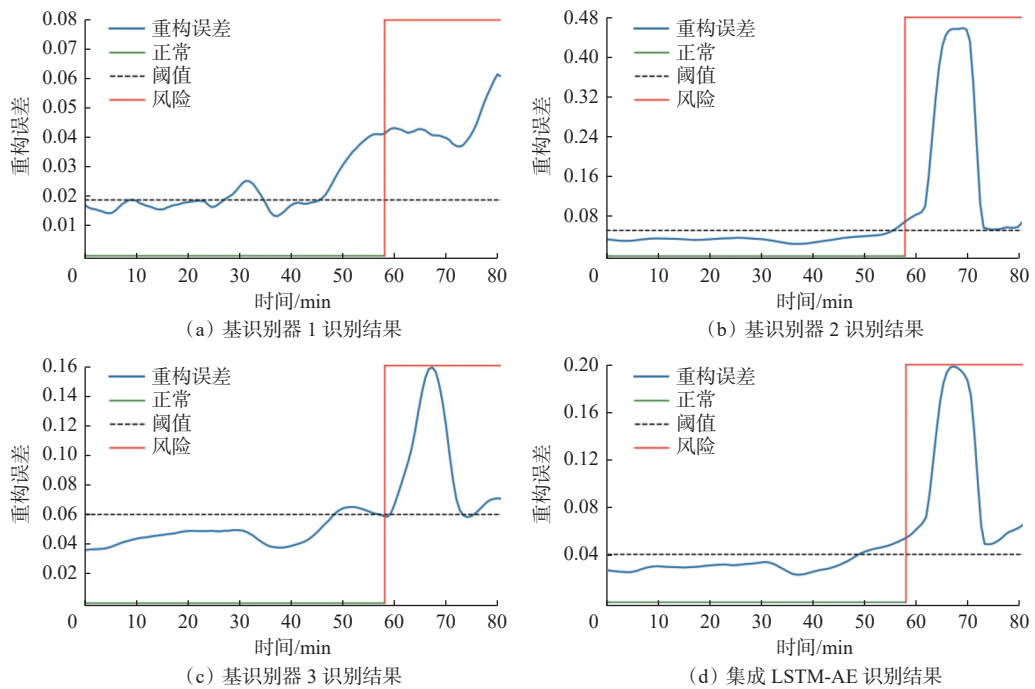


图6 各基识别器与集成 LSTM-AE 识别结果的对比
Fig.6 Comparison of recognition results of integrated LSTM-AE and base recognizers

4 结论与建议

1) 针对因井漏样本数量少导致的传统智能模型识别准确率低的问题, 提出了基于集成 LSTM-AE 的井漏智能识别方法, 能够实现对正常样本与井漏样本的区分。与单一模型相比, 集成 LSTM-AE 模型能够有效降低误报率和漏报率。

2) 集成 LSTM-AE 模型的识别准确率优于 LSTM、SVM 等方法的识别结果, 验证了集成 LSTM-AE 模型在井漏样本数量不足条件下对井漏识别的有效性。

3) 井漏发生时, 往往伴随着多个钻井参数的联动变化, 建议后续通过优选监测参数, 进一步提升

提前了 12 min, 但该模型对参数变化过于敏感, 在 27~35 min 时出现了误报; 基识别器 2 未出现误报的情况, 于 55 min 给出预警, 比现场记录的报警时间提前了 3 min; 基识别器 3 于 48 min 报警, 但随后出现漏报的情况, 表明该模型对监测参数的异常变化不够敏感; 集成 LSTM-AE 未出现误报与漏报的情况, 于 48 min 给出预警, 比现场记录的报警时间提前了 10 min。可见, 集成后的模型在能够及时给出风险预警的同时, 还能够减少误报的情况, 实现了各基识别器的优势结合。

LSTM-AE 模型对井漏的识别准确率。

参 考 文 献

References

[1] 陈钢花, 何宇龙, 邱正松, 等. 钻井过程中井漏特征精细识别方法研究与应用 [J]. 石油钻探技术, 2024, 52(1): 26-31.
CHEN Ganghua, HE Yulong, QIU Zhengsong, et al. Research and application for the fine identification method of lost circulation characteristics during drilling[J]. Petroleum Drilling Techniques, 2024, 52(1): 26-31.

[2] 李中. 中国海油油气井工程数字化和智能化新进展与展望 [J]. 石油钻探技术, 2022, 50(2): 1-8.
LI Zhong. Progress and prospects of digitization and intelligentization of CNOOC's oil and gas well engineering[J]. Petroleum Drilling Techniques, 2022, 50(2): 1-8.

- [3] 徐哲, 李建, 王兵, 等. 基于贝叶斯网络的钻井井漏问题研究 [J]. 石油天然气学报, 2013, 35(12): 125-129.
XU Zhe, LI Jian, WANG Bing, et al. Research on well leakage in drilling based on Bayesian network[J]. Journal of Oil and Gas Technology, 2013, 35(12): 125-129.
- [4] 刘彪, 李宸晓, 李双贵, 等. 基于支持向量回归的井漏预测 [J]. 钻采工艺, 2019, 42(6): 17-20.
LIU Biao, LI Chengxiao, LI Shuangui, et al. Lost circulation prediction based on support vector regression[J]. Drilling & Production Technology, 2019, 42(6): 17-20.
- [5] 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.
SHI Xiaoyan, ZHOU Yingcao, ZHAO Liping, et al. Study on how to determine kicks and losses in real time on basis of random forest method[J]. Drilling & Production Technology, 2020, 43(1): 9-12.
- [6] 杨传书, 李昌盛, 孙旭东, 等. 人工智能钻井技术研究方法及其实践 [J]. 石油钻探技术, 2021, 49(5): 7-13.
YANG Chuanshu, LI Changsheng, SUN Xudong, et al. Research method and practice of artificial intelligence drilling technology[J]. Petroleum Drilling Techniques, 2021, 49(5): 7-13.
- [7] 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 overflow intelligent warning technology based on downhole annulus parameters[J]. Journal of Southwest Petroleum University(Science & Technology Edition), 2023, 45(2): 126-134.
- [8] 孙伟峰, 卜赛赛, 张德志, 等. 基于 DCC-LSTM 的钻井液微量漏失智能监测方法 [J]. 天然气工业, 2023, 43(9): 141-148.
SUN Weifeng, BU Saisai, ZHANG Dezhi, et al. DCC-LSTM based intelligent minor lost circulation monitoring method[J]. Natural Gas Industry, 2023, 43(9): 141-148.
- [9] 张矿生, 宫臣兴, 陆红军, 等. 基于集成学习的井漏智能预警模型及智能推理方法 [J]. 石油钻采工艺, 2023, 45(1): 47-54.
ZHANG Kuangsheng, GONG Chenxing, LU Hongjun, et al. Intelligent early warning model and intelligent reasoning method based on integrated learning for loss circulation[J]. Oil Drilling & Production Technology, 2023, 45(1): 47-54.
- [10] 王钰豪, 郝家胜, 张帆, 等. 钻井溢流风险的自适应 LSTM 预警方法 [J]. 控制理论与应用, 2022, 39(3): 441-448.
WANG Yuhao, HAO Jiasheng, ZHANG Fan, et al. Adaptive LSTM early warning method for kick detection in drilling[J]. Control Theory & Applications, 2022, 39(3): 441-448.
- [11] 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.
- [12] 杨岳毅, 王立德, 陈煌, 等. 基于变分自编码器的 MVB 网络异常检测方法 [J]. 铁道学报, 2022, 44(1): 71-78.
YANG Yueyi, WANG Lide, CHEN Huang, et al. Anomaly detection method for MVB network based on variational autoencoder[J]. Journal of the China Railway Society, 2022, 44(1): 71-78.
- [13] 王晓玉, 刘桂芳, 韩宝坤, 等. 堆叠自编码器在样本不充足下的轴承故障诊断方法 [J]. 噪声与振动控制, 2021, 41(2): 100-104.
WANG Xiaoyu, LIU Guifang, HAN Baokun, et al. Stacked autoencoders for bearing fault diagnosis under the condition of insufficient samples[J]. Noise and Vibration Control, 2021, 41(2): 100-104.
- [14] 秦婉亭, 老松杨, 汤俊, 等. 基于变分自编码器的飓风轨迹异常检测方法 [J]. 系统仿真学报, 2021, 33(9): 2191-2201.
QIN Wanting, LAO Songyang, TANG Jun, et al. Hurricane trajectory outlier detection method based on variational auto-encode[J]. Journal of System Simulation, 2021, 33(9): 2191-2201.
- [15] 田晟, 宋霖. 基于 CNN 和 Bagging 集成的交通标志识别 [J]. 广西师范大学学报(自然科学版), 2022, 40(4): 35-46.
TIAN Sheng, SONG Lin. Traffic sign recognition based on CNN and Bagging integration[J]. Journal of Guangxi Normal University(Natural Science Edition), 2022, 40(4): 35-46.
- [16] 陈江, 单桂军, 李正明. 基于支持向量机集成学习的网络故障诊断方法 [J]. 计算机测量与控制, 2014, 22(12): 3906-3908.
CHEN Jiang, SHAN Guijun, LI Zhengming. Network fault diagnosis based on support vector machine ensemble learning[J]. Computer Measurement & Control, 2014, 22(12): 3906-3908.
- [17] 徐晓芳, 管瑞. 基于神经网络集成学习算法的金融时间序列预测 [J]. 计算机系统应用, 2022, 31(6): 29-37.
XU Xiaofang, GUAN Rui. Financial time series forecasting based on neural network ensemble learning algorithms[J]. Computer Systems & Applications, 2022, 31(6): 29-37.
- [18] YU Yong, SI Xiaosheng, HU Changhua, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.
- [19] 王海彪. 井漏智能识别及处理决策研究 [D]. 成都: 西南石油大学, 2017.
WANG Haibiao. Study of intelligent recognition and management decisions of lost circulation[D]. Chengdu: Southwest Petroleum University, 2017.
- [20] 蔺研锋, 闵超, 代博仁, 等. 基于动态特征和深度神经网络的钻井漏失事故预测 [J]. 西安石油大学学报(自然科学版), 2022, 37(3): 64-69.
LIN Yanfeng, MIN Chao, DAI Boren, et al. Prediction of drilling leakage accident based on dynamic features and deep neural network[J]. Journal of Xi'an Shiyou University(Natural Science Edition), 2022, 37(3): 64-69.
- [21] 孙伟峰, 刘凯, 张德志, 等. 结合钻井工况与 Bi-GRU 的溢流与井漏监测方法 [J]. 石油钻探技术, 2023, 51(3): 37-44.
SUN Weifeng, LIU Kai, ZHANG Dezhi, et al. A kick and lost circulation monitoring method combining Bi-GRU and drilling conditions[J]. Petroleum Drilling Techniques, 2023, 51(3): 37-44.

[编辑 滕春鸣]