

DOI:10.19651/j.cnki.emt.2210660

基于 AMSDAE-BLSTM 的工业过程质量预测^{*}

郭小萍 钟道金 李元

(沈阳化工大学信息工程学院 沈阳 110142)

摘要: 针对具有噪声干扰及延迟等特性的工业过程质量预测,本文提出了一种嵌入注意力机制的堆叠降噪自编码器与双向长短期记忆网络的方法。首先以无监督方式构建自编码器模型,利用高斯噪声对工业数据进行一次重构以实现去噪及去冗余作用;再次嵌入注意力机制对过程变量权重分配进行二次重构以实现深度特征提取;最后采用双向长短期记忆网络学习重构数据的时间序列趋势特征,克服数据间的延迟性,充分挖掘过程变量与质量变量间的潜在关系,实现精准预测。通过脱丁烷过程的单质量变量预测和硫磺回收过程的多变量质量预测仿真实验,验证了本文方法比 BP、LSTM、BLSTM 和 DAE-BLSTM 方法具有更精确的预测效果。

关键词: 质量预测;堆叠降噪自编码器;注意力机制;双向长短期记忆网络

中图分类号: TP183 **文献标识码:** A **国家标准学科分类代码:** 520.2010

Industrial process quality prediction based on AMSDAE-BLSTM

Guo Xiaoping Zhong Daojin Li Yuan

(School of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China)

Abstract: Aiming at the prediction of industrial process quality with noise interference and delay, in this paper, we propose a method of stacking noise reduction auto-encoder embedded with attention mechanism and bidirectional long short-term memory network. Firstly, the AE model is constructed in an unsupervised manner, and the industrial data is reconstructed with Gaussian noise processing to achieve denoising and de-redundancy. Secondly, embed the attention mechanism to reconstruct the weight allocation of process variables to achieve deep feature extraction. Finally, the BLSTM network is used to learn and reconstruct the time series trend characteristics of the data to overcome the delay between the data, and fully explore the potential relationship between the process variables and the quality variables, and finally achieve accurate prediction. Through the simulation experiments of single-mass variable prediction of debutane process and multivariate mass prediction of sulfur recovery process, it is verified that the proposed method has a more accurate prediction effect than other methods such as BP, LSTM, BLSTM and DAE-BLSTM.

Keywords: quality prediction; stacked denoising auto-encoder; attention mechanism; bidirectional long short-term memory network

0 引言

在现代工业过程中,各种先进的监测、控制和优化技术用于保证工业生产的正常运行,优化资源的利用效率,缓解环境污染的压力,但这些技术的开发和实施往往依赖于对重要质量指标的及时测量,然而,质量变量的在线测量仪器^[1]往往是昂贵的,并受到较大的时间延迟和严峻的环境考验。因此,通过易于测量的过程变量对难以测量的质量变量进行实时估计建立质量预测模型还是极为有效的。

现代工业过程采集的数据通常具有复杂关联特性,非线性性和多噪声等特点,这可能导致质量预测模型的预测性

能下降。因此,从大量原始数据中提取有用信息是质量预测建模的重要一步,它促进了特征提取的发展,主成分分析(principal component analysis, PCA)^[2]、偏最小二乘(partial least square, PLS)^[3]、独立成分分析^[4]和流形学习^[5]是传统的特征提取模型,然而这些模型大多只有一个或零隐藏层的浅层学习结构。许多研究表明,在复杂过程中深度网络比浅层网络更有效地提取层次特征。近年来,一些技术被提出用于深度网络训练。例如,自编码器(autoencoder, AE)^[6]利用了隐藏层挖掘其有用的属性使输入基本无损的复制到输出,但并没有强化输出序列与质

收稿日期:2022-07-12

^{*} 基金项目:国家自然科学基金(61673279)、辽宁省教育厅项目(LJ2020021)资助

量变量之间的潜在联系。为了加强学习特征信息的鲁棒性,文献[7]提出了使用降噪自编码器(denoising auto-encoder, DAE)对输入序列进行重构和特征提取,防止了自编码器过拟合问题、同时也剔除了工业数据的冗余信息,重构数据在一定程度上减轻了输入序列与质量变量的代沟。但 DAE 网络通过反向传播算法来进行训练,经多次训练过后误差就会变的极小而没有作用,文献[8]提出堆叠降噪自编码器(stacked denoising auto-encoder, SDAE)来对时间序列进行深度学习特征提取解决了反向传播等问题。尽管如此还是不能精准到输入序列与质量变量的权重分配关系上,文献[9-10]提出了使用注意力机制给特征分配不同权重对发动机和轴承剩余使用寿命可以进行更加精准的预测,文献[11]提出使用注意力机制增强对重要特征的提取能力,可以有效精准对视网膜血管的分割。

先前机器学习算法面对时间序列的延迟性、动态特性以及随机性等特点,无法较好的对时间序列进行大规模训练。最开始尝试解决时间序列的循环神经网络(recurrent neural network, RNN)^[12-13]发展到具有去其糟粕、取其精华的长短期记忆网络(long short-term memory network, LSTM),解决了 RNN 的梯度爆炸和消失问题,所以 LSTM 网络^[14-15]在质量预测领域得到了广泛的运用。但是国内传统的 LSTM 网络只能提取单向时序信息及处理短批次数据,随即有人提出了双向长短期记忆网络(bidirectional long short-term memory network, BLSTM)可以对时间序列进行正逆序的学习其历史数据的趋势特征;文献[16-17]使用 BLSTM 网络对测井重构数据和风电功率进行预测,取得了比 RNN 网络和 LSTM 网络更精确的预测效果。国外倾向于利用 BLSTM 网络做文本分类^[18]或者嵌入优化算法来优化网络基本参数,提高模型预测性能;以及做一些内部控制门的改变来增强记忆网络的中长短期记忆能力,但是也没有突出网络对时间序列的预测能力。

针对上述国内外的研究问题,提出了一种嵌入注意力机制的堆叠降噪自编码器与双向长短期记忆网络(attention mechanism based stacked denoising auto-encoder, AMSDAE-BLSTM)的工业过程质量预测方法, AMSDAE 模型首先将归一化处理后的进行高斯噪声重构,实现去噪和去冗余作用,再次使用注意力机制对过程变量与质量变量进行权重分配,实现二次重构进行深度特征提取,解决了工业数据的非线性问题,还加强了编码器特征学习的鲁棒性;再将 AMSDAE 模型处理过后的重构信息代入到 BLSTM 网络中去进行正逆序学习历史数据中的趋势特征,挖掘其时间序列的深层次关系,突出特征变量对质量变量的影响从而实现精准预测。在脱丁塔过程和硫磺回收过程两个化工过程中验证了所提出方法对比 BP、LSTM、BLSTM 和 DAE-BLSTM 方法后的有效性和灵活性。

1 基础理论

1.1 自编码器

自编码器用输入基本无损复制到输出进行无监督学习,神经元层次由输入层、隐含层、输出层构成,结构上由编码器和解码器组合而成。AE 网络结构如图 1 所示。

输入序列经输入层的编码降维为隐藏序列,隐藏序列被解码器解码后映射输出序列,学习过程如下:

$$h = \sigma(w_1 x + b_1) \quad (1)$$

$$y = \sigma(w_2 x + b_2) \quad (2)$$

式中: w_1 、 w_2 分别为编码器和解码器的权重参数矩阵, b_1 、 b_2 分别为编码器和解码器的偏置参数矩阵; σ 为编码器连接处的激活函数,选择 Sigmoid、Relu 函数。

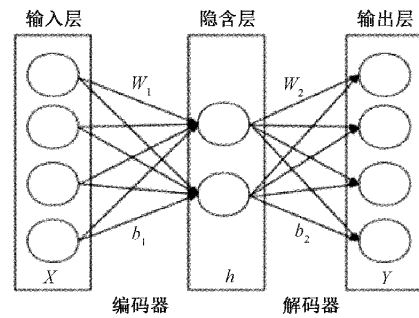


图 1 AE 网络结构示意图

自编码器在迭代过程中通过调节损失函数最小化来优化权重参数和偏置参数矩阵,以达到输入序列与输出序列之间误差的最小值,其损失函数表达式为:

$$J(W, b) = \sum \|y, x\|^2 \quad (3)$$

1.2 注意力机制

注意力机制的主要作用是为网络中隐藏层的重要特征变量赋予不同大小的权重,提高主要信息对输出的影响,提高模型的预测准确率,注意力机制公式如下:

$$e^t = u_a \tanh(W_a h_t + b_a) \quad (4)$$

$$a^t = \frac{\exp(e^t)}{\sum_{i=1}^t e^i} \quad (5)$$

式中: e^t 为注意力概率分布值; u_a 和 w_a 为注意力权重参数矩阵; a_t 为注意力得分; b_t 为注意力偏置参数矩阵; h_t 为 t 时刻神经网络隐藏层状态向量。

1.3 双向长短期记忆网络

BLSTM 网络是 LSTM 网络的基础是改进而来的,由正向 LSTM 层和反向 LSTM 层两个网络组合而成,通过双向学习时间序列的潜藏信息,该网络结构如图 2 所示。

对应的数学表达式如式(6)~(8)所示。

$$\vec{h}_{lstm} = \sigma(W_{x \rightarrow h_t} x_t + W_{h \rightarrow h} \vec{h}_{t-1} + b_h) \quad (6)$$

$$\overleftarrow{h}_{lstm} = \sigma(W_{x \leftarrow h_t} x_t + W_{h \leftarrow h} \overleftarrow{h}_{t-1} + b_h) \quad (7)$$

$$y_t = W_{h \rightarrow y} \vec{h}_t + W_{h \leftarrow y} \overleftarrow{h}_t + b_v \quad (8)$$

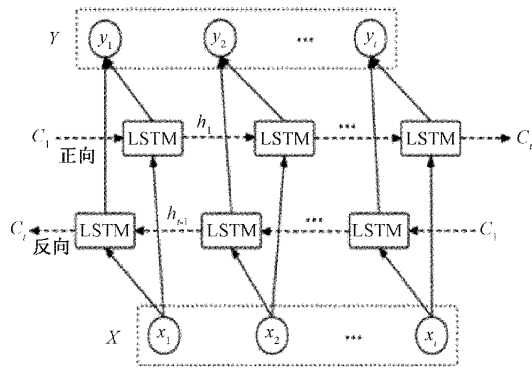


图 2 BLSTM 网络结构示意图

式中： $\overrightarrow{h}_{lstm}$ 、 \overleftarrow{h}_{lstm} 分别为 t 时刻的正向层和逆向层的隐藏向量； W_{xh_t} 为当前输入至下一隐藏层的权重； W_{hh} 为正反 $lstm$ 层中间层的权重； W_{h_o} 为当前输出层的权重矩阵； x_t 、 y_t 分别为 t 时刻的输入与输出； b_h 、 b_o 分别为隐藏层和输出层的偏置矩阵； σ 为本文选用的 *Sigmoid* 和 *Relu* 函数；参数符号上方的箭头代表方向。

1.4 误差指标

本文选取均方根误差 (RMSE)、 R^2 系数两个常用评价指标作为本文预测精度评价。RMSE 是指预测质量值与实际质量值之间的误差，可以用于表示网络在预测中会产生多大的误差，对于较大的误差，系数较高； R^2 系数是指对预测质量值，实际质量值与实际质量均值进行计算，将结果进行了标准化， R^2 系数可以用来评价模型间的差距，当 R^2 接近 1 时，说明预测效果较好，当 R^2 接近 0 时，说明预测效果较差，计算公式如下：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \in [0, +\infty) \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \in [0, +\infty) \quad (10)$$

式中： n 为样本点个数； y_i 和 \hat{y}_i 分别为第 i 个样本点所对应的实际质量值和预测质量值； \bar{y} 为实际质量均值。

2 AMSDAE-BLSTM 质量预测算法

在现代化工厂中获得的工业过程数据往往具有复杂关联性、非线性、动态性、多噪声和延迟性等特性，使得其中深层次关键特征难于挖掘，极大的限制了产品质量的精准预测。

针对预测工业过程中的数据存在各种噪声干扰，冗余性等特点，本文提出了一种基于 AMSDAE-BLSTM 算法的工业过程质量预测方法。

该模型构建如图 3 所示，首先将原始数据加入高斯噪声进行降噪，去冗余处理；其次在自编码器的隐藏层嵌入注意力机制，突出过程变量对质量变量的影响，AMSDAE 网络也增强了编码器的鲁棒性；最终通过 BLSTM 模型消除

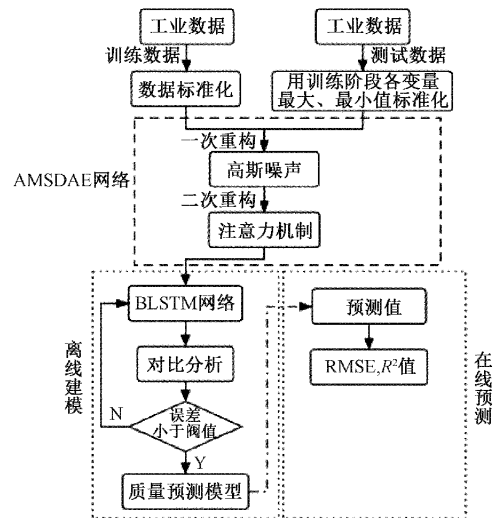


图 3 AMSDAE-BLSTM 算法预测流程

了对时间序列预测的延迟性，具体步骤如下：

离线建模部分：

1) 获取工业过程数据并划分训练集和测试集，将训练数据进行归一化处理；

2) 构建 AMSDAE 模型，设置编码器维度和高斯噪声大小范围，同时将训练重构数据与质量变量进行注意力机制权重分配，根据模型损失函数来验证模型去噪，去冗余的有效性；

3) 构建双向长短期记忆网络模型，设置网络维度、损失函数和迭代次数等，将 AMSDAE 模块输出的重构数据作为输入代入到 BLSTM 模型进行学习训练；

4) 每次训练后根据预测结果进行 RMSE, R^2 值来评估，采用网格搜索方式获取模型最优参数，误差满足阈值时得到预测模型。

测试部分：

1) 提取训练数据归一化处理中各变量的最大值与最小值对测试集进行量纲统一处理；

2) 对处理后的测试数据经过离线建模第二部分进行测试数据重构，去噪去冗余；

3) 将重构数据代入到预测模型进行质量预测并计算评价指标值。

3 案例研究

3.1 AMSDAE-BLSTM 模型主要参数设置

网络主要参数设置如表 1 所示。

3.2 单变量预测结果分析(脱丁烷塔过程)

脱丁烷塔是石油精炼过程中脱硫和石脑油分离装置的不可缺少的部分，其主要目的是解脱丁烷塔底部丁烷 (C_4H_{10}) 含量，这需要实时测量底部丁烷含量，然而， C_4H_{10} 含量是由配备在脱异戊醇装置顶部的气相色谱仪测定的，该装置与脱异戊醇装置之间有一定距离，这样对 C_4H_{10} 含

表 1 网络主要参数设置

网络	参数	设置
AMSDAE	网络结构	8-4-8/7-3-7
	激活函数	RELU
	优化器	Adam
	损失函数	MSE
	网络结构	7-1/5-2
BLSTM	批次处理数量	128
	LSTM 层个数	4
	激活函数	Linear
	优化器	Adam
	损失函数	MAE

量的测量有很大的延迟。

通过预测脱丁烷塔底部 C_4H_{10} 浓度仿真实验来验证本文方法的有效性。选取的过程变量为 7 个如表 2 所示,共采集了 2 394 个样本,将前 1 675 个样本作为训练数据,其余 719 用于测试数据。

表 2 脱丁烷塔中 7 个过程变量的描述

过程变量	描述
U1	塔顶温度
U2	塔顶压力
U3	塔顶回流量
U4	塔顶产品流出量
U5	第六层塔板温度
U6	塔底温度 1
U7	塔底温度 2

为了进一步验证本文方法的有效性,分别采用 BP、LSTM、BLSTM 和 DAE-BLSTM 方法进行浓度预测并进行对比分析。每个模型中神经网络层数、损失函数、迭代次数、激活函数等参数设置一样的情况下,评价指标如表 3、质量预测结果如图 4 所示(预测结果图随机选择了 200 个预测样本)。

表 3 单变量预测结果对比

质量变量 预测方法	C_4H_{10} 浓度	
	RMSE	R^2
BP	0.733	0.630
LSTM	1.550	0.660
BLSTM	1.173	0.051
DAE-BLSTM	0.002	0.705
本文方法	0.001	0.999

通过表 3 和图 4 可以看出:LSTM 和 BLSTM 等方法预测误差较大,拟合度也不高,方法虽然在不断更新改进,

但是对于不同类型的数据预测效果并没有明显变化。本文方法 RMSE 值接近 0, R^2 值接近 1,对单质量变量预测比其它方法有更好的预测精度。

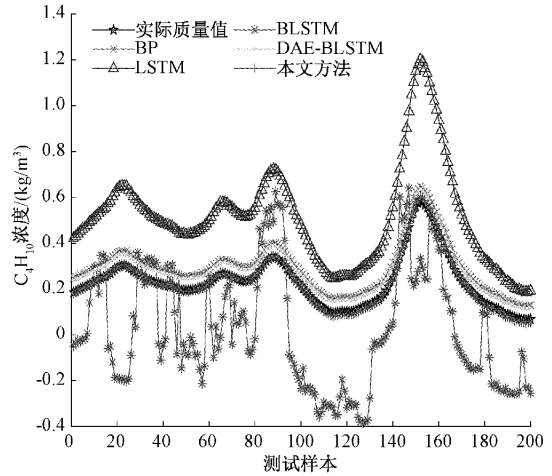


图 4 C_4H_{10} 浓度的预测结果图

3.3 多变量预测结果分析(硫磺回收过程)

硫磺回收的工艺过程是一个从废气中去除硫化物的过程,进入该装置的气体有两种,一种是 H_2S 气体,另一种是 SO_2 气体。两种气体分别被引向两个燃烧炉进行焚烧并以此去除残留的 NH_3 ,冷却后的气体进入硫磺回收和产品制备系统。最后,单体硫被回收,废气在达到排放标准后通过烟囱排放到大气。

采用硫磺回收过程的 2 个质量变量 H_2S 浓度和 SO_2 浓度的预测实验来验证所提方法的有效性。选取的 5 个过程变量如表 4 所示,共采集了 10 081 个样本,将前 9 073 个样本作为训练数据,其余 1 008 个样本作为测试数据。

表 4 硫磺回收过程中变量的描述

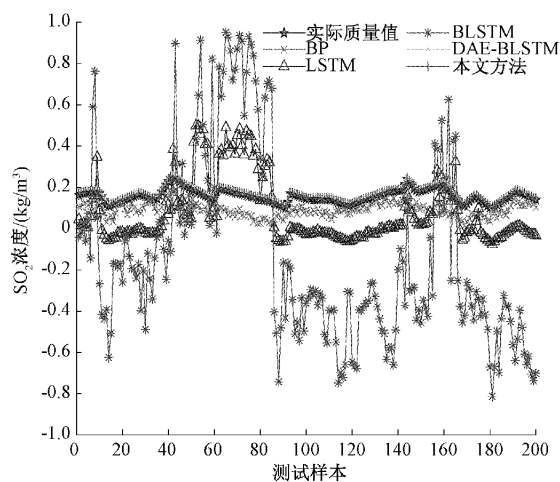
过程变量	描述
U1	MEA_GAS 气体流
U2	AIR_MEA_1 空气流
U3	AIR_MEA_2 空气流
U4	SWS 区域气体流
U5	SWS 区域空气流

为了验证本文方法的有效性,与 BP、LSTM、BLSTM 和 DAE-BLSTM 方法进行对比分析,每个模型测试集的评价指标结果如表 5 所示。

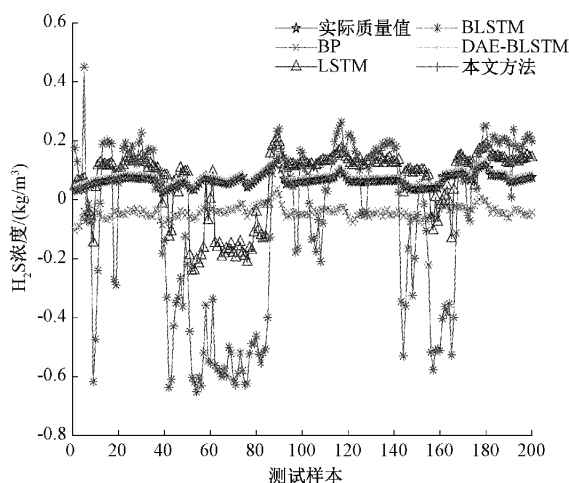
单变量预测验证本文所提算法的有效性后,接着使用硫磺回收过程数据验证了本文方法的多变量预测,由表 5 和图 5 都可以看出:本文方法在 RMSE 值都训练到了 0.001, R^2 达到了 0.999,具有很小的误差和很高的拟合度,而其他方法在 RMSE 值或 R^2 系数方面的效果都表现的不佳,预测的结果都有较大的误差和 R^2 系数,说明本文算法

表5 多变量预测结果对比

质量变量 预测方法	H ₂ S 浓度		SO ₂ 浓度	
	RMSE	R ²	RMSE	R ²
BP	0.741	0.571	0.908	0.208
LSTM	1.505	0.770	1.362	0.780
BLSTM	1.345	0.420	1.207	0.400
DAE-BLSTM	0.001	0.864	0.002	0.903
本文方法	0.001	0.999	0.002	0.999

图5 H₂S 浓度的预测结果图

对多个质量变量同时进行预测也很有效。SO₂ 浓度的预测结果如图6所示。

图6 SO₂ 浓度的预测结果图

4 结 论

本文针对工业数据的多噪声,高延迟性等特性,提出了一种基于 AMSDAE-BLSTM 的工业过程质量预测模型。此模型使用 AMSDAE 网络模块去除了工业数据的噪声干扰和减少了变量的冗余信息及对过程变量进行了权重分

配,突出了重要变量的影响力;接着使用 BLSTM 网络对时间序列进行了潜在特征提取,同时解决了时间序列预测的延迟问题以达到了精准的预测。经过对两个化工过程的单变量和多变量精准预测,验证了本文方法的有效性和稳定性。

参考文献

- [1] SHANG C, YANG F, HUANG D, et al. Data-driven soft sensor development based on deep learning technique [J]. Journal of Process Control, 2014, 24(3): 223-233.
- [2] 吕文玉,丁科,伍永平,等. 综采面支架工作阻力的 PCA-SVR 预测模型[J]. 西安科技大学学报, 2021, 41(6):973-978.
- [3] 林成龙,马义中,肖甜丽,等. 基于 PLS 方法和 Kriging 模型的多目标有效全局优化方法及应用[J]. 系统工程理论与实践, 2021, 41(7):1855-1867.
- [4] 张珂,蔡圣福,石怀涛,等. 基于改进的动态独立成分分析故障检测方法[J]. 电子测量技术, 2021, 44(16): 79-85.
- [5] 亢浏越,黄睿,孙广玲. 基于类属特征的多标签流形学习分类方法[J]. 上海大学学报(自然科学版), 2021, 27(3):525-534.
- [6] 江颀,高甲,陈铁明. 基于 AE-BNDNN 模型的入侵检测方法[J]. 小型微型计算机系统, 2019, 40(8): 1713-1717.
- [7] 高学金,马东阳,韩华云,等. 基于 DAE 和 TCN 的复杂工业过程故障预测[J]. 仪器仪表学报, 2021, 42(6): 140-151.
- [8] 李聪,彭小圣,王皓怀,等. 基于 SDAE 深度学习与多重集成的风电集群短期功率预测[J]. 高电压技术, 2022, 48(2):504-512.
- [9] 赵志宏,李晴,杨绍普,等. 基于 BiLSTM 与注意力机制的剩余使用寿命预测研究[J]. 振动与冲击, 2022, 41(6):44-50.
- [10] 陈泽慧,李博,李博. 考虑加权理论的风电场集群风速预测方法[J]. 国外电子测量技术, 2021, 40(10):34-39.
- [11] 原忠虎,雷莹,韩晓微. 基于密集空洞卷积的注意力机制视网膜血管分割方法[J]. 沈阳大学学报(自然科学版), 2022, 34(4):274-282.
- [12] 向勇,王一军,孔寅恪. 基于深度循环神经网络的大规模风力发电功率预测[J]. 资源信息与工程, 2021, 36(2):136-139.
- [13] 杨丽,吴雨茜,王俊丽,等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2):1-6.
- [14] 刘晴晴,陈华友. 基于赋权 KNN-LSTM 模型的 PM_{2.5} 质量浓度预测[J]. 合肥工业大学学报(自然科学版), 2021, 44(12):1690-1698.
- [15] 陈铁,陈一夫,李咸善,等. 基于 SDS-SSA-LSTM 的变压器油中溶解气体浓度预测[J]. 电子测量技术, 2022,

- 45(12):6-11.
- [16] 张海涛,杨小明,陈阵,等. 基于增强双向长短时记忆神经网络的测井数据重构[J]. 地球物理学进展,2022,37(3):1214-1222.
- [17] 康田雨,覃智君. 基于超参数优化和双重注意力机制的超短期风电功率预测[J]. 南方电网技术,2022,16(5):44-53.
- [18] 李超凡,马凯. 基于注意力机制结合 CNN-BiLSTM 模型的电子病历文本分类[J]. 科学技术与工程,2022,22(6):2363-2370.

作者简介

郭小萍,博士,教授,主要研究方向为基于数据驱动的复杂过程质量预测与故障诊断。

E-mail:gxp2001@sina.com

钟道金,硕士研究生,主要研究方向为工业过程软测量,质量预测。

E-mail:271162130@qq.com

李元,博士,教授,主要研究方向为基于数据驱动的复杂过程故障诊断。

E-mail:li-yuan@mail.tsinghua.edu