

基于神经网络的钻探事故类型判别模型研究

蒲春, 赵阳刚*, 杨斌, 陈映

(中国地质调查局军民融合地质调查中心, 四川成都 610036)

摘要: 钻探孔内事故会造成严重的损失, 若钻探设备能及时判断孔内事故类型, 则可缩短事故处理时间, 遏制事态发展。提出了一种基于神经网络的钻探事故类型判别模型。为了优选不同神经网络在事故类型判别时的正确率, 在Matlab的nntool工具箱中分别构建了BP、RBF两种神经网络模型, 将某矿区施工参数变化趋势作为输入参数, 通过仿真试验发现, BP神经网络中表现最好的是LM、BR算法, RBF神经网络中表现最好的是PNN算法, 三者准确率均可在90%以上, 但BP神经网络容易陷入局部最优, 性能不稳定, 偶有判别错误的现象, 而PNN神经网络无此局限, 且不需要训练。通过对比, PNN算法更适用于事故类型判别模型建立。

关键词: 钻探设备; 孔内事故; 类型判别; Matlab; BP神经网络; RBF神经网络; PNN算法

中图分类号: P634.8 **文献标识码:** A **文章编号:** 2096-9686(2023)S1-0555-06

Research on drilling fault diagnosis model of equipment based on neural network

PU Chun, ZHAO Yanggang*, YANG Bin, CHEN Ying

(Civil-Military Integration Center of China Geological Survey, Chengdu Sichuan 610036, China)

Abstract: Drilling accidents can cause serious economic losses, wasted time, and even threaten life safety. If the drilling equipment can judge the type of accident in time, the accident processing time can be shortened and the development of the situation can be contained. To solve the above problems, this paper proposes a drilling fault diagnosis model of equipment based on neural network. In order to optimize the correct rate of different neural networks in drilling accidents classification, two neural network models of BP and RBF are constructed respectively in nntool of Matlab. Through the simulation test taking the variation trend of construction parameters in a mining area as input parameters, it is found that the best performance in BP neural network is LM and BR algorithm, and the best performance of RBF neural network is PNN algorithm. All three had an accuracy rate of more than 90 percent. But BP neural network is easy to fall into local optimal with unstable performance. On the contrary, PNN neural network has no such limitation, does not require training, and the design process is simple. So PNN algorithm is more suitable for the establishment of drilling fault diagnosis model.

Key words: drilling equipment; in-hole accidents; drilling fault diagnosis; Matlab; BP neural network; RBF neural network; PNN algorithm

0 引言

钻探孔内事故会造成严重的经济损失、时间浪

费, 甚至威胁生命安全, 若钻探设备能及时判断孔内事故类型, 则可缩短事故处理时间, 遏制事态发

收稿日期: 2023-03-01; 修回日期: 2023-06-03 DOI: 10.12143/j.ztgc.2023.S1.091

基金项目: 中国地质调查局军民融合地质调查中心项目“青藏高原寒区资源与环境调查监测与评价”(编号: DD20220881)

第一作者: 蒲春, 男, 汉族, 1988年生, 工程师, 钻井工程专业, 硕士, 长期从事钻探施工及研究工作, 四川省成都市金牛区茶店子路399号, putieren@163.com。

通信作者: 赵阳刚, 男, 汉族, 1988年生, 工程师, 三矿专业, 硕士, 长期从事区域地质矿产调查工作, 四川省成都市金牛区茶店子路399号, ygzha028@163.com。

引用格式: 蒲春, 赵阳刚, 杨斌, 等. 基于神经网络的钻探事故类型判别模型研究[J]. 钻探工程, 2023, 50(S1): 555-560.

PU Chun, ZHAO Yanggang, YANG Bin, et al. Research on drilling fault diagnosis model of equipment based on neural network[J]. Drilling Engineering, 2023, 50(S1): 555-560.

展。随着钻探技术的发展,智能化钻机设备被广泛应用在地质钻探中。通过智能化技术的应用,有助于及时识别井下异常工况,并提高地质钻探作业的安全性。基于此,应逐步建立可靠的地质钻探设备井下事故判别系统,及时发现存在的风险和异常,最大限度地减少损失。根据地质钻探统计资料,在钻进过程中,大约有6%~8%的时间用于处理井下复杂情况和事故^[1]。因此及时进行事故判别,预防和快速处理井下事故,是提高钻进速度、降低成本的重要途径。目前,在针对石油的智能化钻井设备中,常用的方法包括专家系统、人工神经网络等,如张彦龙、李想等利用专家系统构建了完井作业智能化设计与监视系统^[2-3],李彤、张奇志等采用神经网络对卡钻进行预测^[4-5],杨莉、张立刚等采用神经网络对钻速进行预测^[6-7]。上述研究中,大部份是基于BP或RBF神经网络,或者是对BP算法的改进。然而在地质钻探领域,智能化工况类型识别虽有研究^[8-9],但还不多,受专家库知识较少等因素的影响,专家系统较难应用,采用神经网络来解决此问题则较为容易。人工神经网络技术具有自学习能力,在模式识别的领域应用广泛,本文拟通过对比BP、RBF神经网络的不同算法在识别正常、烧钻、卡钻、断钻、漏浆、埋钻、糊钻7种事故类型时的效果,建立判别准确率高、反应迅速的事故判别模型。

1 神经网络模型构建

1.1 BP神经网络

BP神经网络由Rumelhart提出,是神经网络体系的精华,在实际应用中约80%~90%的神经网络模型采用了BP网络或其变化形式^[10]。其结构如图1所示,BP网络采用反向传播学习算法,相邻层之间全连接,激活函数采用Log-Sigmoid或Tan-Sigmoid函数,光滑可微,具有良好的容错性和较高的精度。BP神经网络设计的主要内容包括网络层数、输入节点数、隐含节点数、输出节点数、训练方法。

1.2 RBF神经网络

传统RBF神经网络通常为3层结构:输入层、隐藏层、输出层,如图2所示。其中隐藏层为径向基函数,该函数由Broomhead和Lowe于1988年引入到神经网络中,经实践证明该函数对于非线性网络具有一致逼近的能力,特别适合解决模式分类问题^[11]。

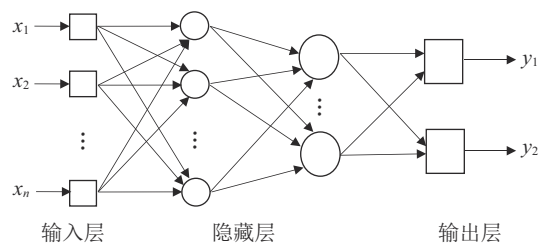


图1 BP神经网络结构

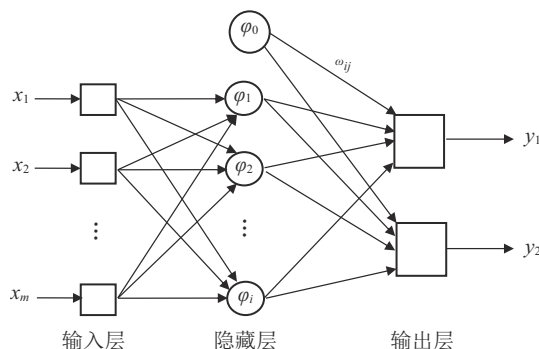


图2 RBF神经网络结构

2 数据准备

2.1 参数获取

在钻进过程中主要观察的参数有转速、扭矩、钻速、泵压、泵量、称重、大钩荷载、负荷、泥浆密度;此外还有跳钻、憋钻、岩屑等,可由中国地质科学院探矿工艺研究所发明的DPI-1型钻参仪获取^[12]。

在钻进过程中,大钩荷载一项只在提、下钻过程中使用,该项参数应舍弃。跳钻、憋钻分别可由钻压、转速两个参数获取,是重复项,应当舍弃。岩屑不能通过传感器获取,不能第一时间反应到智能识别系统中,都应舍弃。

此外,当钻机动力头输出功率一定时,功率=扭矩×转速,扭矩与转速为负相关,因转速获取较为方便,应舍弃扭矩这一参数,选择转速更为妥当;同理泵压与泵量也呈现严格的相关性,由于获取泵压更为灵敏,在这里选择泵压作为关键参数。

再者,钻进现场负荷的主要影响因素是钻机动力头输出功率与泥浆泵功率之和,与孔内工况类型没有严格的对应关系,不必选取。

最后,岩心钻探孔底钻压难以获取,而称重较为容易,在此用称重代替钻压参数。

通过以上分析,排除相关性较高的参数或者无用参数后,选取转速、钻速、泵压、返浆流量、称重、泥浆密度6个独立参数作为神经网络的输入参数^[13]。

2.2 数据来源

近年来甘肃某金矿区先后施工30多个钻孔,发生井故的钻孔有25个,累计断钻10次,卡钻6次,烧钻3次,泥浆漏失18次,孔溢1次,糊钻1次。同一种事故类型不仅在不同孔中发生,还在同一孔的不同深度发生,为提高神经网络模型的泛化能力,须对钻进参数作进一步处理,找出具有普遍适用性的参数变化特征。

2.3 数据处理

异常事故钻进参数变化趋势有两种可能,一是参数突变,如卡钻时的转速、钻速、悬重;二是缓慢变化,如漏浆时,返浆量逐渐变小。数据处理步骤如下:

第一步降噪。为减少噪声干扰,设置安全阈值的方式提高数据信噪比^[14-16]。设原始数据为: $a_1, a_2 \dots a_n$,安全阈值为 K ,即当 $a_n - a_{n-1} \leq K$ 时, a_n 录入在系统中,等待进一步处理;当 $a_n - a_{n-1} > K$ 时, a_n 作为异常数据舍去。 K 值由统计数据获取,置信水平可取95.4%。

第二步获取参数变化趋势——移动平均数^[17]。设通过上一步获取的数据为: $x_1, x_2 \dots x_n$,对每 m 个数求一次平均值,设平均值为 x'_i ,则有:

$$x'_i = \frac{x_i + x_{i+1} + \dots + x_{i+m}}{m} \quad (1)$$

i 为正整数, m 越大 x'_i 越平滑,容易过滤掉有用的突变信息,不利于异常事故判别,当 $m=3\sim 5$ 时,移动平均数具有较强的灵敏性。

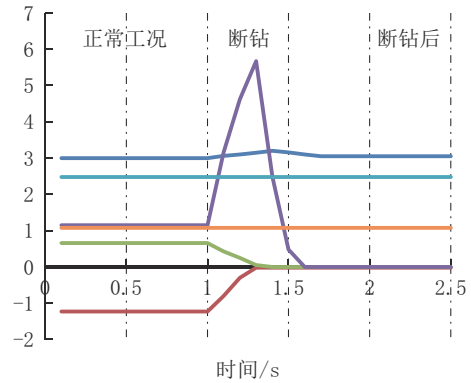
第三步计算绝对偏差。钻进参数的平均值并不能推广到同一矿区的其他钻孔中或是不同深度中,应采用某点参数与平均数的绝对偏差 Δ ,提高泛化能力^[18-20]:

$$\Delta_i = x_i - x'_i \quad (2)$$

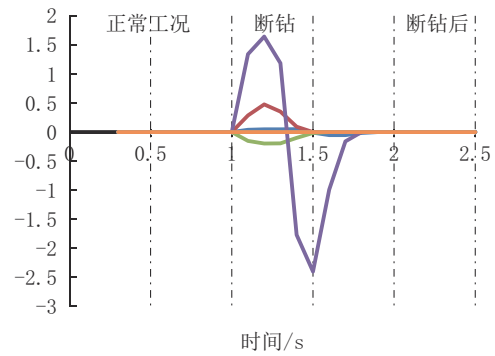
对于激增的钻进参数, Δ_i 会急剧变大;缓增的钻进参数 Δ_i 会缓慢增加,无论是哪一种,神经网络都能较好的识别。

按照上述钻进特征参数趋势获取方法,对矿区钻孔施工过程数据进行处理,选取事故发生过程中某段时间内连续采集的数据用于神经网络模型训练和检验,以正常—断钻工况为例,数据处理前后对比如图3所示。

数据处理后不仅可以使该模型适用于同一钻孔的不同孔深中,还可以适用于不同矿区的不同钻孔



(a) 数据处理前



(b) 数据处理后



图3 数据处理前后对比

中。将处理后的数据分为两部分,其中4/5组数据用于训练神经网络,剩余1/5组数据用于检测神经网络,这两部分数据均包含上述7种事故类型,网络输出采用编码方式,如表1所示。

表1 神经网络输出事故类型表示

正常	断钻	烧钻	漏浆	卡钻	孔溢	糊钻
1000000	0100000	0010000	0001000	0000100	0000010	0000001

3 不同神经网络仿真试验

3.1 BP神经网络仿真

在Matlab GUI的nntool工具箱中选取级联前向神经网络(Cascade-forward BP)和前馈神经网络(Feed-forward BP)建立事故类型判别模型。输入神经元个数为钻进特征参数个数(6个),输出神经元个数为事故类型种类(7个)。

3.1.1 训练函数优选

先假设隐藏层节点个数为3个,分别用不同的

训练函数对 Cascade-forward BP 网络和 Feed-forward BP(见表2)网络进行训练。

表2 Feed-forward、Cascade-forward BP神经网络训练算法对比

BP类型	训练算法	Feed-forward BP			Cascade-forward BP		
		训练时间/s	训练正确率/%	检验正确率/%	训练时间/s	训练正确率/%	检验正确率/%
BFG	拟牛顿反向传播法	<1	17.39	18.18	1~2	47.82	27.27
RP	弹性反向传播法	<1	13.04	0	<1	34.78	9.09
SCG	共轭梯度法	<1	0	0	<1	21.73	9.09
r	随机训练法	126	39.13	18.18	148	69.56	72.72
oss	弦截法	<1	13.04	18.18	<1	60.87	45.45
GDX	自适应、动量因子梯度递减法	<1	0	0	<1	13.04	0
GDA	自适应梯度递减法	<1	17.39	18.18	<1	69.56	45.45
GDM	动量因子梯度递减法	<1	0	0	4~5	0	0
GD	梯度递减反向传播法	<1	0	0	3~4	0	0
CGP	Polak-Ribiere变梯度反向传播法	<1	17.39	9.09	<1	56.52	45.45
CGF	Fletcher-Powell变梯度反向传播法	<1	0	0	<1	39.13	9.01
CGB	Powell-Beale共轭梯度反向传播法	<1	0	0	<1	43.47	27.27
BR	贝叶斯归一化法	1~2	100	72.72	1~2	100	72.72
LM	Levenberg-Marquard反向传播法	<1	100	36.36	<1	91.30	72.72

Feed-forward 与 Cascade-forward BP 神经网络训练算法效果最好的均是BR法,其次是LM法,但 Cascade-forward BP网络性能无论是训练时间、还是检验准确率都优于Feed-forward BP神经网络。

3.1.2 隐藏层神经元个数确定

针对 Cascade-forward BP网络,优化隐藏神经元个数,结果如表3所示。

表3 Cascade-forward BP网络隐藏层神经元个数优化

算法	LM			BP			
	神经元个数	训练时间/s	训练正确率/%	检验正确率/%	训练时间/s	训练正确率/%	检验正确率/%
	2	<1	100	72.72	1~2	96	73.00
	3	<1	100	72.70	1~2	100	81.81
	4	<1	100	90.10	1~2	100	81.82
	5	<1	100	72.72	1~2	100	90.10
	6	<1	100	63.63	2~3	100	81.82
	7	<1	100	90.10	1~2	100	81.82

BR算法准确率和稳定性要优于LM算法,但训练时间比LM算法长,因此如果训练数据集庞大时,应当优先选择LM算法,如果训练数据集较小,应当

优先选用BR算法。LM算法的最优隐藏神经元个数为4个,BR算法最优隐藏神经元个数为5个,两者训练误差曲线如图4、图5所示。

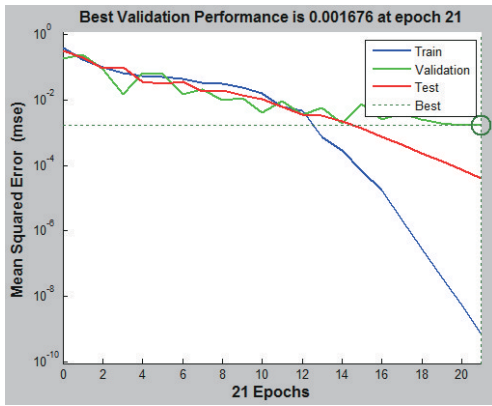


图 4 BR 算法训练误差曲线

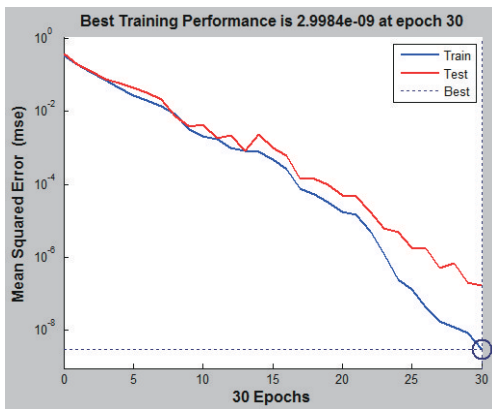


图 5 LM 算法训练误差曲线

此外, BP 神经网络在同样参数条件下每次试验结果不一样, 这是由于 BP 算法初始权值是随机给定的; 同样仿真试验中, 偶有少数训练结果与其他结果差别很大, 这是由于 BP 神经网络容易陷入局部最优引起的。

3.2 RBF 神经网络仿真

GUI 界面设计中 nntool 工具箱包含 4 种 RBF 神经网络, 即概率神经网络 (PNN)、正则化神经网络 (正则化 RBF)、广义神经网络 (广义 RBF) 和广义回归神经网络 (GRNN)。4 种神经网络结构类似, 设置输入、输出神经元个数分别为 6、7 个, 隐藏层神经元个数除广义网络外, 其余 3 种均为训练数据组数, 广义网络隐藏层神经元个数在每次运行时, 会自行增加隐藏层神经元个数直到网络性能达到设置的目标误差。

将 4 种 RBF 神经网络的扩展常数 (Spread constant) 从 1 设置到 0.001 时, 各 RBF 神经网络检测正确率对比如表 4 所示。

性能最好的是 PNN 网络 ($0.01 \leq \text{spread} \leq 0.1$), 以及 GRNN 网络 ($\text{spread} \leq 0.01$)。PNN 的 spread 值

表 4 RBF 神经网络的检测性能

spread	正则化 RBF/ %	广义 RBF/ %	PNN/ %	GRNN/ %
1	54.54	18.18	54.54	54.54
0.1	63.63	63.63	90.91	63.63
0.01	9.09	9.09	90.91	90.91
0.001	18.18	27.27	45.45	90.91

大于 GRNN, 显示其泛化能力更强, 计算时间更少, 因此 PNN 优于 GRNN, 对 PNN 网络进一步优化, 结果表 5 所示。

表 5 PNN 网络性能优化

spread	0.1	0.3	0.5	0.6	0.7	0.9
PNN/%	90.91	90.91	90.91	90.91	72.7	72.7

当 $\text{spread}=0.6$ 时, PNN 网络性能即可达到最优, 检验正确率可达 90.91%。

3.3 事故判别模型实现

BP 神经网络性能不稳定, 且容易陷入局部最优, 判断准确率也不如 RBF 神经网络。RBF 神经网络无需训练, 判断准确率更高, 不存在局部最优和性能不稳定现象, 但由于 RBF 神经网络隐藏层神经元个数等于训练组数, 因此当训练集较少时, RBF 神经网络更具优势。在本文中, 发生事故时间短, 短时间产生的数据较少, 因此训练集也较小, 但也能达到 90.91% 的准确率, 能够满足钻探设备事故判别的智能化需求, 因此本文拟采用 RBF 神经网络 PNN 算法构建钻探设备事故识别模型 (如图 6、图 7 所示), spread 值设为 0.6。

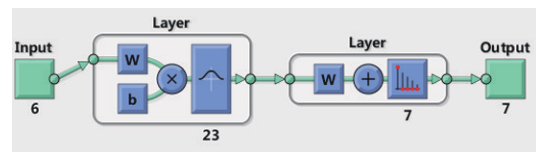


图 6 概率 (PNN) 神经网络结构

在 Matlab 中用 gensim() 函数生成事故判别模型的仿真 (simulink) 模块, 如图 8 所示, 在参数 (constant) 中任意输入一组处理后的事故参数数据, 然后运行模型, 将在显示器 (display) 中显示事故判别结果, 图中所示编码为烧钻。通过仿真试验证明 PNN 模型能够取得较为满意的应用效果。

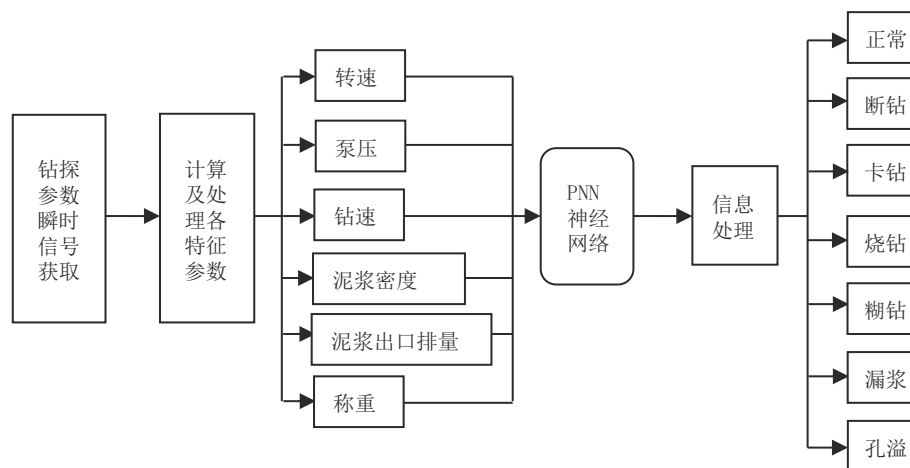


图7 事故判别模型流程

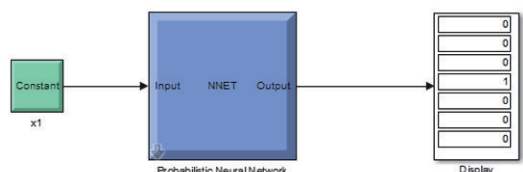


图8 概率(PNN)神经网络模型仿真实现

4 结论

通过上述的建模仿真可以看出,能够应用于钻探设备工况自动化识别的神经网络有很多种,传统的BP神经网络参数设置复杂,计算过程太多,并且需要训练,检测正确率也要明显低于RBF神经网络。同时,由于BP算法容易陷入局部最优,在作事故识别时稳定性较差,容易得到错误结果,耽搁施工进度。而RBF神经网络具有设置参数少的优点,只需要设置扩展参数 spread 即可,其中PNN算法检测正确率可以达到90%以上,具有极高的事故类型判断正确率,说明RBF神经网络PNN算法更适用于钻井设备的事故模式判别。

参考文献:

- [1] 蒋希文. 钻井事故与复杂问题(第二版)[M]. 北京:石油工业出版社,2006.
- [2] 张彦龙,葛云华,张晓林,等. 钻完井工程技术人工智能专家系统初探[C]//西安:2019油气田勘探与开发国际会议,2019:1251-1253.
- [3] 李想. 完井钻方方案辅助设计方法的研究与实现[D]. 大庆:东北石油大学,2015.
- [4] 李彤. 基于神经网络的井下卡钻预测研究[D]. 西安:西安石油大学,2021.
- [5] 张奇志,朱丹,刘光星. 基于神经网络的卡钻预测[J]. 石油石化

- 节能,2013,3(1):5-7,59.
- [6] 杨莉,鹿卓慧,任伟建,等. 基于模糊神经网络的钻速预测方法[J]. 吉林大学学报(信息科学版),2022,40(6):970-978.
- [7] 张立刚,苗振华,黄小刚,等. 基于MESA-BP神经网络的钻井机械钻速预测[J]. 自动化与仪表,2022,37(11):87-92.
- [8] 甘超,曹卫华,王鲁朝,等. 深部地质钻探钻进过程流式大数据分析 with 动态预处理——以辽宁丹东3000 m科学钻探工程为例[J]. 钻探工程,2022,49(4):1-7.
- [9] 文浩贤,杜胜,张典,等. 地质钻探过程状态监测App软件设计与应用[J]. 钻探工程,2022,49(4):41-48.
- [10] 陈明,等. 神经网络原理与实例精解[M]. 北京:清华大学出版社,2013.
- [11] 潘海洋,杨宇,郑近德,等. 基于径向基函数的变量预测模型模式识别方法[J]. 航空动力学报,2017,32(2):500-506.
- [12] 罗光强,胡郁乐. 科学深钻DPI-1智能化多功能钻参仪的研制与应用研究[J]. 地质与勘探,2014(50):777-782.
- [13] 胡远彪,蒲春,刘宝林,等. 基于神经网络的岩心钻探工况判别模型研究[C]//中国地球科学联合学术年会,2016:522-525.
- [14] 李辉,徐伟丞. 循环相关熵和一维浅卷积神经网络轴承故障诊断[J]. 机械科学与技术,2022(10):89-98.
- [15] 陈云. 基于卷积神经网络的地震勘探噪声压制算法研究[D]. 吉林:东北电力大学,2022.
- [16] 张强,刘志恒,王海舰,等. 基于BP神经网络的截齿磨损程度在线监测[J]. 中国机械工程,2017,28(9):1062-1068.
- [17] 马一琳. 基于深度神经网络的股价预测及投资组合研究[D]. 南京:东南大学,2022.
- [18] 杨霞,朱晓冬,刘元宁,等. 分块小波特征结合BP神经网络的虹膜识别方法[J]. 计算机工程与应用,2019,55(18):132-139.
- [19] 张方泽. 基于中位数绝对偏差和自训练卷积神经网络的遥感图像异常训练样本探测[D]. 上海:华东理工大学,2022.
- [20] 赵琳娜,卢姝,齐丹,等. 基于全连接神经网络方法的日最高气温预报[J]. 应用气象学报,2022,33(3):257-269.

(编辑 荐华)