

DOI:10.13296/j.1001-1609.hva.2024.07.018

变压器油中乙炔门控循环单元网络多步预测超参数优化方法

赵 军¹, 高树国¹, 何瑞东¹, 相晨萌¹, 芮逸凡², 王亚林², 尹 毅²

(1. 国网河北省电力有限公司电力科学研究院, 石家庄 050021; 2. 上海交通大学电气工程系, 上海 200240)

摘要: 油中溶解乙炔作为电力变压器中重要的放电程度表征参量之一, 对其进行多步预测可以为变压器故障诊断及预警提供重要依据。现有的状态预测模型主要集中于单步预测, 对于未来更长时期变化趋势的预测手段不足。此外, 基于深度学习的多步预测模型的超参数选择大多基于经验和朴素的单一控制变量法, 超参数之间的耦合关系没有得到充分的研究。文中提出基于多输出策略的门控循环单元(gated recurrent unit, GRU)神经网络多步预测模型, 通过改变模型结构超参数和训练超参数研究超参数之间的耦合关系, 使用多目标灰狼优化算法对不同预测结果倾向的GRU模型进行超参数优化。结果表明, GRU模型可以较为准确的对变压器油中乙炔含量进行30天预测, GRU模型各超参数对输出预测结果的影响规律并不统一且相互影响, 一组超参数无法同时达到多目标最优。多目标灰狼优化算法能够根据预测目标的不同, 优化选择合适的超参数, 为人工智能算法超参数的选取提供参考。

关键词: 变压器; 乙炔; 门控循环单元(GRU); 灰狼算法; 多步预测

Hyper-parameters Optimization Method for Multi-step Prediction of Acetylene in Power Transformer Oil by Gated Cyclic Unit Network

ZHAO Jun¹, GAO Shuguo¹, HE Ruidong¹, XIANG Chenmeng¹, RUI Yifan²,

WANG Yalin², YIN Yi²

(1. State Grid Hebei Electric Power Research Institute, Shijiazhuang 050021, China; 2. Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Multi-step prediction of the acetylene dissolved in oil, one of the important discharge degree characterization parameters in power transformer, can provide an important basis for fault diagnosis and early warning of transformer. The existing state prediction models mainly focus on single-step prediction and are insufficient for the prediction means with long time variation trend in the future. In addition, the hyper parameter selection of multi-step predictive model based on deep learning is mostly based on experience and the naive single control variable method, and the coupling relationship between hyper-parameters has not been fully studied. In this paper, a multi-step prediction model of gated recurrent unit (GRU) neural network based on a multi-output strategy is proposed. The coupling relationship of the hyper parameters is studied by changing the structure hyper parameters and training hyper parameters of the model, and the multi-objective gray wolf optimization algorithm is used for hyper parameter optimization of GRU models with different prediction results tendency. The results show that the GRU model can accurately predict the acetylene content in transformer oil for 30 days. The influence of the hyper parameters of the GRU model on the output prediction results is not uniform and affects each other. The same set of hyper parameters cannot achieve multiple goals optimization simultaneously. The multi-objective gray wolf optimization algorithm can optimize the selec-

收稿日期: 2024-03-01; 修回日期: 2024-05-09

基金项目: 国网河北省电力有限公司科技成本项目(kjcb2020-013)。

Project Supported by Science and Technology Cost Project of State Grid Hebei Electric Power Co., Ltd.(kjcb2020-013).

tion of suitable hyper parameters in accordance with the different prediction targets, which provides a reference for the selection of hyper parameters of artificial intelligence algorithms.

Key words: transformer; acetylene; gated recurrent unit; gray wolf algorithm; multi-step prediction

0 引言

电力变压器是电力系统的重要组成部分之一,其安全稳定运行是保障电力系统可靠性的必要基础。当前针对变压器进行故障诊断的目标主要集中于是否存在故障、故障出现的部位或是故障引起的原因等层面^[1-7]。油中溶解气体分析^[8-12](dissolved gas analysis, DGA)技术是当前用于监测电力变压器缺陷与潜伏性故障最为有效的方法。利用油色谱数据形成的信息时间序列,对变压器油中溶解气体浓度的发展趋势进行预测^[13-15],对于变压器故障的提前预警具有重要参考意义。

目前,业界学者对于时间序列未来变化趋势的预测方法进行了大量的研究,电气行业也正逐步将多种时间序列预测方法用于电力设备的状态趋势预测。对于预测方法主要可以归纳为以下两大类:统计预测方法和智能预测方法^[16]。

统计预测方法主要方法包括多元线性回归(multiple linear regression, MLR)、自回归(auto regression, AR)、自回归移动平均(auto regressive moving average, ARMA)等^[17-18]。该类方法要求具有一定规模的数据且数据呈现的规律性好,才能进行递推预测。但是变压器油色谱的各个特征量受环境、负荷以及老化程度等多重因素的影响,具有较强的随机性,数据趋势的平稳性较差。智能预测方法则是利用计算机人工智能技术,建立用于时间序列数据预测未来发展趋势的模型,并通过大量历史数据的训练分析,不断更新优化模型^[19]。主要的方法包括:BP神经网络(back propagation neural networks, BPNN)^[20]、支持向量机(support vector machine, SVM)^[21-23]等。但是传统的人工智能方法对于时间序列数据的前后关联性不敏感,不能有效捕捉时间维度上的关联性,对于发展趋势预测的精度有限。随着神经网络技术的不断发展,循环神经网络(recurrent neural network, RNN)得到了广泛地运用。RNN最大的特点在于网络内部神经元在某时刻的输出可以再次输入到神经元,如此循环反馈的串联网络结构可以保持数据前后的依赖关系^[24],适合用于时间序列的建模。在此基础上改进得到的门控循环单元(gated recurrent unit, GRU)网络^[25]通过加入门控机制改善了RNN中存在的长期数据依赖性问题,避免RNN

在数据训练过程中出现梯度消失与梯度爆炸等问题。与当前普遍运用于序列预测计算的长短期记忆神经网络相比,GRU简化了结构,减少运算时间^[26]。此外,目前深度学习中多步预测模型的超参数选择大多基于经验或朴素的单一控制变量法,超参数之间的耦合关系没有得到充分的研究^[27]。

文中建立基于多输出策略的GRU神经网络变压器油中溶解乙炔多步预测模型,对乙炔量进行30天时间步长的变化趋势预测,通过改变模型结构超参数和训练超参数研究超参数之间的耦合关系,使用多目标灰狼优化算法对不同预测结果倾向的GRU模型进行超参数优化。

1 循环神经网络

1.1 循环神经网络

RNN是一类用于处理序列数据的神经网络,网络的自反馈机制使其具备记忆功能。输出结果除了受到当前时刻的输入数据影响外,还会受到之前数个时刻输入数据的影响。RNN对于数据在时间上的前后关联性更为敏感,在自然语言处理领域具有十分广泛地运用^[28]。

典型的RNN拓扑结构见图1,由3部分组成,分别为网络的输入层 x 、隐藏层 s 以及输出层 o 。三者之间在 t 时刻的关系见式(1)、(2):

$$s_t = f(Ux_t + Ws_{t-1} + b_h) \quad (1)$$

$$o_t = g(Vs_t + b_o) \quad (2)$$

式(1)、(2)中: f, g 分别表示隐藏层与输出层所运用的激活函数; s_{t-1} 表示隐藏层在 $t-1$ 时刻的状态; U, V 分别表示输入层与隐藏层之间以及隐藏层与输出层之间的权重关系矩阵; W 表示隐藏层自反馈权重矩阵; b 表示隐藏层与输出层的偏置。

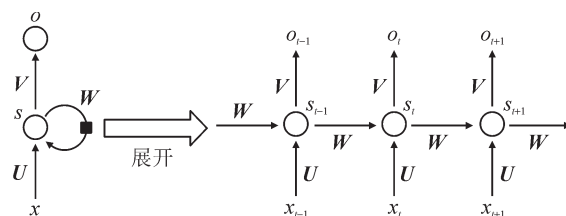


图1 循环神经网络结构

Fig. 1 Architecture of recurrent neural network

实际运用中,RNN对于短期记忆有很好的储存,但是在面对长时间序列时矩阵权重会发生指数级爆炸或是梯度消失问题,不能很好地学习长时间

数据的关联^[29]。

1.2 门控循环单元神经网络

在RNN结构中保留循环反馈机制的基础上,引入门控机制形成门控循环单元神经网络,门控机制包含重置门(reset gate)与更新门(update gate)。其中重置门的作用是控制前一时刻状态信息的忽略程度,重置门的值越大说明历史信息忽略得越少,比重也就越大;更新门则负责控制前一时刻的状态信息延续至当前状态的程度,更新门的值越大说明历史状态信息代入越多。通过门控机制改变序列历史信息的读取与修改,从而做到对于信息权重的控制,并避免梯度爆炸或梯度消失。其拓扑结构见图2。

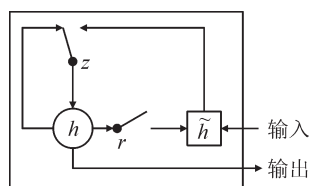


图2 门控循环单元神经网络结构

Fig. 2 Architecture of GRU memory block

在 t 时刻,GRU模块的序列输入、输出分别为 x 和 h , r 与 z 则分别代表GRU中的重置门与更新门,激活函数采用sigmoid函数以保证门的输出固定在0~1之间, $\tilde{h}_t$ 表示经过重置门处理后的输出候选值,如式(3)~(6)所示:

$$r_t = \sigma(W_r h_{t-1} + W_x x_t) \quad (3)$$

$$z_t = \sigma(W_z h_{t-1} + W_x x_t) \quad (4)$$

$$\tilde{h}_t = \tanh[W_h(r_t * h_{t-1}) + W_o x_t] \quad (5)$$

$$h_t = (1 - z_t) * \tilde{h}_t + z_t * h_{t-1} \quad (6)$$

式(3)~(6)中: W_r 、 W_x 表示重置门中的权重关系矩阵; W_h 、 W_z 表示更新门中的权重关系矩阵; W_o 、 W_x 表示输出候选值中的权重关系矩阵;“*”符号表示点乘; σ 表示sigmoid激活函数。

2 基于GRU多输出策略的变压器油中乙炔体积分数多步预测模型

变压器油中溶解乙炔对于油中溶解气体的分析较为重要,乙炔含量与变压器内部放电的程度息息相关,也是变压器运维过程中重点关注的状态量。

2.1 数据预处理

文中对某变电站500 kV电力变压器在线监测油色谱乙炔数据进行多步预测。数据记录长度为1 080天,监测周期为1天,监测数据量共1 080个数据点。按照上一节所述的数据结构构造训练集和验证集,随机打乱顺序避免人为划分造成系统性误

差,以免模型预测效果不具有广泛的适用性。乙炔数据曲线见图3。

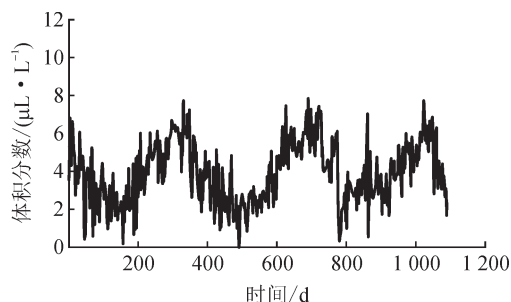


图3 某500 kV变压器乙炔监测数据

Fig. 3 Acetylenedata of a 500 kV transformer

文中GRU预测模型采取梯度求导的反向传播算法对数据进行训练^[30],过大或者过小的数据都会使得迭代算法收敛速度变慢甚至不收敛,以至于难以找到最优解。对数据采用归一化到标准区间的预处理方式可以使得运算收敛速度更快,并提高运算的精确程度。文中采取最小—最大归一化方法将每一个特征量归一化至[0,1]区间内。

2.2 GRU多步预测结构

文中构建的GRU预测算法模型见图4。输入数据首先通过单层GRU层,进行时间步规律学习,再通过退出(Dropout)层防止过拟合。Dropout层是指深度学习训练过程中,对于神经网络训练单元按照一定的概率将其从网络中暂时移除。Dropout层的运行步骤是:首先设定Dropout层的概率,根据相应的概率拿掉一部分的神经元然后开始训练,更新没有被拿掉神经元以及权重的参数,将其保留。参数全部更新之后,又重新根据相应的概率拿掉一部分神经元,然后开始训练,如果新的用于训练的神经元已经在第一次当中训练过,那么继续更新它的参数。而第二次被剪掉的神经元,同第一次已经更新过参数的,保留它的权重,不做修改。Dropout层由于暂时地忽略了部分神经元,减小了运算复杂度,使得模型能够关注曲线整体的趋势。

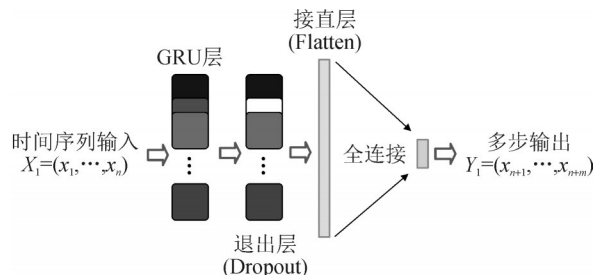


图4 GRU多步预测模型结构

Fig. 4 Multi-step prediction structure of GRU

数据经过Dropout层后将多维输出拉平成一维向量,再经过全连接层后输出预测结果。

为能够预留更多时间对变压器的状态进行预测并防止设备故障,模型需要做出多时间步长的预估和长期预测。通常,多时间步长预测的策略有直接策略、递归策略和多输出策略3种,分别见式(7)-(9):

$$\begin{cases} \hat{y}_{t+1} = \text{Model}_1(x_t, x_{t-1}, \dots, x_{t-n}) \\ \hat{y}_{t+2} = \text{Model}_2(x_t, x_{t-1}, \dots, x_{t-n}) \end{cases} \quad (7)$$

$$\begin{cases} \hat{y}_{t+1} = \text{Model}(x_t, x_{t-1}, \dots, x_{t-n}) \\ \hat{y}_{t+2} = \text{Model}(\hat{y}_{t+1}, x_t, \dots, x_{t-n+1}) \end{cases} \quad (8)$$

$$(\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+m}) = \text{Model}(x_t, x_{t-1}, \dots, x_{t-n}) \quad (9)$$

式(7)-(9)中: x 为实测数据; $\hat{y}$ 为预测数据。直接策略使用不同的模型来预测每个时间步的数据,它的缺点是无法对预测数据之间的依赖关系进行建模。递归策略涉及到多次使用同一个单时间步模型,前一个时间步的预测用作输入以对下一个时间步长进行预测。然而,递归策略难以避免预测误差累积,从而随着预测时间的增加,预测精度会迅速下降。多输出策略只涉及一种模型,该模型能够以一次性方式从历史数据中预测多个时间步的数据。多输出模型更复杂,因此训练速度较慢,需要更多的训练数据才能达到令人满意的精度。

在文中,使用多输出策略来进行多时间步预测,即建立多输入对多输出的模型进行预测。

2.3 GRU神经网络的超参数对预测结果的影响

在机器学习中,超参数是指在开始数据的学习训练过程之前通过人为设定取值的参数,而非在学习过程中通过训练更新的参数。超参数的合适与否会对模型运行的结果有至关重要的影响,因此需要对超参数进行调试优化,以提高学习的性能和效果。

目前大多数研究采用控制变量法研究单一超参数变化对模型预测结果的影响,而忽略了多个超参数之间的耦合关系,进而无法对多个超参数进行多维优化。本节首先研究GRU模型中的多个超参数对模型预测结果的影响规律,第3节中将具体阐述多个超参数的优化方法。

2.3.1 超参数类型

文中GRU多步预测模型的超参数主要可以分为两类:与输入数据有关的模型训练超参数以及预测模型有关的模型结构超参数。

模型的训练超参数包括输入时间序列步长即输入时间步(time-step)和学习率(learning rate)。输入时间步是指模型用来进行预测输出量的输入序列长度,时间步的选择需要综合考虑历史信息的完整性以及时间长度对于预测结果的有效性。学习率是指每次在训练过程中,每个训练参数的修改程

度,学习率过高可能会导致震荡,学习率过低会导致训练效率过低,收敛速度过慢。

模型结构超参数包括Dropout层的Dropout概率和GRU层的单元个数。Dropout层的概率越小,整个网络的误差也就越小,但过小的Dropout概率会使整个网络复杂化,增加网络的训练时间,出现“过拟合”,使得预测结果有较大偏差。GRU层单元数的选取同样需要平衡模型的学习能力与模型复杂程度,单元数过少,模型不具有泛化能力;单元数过多,容易陷入局部最优,出现过拟合。

文中选取两种指标对模型预测的效果进行评估,分别是均方根误差RMSE(root mean square error)和平均绝对误差MAE(mean absolute error),计算式为:

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

$$e_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

式(10)、(11)中, y_i 、 $\hat{y}_i$ 用于表示预测数据的真实值、预测值。

2.3.2 超参数之间的耦合关系

针对图4中的GRU多步预测模型超参数,分别设置:时间序列长度为3、7、15、30;学习率为0.5、0.2、0.1、0.05;Dropout层概率为0.5、0.1、0.01、0.001,GRU单元数量为8、16、32、64,共 $4^4=256$ 组参数。训练上述256组GRU模型,设置输出步长为5,训练100次,对比预测误差。

1)训练超参数对预测结果的影响。由于可以比较的超参数有4组,加上因变量共5维数据,无法通过三维作图来表示相互关系,因此将结构超参数即Dropout概率和GRU层单元的4个极值组合固定,即极大值极小值相互搭配共4组,研究训练超参数即输入时间序列步数和学习率对预测结果的影响,见图5。由于RMSE和MAE的变化趋势类似,图5只展示RMSE的结果。

时间序列步数越大,预测结果整体较为准确。然而时间序列步数的影响规律并不一致,在学习率较低时,如学习率为0.001时,时间序列步数对输出结果的影响没有明显规律。学习率对输出结果的影响也呈现较为分散的特性,当学习率过低或过高时(0.001和0.5),预测结果均较差。

此外,结构超参数的不同也会对训练超参数的效果产生影响。当GRU单元数相同时,不同的Dropout概率下训练超参数的影响规律没有较大的变化,但是Dropout概率相同时,GRU单元数越大,

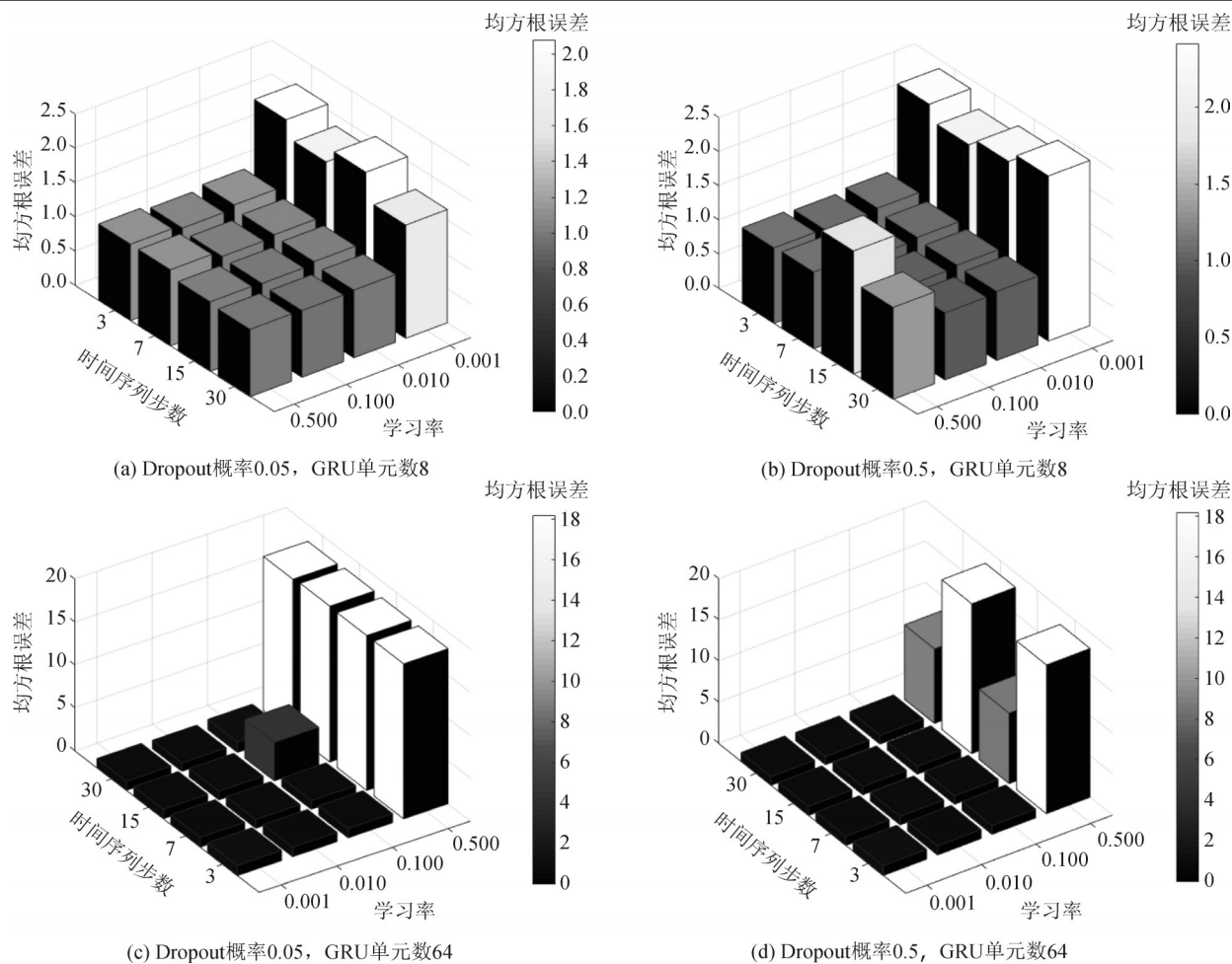


图5 训练超参数对预测误差的影响

Fig. 5 Effect of training hyperparameter on prediction error

学习率的影响程度越大。GRU单元越多,学习率的波动引起输出结果的误差变化越大。这是因为GRU单元数越多,需要学习的参数越多,较大的学习率更容易引起学习结果震荡。

2)结构超参数对预测结果的影响。将训练超参数即学习率和输入时间序列步数的4个极值组合固定,即极大极小值搭配共4组,研究结构超参数即Dropout概率和GRU层单元数对预测结果的影响,见图6。

整体来说,GRU单元数越多,预测结果越准确。然而,GRU单元数的影响规律也受训练超参数的影响:当学习率较高时,如学习率为0.5时,GRU单元数为64的预测误差反而最大,并且远远高于GRU单元数为8~32的预测误差值。GRU单元的影响规律对时间序列步数的敏感度不高,即时间序列步数的变化对GRU单元的规律影响不大。整体来说,Dropout层概率对预测准确率影响不大,可能是因为训练次数不够多,过拟合状态还没有表现出来。但是当学习率较小且GRU单元数较少时,以及学习率较高GRU单元数较多时,Dropout层概率对

预测准确率影响较大,这是因为当学习率较小且GRU单元数较少时,相同的训练次数不足以使得模型学习到复杂的变化规律。而当学习率较高GRU单元数较多时,过高的训练参数变化速率又导致了预测结果的震荡,无法收敛到较佳的结果。

综上所述,可以看出GRU预测模型的表现受到结构超参数和训练超参数的联合控制,结构超参数和训练超参数之间耦合性较强,某些超参数之间存在拮抗关系,需要通过优化算法选取较为合理的超参数构造GRU模型并对其进行训练。

3 基于多目标灰狼算法的GRU模型超参数优化

如上所述,GRU模型的结构超参数和训练超参数对模型训练效果和最终的预测效果具有较大的影响。此外,对于多步预测,不同时间步的预测准确率也有所不同。实际情况下,有时运维人员关注短期预测的准确性,比如监测量快要接近预警值时;有时关注变压器监测量长期的变化趋势,比如监测量处于平稳的状态。因此有必要对GRU模型

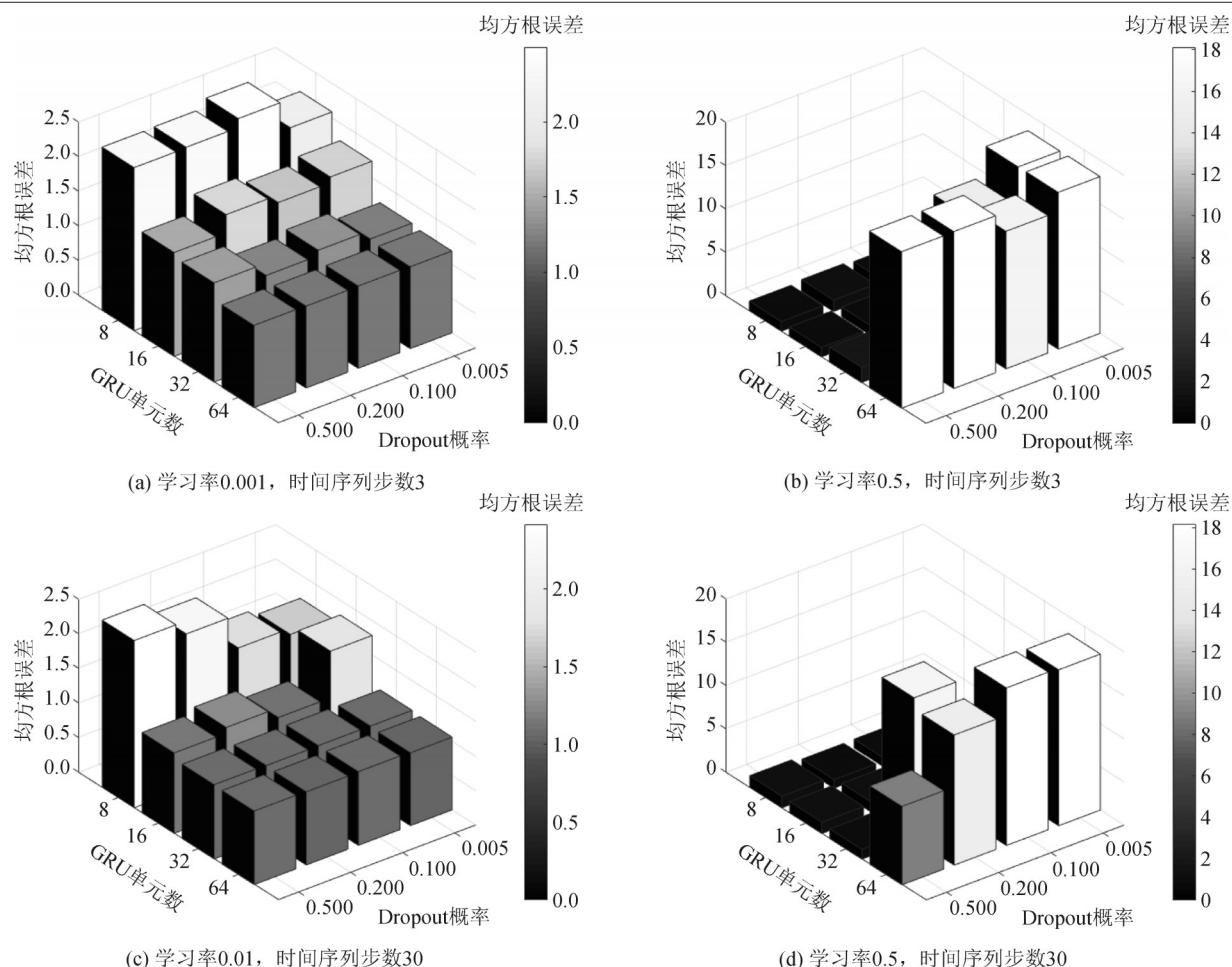


图6 模型超参数对预测误差的影响

Fig. 6 Effect of model hyperparameter on prediction error

在不同时间步的预测效果进行研究,进而为可能的多目标优化提供参考。

3.1 多目标灰狼优化算法

灰狼算法是受到灰狼捕食猎物活动的启发而开发的一种优化搜索方法,它具有收敛性强、参数少、易实现等特点。存在能够自适应调整的收敛因子以及信息反馈机制,能够在局部寻优与全局搜索之间实现平衡,因此在对问题的求解精度和收敛速度方面都有良好的性能^[31]。

灰狼种群具有金字塔形的社会等级层次制度。金字塔第1层为种群中的领导者,称为 α 。金字塔第2层是 α 的智囊团队,称为 β 。金字塔第3层是 δ , δ 听从 α 和 β 的决策命令。适应度不好的 α 和 β 也会降为 δ 。金字塔最底层是 ω ,主要负责种群内部关系的平衡。灰狼的社会等级在群体狩猎过程中发挥着重要的作用,捕食的过程在 α 的带领下完成。假设 α 、 β 和 δ 更了解猎物的潜在位置,则保存迄今为止取得的3个最优解决方案,并利用这三者的位置来判断猎物所在的位置,同时强迫其他灰狼个体(包括

ω)依据最优灰狼个体的位置来更新其位置,逐渐逼近猎物。灰狼的狩猎包括以下3个主要步骤:①跟踪、追逐和接近猎物;②追捕、包围和骚扰猎物,直到它停止移动;③攻击猎物。灰狼算法也模拟灰狼狩猎中的这3个过程,实现逐步逼近最优解的目标。

灰狼围捕猎物的行为定义为^[31]:

$$D = |C \cdot X_p(t) - X(t)| \quad (12)$$

$$X(t+1) = X_p(t) - A \cdot D \quad (13)$$

式(12)、(13)中: t 是目前时刻; A 和 C 是系数向量; X_p 是猎物的位置向量; X 是灰狼种群的位置向量。式(12)表示个体与猎物之间的距离,式(13)是灰狼的位置更新公式。 A 和 C 的计算式为^[31]:

$$A = 2ar_1 - a \quad (14)$$

$$C = 2 \cdot r_2 \quad (15)$$

式(14)、(15)中: a 是从2到0随着迭代次数增加而线性减小的收敛因子向量; r_1 和 r_2 是 $[0, 1]$ 范围的随机向量。

灰狼个体跟踪猎物位置的公式描述如下^[32]

$$\begin{cases} D_{\alpha} = |C_1 \cdot X_{\alpha}(t) - X(t)| \\ D_{\beta} = |C_2 \cdot X_{\beta}(t) - X(t)| \\ D_{\delta} = |C_3 \cdot X_{\delta}(t) - X(t)| \\ X(t+1) = (X_{\alpha} - A_1 \cdot D_{\alpha} + X_{\beta} - A_2 \cdot D_{\beta} + X_{\delta} - A_3 \cdot D_{\delta})/3 \end{cases} \quad (16)$$

式(16)中: D_{α} 、 D_{β} 、 D_{δ} 表示 α 、 β 、 δ 与其他个体间的距离; X_{α} 、 X_{β} 、 X_{δ} 表示 α 、 β 、 δ 的当前位置; A_1 、 A_2 、 A_3 和 C_1 、 C_2 、 C_3 是随机向量; X 代表灰狼种群的位置。式(16)描述了灰狼种群向 α 、 β 、 δ 靠近的过程。

对于多目标优化,通过引入档案群(Archive)机制并制定入选和跳出档案群的规则,可以获得多目标帕累托(Pareto)优化前沿曲线(Pareto front)^[29]。多目标灰狼算法的计算流程见图7。

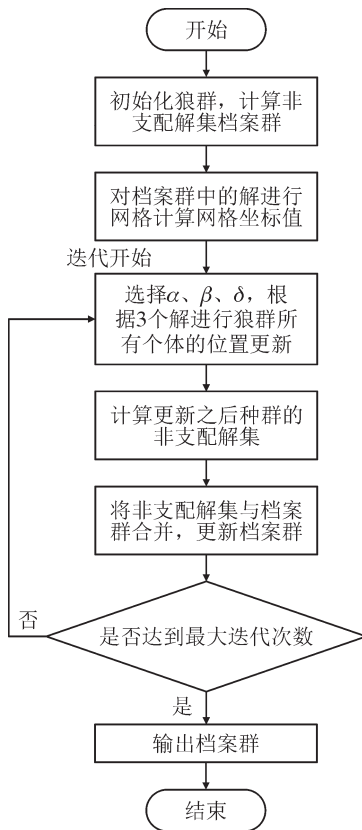


图7 灰狼多目标优化算法流程

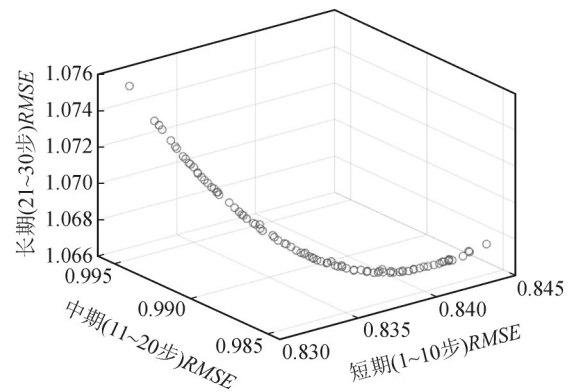
Fig. 7 Procedure of multi-objective grey wolf optimization algorithm

3.2 基于多目标灰狼算法的GRU模型多步预测

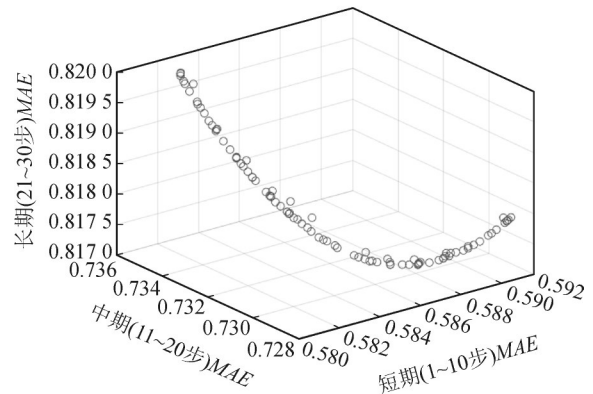
设置GRU多步预测模型的输出时间步为30步,使用多目标灰狼优化算法对学习率(从0.01到0.8)、Dropout层概率(从0.01到0.8)、GRU单元数(从2到64)、输入时间序列步数(从3到30)这两类结构超参数和训练超参数进行优化。定义输出1到10步、11到20步和21到30步预测结果的RMSE和MAE误差为短期、中期和长期误差,将这3种误差作

为多目标函数。

灰狼种群数设置为512,即512组超参数和GRU模型,档案群数据集数量为50,迭代次数设置为500次。将灰狼算法和GRU模型进行混合编程,即灰狼每次迭代后,获得的所有超参数值代入GRU模型中进行训练预测,获得预测RMSE误差和MAE误差,输入到灰狼算法中进行下一步超参数调整,迭代完成后获得的帕累托曲线见图8。对于同一组超参数GRU模型,无法同时做到短期、中期和长期误差都达到最小。相互之间存在拮抗关系。总体来说,短期预测误差最小,中期次之,长期误差最大,这也符合数据预测的一般规律。



(a) 短—中—长期预测RMSE误差帕累托曲线



(b) 短—中—长期预测MAE误差帕累托曲线

图8 GRU模型超参数灰狼多目标优化结果

Fig. 8 Results of multi-objective grey wolf optimization for short-medium-long term GRU prediction

值得注意的是,长期预测误差的RMSE和MAE优化值多于短期预测误差的RMSE和MAE优化值不超过30%,表明GRU模型在短—中—长期预测均表现较为满意的水平。根据多目标灰狼优化算法的结果,可以结合运维和具体情况下的需求,选择不同的超参数进行有所偏重的模型构建与预测。作为示例,文中选取两组帕累托曲线上短期最优和长期最优的超参数,预测效果见图9,各超参数值和误差见表1。图9中实线为实际测量的变压器乙炔数

据,虚线是每次预测的30步乙炔数据。可以看出,超参数的选取对预测结果具有明显的影响。使用长期预测优化超参数时,每次预测的30步曲线更能反映乙炔整体的变化趋势,数据波动较小。使用短期预测优化超参数时,每次预测的30步曲线更贴近初始变化数据,数据整体波动较大。

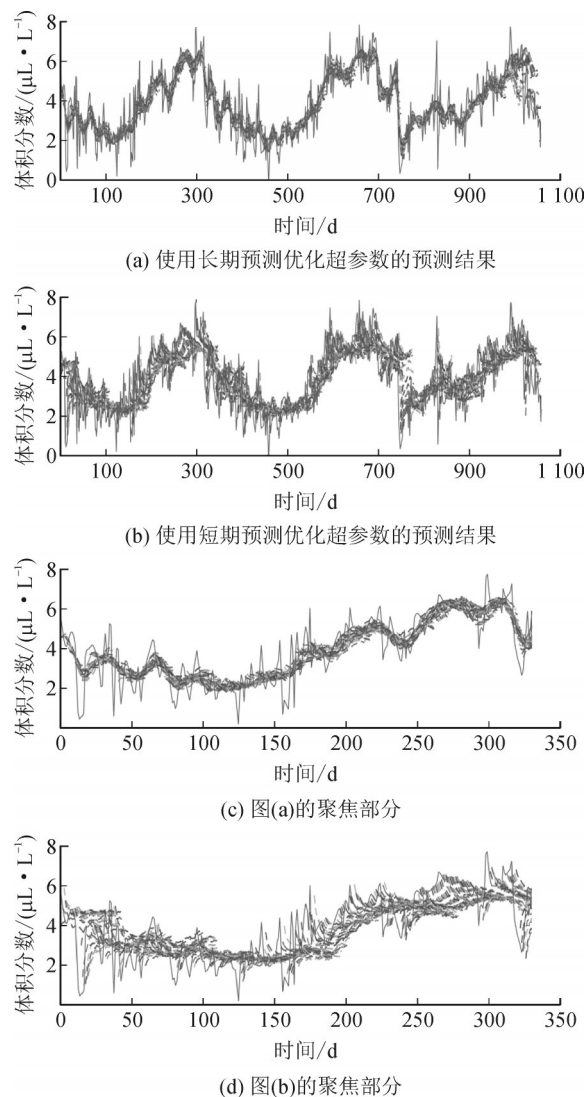


图9 使用不同预测效果倾向超参数的GRU多步预测结果
Fig. 9 Multi-step prediction results of different GRU prediction tendencies

需要指出的是,超参数的选取并没有统一的规律,使用多目标灰狼优化算法可以得出多个帕累托优化解,进而可以根据预测的需求倾向选取不同的超参数对GRU多步预测模型进行建模和训练,以获得有针对性的预测结果。

3.3 GRU预测模型的适用性

为了验证GRU预测模型对不同老化程度变压器的乙炔数据的预测效果,使用上节同样的参数对正常情况下变压器的乙炔监测曲线和逐渐老化出现故障情况下的乙炔曲线进行了30天多步预测,不

设置预测偏好,即学习目标是使得30天的预测误差均值最小。实测曲线和预测曲线见图10。

表1 不同优化超参数的多步预测误差对比

Table 1 Error comparison of multi-step prediction based on different hyperparameters

参数	数值	
	短期	长期
超参数	学习率	0.054
	Dropout 概率	0.26
	GRU 单元数	64
	输入时间步	30
RMSE	短期	0.83
	中期	0.99
	长期	1.07
MAE	短期	0.58
	中期	0.73
	长期	0.81

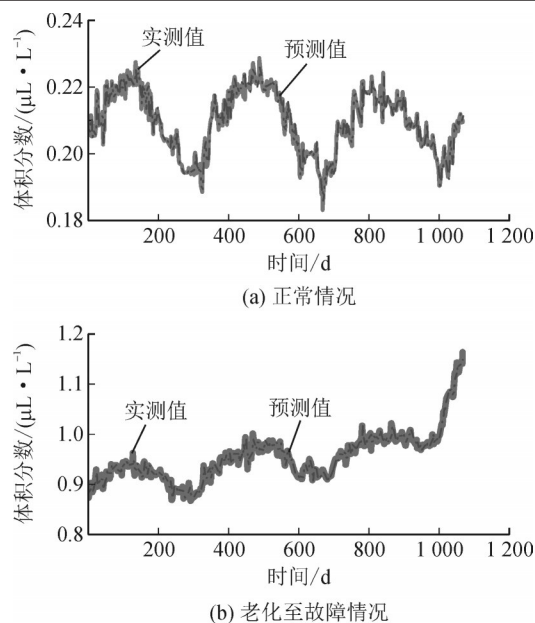


图10 GRU预测模型在不同老化情况下的预测效果
Fig. 10 Prediction results of GRU model for different aging conditions

可以看出GRU多步预测模型对于正常情况和逐步老化情况下乙炔监测曲线均能够较好地预测,正常情况下计算的均方根误差和平均绝对误差分别是0.91和0.72,逐渐老化情况下计算的均方根误差和平均绝对误差分别是0.95和0.78。证明训练好后的算法具有通用性。但需要指出的是,由于数据样本有限,文中采用的数据库均是针对同一厂家生产的变压器油色谱数据进行分析。由于同一厂家生产的变压器采用同种材料和工艺,其老化规律具有近似性,产品可能具有家族共性,因此能够采用

同一模型学习老化规律并进行较准确的预测。同一模型是否能应用其他厂家生产的变压器需要待后续收集其他数据并扩充数据库后来进一步验证。

4 结论

文中采用GRU算法建立了变压器油中溶解乙炔多步预测模型,并且针对超参数之间的耦合关系进行了分析,最后采用多目标灰狼优化算法进行不同目标的GRU模型超参数优化,可以得出如下结论:

1)采用多输出策略构造GRU多步预测模型,能够对变压器乙炔含量进行多步预测,30步预测的RMSE误差不超过1.1,MAE误差不超过0.83。

2)GRU模型的结构超参数和训练超参数如学习率、Dropout层概率、GRU单元数量和输入时间序列长度之间具有耦合关系,各超参数的对输出预测结果的影响规律并不统一且相互影响。总体来说,学习率过高或过低预测效果均不好;GRU单元数量越多越能学习复杂的规律;Dropout层概率越大,模型的输出结果波动越小;输入时间序列长度越长,预测越准确。

3)在30步预测GRU模型中,输出目标划分为短期、中期和长期三类,同一组超参数无法同时达到三目标同时最优。多目标灰狼优化算法能够根据预测目标的不同,优化选择合适的超参数,为人工智能算法超参数的选取提供了参考。

致谢

文中工作受国网河北省电力有限公司科技成本项目(kjcb2020-013)“以人工智能驱动的电力变压器智能诊断与主动预警技术”支持。

参考文献:

- [1] 郑含博. 电力变压器状态评估及故障诊断方法研究[D]. 重庆:重庆大学,2012.
ZHENG Hanbo. Study on condition assessment and fault diagnosis approaches for power transformers[D]. Chongqing:Chongqing University,2012.
- [2] WANG M. Grey-extension method for incipient fault forecasting of oil-immersed power transformer[J]. Electric Power Components and Systems,2004,32(10):959-975.
- [3] 李元,李星辉,孙渭薇,等. 基于多模型级联的油浸式电力变压器故障诊断方法[J]. 智慧电力,2023,51(6):86-92.
LI Yuan, LI Xinghui, SUN Weiwei, et al. Fault diagnosis method of oil-immersed power transformer based on multi-model cascade fusion[J]. Smart Power,2023,51(6):86-92.

- [4] 杜厚贤,刘昊,雷龙武,等. 基于振动信号多特征值的电力变压器故障检测研究[J]. 电工技术学报,2023,38(1):83-94.
DU Houxian, LIU Hao, LEI Longwu, et al. Power transformer fault detection based on multi-eigenvalues of vibration signal[J]. Transactions of China Electrotechnical Society,2023,38(1):83-94.
- [5] 王辉东,姚海燕,郭强,等. 基于多尺度卷积神经网络的变压器故障诊断方法[J]. 电力科学与技术学报,2023,38(4):104-112.
WANG Huidong, YAO Haiyan, GUO Qiang, et al. A transformer fault diagnosis method based on multiscale 1DCNN[J]. Journal of Electric Power Science and Technology,2023,38(4):104-112.
- [6] 刘建锋,刘梦琪,董倩雯,等. 基于改进VMD去噪和优化ELM方法的变压器早期故障诊断[J]. 电力科学与技术学报,2023,38(6):55-66.
LIU Jianfeng, LIU Mengqi, DONG Qianwen, et al. Transformer early fault diagnosis based on improved VMD denoising and optimized ELM method[J]. Journal of Electric Power Science and Technology,2023,38(6):55-66.
- [7] 刘展程,王爽,唐波. 基于布谷鸟搜索算法和DBN模型的变压器故障识别[J]. 电力科学与技术学报,2022,37(2):3-11.
LIU Zhancheng, WANG Shuang, TANG Bo. Transformer fault identification based on the cuckoo search algorithm and DBN model[J]. Journal of Electric Power Science and Technology,2022,37(2):3-11.
- [8] MIROWSKI P, LECUN Y. Statistical machine learning and dissolved gas analysis: A review[J]. IEEE Transactions on Power Delivery,2012,27(4):1791-1799.
- [9] 连鸿松,李文琦,赖永华,等. 变压器绝缘油溶解气体及含气量现场快速检测一体化装置研究[J]. 高压电器,2020,56(3):79-86.
LIAN Hongsong, LI Wenqi, LAI Yonghua, et al. Research on the integration device for on-site rapid detection of dissolved gases and gas content in transformer insulating oil[J]. High Voltage Apparatus,2020,56(3):79-86.
- [10] 邓佳乐,孙辰昊,胡博,等. 基于改进关联规则挖掘的变压器油中溶解气体分析模型[J]. 电力科学与技术学报,2022,37(3):165-172.
DENG Jiale, SUN Chenhao, HU Bo, et al. Evaluation model of the power transformer dissolved gas analysis based on the enhanced association rule mining algorithm[J]. Journal of Electric Power Science and Technology,2022,37(3):165-172.
- [11] 曹旺,万元,唐伟,等. 变压器油中溶解气体在线监测系统设计[J]. 变压器,2023,60(5):31-37.
CAO Wang, WAN Yuan, TANG Wei, et al. Design of on-line monitoring system for dissolved gas in transformer oil

- [J]. Transformer, 2023, 60(5): 31-37.
- [12] 李 佳, 邓 科, 侯玉莲, 等. 基于 GRA-CEEMDAN-BiLSTM 的变压器油中溶解气体浓度预测[J]. 变压器, 2022, 59(6): 42-47.
LI Jia, DENG Ke, HOU Yulian, et al. Prediction of dissolved gas concentration in transformer oil based on GRA-CEEMDAN-BiLSTM[J]. Transformer, 2022, 59(6): 42-47.
- [13] 韩金华, 马云飞, 王立楠, 等. 电力变压器绝缘油色谱氢气超标故障的分类分析[J]. 变压器, 2023, 60(2): 26-31.
HAN Jinhua, MA Yunfei, WANG Linan, et al. Classification and analysis of abnormal Hydrogen gas fault of power transformer oil chromatography[J]. Transformer, 2023, 60(2): 26-31.
- [14] 鲁 永, 赵映宇, 陈 婕, 等. 基于油色谱分析技术的变压器故障分析[J]. 变压器, 2023, 60(6): 64-67.
LU Yong, ZHAO Yingyu, CHEN Jie, et al. Transformer fault analysis based on oil chromatography analysis technology[J]. Transformer, 2023, 60(6): 64-67.
- [15] 沙伟燕, 李秀广, 何宁辉, 等. 基于油色谱时频域信息和残差注意网络的变压器故障诊断方法[J]. 电网与清洁能源, 2022, 38(1): 66-75.
SHA Weiyan, LI Xiuguang, HE Ninghui, et al. Transformer fault diagnosis method based on oil chromatogram time-frequency domain information and residual attention network[J]. Power System and Clean Energy, 2022, 38(1): 66-75.
- [16] 刘云鹏, 许自强, 董王英, 等. 基于经验模态分解和长短期记忆神经网络的变压器油中溶解气体浓度预测方法[J]. 中国电机工程学报, 2019, 39(13): 3998-4007.
LIU Yunpeng, XU Ziqiang, DONG Wangying, et al. Concentration prediction of dissolved gases in transformer oil based on empirical mode decomposition and long short-term memory neural networks[J]. Proceedings of the CSEE, 2019, 39(13): 3998-4007.
- [17] 李 博, 门德月, 严亚勤, 等. 基于数值天气预报的母线负荷预测[J]. 电力系统自动化, 2015, 39(1): 137-140.
LI Bo, MEN Deyue, YAN Yaqin, et al. Bus load forecasting based on numerical weather prediction[J]. Automation of Electric Power Systems, 2015, 39(1): 137-140.
- [18] 郭建鹏, 余颖铃, 温步瀛. 基于 EEMD-ARIMA 模型的地区月负荷量预测[J]. 电力工程技术, 2018, 37(6): 28-32.
GUO Jianpeng, SHE Yingling, WEN Buying. Regional monthly load forecast based on EEMD-ARIMA model[J]. Electric Power Engineering Technology, 2018, 37(6): 28-32.
- [19] 李明伟. 智能预测方法及其在能源领域的应用[M]. 哈尔滨: 哈尔滨工程大学出版社, 2019: 11-23.
LI Mingwei. Intelligent prediction method and its application in energy field[M]. Harbin: Harbin Engineering University Press, 2019: 11-23.
- [20] CHEN Boting, XING Linyuan, ZHAO Lingying, et al. Prediction model of commercial economic index based on BPNN optimization algorithm[C]//2020 International Conference on Computer Engineering and Application (IC-CEA). China: IEEE, 2020: 529-532.
- [21] 邱海枫, 苏 宁, 田松林. 改进支持向量机在电力变压器故障诊断中的应用研究[J]. 电测与仪表, 2022, 59(11): 48-53.
QIU Haifeng, SU Ning, TIAN Songlin. Research on the application of improved support vector machine in power transformer fault diagnosis[J]. Electrical Measurement & Instrumentation, 2022, 59(11): 48-53.
- [22] LIAO Ruijin, ZHENG Hanbo, GRZYBOWSKI S, et al. Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers[J]. Electric Power Systems Research, 2011, 81(12): 2074-2080.
- [23] 咸日常, 范慧芳, 李 飞, 等. 基于改进 GSA-SVM 模型的电力变压器故障诊断[J]. 智慧电力, 2022, 50(6): 50-56.
XIAN Richang, FAN Huifang, LI Fei, et al. Power transformer fault diagnosis based on improved GSA-SVM model[J]. Smart Power, 2022, 50(6): 50-56.
- [24] 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.
YANG Li, WU Yuqian, WANG Junli, et al. Research on recurrent neural network[J]. Journal of Computer Applications, 2018, 38(s2): 1-6.
- [25] 王增平, 赵 兵, 纪维佳, 等. 基于 GRU-NN 模型的短期负荷预测方法[J]. 电力系统自动化, 2019, 43(5): 53-58.
WANG Zengping, ZHAO Bing, JI Weijia, et al. Short-term load forecasting method based on GRU-NN model[J]. Automation of Electric Power Systems, 2019, 43(5): 53-58.
- [26] 牛哲文, 余泽远, 李 波, 等. 基于深度门控循环单元神经网络的短期风功率预测模型[J]. 电力自动化设备, 2018, 38(5): 36-42.
NIU Zhewen, YU Zeyuan, LI Bo, et al. Short-term wind power forecasting model based on deep gated recurrent unit neural network[J]. Electric Power Automation Equipment, 2018, 38(5): 36-42.
- [27] 代杰杰, 宋 辉, 盛戈皞, 等. 采用 LSTM 网络的电力变压器运行状态预测方法研究[J]. 高电压技术, 2018, 44(4): 1099-1106.
DAI Jiejie, SONG Hui, SHENG Gehao, et al. Prediction method for power transformer running state based on LSTM network[J]. High Voltage Engineering, 2018, 44(4): 1099-1106.
- [28] MULDER W D, BETHARD S, MOENS M. A survey on the application of recurrent neural networks to statistical

- [27] WANG Ziyi, CHEN Weigen, SANG Tianyi, et al. Fullerenol-based toxic fluoride gas sensing: A promising way to monitoring Li-ion battery status[J]. Surfaces and Interfaces, 2023(39):102893.
- [28] LU Tian, CHEN Feiwu. Multiwfn: A multifunctional wavefunction analyzer[J]. Journal of Computational Chemistry, 2012, 33(5):580-592.
- [29] 刘宇鹏, 侯文君, 周 渠, 等. 基于密度泛函理论的油中特征气体在钯掺杂 SnP₃单层上吸附及传感性能研究[J]. 中国电机工程学报, 2023, 43(5):2040-2049.
- LIU Yupeng, HOU Wenjun, ZHOU Qu, et al. Investigation on adsorption and sensing performance of characteristic gas in oil on palladium-doped SnP₃ monolayer based on density functional theory[J]. Proceedings of the CSEE, 2023, 43(5):2040-2049.
- [30] SANG T Y, SUN H, HU X, et al. Theoretical exploration of C₄F₇N decompositions on GeSe monolayers for gas sensing based on DFT method[J]. IEEE Sensors Journal, 2022, 22(14):13915-13920.
- 谢醉冰(1988—), 男, 硕士, 高级工程师, 主要研究方向为智能设备在电力系统中的应用、新型电力系统“源—网—荷—储”协同互动等(E-mail: jueyuancailliao001@126.com)。

(上接第 172 页)

- language modeling[J]. Computer Speech & Language, 2015, 30(1):61-98.
- [29] BENGIO Y, BOULANGER-LEWANDOWSKI N, PASCANU R. Advances in optimizing recurrent networks[C]// 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Canada: IEEE, 2013: 8624-8628.
- [30] GIBSON A, PATTERSON J. Deep learning: A practitioner's approach[M]. Boston: O'Reilly Media, 2017.
- [31] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014(69):46-61.
- [32] MIRJALILI S, SAREMI S, MIRJALILI S M, et al. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization[J]. Expert Systems with Applications, 2016(47):106-119.
- 赵 军(1983—), 男, 硕士, 高级工程师, 主要从事变压器类设备的风险隐患管控、健康状态评价、预测性检修、技术监督及重大改造等技术管理(E-mail: 170585815@qq.com)。