DOI:10.19651/j.cnki.emt.2517948

基于 KNN-LASSO-PPC 法的改进 BitCN-LSTM 短期光伏功率预测*

贺宇轩[1] 王 锟[1,2] 曾进辉[1] 刘 颉[1] 周武定[1]
(1.湖南工业大学电气与信息工程学院 株洲 412007;2.株洲高新电业集团有限公司新动力分公司 株洲 412007)

摘 要:针对光伏出力受天气条件随机性和波动性影响的特点,提出一种基于 KNN-LASSO-PCC 法的改进 BitCN-LSTM 神经网络短期光伏功率预测方法。首先,采用 KNN 对数据集进行清洗,再结合 LASSO 与 PCC 进行多层特征筛选;然后,在传统 BitCN-LSTM 方法基础上加入 GRU 与 Elman 神经网络,其中,GRU 解决长时间依赖问题和参数优化问题,Elman 网络增强局部时序建模和记忆能力;最后,在多层特征筛选下选取直角辐射、散角辐射、气温和湿度作为输入变量,选取光伏电站各时段发电功率的预测值作为最终输出,进行为期 1~3 天间隔 15 min 进行一次预测的仿真,所得的最优评估指标平均绝对误差、均方误差以及平均绝对百分比误差分别为 9.976 3%、1.702 9%和 10.626 7%,训练时间和最优测试时间分别为 181.305 1 s 和 0.058 932 s,相较于其他常见的短期光伏预测模型精度更高,速度更快。

关键词:光伏功率预测;多层特征筛选;K 近邻算法;埃尔曼网络;门控循环单元
中图分类号:TP271;TN06 **文献标识码**:A **国家标准学科分类代码**:470.4047

Improved BitCN-LSTM short-term photovoltaic power prediction based on the KNN-LASSO-PPC method

He Yuxuan[1] Wang Kun[1,2] Zeng Jinhui[1] Liu Jie[1] Zhou Wuding[1]
(1. College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412007, China;
2. Zhuzhou High-tech Electric Power Group Co., Ltd., New Power Branch, Zhuzhou 412007, China)

Abstract: The photovoltaic power output is influenced by the randomness and volatility of weather conditions. To address this, an improved BitCN-LSTM neural network-based short-term photovoltaic power forecasting method is proposed using the KNN-LASSO-PCC approach. First, the KNN method is used to clean the dataset. Then, multi-layer feature selection is applied by combining LASSO and PCC. Next, GRU and Elman neural networks are incorporated into the traditional BitCN-LSTM method. Specifically, GRU solves long-term dependency issues and parameter optimization problems, while the Elman network enhances local time-series modeling and memory capacity. Finally, after multi-layer feature selection, global horizontal radiation, diffuse radiation, temperature, and humidity are selected as input variables, and the predicted photovoltaic power output for each time period is selected as the final output. A simulation is conducted for a 1~3 day period with predictions made every 15 minutes. The resulting optimal evaluation metrics are an average absolute error of 9.976 3%, mean squared error of 1.702 9%, and average absolute percentage error of 10.626 7%. The training time and optimal testing time are 181.305 1 s and 0.058 932 s, respectively. Compared to other commonly used short-term photovoltaic forecasting models, the proposed method achieves higher accuracy and faster speed.

Keywords: photovoltaic power prediction; multi-layer feature selection; K-nearest neighbor algorithm; Elman network; gated recurrent unit

0 引 言

车牌自动识别技术在智能交通系统中党的十八大以来我国一直大力发展新能源,推进能源低碳转型。太阳能作为全球分布资源最为广泛的一种可再生能源,因为廉价、清洁等特性受到了越来越多研究人员的关注。但光伏发电受

收稿日期:2025-01-16
* 基金项目:国家自然科学基金(52377185)项目资助

时间、天气等因素影响,昼夜交替、云层覆盖、季节变化等都会导致太阳能的发电量不稳定,这使得大规模接入光伏并网后容易出现电压波动和频率波动的问题[1-2],并且会对电网的保护系统造成影响。精准预测光伏发电功率可以帮助电网进行合理的负荷调度来平滑光伏发电带来的波动,提高电网运行的稳定性和可靠性[3],对我国减少传统化石能源的依赖,推动能源转型和减排目标具有重要意义。

光伏功率预测方法根据预测过程的不同分为直接预测和间接预测[4],直接预测通过使用实测的气象数据和历史光伏数据进行预测,而间接预测则是通过光伏电站位置信息、输出功率曲线还有逆变器系统等技术参数建立模型然后利用未来气象信息进行预测。

当前光伏发电预测领域中的研究热点是基于人工神经网络(artificial neural network,ANN)延展的深度学习预测[5-7]。现在常见的几类人工神经网络中,以卷积神经网络(convolutional neural network,CNN)和长短期记忆网络(long short-term memory,LSTM)为基础的进行改进的神经网络较为流行,文献[8]提出了一种结合卷积神经网络和双向门控循环单元(bidirectional gated recurrent unit,BiGRU)的模型,该方法利用 CNN 提取输入变量与当前负荷间的非线性空间关系,并结合 BiGRU 捕获时间序列的长期依赖关系,取得了较好的预测精度。文中的 CNN 部分适合提取空间特征,但对于时间依赖性较强的数据可能不如其他神经网络模型。文献[9]提出的 VMD-FE-CNN-BiLSTM 方法通过变分模态分解(variational mode decomposition,VMD)分解光伏数据,减少噪声影响;模糊熵(fuzzy entropy,FE)用于子序列特征提取;CNN 提取局部特征,并与双向长短期记忆网络(bidirectional long short-term memory,BiLSTM)网络结合,建立时间特征关系,最终提高了光伏功率预测的精度和稳定性。文中的 VMD 虽然能有效减少噪声,但在处理较长时间序列时,可能导致一些有用信息丢失,而且 BiLSTM 可能在多层次网络中发生过拟合。文献[10]使用的 LSTM-AdaBoost 法通过 LSTM 模型捕获时序信息,结合自适应增强算法(adaptive boosting,AdaBoost)提高模型的准确性。与其他机器学习方法相比,该方法能够在太阳能功率预测中提供更高的准确性。文中的 AdaBoost 部分易受噪声数据影响,进而降低预测稳定性。文献[11]提出的 DPC-LSTM 法通过将密度峰值聚类(density peak clustering,DPC)与 LSTM 的结合,能够有效提高光伏发电功率的预测准确性,特别是在处理光伏发电的随机性和不确定性时,MSE 和 MAE 显著降低。文中用的 DPC 方法在数据密度分布不均匀的情况下,聚类结果可能会受到很大影响,而且在高维数据的情况下该方法产出的结果不够理想。

在短期光伏功率预测任务中,数据的质量直接影响到预测模型的效果。为了提升预测精度和鲁棒性,许多神经网络模型在设计时引入了数据预处理模块,以便有效清洗、转换和提取数据的有用信息。前述模型中,VMD-FE-CNN-BiLSTM 和 DPC-LSTM 这两个模型都包含了特定的数据预处理功能。VMD-FE 模块对原数据进行去噪分解和特征提取,该方法中选取模态数不当会导致信息丢失或过拟合,且在数据出现较多异常值时稳定性不佳;DPC 模块在涉及高维数据时容易出现运算效率低下的问题。文献[12]提出一种以 D-S 证据理论对特征进行综合评分然后用 n 比值法确定阈值进行筛选的数据预处理方法,该方法提高了数据处理的合理性和特征选择的精度,但该方法仍然需要考虑数据噪声的问题,数据噪声带来的不确定性和错误将影响该方法的运行结果。文献[13]提出一种基于辛几何模态分解(symplectic geometry mode decomposition,SGMD)分解时间序列,样本熵(sample entropy,SE)提取特征的数据预处理方法,该方法对数据质量具有依赖性,且在对复杂特征进行提取时易出现过拟合的情况。

处理带有缺失值和异常值这些数据噪声的高维数据集时,针对使用上述文献方法出现的预测精度下降、过拟合等问题,本文提出了一种基于 KNN-LASSO-PCC 和混合神经网络的短期光伏功率预测方法。首先采用 K 近邻插补(K-nearest neighbors imputation,KNN)法对原数据进行数据清洗,找到异常值和缺失值进行替换补全,然后通过 LASSO 回归分析(least absolute shrinkage and selection operator regression,LASSO)与皮尔逊相关系数(Pearson correlation coefficient,PCC)结合的多层特征筛选方法对输入的特征向量进行提取;其次,将优化的数据集输入经由门控循环单元(gated recurrent unit,GRU)模仿埃尔曼网络(elman neural network,Elman)网络优化的 BitCN-LSTM 神经网络,以获得最后的预测结果;最后采用实际光伏发电数据来验证模拟结果的准确性。

1 基于 KNN-LASSO-PPC 法的数据清洗与特征筛选

本研究的实验数据来自陕西省某光伏电厂,时间为 2021 年 1 月 1 日~2021 年 4 月 30 日,每隔 15 min 进行一次采样,共计 11 502 个样本。陕西省的光照资源丰富,太阳辐射年总量高,气候干燥,昼夜温差大,春季天气变化较为频繁,这也使得基于历史数据的光伏发电功率预测需要进行多变量分析,以保证预测结果的准确性。因此本文选取直射辐射,散射辐射,气温,气压,湿度这些相关变量作为输入特征,输出为实际光伏功率。

1.1 K 近邻插补

本文采用的数据清洗方法为 KNN,它作为一种基于相似度的缺失值填充方法[14],其核心思想是利用已有样本的相似性来推测缺失数据,通常用于处理数据集中的缺失值。该方法在单独使用时容易受到噪声数据的影响,且效果会随着数据维度的增加而变差。本文使用 K 近邻插补处理原始数据集中缺失值所用的公式的如式(1)和(2)所示。

$$d(i,j)=\sqrt{\sum_{k=1}^{m}(\boldsymbol{x}_{ik}-\boldsymbol{x}_{jk})^2} \tag{1}$$

其中，$d(i,j)$ 是样本 i 和样本 j 之间的欧氏距离，$\boldsymbol{x}_{ik}$ 和 $\boldsymbol{x}_{jk}$ 分别是样本 i 和样本 j 在第 k 个特征上的值，m 是样本的特征数。

$$\hat{x}=\frac{1}{K}\sum_{i=1}^{K}x_i \tag{2}$$

其中，$\hat{x}$ 为填充缺失值后的结果，K 是选择的邻居数量，x_i 是第 i 个邻居的已知值。

1.2　LASSO 回归分析

LASSO 回归[15]通过引入 L1 正则化来对回归系数施加约束，使得一些特征的系数趋近于零，从而实现特征选择。它的目标是最小化包含 L1 正则化项的损失函数，即在最小化预测误差的同时，惩罚回归系数的绝对值总和。这种正则化方法使得模型更简洁，减少过拟合的风险，并自动选择对输出变量最有影响力的特征。LASSO 回归系数的绝对值越大，表示相关性越强。需要注意的是，当输入特征之间高度相关时，LASSO 可能会随机选择某些特征，而忽略其他重要的特征。计算 LASSO 回归系数涉及到的公式如式(3)所示。

$$\mathcal{L}(\beta)=\frac{1}{2n}\|\boldsymbol{y}-\boldsymbol{X\beta}\|_2^2+\lambda\|\beta\|_1 \tag{3}$$

其中，n 是样本个数，$\frac{1}{2n}\|\boldsymbol{y}-\boldsymbol{X\beta}\|_2^2$ 是损失函数第 1 项，作为普通的最小二乘损失，用来衡量预测值与实际值的差异，$\boldsymbol{y}$ 为目标输出向量，$\boldsymbol{X\beta}$ 为预测值。$\lambda\|\boldsymbol{\beta}\|_1=\lambda\sum_{j=1}^{p}|\boldsymbol{\beta}_j|$ 是损失函数第 2 项，$\|\boldsymbol{\beta}\|_1$ 是 L1 范数，用于引入稀疏性约束，λ 是正则化系数，用于控制稀疏性程度。

1.3　皮尔逊相关系数分析

皮尔逊相关系数 r 作为一种衡量两个变量之间线性关系强度的统计方法[16]，常被用于描述两个变量之间的线性相关性，值的范围在[−1,1]之间。与上方 LASSO 回归分析的判定方法有些类似，皮尔逊相关系数的值越接近 1 或 −1，表示这两个变量之间的关系越强；越接近 0，则表示没有线性关系。需要注意的是，皮尔逊相关系数无法处理多个输入特征之间的共线性问题，可能会高估一些高度相关特征的影响。涉及的计算公式如式(4)所示。

$$r=\frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum(X_i-\bar{X})^2\sum(Y_i-\bar{Y})^2}} \tag{4}$$

其中，X_i 和 Y_i 分别表示 X 和 Y 这两个变量的第 i 个观测值，$\overleftarrow{X}$ 和 $\overleftarrow{Y}$ 分别表示 X 和 Y 这两个变量的样本平均值。

1.4　KNN-LASSO-PCC 法

针对上述 3 种方法单独使用时出现的问题，本文首次提出 KNN-LASSO-PCC 法，该方法分为 KNN 数据清洗、皮尔逊相关系数以及 LASSO 回归系数 3 部分组成，后面两部分为多层特征筛选部分的核心步骤。在 KNN 数据清洗部分，该方法在 1.1 节方法的原步骤上加入了替换异常值和评估填补效果两个改进措施，解决了原 KNN 算法单独使用时对异常值敏感和对高维数据缺失值处理不佳的问题，这样可以进一步提高填补后的数据质量；在多层特征筛选部分，该方法将 1.2 节的皮尔逊相关系数法与 1.3 节的 LASSO 回归系数法进行组合，通过 KNN 考虑多个邻近数据点来减轻多重共线性问题，PCC 识别哪些特征之间具有高度相关性，解决了 LASSO 在单独使用时易受到多个自变量多重共线性影响的问题；通过 KNN 计算最近邻的方式，自适应地捕捉复杂的非线性模式，LASSO 对特征的压缩和筛选，解决了 PCC 在单独使用时无法很好捕捉非线性关系的问题，有助于筛选出相关性更强的特征。具体流程图如图 1 所示。

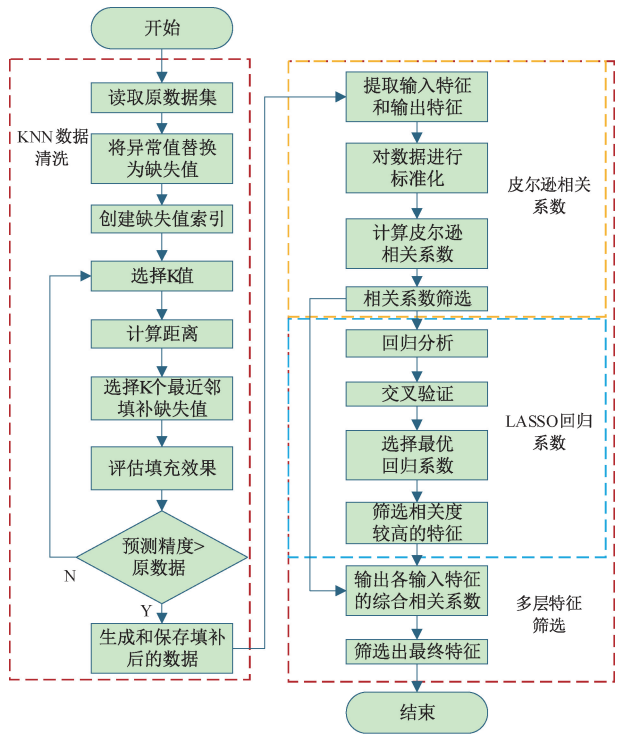

图 1　KNN-LASSO-PCC 法流程图

Fig. 1　Flow chart of KNN-LASSO-PCC method

1)KNN 数据清洗部分：首先识别原数据集中的异常值和缺失值，将异常值替换为缺失值并创建缺失值索引；接着使用式(1)来选择与样本 i 最相似的 K 个样本，并且计算它与所有其他样本之间的距离 $d(i,j)$；随后使用式(2)对 K 个最近邻的已知值去平均来填补缺失值；接下来，对缺失值位置的填补效果进行评估，若评估效果差则会返回重新选择 K 个样本；最后，把生成填补好的数据进行保存，将其输送至皮尔逊相关系数部分。

2)皮尔逊相关系数部分：首先提取数据中的输入和输出特征，对其进行标准化；其次使用式(4)计算每个输入特

征与输出特征之间的皮尔逊相关系数 r;然后使用式(5)设置相关性阈值滤除掉与输出变量关系不大的特征,这步解决了 LASSO 回归分析偶尔忽视重要特征的问题;最后将数据传递到 LASSO 回归系数部分。

$$S=\{f_i \mid P(x_i,y)\mid>t\} \tag{5}$$

3)LASSO 回归系数部分:首先使用式(3)对传入数据进行回归分析,定义 LASSO 的优化目标;其次进行交叉验证评估不同正则化参数下的误差大小,从而选出使交叉验证误差最小的最优值;然后使用最终确定的正则化参数得到最终的回归系数;最后通过回归系数筛选出相关度较高的特征,最后这步解决了皮尔逊系数无法处理多个输入特征之间的共线性问题。

4)输出最终特征部分:使用式(6)综合考虑输入特征与输出特征的线性相关性,结合皮尔逊相关系数和 LASSO 回归系数两部分的分析结果,最后输出如图 2 所示的各输入特征综合相关系数。依照综合相关系数绝对值越大说明与输出变量的相关性越强的原则,筛选出相关性最强的几个特征。

$$S_{\text{final}}=\{f_iL_i\neq 0 \text{ and } \mid P(x_i,y)\mid>t\} \tag{6}$$

其中,S_{final} 表示最终被选中的特征,f_i 表示第 i 个特征,$P(x_i,y)$ 表示第 i 个输入特征和输出特征 y 之间的皮尔逊相关系数。t 表示设置的阈值,文中此处设置的值为 0.1,L_i 表示 LASSO 回归筛选出的系数。

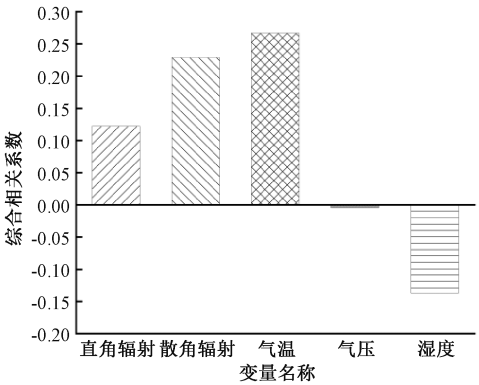

图 2　综合相关系数分析

Fig. 2　Comprehensive correlation coefficient analysis

结合 KNN-LASSO-PCC 法的分析结果,实验数据分析选取前 4 个综合相关系数最高的特征变量(直角辐射、散角辐射、气温和湿度)作为下一步算法模型的输入数据样本变量。这样的特征筛选,一方面减少了特征数量,有助于减少计算的复杂度;另一方面,通过选择相关系数更高的输入特征有助于提高模型的预测精度。

1.5　性能指标

在选用性能指标方面,本文选择训练时间、测试时间加上在回归任务中常用来衡量模型预测性能方面常见的 3 个评估指标:平均绝对误差(mean absolute error,MAE)、平均平方差 (mean squared error,MSE)和平均绝对误差率(mean absolute percentage error,MAPE)[17]。

其中,训练时间和预测时间用于评测模型的计算效率,确保在实际应用中能够快速响应;MAE 和 MSE 主要评测预测误差的大小,帮助评估模型的准确性和鲁棒性;MAPE 用于评测模型相对误差的表现。假定样本总数为 N,真实值为 $y=\{y_1,y_2,y_3,\cdots,y_N\}$,预测值为 $\hat{y}=\{\hat{y}_1,\hat{y}_2,\hat{y}_3,\cdots,\hat{y}_N\}$,3 种指标的计算公式如下:

$$MAE=\frac{1}{N}\sum_{i=1}^{N}\mid y_i-\hat{y}_i\mid \tag{7}$$

$$MSE=\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2 \tag{8}$$

$$MAPE=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-\hat{y}_i}{y_i}\right|\times 100\% \tag{9}$$

1.6　对比分析

选取 BitCN-LSTM 混合神经网络,输入原数据集中部分数据和分别经过 KNN-LASSO-PCC 法,VMD-FE 法,DPC 法,D-S 理论与 n 比值法,SGMD-SE 法进行数据预处理后的数据进行对比测试。

将原数据和优化数据输入神经网络后,得到的结果如表 1 所示。可以看到经由 KNN-LASSO-PCC 法优化过的数据测试结果有着多方面的提升,在与原数据训练时间相差无几的情况下,所用测试时间的明显缩短,且在测试模型预测精度的评测指标上,优化过的数据全方面优于原数据。与其他数据预处理方法相比可以看出,本文方法存在单项稍弱的表现,但在整体评测表现上仍优于其他模型。

表 1　评测指标对比

Table 1　Comparison of evaluation indicators

方法	MSE	MAE	MAPE/%	训练时间/s	测试时间/s
原数据	14.008 4	2.060 8	12.948 5	134.078 0	1.482 20
KNN-LASSO-PCC	11.972 3	1.891 3	11.940 6	139.359 2	0.360 83
VMD-FE	13.062 7	2.039 3	11.893 5	142.569 2	1.283 20
DPC	13.263 8	1.932 4	13.253 6	138.492 0	1.316 93
D-S 理论与 n 比值法	16.318 2	2.190 2	14.062 9	147.823 0	0.672 30
SGMD-SE	11.623 9	1.962 8	15.023 5	150.771 0	0.832 40

2　改进 BitCN-LSTM 神经网络模型

原 BitCN-LSTM 网络由于结合了卷积神经网络和长短期记忆网络的结构[18-19],在训练长序列数据时可能需要花费较大的计算资源,在一些情况下可能会发生梯度消失或者梯度爆炸等问题。虽然该网络可以通过双向卷积提取局部特征,但在长时间局部特征提取方面容易出现信息丢失的问题,且缺乏对时间序列的高效建模。通过引入自注意力机制可以改善模型建模能力,但该方法会消耗大量计

算资源；通过引入集成学习可以减少单个模型的偏差，提高预测的稳定性和准确性，但该方法是多个模型的组合，计算和储存的花费开销大，缺乏单一模型的可解释性。相较于自注意力机制，将 Elman 网络加入到该网络中可以在不损失太多建模能力的情况下，减少计算资源消耗和增强局部时序建模能力。由于 matlab 需要安装相应工具箱才能直接运行 Elman 网络，因此本文首次提出采用 GRU 模仿 Elman 的方法，该方法除了继承上述 Elman 的优点外，还解决了原 Elman 网络容易出现的梯度消失问题，该方法相较于集成学习，所需的计算和储存开销明显更小，可解释性更强。改进的 BitCN-LSTM 神经网络模型如图 3 所示。

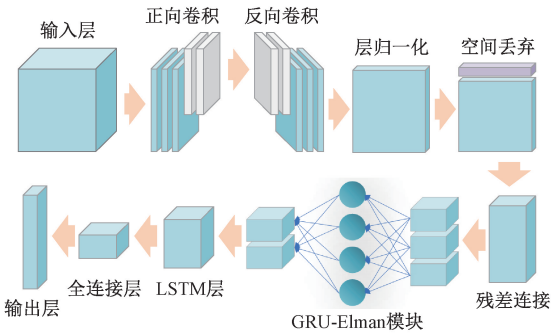

图 3　改进的 BitCN-LSTM 神经网络模型

Fig. 3　Improved BitCN-LSTM neural network modeling

2.1　GRU-Elman 模块

GRU[20]是一种用于处理序列数据的递归神经网络(recurrent neural network，RNN)结构，它主要分为更新门、重置门、候选隐藏状态和最终隐藏状态 4 个部分。GRU 可以通过简化 LSTM 的结构来提高计算效率，同时保留处理长时序依赖的能力，但在处理较为复杂的序列任务时可能遇到困难。

Elman[21]同样是一种递归神经网络，其结构分为输入层，隐层，循环层，输出层四个部分。它具有处理时序数据和“记忆”能力，能够捕捉输入序列中的时序依赖，因其单一的循环结构，很容易在长序列学习中丢失梯度，从而导致长时间依赖问题无法有效捕获。

本文提出的 GRU 模仿 Elman 的 GRU-Elman 模块构建过程如图 4 所示。Elman 网络的核心是其循环结构，即隐层和循环层两部分。

隐层计算并存储当前时刻的隐藏状态，表示网络对输入数据的“记忆”，它通过式(10)将当前输入 x_t 和上一时刻的隐藏状态 h_{t-1} 结合起来计算当前时刻的隐藏状态 h_t。GRU 在这一部分的工作原理是类似的，通过式(11)，使用前一时刻的隐藏状态 h_{t-1} 和当前时刻的候选隐藏状态 $\tilde{h}_t$，然后通过更新门 z_t 加权组合便可得到当前时刻的隐藏状态 h_t。在 GRU-Elman 模块中，GRU 的最终隐藏状态在某种程度上模仿了 Elman 网络中的隐层输出。

$$\boldsymbol{h}_t = \tanh(\boldsymbol{W}_h \cdot \boldsymbol{x}_t + \boldsymbol{U}_h \cdot \boldsymbol{h}_{t-1}) \tag{10}$$

$$\boldsymbol{h}_t = (1 - \boldsymbol{z}_t) \cdot \boldsymbol{h}_{t-1} + \boldsymbol{z}_t \cdot \tilde{\boldsymbol{h}}_t \tag{11}$$

其中，$\boldsymbol{W}_h$ 是一个权重矩阵，负责将当前时刻的输入 $\boldsymbol{x}_t$ 映射到隐藏状态空间；$\boldsymbol{U}_h$ 是另一个权重矩阵，负责将前一时刻的隐藏状态 $\boldsymbol{h}_{t-1}$ 映射到当前隐藏状态空间。

循环层将前一时刻的隐藏状态作为当前时刻的输入之一，实现时序信息的记忆和传递。在 GRU 中，隐藏状态的反馈机制由更新门和重置门控制。通过更新门，GRU 决定了当前时刻的隐藏状态与前一时刻的隐藏状态之间的关系，从而实现了类似 Elman 网络的反馈机制。因此在 GRU-Elman 模块中，GRU 的更新门和重置门共同构成了循环层的功能，它们通过控制历史信息的流动，实现了时间序列信息的传递和记忆。除此之外，GRU 模仿了 Elman 的循环层，通过动态控制前一时刻隐藏状态和当前输入对当前时刻隐藏状态的影响，避免了梯度消失，解决了原 Elman 网络长时间依赖问题无法有效捕获的问题。

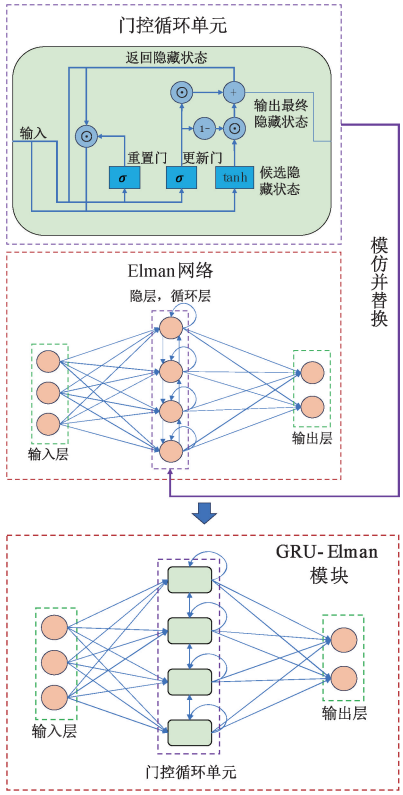

图 4　GRU-Elman 模块组合过程

Fig. 4　GRU-Elman module combination process

2.2　双向时间卷积网络(BitCN)

BitCN 是一种常用于处理时序数据，如时间序列预测、语音识别等任务且结合了双向卷积和时间卷积的神经网络模型[22]。它区别于传统的 TCN 模型的特点在于引入的双向卷积机制。在 BitCN 中，数据不仅通过正向卷积(从过去到未来)进行处理，还通过反向卷积(从未来到过去)进行处理，从而能够同时捕获时间序列中的前后信息。这种双向

处理的方式有助于捕捉数据中的长短期依赖关系,提高对复杂时序数据的建模能力,特别是在某些序列的未来信息对预测有重要影响时,双向卷积提供了更为丰富的上下文信息。BitCN 的网络结构如图 5 所示。

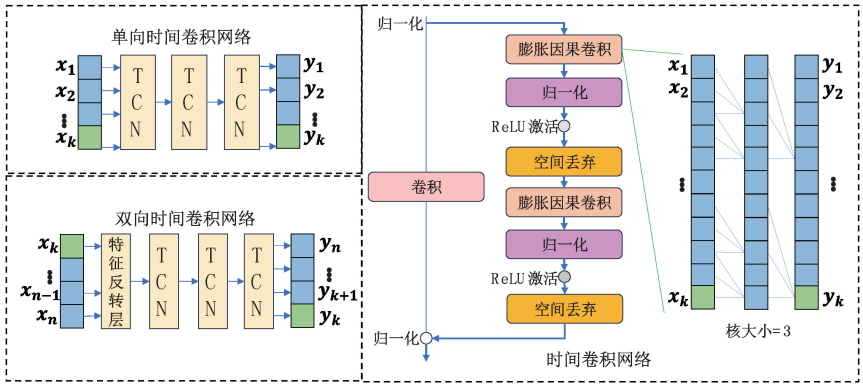

图 5　双向时间卷积网络结构

Fig. 5　Bidirectional temporal convolutional network structure

2.3　长短记忆神经网络(LSTM)

LSTM 网络是一种 RNN 的分支,进行多输入单输出的回归预测特别是在处理时间序列数据时,具有显著优势[23]。它的神经网络结构如图 6 所示。LSTM 网络首先在遗忘门中通过式(12)来决定上一时刻单元状态 $\boldsymbol{C}_{t-1}$ 传到输入门的信息量;然后在输入门中通过式(13)来决定当前时刻的输入 $\boldsymbol{x}_t$ 和上一时刻的输出 $\boldsymbol{h}_{t-1}$ 对单元状态的更新程度;在进入输出门前,LSTM 网络通过式(14)和(15)来进行计算候选记忆和更新单元状态,候选记忆是当前输入 $\boldsymbol{x}_t$ 和上一时刻的输出 $\boldsymbol{h}_{t-1}$ 经过一个 tanh 激活函数计算得到的潜在状态信息 $\tilde{\boldsymbol{C}}_t$,单元状态 $\boldsymbol{C}_t$ 作为 LSTM 网络的核心,通过遗忘门处的输出 $\boldsymbol{f}_t$ 和保留的上一时刻单元状态 $\boldsymbol{C}_{t-1}$ 与输入门处的输出 $\boldsymbol{i}_t$ 和潜在状态信息 $\tilde{\boldsymbol{C}}_t$ 来进行更新;最后在输出门中先通过式(16)计算出 $\boldsymbol{o}_t$ 来决定当前单元状态 $\boldsymbol{C}_t$ 中的信息需要多少输出,再通过式(17)输出门的输出 $\boldsymbol{h}_t$ 与当前单元状态经过 tanh 激活后的值的乘积得到最终输出 $\boldsymbol{h}_t$。涉及的全部计算公式为:

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W}_f \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_f) \tag{12}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_i) \tag{13}$$

$$\tilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_C \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_C) \tag{14}$$

$$\boldsymbol{C}_t = \boldsymbol{f}_t \cdot \boldsymbol{C}_{t-1} + \boldsymbol{i}_t \cdot \tilde{\boldsymbol{C}}_t \tag{15}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \boldsymbol{b}_o) \tag{16}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \cdot \tanh(\boldsymbol{C}_t) \tag{17}$$

其中,σ 是 sigmoid 激活函数,tanh 是双曲正切激活函数,$\boldsymbol{W}$ 是权重矩阵,$\boldsymbol{b}$ 是偏置向量。

3　仿真分析

为验算该算法模型的可行性,本文将 80%的数据作为训练集,剩余的数据作为测试集,这样的划分比例能够为神经网络提供足够的样本来学习数据的特征和规律,避免训练过程中发生过拟合或欠拟合的情况。仿真选择在电脑上进行,CPU 配置为 Intel Core i5-13450HX,主频率和最大频率分别为 2.40 和 4.60 Ghz,显卡为 RTX4050,系统为 win11,程序代码运行环境选择 MATLAB R2023b。

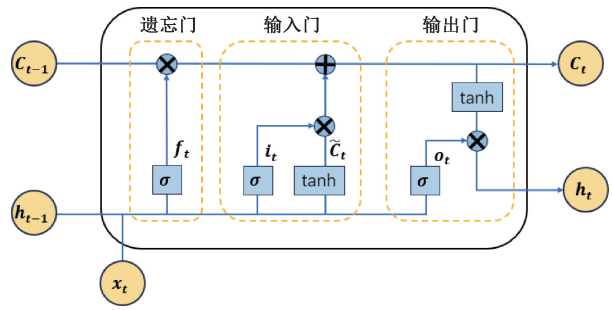

图 6　LSTM 神经网络结构

Fig. 6　LSTM neural network structure

仿真对本文提出的 GRU-Elman-BitCN-LSTM 算法模型进行测试,并且与 CNN-BiGRU-TPA,LSTM-AdaBoost,BitCN-LSTM 3 种不包含数据预处理的新型神经网络模型,以及 VMD-FE-CNN-BiLSTM,DPC-LSTM 两种包含数据预处理的新型神经网络模型,共 5 个组合神经网络模型的预测结果进行比较。本实验在测试模型预测性能指标(平均绝对误差、平均平方差、平均绝对误差率和最大绝对误差)的同时,也对模型测试过程中训练和测试时间进行对比。最大绝对误差代表模型预测值和实际值之间的最大偏差,能够有效体现哪些模型的预测误差最大。对于光伏预测这种时序数据,某些时刻的误差可能远大于其他时刻,选取最大误差可以突出模型的最差预测,帮助展示模型之间的差异性。最大绝对误差涉及的公式如式(18)所示。

$$\mathrm{MAE}_{\max} = \max(|\, y_{\mathrm{true},i} - y_{\mathrm{pred},i} \,|) \tag{18}$$

其中,$y_{\mathrm{true},i}$ 是第 i 个真实值,$y_{\mathrm{pred},i}$ 是第 i 个预测值,$|\, y_{\mathrm{true},i} - y_{\mathrm{pred},i} \,|$ 为第 i 个数据点的绝对误差,max 表示所有误差中的最大值。

3.1　多种模型误差性能对比结果

仿真对 CNN-BiGRU-TPA，VMD-FE-CNN-BiLSTM，LSTM-AdaBoost，DPC-LSTM，BitCN-LSTM 还有本文提出的 GRU-Elman-BitCN-LSTM 6 个混合神经网络模型使用经过 KNN-LASSO-PCC 法预处理的数据进行性能测试。实验测试结果如图 7～9 所示，分别为 6 种模型截取 1 天、2 天、3 天光伏发电功率的预测值和实际值关系图，选取每 15 min 为一个间隔的发电输出功率值，其中棕色线代表实际发电输出功率，其他线代表预测发电输出功率，从预测值线与真实值线的拟合程度来看，经由 GRU 和 Elman 优化过的 BitCN-LSTM 混合神经网络模型实验效果比其他带有数据预处理的两种模型和不带数据预处理的 3 种模型实验效果更为出色，得到的预测值与实际值更为接近；从表 2 可以看出，在比较最大绝对误差方面，可以明显看出本文模型在最差情况下的预测能力远优于其他模型，误差也更小。

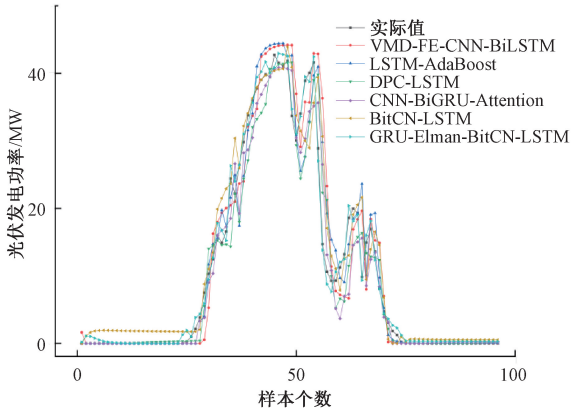

图 7　1 天的光伏功率预测图

Fig. 7　PV power prediction chart for 1 day

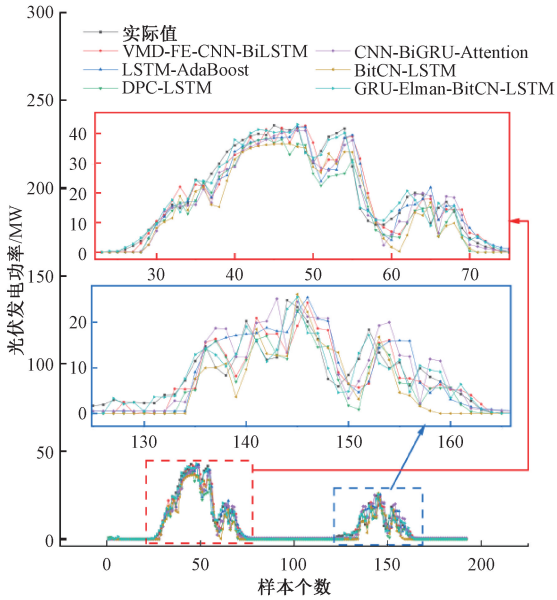

图 8　2 天的光伏功率预测图

Fig. 8　PV power prediction chart for 2 days

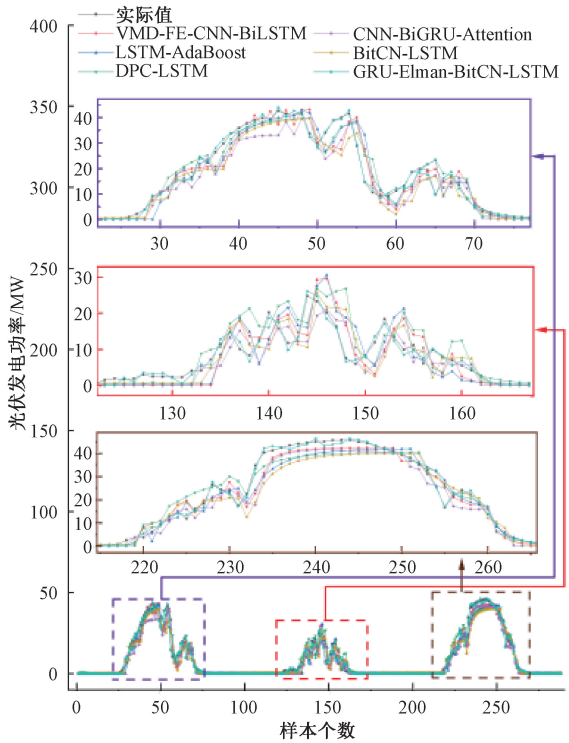

图 9　3 天的光伏功率预测图

Fig. 9　PV power prediction chart for 3 days

表 2　模型最大绝对误差

Table 2　Maximum absolute error of the model

模型	1 天	2 天	3 天
VMD-FE-CNN-BiLSTM	21.603 0	16.758 9	12.581 9
LSTM-AdaBoost	15.101 9	15.074 4	18.604 0
DPC-LSTM	11.278 6	15.104 7	19.247 5
CNN-BiGRU-TPA	11.792 5	13.289 2	17.027 0
BitCN-LSTM	15.949 3	14.539 2	15.841 3
GRU-Elman-BitCN-LSTM	6.882 4	7.539 2	7.722 3

进一步通过具体的预测误差指标分析，如表 3～5 所示，在 MSE，MAE，MAPE 这 3 个指标下，GRU-Elman-BitCN-LSTM 模型的性能指标最好，CNN-BiGRU-TPA 模型次之，而 VMD-FE-CNN-BiLSTM，DPC-LSTM 两种包含数据预处理的模型与本文方法仍有一定的差距。特别在 MSE 指标下，进行 3 天光伏功率预测的 GRU-Elman-BitCN-LSTM 比其他混合神经网络模型最好的结果少了 1.046 8 的误差。

3.2　多种模型运行时间对比结果

训练期间，各神经网络模型的所用时间如表 6 所示。6 个混合神经网络模型 VMD-FE-CNN-BiLSTM、LSTM-AdaBoost、DPC-LSTM、CNN-BiGRU-TPA、BitCN-LSTM、GRU-Elman-BitCN-LSTM 所花费的时间分别为 207.308 1、195.232 1、178.037 4、189.416 7、239.074 2 和 181.305 1 s。

可以看到,在训练时间上本文提出的模型优于除 DPC-LSTM 之外的其余 4 种模型,且与 DPC-LSTM 这个带有数据预处理的神经网络模型的差距不大。

表 3 1 天的模型预测误差指标

Table 3 1-day model prediction error indicator

模型	MAE	MSE	MAPE/%
VMD-FE-CNN-BiLSTM	18.225 4	2.286 8	14.465 3
LSTM-AdaBoost	14.598 3	2.025 7	14.496 3
DPC-LSTM	13.303 6	2.031 3	14.729 2
CNN-BiGRU-TPA	10.687 3	1.803 1	12.034 1
BitCN-LSTM	13.520 6	2.335 7	11.573 0
GRU-Elman-BitCN-LSTM	9.976 3	1.702 9	10.720 2

表 4 2 天的模型预测误差指标

Table 4 2-day model prediction error indicator

模型	MAE	MSE	MAPE/%
VMD-FE-CNN-BiLSTM	11.926 2	1.808 1	17.846 1
LSTM-AdaBoost	12.469 9	2.024 6	17.414 2
DPC-LSTM	13.559 5	1.993 1	20.988 0
CNN-BiGRU-TPA	12.493 9	2.086 7	17.767 4
BitCN-LSTM	15.349 4	2.235 1	21.614 9
GRU-Elman-BitCN-LSTM	11.393 9	1.824 9	15.868 4

表 5 3 天的模型预测误差指标

Table 5 3-day model prediction error indicator

模型	MAE	MSE	MAPE/%
VMD-FE-CNN-BiLSTM	11.077 0	1.923 5	13.535 9
LSTM-AdaBoost	12.953 6	2.172 4	14.268 6
DPC-LSTM	12.920 9	2.215 6	13.771 7
CNN-BiGRU-TPA	11.054 0	1.918 2	12.894 5
BitCN-LSTM	14.409 8	2.211 9	17.094 8
GRU-Elman-BitCN-LSTM	10.007 2	1.735 2	10.626 7

表 6 模型训练时间

Table 6 Model training time

模型	训练时间/s
VMD-FE-CNN-BiLSTM	207.308 1
LSTM-AdaBoost	195.232 1
DPC-LSTM	178.037 4
CNN-BiGRU-TPA	189.416 7
BitCN-LSTM	239.074 2
GRU-Elman-BitCN-LSTM	183.305 1

在测试阶段,这 6 个混合神经网络模型花费的时间如表 7 所示。可以看到在 1 天和 2 天的测试时间中,本文提出的模型优于其余两种带有数据预处理的模型和 3 种不带数据预处理的模型。只在 3 天的测试事件中,该模型稍逊于带有数据预处理的 DPS-LSTM 模型,不过所用时间非常接近。综合以上结果,经由 GRU 和 Elman 优化过的 BitCN-LSTM 混合算法模型更具有优势。

表 7 模型测试时间

Table 7 Model testing time s

模型	1 天	2 天	3 天
VMD-FE-CNN-BiLSTM	0.071 289	0.095 884	0.190 250
LSTM-AdaBoost	0.069 062	0.223 270	0.092 757
DPC-LSTM	0.097 531	0.080 513	0.072 709
CNN-BiGRU-TPA	0.067 425	0.088 346	0.094 893
BitCN-LSTM	0.091 205	0.070 309	0.122 100
GRU-Elman-BitCN-LSTM	0.058 932	0.067 602	0.073 074

4 结 论

为了让光伏电站和电网用户之间的供需关系趋于平衡,提升光伏发电功率预测的准确率,本文提出了基于使用 K 近邻插补、LASSO 与皮尔逊系数组合的多层特征筛选进行数据预处理,GRU 模仿 Elman 网络优化 BitCN-LSTM 混合神经网络模型的方法。通过数据处理和算法模型分析证明可以得到以下结论:

使用 K 近邻插补的数据清洗方法和 LASSO、皮尔逊系数组合的 KNN-LASSO-PCC 法进行数据预处理,有效解决了数据集中出现异常值和缺失值的问题,并且在此过程中去除了回归系数较低的数据,降低了数据维度,使得运算成本减少,为后续提高模型精度奠定了基础;经 GRU 模仿 Elman 网络的 GRU-Elman 模块优化后,得到的 GRU-Elman-BitCN-LSTM 混合神经网络模型结合了 GRU 在处理长时依赖方面的优势以及 Elman 在捕捉时序和抽象信息方面的优势,能够更好地处理复杂的时序数据任务,在 MSE 指标下,它比原 BitCN-LSTM 神经网络模型最高减少了 4.402 6 的误差,比表现最优的神经网络模型最高少了 1.046 8 的误差;从训练时间和测试时间这方面看,GRU-Elman-BitCN-LSTM 所花费的时间要显著低于原 BitCN-LSTM 神经网络模型,并且在绝大多数情况下优于其余 5 个用于对比的神经网络模型。

综上所述,本文提出的方法可以在预测光伏发电功率对光伏并网和电力系统的稳定发展方面起到一定的作用,并为电网调度部门提供一定的参考价值,在与其他神经网络模型进行综合预测性能,运行时间以及测试时间比较后,也验证了本文方法的有效性与准确性。

参考文献

[1] ALAM M S, CHOWDHURY T A, DHAR A, et al. Solar and wind energy integrated system frequency control: A critical review on recent developments[J]. Energies, 2023, 16(2): 812.

[2] IMPRAM S, NESE S V, ORAL B. Challenges of renewable energy penetration on power system flexibility: A survey[J]. Energy Strategy Reviews, 2020, 31: 100539.

[3] AHMED A, KHALID M. A review on the selected applications of forecasting models in renewable power systems [J]. Renewable and Sustainable Energy Reviews, 2019, 100: 9-21.

[4] 吴硕. 光伏发电系统功率预测方法研究综述[J]. 热能动力工程, 2021, 36(8): 1-7.
WU SH. Review of power forecasting methods of photovoltaic power generation system[J]. Journal of Engineering for Thermal Energy and Power, 2021, 36(8): 1-7.

[5] 唐剑飞. 光伏发电预测方法的分类与比较分析[J]. 船电技术,2024,44(9):69-72.
TANG J F. Classification and comparative analysis of forecasting methods for photovoltaic power generation[J]. Marine Electric & Electronic Engineering, 2024, 44(9): 69-72.

[6] 舒胜, 谢应明, 杨文宇, 等. 光伏发电预测方法研究进展[J]. 热能动力工程, 2020, 35(11):1-11.
SHU SH, XIE Y M, YANG W Y, et al. A review of photovoltaic power generation forecasting methods [J]. Journal of Engineering for Thermal Energy and Power, 2020, 35(11): 1-11.

[7] RAIAGUKGUK R A, RAMADHAN R A A, LEE H J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power[J]. Energies, 2020, 13(24): 6623.

[8] 黄宇, 顾智勇, 李永玲, 等. 基于时间模式注意力机制的 CNN-BiGRU 短期负荷预测[J]. 华北电力大学学报(自然科学版),2023,50(6):11-20.
HUANG Y, GU ZH Y, LI Y L, et al. Short-term power load forecasting based on temporal pattern attention mechanism of CNN-BiGRU[J]. Journal of North China Electric Power University (Natural Science Edition), 2023, 50(6): 11-20.

[9] 姜建国, 杨效岩, 毕洪波. 基于 VMD-FE-CNN-BiLSTM 的短期光伏发电功率预测[J]. 太阳能学报, 2024,45(7):462-473.
JIANG J G, YANG X Y, BI H B. Photovoltaic power forecasting method based on VMD-FE-CNN-BiLSTM[J]. Acta Energiae Solaris Sinica, 2024, 45(7): 462-473.

[10] KYEREMEH F, ZHI F, YI Y, et al. Solar PV power forecasting with a hybrid LSTM-AdaBoost ensemble[C]. 2022 IEEE/IET International Utility Conference and Exposition(IUCE). IEEE, 2022: 1-7.

[11] 彭曙蓉, 陈慧霞, 孙万通, 等. 基于改进 LSTM 的光伏发电功率预测方法研究[J]. 太阳能学报, 2024, 45(11):296-302.
PENG SH R, CHEN H X, SUN W T, et al. Research on photovoitaic power predction method based on improved LSTM[J]. Acta Energiae Solaris Sinica, 2024, 45(11): 296-302.

[12] 杨家豪, 张莲, 杨玉洁, 等. 混合特征筛选与分时 Stacking 的无地表辐照度光伏出力预测[J]. 重庆理工大学学报(自然科学),2024,38(10):253-260.
YANG J H, ZHANG L, YANG Y J, et al. Synthetic feature selection and time-division stacking for photovoltaic output forecast without surface solar radiation information [J]. Journal of Chongqing University of Technology (Natural Science), 2024, 38(10): 253-260.

[13] 王文川, 顾淼. 基于 SGMD-SE-AVOA-LSTM 耦合模型的月径流预测[J]. 应用基础与工程科学学报,2024, 32(6):1755-1771.
WANG W CH, GU M. A hybrid monthly runoff prediction model based on SGMD-SE-AVOA-LSTM[J]. Journal of Basic Science and Engineering, 2024, 32(6): 1755-1771.

[14] PUJIANTO U, WIBAWA A P, AKBAR M I. K-nearest neighbor(k-NN) based missing data imputation[C]. 2019 5th International Conference on Science in Information Technology(ICSITech). IEEE, 2019: 83-88.

[15] MOTAMEDI F, PÉREZ-SÁNCHEZ H, MEHRIDEHNAVI A, et al. Accelerating big data analysis through LASSO-random forest algorithm in QSAR studies [J]. Bioinformatics, 2022, 38(2): 469-475.

[16] JEBLI I, BELOUADHA F Z, KABBAJ M I, et al. Prediction of solar energy guided by pearson correlation using machine learning[J]. Energy, 2021, 224: 120109.

[17] TATACHAR A V. Comparative assessment of regression models based on model evaluation metrics[J]. International Journal of Innovative Technology and Exploring Engineering, 2021, 8(9): 853-860.

[18] 王晨阳, 段倩倩, 周凯, 等. 基于遗传算法优化卷积长短记忆混合神经网络模型的光伏发电功率预测[J]. 物理学报,2020,69(10):149-155.
WANG CH Y, DUAN Q Q, ZHOU K, et al. A hybrid model for photovoltaic power prediction of both convolutional and long short-term memory neural networks optimized by genetic algorithm [J]. Acta Physica Sinica, 2020, 69(10): 149-155.

[19] CHEN J, LYU T, CAI S, et al. A novel detection model

for abnormal network traffic based on bidirectional temporal convolutional network[J]. Information and Software Technology, 2023, 157: 107166.

[20] 钱来, 王伟. 一种基于 C-GRU 飞行轨迹预测方法[J].电子测量技术,2022,45(10):87-92.
QIAN L, WANG W. A C-GRU based flight trajectory prediction method [J]. Electronic Measurement Technology, 2022, 45(10): 87-92.

[21] 杨鸿雁, 邢超, 陈仕龙, 等. 基于经验模态分解与 Elman 神经网络的永富直流换相失败故障诊断方法[J].电子测量技术,2020,43(1):169-175.
YANG H Y, XING CH, CHEN SH L, et al. Yongfu DC commutation failure fault diagnosis based on the EMD-Elman neural network [J]. Electronic Measurement Technology, 2020, 43(1): 169-175.

[22] 李大舟, 于沛, 高巍, 等. 基于社交媒体文本信息的金融时序预测[J].计算机工程与设计,2021,42(8):2224-2231.
LI D ZH, YU P, GAO W, et al. Financial time seriers prediction based on social media text information[J]. Computer Engineering and Design, 2021, 42(8): 2224-2231.

[23] 陈蕨峰, 王贺, 李岩, 等. 组合两步分解和 ARIMA-LSTM 的短期风速预测研究[J].太阳能学报,2024,45(2):164-171.
CHEN H F, WANG H, LI Y, et al. Short-term speed prediction by combining two-step decomposition and ARIMA-LSTM[J]. Acta Energiae Solaris Sinica, 2024,45(2):164-171.

作者简介

贺宇轩(通信作者),硕士研究生,主要研究方向为光伏功率预测。

E-mail:imaxhyx@163.com

王锟,硕士生导师,高级工程师,主要研究方向为可再生能源开发利用及能源转换和存储技术应用。

E-mail:wk3622132@163.com