

基于 SSA-BiLSTM-AM 的短期风电功率预测^{*}

章志晃 徐启峰 林 穿

(福州大学电气工程与自动化学院 福州 350108)

摘 要: 风电功率的准确预测可以有效地减少并网波动。现有的风电功率预测模型存在输入特征过多、超参数选择难、时序过长易丢失重要信息等问题。为此,提出了一种麻雀搜索算法(SSA)优化双向长短时记忆(BiLSTM)加注意力机制(AM)的短期风电功率融合预测模型。首先,SSA对BiLSTM神经网络的节点数、学习率和训练次数等超参数进行寻优,确认最佳参数;然后,引入AM对BiLSTM的输入特征分配不同权重,强化关键特征;最后,应用所提模型对新疆210 MW风电场的风电功率进行预测,并与其他模型的预测结果对比。结果表明,SSA-BiLSTM-AM预测模型的均方根误差(RMSE)为5.411 4、平均绝对误差(MAE)为3.674 9,显著优于其他模型的预测精度,证明了SSA优化算法和AM能够有效提高风电机组的短期功率预测精度。

关键词: 风电功率预测;麻雀搜索算法;双向长短时记忆;注意力机制

中图分类号: TM614 **文献标识码:** A **国家标准学科分类代码:** 470.40

Short-term wind power prediction based on SSA-BiLSTM-AM

Zhang Zhihuang Xu Qifeng Lin Chuan

(College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China)

Abstract: Accurate prediction of wind power can effectively reduce grid connection fluctuation. The existing wind power prediction models have some problems, such as too many input features, difficult to select super parameters, and easy to lose important information in long time series. To this end, a short-term wind power fusion prediction model was proposed based on the sparrow search algorithm (SSA) to optimize the bidirectional long short-term memory (BiLSTM) plus attention mechanism (AM). First, SSA optimizes the super parameters such as the number of nodes, learning rate and training times of BiLSTM neural network to confirm the best parameters. Then, AM was introduced to assign different weights to input features of BiLSTM to strengthen key features. Finally, the proposed model is used to predict the wind power of 210 MW wind farm in Xinjiang Province, and the results are compared with those of other models. The results show that the root mean square error (RMSE) of the SSA-BiLSTM-AM prediction model is 5.411 4, and the mean absolute error (MAE) is 3.674 9. The prediction accuracy of SSA optimization algorithm and AM is significantly better than that of other models, which proves that SSA optimization algorithm and AM can effectively improve the short-term power prediction accuracy of wind turbines.

Keywords: wind power prediction; sparrow search algorithm; bidirectional long short-term memory; attention mechanism

0 引 言

近年来,可再生能源发电发展迅速,风能因其零污染、资源分布广的优势成为可再生能源发电的主要形式。但

是,风力发电具有随机性和波动性,风电的输出功率会出现一些波动,在大规模风电并网时,将给电力系统的规划和稳定运行带来巨大挑战^[1]。因此,准的风电出力预测有利于电网调度和安全运行^[2]。

收稿日期:2022-11-14

^{*} 基金项目:国家自然科学基金(51977038)项目资助

国内外很多学者在风电功率预测方面不断进行研究,取得了较为丰富的成果。国外针对风电功率预测的研究起步较早,已经具有比较成熟的风电功率预测方法和系统,且近年来国外开发了一系列风电功率预测系统,如2001年德国太阳能研究所推出了一套风电功率管理系统,该系统一直在德国的四大输电系统运行^[3]。相比于国外,国内的相关研究和软件的开发还比滞后。

风电功率预测模型主要可以分为物理方法、统计方法以及人工智能方法3类。基于严格的数学建模的物理方法最为复杂;基于统计学方法的建模相对简单,平均预测精度也优于物理预测模型^[4];人工智能方法能够更好的反应时间序列的变化趋势,成为了预测方面的更优选择。人工智能方法通过历史数据对模型进行训练,再将训练好的模型进行预测,但单一模型预测精度较低,所以风电功率预测一般采用组合预测模型。文献[5]提出了一种改进的长短时记忆(long short-term memory, LSTM)的短期风电功率混合预测模型,相比其他模型具有更好的鲁棒性和准确性。文献[6]提出一种经验模态分解(empirical mode decomposition, EMD)与径向基函数(radial basis function, RBF)神经网络的短期风功率预测模型。EMD能自主获取输入数据特性,具有自适应性和多分辨率,缺点是会产生模态混叠现象,使预测精度降低^[7]。文献[7]建立了基于变分模态分解(variational mode decomposition, VMD)预测模型,VMD可以有效解决风电功率数据存在随机性和非线性等问题。文献[8]使用寻优能力强、收敛速度快的麻雀搜索算法(sparrow search algorithm, SSA)对LSTM进行参数优化,构建了SSA-LSTM模型,提高了预测精度。LSTM虽然可以学习时间序列数据的长时间相关性,但单向LSTM未考虑到时间序列较长且存在双向信息流。同时,上述文献未考虑到不同输入特征对输出结果的影响程度,没有对输入特征赋予不同的权重,减少复杂特征过多对预测结果的影响。

综上所述,本文提出一种基于注意力机制(attention mechanism, AM)的麻雀搜索算法优化双向长短时记忆(bidirectional long short-term memory, BiLSTM)的短期风电功率预测模型。通过麻雀算法对BiLSTM神经网络的学习率和迭代次数等参数进行优化,并引入注意力机制突出关键特征。最后采用新疆风电场的风电功率数据进行预测,结果表明,通过SSA对网络参数进行优化选择并引入AM的模型可以显著提高预测精度。

1 算法原理

1.1 麻雀搜索算法

麻雀搜索算法^[9-12]是2020年提出的优化算法,该算法寻优能力强、收敛速度快。在SSA算法中,初始化麻雀种群的位置矩阵为:

$$X = \begin{bmatrix} X_1^1 & X_1^2 & \cdots & X_1^d \\ X_2^1 & X_2^2 & \cdots & X_2^d \\ \vdots & \vdots & \ddots & \vdots \\ X_n^1 & X_n^2 & \cdots & X_n^d \end{bmatrix} \quad (1)$$

式中: n 表示麻雀种群的数量; d 表示要优化的变量的维度。

在SSA中,发现者具有良好的适应度值,可以优先找到食物,在搜索过程中的位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \exp\left(-\frac{i}{\alpha I_{\max}}\right), & R_2 < S_T \\ X_{i,j}^t + QL, & R_2 \geq S_T \end{cases} \quad (2)$$

式中: j 为当前维度, $j \in [1, 2, \dots, d]$; $X_{i,j}^t$ 表示第 t 次迭代时,第 i 只麻雀的第 j 维度的位置; $I_{\max}$ 是最大迭代次数; $\alpha \in [0, 1]$ 是一个随机数; Q 属于服从正态分布的随机数; L 为 $1 \times d$ 的单位矩阵; R_2 为预警值; S_T 为安全值。当 $R_2 < S_T$ 时,意味着周围没有天敌,发现者可以进行广泛搜索。当 $R_2 \geq S_T$ 时,意味着周围存在捕食者,且一部分麻雀已经发现了捕食者,所有麻雀需要立即飞往其他安全区域。

在麻雀算法中,追随者适应度值较差,当追随者察觉发现者找到食物时,会前去争夺。如果争夺失败,则继续到其他地方搜寻食物。追随者位置更新过程如下:

$$X_{i,j}^{t+1} = \begin{cases} Q \exp\left(\frac{X_{\text{worse}}^t - X_{i,j}^t}{i^2}\right), & i > n/2 \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| A^+ L, & i \leq n/2 \end{cases} \quad (3)$$

式中: X_p^{t+1} 是第 $t+1$ 次迭代时发现者所在的最优位置; X_{worse}^t 表示第 t 次迭代时,当前种群的最差位置; A 为一个 $1 \times d$ 的矩阵,矩阵中每个元素的值随机分配为1或-1, $A^+ = A^T(AA^T)^{-1}$; n 为麻雀的总数。当 $i > n/2$ 时,代表适应度较低的第 i 个追随者没能获取到食物,需要前往其他区域寻找食物;当 $i \leq n/2$ 时,第 i 个追随者在发现者寻找食物的中心行动,在中心位置附近寻找食物。

在算法中,一般侦查者的数量占总种群的10%~20%,更新位置的公式如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta |X_{i,j}^t - X_{\text{best}}^t|, & f_i > f_g \\ X_{i,j}^t + K \left(\frac{|X_{i,j}^t - X_{\text{worse}}^t|}{(f_i - f_w) + \epsilon} \right), & f_i = f_g \end{cases} \quad (4)$$

式中: X_{best}^t 是 t 次迭代时全局最优位置; β 是步长控制参数,其值是服从均值为0,方差为1的正态分布的随机数; $K \in [-1, 1]$; ϵ 是避免分母为0的最小常数; f_i 为当前麻雀的适应度值; f_g 和 f_w 为全局最佳适应度值和最差适应度值。当 $f_i > f_g$ 时,说明麻雀处于种群的边缘位置,可能遭遇捕食者;当 $f_i = f_g$ 时,说明处于种群中心的麻雀意识到了危险,需要向其他麻雀靠拢。

综上,麻雀算法的主要步骤如下。

步骤1)设置最大迭代次数、麻雀种群数目、发现者数

量、侦查者数量以及预警值。

步骤2)计算每只麻雀的适应度值并找到最佳适应度和最差适应度值。

步骤3)分别利用式(2)~(4)计算得到发现者、追随者、侦查者的新位置,若新位置的适应度值优于之前的,则更新位置。

步骤4)重复步骤3),不断地进行迭代更新最佳麻雀位置。当迭代次数达到步骤1)设置的最大迭代次数时,迭代停止。最优解为所有次迭代中适应度值最低的那只麻雀。

1.2 BiLSTM 神经网络

LSTM 是循环神经网络(recurrent neural network, RNN)的一种。LSTM 结构的设计是为了避免 RNN 出现的长期依赖性问题,且 LSTM 在处理时间序列问题时,具有更强的时间序列学习能力及信息选择能力,能够更好地协调历史记忆单元中的信息分配,具有比 RNN 更好协调记忆的能力^[13]。LSTM 主要的改进是在 RNN 的结构基础上引入3个门,为遗忘门、输入门、输出门,解决 RNN 无法有效处理长距离信息,易出现梯度消失和梯度爆炸问题^[14]。LSTM 网络结构如图1所示。

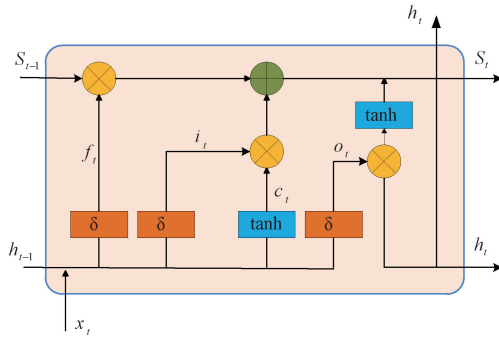


图1 LSTM网络结构

遗忘门、输入门、记忆单元状态和输出门的计算公式如下:

$$f_t = \delta(W_f h_{t-1} + U_f x_t + b_f) \quad (5)$$

$$i_t = \delta(W_i h_{t-1} + U_i x_t + b_i) \quad (6)$$

$$c_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \quad (7)$$

$$o_t = \delta(W_o h_{t-1} + U_o x_t + b_o) \quad (8)$$

式中: x_t 表示 t 时刻的输入; h_{t-1} 表示 $t-1$ 时刻的隐藏层状态值; W_f 、 W_i 、 W_o 和 W_c 分别表示遗忘门、输入门、输出门和特征提取过程中 h_{t-1} 的权重系数; U_f 、 U_i 、 U_o 和 U_c 分别表示遗忘门、输入门、输出门和特征提取过程中 x_t 的权重系数; b_f 、 b_i 、 b_o 和 b_c 分别表示遗忘门、输入门、输出门和特征提取过程中的偏置值; $\tanh$ 和 δ 分别表示正切双曲函数和激活函数。

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (9)$$

$$\delta(x) = \frac{1}{1 + e^{-x}} \quad (10)$$

由图1可得 t 时刻的细胞状态 S_t 为:

$$S_t = S_{t-1} \otimes f_t + i_t \otimes c_t \quad (11)$$

最后,可知 t 时刻的隐藏层状态值 h_t 为:

$$h_t = o_t \otimes \tanh(S_t) \quad (12)$$

BiLSTM 是在 LSTM 基础上的一种优化改进。由于 LSTM 训练数据时,只能够利用从正向序列中传来的历史信息,从而会导致对于较早学习过的内容的遗忘。而 BiLSTM 不同的是,由两个独立的正反 LSTM 构成,既有利于前向序列信息输入又有利于后向序列信息输入,可以考虑历史数据的全局信息,更好的学习和训练历史数据,有利于提高预测精度^[15]。BiLSTM 网络结构如图2所示。

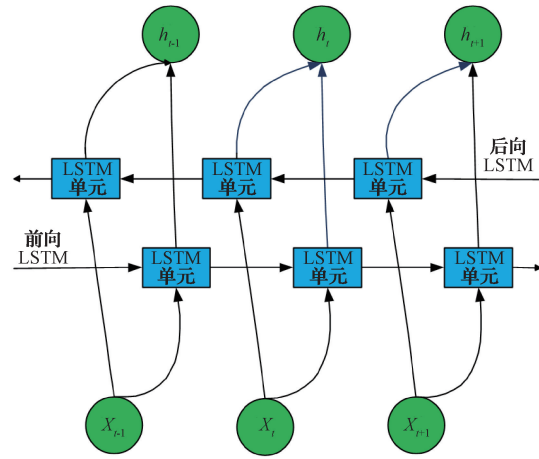


图2 BiLSTM网络结构

1.3 注意力机制

注意力机制是一种模仿人脑注意力的模型。人脑在处理信息时,通常会将更多的注意力集中在关键区域,以获得重要信息^[16]。注意力机制对输入特征分配不同权重,以此来突出强相关性因素影响,减少弱相关性因素影响,增强对历史数据关键信息的注意程度,提高预测精度^[17]。其结构如图3所示。其中 h_1 、 h_{t-1} 、 h_t 、 h_n 是输入序列 X_1 、 X_{t-1} 、 X_t 、 X_n 的隐藏层状态值; a_1 、 a_{t-1} 、 a_t 、 a_n 为隐藏层输出值 h_1 、 h_{t-1} 、 h_t 、 h_n 对当前输入的注意力权重值。

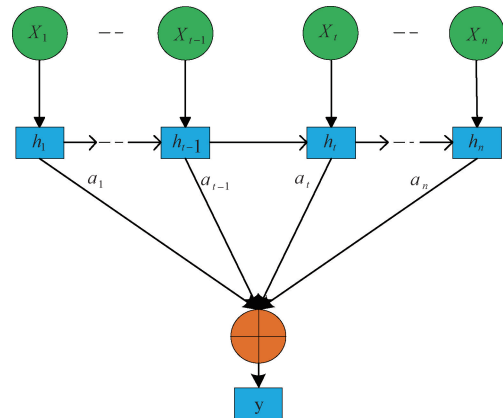


图3 注意力机制结构

双向长短时记忆结合注意力机制,不仅可以长时间保留历史信息,还能判断出重要时间序列信息,增加重要时间序列信息对结果的影响权重,以此提高预测精度^[18]。

2 SSA-BiLSTM-AM 模型

由于原始的风电功率、风向、风速、气压等数据的数量单位和数量级都不同,直接采用原始风电功率数据,会导致预测误差偏大。因此,对原始数据采用归一化处理,可以提升算法的精度和收敛速度。

LSTM 作为 RNN 的改进,可以解决 RNN 的长期依赖性问题,但在时间序列过长时会丢失某些信息,而双向 LSTM 可以很好解决序列过长而导致信息的丢失。

注意力机制可以关注信息的分配并且对重要信息给予更大关注程度,突出影响结果的关键特征,使预测效果更好。

神经网络模型在预测过程中,参数的选择是凭借经验进行设定的,而不同的参数选取对预测结果影响较大。为此,本文采用收敛度和精度都表现良好的麻雀算法对 BiLSTM-AM 网络的学习率、训练次数、Batchsize、第 1 隐层节点数、第 2 隐层节点数、全连接层节点数等 6 个超参数进行优化选择,以均方差函数作为适应度函数,为 BiLSTM-AM 网络寻得最优参数,解决了超参数的选择问题。

综上,本文构建了基于 SSA-BiLSTM-AM 的短期风电功率预测模型,整个预测流程如图 4 所示。

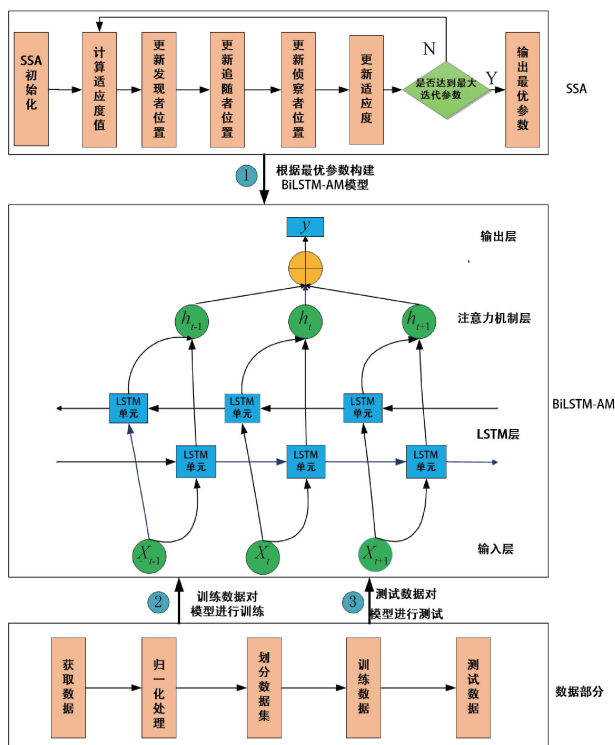


图 4 SSA-BiLSTM-AM 预测流程

3 算例分析

3.1 数据描述及预处理

为了验证本文所提预测模型的准确性,选取了新疆风电场 2019 年 1 月 1 日~31 日的风电功率数据,采样周期为 15 min。取前 30 d 共 2 880 条数据作为训练集,最后一天共 96 条数据作为测试集。数据集包含风速、风向、温度、湿度以及气压等特征。

原始数据归一化处理可以加快损失函数的收敛速度,提高预测精度,因为真实数据通常包含多种不同量纲的数据,需要对数据进行归一化处理^[15]。本文采用 min-max 标准化,归一化结果如图 5 所示。计算公式如式(13)所示。

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (13)$$

式中: x 为原始数据; x^* 为归一化后的值; $x_{\max}$ 和 $x_{\min}$ 分别为数据的最大值和最小值。

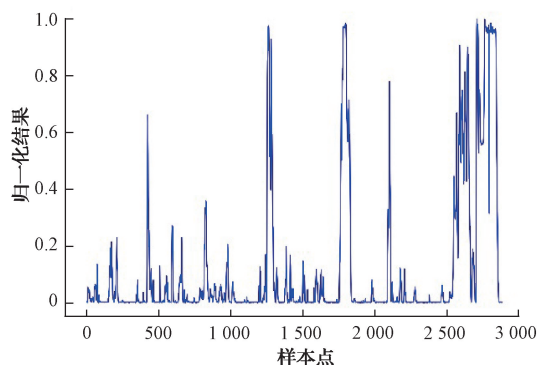


图 5 数据归一化结果

3.2 模型评估标准

为了验证模型的可靠性,本文将平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)和拟合系数 R^2 3 个指标依据各个模型预测结果进行对比分析^[19-20],公式如下:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (14)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (15)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (16)$$

式中: m 为样本数据总数; y_i 真实值; $\hat{y}_i$ 为预测值; $\bar{y}$ 为真实值的平均值。

3.3 基于 SSA 优化 BiLSTM-AM 的结果

通过 SSA 优化得到的 6 个超参数如图 6 所示。由图 6 可得优化后的训练次数、Batchsize、学习率、第 1 隐层

节点数、第2隐含层节点数和全连接层节点数分别为91、42、0.0048、7、16和99。

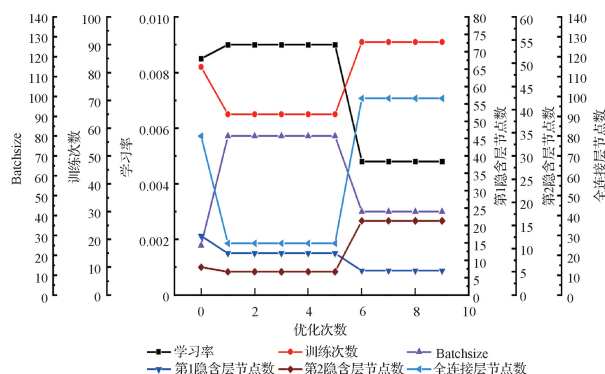


图6 参数优化结果

将SSA优化后的参数输入BiLSTM-AM模型进行训练和预测,得到SSA-BiLSTM-AM模型的最终预测曲线如图7所示。由图7可知,SSA-BiLSTM-AM模型预测结果接近实际值,预测效果较好。

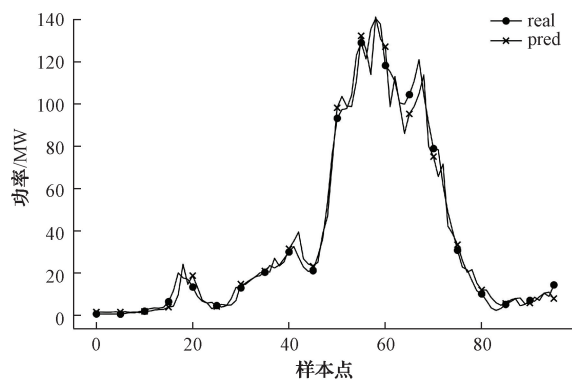


图7 SSA-BiLSTM-AM预测曲线

3.4 预测结果对比

为了体现本文所提的短期风电功率预测模型的效果,将本文模型与LSTM、BiLSTM以及BiLSTM-AM模型进行比较,各模型预测曲线如图8所示。

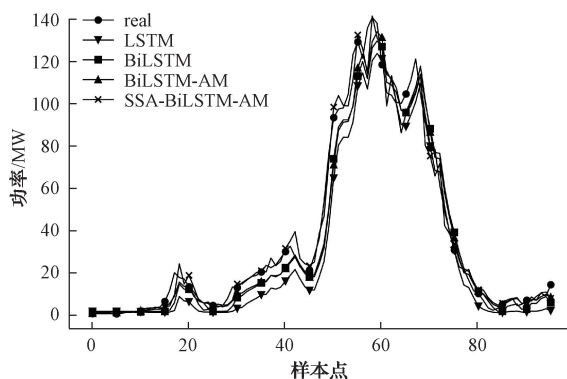


图8 不同模型预测结果

各模型评价指标如表1所示。由表1可知,相较于LSTM,BiLSTM的RMSE、MAE分别降低了3.2648、2.6505, R^2 提升3.68%,说明BiLSTM的双向时序特征提取可以考虑历史数据的全局信息,有利于提高预测结果的准确性;BiLSTM-AM与BiLSTM相比,其RMSE、MAE分别降低0.3608、0.6147, R^2 提升0.3%,可以看出通过注意力机制对不同特征分配不同的权重,可以提高预测精度;SSA-BiLSTM-AM较于BiLSTM-AM,其RMSE、MAE分别降低了2.1905、1.3726, R^2 提升1.61%,表明SSA对BiLSTM-AM的参数进行了优化选择,解决了参数选择困难问题,可以有效提高预测结果,从而体现了本文所提模型的可行性。

表1 不同模型预测精度

预测模型	RMSE	MAE	R^2
LSTM	11.227 5	8.312 7	0.930 9
BiLSTM	7.962 7	5.662 2	0.965 2
BiLSTM-AM	7.601 9	5.047 5	0.968 3
SSA-BiLSTM-AM	5.411 4	3.674 9	0.983 9

4 结论

为了提高风电功率的预测精度,本文提出了SSA-BiLSTM-AM模型,得出以下结论。LSTM作为RNN神经网络的一种改进,可以有效学习时序的长时间性,但忽略了序列存在的双向信息,而BiLSTM神经网络可以挖掘历史序列的正、反向信息,相比LSTM预测模型,BiLSTM预测模型的误差指标RMSE、MAE均减小。AM通过对不同特征分配不同的权重,获取不同特征的贡献程度,突出对结果影响较大的特征,可以提高模型的预测精度。针对网络参数选择困难,采用SSA优化BiLSTM-AM模型参数,找出一组最优参数,使得模型的误差最小,相比于LSTM、BiLSTM、BiLSTM-AM模型,预测精度有明显提升,验证了本文模型的优越性。

由于数据采集的时候不可避免的会存在异常数据等问题,后续会在此基础上考虑进行特征筛选。

参考文献

- [1] 王慧莹,吴亮红,梅盼盼,等. 果蝇优化广义神经网络的风电功率短期预测[J]. 电子测量与仪器学报, 2019,33(6):177-183.
- [2] 吴问足,乔颖,鲁宗相,等. 风电功率概率预测方法及展望[J]. 电力系统自动化,2017,41(18):167-175.
- [3] 丁婷婷. 考虑气象分类的短期风电功率组合预测方法研究[D]. 济南:山东大学,2022.
- [4] 赵凌云,刘友波,沈晓东,等. 基于CEEMDAN和改进时间卷积网络的短期风电功率预测模型[J]. 电力系统保护与控制,2022,50(1):42-50.
- [5] SON N, YANG S, NA J. Hybrid forecasting model

- for short-term wind power prediction using modified long short-term memory[J]. *Energies*, 2019, 12(20): 3901.
- [6] 王佑宣, 邓斌, 王江. 基于经验模态分解与 RBF 神经网络的短期风功率预测[J]. *电力系统及其自动化学报*, 2020, 32(11): 109-115.
- [7] 史加荣, 赵丹梦, 王琳华, 等. 基于 RR-VMD-LSTM 的短期风电功率预测[J]. *电力系统保护与控制*, 2021, 49(21): 63-70.
- [8] 赵婧宇, 池越, 周亚同. 基于 SSA-LSTM 模型的短期电力负荷预测[J]. *电工电能新技术*, 2022, 41(6): 71-79.
- [9] 曾亮, 雷舒敏, 王珊珊, 等. 基于 OVMD-SSA-DELM-GM 模型的超短期风电功率预测方法[J]. *电网技术*, 2021, 45(12): 4701-4712.
- [10] 常东峰, 南新元. 基于混合麻雀算法改进反向传播神经网络的短期光伏功率预测[J]. *现代电力*, 2022, 39(3): 287-298.
- [11] 郭建帅, 崔双喜, 郭建斌, 等. 基于 VMD-SSA-HKELM 的超短期负荷预测[J]. *国外电子测量技术*, 2022, 41(6): 105-111.
- [12] 杨邓, 杨俊杰, 胡晨阳, 等. 基于改进 LSSVM 的短期电力负荷预测[J]. *电子测量技术*, 2021, 44(18): 47-53.
- [13] 王伟, 刘宏伟, 陈永杰, 等. 基于 LSTM 循环神经网络的风力发电预测[J]. *可再生能源*, 2020, 38(9): 1187-1191.
- [14] 沙骏, 徐雨森, 刘冲冲, 等. 基于变分模态分解和分位数卷积-循环神经网络的短期风功率预测[J]. *中国电力*, 2022, 55(12): 61-68.
- [15] 任建吉, 位慧慧, 邹卓霖, 等. 基于 CNN-BiLSTM-Attention 的超短期电力负荷预测[J]. *电力系统保护与控制*, 2022, 50(8): 108-116.
- [16] 邵必林, 严义川, 曾卉玢. 注意力机制下的 VMD-ID-BiGRU 负荷预测模型[J]. *电力系统及其自动化学报*, 2022, 34(10): 120-128.
- [17] 王继东, 杜冲. 基于 Attention-BiLSTM 神经网络和气象数据修正的短期负荷预测模型[J]. *电力自动化设备*, 2022, 42(4): 172-177, 224.
- [18] 马磊, 黄伟, 李克成, 等. 基于 Attention-LSTM 的光伏超短期功率预测模型[J]. *电测与仪表*, 2021, 58(2): 146-152.
- [19] 张淑清, 要俊波, 张立国, 等. 基于改进深度稀疏自编码器及 FOA-ELM 的电力负荷预测[J]. *仪器仪表学报*, 2020, 41(4): 49-57.
- [20] 王永超, 唐求, 马俊, 等. 基于复合核 SVM 的智能电表基本误差预测方法[J]. *电子测量与仪器学报*, 2021, 35(10): 209-216.

作者简介

章志晃, 硕士研究生, 主要研究方向为电力负荷预测和机器学习。

E-mail: 442332946@qq.com

徐启峰, 教授, 博士生导师, 主要研究方向为智能电网测量与控制新技术。

E-mail: ranger123098@163.com

林穿, 博士研究生, 主要研究方向为电力设备故障诊断、人工智能与电力大数据分析。

E-mail: 492456598@qq.com