

DOI:10.19651/j.cnki.emt.2518051

# 基于聚类的 AW-CNN-LSTM 光伏功率预测方法<sup>\*</sup>

刘丽丽<sup>1</sup> 谢梦<sup>1</sup> 王艳<sup>1</sup> 杨春蕾<sup>1</sup> 顾明剑<sup>1,2</sup>

(1. 中科技术物理苏州研究院 苏州 215000; 2. 云遥动力科技(苏州)有限公司 苏州 215000)

**摘要:** 由于光伏发电的波动性和随机性,传统模型难以对其进行准确预测。针对这一问题,在聚类的基础上建立一种自适应权重的 CNN-LSTM 网络模型。首先,对光伏电站历史数据进行预处理,并采用基于肘部法的 K-means 算法进行聚类;其次,根据同聚类簇的训练样本与对应测试样本的特征中心间的距离建立自适应权重;然后,根据聚类结果和自适应权重建立适用于不同聚类类别的 AW-CNN-LSTM 网络模型,其中,CNN 用于捕获不同特征间的关系,LSTM 用于捕获时序特征;最后对各模型预测结果进行整合得到最终预测结果。在澳大利亚沙漠太阳能研究中心的光伏电站数据集上进行试验,证明了本文所提方法的有效性。

**关键词:** 光伏功率预测;自适应权重;K-means 聚类;卷积神经网络;长短期记忆网络

中图分类号: TN06; TP271 文献标识码: A 国家标准学科分类代码: 470.4047

## AW-CNN-LSTM photovoltaic power prediction method based on clustering

Liu Lili<sup>1</sup> Xie Meng<sup>1</sup> Wang Yan<sup>1</sup> Yang Chunlei<sup>1</sup> Gu Mingjian<sup>1,2</sup>

(1. Suzhou Institute of Science and Technology Physics, Suzhou 215000, China;

2. Yunyao Power Technology (Suzhou) Co., Ltd., Suzhou 215000, China)

**Abstract:** Due to the volatility and randomness of photovoltaic power generation, it is difficult for traditional models to accurately predict it. To solve this problem, a prediction model of AW-CNN-LSTM is established based on clustering. First, the photovoltaic power plant historical data set is pre-processed and clustered using the K-means clustering algorithm based on the elbow method; secondly, an adaptive weight is established based on the distance between the training samples and the feature center of test samples of the same clustering category; then, an AW-CNN-LSTM network model suitable for different clustering categories is established based on the clustering results and adaptive weights. CNN are used to capture the relationships between different features, while LSTM are used to capture temporal features. Finally, the forecast results of each model are integrated to get the final forecast results. Experiments on the data set of photovoltaic power stations in the Australian Desert Solar Energy Research Center demonstrate the effectiveness of the proposed method.

**Keywords:** photovoltaic power prediction; adaptive weight; K-means clustering; convolutional neural networks; long short term memory networks

## 0 引言

光伏发电是最易获得、成本低且最有前途的可再生能源之一。然而,由于光伏发电的间歇性和随机性,现有电力系统的运行和规划对光伏功率的预测精度提出了挑战<sup>[1]</sup>。可靠的预测结果有助于提高系统稳定性<sup>[2]</sup>,因此,如何提高光伏功率的预测精度是研究的难点。

目前,主流的光伏功率预测模型分为物理模型、统计模型和深度学习模型 3 种,其中,深度学习模型是目前最有效

的预测模型<sup>[3]</sup>。物理模型方法主要根据数值天气预报(numerical weather prediction, NWP)内的各种气象资料预测电场的输出功率,但是该方法的预测精度较低,且仅适用于中长期预测<sup>[4]</sup>。统计模型方法是基于数理统计理论对时间序列数据进行曲线拟合和参数估计并建立数学模型,但该方法只适用于平滑的能源数据。深度学习方法是通过构建多层神经网络,自动提取和学习复杂特征与模式,从而显著提高光伏功率预测精度的一种先进技术。

近年来,深度学习模型展现出更显著的优势,在光伏功

收稿日期:2025-02-11

\* 基金项目:苏州市科技项目(SYG202135)、苏州市社会发展项目(2023ss17)资助

率预测领域得到更广泛的应用。文献[5]提出了一种结合分解算法和深度学习的风电和光伏超短期智能预测系统,该系统将卷积神经网络(convolutional neural network,CNN)作为基准模型,具有高度的泛化性和自适应能力。文献[6]提出了基于 CNN 和长短期记忆神经网络(long short term memory networks,LSTM)的混合神经网络模型,并将该混合模型和常用的单个基准模型相比,证明了混合神经网络的有效性。文献[7]提出基于 CNN 和 LSTM 的混合网络模型,CNN 将天气状况进行二分类,LSTM 对不同天气状况学习发电模式,并利用该模型进行发电预测,结果表明该模型可根据发电模式的功率变化进行准确的发电预测。文献[8]提出了基于深度学习的 CNN-LSTM 双流融合网络,CNN 学习空间信息,LSTM 进行时态特征提取,最终融合时间和空间特征。文献[9]利用 CNN 对不同区域的空间相关性进行分析,并通过 LSTM 捕获发电数据的时间序列特征,结果表明,与传统 LSTM 模型相比,CNN-LSTM 模型有很大的改进。以上方法表明,与单一模型相比,CNN-LSTM 混合深度学习模型能同时提取空间和时间特征,在光伏功率预测方面有很大的成效。尽管混合深度神经网络表现出色,却忽略了光伏电站历史数据中天气类别的不平衡问题。通过将具有相似特征的天气数据归为一类并建立对应的多个预测模型,可以有效解决数据不平衡问题。文献[10]通过对历史功率进行聚类,提出一种基于灰色关联度分析的未来波动天气事件识别匹配方法,从而实现天气过程的场景划分。文献[11]利用平衡迭代规约和聚类模块挖掘数据中的信息,划分特征库,并根据光伏功率的波动性建立对应的双向长短记忆网络预测模型。文献[12]提出一种基于莱维飞行天牛须搜索算法的相似日选择方法,以选择相似于预测日的历史日作为输入数据集,建立一组深度学习光伏功率预测模型,将每个子序列预测结果叠加得到光伏功率最终预测结果。以上方法均对天气过程进行划分,克服了光伏电站历史数据类别不平衡的部分影响,但所有样本对模型训练具有相同的权重,均未解决相似点划分错误对模型性能的影响。如何减少相似点划分错误带来的负面影响,以提高模型在多变天气条件下的预测准确性,依然是亟待解决的难题。

针对以上难题,本文结合聚类、卷积神经网络、长短期记忆网络的优点,在聚类的基础上建立一种自适应权重的 CNN-LSTM 网络模型(a hybrid model with adaptive weights based on convolutional neural networks and long short-term memory network, AW-CNN-LSTM)光伏功率预测模型,采用基于肘部法的 K-means 聚类算法将相似天气类型的样本进行提取,并考虑了不同样本的权重差异,设置的自适应权重为每个训练样本自适应地分配不同权重,构建适用于不同聚类类别的 AW-CNN-LSTM 的光伏功率预测模型。该方法期望降低后续预测模型的复杂性,自动确定样本重要性并赋予样本权重,减少了聚类错误对模型

性能的影响,旨在提高光伏功率的预测精度。

## 1 方法介绍

本文提出一种基于聚类的混合神经网络模型,基于光伏电站历史数据进行预处理及样本划分,获得训练样本和测试样本;对训练样本和测试样本进行聚类处理,确定聚类结果;根据同聚类簇中训练样本与对应测试样本的特征中心间的距离设置自适应权重;根据聚类结果以及自适应权重,构建适用于不同聚类类别的 AW-CNN-LSTM 的光伏功率预测模型;基于 AW-CNN-LSTM 的光伏功率预测模型,生成光伏电站的光伏功率预测结果,最终将多模型预测结果进行整合得到最终预测结果。具体流程图如图 1 所示。

### 1.1 基于肘部法的 K-means 聚类

对于超短期功率预测, $T+1$  时刻与  $T$  时刻天气特征和光伏功率关系紧密。为了更好地预测  $T+1$  时刻的光伏功率,将  $T$  时刻的天气特征及光伏功率进行聚类,从而将相似的天气聚为一簇,为后续基于聚类的 AW-CNN-LSTM 深度学习模型提供基础。

聚类是一种无监督的机器学习方法,聚类的目标是将数据点划分为若干个簇,使得同一簇内的数据点相似度高、不同簇间的数据点相似度低<sup>[3]</sup>。常用的聚类方法有 K-means 聚类、层次聚类、密度聚类等,其中 K-means 聚类是广泛认可和最常用的聚类方法<sup>[4]</sup>。K-means 聚类将聚成预先指定数量的簇,通过优化迭代簇内数据点的均值来达到最小化样本与簇中心之间距离的目标<sup>[5]</sup>。K-means 聚类的簇数  $K$  不同聚类结果差异较大,为了克服  $K$  均值聚类对  $K$  的敏感度,采用肘部法确定聚类最优数。

肘部法的核心思想是随着簇数目的增加,簇内平方和(sum of squared errors, SSE)会逐渐减小。SSE 是衡量聚类效果的一个重要指标,它表示每个数据点到其所属簇中心的距离的平方和<sup>[6]</sup>。在肘部法中,根据不同数量的簇并计算每个簇对应的 SSE 值。然后,通过绘制簇数目与对应 SSE 的图表,观察 SSE 随簇数目增加而减小的趋势,找到拐点即可确定最佳的簇数目<sup>[7]</sup>。图 2 是基于肘部法的 K-means 聚类流程图。

### 1.2 自适应权重

为了得到更好的预测效果,根据聚类结果计算每个训练样本距离测试样本的特征中心间的距离,自适应调整训练样本权重。这里将所有元素的特征向量的平均值作为该聚类簇的特征中心。通过为训练样本设置权重,可以指定某些样本对于模型的学习具有更高的影响力,从而更好地适应测试样本的分布,这对于处理样本不平衡或关注特定聚类类别的情况非常有用。

为了得到最佳权重,使得训练的模型偏向同聚类类别的测试样本。采用训练样本与测试样本特征中心距离的倒数作为权重系数,从而将与测试样本较相似的训练样本赋

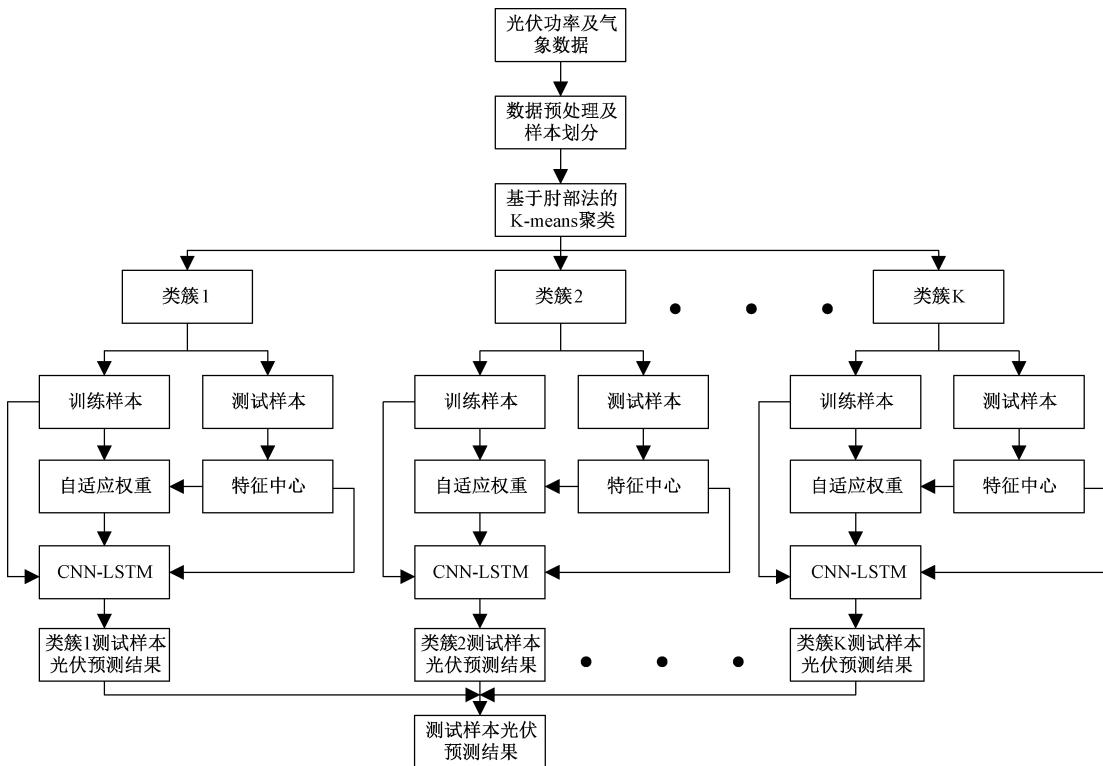


图 1 光伏功率预测流程图

Fig. 1 Photovoltaic power forecast flow chart

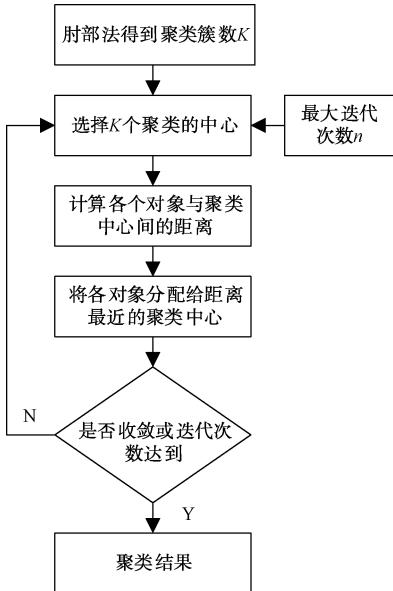


图 2 基于肘部法的 K-means 聚类

Fig. 2 K-means clustering based on elbow method

予大的权重,与测试样本差异小的训练样本赋予较小的权重。设测试样本中聚类类别为  $m$  的样本个数为  $n$ ,  $x_i$  为测试样本中第  $i$  个样本的特征向量,则聚类类别为  $m$  的测试样本的特征中心  $C_m$  如式(1)所示。聚类类别为  $m$  的第  $j$  个训练样本的权重为  $w_{m,j}$ , 对应的特征向量为  $x'_j$ , 则该样

本的权重如式(2)所示。

$$C_m = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

$$w_{m,j} = \frac{1}{\sqrt{(x'_j - C_m)^2}} \quad (2)$$

### 1.3 CNN-LSTM 模型

根据 CNN 和 LSTM 的特点,本文建立了基于聚类的 AW-CNN-LSTM 的光伏功率预测模型。一维 CNN 层使用了一种提高数据效率和表示模块化的卷积技术<sup>[18]</sup>,通过卷积运算<sup>[19]</sup>从包含因子相关性信息的输入数据中提取有价值的特征<sup>[20]</sup>。LSTM 层进行时间序列学习,提取序列模式信息以及短期和长期依赖关系<sup>[21]</sup>。因此,利用 CNN-LSTM 混合深度学习模型有望提高光伏功率预测的准确性。

基于 CNN-LSTM 的深度学习模型架构如图 3 所示。特征提取模块由 3 个一维 CNN 层组成,在深度学习框架中堆叠多个卷积层使初始层能够学习输入的低级特征。在时间序列学习模块中使用了两个 LSTM 层,前一个 LSTM 层的返回序列被设置为 true,以便网络将输出隐藏状态的完整序列后一个 LSTM 层,返回序列设为 false,使网络在最后的时间步长输出隐藏状态。时间序列学习后有两个全连接层,最后的输出层中的神经元数量根据输出特征数而定。激活函数用于增强模型学习复杂结构的能

力,本文采用的线性整流函数(rectified linear unit,ReLU)可减弱梯度消失问题,已被各种研究人员广泛应用,使网络更具可训练性<sup>[22]</sup>。在本实验中,采用平均绝对误差作为损失函数,并使用了著名的优化器,即适应性矩估计

(adaptive moment estimation,Adam)。Adam 是一种常用的随机梯度下降算法,它结合了适应性梯度算法和动量法两种优化算法的优点,具有较快的收敛速度和较好的泛化性能<sup>[23]</sup>。CNN-LSTM 网络框架的参数设置如表 1 所示。

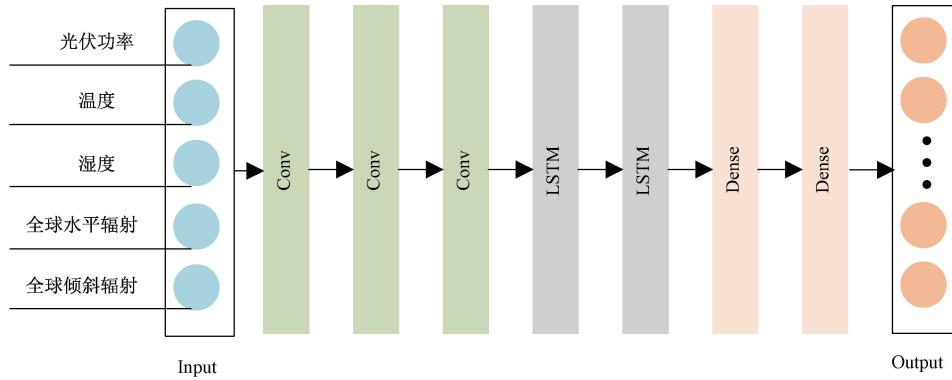


图 3 CNN-LSTM 网络

Fig. 3 CNN-LSTM network

表 1 CNN-LSTM 网络参数

Table 1 CNN-LSTM network parameters

网络层	输出大小	参数数量
Conv	(None, 13, 32)	128
Conv	(None, 11, 64)	6 280
Conv	(None, 9, 128)	24 704
Lstm	(None, 9, 100)	91 600
Lstm	(None, 100)	80 400
Dense	(None, 64)	6 464
Dense	(None, 32)	2 080
Output	(None, 1)	33

## 2 实验结果与分析

### 2.1 数据预处理

本文选取公开发布的澳大利亚太阳能中心的艾丽丝泉光伏数据和天气数据。数据 1 为第 25 站点的 2016 年 7 月 5 号~2023 年 10 月 6 号间的光伏功率和相关天气数据。数据 2 为第 32 站点 2016 年 11 月 9 号至 2023 年 9 月 30 日的光伏功率和相关天气数据,数据分辨率由 5 min 降采样为 15 min。

由于数据的质量严重影响着深度学习的性能,为了实现准确的预测,需对数据进行预处理,包括时间段选取、异常值去除、缺失值填补、数据归一化及特征选择。本文选取 7:30~18:00 之间的数据,对选取的数据去除异常值,并用相邻非离群数据的平均值进行缺失值填补。在输入缺失值后,需要进行数据归一化,并采用皮尔逊相关系数进行特征选取。其中皮尔逊相关系数用于衡量光伏功率与其它变量之间的线性相关关系,取值范围为 [-1, 1],相关性越强越接近于 1 或 -1, 相关性越弱越接近于 0, 值为负

值表示负相关,值为正表示正相关。公式如式(3)所示,其中  $r$  为皮尔逊相关系数,代表变量  $X$  和  $Y$  之间的线性相关程度,  $n$  为变量的样本个数,  $X_i$  和  $Y_i$  分别是变量  $X$  和  $Y$  的第  $i$  个样本值,  $\bar{X}$  是  $X$  的样本平均值,  $\bar{Y}$  是  $Y$  的样本平均值。

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3)$$

表 2 为数据 1 和数据 2 的光伏功率与各变量之间的皮尔逊相关系数。由表 2 可知,数据 1 和数据 2 分别选择皮尔逊相关系数较大的变量,包括光伏功率、温度、湿度、全球水平辐射和全球倾斜辐射 5 个变量。

表 2 光伏功率与各变量间的皮尔逊相关系数

Table 2 Pearson correlation coefficient between photovoltaic power and meteorological factors

变量	数据 1	数据 2
光伏功率	1.000	1.000
温度	0.202	0.129
湿度	-0.335	-0.307
全球水平辐射	0.921	0.806
漫射水平辐射	0.107	0.086
风向	0.011	0.060
降水量	-0.095	-0.101
全球倾斜辐射	0.899	0.787
漫射倾斜辐射	0.123	0.094

### 2.2 评价指标

本文选择均方根误差(root mean square error,

RMSE)、平均绝对误差(mean absolute error, MAE)和相关系数(correlation coefficient, CC)作为光伏功率评价指标。其中 RMSE 是衡量实际观测值和模型预测值之间的差异, 它表示模型预测值和真实值之间的平均偏差。MAE 是预测值和真实值间绝对误差的平均值, 可更好地反映预测值误差的真实情况。CC 用于衡量两个变量之间的相关性, 越接近于 1 表示效果越好。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - O_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - O_i| \quad (5)$$

$$CC = \frac{\sum_{i=1}^n (O_i - \bar{O})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (f_i - \bar{f})^2}} \quad (6)$$

式(4)、(5)、(6)中,  $n$  是样本的总数,  $O_i$  和  $f_i$  是第  $i$  对被检验的真实光伏功率和预测光伏功率值,  $\bar{O}$  和  $\bar{f}$  分别是它们对应的平均值。

## 2.3 实验结果

### 1) 基于肘部法的 K-means 聚类

通过观察肘部法曲线, 在拐点处选择簇数  $K$ , 即所谓的“肘部”位置。这个位置代表了在簇数增加到一定程度后, 继续增加簇数对聚类性能提升的贡献已经不显著, 因此选择拐点对应的簇数作为最优的聚类数目。

图 4 的纵坐标表示簇内平方和, 横坐标表示簇类数量, 通过“拐点”找出最佳聚类簇数, 即当  $k=3$  时为肘部, 数据 1 和数据 2 的最佳聚类簇数为 3, 因此, 采用 K-means 聚类将所有样本分为 3 个类簇。

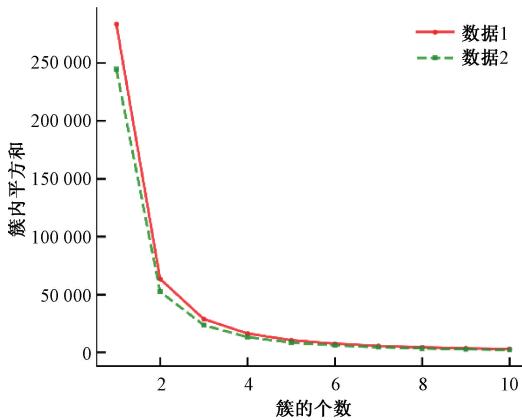


图 4 肘部法曲线

Fig. 4 Elbow method curve

### 2) AW-CNN-LSTM 光伏功率预测

在目前的研究中, 测试了深度模型的各个超参数的不同组合, 以优化它们的学习参数。经多次实验发现当激活函数为 ReLU, 优化器为 Adam, 损失函数为 RMSE, 迭代次数为 100, batch size 为 20, 学习率为 0.001 时, 模型的性

能最佳。按照时间顺序将数据 1 和数据 2 的前 80% 作为训练样本, 后 20% 的时间样本作为测试样本, 预测未来 15 min 的光伏功率。

图 5 和图 6 分别为数据 1 和数据 2 的光伏功率部分预测结果, 结果中展示了连续 10 天的光伏功率预测值, 其中本文方法简称为 AW-CNN-LSTM, 真实功率为真实值。从图 5 和 6 中可以看出本文方法在不同数据集上预测曲线拟合良好, 说明预测功率和真实功率极为相近, 研究提出的 AW-CNN-LSTM 模型性能非常优异。

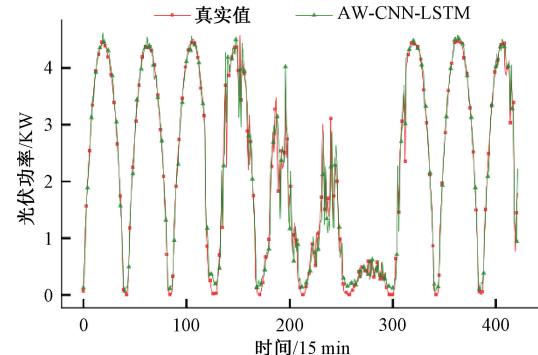


图 5 本文方法光伏功率预测结果(数据 1)

Fig. 5 The results of photovoltaic power prediction in this paper(data 1)

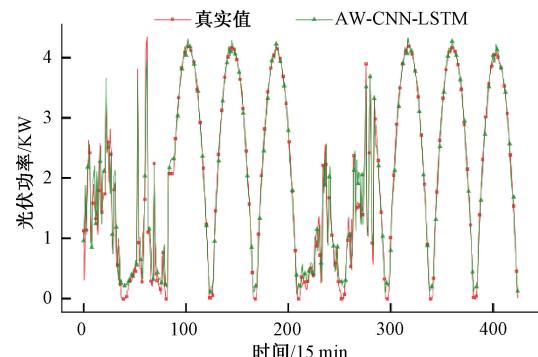


图 6 本文方法光伏功率预测结果(数据 2)

Fig. 6 The results of photovoltaic power prediction in this paper(data 2)

## 2.4 对比方法

为了证明本文方法的有效性, 采用几种深度学习方法作为对比, 包括 LSTM、CNN-LSTM 及门控循环神经网络(gated recurrent unit, GRU)网络。CNN-LSTM 和本文方法 AW-CNN-LSTM 的网络中的 CNN-LSTM 的结构一致, GRU 是 LSTM 的一种变体, CNN-BiLSTM-Attention 方法<sup>[24]</sup>结合了 CNN 和双向长短期记忆网络(bi-directional long and short-term memory network, BiLSTM)以及注意力机制。所有模型均在相同的数据集和环境下进行训练和测试。

表 3 和 4 分别为两个数据集的光伏功率预测精度对

比。从表 2 中可以看出,数据 1 采用本文方法的精度最高,GRU 方法的精度最低,相比 LSTM、CNN-LSTM、GRU 和 CNN-BiLSTM-Attention 方法,本文方法的 MAE 分别降低了 40.82%、32.87%、43.14% 和 34.68%,RMSE 分别降低了 49.29%、48.43%、50.35% 和 47.41%,CC 分别提高了 2.48%、2.37%、2.59% 和 2.16%。从表 3 可以看出,数据 2 采用本文方法的效果最好,对应的 MAE 和 RMSE 最小,相关系数 CC 的值最大,与 LSTM、CNN-LSTM、GRU 和 CNN-BiLSTM-Attention 方法相比,本文方法的 MAE 分别降低了 40.43%、33.01%、42.62% 和 37.50%,RMSE 分别降低了 49.76%、48.89%、51.52% 和 48.00%,CC 提高了 3.01%、2.91%、3.12% 和 2.80%。与对比方法对比,本文方法在两个数据集上均取得了显著的性能提升,证明了所提模型具有较好的泛化性能。

表 3 光伏功率预测方法对比(数据 1)

Table 3 Comparison of photovoltaic power prediction methods(data 1)

方法	MAE	RMSE	CC
LSTM	0.245	0.420	0.968
CNN-LSTM	0.216	0.413	0.969
GRU	0.255	0.429	0.967
CNN-BiLSTM-Attention	0.222	0.405	0.971
AW-CNN-LSTM	<b>0.145</b>	<b>0.213</b>	<b>0.992</b>

表 4 光伏功率预测方法对比(数据 2)

Table 4 Comparison of photovoltaic power prediction methods(data 2)

方法	MAE	RMSE	CC
LSTM	0.235	0.414	0.962
CNN-LSTM	0.209	0.407	0.963
GRU	0.244	0.421	0.961
CNN-BiLSTM-Attention	0.224	0.400	0.964
AW-CNN-LSTM	<b>0.140</b>	<b>0.208</b>	<b>0.991</b>

图 7 和 8 分别为两个数据采用五种对比方法的光伏功

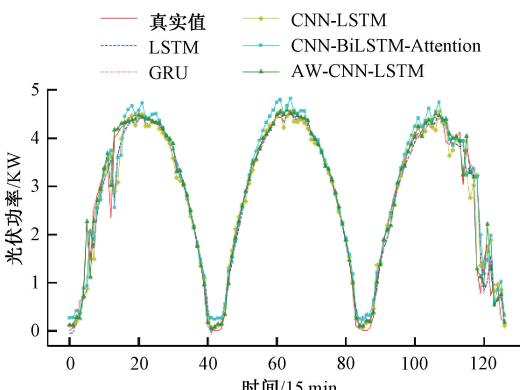


图 7 光伏功率预测方法对比(数据 1)

Fig. 7 Comparison of photovoltaic power prediction methods(data 1)

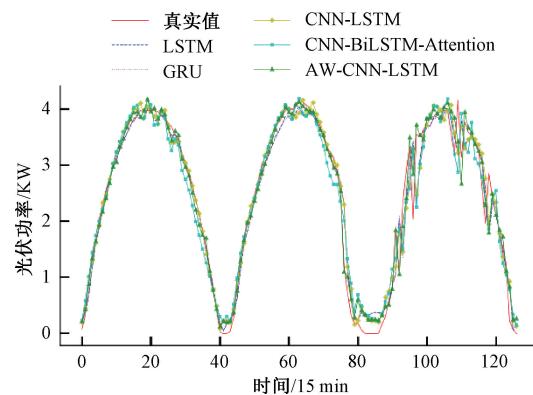


图 8 光伏功率预测方法对比(数据 2)

Fig. 8 Comparison of photovoltaic power prediction methods(data 2)

率预测结果曲线,图中展示了 3 天的光伏预测功率对比效果。从图中可以看出本文方法的曲线拟合真实功率的效果最好,其次是 CNN-BiLSTM-Attention 和 CNN-LSTM 方法,GRU 方法的预测效果最差,说明本文方法的精度最高。

### 3 结 论

针对目前光伏系统发电波动性和随机性导致预测精度不足的问题,提出 AW-CNN-LSTM 光伏功率预测模型,通过在两个公开的数据集上验证其有效性和可行性,本文主要贡献为:1)利用基于肘部法的 K-means 聚类将历史天气条件及光伏功率数据进行聚类,建立多个混合深度学习模型,从而针对每个聚类类别进行专门的模型训练和优化,使得模型对各类数据的特征捕捉更加精准。2)提出的基于聚类的自适应权重机制允许模型根据数据分布情况动态调整权重。将训练样本与同类簇测试样本特征中心的距离作为衡量该样本权重的尺度,使得训练的模型偏向该类别的测试样本,从而提高模型的预测精度。3)利用 AW-CNN-LSTM 网络在两个公开的数据集上进行光伏发电功率预测。结果表明,所提方法在数据 1 和数据 2 上的表现均优于现有的 CNN-BiLSTM-Attention、LSTM、CNN-LSTM、GRU 方法,显著提升了光伏功率预测的准确度。

### 参 考 文 献

- [1] 白隆,俞斌,高峰,等.基于 ICEEMDAN 和 TCN-AM-BiGRU 的短期光伏功率预测[J].电子测量技术,2024,47(9):61-69.  
BAI L, YU B, GAO F, et al. Short-term photovoltaic power prediction based on ICEEMDAN and TCN-AM-BiGRU [J]. Electronic Measurement Technology, 2024, 47(9): 61-69.
- [2] 袁建华,蒋文军,李洪强,等.基于 SSA-BiLSTM 非线性组合方法的光伏功率预测[J].电子测量技术,2023,46(21):63-71.

- YUAN J H, JIANG W J, LI H Q, et al. Photovoltaic power forecasting based on SSA-BiLSTM nonlinear combination method [J]. Electronic Measurement Technology, 2023, 46(21):63-71.
- [3] ASLAM S, HERODOTOU H, MOHSIN S M, et al. A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids[J]. Renewable and Sustainable Energy Reviews, 2021, 144: 110992.
- [4] YI Y, WANG L, CHEN Z. Adaptive global kernel interval SVR-based machine learning for accelerated dielectric constant prediction of polymer-based dielectric energy storage [J]. Renewable Energy, 2021, 176: 81-88.
- [5] ZHANG Y, PAN Z, WANG H, et al. Achieving wind power and photovoltaic power prediction: An intelligent prediction system based on a deep learning approach[J]. Energy, 2023, 283: 129005.
- [6] TOVAR M, ROBLES M, RASHID F. PV power prediction, using CNN-LSTM hybrid neural network model. Case of study: Temixco-morelos, méxico[J]. Energies, 2020, 13(24): 6512.
- [7] LIM S C, HUH J H, HONG S H, et al. Solar power forecasting using CNN-LSTM hybrid model [J]. Energies, 2022, 15(21): 8233.
- [8] ALHARKAN H, HABIB S, ISLAM M. Solar power prediction using dual stream CNN-LSTM architecture[J]. Sensors, 2023, 23(2): 945.
- [9] XU C, YU J, CHEN W, et al. Deep learning in photovoltaic power generation forecasting: Cnn-lstm hybrid neural network exploration and research[C]. The 3rd International Scientific and Practical Conference, 2024, 363: 295.
- [10] 葛俊雄,蔡国伟,姜柳,等.基于天气变化自适应分型与匹配的分布式光伏短期功率预测方法[J].激光与光电子学进展,2024,61(15): 1-12.
- GE J X, CAI G W, JIANG L, et al. Method for distributed photovoltaic short-term power prediction based on weather change adaptive fractal and matching [J]. Laser & Optoelectronics Progress, 2024,61(15): 1-12.
- [11] 张晓珂,张辉,戴小然,等.基于层次聚类和BiLSTM的光伏短期功率预测模型[J].智慧电力,2024,52(9): 41-48.
- ZHANG X K, ZHANG H, DAI X R, et al. Photovoltaic short-term power forecasting model based on hierarchical clustering & BiLSTM [J]. Smart Power, 2024,52(9):41-48.
- [12] 刘源延,孔小兵,马乐乐,等.基于小波包变换与深度学习的超短期光伏功率预测[J].太阳能学报,2024,45(5):537-546.
- LIU Y Y, KONG X B, MA L L, et al. Ultra-short-term photovoltaic power forecasting based on wavelet packet transform and deep learning[J]. Acta Energiae Solaris Sinica, 2024,45(5):537-546.
- [13] PITAFI S, ANWAR T, SHARIF Z. A taxonomy of machine learning clustering algorithms, challenges, and future realms[J]. Applied Sciences, 2023, 13(6): 3529.
- [14] MIRAFTABZADEH S M, COLOMBO C G, LONGO M, et al. K-means and alternative clustering methods in modern power systems[J]. IEEE Access, 2023.
- [15] JAHA S, SETU A A, MUNTAHA S, et al. Adaptive cluster center initialization using density peaks for geodesic distance-based clustering [J]. Numerical Algebra, Control and Optimization, 2025, 15(3): 601-618.
- [16] UMARGONO E, SUSENO J E, GUNAWAN S K V. K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula[C]. The 2nd International Seminar on Science and Technology (ISSTEC 2019). Atlantis Press, 2020: 121-129.
- [17] SCHUBERT E. Stop using the elbow criterion for k-means and how to choose the number of clusters instead[J]. ACM SIGKDD Explorations Newsletter, 2023, 25(1): 36-42.
- [18] KIM J, MOON J, HWANG E, et al. Recurrent inception convolution neural network for multi short-term load forecasting [J]. Energy and Buildings, 2019, 194: 328-341.
- [19] 王愈轩,刘尔佳,黄永章.基于CNN-LSTM-lightGBM组合的超短期风电功率预测方法[J].科学技术与工程,2022,22(36): 16067-16074.
- WANG Y X, LIU ER J, HUANG Y ZH. An ultra-short-term wind power prediction method based on CNN-LSTM-lightGBM combination [J]. Science Technology and Engineering, 2022, 22(36): 16067-16074.
- [20] 曹超凡,罗泽南,谢佳鑫,等. MDT-CNN-LSTM 模型的股价预测研究[J].计算机工程与应用,2022,58(5): 280-286.
- CAO CH F, LUO Z N, XIE J X, et al. Stock price prediction based on MDT-CNN-LSTM model [J]. Computer Engineering and Applications, 2022, 58(5): 280-286.
- [21] 张弛,朱宗玖.基于误差修正的短期光伏功率预测[J].科学技术与工程,2023,23(27): 11664-11672.

- ZHANG CH, ZHU Z J. Short-term PV power prediction based on error correction [J]. Science Technology and Engineering, 2023, 23 ( 27 ): 11664-11672.
- [22] JAHAN I, AHBED M F, ALI M O, et al. Self-gated rectified linear unit for performance improvement of deep neural networks[J]. ICT Express, 2023, 9(3): 320-325.
- [23] REYAD M, SARHAN A M, ARAFA M. A modified adam algorithm for deep neural network optimization[J]. Neural Computing and Applications, 2023, 35(23): 17095-17112.
- [24] ZHANG J, YE L, LAI Y. Stock price prediction using CNN-BiLSTM-Attention model[J]. Mathematics, 2023, 11(9): 1985.

### 作者简介

刘丽丽,硕士研究生,工程师,主要研究方向为人工智能、新能源技术与应用。

E-mail:2932361426@qq.com

谢梦(通信作者),硕士研究生,工程师,主要研究方向为人工智能、新能源发电与预测。

E-mail:15262036926@163.com