

基于随机森林-卷积神经网络混合集成模型的风力发电功率预测

李 桓 滕云雷

(国网山东省电力公司临沂供电公司, 山东 临沂 276000)

摘要 风电功率预测对于确保风能可靠地接入电网起着关键作用。本文提出一种随机森林 (RF)-卷积神经网络 (CNN) 混合模型, 用于短期风电功率预测。该模型融合 RF 的集成技术、随机选择属性与 CNN 捕获风电时空特征的优势, 增强预测的准确性和稳健性。首先, 通过分析决策树与 CNN 的类比等效性, 明确 RF 与 CNN 结合的理论依据; 然后, 构建包含方均根误差 (RMSE)、决定系数和 Spearman 相关系数的风电功率预测模型评估指标体系; 最后, 基于欧洲地区风电场的 3 个开源数据集进行模型有效性验证。结果表明: 与其他 5 种模型相比, RF-CNN 模型表现最优, 验证了该模型进行风电功率预测的有效性和准确性。

关键词: 风电功率预测; 随机森林; 卷积神经网络; 混合模型; 误差指标

Wind power prediction based on hybrid integrated model of random forest-convolutional neural network

LI Huan TENG Yunlei

(State Grid Shandong Electric Power Company Linyi Power Supply Company, Linyi, Shandong 276000)

Abstract Wind power prediction plays a crucial role in ensuring the reliable integration of wind energy into the grid. This study proposes a novel hybrid model combining random forest (RF) and convolutional neural network (CNN), referred to as the RF-CNN model, specifically designed for short-term wind power prediction. The model integrates the advantages of RF integration technology, random selection of attributes, and CNN capturing the spatiotemporal characteristics of wind power, to enhance prediction accuracy and robustness. Firstly, by analyzing the analog equivalence between decision trees and CNNs, the theoretical basis for combining RF and CNN is established. Next, an evaluation system for wind power prediction models that includes root mean square error (RMSE), determination coefficient, and Spearman correlation coefficient is introduced. Finally, validatinos are conducted using three open-source wind power datasets from European wind farms. The results demonstrate that, compared to other five models, the RF-CNN model outperforms in all three datasets, thus confirming the model's effectiveness and accuracy for wind power prediction.

Keywords: wind power prediction; random forest; convolutional neural network; hybrid model; error index

0 引言

近年来, 风电装机容量显著增长^[1]。然而, 风力发电的间歇性导致风机输出功率出现大幅波动, 给电网稳定性带来了巨大挑战。因此, 实现准确的风电功率预测对于保障电网的可靠、稳定运行

至关重要。

近年来, 各种风电功率预测模型不断涌现, 包括利用历史数据模式的统计模型、能够学习复杂关系的机器学习模型和协同组合多种技术的混合模型。直接预测方法和间接预测方法是风电功率预测常用的两种方法。直接预测方法是指基于对风电机

组历史数据、气象数据等的分析,直接预测未来的风电功率;间接预测方法首先预测风速、风向等,然后基于这些预测数据,根据风力发电机的功率曲线或其他功率模型来估计风电功率。ZHOU Bowen等利用历史风电功率时间序列数据来直接预测未来输出功率^[2];苏向敬等运用长短期记忆(long short term memory, LSTM)神经网络对风力发电的时间依赖性进行建模,取得了较好的预测准确率,但是该模型在处理风力突然变化和间歇性方面存在局限,这是短期预测中的常见难题,影响了其在极端天气事件中或风速突然波动期间的适应性^[3];SHI Jie等提出的混合模型结合小波变换和支持向量回归(support vector regression, SVR),在捕捉数据中的细微模式方面展现出潜力,但该模型在快速变化的风模式下表现不佳,限制了其实时预测性能^[4]。与此同时,间接预测方法也备受关注。胡锐等将风速预测作为估算功率的前置步骤,并在集成模型中采用数据驱动方法,利用历史风速数据和大气预测因子,证明了该方法在预测风速和将预测数据转换为输出功率方面的有效性,但是该方法在精确捕捉局部风现象和适应突然的小气候变化方面存在不足,特别是在地形复杂或天气模式快速变化的地区^[5]。高晨等强调通过混合模型弥补直接和间接方法各自不足的必要性,虽然提出了整合两种方法预测能力的建议,但是缺乏解决短期风电功率预测动态性的具体方法,使模型对意外的风行为和复杂大气条件的适应性存在不确定性^[6]。

作为学习影响风电功率输出的各种变量之间复杂非线性关系的智能工具,人工智能方法是现有风电功率预测方法的研究热点。张玉典等提出一种将小波变换与人工智能技术相结合的风电功率预测混合模型,该模型实现了更高的预测准确性和稳定性^[7]。何厚桦等提出一种将季节性分解与集合经验模态分解(ensemble empirical mode decomposition, EEMD)和LSTM神经网络相结合的混合模型,通过有效捕捉风电数据的季节性和长期依赖性提高了模型预测的准确性^[8]。JIANG Ping等提出一种将神经网络与模糊时间序列相结合的混合模型,该模型有效地捕捉了风速和风电功率之间的非线性关系,提升了预测准确性^[9]。

然而,上述研究仍存在不足:现有模型无法有效捕捉数据中固有的复杂时空关系,导致预测不准确;现有模型对训练数据的变化较为敏感,容易导致过拟合和泛化性能下降。鉴于此,本文提出一种

新型的随机森林(random forest, RF)-卷积神经网络(convolutional neural network, CNN)混合模型,通过引入RF集成方法减少方差并提升模型的整体泛化性能。

本文首先介绍RF、CNN的组成结构及混合模型RF-CNN的计算流程,然后基于3个风电场的真实风电功率数据来验证所提RF-CNN模型的有效性,构建包含方均根误差(root mean square error, RMSE)、决定系数和Spearman相关系数的风电功率预测模型评估指标体系,最后将RF-CNN模型与5种现有风电功率预测模型进行对比,以证明RF-CNN模型的高效性和准确性。

1 随机森林-卷积神经网络模型

1.1 随机森林

RF是一种强大的集成学习技术,通过组合多个决策树模型的预测结果来提高整体预测的准确性和稳健性^[10]。RF在机器学习和数据挖掘领域应用广泛,包括分类、回归和预测等。RF采用“袋装法”集成学习方法。“袋装法”的核心思想是从原始训练数据集中随机抽取多个子集用作训练集,每次抽取是有放回的,即允许重复抽样,袋外数据(未选中的数据)可以作为验证集,保证了风电时间序列数据的充分利用。在构建每一棵决策树时,随机森林还会对特征进行随机选择。每次决策树分裂时,不是考虑所有特征,而是从特征子集中随机选取部分特征进行分裂。这一机制进一步增加了模型的多样性,从而增强了泛化能力。

利用随机森林进行数据预测的步骤如图1所示,具体如下。

1) 数据准备:包含输入特征和目标变量的标签训练数据集。

2) 基模型创建:基模型选择为决策树。从训练数据集中有放回地随机抽取多个子集,用于训练每一棵决策树。每棵决策树的训练集中大部分数据来自原始训练集,由于是有放回的抽样,因此每棵树的训练集通常包含重复数据,且会遗漏一些数据。

3) 训练:从原始训练数据集中随机抽取多个子集(每个子集的大小与原始数据集相同,但允许重复抽样),然后一个子集训练一棵决策树。决策树基于选定特征的某个分裂点进行训练,直到满足停止条件,如达到最大深度、节点中的样本数小于某个阈值或无法继续分裂。

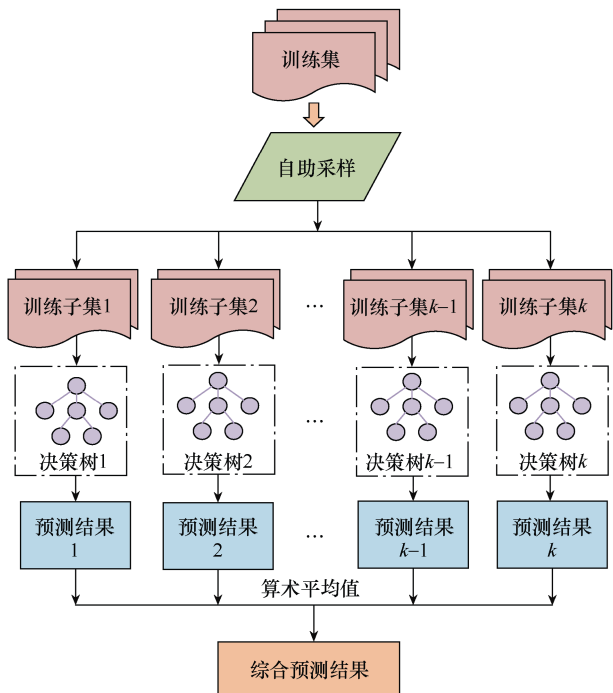


图 1 随机森林进行数据预测步骤

4) 预测聚合: 使用投票(分类)或平均(回归)方案聚合所有训练决策树的预测结果。在分类任务

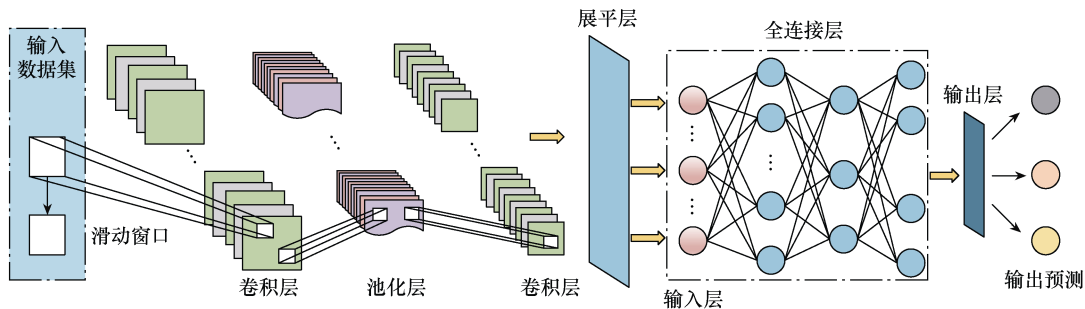


图 2 卷积神经网络架构

1.3 随机森林-卷积神经网络

图 3 所示为本文所提 RF-CNN 模型框架。从图 3 可以看出, 本文将 RF 中的决策树基模型更换为 CNN 模型, 利用 RF 的“袋装法”集成学习方法, 通过自助随机选取数据样本, 结合 CNN 多层滤波器的深入提取, 模型能够逐渐学习到更抽象和复杂的时空信息, 从而提高集成性能。最后, 对多个基模型的结果求取算术平均后, 聚合得到最终预测结果。

在 RF 集成中选择 CNN 模型而非决策树作为基模型具有显著优势, 具体如下:

1) CNN 在处理复杂的数据结构方面表现优异, 特别是那些具有时空属性的数据结构, 这在风电功率预测等应用中较为常见, 是决策树可能难以企及的。

中, 决策树中的多数投票决定最终的预测类标签。在回归任务中, 最终预测通过取决策树预测值的算术平均值获得。

1.2 卷积神经网络

CNN 算法遵循特定的训练和预测方法^[11-13]。CNN 的组成架构如图 2 所示。首先进行数据预处理, 包括图像归一化、调整大小和增强技术等任务, 提高数据集的多样性和稳健性。然后, 将输入数据输入卷积层, 提取有意义的特征。每个卷积层由多个可学习的滤波器组成, 滤波器对输入数据进行卷积操作, 捕捉空间模式并检测相关特征。在每个卷积层后应用非线性激活函数, 如修正线性单元, 以引入非线性并增强模型的表达能力。此外, 池化层, 如最大池化或平均池化, 对特征图进行下采样, 降低计算复杂度并提高平移不变性。卷积层的输出被展平并连接到一个或多个全连接 (full-connected, FC) 层, 这些层通过反向传播和梯度下降技术学习所提取的特征与目标输出之间的复杂关系, 最后的全连接层根据具体任务使用适当的激活函数将可用特征映射到期望的结果, 例如多分类使用 softmax 激活函数、回归使用线性激活函数。

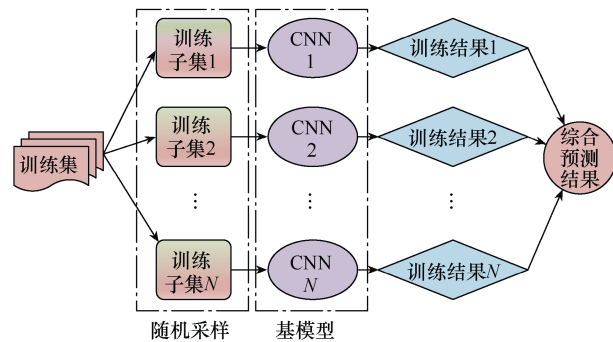


图 3 RF-CNN 模型框架

2) CNN 具备特征学习能力, 能够自主识别原始数据中有意义的特征。这种特征提取过程减少了对手动特征工程的依赖, 在处理时空数据时是一个

重要优势。相比之下,决策树通常需要精心的特征提取才能达到类似效果。

3) CNN 在面对噪声和数据波动时表现出较强的稳健性。RF 算法通过在自助数据样本上训练多个模型来减轻过拟合,它与 CNN 相辅相成,使 CNN 在面对噪声数据时也能进行泛化。这种稳健性增强了 CNN 集成到 RF 中的适用性。

4) CNN 擅长从数据中提取层次特征。随着深入 CNN 的层,模型能够逐渐学习到更抽象和复杂的信息。在将 CNN 与 RF 结合时,可以利用这种层次特征提取能力提高集成性能。

1.4 决策树与 CNN 的类别有效性

前文的分析明确了 CNN 相较于决策树作为基模型的优势所在。但是,RF 模型与 CNN 模型能否匹配共同组成 RF-CNN 混合模型是本文研究的关键。因此,本文从决策树与 CNN 的类别有效性角度,分析 RF-CNN 混合模型的可行性^[14]。决策树与 CNN 的关联关系如图 4 所示。

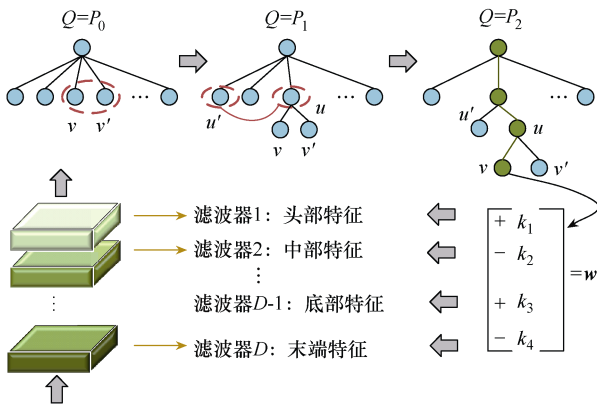


图 4 决策树与 CNN 的关联关系

1) CNN 预测原理。给定输入图像 I , 设 $\mathbf{x} \in \mathbf{R}^{L \times L \times D}$ 表示经过线性整流函数 (rectified linear unit, ReLU) 运算后最上层卷积层的特征映射, 其中 L 为特征映射的尺度, D 为滤波器编号。设 y 表示在 softmax 激活函数操作前某一类别的分类得分, 则任务为用 \mathbf{x} 来表示 y 的基本原理。这里可以使用分段线性表示等效级联的 FC 层和 ReLU 层的函数, 有

$$y = f_{\text{fc-n}}(f_{\text{relu}}(\cdots f_{\text{fc-1}}(\mathbf{x}))) = \sum_{h,w,d} g^{(h,w,d)} x^{(h,w,d)} + b \quad (1)$$

式中: $x^{(h,w,d)} \in \mathbf{R}$ 为第 d 个通道位置 (h,w) 处的元素; $g^{(h,w,d)}$ 为描述 $x^{(h,w,d)}$ 对预测 I 的重要性的权重; $f_{\text{fc-n}}$ 和 f_{relu} 分别为第 n 号 FC 层滤波器和 ReLU 层滤波器; b 为学习参数, $b = y - \mathbf{g} \otimes \mathbf{x}$, \mathbf{g} 为权重向量, \otimes

为卷积。

然后, 考虑利用 \mathbf{x} 和 \mathbf{g} 来表示全连接层的预测原理, 有

$$y \approx \mathbf{g}^T \mathbf{x} + b \quad (2)$$

向量 \mathbf{x} 的不同维度测量不同对象部分的标量信号强度, 因为滤波器可能代表某个对象部分; \mathbf{g} 对应于 CNN 预测对象部分的选择。

2) 学习决策树。与层次聚类一样, 学习决策树的基本思想是从不同图像的特定决策模式中总结出共同的通用决策模式。首先, 设置 $\bar{\mathbf{g}} = \mathbf{g}_i$ 和 $\alpha = 1$, 其中 $\bar{\mathbf{g}}$ 为单位向量, 反映图像共享的原理, \mathbf{g}_i 为第 i 个正图像对于预测的权重向量, α 为滤波器的二进制选择系数。将每个正图像 (某一关注类别的图像, 其他图像为负图像) I_i 的决策模式 \mathbf{g}_i 初始化为一个叶节点。因此, 本文构建如图 4 所示的初始树 Q , 其中根节点将所有正图像的决策模式作为子节点。在每一步中, 选择和合并第二树层中的两个节点 $v, v' \in V$ (即根节点的子节点), 以获得一个新的节点 u , 其中 V 表示根的子集。 u 成为根节点的新子节点, v 和 v' 被重新指定为 u 的子节点。这样, 逐步将初始树 $Q = P_0$ 向 T 次归并后的最终树修正为

$$Q = P_0 \rightarrow P_1 \rightarrow P_2 \rightarrow \cdots \rightarrow P_T = \hat{P} \quad (3)$$

在第 t 步中, 合并 P_{t-1} 的第二树层中的两个节点 $v, v' \in V$, 得到一个新的节点 u , 从而得到一个新的树 P_t 。因此, 通过贪婪策略来学习决策树。

3) CNN 的类比解释。给定测试图像 I_i , CNN 做出的预测标记为 y_i 。在推理过程中, 以自顶向下的方式从根节点开始推断解析树。图 4 中显示了一个解析树。当选择节点 u 中的决策模式作为基本原理时, 可以进一步选择与最具体的基本原理 \mathbf{g}_i 兼容最大的子模型 v 作为更精细的模式, 有

$$\bar{v} = \arg \max_{v \in C(u)} \cos(\mathbf{g}_i, \mathbf{w}_v) \quad (4)$$

式中: \bar{v} 为 u 的子节点; \mathbf{w}_v 为子节点 v 决策模式的基本原理; $C(u)$ 为节点 u 的子节点集合。

解析树中的节点 v 提供了在特定细粒度级别上对图像 I_i 进行预测的基本原理。通过计算向量 \mathbf{J}_i 和 \mathbf{K}_i 来评估不同滤波器和不同目标部分的贡献。

$$\begin{cases} \mathbf{J}_i = \mathbf{w}_v \circ \mathbf{x}_i \\ \mathbf{K}_i = \mathbf{A} \mathbf{J}_i \end{cases} \quad (5)$$

式中: \circ 为元素乘法; \mathbf{A} 为卷积层元素矩阵; \mathbf{J}_i 为不

同滤波器的预测贡献； K_i 为不同对象部分的预测贡献。 $J_i^{(d)}$ 为第 d 个滤波器对 CNN 预测的贡献。如果 $J_i^{(d)} > 0$ ，则第 d 个滤波器的贡献为正；如果 $J_i^{(d)} < 0$ ，则第 d 个滤波器的贡献为负。

2 误差评估指标

本文选择 3 个指标来评估风电功率预测模型的性能：方均根误差 e_{RMSE} 、决定系数 c_{r2score} 和 Spearman 相关系数 c_{LRxy} 。

1) 方均根误差 e_{RMSE}

RMSE 给出估计值和原始数据之间的总体一致性，避免了误差补偿，量化了预测值与实际值之间的误差程度，计算公式为

$$e_{\text{RMSE}} = \sqrt{\frac{1}{m} \sum_{i=1}^m (p - p_e)^2} \quad (6)$$

式中： p 为原始功率； p_e 为估计功率； m 为数据序列数目。

2) 决定系数 c_{r2score}

决定系数用来表示模型拟合数据的程度，反映因变量的全部变动能通过回归关系被自变量解释的比例，计算式为

$$c_{\text{r2score}} = 1 - \frac{\sum |p - p_e|}{\sum |p - p_0|} \quad (7)$$

式中， p_0 为平均值。

3) Spearman 相关系数 c_{LRxy}

Spearman 相关系数是一种用于衡量两个变量之间单调关系（即变量之间是否有一致的增减趋势）的统计量。当 $c_{\text{LRxy}}=1$ 时表示完全正相关，两个变量的秩完全一致，随着一个变量的增加，另一个变量也不断增加；当 $c_{\text{LRxy}}=-1$ 时表示完全负相关，随着一个变量的增加，另一个变量不断减少；当 $c_{\text{LRxy}}=0$ 时表示没有单调关系，两个变量之间没有显著的单调趋势。具体计算公式为

$$\begin{cases} D_x = \frac{1}{N} \sum_{k=0}^{N-1} \left(X(k) - \frac{1}{N} \sum_{k=0}^{N-1} X(k) \right)^2 \\ C_{xy} = \frac{1}{N} \sum_{k=0}^{N-1} \left(X(k) - \frac{1}{N} \sum_{k=0}^{N-1} X(k) \right) \times \left(Y(k) - \frac{1}{N} \sum_{k=0}^{N-1} Y(k) \right) \\ c_{\text{LRxy}} = \frac{C_{xy}}{\sqrt{D_x D_y}} \end{cases} \quad (8)$$

式中： D_x 、 D_y 分别为两个序列的方差； $X(k)$ 、 $Y(k)$

分别为长度为 N 的风电功率实测幅值序列、预测幅值序列； C_{xy} 为两个序列的协方差。

3 预测模型验证

3.1 数据集选取

本文获取了欧洲地区 3 个开源数据集来训练和测试所提风电功率预测混合模型^[15]，数据集包含实际发电功率和气象数据（风速、风向、温度、气压、湿度），因此模型的输入数据集为

$$F_w(t) = [W_S \ W_D \ W_T \ W_A \ W_H \ W_P] \quad (9)$$

$$\begin{cases} W_S = [W_S(t) \ W_S(t-1) \ W_S(t-2) \ \dots \ W_S(t-24)]^T \\ W_D = [W_D(t) \ W_D(t-1) \ W_D(t-2) \ \dots \ W_D(t-24)]^T \\ W_T = [W_T(t) \ W_T(t-1) \ W_T(t-2) \ \dots \ W_T(t-24)]^T \\ W_A = [W_A(t) \ W_A(t-1) \ W_A(t-2) \ \dots \ W_A(t-24)]^T \\ W_H = [W_H(t) \ W_H(t-1) \ W_H(t-2) \ \dots \ W_H(t-24)]^T \\ W_P = [W_P(t-1) \ W_P(t-2) \ W_P(t-3) \ \dots \ W_P(t-24)]^T \end{cases} \quad (10)$$

式中， W_S 、 W_D 、 W_T 、 W_A 、 W_H 、 W_P 分别为输入以 t 为时间间隔的风速、风向、温度、气压、湿度和风电功率特征。 t 一般以 h 为单位，时间跨度至过去 24h。数据的 70% 作为训练集，30% 作为测试集。

3.2 模型参数设置

为确定最优参数，利用网格搜索交叉验证（grid search cross validation, GridSearchCV）技术分析不同的参数组合，并根据不同数据集确定 RF-CNN 模型的最佳配置，模型最优参数见表 1，其中 RF 基模型的数量设定为 10、12，训练周期数量为 10、15，批量大小为 30、64；CNN 的滤波器数量设定为 510，核大小为 3，密集层数量为 64，激活函数为 ReLU，其他参数均采用默认设置，具体参数组合见表 1。GridSearchCV 技术可以详尽地搜索指定的参数网格，并评估每个组合的模型性能，这不是本文重点，此处不再详细分析，具体参看文献[16-18]。

表 1 RF-CNN 模型最优参数

模型优化参数	输入集	优化参数数值		
		数据集 1	数据集 2	数据集 3
基模型数量	10, 12	10	10	12
训练周期	10, 15	15	10	10
批量大小	30, 64	30	64	64

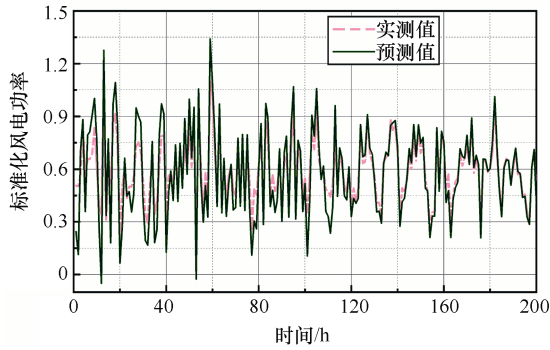
3.3 风电功率预测结果

RF-CNN 模型风电功率预测值与实测值对比如

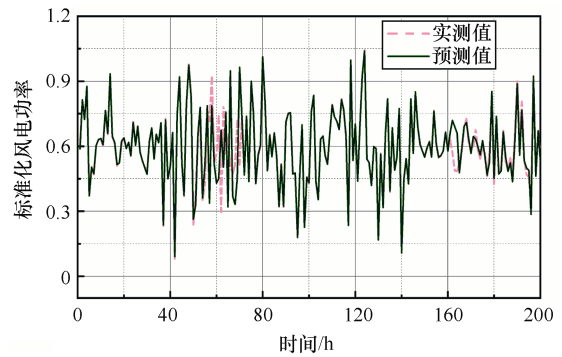
图 5 所示, 图中功率数据均为归一化值。从图 5 可以看出, 预测值与实测值的相似性较高。除少数峰值处的模型预测值有偏差外, 所有训练周期范围内的预测值都是可接受的。

为定量分析模型预测值与实测值的偏差, 基于第 2 节的 3 种误差评估指标, 对比分析支持向量机 (support vector machine, SVM)^[19]、RF^[20]、CNN^[21] 3 种单一模型, 以及鲸鱼优化算法 (whale optimization algorithm, WOA)-LSTM^[22]、粒子群优化算法-支持向量回归 (particle swarm optimization-support vector regression, PSO-SVR)^[23-24] 2 种混合模型。不同模型的误差指标见表 2。

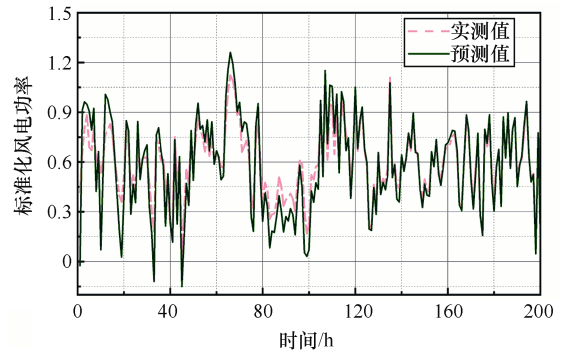
首先, 比较 RF-CNN 模型与 3 种单一模型。对于数据集 1、数据集 2 和数据集 3, RF-CNN 模型的 e_{RMSE} 均优于其他模型。以数据集 1 为例, RF-CNN



(a) 数据集 1



(b) 数据集 2



(c) 数据集 3

图 5 RF-CNN 模型风电功率预测值与实测值对比模型的 $c_{r2score}$ 比 SVM、RF、CNN 3 种单一模型分别提高了 74%、46%、46%; 对于数据集 2 和数据集 3, 也均有提升。以数据集 2 为例, RF-CNN 模型的 c_{LRxy} 相较于 3 种单一模型分别提高了 41%、45%、30%。

表 2 不同模型的误差指标

模型	数据集 1			数据集 2			数据集 3		
	e_{RMSE}	$c_{r2score}$	c_{LRxy}	e_{RMSE}	$c_{r2score}$	c_{LRxy}	e_{RMSE}	$c_{r2score}$	c_{LRxy}
RF-CNN	0.124 4	0.894 3	0.913 4	0.108 5	0.891 2	0.915 0	0.125 1	0.905 0	0.952 7
SVM	0.164 7	0.513 9	0.745 2	0.167 4	0.510 5	0.648 9	0.162 9	0.515 4	0.744 6
RF	0.181 4	0.614 3	0.734 6	0.182 2	0.617 8	0.633 1	0.182 1	0.615 2	0.735 3
CNN	0.143 0	0.612 1	0.805 6	0.142 5	0.610 7	0.703 4	0.142 7	0.613 0	0.804 4
WOA-LSTM	0.1387	0.724 7	0.841 5	0.128 2	0.725 4	0.740 3	0.137 8	0.623 2	0.712 3
PSO-SVR	0.131 8	0.703 4	0.816 3	0.150 5	0.681 2	0.824 7	0.130 2	0.702 0	0.844 9

其次, 比较 RF-CNN 模型与其他 2 种混合模型。对于数据集 1、数据集 2 和数据集 3, RF-CNN 模型的 $c_{r2score}$ 比 WOA-LSTM 模型分别提高了 23%、23% 和 45%, RF-CNN 模型的 c_{LRxy} 比 WOA-LSTM 模型分别提高了 9%、24% 和 34%。对于数据集 1、数据集 2 和数据集 3, RF-CNN 模型的 $c_{r2score}$ 比 PSO-SVR 模型分别提高了 27%、31% 和 29%, RF-CNN 模型的 c_{LRxy} 比 PSO-SVR 模型分别提高了 12%、11% 和

13%。针对 e_{RMSE} 指标, RF-CNN 模型与其他 2 个混合模型的偏差不大, 但均保持正向改进, 其中相较于 PSO-SVR 模型最高减小了 28%。

此外, 本文还对比了各个模型的算法复杂度, 利用时间成本定量表征。为防止其他因素干扰, 设备均使用同一台计算终端, 且采用重复测试消除不确定性干扰, 测试 5 次。经计算, SVM、RF 和 CNN 3 个单一模型的时间成本均在 1min 内, 其中 SVM

的时间成本最小,为30s; RF-CNN、WOA-LSTM和 PSO-SVR 3个混合模型较单一模型耗时长,但均<2min。综合来看,各模型时间成本差别很小,所有模型的时间成本均在可接受范围内。

综上所述,由于上述单一模型和其他混合模型会过拟合于某些特定的数据特征,导致误差较大,而 RF-CNN 模型结合机器学习及多个深度学习模型,并引入随机性,通过多种结果得出最终预测结果,因此就本文的3种误差指标来看,RF-CNN 模型最优。

4 结论

本文提出了一种结合随机森林和卷积神经网络优势的新算法,用于200h内的风电功率预测。首先通过CNN与决策树的类比有效性论述了RF与CNN组合的理论可行性,然后基于RF的集成技术与随机属性,与CNN的特征提取能力相结合,利用本文的误差指标体系,与其他风电功率预测模型进行对比,验证了本文所提模型的有效性,得到如下结论:

1) RF-CNN模型比现有常用的风电功率预测模型具有更优越的性能。基于本文的误差指标体系,与其他模型相比,RF-CNN模型的RMSE值始终保持最低水平;以数据集1为例,RF-CNN模型的决定系数比3种单一模型分别提高了74%、46%、46%,对于数据集2和数据集3,也均有提升;与WOA-LSTM和PSO-SVR混合模型相比,针对决定系数指标,对于数据集1、数据集2和数据集3均提高了至少23%。

2) 总体而言,本文工作证明了所提 RF-CNN模型在风电功率预测中的有效性和优越性。研究结果有助于推进风电功率预测技术的应用,实现风能可靠、高效地并网。

参考文献

- [1] 王小明,徐斌,尹元亚,等.基于空间相关性与Stacking集成学习的风电功率预测方法[J].电力工程技术,2024,43(5):224-232.
- [2] ZHOU Bowen, MA Xiangjin, LUO Yanhong, et al. Wind power prediction based on LSTM networks and nonparametric kernel density estimation[J]. IEEE Access, 2019, 7: 165279-165292.
- [3] 苏向敬,程子凡,聂良钊,等.基于AGCN-LSTM模型的海上风电场功率概率预测[J].电力系统自动化,2024,48(22):140-149.
- [4] LIU Yongqian, SHI Jie, YANG Yongping, et al. Short-term wind-power prediction based on wavelet transform-support vector machine and statistic-characteristics analysis[J]. IEEE Transactions on Industry Applications, 2012, 48(4): 1136-1141.
- [5] 胡锐,乔加飞,李永华,等.基于WOA-VMD-SSA-LSTM的中长期风电预测[J].太阳能学报,2024,45(9):549-556.
- [6] 高晨,赵勇,汪德良,等.海上风电机组电气设备状态检修技术研究现状与展望[J].电工技术学报,2022,37(增刊1):30-42.
- [7] 张玉典,王宇驰.离散小波变换结合BP神经网络的光伏发电功率预测[J].现代工业经济和信息化,2024,14(12):262-263,266.
- [8] 何厚桦,王仲平,崔萌.基于EEMD-LSTM-WOA的风速预测混合模型[J].应用数学进展,2024,13(10):4486-4497.
- [9] JIANG Ping, YANG Hufang, HENG Jiani. A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting[J]. Applied Energy, 2019, 235: 786-801.
- [10] 董志文,苏晶晶.基于变分模态分解能量熵混合时域特征和随机森林的故障电弧检测方法[J].电气技术,2024,25(1):1-7.
- [11] 衣思彤,刘雅浓,马耀滢,等.基于贝叶斯优化-卷积神经网络-双向长短期记忆神经网络的锂电池健康状态评估[J].电气技术,2024,25(5):1-10,21.
- [12] 王立辉,柯泳,苏如开.基于1DCNN-BiLSTM-BiGRU的电能质量扰动分类方法[J].电气技术,2024,25(5):51-56,64.
- [13] 涂彦昭,高伟,杨耿杰.一种基于卷积神经网络和长短期记忆神经网络的光伏系统故障辨识方法[J].电气技术,2022,23(2):48-54.
- [14] ZHANG Quanshi, YANG Yu, MA Haotian, et al. Interpreting CNNs via decision trees[EB/OL]. <https://arxiv.org/abs/1802.00121v2>.
- [15] ZAMEER A, ARSHAD J, KHAN A, et al. Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks[J]. Energy Conversion and Management, 2017, 134: 361-372.
- [16] 王昭栋,王自法,李兆焱,等.基于机器学习-网格搜索优化的砂土液化预测[J].振动与冲击,2024,43(5):82-93.

支路,通过控制开关元件实现电容放电功率的保持,提高了放电效率。

3) 放电回路参数对放电速率有较大影响,实际应用时应根据设备、所采用元件的性能选择放电回路元件,以使放电速率达到最佳。

4) 本文针对高压电容放电的研究结果可为电力生产、新能源汽车、超级电容储能等相关领域提供参考。

参考文献

- [1] 谢翔杰,魏桃胜,贺星昶,等. 变电站蓄电池开路故障定位及集中式续流方法[J]. 电气技术, 2024, 25(3): 63-67, 73.
- [2] 陈华文. 一次电路电容放电实验方法浅析[J]. 日用电器, 2014(7): 25-27, 33.
- [3] 殷昂. 折叠式低压放电棒的研制与应用[J]. 山东煤炭科技, 2016(5): 126-127.
- [4] 高惠新,周刚,尚浩志,等. 现场放电棒及试验线检测装置的研制[C]//2016 年第二届今日财富论坛论文集,北京,2016.
- [5] 金雍奥,曲金秋,周志昊,等. 变电设备高压试验安全防护措施[J]. 电气技术, 2015, 16(5): 85-88, 94.
- [6] 中国南方电网有限责任公司电力安全工作规程: Q/CSG 51001—2015[S].
- [7] 李杨,顾志斌,陈浩,等. 电容放电装置研究[J]. 电力电子技术, 2022, 56(2): 1-3.
- [8] 徐旭东,汪源,彭庆辉,等. 电力互感器室内检定自放电装置的研制[J]. 科技创新与应用, 2019(4): 9-10.
- [9] 张采龙. 岸边集装箱起重机驱动器电容放电装置设计[J]. 港口装卸, 2024(1): 31-32, 70.
- [10] 张晓军,杨家强,周宇晨. 计及总损耗功率的电动汽车母线电容主动快速放电方法[J]. 电工技术学报, 2024, 39(6): 1737-1748.
- [11] 林耿明. 一种新型 35kV 放电棒[J]. 电子测试, 2022, 36(11): 109-111.
- [12] 左涛. 一起化工厂 10kV 开关柜过电压保护器热崩溃事故分析[J]. 电气技术, 2024, 25(2): 68-73.
- [13] 赵永红. 本安电路电容放电模型与仿真分析[J]. 测试技术学报, 2021, 35(5): 392-397.
- [14] 王胜利,李阳,王亚杰. 电容器声光报警专用放电棒的研制与应用[C]//2023 年电力行业技术监督工作交流会暨专业技术论坛论文集(上册), 南宁, 2023.
- [15] 刘应敏. 试析 RC 和 RL 电路充、放电过程中能量的变化关系[J]. 焦作师范高等专科学校学报, 2008(1): 74-75.
- [16] 赵强,孙殿平,郭超修. 用电路模型等效变换的方法教学 RL 暂态过程[J]. 大学物理, 2017, 36(7): 56-58.
- [17] 朱小全,侯鹏辉,侯锦涛,等. 带有 LC 串联谐振功率自均衡单元的模块化光伏直流升压变换器[J]. 电工技术学报, 2024, 39(4): 1087-1102.
- [18] 李尊,张小科,张少峰,等. 基于改进 LSTM 的风电功率预测因素分析[J]. 计算机仿真, 2024, 41(10): 63-68, 143.
- [19] 徐炜君,裴欢,魏勇. 基于粒子群优化的 SVR 风电功率超短期预测[J]. 东北师大学报(自然科学版), 2017, 49(1): 73-77.
- [20] 赵倩,陈芳芳,甘露. 基于改进粒子群算法优化支持向量机的风电功率预测[J]. 电气技术, 2020, 21(12): 12-16.
- [21] 刘坤. 基于随机森林的风力发电系统输出功率预测方法[J]. 光源与照明, 2022(7): 165-167.
- [22] 张维通,闫正兵,张正江,等. 基于物理信息的时间卷积神经网络风电功率预测[J]. 计算机测量与控制, 2024, 32(11): 101-108, 117.

(上接第 33 页)

- [17] 陈光宇,袁文辉,徐晓春,等. 基于残差图卷积深度网络的电网无功储备需求快速计算方法[J]. 电工技术学报, 2023, 38(17): 4683-4700.
- [18] 刘笑,杨建,李力,等. 基于机器学习的带被动阻尼直流微电网系统的稳定性检测[J]. 电工技术学报, 2024, 39(8): 2281-2293, 2324.
- [19] 刘金朋,邓嘉明,高鹏宇,等. 基于 VMD-IMPACT-SVM 的超短期风电功率预测[J]. 智慧电力, 2024, 52(7): 24-31, 79.
- [20] 刘坤. 基于随机森林的风力发电系统输出功率预测方法[J]. 光源与照明, 2022(7): 165-167.
- [21] 张维通,闫正兵,张正江,等. 基于物理信息的时间卷积神经网络风电功率预测[J]. 计算机测量与控制, 2024, 32(11): 101-108, 117.

收稿日期: 2024-11-25

修回日期: 2025-01-02

作者简介

蔡伟杰(1997—),男,广东潮州人,学士,助理工程师,从事电力设备高压试验工作。

收稿日期: 2025-01-20

修回日期: 2025-02-27

作者简介

李 桓(1994—),男,山西省大同市人,硕士,中级工程师,主要从事新能源发电及负荷预测研究工作。