Paper ID: 534

XAI-SP: Explainable AI Approach for SPF using CNN-LSTM-Attention Model

Md Rasel Sarkar [1], Sreenatha G. Anavatti[1], Md Meftahul Ferdaus[2] and Tanmoy Dam[3]
[1] School of Engineering & Technology University of New South Wales, Canberra, Australia
[2]Department of Computer Science, University of New Orleans, New Orleans, USA
[3]Department of Biomedical Engineering, Emory University, Atlanta GA, USA
E-mail: [1]raselbdeee@gmail.com, [1]agsrenat@adfa.edu.au, [2]mferdaus@uno.edu, [3]tanmoy.dam@yahoo.com

Abstract

Accurate solar power forecasting (SPF) faces challenges due to the complex nature of solar power generation data. While deep learning (DL) techniques like convolutional neural networks (CNNs) and and long-short-term memory (LSTM) show promise, their lack of transparency creates trust issues. We propose a hybrid model (CNN-LSTM-Attention) combining CNNs for spatial features, LSTMs for temporal dependencies, and attention mechanisms for key time steps. To enhance transparency, we employ feature ablation through explainable artificial intelligence (XAI) solar power (XAI-SP) analysis to identify critical prediction factors. Testing on real-world data from two solar stations shows the hybrid model improves the accuracy over baseline approaches. The analysis shows total solar irradiance and direct normal irradiance as crucial features. These perceptions can guide data collection strategies and feature selection, ultimately improving solar energy system integration and management.

Keywords: Solar power forecasting, CNN-LSTM-Attention mechanism, Explainable AI, and Feature ablation.

1. Introduction

Solar power (SP) represents a viable and promising solution in the pursuit of a sustainable energy future, as it harnesses the renewable energy derived from the sun's radiant emissions. A key advantage of SP is its capability to produce electrical energy without releasing greenhouse gas emissions, in contrast to conventional fossil fuel-based generation methods [1]. Ac- curate prediction of SP is crucial for the efficient integration and maximum utilization of this intermittent renewable energy source. SPF involves determining the level of solar radiation expected at a specific location and time, considering different atmospheric conditions and weather patterns. This information is essential for grid operators and utility companies to maintain balanced supply and demand and ensure grid reliability [2]. There are some commonly used statistical time series forecasting (TSF) namely, autoregressive (AR) [3], moving average (MA) [4], and the combined autoregressive integrated moving average (ARIMA) models [5] which are fitted individually to each time series dataset. Nonetheless, statistical approaches deal with several key limitations: diminished accuracy over longer forecasting horizons, and inadequate handling of seasonal dependencies. To overcome these drawbacks, DL networks [6] can learn complex patterns from data. Recurrent neural networks (RNNs) [7], are particularly adept at sequential time series data in an autoregressive manner. RNN-based networks such as LSTM [8] and gated recurrent units (GRU) [9] can effectively capture long-range temporal dependencies from historical sequences. To overcome issues like vanishing/exploding gradients faced by conventional RNNs, these models enable more stable training and knowledge extraction over long periods. However, processing data sequentially limits parallel training capabilities and the range of dependencies that RNNs can develop. Attention mechanisms [10] have been integrated into neural architectures to selectively focus on the most relevant parts of the input sequence when making predictions and overcome the challenges of the RNN-based networks. Despite their robust predictive capabilities, DL models often shortage of significance, making it difficult to explain their predictions and understand which features are most influential. XAI techniques [11], such as feature ablation studies [12], can assist shed light on these "black box" models by thoroughly removing or distressing input features to compute their impact on model outputs. Such XAI methods can provide valuable visions into the key drivers of SPF, promoting model explanation, and feature selection, and possibly improving prediction accuracy. By finding the most relevant input variables, feature ablation allows [13] for a better understanding of the model's decision-making process, increasing precision and trust in the predictions. Moreover, this study can reduce the effort of more feature collection for the prediction model inputs.

Consequently, FA has the potential to enhance the reliability of DL prediction of SP, an area of critical importance for effective renewable energy management. There are several notable studies in the area of SPF below.
Gigoni, L [14] carried out research that compared different machine-learning techniques for forecasting PV power output. Their evaluation included the Gray-Box model, shallow artificial neural networks (ANNs), support vector machines (SVMs), K-nearest neighbors (KNN) regression, quantile random forests, and ensemble learning approaches. They analyzed data from 32 PV plants while considering the impact of weather conditions. The findings revealed that the hourly predictions on the test dataset showed minimal discrepancies across the various models examined. Huang et al. [15] implemented a three-layer CNN for day-ahead hourly PV output forecasting. CNN was trained on hourly PV power and weather data from the previous five days. Their approach demonstrated high prediction accuracy and robustness. In addition, Zang et al. [16] employed Variational Mode Decomposition (VMD) to derive nine new time-series features from historical PV power output data. These features were used to enhance the one-hour-ahead prediction performance of a two-dimensional CNN model when compared to single CNN, SVM, VMD-SVM, and VMD one-dimensional CNN models. In a subsequent study, Zang et al. [17] integrated residual networks (ResNets) into their previous model and applied meta-learning techniques to enable both point and probabilistic forecasting. However, their approach assumed prior knowledge of the day-ahead weather type, rendering the model's performance heavily reliant on the accuracy of a separate weather prediction model. RNNs, particularly the LSTM architecture proposed by Hochreiter and Schmidhuber [18], have gained significant traction among researchers for processing and forecasting long time-series data with temporal dependencies. The LSTM's ability to selectively remember and forget information through its gating mechanism allows it to effectively capture temporal correlations [19]. Utilizing this capability, Srivastava and Lessmann [20] employed LSTM models to forecast day-ahead Global Horizontal Irradiance (GHI) at 3-hour intervals across 21 locations, utilizing satellite data as input. In order to improve the prediction of day-ahead hourly PV power, Gao et al. [21] presented an LSTM model that integrated estimates of total daily power output from a Discrete Grey Model (DGM). Their method also distinguished between ideal and non-ideal weather conditions based on forecasted meteorological data for the next day. Meanwhile, Ospina et al. [22] developed a hybrid model combining LSTM and deep neural networks (DNNs), employing wavelet transform (WT) and statistical feature extraction techniques. Their model aimed to predict PV output at a 30-minute resolution for the next 12 hours. However, the authors acknowledged the need to improve the stability of their recursive multi-step forecasting approach. Peng et al. [23] proposed a methodology that applied complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to decompose solar radiation time-series data into 11 sub-modes. Subsequently, a bidirectional LSTM (BiLSTM) optimized by the sine cosine algorithm (SCA) was employed to predict the next hours. The study aims to develop a hybrid model for SPF where the input features are multi-variable and incorporate XAI techniques to understand the impacts of different input features in achieving accurate forecasts.

2. CNN-LSTM-Attention-based proposed method: XAI-SP

All The CNN-LSTM-Attention model is a hybrid architecture that has the strengths of each component to capture both local patterns and long-term dependencies in the time series data. The CNN layers are used to extract local features and capture short-term temporal dependencies in the input sequence [24]. The LSTM layers are then employed to model the long-term dependencies and temporal dynamics of the extracted features [25]. Finally, the attention mechanism is applied to selectively focus on the most relevant time steps and assign higher importance to them when making predictions. Let's denote the input time series as $X_t = [X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(M)}$ be the multivariable input data at time step t, and T is the length of the input sequence. CNN Layer: The CNN layer applies convolutional operations to the input sequence X_t . Let $W_c \in R^{k \times d \times f}$ be the convolutional kernel, where k is the kernel size, d is the input feature dimension, and f is the number of filters. The convolutional operation is applied as follows [24]:

$$c_t = ReLU\ (W_c * X_{t:t+k-1} + b_c) \quad (1)$$

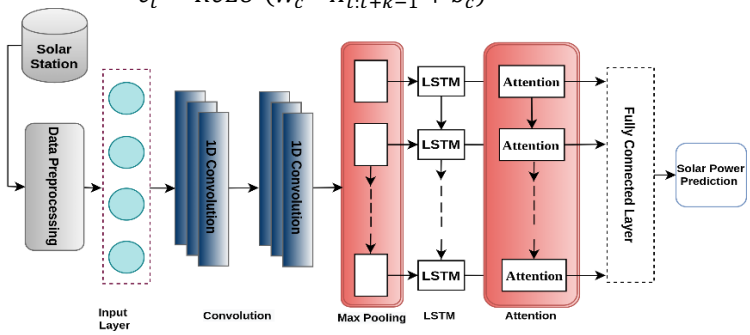


Fig. 1. Work-flow of CNN-LSTM-Attention Module in XAI-SP for SPF.

$c_t \in \mathbb{R}^f$ represents the output of the convolutional operation at time step t,$*$ denotes the convolution operation, and $b_c \in \mathbb{R}^f$ in is the bias term. The CNN layer outputs a sequence of feature maps $C = [c_1, c_2, \ldots., c_{T-k+1},]$.
STM Layer: The LSTM layer takes the CNN output C as input and models the long-term dependencies. The LSTM equations are as follows [25].

$$i_t = \sigma(W_i \cdot [h_{t-1}, c_t] + b_i) \quad (2)$$
$$f_t = \sigma(W_f \cdot [h_{t-1}, c_t] + b_f) \quad (3)$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, c_t] + b_o) \quad (4)$$
$$\tilde{C_t} = tanh(W_C \cdot [h_{t-1}, c_t] + b_C) \quad (5)$$
$$C_t = \sigma(f_t * C_{t-1} + i_t * \tilde{C_t}) \quad (6)$$
$$h_t = o_t * \tanh(C_t) \quad (7)$$

i_t, f_t, and o_t are the input gate, forget gate, and output gate, respectively. C_t is the cell state, and h_t is the hidden state at time step t. W_i, W_f, W_f and W_C are the weight matrices, and b_i, b_f, b_o and b_C are the bias terms. The LSTM layer outputs a sequence of hidden states $H = [h_1, h_2, \ldots., h_{T-k+1},]$.
Attention Layer: The attention layer computes attention weights to focus on relevant time steps. The attention mechanism is defined as follows [26]:

$$e_t = v^T \tanh(W_a h_t + b_a) \quad (8)$$
$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^{T-k+1} \exp(e_i)} \quad (9)$$
$$\tilde{h} = \sum_{i=1}^{T-k+1} \alpha_t h_t \quad (10)$$

e_t is the attention score for time step t, W_a and b_a are the attention weight matrix and bias term, respectively.
α_t is the normalized attention weight, and h_t is the context vector obtained by the weighted sum of hidden states.
Output Layer: The context vector $\tilde{h}$ is fed into a fully connected layer for final prediction:

$$\hat{y} = W_o \tilde{h} + b_o \quad (11)$$

$\hat{y}$ is the predicted value, W_o and b_o are the output weight matrix and bias term, respectively. The block diagram of hybrid model is shown in Fig. 1 which is trained by minimizing a suitable loss function, such as mean squared error (MSE), between the predicted values and the actual values. In addition, the output of SPF from the proposed model gets into the XAI-based feature ablation analysis to determine the features' effectiveness in the model.

3. Feature ablation analysis

While DL models like CNN-LSTM-Attention excel in capturing complex relationships within solar power data, understanding the contribution of individual features to the model's predictions remains crucial for interpretability and trust, which is why our model, XAI-SP, incorporates feature importance analysis via feature ablation [27] to evaluate the impact of each input feature on its forecasting accuracy. This methodology involves systematically removing or modifying specific features of the model and evaluating the resulting impact on its predictive performance, typically measured by the impact value of the feature. By conducting rigorous ablation experiments, researchers can gain valuable insights into the relative importance and roles of each component in the model's architecture. Through carefully designed ablation studies, researchers can determine the key components that significantly contribute to the model's forecasting accuracy and potentially optimize the architecture by removing less influential components. Hence, improving computational efficiency and model illustratable [28]. However, it is crucial to conduct feature ablation on diverse datasets representing different geographical locations and weather conditions to ensure the robustness and generalizability of the findings. Moreover, the importance of each component may vary depending on the specific characteristics of the solar power data and the forecast horizon.

4. Experiment

The purpose of this experimental study is to assess the effectiveness of a proposed model. The study examines two different DL networks: LSTM, and CNN-LSTM. These models are tested using real-world datasets. The models are implemented using the *PyTorch* library with Python 3.9 programming language. The computer hardware used for training the models consists of a single NVIDIA Corporation / Mesa Intel® UHD Graphics (TGL GT1) graphics processing unit. Firstly, the dataset was divided into three parts: a training set (70%), a test set (20%), and a validation set (10%). The proposed model, along with two baseline models, is trained on the training set and evaluated on the test and validation sets. After trial and error with different numbers of epochs, 100 epochs have been set for training the models. It is observed that after 100 epochs, there was no further change in the mean squared error (MSE), which measures the model's performance. Then, the parameters of the networks were set as follows: A one-dimensional convolutional layer was used with 64 filters, a kernel size of 3, a stride of 1, the '*RuLU*' activation function, and 'same' padding. A one-dimensional max pooling layer with a size of 2 was included. The LSTM layer had 64 units, and the return sequences parameter was set to True. The attention mechanism had a shape of 64 and used the *'softmax'* activation function. The output layer was a dense layer with

one unit. The same parameter settings were used for the other baseline models. Moreover, the proposed models were applied to two different datasets.

4.1 Datasets

SP generation data were gathered from a source that covers eight different locations [29]. Out of these eight sites, two datasets were selected for the experiment, specifically the data from solar station site 1 and solar station site 2. The data collected included information about the weather conditions as well as the actual power generation levels recorded in real-time. The data length is a two-year period from 2019 to 2020, with recorded every 15 minutes intervals. The solar generation capacity at these sites ranges from 50 megawatts (MW) to 130 MW. Every data point in the dataset consisted of six different variables. These parameters were recorded at regular intervals of 15 minutes. In total, the dataset contained 70,176 of these time-step measurements across the six parameters. Detailed statistical information about each solar station can be found in Table 1.

Table 1: Statistical Analysis of the Solar Power Station

Solar station	Statistics	Power output (MW)	Total solar irradiance (W/m²)	Direct normal irradiance (W/m²)	Global horizontal irradiance (W/m²)	Air temperature (°C)
Solar station site 1	Mean	9.7	266.4	93.3	67.7	13.1
	Minimum	0.0	0.0	0.0	0.0	0.0
	Maximum	48.3	1359	980	989	41.2
	Standard deviation	13.7	368	200.8	111.2	14.3
Solar station site 2	Mean	19.6	169.6	122.4	78.3	13.7
	Minimum	0.0	0.0	0.0	0.0	0.0
	Maximum	109.4	1041.9	751.8	561.8	40.5
	Standard deviation	28	148.4	179.2	117.6	12.1

5. Results and Discussion

SPF is crucial for optimizing energy storage operations, balancing supply and demand, reducing operational costs, and ensuring the stability and efficiency of the overall electrical grid with increasing solar power penetration. Multi-step SPF, which involves predicting solar power output over multiple future time steps, is effective for solar energy management and grid integration. The 24 steps have been considered in the experiment. Fig. 2 and Fig. 3 represent the SPF with different models in terms of solar stations 1 and 2 respectively. From Fig. 2, the predicted line of the proposed hybrid model is near the actual value as compared to baseline models. Similarly, predicted and actual SPF are very close to each other shown in Fig. 3. It can be concluded that the proposed hybrid model is enhancing the performance in comparison to the baseline models.

Table 2 presents an evaluation of the performance of the model for SPF at two different solar stations. The performance indicators considered mean absolute error (MAE), MAPE, root mean square error (RMSE), and co-efficient of determination (R^2) for the accuracy of the proposed hybrid compared to the baseline models. For Solar Station 1, hybrid networks displayed the best performance, with the lowest MAE of 0.053, MAPE of 0.164, and RMSE of 0.119, accompanied by the highest R^2 value of 0.93. It can be concluded that the proposed network achieved promising accuracy for Solar Station 1. The LSTM model showed the least satisfactory performance in terms of the performance indicators. For Solar Station 2, the proposed network shows a better performance in terms of the lowest MAE of 0.046, MAPE of 0.147, and RMSE of 0.147, along with the highest R^2 of 0.95. In addition, the baseline networks exhibited comparable performance. Across both solar stations, the proposed model is consistently seen as the superior one in terms of performance. The integration of these architectures yielded more accurate and reliable forecasts, minimizing errors and maximizing performance.

Table 2: Performance Indicators of SPF.

`Stations	Model	MAE	MAPE	RMSE	R²
Solar station 1	LSTM	0.064	0.193	0.181	0.87
	CNN-LSTM	0.058	0.164	0.139	0.90
	XAI-SP	0.053	0.158	0.119	0.93
Solar station 2	LSTM	0.061	0.175	0.143	0.91
	CNN-LSTM	0.057	0.169	0.135	0.91
	XAI-SP	0.046	0.147	0.104	0.95

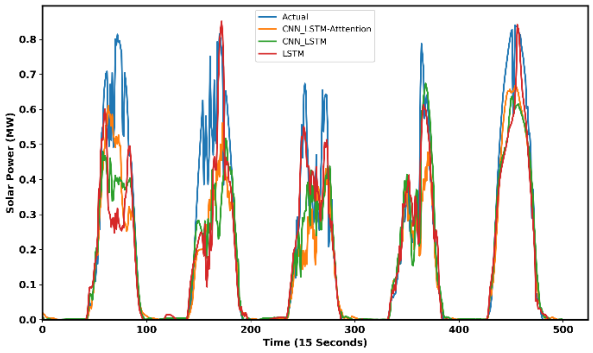


Fig. 2. SPF at Solar Station Site 1 (50 MW Capacity).

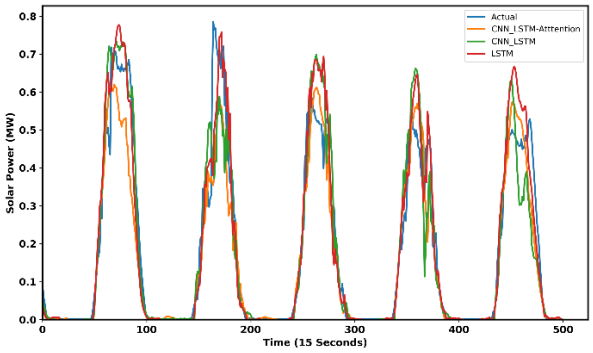


Fig. 3. SPF at Solar Station Site 2 (130 MW Capacity).

5.1 Feature ablation analysis

Fig. 4 and Fig. 5 present the results of the feature ablation analysis conducted on the SPF for both solar stations. The analysis aims to evaluate the impact of individual features on the model's output. The feature ablation methodology was applied to investigate the performance of the three networks. The feature analysis of the proposed model output of datasets are shown in Fig. 4 and Fig. 5. Fig. 4 shows “Total solar irradiance (W/m2)" has the most significant impact on the model's output, with an impact value of 0.1157094. This suggests that solar irradiance plays a crucial role in the model's forecasting capabilities. To modify this feature would lead to the largest change in the model's predictions compared to the other features.

The second most influential feature is "Atmosphere (hpa)" with an impact value of 0.07694051. This indicates that atmospheric pressure also has a substantial effect on the model's output, although less than solar irradiance. "Air temperature (°C)" and "Global horizontal irradiance (W/m2)" have similar impact values of 0.04363717 and 0.04108186, respectively. These features contribute moderately to the model's predictions, with air temperature having a slightly higher influence than global horizontal irradiance. "Direct normal irradiance (W/m2)" has the lowest impact value of 0.02045121, suggesting that it has the least influence on the model's output among the analyzed features. However, it still contributes to the model's performance, as evident from the non-zero impact value. Moreover, from Fig. 5, total solar irradiance and direct normal irradiance have the highest impacts at 0.10689645 and 0.09172923, respectively, indicating their crucial role in accurate forecasting. Air temperature (0.07142675) and atmosphere (0.04309148) have moderate influences, while global horizontal irradiance has the least impact at 0.03129654 among the analyzed features. It can be concluded that feature ablation analysis is a way to determine which features are the most crucial in SPF. It helps researchers make better decisions when enhancing their SPF models by focusing on the most important features and potentially the model, making it faster and more accurate. However, it's important to keep in mind that the results may vary depending on the dataset and how the features interact with each other. Accurate SPF can contribute to the integration of solar energy into the power grid and help with planning and managing solar resources more effectively.

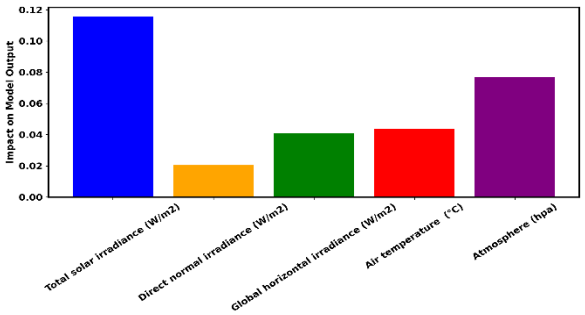


Fig. 4. Key Feature Contributions to SPF at Site 1.

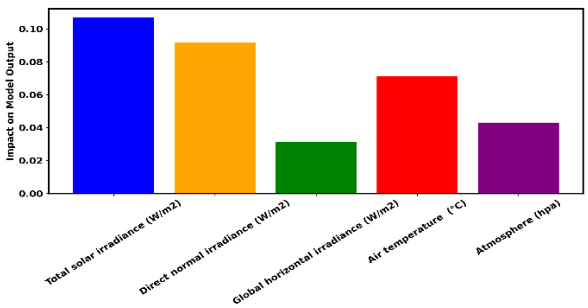


Fig. 5. Key Feature Contributions to SPF at Site 2.

6. Conclusion

This study presents a novel XAI-SP forecasting approach that utilizes the strengths of CNNs, LSTMs, and attention mechanisms to deliver accurate and transparent explanations. By integrating spatial feature extraction, temporal dependency capture, and focused attention on key time steps. The proposed XAI-SP model provides a comprehensive understanding of SPF by the feature ablation study, reveals that total solar irradiance and direct normal irradiance are critical features for achieving accurate forecasts, informing data acquisition strategies and feature selection for optimizing solar forecasting models. Notably, the model enriches performance better than

baseline models across two real-world datasets, in terms of the lower MAE and highest R^2. The proposed model's versatility and generalizability are demonstrated through its application to two distinct solar stations, showcasing its potential for widespread adoption in the renewable energy sector.

References

[1] A. H. Alami *et al.*, "Concentrating solar power (CSP) technologies: Status and analysis," vol. 18, p. 100340, 2023.

[2] L. Cheng *et al.*, "Prediction of non-stationary multi-head cloud motion vectors for intra-hourly satellite-derived solar power forecasting," vol. 39, no. 2, pp. 4010-4019, 2023.

[3] P. Bacher, H. Madsen, and H. A. J. S. e. Nielsen, "Online short-term solar power forecasting," vol. 83, no. 10, pp. 1772-1783, 2009.

[4] M. Louzazni, H. Mosalam, A. Khouya, K. J. S. E. T. Amechnoue, and Assessments, "A non-linear auto-regressive exogenous method to forecast the photovoltaic power output," vol. 38, p. 100670, 2020.

[5] A. Phinikarides, G. Makrides, N. Kindyni, A. Kyprianou, and G. E. Georghiou, "ARIMA modeling of the performance of different photovoltaic technologies," in *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, 2013, pp. 0797-0801: IEEE.

[6] K. Wang, X. Qi, and H. J. A. E. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," vol. 251, p. 113315, 2019.

[7] S. Mishra and P. Palanisamy, "Multi-time-horizon solar forecasting using recurrent neural network," in *2018 IEEE energy conversion congress and exposition (ECCE)*, 2018, pp. 18-24: IEEE.

[8] J. Zhang, Y. Chi, and L. Xiao, "Solar power generation forecast based on LSTM," in *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, 2018, pp. 869-872: IEEE.

[9] W. Zheng and G. J. I. T. o. C. Chen, "An accurate GRU-based power time-series prediction approach with selective state updating and stochastic optimization," vol. 52, no. 12, pp. 13902-13914, 2021.

[10] H. Abbasimehr, R. J. J. o. A. I. Paki, and H. Computing, "Improving time series forecasting using LSTM and attention models," vol. 13, no. 1, pp. 673-691, 2022.

[11] D. Gunning and D. J. A. m. Aha, "DARPA's explainable artificial intelligence (XAI) program," vol. 40, no. 2, pp. 44-58, 2019.

[12] K. J. D. Litkowski, MD, USA: CL Research, "Feature ablation for preposition disambiguation," 2016.

[13] H. G. Ramaswamy, "Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization," in *proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2020, pp. 983-991.

[14] L. Gigoni *et al.*, "Day-ahead hourly forecasting of power generation from photovoltaic plants," vol. 9, no. 2, pp. 831-842, 2017.

[15] C.-J. Huang and P.-H. J. I. a. Kuo, "Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting," vol. 7, pp. 74822-74834, 2019.

[16] H. Zang *et al.*, "Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network," vol. 12, no. 20, pp. 4557-4567, 2018.

[17] H. Zang *et al.*, "Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning," vol. 118, p. 105790, 2020.

[18] S. J. N. C. M.-P. Hochreiter, "Long Short-term Memory," 1997.

[19] S. mahdi Miraftabzadeh, M. Longo, and F. Foiadelli, "A-day-ahead photovoltaic power prediction based on long short term memory algorithm," in *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, 2020, pp. 1-6: IEEE.

[20] S. Srivastava and S. J. S. E. Lessmann, "A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data," vol. 162, pp. 232-247, 2018.

[21] M. Gao, J. Li, F. Hong, and D. J. E. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM," vol. 187, p. 115838, 2019.

[22] J. Ospina, A. Newaz, and M. O. J. I. R. P. G. Faruque, "Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model," vol. 13, no. 7, pp. 1087-1095, 2019.

[23] T. Peng, C. Zhang, J. Zhou, and M. S. J. E. Nazir, "An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting," vol. 221, p. 119887, 2021.

[24] A. Agga, A. Abbou, M. Labbadi, Y. El Houm, and I. H. O. J. E. P. S. R. Ali, "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production," vol. 208, p. 107908, 2022.

[25] S. M. J. Jalali, S. Ahmadian, A. Kavousi-Fard, A. Khosravi, S. J. I. T. o. S. Nahavandi, Man,, and C. Systems, "Automated deep CNN-LSTM architecture design for solar irradiance forecasting," vol. 52, no. 1, pp. 54-65, 2021.

[26] S. Sharda, M. Singh, and K. J. I. T. o. S. E. Sharma, "RSAM: Robust self-attention based multi-horizon model for solar irradiance forecasting," vol. 12, no. 2, pp. 1394-1405, 2020.

[27] J. An, Y. Lai, and Y. Han, "Logic Rule Guided Attribution with Dynamic Ablation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, vol. 36, no. 1, pp. 77-85.

[28] F. Zhao, X. Li, Y. Gao, Y. Li, Z. Feng, and C. J. E. S. w. A. Zhang, "Multi-layer features ablation of BERT model and its application in stock trend prediction," vol. 207, p. 117958, 2022.

[29] Y. Chen and J. J. S. D. Xu, "Solar and wind power data from the Chinese state grid renewable energy generation forecasting competition," vol. 9, no. 1, p. 577, 2022.