SCOPUA
BRIDGING IDEAS, BUILDING SOLUTIONS

Journal of Engineering, Science and Technological Trends (JESTT)
Vol. 1, Issue 2, August 2024
DOI: https://doi.org/10.48112/jestt.v1i2c.5

Article

Improved Solar Power Prediction Using CNN-LSTM Models for Optimized Smart Grid Performance

Aamina Ahsan[1], Ahsan Zafar[2*], Muhammad Adeel Afzal[3], Mubashar Javed[4], Daniyal Zafar[5] and Mubashar Ali[6]

[1]Advanced Control System Research Group, Flinders University, Clovelly Park, Australia

[2]School of Electrical and Information Engineering, Tianjin University, Tianjin, China

[3]School of Economics and Management Science, China Three Gorges University, Yichang, China

[4]Department of Electrical and Computer Engineering, COMSATS University, Islamabad, Pakistan

[5]Faculty of Engineering, Abasyn University, Peshawar, Pakistan

[6]Department of Physics, University of Education, Lahore, Pakistan

* Corresponding Email: ahsanuetp@hotmail.com (A. Zafar)

Received: 05 July 2024 / Revised: 17 August 2024 / Accepted: 18 August 2024 / Published Online: 31 August 2024

ABSTRACT

During the fourth energy revolution, the integration of Artificial Intelligence (AI) across various technological fields is critical to meet rising energy demands and address the depletion of fossil fuel reserves, leading to the adoption of smart grids. This study aims to enhance power generation capacity and minimize losses in smart grids by accurately predicting parameters. Traditional power grid stations transitioning to smart grids require precise parameter predictions. To achieve this, we employed AI-based machine learning models, specifically Random Forest (RF) and Long Short-Term Memory (LSTM), to predict the parameters of a solar power plant. After initial analysis through graphical visualization, we further refined the LSTM model using an advanced technique: Convolutional Neural Network (CNN-LSTM). Comparative results indicate that the CNN-LSTM model outperforms both the LSTM and RF models. For daily power generation, the CNN-LSTM achieved the lowest Mean Absolute Error (MAE) of 0.1335 and Mean Squared Error (MSE) of 0.0497. Consequently, the application of AI in this study significantly improves the accuracy of parameter prediction, enhancing the performance of basic machine learning models. This advancement supports the development of a robust and efficient power system that reduces power losses and boosts production capacity within the framework of smart grids.

Keywords*:* Convolutional Neural Network; Long Short-Term Memory; Machine Learning; Solar Power; Renewable Energy

1. Introduction

Renewable energy is becoming increasingly critical as fossil fuel resources continue to dwindle. Integrating green energy sources into smart grids is essential to maintain a reliable and efficient electricity supply. Achieving this integration necessitates the use of advanced technologies like artificial intelligence (AI) [1]. Among renewable sources, solar power plants are significant contributors to meeting energy demands, generating electricity daily for national grids [2]. However, accurately predicting energy generation in solar facilities is challenging [3]. Although many established techniques exist for forecasting power generation, improving their accuracy remains a priority [4]. Recent research has utilized various machine learning (ML) models derived from diverse energy plants, demonstrating success in predicting numerous factors [5]. For instance, the LSTM model has shown lower error rates than other time-series forecasting models like Linear Regression, ARIMA, and SARIMA for predicting photovoltaic (PV) power generation output [8]. However, these models still require fine-tuning and ensemble techniques to enhance efficiency.

Machine learning models have been widely implemented not only in smart grid applications but also in energy storage, frequency modulation, and voltage stability, among other industrial and medical applications [9],[10],[11],[12]. In the medical field, a novel deep-learning method was developed to recover high-quality cardiac MR images. This method effectively models the recurrence of iterative reconstruction stages and learns spatio-temporal dependencies, although it is limited to cardiac MR images and lacks interpretability analysis [13]. The CNN-LSTM hybrid model has demonstrated superior performance in predicting PV farm variables [14]. Despite the model's advantages, data availability and quality remain significant challenges. Similarly, the proposed CNN-LSTM model outperforms traditional ML and single deep learning models in precision, and stability, as evaluated by error metrics like MAE, MAPE, and RMSE [15]. A hybrid framework combining a CNN for local correlations, an A-LSTM network for nonlinear time-series characteristics, and an Auto-Regression model for linear time-series characteristics has shown superior accuracy in forecasting power generation from multiple renewable energy sources [16].

In weather forecasting, deep learning encoder-decoder architectures have been employed to enhance the prediction of tropical cyclone tracks and intensity. The HurriCast structure utilizes multiple ML encoder-decoder methods and data sources to achieve comparable accuracy to operational forecast models, though it is limited to specific regions and a 24-hour anticipation period [17]. Similarly, a comprehensible dam movement prediction model using a mixed attention mechanism LSTM (MAM-LSTM) adaptively selects influential factors and extracts key time segments, providing physical interpretation through the measurement and display of interest weights [18]. However, the practical application of ML models in this field faces challenges like static modelling methods and a lack of adaptive differentiation of segments and influencing factors [19]. For dynamic energy prediction at thermal energy facilities, a hybrid ML encoder-decoder architecture has been used, focusing on a new model based on physics architecture. This model demonstrates the highest accuracy in predicting heat collected by steam and water in a boiler, although further work is needed to expand the dataset and models investigated [20]. A comparison with LSTM auto-encoder models has shown that the latter are superior, albeit with a slightly higher RMSE [21].

Studies utilizing computer vision algorithms alongside ML techniques have successfully estimated wind turbine angular velocity and extracted acceptable data structures for compression, with autoencoders outperforming other feature extraction methods [22],[23]. Autoencoder techniques have also been used to accurately forecast day-ahead solar plant output, managing uncertainties and data noise, with hybrid models outperforming existing methods [24]. Similar success has been found in forecasting photovoltaic system power output using hybrid AE-LSTM models, which demonstrate higher accuracy by identifying complex temporal patterns and relationships in the data [25-28]. The hybrid method combining LSTM neural networks and autoencoders has also shown superior predictive capabilities by capturing both temporal and spatial features in the data [30]. Finally, the LSTM autoencoder (AE) model introduced for photovoltaic power forecasting has outperformed benchmark deep learning methods in various performance measures using a dataset from a 23.40 kW PV power plant in Australia [31]. These models utilize diverse input features such as panel surface temperature, accumulated energy, solar radiation, humidity, irradiance, and past solar energy to effectively forecast solar energy generation.

Accurately predicting solar energy generation is critical for effective smart grid integration. However, due to the inherent variability and instability of solar energy production, this task is challenging. This study proposes a novel approach combining a Convolutional Neural Network (CNN) with a Long Short-Term Memory (LSTM) neural network to address this challenge. The study focuses on forecasting two vital constraints: daily power generation (DPG) and radiance (Rad). The predictive capability is achieved through an LSTM model, further enhanced using CNN LSTM techniques. The main contributions of this study are as follows:

- The proposed CNN LSTM model significantly outperforms the reference models (LSTM and RF) in forecasting solar power generation across several parameters. This innovative architecture leverages the CNN layers to extract spatial features and LSTM layers to capture temporal dependencies.
- The study employs performance metrics such as Mean Absolute Error (MAE), Mean Square Error (MSE), and Root Mean Square Error (RMSE) to evaluate the models' effectiveness. Results consistently show that the CNN LSTM model surpasses both LSTM and RF models in predicting DPG and Rad.
- Using the CNN LSTM model to anticipate solar energy production more accurately is a significant contribution. This improvement is essential for maximizing financial choices about resource allocation and PV plant operations.
- The CNN-LSTM model enhances the accuracy of parameter predictions. AI contributes to the development of a resilient and efficient power system, supporting sustainability goals.

Section 2 of the paper details the proposed methodology, outlining the structure, benefits, and limitations of the prediction models used. Section 3 presents the case study, discussing the data collection process and the nature of the data using graphical visualizations. The results and associated errors are discussed in Section 4. The study is concluded in Section 5.

2. Methodology

This study employed three machine learning models: RF, LSTM, and CNN LSTM to analyze a static time series dataset. The dataset comprised daily power output data from a large-scale solar power facility collected over one year. Preprocessing ensured the data met the criteria for a stationary time series. The dataset was divided into training, sample, and validation sets. The models were trained on the training set, and their performance was evaluated on the test set, with adjustments made to enhance precision. The models' predictions were validated and used to forecast power production for the following year. These predictions, based on the models' extrapolations from the time series data, aimed to assist solar power plant operators in resource allocation and pricing decisions. The LSTM model focused on making accurate predictions about future power generation. Incorporating AI into smart grids offers several benefits, such as improved energy management, precise power output predictions, enhanced grid stability, and better integration of renewable energy sources. This study found that adding a CNN to the LSTM model improved accuracy and reduced both MAE and MSE compared to the LSTM and RF models.

The methodology involves collecting one year's worth of real-time data from a solar farm, focusing on two key metrics: daily power generation and radiance. Each metric includes approximately 365 values, representing a year of data. Three AI techniques: RF, LSTM, and CNN LSTM were developed using

Python. The data underwent initial analysis to ensure quality and consistency, verifying numerical format without inconsistencies. The data was split, with 80% (10 months) used for training the models and 20% (two months) reserved for final predictions and visualization s. These visualizations provided graphical representations of the predicted results. Comparative analysis showed that the CNN LSTM model had lower RMSE and MAE values when applied to real-time solar farm data, indicating higher accuracy in predictions compared to the LSTM and RF models. Obtaining the forecasted visual outcomes from the models is depicted in detail in the flowchart shown in Figure 1.

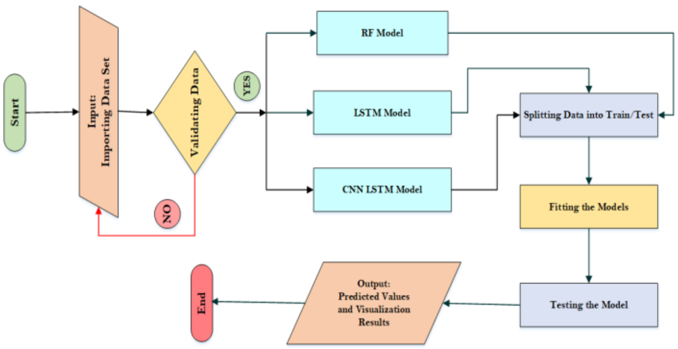


Figure 1: *Proposed Methodology Flowchart to Get Final Visualizations from LSTM, CNN LSTM and Auto-encoder LSTM.*

2.1. Functional Process of RF

The random forest model, which is well-known for its effectiveness and higher accuracy in regression when compared to other approaches, is a frequently used methodology for classification and regression in decision tree learning. The forecast is obtained by averaging the outputs of many uncorrelated decision trees that are constructed during training. The bagging approach, which fits decision trees using the Gini impurity and repeatedly chooses bootstrap samples of the training set, is used to train each tree. The outputs of all the trees are averaged, as in Equation 1, to make predictions for fresh data. Random forest as illustrated in Figure 2 produces forecasts that are more reliable and precise since it models numerous trees rather than just one.

$$Y = \frac{1}{B}\sum_{b=1}^{B} t_b(\mathrm{x}) \qquad (1)$$

Where B is the number of trees and $t_b(\mathrm{x})$ represents the prediction of each tree.

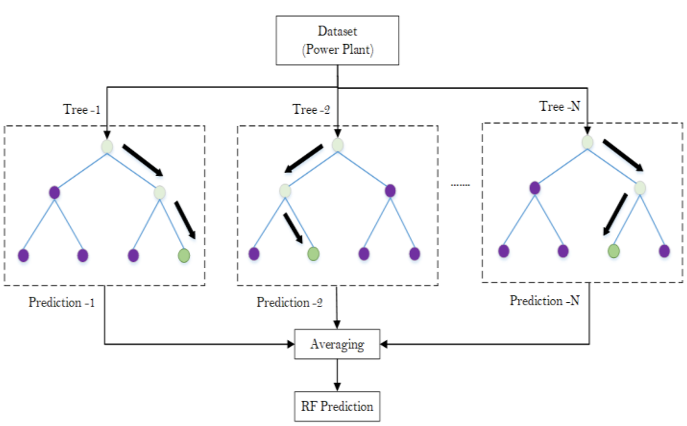


Figure 2: *RF Structure.*

A development of decision tree techniques that addresses their propensity for overfitting is random forest. Though they are prone to collecting noise, decision trees are essential in classification and regression problems. By generating an ensemble of trees trained on randomized data subsets, random forest prevents overfitting. It improves resilience and accuracy by averaging forecasts.

2.2. Functional Process of LSTM

Long Short-Term Memory (LSTM) cells are specialized units within recurrent neural networks (RNNs) designed to handle long-term dependencies in data sequences. These cells feature interconnected components that manage the selective retention or forgetting of information over time. The primary elements of an LSTM cell include the input gate, forget gate, output gate, and memory cell as shown in Figure 3. The forget gate decides which information to discard from the memory cell, applying it to remove unnecessary data. The memory cell itself stores information persistently, continuously updated by the input and forget gates and maintains its state through a self-loop connection for long-term retention. Utilizing an activation function that is sigmoid and squeezing the values obtained from the activation function that is tanh to stay within a predetermined range, the gate of output regulates the information flow from the cell of memory to the subsequent state that is hidden.

By orchestrating these elements, LSTM cells can effectively update memory, hold onto important information, eliminate unnecessary data, and produce accurate outputs, making them particularly adept at identifying long-term relationships in sequential data [8].

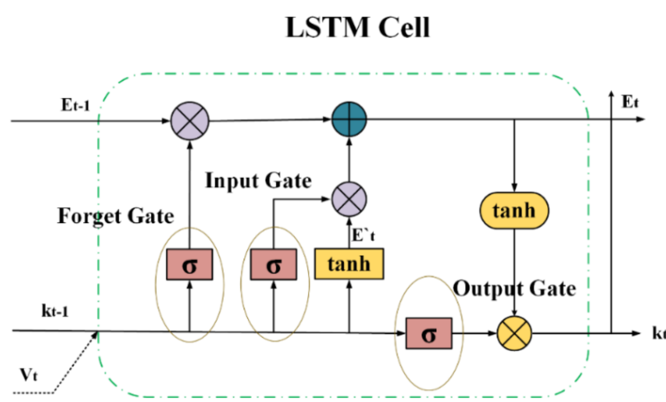


Figure 3: *Structure of LSTM.*

Three gates make up the LSTM cell structure: the gate of forget, the gate of input, and the gate of output, which controls the flow of information within the LSTM cell. Here, we provide a professional explanation of the equations mentioned in the research paper [14]: Equation (2) shows the forget gate (G_f)

$$G_f = \sigma\{X_f\,(k_{t-1}, v_t)\} \qquad (2)$$

Here σ is the sigma function, weighted X_f the sum of the inputs in the gate of forget, k_{t-1} is earlier step's time in a state of hidden and v_t is present input.

The gate layer of input (G_i), calculated using the function of sigmoid, selects the values that need to be updated. Equation (3), presents the tanh layer that produces new candidate (E'_t) values which might be appended to the state of the cell.

$$G_i = \sigma\{X_i(k_{t-1}, v_t)\} \quad (3)$$
$$E_t' = tanh\{X_c * (k_{t-1}, v_t)\} \quad (4)$$

The state of the cell (E_t) is modified. Equation (5) describes this update process:

$$E_t = (G_f * E_{t-1}) + E_t * E_t' \quad (5)$$

Equation (6) represents this calculation for the state of hidden (S_h).

$$G_o = \sigma\{X_o * (k_{t-1}, v_t)\} \quad (6)$$
$$S_h = G_o * tanh(E_t) \quad (7)$$

In the context of the above LSTM equations, the recurrent weights are denoted by (X_f, X_i, X_c, X_o). At time step$t$, the input, hidden state, and cell state are represented by v_t , k_t and E_t , respectively.

2.3. Functional Process of CNN LSTM:

Convolutional Neural Networks (CNNs) are structured with layers including an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer, designed to directly detect visual patterns in pixel images. The CNN LSTM model, depicted in Figure 4, integrates CNN layers for feature extraction with Long Short-Term Memory (LSTM) layers for sequence prediction. This architecture is advantageous for tasks such as activity recognition, labelling images and videos, time series forecasting, and creating textual annotations from image sequences. The CNN-LSTM framework starts with an input layer for extracting features, followed by LSTM layers that capture temporal dependencies to enhance prediction. However, the CNN part extracts spatial features, while the LSTM part models sequential patterns, making the structure adept at handling datasets that are time series. This integrated approach is especially useful for forecasting time series data with intricate correlations, where understanding both spatial and temporal patterns is essential.

The flexibility of CNNs is demonstrated in various applications, including solar energy forecasting, where 1D convolution simplifies time series analysis. This versatility allows CNNs to excel in diverse tasks. The operation is defined by the equation (8)

$$z = \sigma(W * x + b) \quad (8)$$

Where z is the output feature map, $W * x$ denotes the convolution of the filter weights with the input data, and σ is the activation function. Additionally, the pooling layer's output z' is given by

$$z' = p(z) \quad (9)$$

Where z the input is the feature map and p represents the pooling operation.

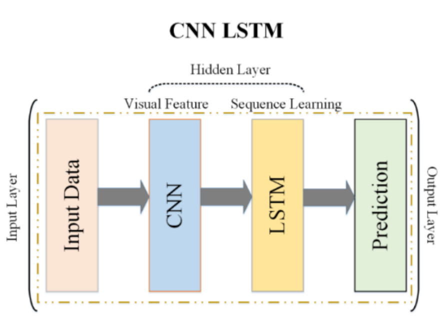


Figure 4: *Structure of CNN LSTM.*

Section 4 delves deeper into this enhanced performance, emphasizing the CNN LSTM model's efficacy in contrast to other models such as LSTM.

3. CASE STUDY

3.1 Solar Plant's Structure:

The Zhenfa Energy Group Solar PV Park, located in Punjab, Pakistan, is a notable renewable energy initiative developed and owned by Zhenfa New Energy. The construction of the Zhenfa Energy Group Solar PV Park marked a significant milestone in the country Figure 5. It shows the structure of the solar plant from where the data is collected.

This substantial project has a capacity of 100 megawatts (MW) and spans 650 acres. Operational since April 2022, the plant is equipped with over 400,000 solar panels and connects to the national grid via a 132 kV transmission line. The solar power plant, generating around 165 GWh of electricity annually, contributes to meeting Pakistan's growing energy needs while reducing reliance on fossil fuels and supporting climate change mitigation efforts.

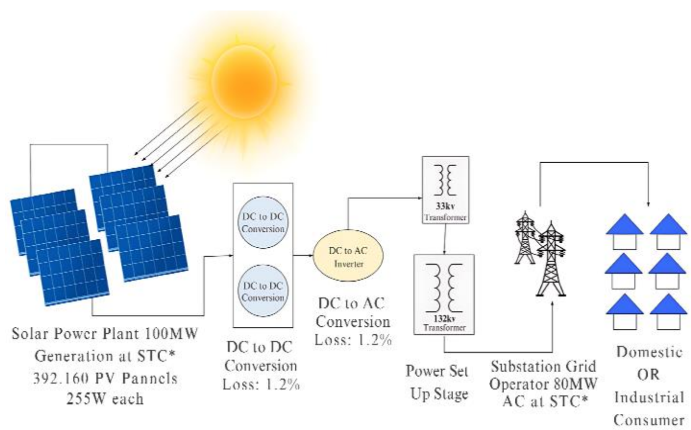


Figure 5: *Single Line Diagram to Express the Structure of Solar Plant.*

3.3. Data Analysis

The dataset for Solar Energy Variables (SEVs) comprises two categories of variables within the 100MW solar plant. In this simulation utilizing deep learning algorithms, the independent variable under consideration is the Radiance, while the dependent variable being analyzed is the Daily power generation.

Daily power generation is crucial for optimizing solar power plants, typically producing 400MW to 500MW daily. Monitored through energy meters and a SCADA system, it provides data for simulations, reflecting the plant's performance influenced by environmental factors and operational activities. This metric helps assess the plant's efficiency and its contribution to meeting energy demands and renewable energy supply. Radiance (MJ/m²) in a solar power plant measures the solar energy received on a surface area per unit of time, crucial for evaluating energy capture efficiency. It accounts for factors like panel orientation, atmospheric conditions, and sunlight angle, directly affecting the plant's electricity generation. This metric is essential for optimizing panel alignment and overall plant performance.

Box plots provide a concise and effective method for summarizing large datasets and understanding their distribution. Figure 6 illustrates the interquartile range (Q1–Q3), median (Q2), and extreme or outlier values for three features from a large-scale power plant: Daily Generation, and Radiance. The whiskers stretch to the least and greatest scores between a range of 1.5 times the interquartile range from Q1 and Q3, while the dimension that is vertical of the box reflects the middle 50% of the information. Relevant outliers are indicated by separate dots or circles for points of data, not within the specified range.

The heat map in Figure 7 illustrates the correlation of the three parameters in which three columns and three rows, with each cell representing the correlation coefficient between two parameters. The diagonal cells show the correlation of each parameter with itself, which is always 1. A more positive association is shown by higher numbers and lower numbers indicate a weaker negative correlation among the two variables. The values in the cells vary between 0 to 1. As an illustration, a positive correlation is denoted by a correlation factor of 1, a negative correlation is implied by a factor of -1, and no correlation is shown by a factor of 0. The range of correlation coefficients is represented by a colour bar on the right side of the heat map. The colour bar has a range between 0.6-1.0, deeper hues correspond to higher positive correlation and lighter hues to lower positive correlations.

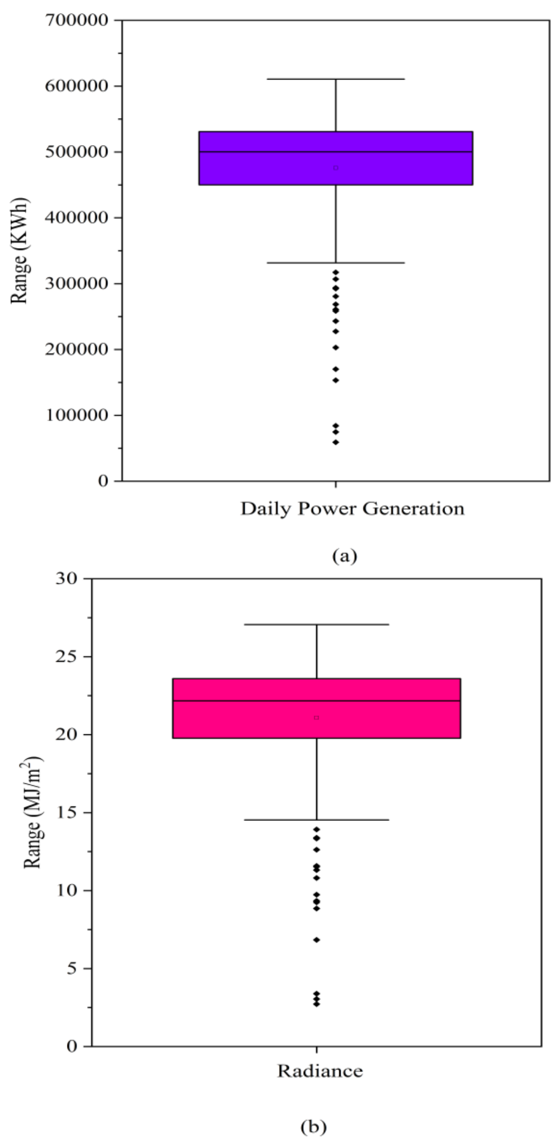


Figure 6: *Analysis of Data via Box Plot.*

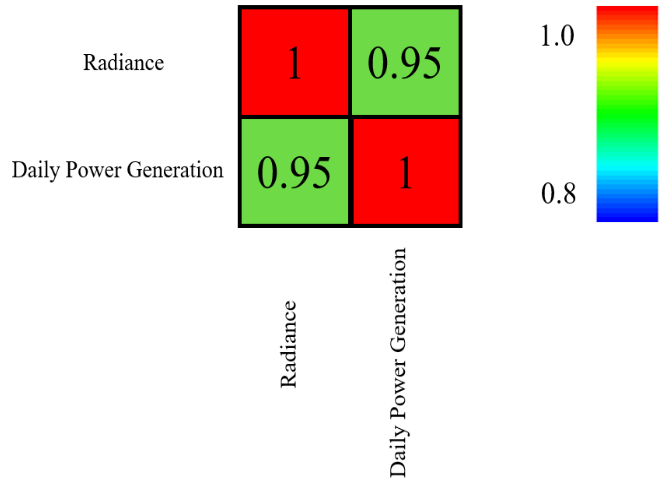


Figure 7: *Heat Map Data Analysis.*

The findings obtained from the models are presented in graphical depiction form in the following section 4.

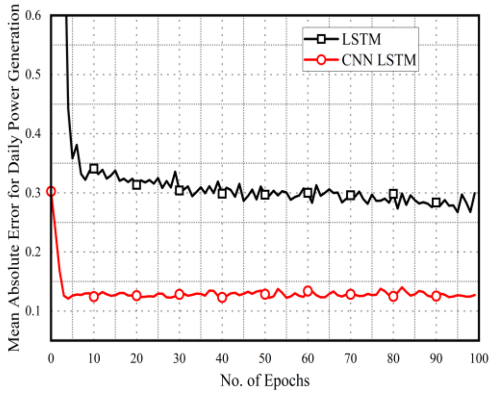
Table 1:
Model Performance indicators.

Performance Indicator	Formula
Mean Absolute Error	$MAE = \frac{1}{N}\sum_{n=1}^{N} \lvert \overline{x_n} - x_n \rvert$
Mean Square Error	$MSE = \frac{1}{N}\sum_{n=1}^{N} (\overline{x_n} - x_n)^2$

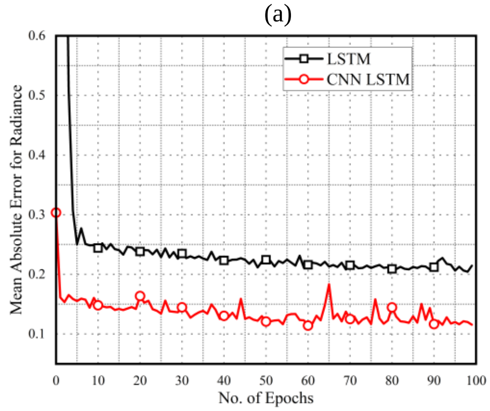
4. Results and Discussion:

This section presents the findings derived from analyzing a year's worth of real-time data from a solar power plant, consisting of three essential parameters: “daily power generation”, and “radiance”. Our objective was to develop machine learning models for future predictions based on this dataset. After an extensive review of the literature, we opted for the RF and LSTM model due to its promising track record in similar applications. The LSTM model implementation yielded satisfactory results. However, in our pursuit of enhancing prediction accuracy, we introduced the CNN LSTM. The results are represented visually, and it is clear that the CNN LSTM model outperformed both LSTM as well as RF models, exhibiting a decreased error rate and better forecasting accuracy.

Figure 8(a & b) depicts the comparison of validation MAE among LSTM and CNN-LSTM models for the DPG and Rad parameters, the CNN-LSTM model demonstrates superior performance. For the DPG parameter, the LSTM model yields an MAE of 0.162, whereas the CNN-LSTM model achieves a lower MAE of 0.127, indicating better predictive accuracy. The difference is even more pronounced with the Rad parameter, where the LSTM model records an MAE of 0.214, compared to the significantly lower MAE of 0.115 for the CNN-LSTM model. This suggests that the CNN-LSTM model is an accurate parameter prediction.



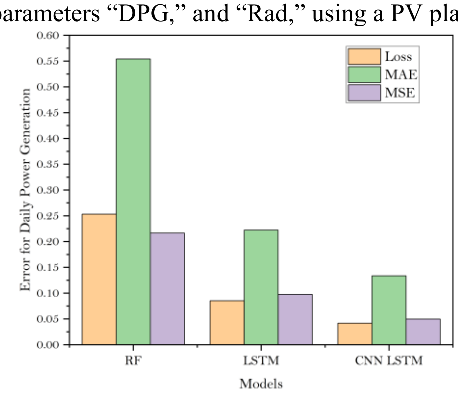
(a)



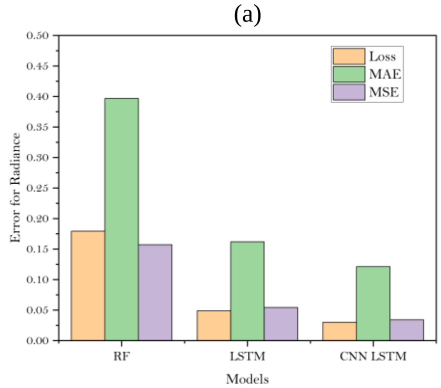
(b)

Figure 8: ***(a)***. *MAE of Daily Power Generation through LSTM & CNN LSTM and* ***(b)***. *MAE of Radiance through LSTM & CNN LSTM*

Figure 9(a & b) present the results of the loss, MAEs and MSE comparison between the RF, LSTM and the CNN LSTM model using the parameters "DPG," and "Rad," using a PV plant.



(a)

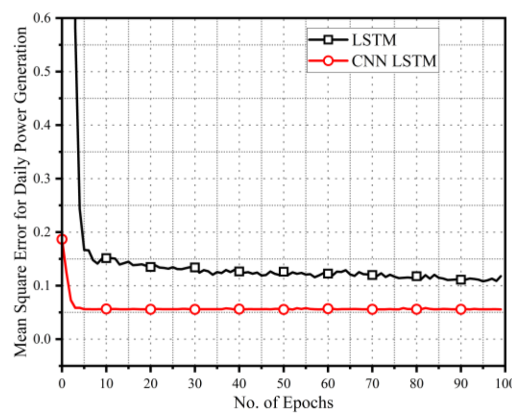


(b)

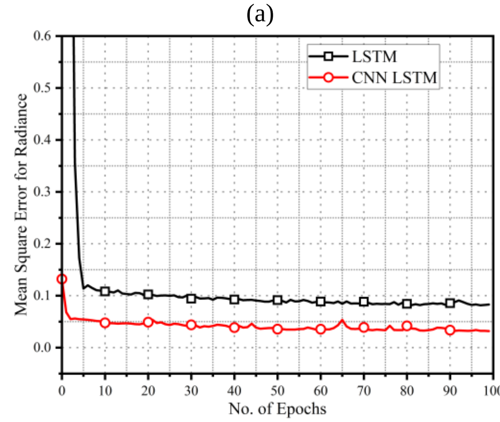
Figure 9: ***(a)***. *Error Comparison of Daily Power Generation through RF, LSTM & CNN LSTM and* ***(b)***. *Error Comparison of Radiance through RF, LSTM & CNN LSTM*

Table 2 compares the performance of three models: Random Forest (RF), Long Short-Term Memory (LSTM), and Convolutional Neural Network with LSTM (CNN LSTM) across two datasets: Daily Power Generation and Radiance. For the Daily Power Generation dataset, the CNN LSTM model outperforms both RF and LSTM, achieving the lowest Loss (0.0418), Mean Absolute Error (MAE) (0.1335), and Mean Squared Error (MSE) (0.0497). In contrast, RF shows the highest values for Loss (0.2532), MAE (0.554), and MSE (0.2165), indicating weaker performance. LSTM performs better than RF with a Loss of 0.0852, MAE of 0.2223, and MSE of 0.0976, but it still lags behind the CNN LSTM. Similarly, for the Radiance dataset, CNN LSTM again demonstrates superior performance with a Loss of 0.0301, MAE of 0.1212, and MSE of 0.0341. RF has the highest metrics, with a Loss of 0.1793, MAE of 0.3968, and MSE of 0.1571. The LSTM model performs moderately with a Loss of 0.0489, MAE of 0.16188, and MSE of 0.05427. Overall, CNN LSTM consistently provides the most accurate predictions across both datasets, highlighting its effectiveness in time-series forecasting tasks.

In Figure 10(a & b), the results of validation MSE between the LSTM and CNN LSTM models for the DPG and Rad parameters, the CNN-LSTM model consistently shows better performance. For the DPG parameter, the LSTM model achieves an MSE of 0.068, while the CNN-LSTM model improves upon this with a lower MSE of 0.055. The difference is even more significant for the Rad parameter, where the LSTM model records an MSE of 0.083, compared to the substantially lower MSE of 0.031 for the CNN-LSTM model. These results indicate that the CNN-LSTM model is more effective in reducing prediction errors.



(a)



(b)

Figure 10*:* ***(a)***. *MSE of Daily Power Generation through LSTM & CNN LSTM and* ***(b)***. *MSE of Daily Power Generation through LSTM and CNN LSTM*

Table 2:
Comparison Table between RF, LSTM, and CNN LSTM Results

Parameter	RF			LSTM			CNN LSTM		
	LOSS	MAE	MSE	LOSS	MAE	MSE	LOSS	MAE	MSE
Daily Power Generation	0.2532	0.554	0.2165	0.0852	0.2223	0.0976	0.0418	0.1335	0.0497
Radiance	0.1793	0.3968	0.1571	0.0489	0.16188	0.05427	0.0301	0.1212	0.0341

In addition, a close examination of each model's graphical representation shows plenty of resemblance, suggesting that the accuracy of data is higher than 95% and that none of the outliers have been found. These results provide even more evidence for the applicability and accuracy of the power plant data, offering a strong basis for further investigation and making decisions.

The practical application of parameters like daily power generation (kWh) and radiance (MJ/m²) is essential for effectively forecasting solar power plant performance. These parameters play critical roles in solar energy production, influencing multiple areas. Daily power generation is crucial for analyzing plant efficiency, planning maintenance, and estimating financial returns, making it indispensable for operators and owners. Radiance is vital for predicting energy output, tracking performance, planning maintenance, and evaluating site suitability, which together enhance the financial feasibility and dependability of solar projects. These parameters are key to maximizing efficiency, sustainability, and reliability in solar power generation. Moreover, by facilitating the production of cleaner energy and decreasing reliance on fossil fuels, these models support environmental sustainability goals. Finally, their integration into smart grids improves grid performance and stability, optimizing the use of renewable energy and advancing the development of more sustainable and efficient power systems.

Figure 11 illustrates the "Daily Power Generation (kWh)" data over a year, with the y-axis ranging from 2×10^5 to 7×10^5 units also the x-axis representing the days in number. This graph provides a clear visualization of the precision and performance of various ML models used for prediction. The closer the alignment between the test data and the forecast data, the more accurate the models are in predicting this parameter. By comparing the predictions from RF, LSTM, and CNN LSTM models with actual test data over 60 days, it is clear that the CNN LSTM model operates remarkably effectively. Its predictions closely match the test data with minimal discrepancies. When these models are used to forecast daily power generation over the next 10 months (300 days), the CNN LSTM model consistently outperforms the other two models. It demonstrates the lowest prediction errors compared to the actual test data, indicating its superior accuracy in forecasting daily power generation for the upcoming year. The LSTM model follows, showing better performance than the basic RF model. This suggests that the LSTM model captures more intricate patterns and dependencies in the data compared to the RF model.

In conclusion, the analysis of Figure 11 indicates that the CNN LSTM model is the most precise for predicting DPG over the next year, followed by the LSTM model. The RF model lags in terms of predictive accuracy. This information is crucial for decision-making in the energy sector, as it can help stakeholders make more informed choices regarding power generation and distribution.

Figure 12 presents the "Radiance" parameter, with values ranging from 0 to 35 MJ/m² on the y-axis and the number of days displayed on the x-axis. This figure visually showcases predictions for radiance levels over the next year using three different models: RF, LSTM, and CNN LSTM. These models are employed to forecast "Radiance" for the remaining 10 months, as shown in Figure 12. The test data graph reflects actual values from a solar power plant, while the prediction graphs display estimated values generated by the RF, LSTM, and CNN LSTM models. Over a 60-day test period, the predictions from the CNN LSTM model closely match the actual test data, with only minor deviations at certain points. A thorough analysis indicates that the CNN LSTM model consistently outperforms both the LSTM and the basic RF models. The predictions made by the CNN LSTM align closely with the test

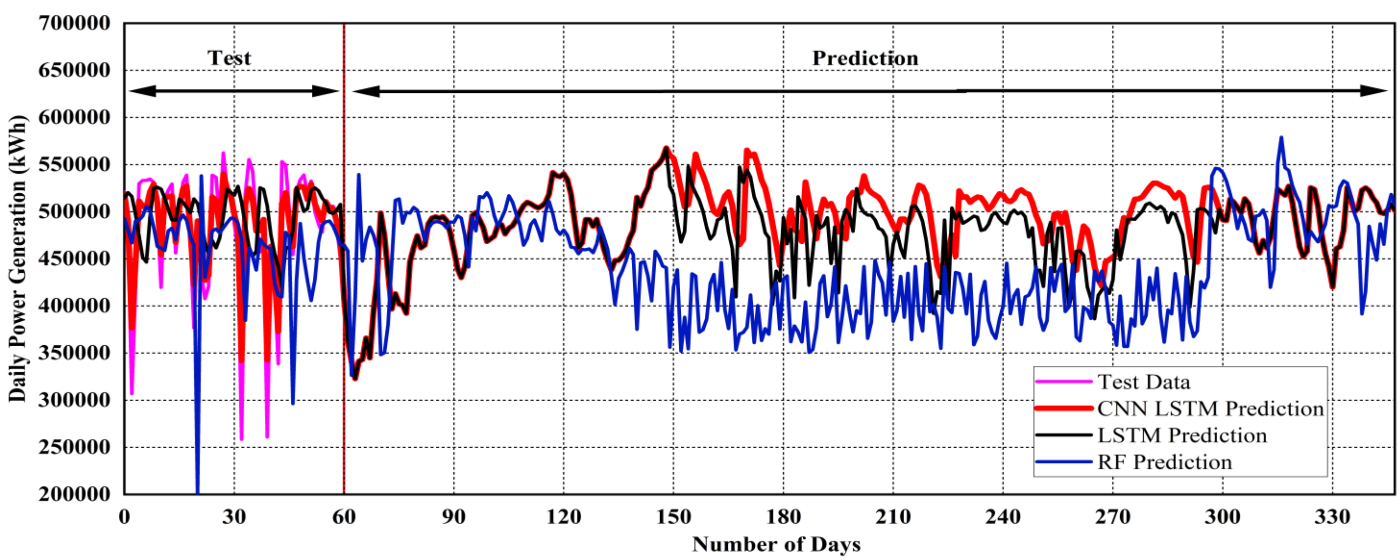


Figure 11: An Analysis of RF, LSTM, and CNN LSTM Models for Predicting Solar Plant Daily Power Generation.

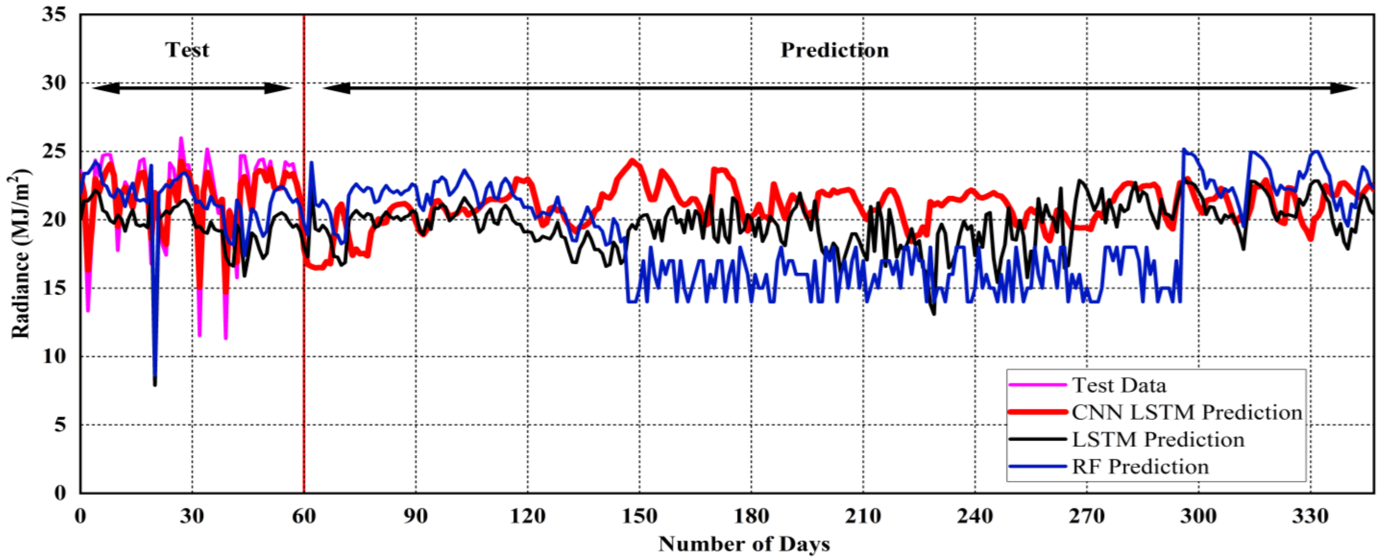


Figure 12: An Analysis of RF, LSTM, and CNN LSTM Models for Predicting Solar Plant Radiance.

data, demonstrating high accuracy in radiance forecasting. Although there are slight discrepancies, they are significantly fewer compared to those observed in the other models. The LSTM model also shows better predictive capabilities than the basic RF, although it falls slightly short of the accuracy achieved by the CNN LSTM.

These results highlight the importance of using advanced techniques, particularly the CNN LSTM, for precise radiance predictions. This data is invaluable for decision-makers in renewable energy, enabling them to make informed choices and improve their planning for radiance-related applications in the upcoming year.

5. Conclusion

In this comprehensive study, we thoroughly analyzed DPG and Rad data via a PV plant over a year. We carefully selected, extensively trained, and rigorously tested three machine learning models: RF, LSTM, and CNN LSTM to predict these crucial parameters. To evaluate the discrepancies between predicted and actual values, we used MAE and MSE as metrics. The visual representations of the results demonstrated the superiority of the CNN LSTM model over the LSTM and RF models in all scenarios. The CNN LSTM consistently exhibited reduced error scores and greater data similarity compared to the other models. Specifically, for DPG, the CNN LSTM model achieved a notably low Root Mean Square Error (RMSE) of 0.222, while the LSTM model recorded an RMSE of 0.312, and the RF model had a higher RMSE of 0.465. A similar pattern was observed for the Rad parameter, where the CNN LSTM model excelled with an RMSE of 0.184, outperforming the LSTM model (RMSE of 0.232) and the RF model (RMSE of 0.396). These results underscore the CNN LSTM's exceptional capability to accurately predict these parameters. Additionally, it's important to note that the data visualizations were based on a combination of 20% experimental data and 80% predictive data. The CNN LSTM consistently demonstrated its ability to capture underlying patterns and trends, resulting in more precise predictions across all parameters.

Further enhancing these models through hybrid or ensemble techniques could establish a solid foundation for future advancements in renewable energy technologies and their integration into existing power infrastructures. These optimized models have the potential to significantly contribute to the overall stability, reliability, and economic viability of renewable energy systems, paving the way for a greener and more sustainable future.

Data availability statement: The whole data of this research is included in this article.

Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Pal, R., Chavhan, S., Gupta, D., Khanna, A., Padmanaban, S., Khan, B., & Rodrigues, J. J. (2021). A comprehensive review on IoT-based infrastructure for smart grid applications. *IET renewable power generation*, *15*(16), 3761-3776.
2. Qayum, M., Xu, Z., Qayyum, S., & Batool, A. (2024). New Aspect of Management Engineering: Connotation between FDI Inflow, Gender Gap, Educational Attainment and Skilled Workforce. *Journal of Engineering, Science and Technological Trends*, *1*(1), 16-26.
3. Başaran, K., Bozyiğit, F., Siano, P., Yıldırım Taşer, P., & Kılınç, D. (2020). Systematic literature review of photovoltaic output power forecasting. *IET Renewable Power Generation*, *14*(19), 3961-3973.
4. Sudharshan, K., Naveen, C., Vishnuram, P., Krishna Rao Kasagani, D. V. S., & Nastasi, B. (2022). Systematic review on impact of different irradiance forecasting techniques for solar energy prediction. *Energies*, *15*(17), 6267.
5. Abdulnabi, G. (2023). Preparation and Characterization of (CuO/TiO2/α-Fe2O3) Ternary Nanocomposite and Application in a Solar Cell. *Journal of Engineering, Science and Technological Trends*, *1*(1), 01-13.
6. Kondaiah, V. Y., Saravanan, B., Sanjeevikumar, P., & Khan, B. (2022). A review on short-term load forecasting models for micro-grid application. *The Journal of Engineering*, *2022*(7), 665-689.
7. Nishtar, Z., & Afzal, J. (2024). Seq2Seq-Based-Day-Ahead Scheduling for SCUC in Islanded Power Systems with Limited Intermittent Generation . *Journal of Engineering, Science and Technological Trends*, *1*(1), 43–50.
8. Kim, E., Akhtar, M. S., & Yang, O. B. (2023). Designing solar power generation output forecasting methods using time series algorithms. *Electric Power Systems Research*, *216*, 109073.
9. Alahmar, H. T. M. (2023). An Evaluation of Algorithms Applied to the Lattice Boltzmann Methods Technique for Segmentation of Medical Images: A Review. *Journal of Engineering, Science and Technological Trends*, *1*(1), 22–32.
10. Zafar, A. (2019). OFFSHORE WIND ENERGY CONNECTED TO HVDC SYSTEM, VSC CONTROL. *City University International Journal of Computational Analysis*, *3*(1), 29-40.
11. Shah, N., & Zafar, A. (2022). Improved Performance of Silicon-Germanium Solar Cell Based on Optimization of Layer

Thickness. *City University International Journal of Computational Analysis*, *5*(1), 1-10.
12. Farooq, M. U., Ghani, M. U., & Zafar, A. (2016, December). Survey on driving behavior and motivational factors causing aggressive driving: A case study of Peshawar, Pakistan. In *Proceedings of the 2nd International Conference on Emerging Trends in Engineering, Management & Sciences (ICETEMS-2016), Peshawar, Pakistan* (pp. 28-30).
13. Zafar, A., Che, Y., Rasool, S., Afzal, U., & Aamina, A. (2023, December). Evaluation of Machine Learning Models for Predicting Smart Grid Parameters. In *Proceedings of the 5th International Conference on Emerging Trends in Engineering, Management & Sciences (ICETEMS-2023), Peshawar, Pakistan* (pp. 15-25).
14. Jamil, I., Lucheng, H., Iqbal, S., Aurangzaib, M., Jamil, R., Kotb, H., ... & AboRas, K. M. (2023). Predictive evaluation of solar energy variables for a large-scale solar power plant based on triple deep learning forecast models. *Alexandria Engineering Journal*, *76*, 51-73.
15. Agga, A., Abbou, A., Labbadi, M., El Houm, Y., & Ali, I. H. O. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. *Electric Power Systems Research*, *208*, 107908.
16. Zheng, J., Du, J., Wang, B., Klemeš, J. J., Liao, Q., & Liang, Y. (2023). A hybrid framework for forecasting power generation of multiple renewable energy sources. *Renewable and Sustainable Energy Reviews*, *172*, 113046.
17. Boussioux, L., Zeng, C., Guénais, T., & Bertsimas, D. (2022). Hurricane forecasting: A novel multimodal machine learning framework. *Weather and forecasting*, *37*(6), 817-831.
18. Ren, Q., Li, M., Li, H., & Shen, Y. (2021). A novel deep learning prediction model for concrete dam displacements using interpretable mixed attention mechanism. *Advanced Engineering Informatics*, *50*, 101407.
19. Liang, J., & Tang, W. (2022). Ultra-short-term spatiotemporal forecasting of renewable resources: An attention temporal convolutional network-based approach. *IEEE Transactions on Smart Grid*, *13*(5), 3798-3812.
20. Machalek, D., Tuttle, J., Andersson, K., & Powell, K. M. (2022). Dynamic energy system modeling using hybrid physics-based and machine learning encoder–decoder models. *Energy and AI*, *9*, 100172.
21. Mukhoty, B. P., Maurya, V., & Shukla, S. K. (2019, June). Sequence to sequence deep learning models for solar irradiation forecasting. In *2019 IEEE Milan PowerTech* (pp. 1-6). IEEE.
22. Bahaghighat, M., Abedini, F., Xin, Q., Zanjireh, M. M., & Mirjalili, S. (2021). Using machine learning and computer vision to estimate the angular velocity of wind turbines in smart grids remotely. *Energy Reports*, *7*, 8561-8576.
23. Das, L., Garg, D., & Srinivasan, B. (2020). NeuralCompression: A machine learning approach to compress high frequency measurements in smart grid. *Applied Energy*, *257*, 113966.
24. Zhang, Y., Qin, C., Srivastava, A. K., Jin, C., & Sharma, R. K. (2020). Data-driven day-ahead PV estimation using autoencoder-LSTM and persistence model. *IEEE Transactions on Industry Applications*, *56*(6), 7185-7192.
25. Neo, Y. Q., Teo, T. T., Woo, W. L., Logenthiran, T., & Sharma, A. (2017, November). Forecasting of photovoltaic power using deep belief network. In *Tencon 2017-2017 IEEE Region 10 Conference* (pp. 1189-1194). IEEE.
26. Li, L. L., Cheng, P., Lin, H. C., & Dong, H. (2017). Short-term output power forecasting of photovoltaic systems based on the deep belief net. *Advances in mechanical engineering*, *9*(9), 1687814017715983.
27. Gensler, A., Henze, J., Sick, B., & Raabe, N. (2016, October). Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks. In *2016 IEEE international conference on systems, man, and cybernetics (SMC)* (pp. 002858-002865). IEEE.
28. Abdel-Nasser, M., & Mahmoud, K. (2019). Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural computing and applications*, *31*, 2727-2740.
29. Zafar, A., Che, Y., Ahmed, M., Sarfraz, M., Ahmad, A., & Alibakhshikenari, M. (2023). Enhancing Power Generation Forecasting in Smart Grids using Hybrid Autoencoder Long Short-Term Memory Machine Learning Model. *IEEE Access*.
30. Said, Y., & Alanazi, A. (2023). AI-based solar energy forecasting for smart grid integration. *Neural Computing and Applications*, *35*(11), 8625-8634.
31. Zafar, A., Che, Y., Faheem, M., Abubakar, M., Ali, S., & Bhutta, M. S. (2024). Machine learning autoencoder-based parameters prediction for solar power generation systems in smart grid. *IET Smart Grid*.