IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 August 2024, accepted 26 August 2024, date of publication 2 September 2024, date of current version 10 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3452778

RESEARCH ARTICLE

Explainable Deep Learning Model for Grid-Connected Photovoltaic System Performance Assessment for Improving System Reliability

IMAD HASSAN[1], IBRAHIM ALHAMROUNI[1], ZAHRAOUI YOUNES[2], NURUL HANIS AZHAN[1], SAAD MEKHILEF[3,4], (Fellow, IEEE), MEHDI SEYEDMAHMOUDIAN[3], AND ALEX STOJCEVSKI[5]

[1]Electrical Engineering Section, British Malaysian Institute, University of Kuala Lumpur, Kuala Lumpur 53100, Malaysia
[2]Department of Engineering Cybernetics, Norwegian University of Science and Technology (NTNU), 7034 Trondheim, Norway
[3]Department of Engineering Technologies, Swinburne University of Technology, Melbourne, VIC 3122, Australia
[4]Department of Electrical and Electronics Engineering, Presidency University, Bengaluru, Karnataka 560064, India
[5]Curtin Singapore, Singapore 117684

Corresponding authors: Ibrahim Alhamrouni (ibrahim.mohamed@unikl.edu.my) and Saad Mekhilef (smekhilef@swin.edu.au)

ABSTRACT Solar power is an important renewable resource in our journey towards a sustainable energy future; however, integrating it with existing grids, especially in dust-prone environments, presents challenges, such as power reduction and financial impact. Regular performance assessment is crucial for identifying issues and maximizing energy production. Therefore, the development of an accurate and reliable predictive model is essential. Such a model should not only predict photovoltaic (PV) system performance but also offer insights into various factors influencing system efficiency. In this regard, this study presents the development of an interpretable deep learning model for the assessment of photovoltaic (PV) system performance. This model focuses on predicting the essential key performance indicator (KPI) performance ratio, which is crucial for PV system evaluation. A feedforward neural network (FFNN) architecture enhanced by a univariate linear regression approach was employed to comprehend the coefficient weights for interpretability. To optimize the model, various optimizers were explored during model training. Furthermore, Local Interpretable Model-agnostic Explanations (LIME) were utilized to determine the influence of specific factors on each prediction made by the FFNN model, enhancing its explainability. The performance of the model was evaluated using standard metrics, such as R-squared (R2)(0.9965), Mean Absolute Error (MAE)(0.0036), Mean Squared Error (MSE)(0.0001), and Root Mean Squared Error (RMSE)(0.0078). The results indicate that the proposed model outperforms conventional deep-learning models, demonstrating promising accuracy and interpretability for PV system performance assessments. By providing insights into the factors affecting PV system performance, our model aims to assist operators and stakeholders in making informed decisions to optimize solar energy utilization.

INDEX TERMS PV system, performance ratio, machine learning, soiling loss, sustainability.

I. INTRODUCTION

The global energy landscape is undergoing a transformative shift, with increasing emphasis on the sustainability and adoption of renewable energy sources [1], [2], [3]. Governments, industries, and communities are aligning their priorities to reduce the carbon footprint and transition towards cleaner alternatives [4], [5], [6]. Amid this broader movement, solar energy has emerged as a promising, abundant, and environmentally friendly power sourcesource [8], [9],

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy.

[10]. The primary aim of all PV systems is to effectively convert sunlight into inexpensive power without wasting energy [10], which is referred to as the PV system efficiency or performance. Monitoring the functionality of these systems is crucial for several reasons. First, it ensures that solar installations operate optimally, thereby maximizing energy production and economic returns [11], [12]. Through performance evaluation, system owners can pinpoint and rectify issues, such as equipment malfunctions, shading, or soiling, which may impede efficiency [13], [14]. The performance of a solar PV power facility is influenced by various weather conditions such as solar radiation, temperature, wind speed, precipitation, humidity, dust accumulation, atmospheric pressure [15], [16], [17], [18], and technical parameters such as inverter loss and PV array losses [19], [20]. The accumulation of dust and debris on solar panels, referred to as soiling loss, is a significant factor influencing solar power generation. The effect of soiling loss on the solar energy output is substantial, introducing a dynamic that directly affects the performance of solar installations. Understanding and mitigating these losses are essential to ensure the consistent and optimal performance of solar energy systems [21], [22], [23]. To recognize the crucial role of photovoltaic (PV) systems in sustainable energy production, a growing number of researchers have focused on evaluating their performance. Recent strides in artificial intelligence (AI) have expanded its applications across various domains, positioning AI techniques as highly promising tools for assessing the performance of PV systems. By leveraging AI algorithms and machine-learning methodologies, researchers aim to enhance the efficiency, reliability, and overall effectiveness of PV systems. By harnessing the power of AI, these studies seek to optimize energy output, mitigate operational inefficiencies, and address maintenance issues, thereby significantly contributing to the advancement and widespread adoption of renewable energy technologies [24], [25], [26], [27]. In recent years, several studies have been conducted globally to enhance the understanding of the photovoltaic (PV) system performance under varying environmental conditions. For instance, a study in Kerala, India [28] investigated a 2 MW system using an Adaptive Neuro-Fuzzy Inference System (ANFIS), Artificial Neural Networks (ANN), and Response Surface Methodology (RSM), focusing on solar irradiance, wind speed, and ambient temperature. AI tools have been employed to predict the power generation and performance ratios in solar PV systems. The ANFIS model exhibited the highest predictive efficiency, demonstrating low errors and a robust performance for power generation prediction. However, limitations such as high RMSE values and the black-box nature of the machine learning models were noted. Another study in Busan, Korea [29], investigated a hybrid model combining a Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) networks to forecast the PV power generation of a 2.5 kW system based on qualitative and quantitative evaluation. Qualitative evaluation assessed the model's ability to reflect real-world trends and fluctuations, whereas quantitative evaluation measured statistical accuracy using metrics such as RMSE, MAE, and MAPE. This study aimed to address the challenges posed by environmental factors such as solar radiation, temperature, humidity, cloud cover, and wind speed, which significantly impact PV power output. The CNN component classifies weather conditions, whereas the LSTM component learns the power generation patterns based on these conditions. However, limitations such as the non-explainable nature of the hybrid model and the exclusive focus on power generation without incorporating other performance metrics, such as the performance ratio or reference yield, were observed. Another study in Greece focused on a 100 kWp grid-connected PV park [30] by applying Artificial Neural Networks (ANNs) to handle the uncertainties associated with solar radiation. This study highlighted the capability of the ANN model to decouple the fluctuating effects of PV panel soiling, which interferes with the efficiency degradation process. This methodology aims to quantify degradation effects and serves as a fault diagnosis tool for long-term analysis. Similarly, a study conducted at King Khalid University in Abha, Saudi Arabia [31] employed various ML algorithms to predict the power outputs of residential PV systems. This study utilized data from a PV system and weather station and applied a backward feature elimination technique to identify the most relevant features for accurate prediction. Among the tested models, the deep-learning-based model provided the lowest prediction errors with a minimal feature set of approximately seven features. When the feature set exceeded ten features, the polynomial regression model demonstrated superior prediction performance with minimal errors. Conversely, the linear regression model exhibited the highest prediction errors. A notable study in Jordan [32] utilized an ensemble of optimized and diversified Artificial Neural Networks (ANNs) for 24-hour ahead PV power production predictions. The key findings indicated that the ensemble approach significantly outperformed the benchmark models, including a smart persistence model and a single optimized ANN model, achieving performance gains of up to 11% in RMSE, 12% in MAE, and 9% in WMAE. The optimized and diversified ANNs provided a reduction in prediction errors and a robust mechanism for handling the variability in the solar power output. Additionally, a study conducted in Gandhinagar, Gujarat, India [33] focused on the long-term seasonal performance of monocrystalline (m-Si), polycrystalline (p-Si), and amorphous silicon (a-Si) PV modules within a 1 MW solar plant. This study utilized key performance indicators (KPIs) such as daily power generation, final yield (Yf), reference yield (Yr), total energy loss (TEL), and performance ratio (PR) to evaluate the modules. The research incorporated an Explainable AI (XAI) model to predict KPIs and evaluate the performance of several learning algorithms, using metrics such as R-value, MAE, iteration count, and execution time. The Levenberg-Marquardt algorithm achieved prediction

TABLE 1. Overview of the literature on performance evaluation of PV systems using machine learning methodologies.

Reference	Year	Capacity	Location	ML Models	Parameters	Feature Selection	PV system assessment KPI	Limitation
[33]	2024	1 MW	Gujrat, India	XAI	ambient temperature, relative humidity, surface pressure, wind speed, wind direction, precipitation, solar irradiation	×	final yield, reference yield, performance ratio	Lack of interpretability for individual predictions, high prediction error for monocrystalline system
[28]	2022	2 MW	Kerala, India	ANFIS, ANN, RSM	solar irradiance, wind speed, ambient temperature	×	energy yield, performance ratio	Black box models, high RMSE
[29]	2022	2.5 KW	Busan, Korea	CNN-LSTM Hybrid Model	Weather data	×	power generation	Black box models
[30]	2022	100 KW	Greece	ANN	Ambient Temperature, Solar Irradiance	×	output power	Black box model
[34]	2022	15 KW	Faisalabad, Pakistan	Logistic Regression	Solar irradiance,relative humidity, air temperature, wind speed	×	final yield, reference yield, performance ratio, capacity factor	Black box model, no consideration for factors like dust affecting PV performance
[35]	2022	100 MW	Bahawalpur, Pakistan	ARIMA	historical data from the PV plant SCADA system	×	performance ratio, production amount, plan of array	Black box model, lack of comparison with other machine learning models
[31]	2021	5.8 KW	Abha, Saudi Arabia	LASSO, LR, PR, SVM, RF, XGBoost, NN	Weather station data, pyranometer sensor data	backward feature-elimination technique	output power	Black box models
[36]	2020	954 KW	Nouakchott, Mauritania	Linear Models	Insolation, Module Temperature, Wind Speed	×	generated energy, array output	Black box model
[37]	2020			ARIMA, SVM, MLR	IoT system collected data from PV plant	×	performance ratio	Black box model
[32]	2019	231 KW	Jordan	Ensemble model	Ambient temperature, Relative humidity,Wind speed, wind direction, pressure	×	output power	Black box model
[38]	2019	-	-	PCA-SVM	Weather data	×	performance ratio	Black box model, Low R2 value
[39]	2019	-	-	Regression analysis	global horizontal irradiance, ambient temperature , wind speed , relative humidity	×	output Power	Black box model. Limited discussion on the impact of outliers in model accuracy.

accuracies of 98.63% for p-Si, 98.58% for am-Si, and 90.09% for m-Si. However, despite the use of sophisticated training algorithms, the model lacks interpretability, making it difficult to detect biases and provide clear explanations of individual predictions. It also exhibits high prediction errors for monocrystalline silicon systems. A summary of the existing literature on the performance assessment of PV systems utilizing various machine learning methodologies and parameters based on key performance indicators is presented in Table 1.

TABLE 2. Location and orientation of PV power plant.

Parameter	Values
Field Type	Fixed Tilted Plan
Optimum tilt of PV modules	27^0 / 180^0
Yearly Global tilted irradiation at optimum angle	2461 kWh/m^2
Latitude	24.7^0 N
Longitude	46.8^0 E

A. CONTRIBUTION AND NOVELTY

The literature presents a consensus on the critical role of advanced PV power prediction models in facilitating the integration of solar energy into the grids. The development of models that can provide reliable forecasts of the PV system performance is a key area of focus. Multi-source data fusion and machine learning techniques are among the strategies that have been found to improve forecasting performance. Continuous evolution of these models is essential for optimizing grid operations and maximizing the use of renewable energy sources. Despite this progress, several challenges remain to be overcome. The majority of existing research relies on "black box" models for PV system performance assessment, lacking transparency in their decision-making processes, highlighting a significant opportunity to develop interpretable machine learning models that provide clear explanations for their predictions, thereby enhancing the understanding of the factors influencing PV system performance. Additionally, many studies do not employ rigorous feature selection techniques, potentially overlooking critical parameters that impact PV system performance and hampering transparency and the ability to detect multicollinearity, leading to unstable models and unreliable interpretations. While some studies have utilized machine learning and XAI models, there remains the potential to explore advanced, interpretable deep learning architectures to capture temporal dependencies and complex relationships within the PV system data, thereby enhancing assessment accuracy and reliability. Furthermore, most studies have focused predominantly on weather data, often neglecting important technical parameters such as inverter losses, soiling loss factor, array voltage and current, and generated energy, integrating these parameters could result in more accurate and comprehensive PV system assessment models. Finally, there has been limited exploration of advanced optimization algorithms during model training, and investigating these optimizers could significantly improve the model performance and convergence. According to the existing literature review, this study is the first to develop an explainable deep learning model for PV system performance assessment by integrating univariate linear regression for interpretability and Local Interpretable Model-Agnostic Explanations (LIME) for explainability within an FFNN architecture, utilizing the ADAM optimizer for model training. This integration not only provides accurate predictions, but also offers unparalleled insights into the underlying factors affecting PV system performance.

To address the gaps in the existing literature, our study aimed to achieve the following objectives:

- New Deep Learning-based Architecture: The primary contribution lies in the development and application of an explainable deep learning model tailored to assess the performance of a 5 MW grid-connected PV system in a dust-prone environment based on environmental and operating conditions, such as the soiling loss factor and inverter losses, and provides insights into the impact of these factors, along with other environmental and technical factors on the model output, thereby providing a new methodology for predicting PV system performance ratios with high accuracy.
- Enhanced Interpretability and explainability: Interpretability and explainability are two key features achieved through specific methodologies. Univariate linear regression was used to understand the weights of each input, thereby enhancing interpretability. Additionally, the application of LIME analysis allows us to comprehend the effect of each input on the model's predictions, thereby enhancing explainability.
- Model Optimization: Through a thorough analysis of various training algorithms, we identified the ADAM optimizer as the most suitable model for predicting the Performance Ratio of the PV system. This selection was based on the final test score loss and the iteration time.
- Superior Performance: We compared the performance metrics of our explainable deep learning model with those of conventional deep learning models. The

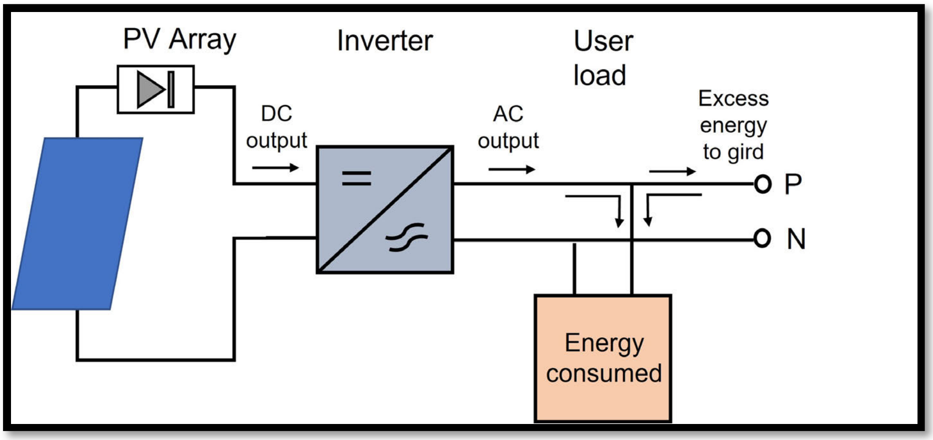

FIGURE 1. Simplified sketch of the grid connected PV system.

results clearly demonstrate that our proposed model not only offers enhanced explainability but also achieves superior accuracy. This comparison demonstrated the effectiveness and practical benefits of the proposed approach, highlighting its potential for real-world applications.

- Real-world Applicability: By providing detailed insights into the factors affecting PV system performance, our model assists operators and stakeholders in making informed decisions to optimize solar energy utilization in challenging environments.

II. SYSTEM DESCRIPTION AND PARAMETERS

Riyadh, the capital city of Saudi Arabia, was the proposed site for installing the solar system. It is located at approximately 24.7136° N latitude and 46.8° E longitude at an altitude of approximately 612 m above sea level. These geographical coordinates make Riyadh well suited for solar energy capture because of its ample sunlight throughout the year. To optimize the efficiency of the solar panels, they were positioned at specific angles. The azimuth angle, set at 180° (facing south), ensured that the panels faced the path of the sun directly. The tilt angle was adjusted to approximately 27°, allowing for maximum exposure to sunlight and efficient energy generation from the panels. This configuration resulted in a yearly transposition factor of 1.1% and a yearly global irradiation of 2461 kW/m ^{2} on the collector, as shown in Table 2. These figures indicate the effectiveness of the solar power system in Riyadh, where geographical conditions are conducive to efficiently harnessing solar energy.

A standard grid-connected PV system with a 5 MW capacity was used for the simulation, as depicted in Figure 1. The PV system, composed of Swiss Solar's IBEX-144MHC-Cosmos-455Wp-144 cell modules, consists of 10,985 units with a combined nominal output of 4998 kWp, spanning an area of 24,441 square meters. These high-efficiency monosilicon modules were arranged in 845 strings, with each string comprising 13 modules, to optimize power generation. Operating at 50°C, this configuration delivered a total power of 4552 kWp at 493 volts and 9232 A. Its functionality was enhanced by five Danfoss QLX 1000 inverters individually rated at 1000 kWac, providing a collective capacity of 5000 kWac. These inverters directly converted the DC power from the modules to AC power, while maintaining a nominal ratio of 1.00. Operating within a voltage range of 450-950 volts, this system demonstrates efficient energy conversion, resulting in an impressive annual energy injection of 8.31 GWh into the grid.

Riyadh experiences a desert climate, classified as BWh, according to the Köppen-Geiger climate classification [40]. The city's climate is characterized by extremely hot summers and mild winters, with very low precipitation throughout the year. The city also experiences dust storms, particularly during the transitional periods. The soiling loss factor in Riyadh varies throughout the year. During dusty and windy seasons, typically from March to September, the soiling loss factor can result in performance reduction ranging from 5% to as high as 40% per month, which is equivalent to approximately 0.17%–1.33% of power drop per day. These months saw the highest soiling rates, owing to the increased dust and sand in the air. In contrast, during less dusty months, such as October to February, the soiling loss factor may decrease to approximately 2%–5% per month or approximately 0.07%–0.17% per day. This variation in soiling loss corresponds to seasonal changes in dust levels and wind intensity [41], [42] [43].

Table 3 shows the percentage of monthly soiling loss values induced in this study. These values represent the percentage of efficiency reduction in the photovoltaic (PV) system owing to accumulated dust and dirt on the solar panels, resulting in a corresponding loss in the energy generated. Figure 2 shows the differences in power generation with and without soiling loss. Power generation was calculated hourly for an entire year, resulting in 8,760 samples for both scenarios. The data illustrate that the average soiling loss fraction over the year was 24.2%. This average was influenced by varying soiling rates across different months. For instance, in January and February, the soiling losses were relatively low at 5.0%

TABLE 3. Monthly soiling loss induced in the system.

Month	Soiling Loss
January	5%
February	10%
March	15%
April	20%
May	25%
June	30%
July	35%
August	40%
September	50%
October	25%
November	20%
December	15%

and 10.0%, respectively. However, as the year progressed into the dustier months from March to August, soiling losses increased significantly, reaching peaks in September at 50.0% and August at 40.0%. This resulted in a noticeable dip in the energy generation graph, indicating a corresponding decrease in energy injected into the grid. These months typically experience higher levels of dust and sand in the air, leading to a notable decrease in PV efficiency. As the year transitioned to less dusty conditions from October to December, soiling losses decreased to 25.0%, 20.0%, and 15.0%, respectively, resulting in improved energy generation.

The careful selection of input variables for this study was conducted with rigorous attention to detail, focusing on environmental factors essential for understanding photovoltaic (PV) power plant performance, as well as technical parameters of the PV system that can impact its efficiency. Various environmental factors were chosen to provide a comprehensive overview of the external conditions affecting PV plant efficiency. These factors include the wind speed, which influences the cooling of PV modules, and their temperature coefficients; solar irradiance, a critical determinant of available sunlight for energy conversion; ambient temperature, which directly affects the efficiency of both PV modules and inverters; and soiling loss factors, which account for the gradual accumulation of dirt and debris on PV surfaces, resulting in a significant reduction in the output over time. The technical parameters include inverter losses that occur during energy conversion from DC to AC, array output, and the energy generated. Incorporating these variables into the study enables a holistic understanding of the complex relationship between environmental conditions and PV plant performance, which is essential for precise modeling and analysis. Table 4 presents the selected features, along with the minimum and maximum value ranges for all variables.

III. METHODOLOGY

The methodology for this research project aims to develop an explainable deep-learning model for predicting the performance of photovoltaic (PV) systems. The process involved two main steps: first, identifying the key performance indicator (KPI) essential for evaluating the PV system performance, which included the Performance Ratio (PR). Second, a transparent framework was constructed using a feed-forward neural network (FFNN) architecture to predict the identified KPI by incorporating meteorological, environmental, and technical parameters. The methodology encompassed the collection of PV output power data, followed by extensive pre-processing to handle missing values and standardize the dataset. Feature selection and extraction techniques were employed to refine the dataset for interpretability. The refined dataset was trained using five different optimizers (Stochastic Gradient Descent, ADAM, AdaGrad, RMSProp, ADAdelta) to determine the best learning rate. Local Interpretable Model-agnostic Explanations (LIME) were utilized to enhance model explainability, allowing for the determination of each input's contribution to the prediction. Finally, a comparative analysis was conducted against three conventional deep learning techniques (CNN, RNN, and LSTM) based on the performance metrics, thereby demonstrating the superior predictive capabilities of the proposed model for PR. The key steps involved in implementing this methodology for predicting PV system performance are illustrated in Figure 3, providing a visual representation of the sequential processes.

A. IDENTIFYING KPI TO EVALUATE THE PERFORMANCE OF PV SYSTEM

When designing an effective power system, it is critical to ensure that system performance is reliably and accurately assessed. The performance of a photovoltaic (PV) system is an important benchmark for determining the soundness of its design and reliable integration of PV components. According to the criteria proposed by the International Electrotechnical Commission (IEC) Standard 61724, the specified performance parameters for PV-powered systems play a critical role in ensuring increased system efficiency. Based on the investigations conducted in [44], the Performance Ratio (PR) was selected to assess a solar PV power system in this study.

The performance ratio (PR) is a key metric used to assess the efficiency of a solar PV system by comparing its actual energy output with the Reference Yield. It was calculated as the ratio of the actual energy produced (Final Yield) to the expected energy output (Reference Yield) under the optimal conditions. A higher PR indicates that the system operates more efficiently with less energy loss owing to factors such as shading, soiling, or equipment degradation. Monitoring the PR over time helps identify any deviations from the expected performance and allows targeted maintenance or optimization efforts to improve the system's overall efficiency and output. In Figure 4, which shows the initial 1000 instances of the calculated PR for the study, the PR remained relatively stable between 70% and 90%. However, occasional dips in

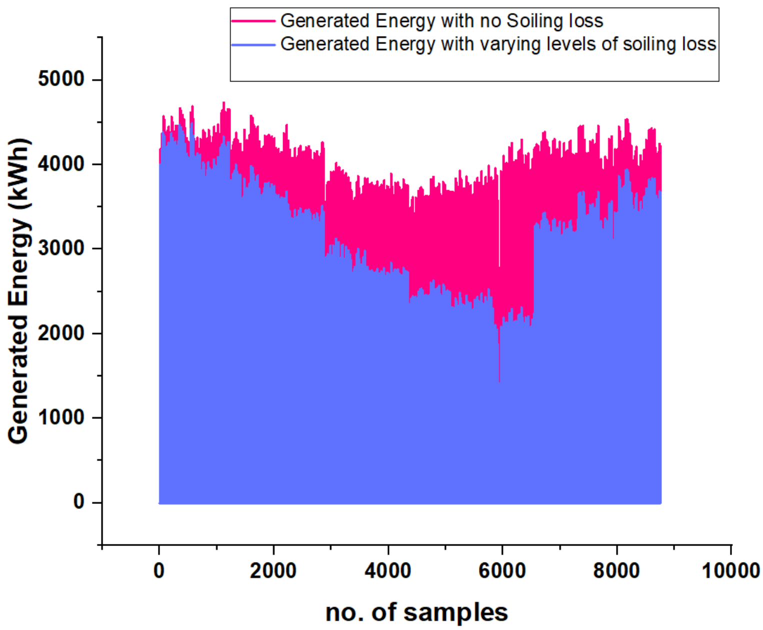

FIGURE 2. Energy generated with and without soiling loss.

TABLE 4. Descriptive statistics of parameters.

	Role	Mean	Std. Deviation	Minimum	Maximum
Solar Irradiance	Feature	496.093	223.808	43.49	1007.8
Generated Energy	Feature	$2.200\times10^{+6}$	995886.805	155358	$4.490\times10^{+6}$
Wind Speed	Feature	3.873	2.048	0.1	12.7
Ambient Temperature	Feature	30.909	9.436	4.9	46.82
Array Current	Feature	4539.933	2054.06	386.13	9230.6
Array Voltage	Feature	496.453	19.579	461.98	559.27
Soiling Loss Factor	Feature	0.758	0.124	0.5	0.95
Array Output	Feature	$2.246\times10^{+6}$	$1.011\times10^{+6}$	189029	$4.593\times10^{+6}$
Inverter Losses	Feature	45484.553	16225.384	29902	102369
Module Temperature	Feature	43.408	10.151	9.511	61.516
PR	Target	0.658	0.126	0.357	0.917

the PR are noticeable, which are likely attributed to low solar irradiance during cloudy or rainy weather, soiling losses, or system inefficiencies including inverter losses.

B. MODEL DEVELOPMENT FOR PREDICTING PV SYSTEM PERFORMANCE INDICATORS

The process of constructing an interpretable deep learning model to predict performance metrics for a 5 MW grid-connected solar power plant involved several key steps. These steps include data collection, data preparation, parameter selection, and model development and evaluation. The dataset created to evaluate the performance of the 5 MW solar plant was characterized by its nonlinear and complex nature. This complexity has led us to explore different deep learning techniques. FFNNs are beneficial for predicting PV system performance in dusty environments because of their ability to capture nonlinear relationships, learn relevant features, handle noise, model complex interactions, adapt to

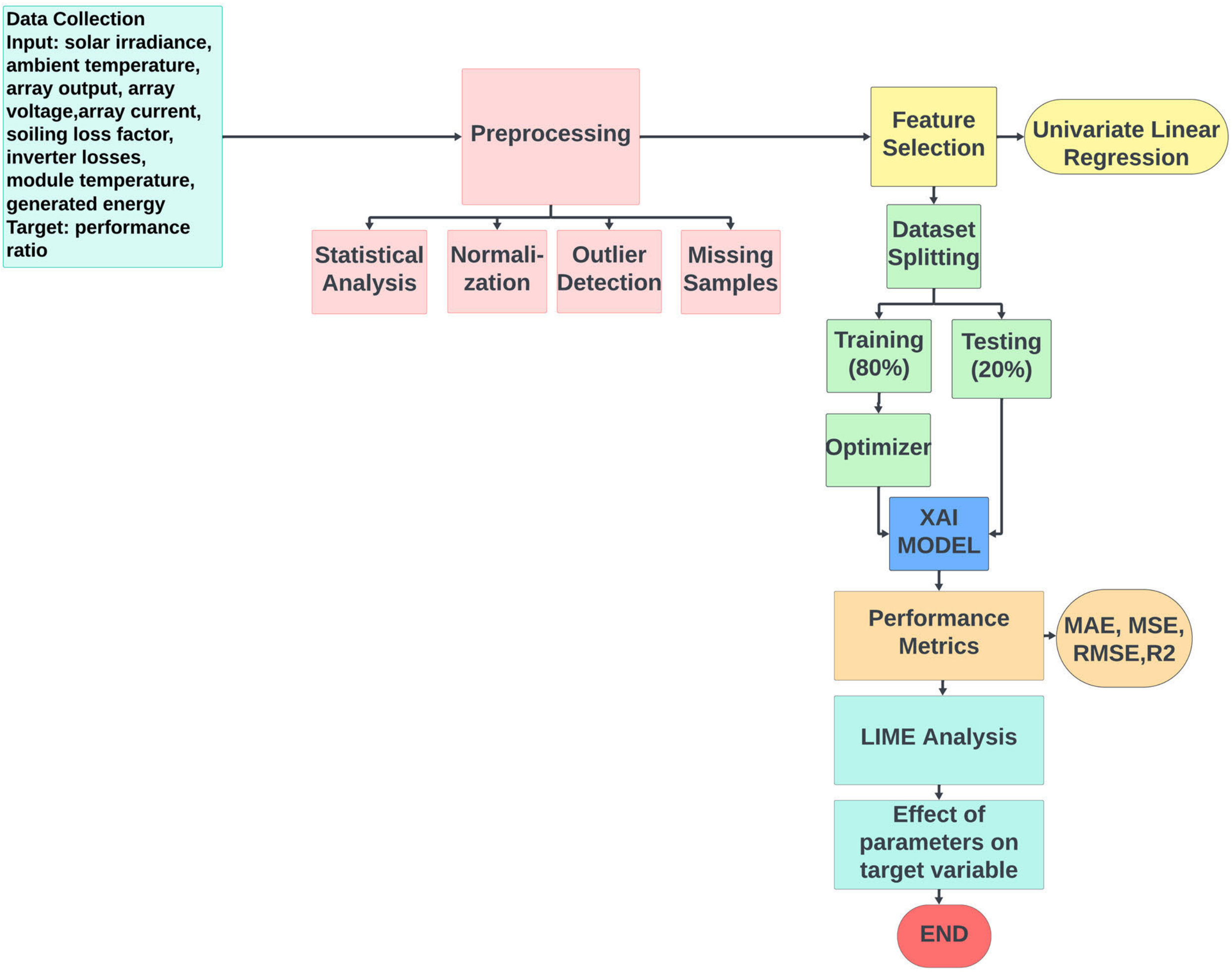

FIGURE 3. Workflow diagram of the proposed methodology.

changing conditions, and enable performance optimization strategies. This technique was carefully evaluated for its ability to predict the unique characteristics of solar plant data.

C. DATA COLLECTION AND ANALYSIS

The data for all parameters were gathered from a 5 MW simulated solar power plant, comprising mainly meteorological, environmental, and technical factors used as input, and the performance ratio was used as the target variable. The simulation was conducted using PVSYST, a widely recognized commercial software known for its effectiveness in PV system design, using real-time meteorological data from the Meteonom database. The proposed model was implemented using Python code in a Google Colab environment. This dataset spans from January 2021 to December 2021, covering a full year. The data for these parameters were recorded hourly throughout the year for the power plant, resulting in a total of 8760 samples. This extensive dataset provides a thorough and representative basis for comprehensive analysis. By including data from the entire year, various weather conditions were accounted for, creating a robust dataset that encompasses the full range of environmental fluctuations and their effects on the PV system performance. To develop the model, data collected at night were disregarded, because there was no solar production during these hours. Therefore, only data from 7 a.m. to 4 p.m. were considered, resulting in a total of 3689 samples for model development. Following data collection, data cleansing techniques were employed for preprocessing and normalization to improve data quality in anticipation of developing the prediction model. The features were normalized using the min-max scaling method to bring them within the range of 0-1. The Interquartile Range (IQR) method was used to detect the outliers. The IQR for each feature representing the middle 50% of the data values was calculated. Outliers were identified using the threshold defined by Equation 1 as the lower bound, and Equation 2 as the upper bound.

$$\text{Lower threshold} = Q1 - 1.5 \times IQR \tag{1}$$

$$\text{Upper threshold} = Q3 + 1.5 \times IQR \tag{2}$$

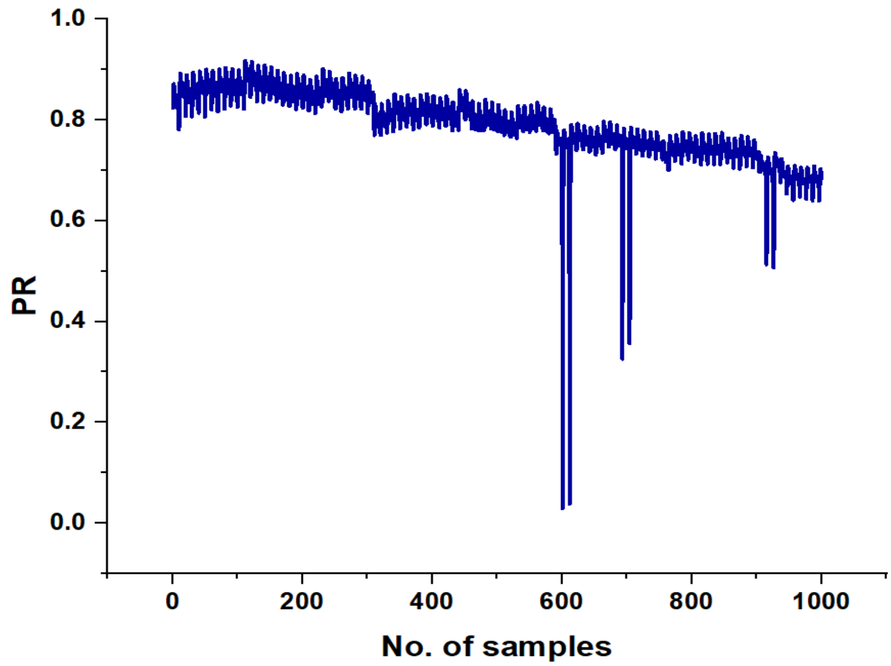

FIGURE 4. Calculated PR first 1000 instances.

In Figure 5, a visualization is crafted to pinpoint outliers within the dataset, which are data points that deviate from the norm. These outliers are highlighted in a line plot illustrating various parameters. The outliers of each parameter are marked with different colors. The provided outlier data exemplify instances in which these parameters deviate from the norm. However, it is important to note that most data points lie within the norm, and the identified deviations are not noise but actual data points. Therefore, these deviations do not require any further manipulation, and visualization serves as a valuable tool for spotting noteworthy anomalies in the dataset.

D. FEATURE SELECTION

The process of condensing a significant volume of data into a concise set of closely interconnected elements is commonly referred to as feature extraction. Feature selection, however, entails the selection of a subset of relevant features to be incorporated into statistical and machine learning models.

This strategy of reducing data volume expedites machine learning processes and streamlines model development with reduced complexity. Within the scope of this study, univariate linear regression was employed as a feature-selection technique. Univariate linear regression serves as a statistical tool for modeling the relationship between the single independent variable and a single dependent variable. This assists in identifying the most influential independent variables (features), which substantially affect the prediction of the dependent variable (target). The algorithm initially postulates a linear relationship between the independent and dependent variables.

The primary objective was to ascertain the coefficients (weights) associated with each independent variable to minimize the disparity between the predicted and actual observed values. Mathematically, the univariate linear regression model is represented by Equation (3).

$$y = \beta_0 + \beta_1 x + E \tag{3}$$

where y is the dependent variable; $\beta 0$ is the intercept; $\beta 1$ the coefficients (weights) associated with the independent variables; x is the independent variable (features); and ε represents the error term accounting for unexplained variability.

The dataset was efficiently refined through the effective application of univariate linear regression, resulting in a more manageable selection of variables for the target variable. The selected variables are identified based on their computed weights, as shown in Figure 6. In the context of PR, Array Voltage emerged as the most influential parameter, with a coefficient of 0.732723, suggesting its significant impact on PR. Following closely are The Soiling Loss Factor and Ambient Temperature, each with coefficients above 0.4, indicate their considerable roles in PR determination. The inverter Losses, Array Output, and Generated Energy were also relevant, although to a slightly lesser extent, with coefficients ranging from 0.3 - 0.4.

E. MODEL OPTIMIZATION

Optimizers play a crucial role in ensuring convergence and efficiency of the training process in machine learning and neural network training. These algorithms are designed to modify the attributes of the neural network, such as the weights and learning rates, to minimize the error between the

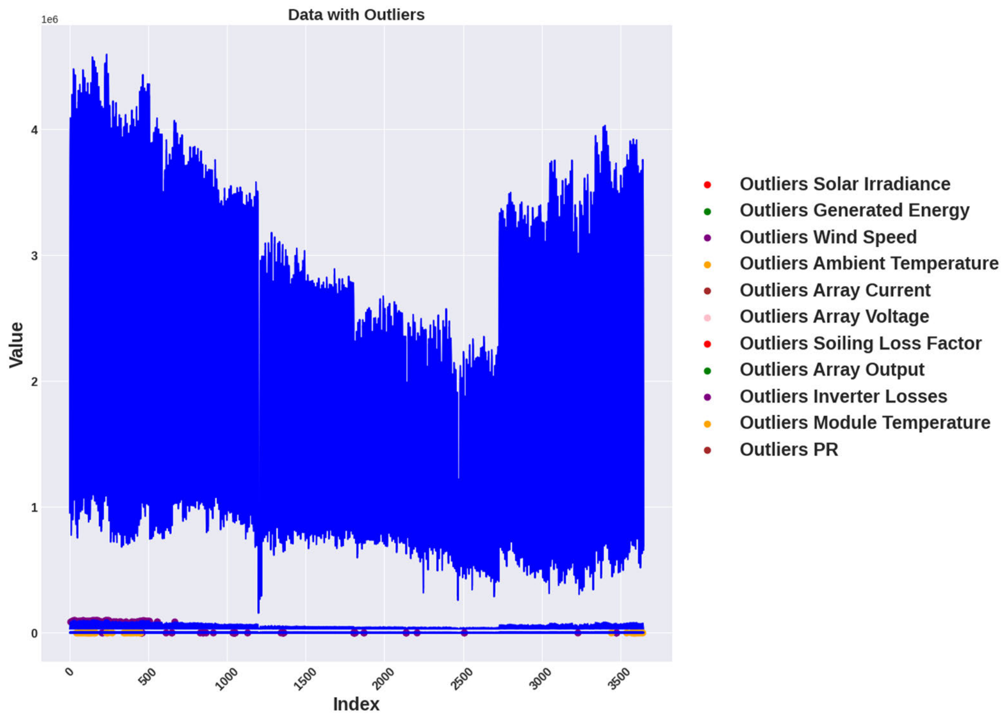

FIGURE 5. Variable outliers.

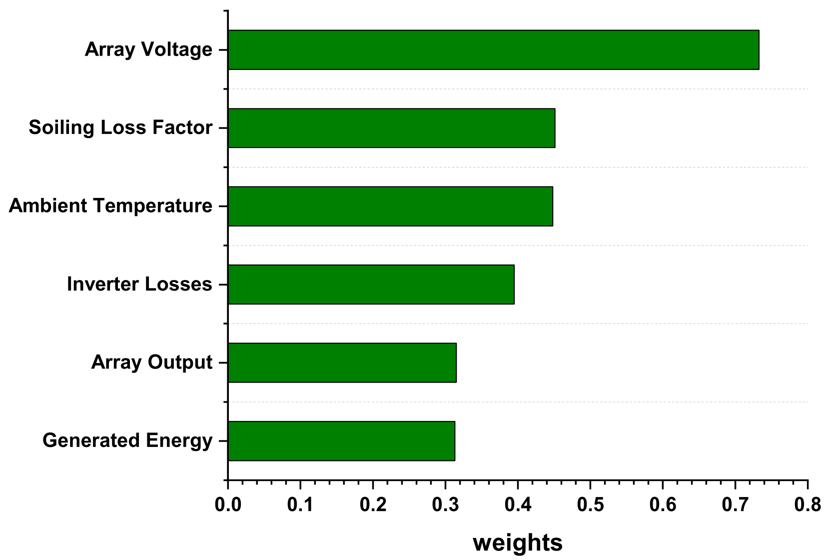

FIGURE 6. Univariate linear regression coefficient weights.

predicted and actual outputs during training. In our proposed model training, we employed a range of optimizers including Adam, Adagrad, RMSprop, and Adadelta. This diverse set of optimizers can potentially enhance the convergence speed, stability, and performance of a model. The following is a concise definition of each optimizer.

1) STOCHASTIC GRADIENT DESCENT (SGD)

This is one of the most fundamental optimizers. The model parameters are updated in the direction opposite the gradient of the loss function with respect to the parameters. SGD is known for its simplicity, but it can be sensitive to the choice of the learning rate [45].

2) ADAPTIVE MOMENT ESTIMATION(ADAM)

Adam is an algorithm for adaptive learning rate optimization. It estimates the first and second moments of the gradients, and uses these values to construct individual adaptive learning rates for various parameters. This approach is effective in many applications and computationally efficient [46].

3) ADAPTIVE GRADIENT ALGORITHM(ADAGRAD)

Adagrad adapts the learning rate to the parameters by performing larger updates for infrequent parameters and smaller updates for frequent parameters. This method is particularly effective when dealing with sparse data [47].

4) ROOT MEAN SQUARE PROPAGATION (RMSPROP)

RMSprop is an adaptive learning rate optimization algorithm that divides the learning rate by an exponentially decaying average of squared gradients. This helps to scale the learning rate differently for each parameter, often leading to better convergence in deep neural networks [48].

5) ADADELTA

Adadelta is an extension of Adagrad that seeks to reduce its aggressive monotonically decreasing learning rate. It uses a moving window for the RMS of parameter updates, allowing it to converge smoothly to an optimal solution [49].

F. PERFORMANCE METRICS

To evaluate the effectiveness of our model, we employed a comprehensive set of performance metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R2). MSE provides an average of the squared differences between the predicted and actual values, placing more emphasis on larger errors and making it particularly sensitive to outliers. MAE, on the other hand, calculates the average of the absolute differences, offers a clearer understanding of the error in the same unit as the target variable, and is less affected by the outliers. The RMSE, which is the square root of the MSE, provides a more interpretable measure of error, aiding direct comparison with the original data. Finally, R2 assesses goodness of fit by determining the proportion of variance in the dependent variable accounted for by the independent variables, with a perfect fit indicated by an R2 value of 1. These metrics collectively offer a robust evaluation of the performance of the proposed models for prediction tasks. Equations 4–6 present the formulae for the performance indices:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Yi - Yp)^2 \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Yi - Yp)^2}{n}} \tag{5}$$

$$MAE = \frac{1}{n} \times \sum\nolimits_{i=1}^{n} |Yi - Yp| \tag{6}$$

G. LIME FOR ENHANCING MODEL TRANSPARENCY

To enhance the interpretability of our prediction model, we integrated Local Interpretable Model-agnostic Explanations (LIME) into our analysis. Although several interpretability models are available, such as SHAP, Partial Dependence Plots (PDP), and traditional Feature Importance techniques, LIME was chosen because of its specific advantages in our study. The selection of LIME was based on its ability to provide local interpretability, which is crucial for understanding the influence on individual predictions in our deep-learning-based PV system performance model. Unlike SHAP, which can be computationally intensive and focused on the overall model behavior, LIME's model-agnostic approach allows us to seamlessly apply it to our complex deep -learning model. Additionally, its simplicity and intuitive explanations make it easier to communicate and comprehend the effects of various input features on predictions. By utilizing the LIME, we were able to gain deeper insights into the specific attributes driving our PV system performance forecasts, thereby enhancing the overall transparency and trustworthiness of our results [50].

IV. RESULTS AND DISCUSSION

The Results and Discussion section provides a detailed analysis of the findings of the study, focusing on the performance of the proposed Explainable Artificial Intelligence (XAI) model for predicting the Performance Ratio (PR) of a grid-connected photovoltaic (PV) system. Additionally, it discusses LIME analysis, which highlights the factors that positively and negatively impact the prediction model. This section also discusses the implications and significance of the study's findings.

A. MODEL EVALUATION

To evaluate the system performance, the PR was calculated on an hourly basis over the course of one year using operational data. The developed XAI model aims to interpret the ''black-box'' machine learning model, specifically the Feedforward Neural Network (FFNN), which serves as the base model for this study.

The FFNN model was trained using the training data described in the previous section, and its hyperparameters were carefully tuned to achieve maximum accuracy. The hyperparameters of the FFNN model play a critical role in determining its structure and optimization strategy. With two hidden layers, ReLU activation for the hidden layers and linear activation for the output layer, the model controls the flow of information and introduces nonlinearity. The number of hidden units in each layer (128 for the first hidden layer and 64 for the second) defines the complexity and capacity of the model to learn from data. In this study, five optimizers were employed to train the FFNN model: Stochastic Gradient Descent (SGD), Adam, Adagrad, RMSprop, and AdaDelta. The final test loss and iteration time were calculated for each optimizer. The scores for the iteration time and final test loss for all five optimizers are shown in Figure 7. The best-performing optimizer was selected based on the final test loss, providing insight into which optimization strategy was the most effective for predicting the PR of the PV system. Based on these results, the Adam optimizer emerged as the most effective choice for training the FFNN model to predict the PR of the PV systems. The combination of a low final test loss (0.0001) and reasonable iteration time (4.63 s) makes it a suitable optimizer for this application. The FFNN model

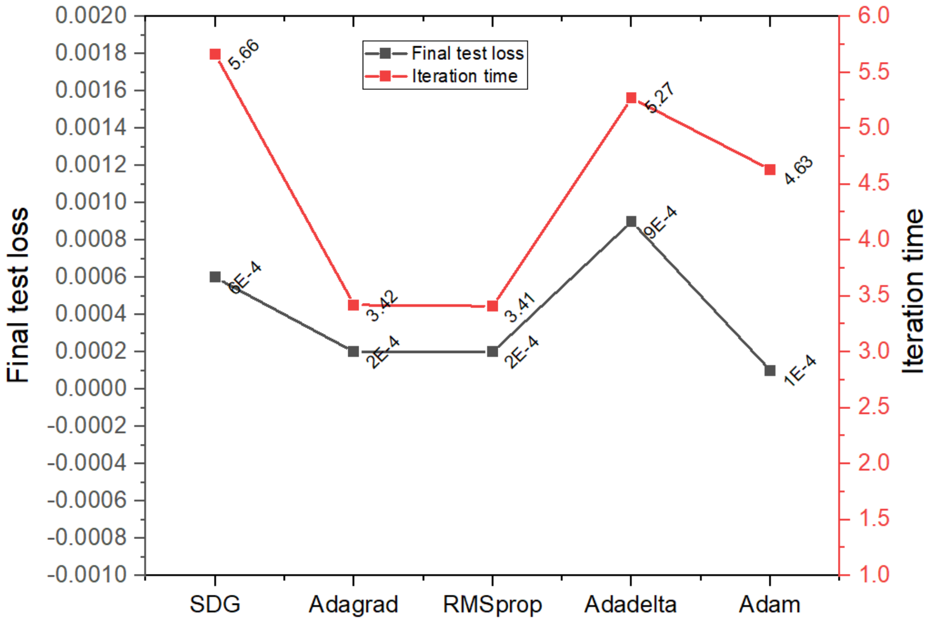

FIGURE 7. Final test loss score of each optimizer.

trained with Adam as the optimizer demonstrated an outstanding performance in predicting the PR of the PV system. The low MSE(0.0001), MAE(0.0036), and RMSE(.0078) values, coupled with the high R^2 (0.9965) score, confirmed the model's accuracy and ability to generalize well to unseen data.

These results validate the effectiveness of the FFNN model with the Adam optimizer for PR prediction in PV systems. Figure 8 illustrates the performance metrics and predicted PR outcomes resulting from optimal optimizer selection.

B. LIME ANALYSIS

The application of LIME analysis to our deep learning model provided a valuable layer of transparency and explainability for our findings. In Figures 9-DCLfiglabel12, we present the local explanations for two distinct days: the 7th and 13th hour, and the 8th and 12th hour of 1st January and 4th January respectively. Each figure illustrates three key components of the LIME results.

Prediction Probabilities: The left part of each figure displays the prediction probabilities for PR prediction focusing on a binary classification problem with ''PR'' and ''NOT PR'' classes. The ''Prediction probabilities'' section indicates the model's confidence scores for each class.

LIME Explanation: The middle part of the figures presents the LIME explanations of the selected features.

Feature Importance: The right side of the figures shows the original feature values ranging from 0 to 1, where 0 represents the minimal effect and 1 represents the highest effect.

Across four instances of lime analysis on January 1st and 4, 2021, the model consistently predicted the probability of ''PR'' values with varying degrees of confidence at different times. On January 1st, at 7:00 AM, the model predicted a ''PR'' with a probability of 0.83, closely matching the actual predicted value of 0.8265. The Lime analysis highlighted that features such as ''soiling loss factor,'' ''array voltage,'' ''generated energy,'' and ''array output'' positively influenced this prediction, with importance scores of 0.89, 0.74, 0.82, and 0.82, respectively. However, ambient temperature (−0.27) and inverter losses (−0.71) had a significant negative effect, indicating its decrease in influence towards a ''PR'' prediction. At 1:00 PM on the same day, the model predicted a ''PR'' with a probability of 0.81, aligning closely with the actual value of 0.8064. Similar features were observed, with ''soiling loss factor,'' array voltage, ''generated energy,'' and ''array output, which play significant roles, whereas inverter losses and ambient temperature have a negative influence. On January 4th, at 8:00 AM, the model predicted a ''PR'' with a probability of 0.88, closely matching the actual predicted value of 0.8794. The Lime analysis indicated strong positive impacts from the ''soiling loss factor,'' array voltage, ''generated energy,'' and ''array output,'' whereas ''ambient temperature'' had a less negative effect (−0.25). Finally, at 12:00 PM on the same day, the model predicted a ''PR'' with a probability of 0.72, closely matching the actual predicted value of 0.7187. The Lime analysis highlighted the positive effects of ''soiling loss factor,'' ''array voltage,'' and ''generated energy,'' while ''ambient temperature'' had a more significant negative impact (−0.52) where as inverter losses had relatively lesser (−0.38) negative impact. These Lime analyses collectively demonstrate how various features contribute to a model's predictions, with certain factors consistently showing importance in various instances.

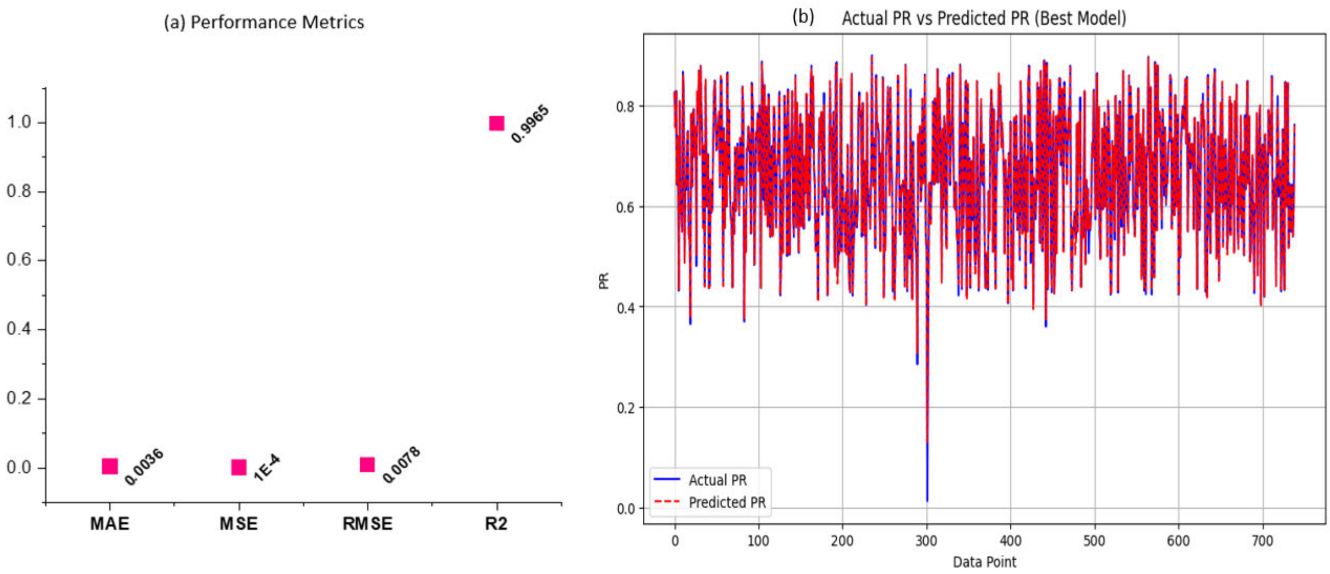

FIGURE 8. Performance metrics and predicted PR.

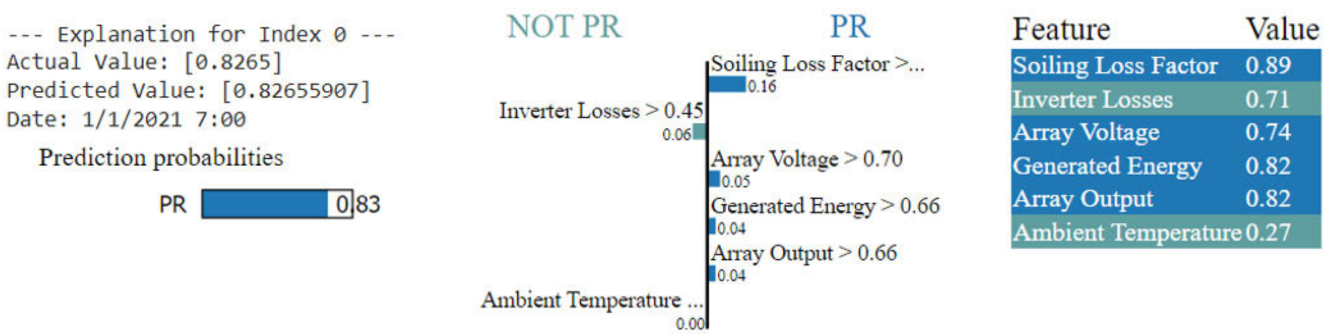

FIGURE 9. Lime analysis for hour 7 of a day.

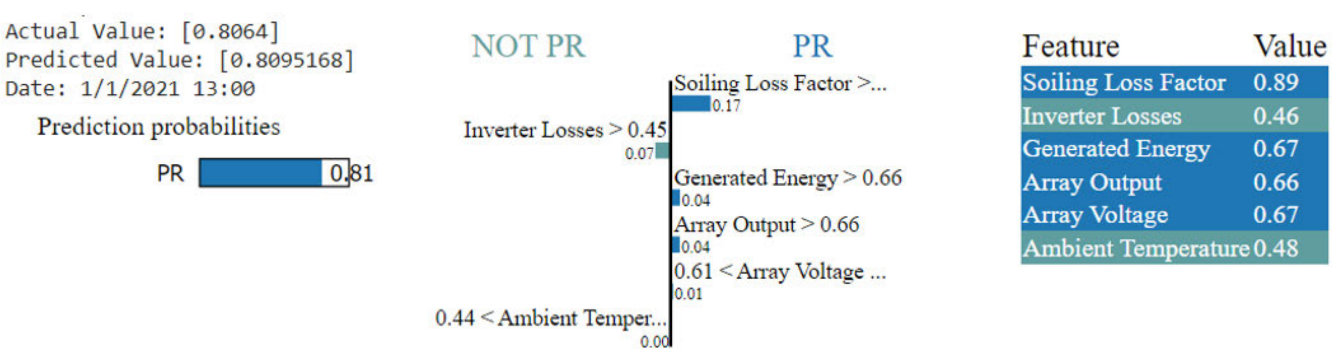

FIGURE 10. Lime Analysis for hour 13 of a day.

C. STUDY FINDINGS AND IMPLICATIONS

- This study introduced a deep learning model to predict the Performance Ratio (PR) of PV systems and provide interpretability through eXplainable Artificial Intelligence (XAI) techniques. XAI aids in understanding the significance and impact of different features on the output of the system. It assigns weights to the features, indicating their importance in determining the performance of the system and identifying the conditions that lead to significant performance degradation.
- The LIME explainer was used to analyze how individual features contributed to the PR. The results showed that

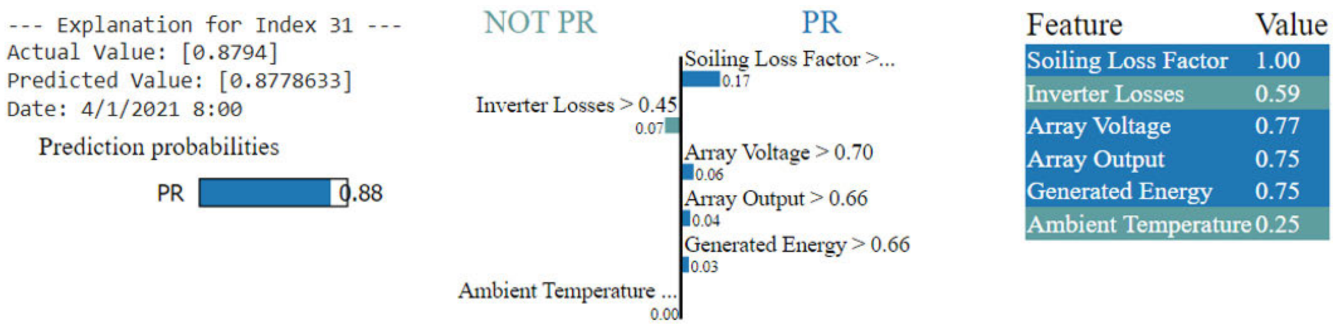

FIGURE 11. Lime analysis for hour 8 of a day.

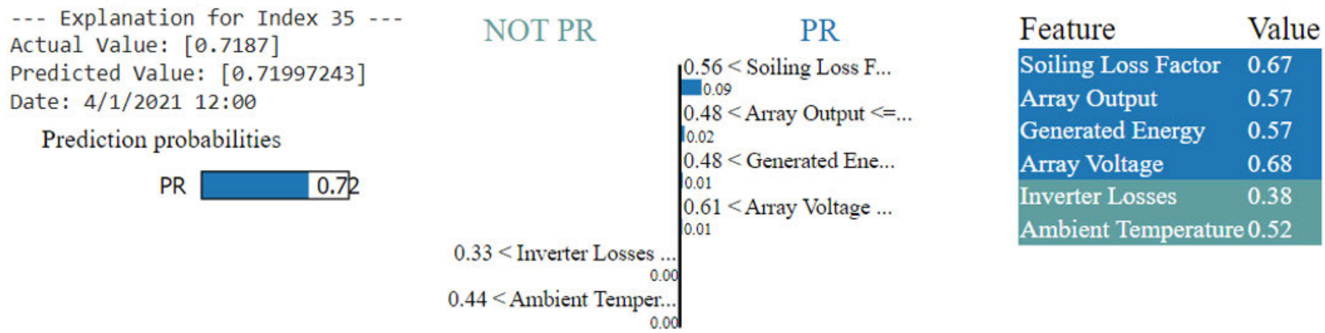

FIGURE 12. Lime analysis for hour 12 of a day.

factors such as the soiling loss factor, array voltage, array current, and generated energy had a positive impact, whereas the ambient temperature and inverter losses had a negative impact.

- These findings, consistent with previous research, emphasize that temperature and inverter losses negatively affect PR, whereas factors such as the soiling loss factor, array voltage, current, and generated energy positively impact the performance. This underlines the importance of using XAI methods to understand the relationships between the input variables and system performance.
- Identifying key features and their effects can lead to improved decision making and optimization of the performance of solar PV systems. The high weighting of the soiling loss factor, as indicated in this study, underscores the importance of regular cleaning schedules in areas prone to dust accumulation, which leads to enhanced performance. Additionally, scheduling the regular monitoring of inverters for faults can decrease losses and improve the system efficiency. Additionally, the negative impact of ambient temperature on the predictions aligns with the observations of seasonal variations influencing dust accumulation. Recognizing these seasonal patterns, as highlighted in Lime analysis, is crucial for developing adaptive cleaning schedules to maintain the long-term efficiency of PV systems. These insights can contribute to enhancing the overall system efficiency and performance management.

V. COMPARISON OF THE PROPOSED XAI MODEL WITH CONVENTIONAL DEEP LEARNING MODELS

To evaluate the accuracy and robustness of the model, the proposed XAI model was compared with several well-known conventional deep learning models, such as CNN, RNN, and LSTM, chosen for their distinct advantages and widespread use in various predictive modeling tasks. CNNs were selected for their ability to excel in spatial data interpretation, making them a strong baseline for performance comparisons. RNNs were included because of their proficiency in handling sequential data, which is critical for time-series analysis, such as PV system performance. LSTMs, an advanced variant of RNNs, were chosen for their capability to manage long-term dependencies, thus addressing some of the limitations of standard RNNs. The proposed Explainable Deep learning model demonstrated superior performance based on several metrics compared with the CNN, RNN, and LSTM models as shown in Figure 13. The Mean Squared Error (MSE) of the XAI model was 0.0001, indicating that it had the lowest average squared difference between the predicted and actual values among the models. The Mean Absolute Error (MAE) was 0.0036, indicating that it had the smallest average absolute difference between the predicted and actual values. The Root Mean Squared Error (RMSE) was 0.0078, signifying that the XAI model's predictions were closest to the actual values on average compared to those of the CNN, RNN, and LSTM models. The R2 score of 0.9965 for the XAI model highlights its exceptional explanatory power, which is nearly perfect for capturing variance in the data.

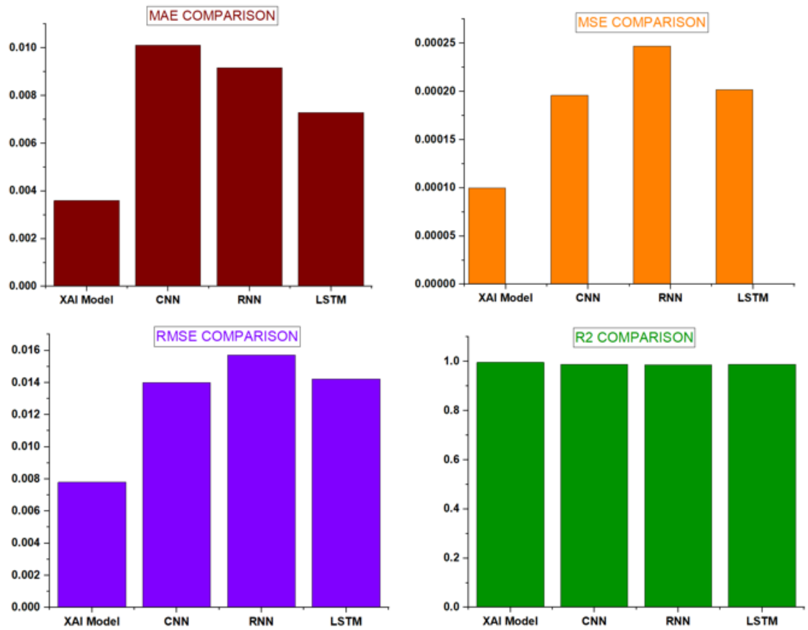

FIGURE 13. Performance metrics of the proposed and conventional deep learning models.

In comparison, the CNN, RNN, and LSTM models had higher MSE, MAE, and RMSE values and lower R2 scores, indicating that they were less accurate and less effective at explaining the variability in the PV system data. Specifically, the CNN model has an MSE of 0.000196, MAE of 0.010116, RMSE of 0.013988, and R2 score of 0.988642; the RNN model has an MSE of 0.000247, MAE of 0.009162, RMSE of 0.015715, and R2 score of 0.985665; and the LSTM model has an MSE of 0.000202, MAE of 0.007285, RMSE of 0.014224, and R2 score of 0.988256. Collectively, these metrics underscore the superior performance of the proposed XAI model. This performance enhancement of the proposed model is driven by optimized training using a variety of optimizers, such as Stochastic Gradient Descent (SGD), Adaptive Moment Estimation (Adam), Adaptive Gradient Algorithm (Adagrad), Root Mean Square Propagation (RMSprop), and Adadelta. This optimization ensured effective model convergence and improved the performance metrics. Additionally, feature selection via univariate linear regression significantly enhanced prediction accuracy, providing a clear understanding of the impact of each feature on the performance ratio (PR). The model's enhanced interpretability, achieved through univariate linear regression to analyze coefficient weights, further contributes to its ability to explain the factors that influence performance. In contrast, conventional models have limitations that hinder their performance on this task. RNNs often struggle with vanishing and exploding gradients, which makes them less effective for long-term dependencies and continuous data. CNNs, while excelling in spatial data interpretation, are not inherently designed for time-series data or continuous inputs, which are crucial for PV system performance assessment. Although LSTMs are better at handling long-term dependencies than RNNs, they can be computationally intensive and require a significant training time. Furthermore, without proper interpretability mechanisms, LSTM predictions can be perceived as black-box outputs, making it difficult to gain insight into the influencing factors. The proposed model not only excels in accuracy but also provides real-time, interpretable insights into the factors influencing PV system performance, which is essential for timely decision-making and system optimization. By integrating advanced optimization techniques and rigorous feature selection, this model has emerged as an invaluable tool for optimizing PV system performance and guiding decisions.

VI. CONCLUSION

Grid-connected solar PV system performance assessment is a formidable challenge, particularly in environments prone to dust, where factors such as temperature, soiling loss, and solar irradiation, along with technical parameters such as the array output and inverter losses, play significant roles. Therefore, understanding these factors and their impacts on performance requires careful consideration. The proposed work addresses this challenge by offering an AI-based solution for the evaluation of PV system performance, providing an alternative to labor-intensive human analysis. Using data from a 5 MW grid-connected solar PV system, we simulated the Performance Ratio based on a diverse set of environmental and technical parameters. The Performance Ratio was then predicted using an explainable deep learning model. With a remarkable regression score of 0.9968, the Feedforward Neural Network (FFNN) was chosen for explainability through LIME analysis. This study is versatile and can be applied to both grid-connected and off-grid systems by leveraging AI power. Our approach offers real-time insights supported by explanations provided by univariate linear regression, the ADAM optimizer, and LIME local surrogacy for each data instance in the dataset. Through feature identification, the proposed method aids the enhancement and management of solar PV system maintenance. The AI model achieved an R2-score of 0.9968, indicating an exceptional fit to the data with the ability to accurately predict PV system performance. This was reinforced by the low Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) values, thereby validating the effectiveness of the model in this critical domain. The proposed AI model not only offers high predictive accuracy, but also provides insights into the influential factors driving PV system performance, making it a valuable tool for solar energy management and maintenance. In the future, expanding the dataset to include a wider range of environmental conditions and technical parameters could enhance the robustness and applicability of the model. Furthermore, investigating the scalability of the model to larger PV systems and its adaptability to varying geographical locations would be beneficial for broader implementations. Continuous refinement of the AI model, perhaps through ensemble learning techniques or hybrid models, can lead to more accurate and reliable predictions, thereby improving the efficiency and longevity of the PV systems. Additionally, integrating the model into user-friendly software platforms can facilitate its adoption by technicians and engineers in the field, thereby streamlining the PV system evaluation and maintenance processes.

CONFLICT OF INTEREST DECLARATION

The authors declare no potential conflicts of interest.

REFERENCES

[1] G. Li, M. Li, R. Taylor, Y. Hao, G. Besagni, and C. N. Markides, "Solar energy utilisation: Current status and roll-out potential," *Appl. Thermal Eng.*, vol. 209, Jun. 2022, Art. no. 118285, doi: 10.1016/j.applthermaleng.2022.118285.

[2] K. Obaideen, M. Nooman AlMallahi, A. H. Alami, M. Ramadan, M. A. Abdelkareem, N. Shehata, and A. G. Olabi, "On the contribution of solar energy to sustainable developments goals: Case study on mohammed bin rashid Al maktoum solar park," *Int. J. Thermofluids*, vol. 12, Nov. 2021, Art. no. 100123, doi: 10.1016/j.ijft.2021.100123.

[3] M. K. H. Rabaia, M. A. Abdelkareem, E. T. Sayed, K. Elsaid, K.-J. Chae, T. Wilberforce, and A. G. Olabi, "Environmental impacts of solar energy systems: A review," *Sci. Total Environ.*, vol. 754, Feb. 2021, Art. no. 141989, doi: 10.1016/j.scitotenv.2020.141989.

[4] M. Doğan, S. Raikhan, N. Zhanar, and B. Gulbagda, "Analysis of dynamic connectedness relationships among clean energy, carbon emission allowance, and BIST indexes," *Sustainability*, vol. 15, no. 7, p. 6025, Mar. 2023, doi: 10.3390/su15076025.

[5] N. Saqib and G. Dincă, "Exploring the asymmetric impact of economic complexity, FDI, and green technology on carbon emissions: Policy stringency for clean-energy investing countries," *Geosci. Frontiers*, vol. 15, no. 4, Jul. 2024, Art. no. 101671, doi: 10.1016/j.gsf.2023.101671.

[6] G. E. Halkos and A. S. Tsirivis, "Electricity production and sustainable development: The role of renewable energy sources and specific socioeconomic factors," *Energies*, vol. 16, no. 2, p. 721, Jan. 2023, doi: 10.3390/en16020721.

[7] N. Ahmad, Y. G. Ghadi, M. Adnan, and M. Ali, "From smart grids to super smart grids: A roadmap for strategic demand management for next generation SAARC and European power infrastructure," *IEEE Access*, vol. 11, pp. 12303–12341, 2023, doi: 10.1109/ACCESS.2023.3241686.

[8] B. Han, Y. Zahraoui, M. Mubin, S. Mekhilef, M. Seyedmahmoudian, and A. Stojcevski, "Home energy management systems: A review of the concept, architecture, and scheduling strategies," *IEEE Access*, vol. 11, pp. 19999–20025, 2023, doi: 10.1109/ACCESS.2023.3248502.

[9] J. S. Saranyaa, "A comprehensive survey on the current trends in improvising the renewable energy incorporated global power system market," *IEEE Access*, vol. 11, pp. 24016–24038, 2023, doi: 10.1109/ACCESS.2023.3252574.

[10] M. J. Burke and J. C. Stephens, "Political power and renewable energy futures: A critical review," *Energy Res. Social Sci.*, vol. 35, pp. 78–93, Jan. 2018, doi: 10.1016/j.erss.2017.10.018.

[11] L. Alhmoud, "Why does the PV solar power plant operate ineffectively?" *Energies*, vol. 16, no. 10, p. 4074, May 2023, doi: 10.3390/en16104074.

[12] K. Obaideen, A. G. Olabi, Y. Al Swailmeen, N. Shehata, M. A. Abdelkareem, A. H. Alami, C. Rodriguez, and E. T. Sayed, "Solar energy: Applications, trends analysis, bibliometric analysis and research contribution to sustainable development goals (SDGs)," *Sustainability*, vol. 15, no. 2, p. 1418, Jan. 2023, doi: 10.3390/su15021418.

[13] A. K. Hamid, N. T. Mbungu, A. Elnady, R. C. Bansal, A. A. Ismail, and M. A. AlShabi, "A systematic review of grid-connected photovoltaic and photovoltaic/thermal systems: Benefits, challenges and mitigation," *Energy Environ.*, vol. 34, no. 7, pp. 2775–2814, Nov. 2023, doi: 10.1177/0958305x221117617.

[14] L. B. Bosman, W. D. Leon-Salas, W. Hutzel, and E. A. Soto, "PV system predictive maintenance: Challenges, current approaches, and opportunities," *Energies*, vol. 13, no. 6, p. 1398, Mar. 2020, doi: 10.3390/en13061398.

[15] M. Tawalbeh, A. Al-Othman, F. Kafiah, E. Abdelsalam, F. Almomani, and M. Alkasrawi, "Environmental impacts of solar photovoltaic systems: A critical review of recent progress and future outlook," *Sci. Total Environ.*, vol. 759, Mar. 2021, Art. no. 143528, doi: 10.1016/j.scitotenv.2020.143528.

[16] H. Salimi, H. A. D. Ashtiani, A. M. Lavasani, and R. Fazaeli, "Experimental analysis and modeling of weather condition effects on photovoltaic systems' performance: Tehran case study," *Energy Sources A, Recovery, Utilization, Environ. Effects*, vol. 42, 2020, doi: 10.1080/15567036.2020.1765902.

[17] H. A. Kazem, M. T. Chaichan, A. H. A. Al-Waeli, and A. Gholami, "A systematic review of solar photovoltaic energy systems design modelling, algorithms, and software," *Energy Sources, A, Recovery, Utilization, Environ. Effects*, vol. 44, no. 3, pp. 6709–6736, Sep. 2022, doi: 10.1080/15567036.2022.2100517.

[18] L. D. Jathar, S. Ganesan, U. Awasarmol, K. Nikam, K. Shahapurkar, M. E. M. Soudagar, H. Fayaz, A. S. El-Shafay, M. A. Kalam, S. Bouadila, S. Baddadi, V. Tirth, A. S. Nizami, S. S. Lam, and M. Rehan, "Comprehensive review of environmental factors influencing the performance of photovoltaic panels: Concern over emissions at various phases throughout the lifecycle," *Environ. Pollut.*, vol. 326, Jun. 2023, Art. no. 121474, doi: 10.1016/j.envpol.2023.121474.

[19] A. D. Dhass, N. Beemkumar, S. Harikrishnan, and H. M. Ali, "A review on factors influencing the mismatch losses in solar photovoltaic system," *Int. J. Photoenergy*, vol. 2022, pp. 1–27, Feb. 2022, doi: 10.1155/2022/2986004.

[20] N. Ketjoy, W. Chamsa-Ard, and P. Mensin, "Analysis of factors affecting efficiency of inverters: Case study grid-connected PV systems in lower northern region of Thailand," *Energy Rep.*, vol. 7, pp. 3857–3868, Nov. 2021, doi: 10.1016/j.egyr.2021.06.075.

[21] A. E. Cosgun and H. Demir, "The experimental study of dust effect on solar panel efficiency," *Politeknik Dergisi*, vol. 25, no. 4, pp. 1429–1434, Dec. 2022, doi: 10.2339/politeknik.903989.

[22] M. Katoch, K. Kumar, and V. Dahiya, "Dust accumulation and reduction in electrical performance of solar PV panels," *Mater. Today, Proc.*, vol. 46, pp. 6608–6612, Jan. 2021, doi: 10.1016/j.matpr.2021.04.082.

[23] M. Jaszczur, A. Koshti, W. Nawrot, and P. Sedor, "An investigation of the dust accumulation on photovoltaic panels," *Environ. Sci. Pollut. Res.*, vol. 27, no. 2, pp. 2001–2014, Jan. 2020, doi: 10.1007/s11356-019-06742-2.

[24] B. Zazoum, "Solar photovoltaic power prediction using different machine learning methods," *Energy Rep.*, vol. 8, pp. 19–25, Apr. 2022, doi: 10.1016/j.egyr.2021.11.183.

[25] G. M. Tina, C. Ventura, S. Ferlito, and S. De Vito, "A state-of-art-review on machine-learning based methods for PV," *Appl. Sci.*, vol. 11, no. 16, p. 7550, Aug. 2021, doi: 10.3390/app11167550.

[26] B. Carrera and K. Kim, "Comparison analysis of machine learning techniques for photovoltaic prediction using weather sensor data," *Sensors*, vol. 20, no. 11, p. 3129, Jun. 2020, doi: 10.3390/s20113129.

[27] U. Munawar and Z. Wang, "A framework of using machine learning approaches for short-term solar power forecasting," *J. Electr. Eng. Technol.*, vol. 15, no. 2, pp. 561–569, Mar. 2020, doi: 10.1007/s42835-020-00346-4.

[28] A. Gopi, P. Sharma, K. Sudhakar, W. K. Ngui, I. Kirpichnikova, and E. Cuce, "Weather impact on solar farm performance: A comparative analysis of machine learning techniques," *Sustainability*, vol. 15, no. 1, p. 439, Dec. 2022, doi: 10.3390/su15010439.

[29] S.-C. Lim, J.-H. Huh, S.-H. Hong, C.-Y. Park, and J.-C. Kim, "Solar power forecasting using CNN-LSTM hybrid model," *Energies*, vol. 15, no. 21, p. 8233, Nov. 2022, doi: 10.3390/en15218233.

[30] E. Roumpakias and T. Stamatelos, "Prediction of a grid-connected photovoltaic Park's output with artificial neural networks trained by actual performance data," *Appl. Sci.*, vol. 12, no. 13, p. 6458, Jun. 2022, doi: 10.3390/app12136458.

[31] M. Mohana, A. S. Saidi, S. Alelyani, M. J. Alshayeb, S. Basha, and A. E. Anqi, "Small-scale solar photovoltaic power prediction for residential load in Saudi Arabia using machine learning," *Energies*, vol. 14, no. 20, p. 6759, Oct. 2021, doi: 10.3390/en14206759.

[32] S. Al-Dahidi, O. Ayadi, M. Alrbai, and J. Adeeb, "Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction," *IEEE Access*, vol. 7, pp. 81741–81758, 2019, doi: 10.1109/ACCESS.2019.2923905.

[33] V. Shukla, A. Sant, P. Sharma, M. Nayak, and H. Khatri, "An explainable artificial intelligence based approach for the prediction of key performance indicators for 1 megawatt solar plant under local steppe climate conditions," *Eng. Appl. Artif. Intell.*, vol. 131, May 2024, Art. no. 107809, doi: 10.1016/j.engappai.2023.107809.

[34] A. Yar, M. Y. Arshad, F. Asghar, W. Amjad, F. Asghar, M. I. Hussain, G. H. Lee, and F. Mahmood, "Machine learning-based relative performance analysis of monocrystalline and polycrystalline grid-tied PV systems," *Int. J. Photoenergy*, vol. 2022, pp. 1–18, Aug. 2022, doi: 10.1155/2022/3186378.

[35] M. Abubakar, Y. Che, L. Ivascu, F. M. Almasoudi, and I. Jamil, "Performance analysis of energy production of large-scale solar plants based on artificial intelligence (machine learning) technique," *Processes*, vol. 10, no. 9, p. 1843, Sep. 2022, doi: 10.3390/pr10091843.

[36] M. E. H. Jed, R. Ihaddadene, N. Ihaddadene, C. E. E. Sidi, and M. El Bah, "Performance analysis of 954,809 kWp PV array of Sheikh Zayed solar power plant (Nouakchott, Mauritania)," *Renew. Energy Focus*, vol. 32, pp. 45–54, Mar. 2020, doi: 10.1016/j.ref.2019.11.002.

[37] G. G. Kim, J. H. Choi, S. Y. Park, B. G. Bhang, W. J. Nam, H. L. Cha, N. Park, and H.-K. Ahn, "Prediction model for PV performance with correlation analysis of environmental variables," *IEEE J. Photovolt.*, vol. 9, no. 3, pp. 832–841, May 2019, doi: 10.1109/JPHOTOV.2019.2898521.

[38] S. Bandong, E. Leksono, A. Purwarianti, and E. Joelianto, "Performance ratio estimation and prediction of solar power plants using machine learning to improve energy reliability," in *Proc. 6th Int. Conf. Instrum., Control, Autom. (ICA)*, Jul. 2019, pp. 36–41, doi: 10.1109/ICA.2019.8916687.

[39] S. Bandong, E. Joelianto, E. Leksono, A. Purwarianti, and I. N. Haq, "One-step and multi-step performance ratio prediction of solar power plants using time series ARIMA," *Internetwork. Indonesia J.*, vol. 12, pp. 39–45, Feb. 2020.

[40] A. Abulibdeh, "Analysis of urban heat island characteristics and mitigation strategies for eight arid and semi-arid Gulf region cities," *Environ. Earth Sci.*, vol. 80, no. 7, p. 259, Apr. 2021, doi: 10.1007/s12665-021-09540-7.

[41] A. Modaihsh, A. Ghoneim, F. Al-Barakah, M. Mahjoub, and M. Nadeem, "Characterizations of deposited dust fallout in Riyadh City, Saudi Arabia," *Polish J. Environ. Stud.*, vol. 26, no. 4, pp. 1599–1605, Jul. 2017, doi: 10.15244/pjoes/68565.

[42] A. Labban and A. Farahat, "Effect of major dust events on ambient temperature and solar irradiance components over Saudi Arabia," *Atmosphere*, vol. 14, no. 2, p. 408, Feb. 2023, doi: 10.3390/atmos14020408.

[43] H. Z. Al Garni, "The impact of soiling on PV module performance in Saudi Arabia," *Energies*, vol. 15, no. 21, p. 8033, Oct. 2022, doi: 10.3390/en15218033.

[44] B. Marion, J. Adelstein, K. Boyle, H. Hayden, B. Hammond, T. Fletcher, B. Canada, D. Narang, A. Kimber, L. Mitchell, G. Rich, and T. Townsend, "Performance parameters for grid-connected PV systems," in *Proc. 31st IEEE Photovoltaic Specialists Conf.*, Jan. 2005, pp. 1601–1606, doi: 10.1109/PVSC.2005.1488451.

[45] B.-C. Zhou, C.-Y. Han, and T.-D. Guo, "Convergence of stochastic gradient descent in deep neural network," *Acta Mathematicae Applicatae Sinica, English Ser.*, vol. 37, no. 1, pp. 126–136, Jan. 2021, doi: 10.1007/s10255-021-0991-2.

[46] E. Okewu, S. Misra, and F.-S. Lius, "Parameter tuning using adaptive moment estimation in deep learning neural networks," in *Proc. Int. Conf. Comput. Sci. Appl.*, 2020, pp. 261–272, doi: 10.1007/978-3-030-58817-5_20.

[47] Y. Gui, D. Li, and R. Fang, "A fast adaptive algorithm for training deep neural networks," *Int. J. Speech Technol.*, vol. 53, no. 4, pp. 4099–4108, Feb. 2023, doi: 10.1007/s10489-022-03629-7.

[48] Y.-L. Peng and W.-P. Lee, "Practical guidelines for resolving the loss divergence caused by the root-mean-squared propagation optimizer," *Appl. Soft Comput.*, vol. 153, Mar. 2024, Art. no. 111335, doi: 10.1016/j.asoc.2024.111335.

[49] S. Sveleba, I. Katerynchuk, I. Kunyo, V. Brygilevych, V. Kotsun, and Y. Sidelnyk, "Dynamics of the learning process of a multilayer neural network when using the AdaDelta optimization method," in *Proc. 24th Int. Conf. Comput. Problems Electr. Eng. (CPEE)*, Sep. 2023, pp. 1–4, doi: 10.1109/cpee59623.2023.10285140.

[50] M. Kuzlu, U. Cali, V. Sharma, and Ö. Güler, "Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools," *IEEE Access*, vol. 8, pp. 187814–187823, 2020, doi: 10.1109/ACCESS.2020.3031477.

IMAD HASSAN received the bachelor's degree in electrical engineering from the University of Engineering and Technology, Peshawar, and the master's degree in electrical engineering from Staffordshire University, U.K. He is currently pursuing the Ph.D. degree in electrical engineering with the University of Kuala Lumpur, Malaysia. He is also a Lecturer with the Energy Engineering Technology Department. He holds professional electrical engineering status from Saudi Council of Engineers. His research interests include power electronics, renewable energy, power system analysis and control, power generation, clean energy, and environmental protection.

IBRAHIM ALHAMROUNI received the B.Eng. degree in electrical engineering from Elmergib University, Al-Khums, Libya, in 2008, the M.Sc. degree in electrical power engineering from Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor, Malaysia, in 2011, and the Ph.D. degree from the Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), in 2015. He is currently a Senior Lecturer and a Researcher with British Malaysian Institute, University of Kuala Lumpur, Malaysia. His research interests include power system planning and operation, the deregulation and restructuring of power systems, and power line maintenance. He is also involved in microgrid technologies and the applications of power electronics in power systems.

ZAHRAOUI YOUNES received the B.S. and M.S. degrees in electrical engineering from the Faculty of Hydrocarbons and Chemistry (INH), University of Boumerdes, Algeria, in 2015 and 2017, respectively, and the Ph.D. degree from British Malaysian Institute, Universiti Kuala Lumpur, Malaysia, in 2022. He is currently a Postdoctoral Researcher with NTNU, Norway. His research interests include motor drives, autonomous technologies, microgrid controls, energy management systems, hybrid renewable energy systems, microgrid systems, energy efficiency, energy markets, artificial intelligence applications, and electric vehicle.

NURUL HANIS AZHAN received the B.Eng. and M.Eng. degrees in electrical and electronics engineering from Tokai University, Japan, in 2012 and 2014, respectively, and the Engineering degree from the Faculty of Electrical and Electronics Engineering, Tokai University, in 2017. She is certified as the Energy Manager. She is currently the Head of the Electrical Section and a Senior Lecturer and a Researcher with British Malaysian Institute, University of Kuala Lumpur, Malaysia. Her research interests include material characterization, thin-film and nanotechnology, thin-film fabrication, solar power generation, renewable energy, and nanomaterials.

SAAD MEKHILEF (Fellow, IEEE) is currently a Distinguished Professor with the School of Science, Computing, and Engineering Technologies, Swinburne University of Technology, Australia, and a honorary Professor with the Department of Electrical Engineering, University of Malaya. He has authored and co-authored more than 600 publications in academic journals and proceedings, five books with more than 49 000 citations, and more than 80 Ph.D. students who graduated under his supervision. He is actively involved in industrial consultancy for major corporations in power electronics and renewable-energy projects. His research interests include power conversion techniques, power converter control, maximum power point tracking (MPPT), renewable energy, and energy efficiency. He is an IET Fellow. He serves as an Editorial Board Member for many top journals, such as IEEE Transactions on Power Electronics, IEEE Open Journal of Industrial Electronics Society, *IET Renewable Power Generation*, *Journal of Power Electronics*, and *International Journal of Circuit Theory and Applications*. He has been listed by Thomson Reuters (Clarivate Analytics) as one of the world's most cited (top 1%) engineering researchers.

MEHDI SEYEDMAHMOUDIAN received the B.Sc., M.Eng., and Ph.D. degrees in electrical and electronics engineering. He was the Deputy Director of the CADET-VR Laboratory, Deakin University. He is currently the Discipline Director of the Electrical and Electronics, School of Science Computing and Engineering Technologies, and the Director of the Siemens Swinburne Energy Transition Hub, Swinburne University of Technology, Australia. He strongly believes in industry-engaged research initiatives and has established several key research partnerships with leading external industries and research centers during his academic career. He has been the author/co-author of $\sim$180 research articles, including more than 200 highly ranked journal articles, cited more than 7000 times in prestigious databases, including *Nature*. He has been a chief/partner investigator around \$21m funding application submission, including $\sim$\$10m successfully funded projects. He is among the top 2% of researchers in the world, based on his research performance within the electrical engineering and energy disciplines. His research interests include renewable energy systems, microgrids, smart grids, and AI-integrated energy systems. In addition, he serves as an editor/assimilate editor for prestigious academic journals and invited keynote speaker to numerous national and international energy-related conferences and symposiums.

ALEX STOJCEVSKI received the B.Eng. degree in electrical and electronics engineering, the M.Eng. degree in engineering and artificial intelligence, and the Ph.D. degree in electrical engineering from Victoria University in 1998, 2000, and 2004, respectively. To date, he has authored over 350 papers published in international journals and conference proceedings. His research focuses on microgrids, renewable energy sources, power converters, and power systems. He currently holds the position of Pro Vice Chancellor and President at Curtin Singapore.