Energy 302 (2024) 131709

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Solar power forecasting using domain knowledge

Rakesh Mondal *, Surajit Kr Roy, Chandan Giri

Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India

ARTICLE INFO

Keywords:
Solar power prediction
Machine learning
Deep learning
PV
Energy

ABSTRACT

Integrating renewable energy into the existing energy market is crucial. Solar power forecasting is essential since it depends on weather parameters and must integrate with the central grid to use the produced solar power effectively. Contemporary studies indicate that machine learning has the potential to predict the future generation of solar energy based on past data. This research demonstrates a broad range of solar power forecasting, combining the one-year time series solar power generation data, solar panel physical features, and weather information with the help of machine learning and deep learning tools with domain knowledge. The dataset is curated and preprocessed. We propose a deep learning ensemble model based on BI-LSTM. The proposed model can forecast well regardless of geographical position and is able to predict both short-term and long-term time horizons. We compared the results of the proposed model with the existing dataset and multiple standard deep learning models and found that our model produced better performance than traditional models. We also validated our model using different solar plants in Durgapur, India. For long-term forecasting, our model also outperformed the base model.

1. Introduction

Since civilization began, we have exploited Mother Nature in many ways to meet our regular needs. As we modernize our society proportionally, our energy requirements also increase, and fossil fuels fulfil most of the energy requirements, which are very limited in quantity and our environment, which continuously deteriorates due to excessive use of fossil fuels. It is evident that the use of conventional energy negatively impacts our environment and society. We are rapidly deploying sustainable energy in our central energy transmission grid to overcome the above mentioned issues in the last few years. Solar power is one of the most important sustainable energy sources since it is easily available and abundant in quantity. According to information from the International Energy Organization, by 2027, the capacity-wise contribution of photovoltaic solar power will beat coal and become a significant energy source; at present, photovoltaic solar power has only a 12.8% contribution to global energy supply. Fig. 1 demonstrates the estimated contribution of the worldwide energy supply of different energy sources. So, there is ample opportunity to work in the photovoltaic solar energy sector.

However, when putting solar energy into practice and employing it as a significant source of energy resources, there are also some difficulties. This is because there is uncertainty in electricity production as it highly depends on weather conditions. To use solar energy with maximum efficiency, we must connect the solar plant to the central electricity transmission grid. To deal with the uncertainty and conveniently transmit electricity through the grid, we need to forecast solar power production precisely in a particular solar plant in advance [2]. We can generalize the solar power forecasting technique mainly on three categories: physical model, statistical model and machine learning model. The physical forecasting model can be subdivided into numerical weather prediction(WNP) [3], Total-sky Image(TSI) [4], and satellite image [5], and these physical models depend on the relation between the solar irradiance and the law of physics on the environment [6]. On the other hand, the statistical models like Grey theory [7], Autoregressive [8], Markov chain [9] and Regression model [10]are used to predict solar power. These Statistical models rely on the previous historical data to predict the future time series. Machine learning models also use historical data and forecast future results, and they can map straight from input to output and efficiently extract complex nonlinear characteristics with large dimensionality [11]. There are several machine learning models; for example, linear regression [12], support vector machine(SVM) [13], and Artificial neural network (ANN) [14]. For Time series data, people have commonly used ANN for a long time [15,16]. Still, with the development of other deep learning models, people intended to use the recurrent neural network and another advanced model like gated recurrent neural network.

* Corresponding author.
E-mail addresses: rakeshmondal1991@gmail.com (R. Mondal), suraroy@gmail.com (S.K. Roy), chandangiri@gmail.com (C. Giri).

https://doi.org/10.1016/j.energy.2024.131709
Received 19 January 2024; Received in revised form 19 March 2024; Accepted 18 May 2024
Available online 22 May 2024

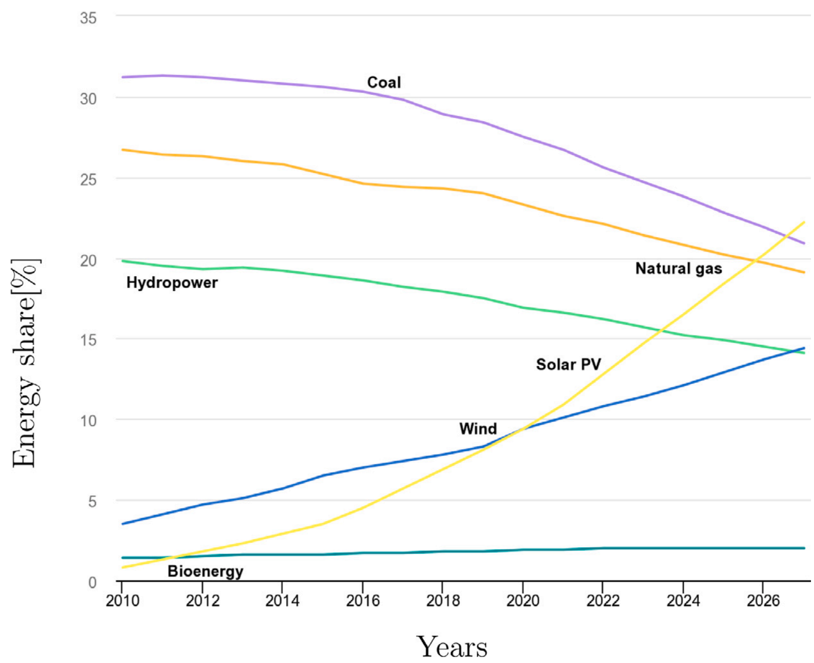


Fig. 1. Contribution of energy distribution over time [1].

Table 1
Result of different ML techniques with different time.

Previous work	Used models	Time span	Evaluation metrics	Obtained result
Yuchi et al. [22]	CNN	15 Min	RMSE	2.51
Rana et al. [29]	ANN, SVR	15 Min	MAE	67.11, 72.89
Chu et al. [18]	Hybrid	15 Min	MAE	51.8
Shi et al. [21]	SVM	24 h-ahead	MRE	8.64
Khan et al. [19]	Ensemble	15 Min	RMSE, MAE	0.74, 0.47
Pombo et al. [30]	Hybrid	24 h-ahead	RMSE	8.06

1.1. Machine learning (ML) based forecasting model

There are multiple options for machine learning algorithms and techniques, so they are also applicable to forecasting techniques [17–22]. In Table 1, we mention some previous works with different forecasting techniques. There is no single forecasting technique, and its result varies with the paradigm and dataset; observing different results on different techniques with a single dataset is very common. Sometimes, we can see that the same algorithm with the same dataset gives different results only because of different model inner parameter settings. When we talk about the machine learning model, a well-trained model can also provide different results when we slightly change input data. The result of machine learning-based forecasting also varied with the time frame; some models can provide good accuracy when we predict for a short timeframe; on the other hand, the model fails to produce acceptable predictions within an extended time frame. Several research works [23,24] deal with the shortcomings of machine learning-based forecasting, and we identified some of the most suitable approaches to predicting solar power more accurately using machine learning algorithms. Out of these algorithms, Random Forest [25], XGBoost [26] provides promising results compared to other models. The ensemble approaches show more stability than individual models [27]. Munawar et al. [28]created an environment to assess several feature selection techniques and machine learning models for short-term PV power forecasting. They admitted that the XGBoost approach performs better than different machine learning techniques.

1.2. Deep learning (DL) based forecasting

Due to the new paradigm shifts towards deep learning based forecasting techniques, deep learning models can provide more accurate results compared to simple machine learning techniques, researchers are using deep learning models, and it is evident that the Deep learning models are more reliable when slight changes in the input dataset exist. There are a number of deep learning models, from simple Artificial neural networks (ANN) [31] to complex Long short-term memory (LSTM) [32], used in renewable energy forecasting. Abuella et al. [12] Proposed an ANN based model to predict solar power generation. They selected the most important features based on sensitivity analysis of the features. Florencia et al. [33] Proposed a dynamic ANN approach to predict solar power one hour ahead of a PV plant, which is able to produce good accuracy.

1.3. Ensemble-based forecasting model

The ensemble of deep learning models is one step more advanced than the simple deep learning model [34]. A single or multiple ML and DL models are used to produce combined results. An ample amount of research [35] already exists, which shows that the ensemble based model is more stable than a single-layer model. Pierro et al. [36] introduced an ensemble with a combination of ANN, SVM, and statical models. They compared their multi-model ensemble with other data-driven prediction models and found their model outperformed others. Kumari et al. [37] proposed an ensemble with a combination of two base models named Extreme Gradient Boosting Forest and Deep Neural Network(XGBF-DNN) and compared their model with some previous standard models and found that their ensemble model is more stable than other base models.

1.4. Limitations of current existing techniques

This section addresses the shortcomings of the PV power prediction model. Most of the existing research on solar power generation prediction considers the task to be a simple regression task. They tried to implement various ML and DL-based individual and ensemble models to predict solar power from the weather parameters and previous generation data. The dynamic behaviour of solar power generation makes it more complex to predict accurately using only the computational power. We have worked on the input data and prediction models to eliminate these issues. Concerning input data, most of the prediction models use weather parameters and solar generation data as input to train the model. It is also an established fact, and authors have shown in [32,38] that solar power generation also depends on the physical characteristics of the solar panels like the number of cells in a solar panel, the maximum working temperature of the solar panel, the material type of the solar panel, ambience temperature etc., None of the existing techniques have considered these parameters during prediction of the solar power. Hence, in this work, we have included these features that enhance the accuracy of predicting solar power. This work considers an ensemble with BI-LSTM [39] as the base model to predict solar power generation.

1.5. Novelty of the work

The main contribution of this work is summarized below:

- First of all, we prepared our dataset by combining weather parameters and solar power generation data .
- We have uniquely augmented our dataset by combining meteorological data and physical characteristics of solar panels deployed in the respective solar plant.
- We propose an ensemble model with BI-LSTM as the base model, which can predict the future solar power generation of a specific solar plant on both short and long horizons regardless of the geographical position of the solar plant.
- We have evaluated the model with respect to different error metrics and found and analysed the performances with the base models; we have also validated our model with another published work data set.
- We also compared the model results with some individual standards models and found our model outperforms others.
  The rest of the work is organized as follows: Section 2 describes the problem statement. Section 3 explains the methodology used to prepare the dataset and model development. In Section 4, we described the result of our model with different evaluation metrics. Finally, Section 5 concludes the work and provides a direction for the future.

2. Problem statement

Provided a time series of photovoltaic solar plant outputs and ambient data up to time $t : x = x_1, x_2, x_3 \dots x_t$, the target is to predict the multiple of following n future values (i.e., $x_{t+1}, x_{t+2}, \dots, x_{t+n}$) of the time series x. In our experiment, we use a time series data set of 15-min intervals and are trying to forecast the next 15 min to 1 h in the short range and one day to seven days in the long range.

3. Proposed methodology

The flow of our proposed methodology is presented in Fig. 2. In the very first step, we have collected weather information and photovoltaic plant generation data. Then, we concatenated these data and expanded the dataset with the knowledge of PV panel physical characteristics. We prepare the dataset accordingly to fit in our proposed deep learning ensemble module. We split the dataset accordingly and trained with the proposed model. Here, we proposed a BI-LSTM ensemble model to predict a solar plant's future solar power generation. Since BI-LSTM has the capacity to retain information in both forward and backward directions, it is very suitable to work with time series data. Then, we evaluate the performances of our model and predict the solar power generation from various other PV plants also.

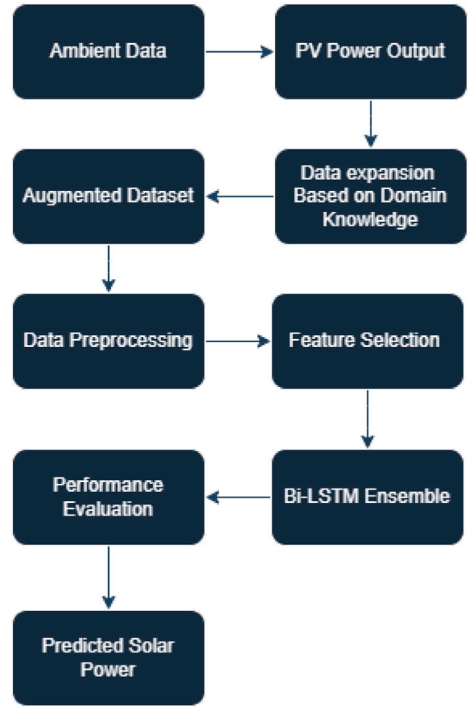


Fig. 2. Proposed methodology.

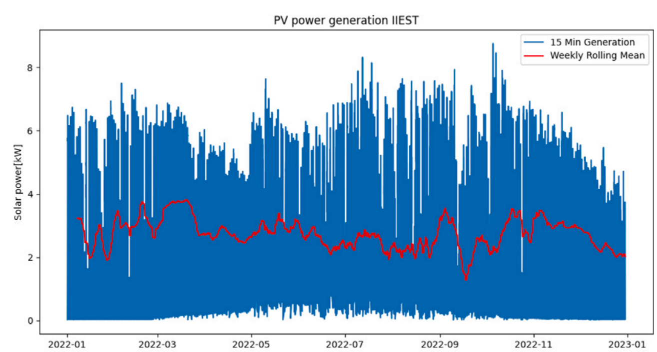


Fig. 3. PV power generation of IIEST solar plant.

3.1. Data description

3.1.1. PV data

In this work, we used solar data from the Indian Institute of Engineering Science and Technology (IIEST) Shibpur rooftop solar plant. The plant's capacity is 10 kW, which contains thirty-eight solar panels, each with a max power capacity of 260 W. We have collected data with a frequency of fifteen minutes from 1 January 2022 to 31 December 2022. Initially, we keep a record of data with different time frames based on seasons. We have taken an 11-hour window (6:00 a.m. to 5:00 p.m.) as most of the data outside this window is near zero due to very low solar irradiance. Fig. 3 shows the PV plant produced energy with its weekly rounded-off mean values.

3.1.2. Meteorological data

In this work, we also have used meteorological data from the Central Pollution Control Board of India (CPCBI), Kolkata [40] data collection centre near IIEST, Shibpur.

3.2. Data preparation

Data preparation is essential to use in any machine learning or deep learning model. In this section, we describe in detail the data pre-processing action taken care of. First, we checked for null values from the data set and replaced those with a mean value of the previous and following three days of the same timestamp. We also checked for values that were larger than the max capacity of the plant and, if found, replaced them with the same strategy as taken for null values. Missing values are checked and filled with the same approach as mentioned earlier. For detecting outliers, we took two measures of negative values for generated data and the values larger than the max capacity of the plant. With some visualization observations, we have taken care of those outliers. Likewise, other features are also taken care of to improve the quality of the dataset.

3.3. Expansion of dataset

We expand our dataset based on the knowledge of the physical characteristics of solar panels used in the considered solar plant and the essential weather parameters like ambient temperature and solar radiation. Preparing the dataset for a machine learning model by introducing new features based on domain knowledge like heat generated in a single cell of a solar panel, electricity produced in a single cell of a solar panel, and total heat generation in a solar panel. Using domain knowledge is a beneficial trending approach for preparing various machine learning models. Here, we generated new features for forecasting the production of solar energy from the background knowledge of scientific relations of the production of solar energy, weather parameters and solar panel characteristics. The work of King et al. [38] inspired us to expand the dataset by the generic "Power temperature coefficient model", which needs a minimal amount of information and provides a good amount of accuracy. The essential requirement for the asserted model is the Plane of Array Irradiance or E_{POA} [W/m^2], which is mentioned in Eq. (1) where; E_b and E_d stand for beam and defuse irradiance, and E_g denotes the irradiance reflected on the surroundings. In this work, we consider Solar radiation data E_{POA} and further calculations were done as per the predefined formulas as follows:

$$E_{POA} = E_b + E_d + E_g \tag{1}$$

$$T_{module} = T_a + E_{POA} exp(a + bW_s) \tag{2}$$

$$T_{cell} = T_{module} + \frac{E_{STC}}{E_{POA}} \Delta T \tag{3}$$

$$P_{DC,Panel} = \frac{E_{STC}}{E_{POA}} P_{mp,STC} [1 + \gamma_{mp}(T_c + T_{STC})] \tag{4}$$

$$P_{DC} = P_{DC,Panel} N_s N_p \tag{5}$$

The temperature of the module(T_{module}) in [°C] can be estimated using the Eq. (2). In this equation, T_a stands for the ambient temperature, and W_s stands for wind speed a and b are two coefficients related to module construction parameters and materials as described in [38]. The necessary characteristics of the used photovoltaic panel are presented in Table 2. Eq. (3) derives the induced temperature of individual cells of solar panels. Here E_{STC} determines the irradiance in standard condition(1000 W/m^2), and ΔT is taken from the PV module datasheet, measuring the temperature gap between the module and individual cell. The DC power of individual panels is calculated using Eq. (4); in this equation, $P_{mp,STC}$, T_{STC} and γ_{mp} stand for the panel's peak power [W], the temperature measured in standard test condition and the normalized temperature coefficient of peak power, respectively. Total power is calculated in Eq. (5) by multiplying the DC power of the panel by all parallel and series-connected PV modules, which are denoted by N_p, N_s.

Table 2
Photovoltaic system information.

Parameters	Values
a	−3.56
b	−0.0750
Δ T	3
$P_{mp,STC}$	260 W
$\gamma_m p$	−0.037
N_p	2
N_s	19
P_{ACmax}	10 kW
Inverter	—ABB PVI-10.0-TL-OUTD-5Y [41]

Table 3
Statical information of dataset.

Unit	SR	AT	WS	WD	RH	BP	SP
max	1.172	41.59	10.07	359.99	99.360	10.950	8.75
min	0.0150	11.77	0.020	0.010	17.07	9.095	0.0148
mean	0.35	28.09	3.051	203.29	74.69	9.91	2.73
std	0.272	5.305	1.80	103.35	24.01	0.204	1.96

3.4. Analyses of the dataset

Our final dataset has fifteen features except the time information, which is taken as the index of our dataset. Out of fifteen features, we have six weather information: solar radiation (SR), average temperature (AT), relative humidity (RH), wind speed (WS), wind direction (WD), and air pressure (BP). On the other hand, we have other information like solar power (SP) produced as kilowatt-hours and the amount of current flow from the inverter of the IIEST solar plant. As mentioned earlier, features like cell temperature, module temperature, power produced in a single temperature and total produced power are augmented. In Fig. 4, we demonstrate the correlation between the features, which clearly shows some of the features have very low correlation values, but we take them in our final dataset because by including them, we are getting more accurate results. A statistical description of the dataset is given in Table 3.

3.5. Data scaling

The collected dataset is very noisy in terms of solar power generation value as it fluctuates very high there. So, we normalized the data based on Scaling.

3.6. Training and testing

We split our entire dataset into two parts. We took 75% data for training and the other 25% for testing purposes.

3.7. Performance measures

Four widely used standard error metrics in the literature are used to estimate the model errors in order to assess the performance of the proposed model: the coefficient of determination mean absolute error (MSE), $R2\_score$, root mean square error (RMSE), and mean absolute error (MAE) [42–45] here we defined all the performance evaluation metrics used in this work:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

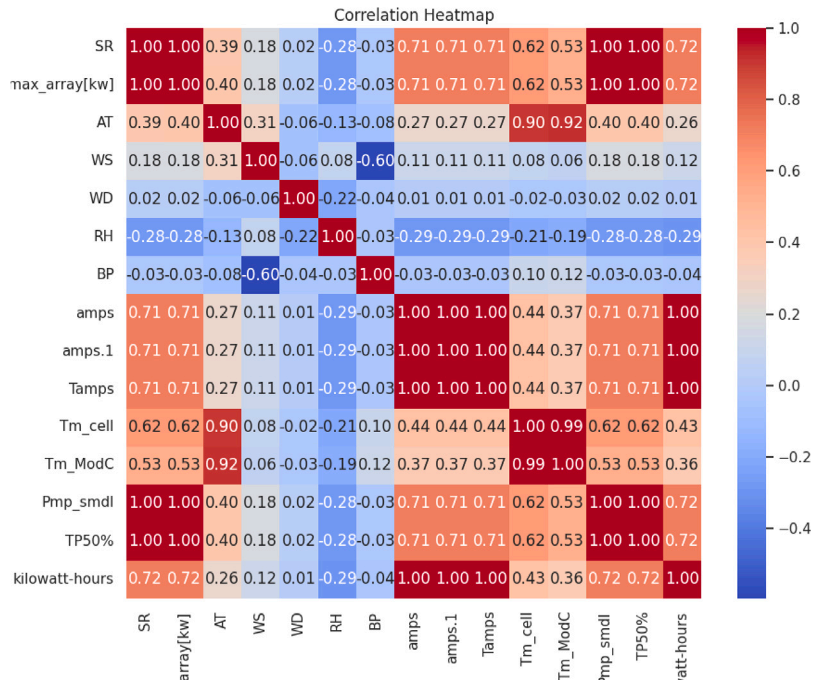


Fig. 4. Correlation matrix of the dataset.

$$R2\_Score = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \overline{y})^2}$$

Here in the above formulas, y represents the actual value, $\hat{y}$ represents the predicted value, $\overline{y}$, represents the mean value of the sample and n corresponds to the number of observations in the dataset.

3.8. LSTM

Long short term memory network (LSTM) model [46] is an improved version of recurrent neural network (RNN) [47]. LSTM resolves the vanishing gradient problem of RNN, and this module is capable of working with long-term sequential information. A basic internal structure of the LSTM module is depicted in Fig. 5. The most important part of LSTM is its cell states, the horizontal line mentioned at the top of the diagram. The LSTM has a gate mechanism which enables it to persist or forget information about the cell state. There are three gates, namely the forget gate f_t, input gate i_t, and output gate o_t.

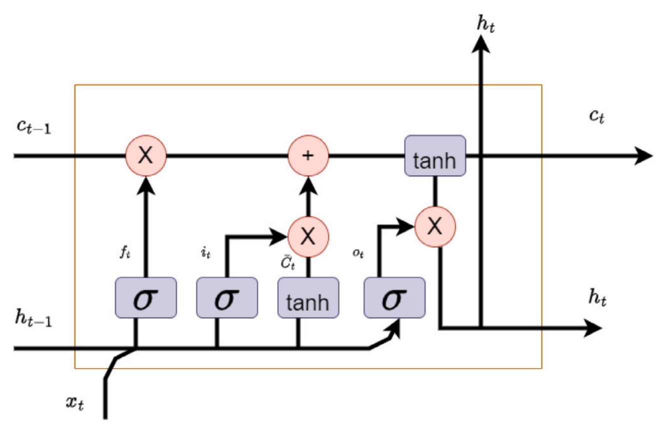


Fig. 5. Internal structure of LSTM.

3.9. Bidirectional LSTM (BI-LSTM)

BI-LSTM [39] is one step more advanced than generic LSTM. BI-LSTM has the ability to store and process information in both forward and backward directions, and this feature makes it more suitable to work with time series data. This contains two LSTM layers. One carries information forward and another backwards. A simple diagram of the BI-LSTM module is given in Fig. 6. The BI-LSTM module takes input as a sequence, and this information passes through both forward and backward LSTM chains. During the forward pass, the LSTM layer captures information from the past (previous time steps), while during the backward pass, it captures information from the future (following time steps). This bidirectional processing allows the model to capture long-term dependencies in the input sequence effectively. One copy of this processed information from both of these cells is concatenated and passes through a sigmoid activation function, which produces the output of that state, and another copy is passed to the following LSTM cells.

3.10. BI-LSTM ensemble

We have created a BI-LSTM ensemble model to forecast solar power. An easy-to-understand block diagram of our ensemble model is given in Fig. 7. Our model has an input layer, then two BI-LSTM layers, an average layer and final output. Our model takes input with the input shape of the dataset. Two BI-LSTM layers have different initial weights but are trained with the same input. A dropout layer follows each layer's first bidirectional LSTM module; again, a bidirectional LSTM module takes the first module output as input, then it passes through a dropout layer, and the output of this layer is fed to another bidirectional LSTM module, and finally the output passes through a dense layer. Likewise, another worked, and the final output is produced by averaging the output of both layers. We have used LeakyReLu in the LSTM module and the ReLU activation function in the dense layer.

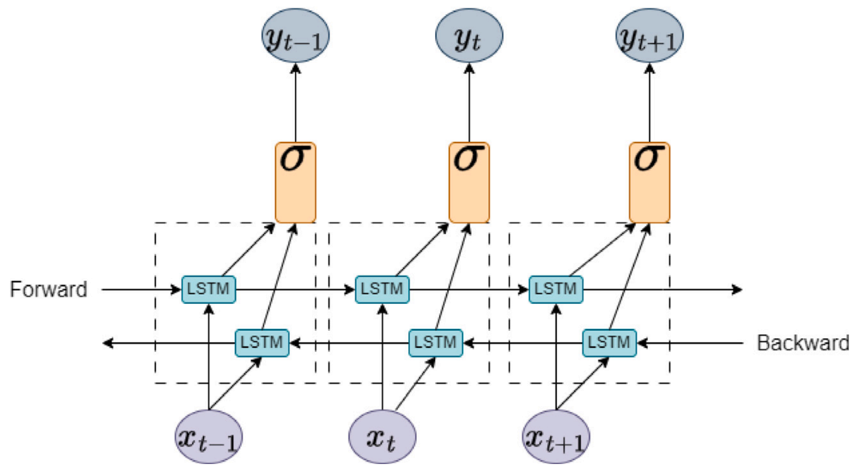


Fig. 6. Structure of BI-LSTM.

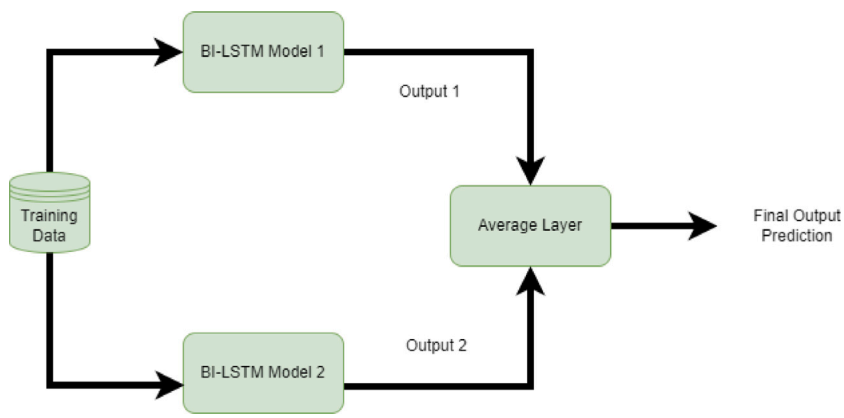


Fig. 7. Description of ensemble.

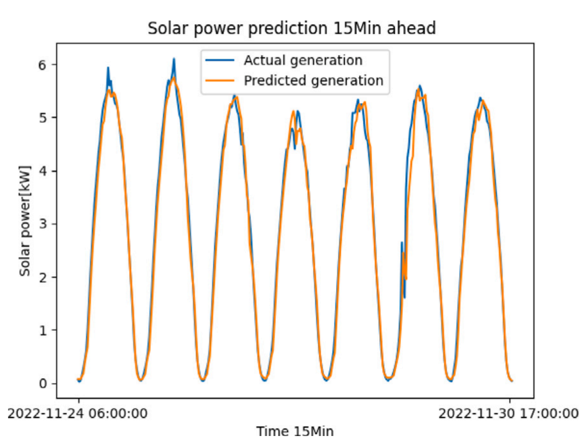


Fig. 8. Solar power prediction.

3.11. *Development setup*

We have simulated the experiment using the Google Colab platform and Python programming language. We used Keras to build and train our ensemble with the help of TensorFlow, scikit-learn, Numpy and Matplotlib libraries. We curate the dataset used in this work, and the solar power generation data is collected from IIEST, Shibpur solar plant, and the weather-related data is collected from CPCBI, Kolkata [40], the solar panel characteristics are collected from manufacturer datasheet.

Table 4
Comparison of model performances.

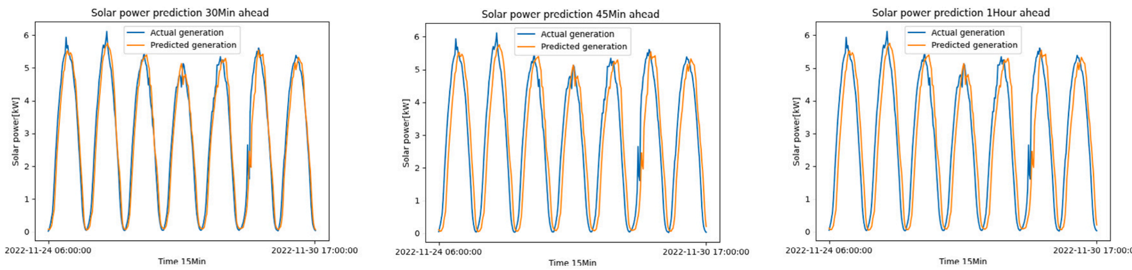
Evaluation metrics	Proposed model	LSTM	GRU	RNN	ELM	SVR
MSE	**0.41**	0.63	0.83	0.83	0.56	0.60
RMSE	**0.64**	0.8	0.91	0.91	0.75	0.77
MAE	**0.36**	0.45	0.51	0.51	0.50	0.52
R2_Score	**0.90**	0.84	0.78	0.78	0.89	0.85

The dataset used in this work is not publicly available, but it can be obtained on request.

4. Result and discussion

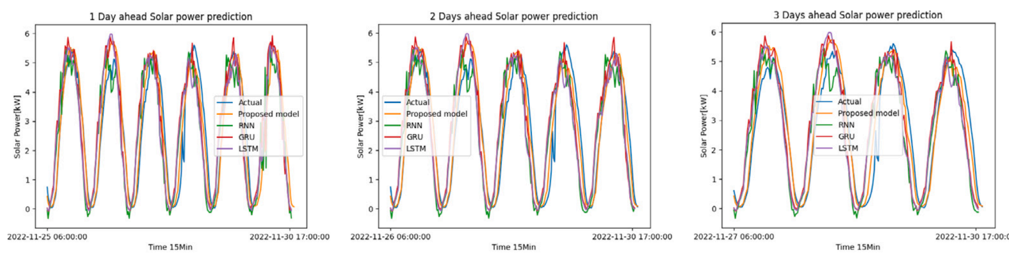
This section provides a detailed performance evaluation of the proposed model. The proposed model is eligible to predict PV power generation in both the short-term and long-term periods. For short-term prediction, we can predict the generation of solar power for fifteen minutes to one hour ahead, and for long-term prediction, we are able to predict PV power generation for one day to three days ahead with noticeable accuracy.

Table 4 demonstrates the performance metrics of the proposed model for 15 min ahead prediction and compares the result with three other standard standalone deep learning models, namely recurrent neural network (RNN), gated recurrent unit (GRU) [48] and long short-term memory (LSTM), extreme learning machine (ELM) [49], and support vector regression (SVR)[50] and we found that our model outperformed all the model results in all aspects.



(a) 30 Minute ahead prediction (b) 45 Minute ahead prediction (c) 1Hour ahead prediction

Fig. 9. Short term forecasting.



(a) One day ahead prediction (b) Two days ahead prediction (c) Three days ahead prediction

Fig. 10. Model comparison for long term prediction.

Table 5
Comparison of RMSE value with base model.

Time horizon (h)	Base model [30] [$RMSE$]	Proposed model [$RMSE$]
24	8.06	1.26
48	8.69	1.36
72	8.96	1.36

Table 6
Comparison of dataset performances.

Name of dataset	Time horizon	MSE	RMSE	MAE	R2_Score
IIEST	15 Min	0.41	0.64.	0.36	0.90
Pombo [51]	15 Min	0.58	0.76	0.51	0.87

Fig. 8 shows the prediction result for a 15 min time horizon for seven consecutive days for the month of November of the year 2022. In Fig. 9, we have presented the prediction result for short-term periods. Fig. 9(a), (b) and (c) present the results for 30 min, 45 min and one hour ahead, respectively.

In Table 5, we have shown a comparison of the RMSE values for different time horizons of our model with the previously published work of Pombo et al. [30] and found that our model noticeably beats the base model results.

Fig. 10 presents the comparisons of the proposed model for long-term prediction with three other standalone models, RNN, GRU, and LSTM, and it shows that the proposed model can predict more accurately than other mentioned models.

Since we train our forecasting model with our own curated dataset, we tried to validate our model with other similar datasets used in [30], and the result is demonstrated in Table 6.

At last, we have also tried to validate our prediction model on some solar plants situated more than 150 km from our base plant and got results as per Table 7. These plants' solar power generation capacity is multiple times more than our base solar power plant. We have demonstrated the result on normalized data of respective solar plants. The functionality of the proposed prediction model in diverse situations is quite impressive.

Table 7
Performance analyst to different solar plant in Durgapur India.

Plant name	Time horizon	RMSE	MAE	R_2Score	Capacity [kW]
IIEST	15 Min	0.07	0.05	0.95	10.0
DGP8	15 Min	0.1	0.06	0.87	80.0
DGP9	15 Min	0.1	0.06	0.87	100.0
DGP10	15 Min	0.1	0.06	0.86	75.0
DGP11	15 Min	0.1	0.06	0.87	100.0
DGP12	15 Min	0.1	0.06	0.87	100.0
DGP14	15 Min	0.09	0.06	0.87	50.0
DGP15	15 Min	0.09	0.06	0.88	25.0

5. Conclusion

Most of the previous work used only meteorological data and the generated solar power to train their corresponding model. As we augmented our dataset with the physical characteristics of Photovoltaic panels, it enriched our model to predict more accurately.

In this work, we propose an approach to prepare the dataset for a deep learning model that is able to produce improved prediction and requires minimum information, which can be found in the PV panel manufacturer datasheet. On the other hand, we develop a BI-LSTM ensemble model which is able to predict solar power more accurately than various previous published works. We also validate our work with a published dataset, which originated in Denmark, and our proposed model is trained with the data of IIEST, Shibpur, West Bengal, India. We also validated the proposed model with the data from multiple solar power plants situated in Durgapur, India, and the generation capacity of those plants is multiple times higher than our IIEST solar plant. Though there is diversity in capacity and geographical position, the prediction of our model is quite good. So, we can say that our model can produce good predictions regardless of the geographical position of the solar plant.

CRediT authorship contribution statement

Rakesh Mondal: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Surajit Kr Roy:**

Writing – review & editing, Supervision. **Chandan Giri:** Validation, Supervision, Investigation, Writing – review & editing.

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.

Data availability

Data will be made available on request.

References

[1] International energy egency. 2023, https://www.iea.org/energy-system/renewables/solar-pv, [Accessed 26 October 2023].
[2] Blaga R, Sabadus A, Stefu N, Dughir C, Paulescu M, Badescu V. A current perspective on the accuracy of incoming solar energy forecasting. Prog Energy Combust Sci 2019;70:119–44.
[3] Lopes FM, Silva HG, Salgado R, Cavaco A, Canhoto P, Collares-Pereira M. Short-term forecasts of GHI and DNI for solar energy systems operation: Assessment of the ECMWF integrated forecasting system in southern Portugal. Sol Energy 2018;170:14–30.
[4] Alonso-Montesinos J, Batlles F, Portillo C. Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images. Energy Convers Manage 2015;105:1166–77.
[5] Miller SD, Rogers MA, Haynes JM, Sengupta M, Heidinger AK. Short-term solar irradiance forecasting via satellite/model coupling. Sol Energy 2018;168:102–17.
[6] Sobri S, Koohi-Kamali S, Rahim NA. Solar photovoltaic generation forecasting methods: A review. Energy Convers Manage 2018;156:459–97.
[7] Hou Z-X, Huang G-H, Gong C, Zheng Y. The relation between tension, flange thickness and prying force for T-type bolt joints. In: Resources, environment and engineering II: proceedings of the 2nd technical congress on resources, environment and engineering (CREE 2015, hong kong, 25-26 September 2015). CRC Press; 2015, p. 53.
[8] Agoua XG, Girard R, Kariniotakis G. Short-term spatio-temporal forecasting of photovoltaic power production. IEEE Trans Sustain Energy 2018;9(2):538–46. http://dx.doi.org/10.1109/TSTE.2017.2747765.
[9] Miao S, Ning G, Gu Y, Yan J, Ma B. Markov chain model for solar farm generation and its application to generation performance evaluation. J Clean Prod 2018;186:905–17.
[10] Massidda L, Marrocu M. Use of multilinear adaptive regression splines and numerical weather prediction to forecast the power output of a PV plant in Borkum, Germany. Sol Energy 2017;146:141–9.
[11] Hossain M, Mekhilef S, Danesh M, Olatomiwa L, Shamshirband S. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. J Clean Prod 2017;167:395–405.
[12] Abuella M, Chowdhury B. Solar power probabilistic forecasting by using multiple linear regression analysis. In: SoutheastCon 2015. IEEE; 2015, p. 1–5.
[13] Barman M, Choudhury NBD. Season specific approach for short-term load forecasting based on hybrid FA-SVM and similarity concept. Energy 2019;174:886–96.
[14] Rodríguez F, Fleetwood A, Galarza A, Fontán L. Predicting solar energy generation through artificial neural networks using weather forecasts for microgrid control. Renew Energy 2018;126:855–64.
[15] Voyant C, Notton G, Kalogirou S, Nivet M-L, Paoli C, Motte F, Fouilloy A. Machine learning methods for solar radiation forecasting: A review. Renew Energy 2017;105:569–82.
[16] Yagli GM, Yang D, Srinivasan D. Automatic hourly solar forecasting using machine learning models. Renew Sustain Energy Rev 2019;105:487–98.
[17] Bogner K, Pappenberger F, Zappa M. Machine learning techniques for predicting the energy consumption/production and its uncertainties driven by meteorological observations and forecasts. Sustainability 2019;11(12):3328.
[18] Chu Y, Li M, Pedro HT, Coimbra CF. Real-time prediction intervals for intra-hour DNI forecasts. Renew Energy 2015;83:234–44.
[19] Khan W, Walker S, Zeiler W. Improved solar photovoltaic energy generation forecast using deep learning-based ensemble stacking approach. Energy 2022;240:122812.
[20] Lai J-P, Chang Y-M, Chen C-H, Pai P-F. A survey of machine learning models in renewable energy predictions. Appl Sci 2020;10(17):5975.
[21] Shi J, Lee W-J, Liu Y, Yang Y, Wang P. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Trans Ind Appl 2012;48(3):1064–9.
[22] Sun Y, Venugopal V, Brandt AR. Short-term solar power forecast with deep learning: Exploring optimal input and output configuration. Sol Energy 2019;188:730–41.
[23] Ahmed R, Sreeram V, Mishra Y, Arif M. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renew Sustain Energy Rev 2020;124:109792.
[24] Wang H, Liu Y, Zhou B, Li C, Cao G, Voropai N, Barakhtenko E. Taxonomy research of artificial intelligence for deterministic solar power forecasting. Energy Convers Manage 2020;214:112909.
[25] Ledmaoui Y, El Maghraoui A, El Aroussi M, Saadane R, Chebak A, Chehri A. Forecasting solar energy production: A comparative study of machine learning algorithms. Energy Rep 2023;10:1004–12.
[26] Wang J, Li P, Ran R, Che Y, Zhou Y. A short-term photovoltaic power prediction model based on the gradient boost decision tree. Appl Sci 2018;8(5):689.
[27] Leva S, Dolara A, Grimaccia F, Mussetta M, Ogliari E. Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power. Math Comput Simul 2017;131:88–100.
[28] Munawar U, Wang Z. A framework of using machine learning approaches for short-term solar power forecasting. J Electr Eng Technol 2020;15:561–9.
[29] Rana M, Koprinska I, Agelidis VG. Univariate and multivariate methods for very short-term solar photovoltaic power forecasting. Energy Convers Manage 2016;121:380–90.
[30] Pombo DV, Bindner HW, Spataru SV, Sørensen PE, Bacher P. Increasing the accuracy of hourly multi-output solar power forecast with physics-informed machine learning. Sensors 2022;22(3):749.
[31] Elsheikh AH, Sharshir SW, Abd Elaziz M, Kabeel AE, Guilan W, Haiou Z. Modeling of solar energy systems using artificial neural network: A comprehensive review. Sol Energy 2019;180:622–39.
[32] De V, Teo TT, Woo WL, Logenthiran T. Photovoltaic power forecasting using LSTM on limited dataset. In: 2018 IEEE innovative smart grid technologies-Asia. ISGT Asia, IEEE; 2018, p. 710–5.
[33] Almonacid F, Pérez-Higueras P, Fernández EF, Hontoria L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. Energy Convers Manage 2014;85:389–98.
[34] Basaran K, Özçift A, Kılınç D. A new approach for prediction of solar radiation with using ensemble learning algorithm. Arab J Sci Eng 2019;44:7159–71.
[35] Alkhayat G, Mehmood R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. Energy AI 2021;4:100060.
[36] Pierro M, Bucci F, De Felice M, Maggioni E, Moser D, Perotto A, Spada F, Cornaro C. Multi-model ensemble for day ahead prediction of photovoltaic power generation. Solar Energy 2016;134:132–46.
[37] Kumari P, Toshniwal D. Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. J Clean Prod 2021;279:123285.
[38] King DL, Kratochvil JA, Boyson WE. Photovoltaic array performance model. Vol. 8, Citeseer; 2004.
[39] Subbiah SS, Paramasivan SK, Arockiasamy K, Senthivel S, Thangavel M. Deep learning for wind speed forecasting using Bi-LSTM with selected features. Intell Autom Soft Comput 2023;35(3).
[40] Central pollution control board of India. 2023, https://airquality.cpcb.gov.in, (Accessed 26 July 2023).
[41] ABB PVI-10.0-TL-OUTD. 2023, https://new.abb.com/products/6AGC003526/3n81990s000a, (Accessed 28 July 2023).
[42] Lin Y, Koprinska I, Rana M, Troncoso A. Pattern sequence neural network for solar power forecasting. In: Neural information processing: 26th international conference, ICONIP 2019, sydney, NSW, Australia, December 12–15, 2019, proceedings, part v 26. Springer; 2019, p. 727–37.
[43] Gensler A, Henze J, Sick B, Raabe N. 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, IEEE; 2016, p. 002858–65.
[44] Wang D, Luo H, Grunder O, Lin Y, Guo H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. Appl Energy 2017;190:390–407.
[45] Bouzerdoum M, Mellit A, Pavan AM. A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Solar Energy 2013;98:226–35.
[46] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.
[47] Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci 1982;79(8):2554–8.
[48] Weerakody PB, Wong KW, Wang G, Ela W. A review of irregular time series data handling with gated recurrent neural networks. Neurocomputing 2021;441:161–78.
[49] Lin L, Wang F, Xie X, Zhong S. Random forests-based extreme learning machine ensemble for multi-regime time series prediction. Expert Syst Appl 2017;83:164–76.
[50] Lu C-J, Lee T-S, Chiu C-C. Financial time series forecasting using independent component analysis and support vector regression. Decis Support Syst 2009;47(2):115–25.
[51] Pombo DV, Gehrke O, Bindner HW. SOLETE, a 15-month long holistic dataset including: Meteorology, co-located wind and solar PV power from Denmark with various resolutions. Data Brief 2022;42:108046.