ISSN: 2584-0495 Vol. 3, Issue 1, pp. 1526-1531

International Journal of Microsystems and IoT

ISSN: (Online) Journal homepage: https://www.ijmit.org

Multivariate Deep Learning Bidirectional LSTM Model for Forecasting Solar Radiation considering Different Climate Zones over India

Lipika Datta, Shubham Anand

Cite as: Datta,, L., & Anand, S. (2025). Multivariate Deep Learning Bidirectional LSTM Model for Forecasting Solar Radiation considering Different Climate Zones over India. International Journal of Microsystems and IoT, 3(1), 1526–1531. https://doi.org/10.5281/zenodo.15493941

Published online: 20 January 2025

DOI: https://doi.org/10.5281/zenodo.15493941

International Journal of Microsystems and IoT
Vol.3, Issue 1, pp.1526-1531; DOI: https://doi.org/10.5281/zenodo.15493941

Multivariate Deep Learning Bidirectional LSTM Model for Forecasting Solar Radiation considering Different Climate Zones over India

Lipika Datta, Shubham Anand
Department of Electrical Engineering, Galgotias College of Engineering & Technology, Greater Noida, India

ABSTRACT

The growing need for sustainable energy is leading to the widespread development of solar power plants worldwide. The power generation of solar plants depends on solar radiation, which is impacted by weather conditions. Therefore, understanding and predicting weather patterns is crucial for estimating solar radiation availability and ensuring reliable energy production. Precisely predicting solar radiation is essential for the smooth and efficient integration of solar power plants into the grid. This study focuses on designing a multivariate deep learning (DL) Bidirectional Long Short-Term Memory (Bi-LSTM) model to predict solar radiation. The model is developed for various climatic conditions on different climate zone over India. To build the DL model, the wind speed and ambient temperature along with solar radiation have been taken as input. The presented DL model demonstrates superior performance metrics, for hot and dry climate zones with RMSE of 0.0051, R^2 values of 0.92, MAE of 0.0204, and MSE of 0.0018. The high-quality meteorological data for different climate zones are used to develop the model. A comparative analysis is carried out with other DL models to validate the obtained results. The proposed model is found to be accurate, and it can be useful for smart grid applications to maintain grid stability, reliability, and power quality.

KEYWORDS

Renewable Energy Resources. Solar Radiation; Deep Learning; Gated Recurrent Unit; Bi-Directional Long Short-Term Memory.

1. INTRODUCTION

The depletion of non-renewable energy reserves and their environmental impact have driven advancements in renewable energy generation systems. Among these alternatives, solar power plants are becoming increasingly important due to the abundant availability of solar energy worldwide [1]. The reduction in capital and maintenance costs of solar power plants is further driving the momentum of this growth trend. However, incorporating solar power generation systems presents a challenge, as energy output varies significantly with fluctuations in solar radiation.

Due to the dynamic movement of clouds, the solar irradiance received by PV modules varies significantly on cloudy days. High fluctuation over time creates lots of uncertainty in economic growth. Moreover, uncertainty can also affect the stability of the grid [2]. A literature review reveals that most existing regression models are designed for sunny conditions but are ineffective under foggy or cloudy skies. Various solutions have been implemented to address these challenges, but some are costly, while others may not be practical in different scenarios. One of the promising solutions is to predict solar energy for different climate zones. Accurately predicted solar energy with minimum uncertainty can help grid operators in managing demand and supply. One of the most challenging aspects of the renewable energy system is to improve the forecasting models [3]. It has been found that the primary factor influencing PV power output is the amount of solar radiation hitting the PV panel.

However, accurately predicting solar radiation is challenging due to the unpredictable nature of solar energy, which is influenced by various meteorological, spatial, and temporal factors. Recently to address solar radiation prediction many statistical and machine-learning techniques have been used. In[] developed a meteorological parameter-based regression models to predict solar energy in Turkey. The findings indicate that the Veeran and Chegaar models maintain an acceptable error margin of ±10% (Bayrakci, Demircan, and Kecebas, 2018). In India, where over 80 days per year are cloudy, achieving accurate solar energy forecasts using multiple regression analysis approaches becomes challenging.

Since intelligent modeling techniques have been introduced to address the uncertainties caused by weather variations. ML models like Artificial neural networks (ANN), Decision Trees (DT), Support Vector Machines (SVM) [5], and Linear Regression (LR) deal with nonlinear data to predict the output [6]. However, due to the unpredictable behavior of solar radiation, traditional artificial intelligence (AI) algorithms may struggle with local minima. So, DL models come into the picture which gives accurate output as compared to traditional AI models. DL is a more sophisticated variant of a neural network. Recently, various DL models have gained importance in solving nonlinear time series problems over data-driven prediction models. Recently, Convolutional neural network (CNN) models have gained more attention in short-term prediction as they effectively capture the features from the input sequence [7]. Recurrent neural networks (RNN) have the issue of gradient vanishing and exploding, despite having internal memory to store the prior value [8]. It stops updating the weight in backward propagation due to this gradient vanishing problem, and it modifies the weight value in case of gradient exploding. To solve these problems, long short-term memory (LSTM) and Gated recurrent unit (GRU)-based models have come into the picture which has some extra control gates [9]. While in most cases, LSTM is used for the univariate model, multivariate LSTM models with all input parameters such as Global Horizontal Irradiance (GHI),

2. FORECASTING TECHNIQUES AND METHODOLOGIES

In this article, the proposed Bi-directional LSTM model compared with two other deep learning models is described below. It is an improved version of traditional RNN architecture. It is specifically designed to address the vanishing gradient problem. This enhancement improves its ability to capture long-term dependencies [13]. It has a

2.1 Long Short-term Memory

It is a variant of the RNN architecture [11]. Unlike Conventional RNNs, LSTM networks are specifically designed to mitigate the vanishing gradient problem, which commonly occurs when training neural networks on long data sequences. LSTMs incorporate a memory cell that retains information over time, along with gates that regulate the flow of information into and out of the cell, as illustrated in Fig. 1. These gates enable the network to learn which information to retain or discard, helping LSTMs overcome the vanishing gradient issue [12]. The mathematical formulation of the LSTM model is provided in Eq. (1-6).

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \quad (1)$$

$$\hat{c}_t = \tanh(w_c[h_{t-1}, x_t] + b_c) \quad (2)$$

GRU cell consisting of an update gate and a reset gate. The update gate controls how much of the previous hidden state is preserved [14], while the reset gate determines the extent to which past information is forgotten and how much of the new input contributes to updating the hidden state. The fundamental structure of the GRU is illustrated in Fig. 2, and its mathematical formulation is provided below.

$$r_t = \sigma(w_{rh}h_{t-1} + w_{rx}x_t + b_r) \quad (7)$$

$$z_t = \sigma(w_{zh}h_{t-1} + w_{zx}x_t + b_z) \quad (8)$$

$$\tilde{h}_t = \varphi(\tilde{w}_{hh}h_{t-1}) + \tilde{w}_{hx}x_t + b_z) \quad (9)$$

$$h_t = (1 - z_t)h_{t-1} + z_t\tilde{h}_t \quad (10)$$

Here r_t,z_t,x_t & h_t represent the reset gate, update gate, and input and output gate respectively.

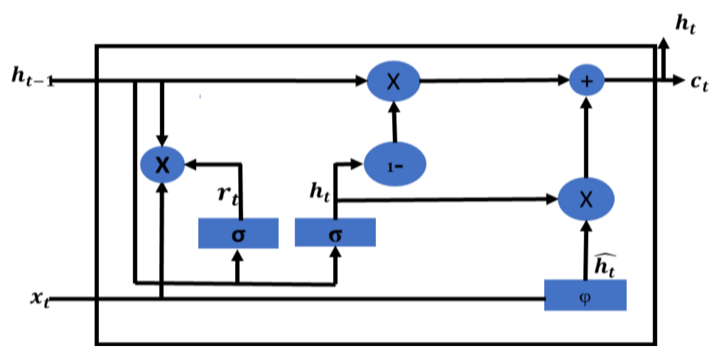

Fig. 2. GRU Architecture

1.2 Bi-Directional Long Short-Term Memory

It is a variant of the RNN architecture widely used for sequential data processing tasks, including natural language processing (NLP) and speech recognition [15]. In a Bi-LSTM, the input sequence is analyzed in both forward and backward directions, as illustrated in Fig. 3. The outputs from each direction are then concatenated to generate the final output sequence [16]. By utilizing information from both past and future contexts, Bi-LSTM enhances prediction accuracy by considering the entire sequence rather than relying solely on the current input [17].

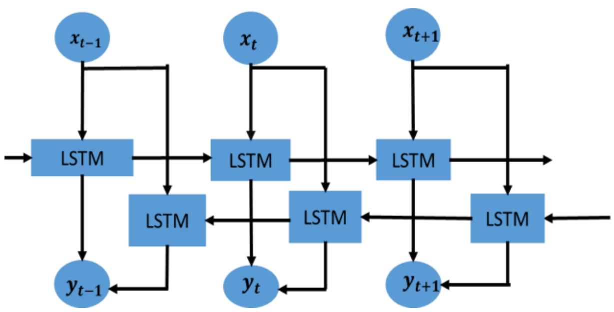

Fig. 3. Basic Architecture of Bi-LSTM

3. VARIABLE INPUT AND DATA PREPROCESSING

This section describes the insights into input parameters and data processing techniques that have been used in this research.

3.1 Climate Zone

A study by Bansal and Minke in 1988 identified five climate zones from 233 meteorological stations across India by analyzing average monthly data. Prevailing weather conditions for more than six months is the key criterion for assigning a site to these zones. These classified zones are hot & dry, cold & cloudy, moderate, composite, and warm & humid as shown in Table 1.

3.2 Collection of Meteorological Data and Model Inputs

The site selection and input variables play an important role in the performance of the solar radiation forecasting model. Solar radiation is highly correlated with different weather variables. Yet, all weather parameters do not have the same correlation and significance as the forecasted variables. Proper selection of adequate input gives a better-performing forecasting model. Improper selection of input parameters that have weak correlation factors results in a complex forecasting model. In this study, 1 year of hourly data for five climate zones has been utilized, sourced from the National Renewable Energy Laboratory (NREL) site. The primary input variables considered for forecasting solar radiation are GHI, temperature & wind speed for the multivariate prediction model. The daily average solar radiation and temperature for the specified region collected from the NREL site are depicted in Fig.4 and Fig.5. The daily average wind speed for the specified region is shown in Fig.6.

Table 1. Topographical features of meteorological stations with different climatic conditions.

Climate Zone	Meteorological Stations	$T_{amb}(^0C)$	Relative humidity	No of Sunny Days
Hot & Dry	IIT Jodhpur, Rajasthan	>30	<55	>20
Cold & Cloud	Meghalaya	<25	>55	<20
Moderate	Karad, Pune	25-30	<75	<20
Composite	Gurgaon, Haryana	This condition occurs when six months or more does not fit into any of the classifications mentioned below.		
Warm & Humid	CWET Chennai, Tamil Nadu	>30	>55	<20

3.3 Data Preprocessing

A normalization strategy has been implemented in this model to ensure the prediction model remains well-conditioned. Normalization enhances the stability and linearity of the output while handling non-linear input datasets, as described by:

$$x_{norm} = \frac{x_i - x_{mn}}{x_{mx} - x_{mn}} \quad (11)$$

Here, x_i denotes the ith component of the original data input, while x_{mx} and x_{mn} represent the maximum and minimum values of the x dataset, respectively. After data processing, the dataset, which includes observations of global solar irradiance, ambient temperature, and wind speed at each time step, is transformed into a supervised learning format.

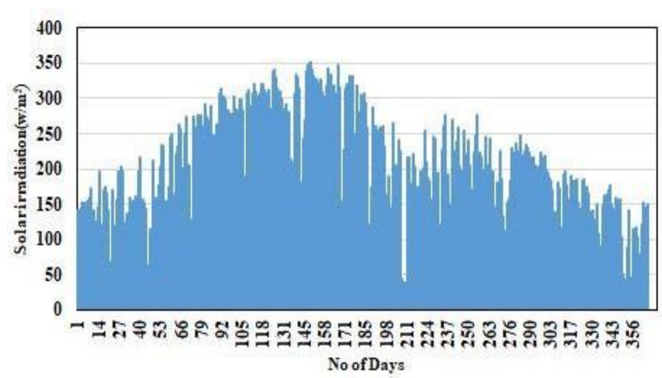

Fig. 4. Daily Average Solar Irradiance

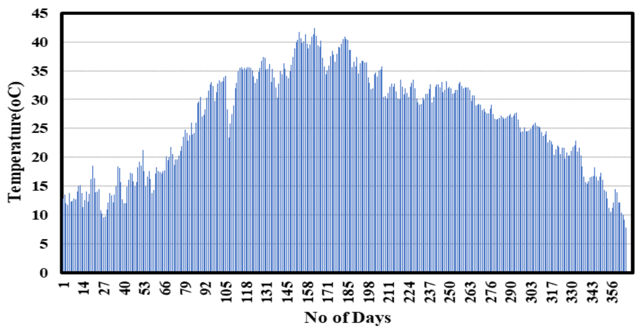

Fig. 5. Daily Average Temperature

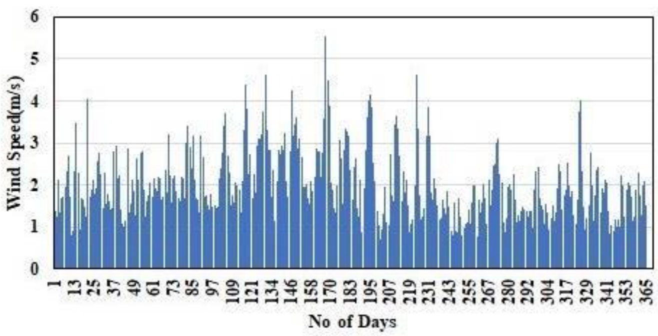

Fig. 6. Daily Average Wind Speed

3.4 Design and Development of Deep Learning Model

To construct the DL model, the data must be formatted as a three-dimensional (3D) array, consisting of batch size, step size, and the number of input variables. For instance, a 3D data array with a shape of (72,2,25) indicates a batch size of 72, two input variables, and 25-time steps. Certain hyperparameters were determined through exploratory studies, including batch size, the number of epochs, window size, and the number of hidden layers. The hyperparameters used for these models are presented in Table 2.

Table 2. Hyperparameter used for Bi- LSTM Model

Hyper Parameter	Values
Activation Function	tanh
No of hidden layers	50
Batch size	32
Epoch	50
Optimizer	Adam
Learning Rate	0.01

4. RESULT ANALYSIS AND DISCUSSION

Each mentioned forecasting model is designed and simulated in Python software using different optimization parameters. In this section of the work LSTM, GRU, and Bi-LSTM-based deep learning models are tested for the multivariate forecasting models for five different climate zones. The performance evaluation and comparison of the proposed deep learning model for multivariate forecasting models for every climate zone as shown in Table 3 - Table 7. The Bi- LSTM-based DL model has the RMSE and R^2 of 50.246 and 0.97 for hot and dry weather climate zones. It has RMSE and R^2 of 77.528 and 0.91 for cold and cloud weather climate zones. The Bi- LSTM-based DL model has the RMSE and R^2 for composite weather is 78.89 and 0.91. The Bi- LSTM-based DL model has the RMSE and R^2 of 48.08 and 0.97 for warm and humid weather climate zones. The RMSE and R^2 of 60.867 and 0.96 for moderate weather climate zones. The MAE values for hot & dry, cold & cloud, composite, warm & humid, and moderate weather zones are 0.0204,0.0406,0.0356,0.0225 and 0.0291.

The MSE values for hot & dry, cold & cloud, composite, warm & humid, and moderate weather climate zones are 0.0018,0.0060,0.0056,0.0020 and 0.0038. The Bi-LSTM-The DL model has the MAE and MSE of .0204 and 0.0018 for hot and dry weather climate zones is the lowest error value as compared to the other climate zones. In multivariate analysis, Bi-LSTM has the lowest error and convergence for hot and dry weather climate zones. The performance evaluation and comparison of different DL models with the proposed deep learning model for hot and dry weather climate zones as shown in Fig.7 & Fig.8. After thoroughly analyzing all the results, it is evident that the proposed Bi-LSTM model outperforms the other deep learning models in both univariate and multivariate forecasting predictions. The predicted response plot of the proposed model is shown in Fig.9. From Fig.9 it can be observed that the difference between the actual output and predicted output is very small with minimum errors.:

Table 3. Performance Errors for Multivariate Solar Radiation Prediction Hot and Dry Climate Zone

TECHNIQUE	R^2	MAE	MSE	RMSE
GRU	0.976	0.0253	0.0029	50.246
LSTM	0.977	0.0224	0.0027	50.63
Bi-LSTM	0.98	0.0204	0.0018	50.60

Table 4. Performance Errors for Multivariate Solar Radiation Prediction Cold and Cloud Climate Zone

TECHNIQUE	R^2	MAE	MSE	RMSE
GRU	0.887	0.0426	0.0063	82.55
LSTM	0.912	0.0399	0.0061	77.21
Bi-LSTM	0.914	0.0406	0.0060	77.18

Table 5. Performance Errors for Multivariate Solar Radiation Prediction Composite Climate Zone

TECHNIQUE	R^2	MAE	MSE	RMSE
GRU	0.912	0.0375	0.0059	78.89
LSTM	0.932	0.0359	0.0057	76.14
Bi-LSTM	0.922	0.0356	0.0056	76.76

Table 6. Performance Errors for Multivariate Solar Radiation Prediction Warm and Humid Climate Zone

TECHNIQUE	R^2	MAE	MSE	RMSE
GRU	0.94	0.0261	0.0024	69.46
LSTM	0.94	0.0217	0.0021	70.01
Bi-LSTM	0.95	0.0225	0.0020	60.27

Table 7. Performance Errors for Multivariate Solar Radiation Prediction Moderate Climate Zone

TECHNIQUE	R^2	MAE	MSE	RMSE
GRU	0.96	0.0324	0.0041	61.59
LSTM	0.95	0.0288	0.0039	66.13
Bi-LSTM	0.96	0.0291	0.0038	60.87

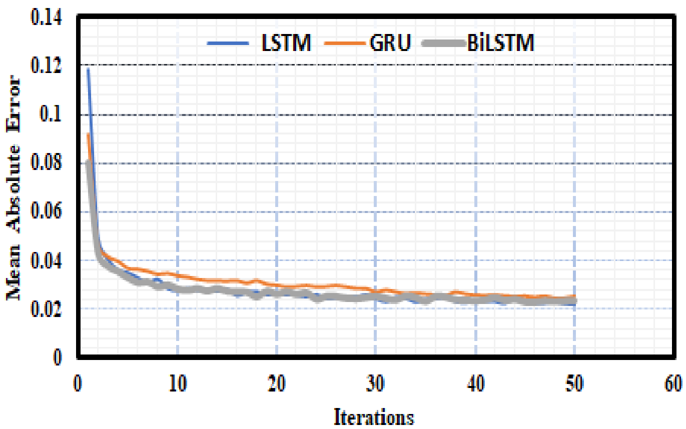

Fig. 7. Mean Absolute error Minimization for Multivariate Solar Radiation Prediction.

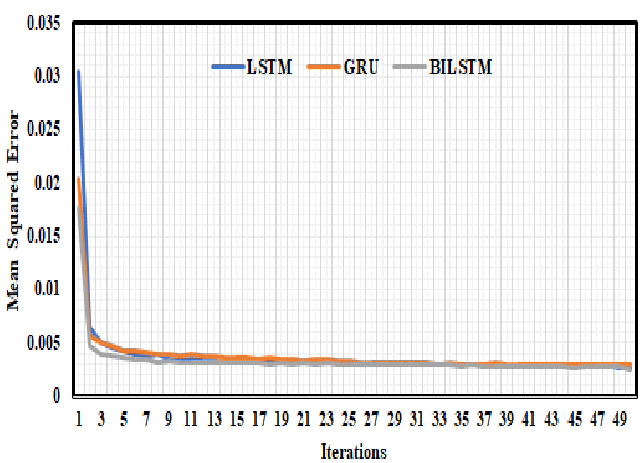

Fig. 8. Mean Square Error Minimization for Univariate Solar Radiation Prediction.

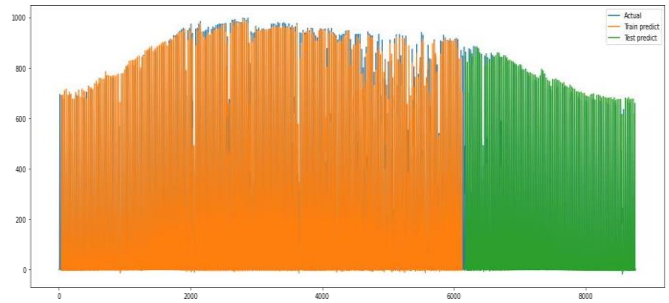

Fig. 9. Prediction Response & comparison of the proposed model with the actual response

5. CONCLUSION AND FUTURE SCOPE

An accurate forecasting model is beneficial for the reliable operation of the grid and for increasing energy sustainability. This paper gives insight into a novel multi-variate and univariate solar radiation forecasting model based on deep learning. The model was developed using topographical and meteorological information for the Delhi, NCR region. The forecasting model was developed using irradiance as the input for the univariate system, while the multivariate system utilized irradiance, wind speed, and temperature as inputs. The proposed model was assessed alongside two other DL-based models during the training and testing phases. The statistical indicators considered for the effective performance of the model are RMSE, R^2, MAE, and MSE. The proposed Bi-LSTM method gives a better result as compared to another DL model. During the testing phase RMSE, R^2, MSE, and MAE are 50.60,0.98,0.0018 and 0.0.0204 for multivariate forecasting models for hot and dry weather climate zones. Respectively among all the applied models the proposed Bi-LSTM-based DL model outperforms another DL model for multivariate analysis for hot and dry weather climate zones in terms of the loss function. The findings can be utilized for the deployment of grid-integrated SPV systems in different regions for varying climate zones. The proposed model will contribute to energy management applications in smart grids. For future research, this work can be expanded in several directions: (i) integrating the BiLSTM model with other DL models for solar radiation prediction in different locations, and (ii) enhancing performance through advanced hyperparameter tuning.

REFERENCES

1. X. Huang, S. Han, W. Huang, and X. Liu, "Enhancing solar cell efficiency: The search for luminescent materials as spectral converters," *Chem Soc Rev*, vol. 42, no. 1, pp. 173–201, Dec. 2013,doi: 10.1039/c2cs35288e
2. F. Almonacid, P. J. Pérez-Higueras, E. F. Fernández, and L. Hontoria, "A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator," *Energy Convers Manag*, vol. 85, pp. 389–398, 2014, doi: 10.1016/j.enconman.2014.05.090
3. R. L. Cavalcante, T. O. Costa, M. P. Almeida, S. Williamson, M. A. B. Galhardo, and W. N. Macêdo, "Photovoltaic penetration in isolated thermoelectric power plants in Brazil: Power regulation and specific consumption aspects," *International Journal of Electrical Power and Energy Systems*, vol. 129, Jul. 2021, doi: 10.1016/j.ijepes.2020.106648

4. Colak I., Yesilbudak M., Genc N., and Bayindir R.,. Multi-period prediction of solar radiation using ARMA and ARIMA models. In: Proc. *IEEE 14th Int. Conf. Mach. Learn. Appl.* (ICMLA); December 2015: 1045–1049.
5. Liu G, Zhou J, Jia B, He F, Yang Y, Sun N. Advance short-term wind energy quality assessment based on instantaneous standard deviation and variogram of wind speed by a hybrid method. *Appl Energy*. 2019;238:643-667.
6. Kashyap Y, Bansal A, Sao AK. Solar radiation forecasting with multiple parameters neural networks. *Renew Sustain Energy* Rev. 2015; 49:825-835.
7. V. Suresh, P. Janik, J. Rezmer, and Z. Leonowicz, "Forecasting solar PV output using convolutional neural networks with a sliding window algorithm," *Energies (Basel)*, vol. 13, no. 3, 2020, doi: 10.3390/en13030723.
8. Cho K, Merriënboer B.V. Gulcehre C, Bahdanau D., Bougares F., Schwenk H., Bengio Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation; 2014. arXiv preprintarXiv:1406.1078.
9. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-1780.
10. Siami-Namini S., Tavakoli N., and Namin A.S., The Performance of LSTM and BiLSTM in Forecasting Time Series. 2019 *IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA; 20.
11. T. Sagheer, M. Kotb, Unsupervised Pre-training of a Deep LSTM-based Stacked autoencoder for multivariate time series forecasting problem, *Sci. Rep*. 9 (2019) 19038.
12. M. Husein, I.Y. Chung, Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: a deep learning approach, *Energies* 12 (2019) 1–21.
13. H. Eivazi, L. Guastoni, P. Schlatter, H. Azizpour, R. Vinuesa, Recurrent neural networks and Koopman-based frameworks for temporal predictions in a low-order model of turbulence, Int. J. Heat Fluid Flow 90 (2021).
14. S. Wang, H. Chen, A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network, *Appl. Energy* 235 (2019) 1126–1140.
15. Huang, Z., Xu, W., & Yu, K. (2015). Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:1508.01991.
16. Yao, Y., & Huang, Z. (2016, October). Bi-directional LSTM recurrent neural network for Chinese word segmentation. In *International Conference on Neural Information Processing* (pp. 345-353). Springer, Cham.
17. Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2673-2681.

AUTHORS

Lipika Datta received her BTech degree from NIT, Agartala, India in 2007 and her MTech degree in Power Electronics and Drives from NIT Agartala, India in 2010. She is currently pursuing PhD at the Department of Electrical Engineering, at Delhi Technological University. Currently, she is working at Galgotias College of Engineering and Technology, Greater Noida. Her areas of interest are Renewable Energy, Forecasting, and Machine learning.

E-mail: lipika.datta@galgotiacollege.edu

Shubham Anand received his diploma in electrical engineering from Amity University Greater Noida Uttar Pradesh in 2021. He is currently pursuing a B.Tech in electrical engineering from Galgotias College of Engineering and Technology Greater Noida.

E-mail: shubhdeepanand007@gmail.com