LSTM BASED DEEP LEARNING MODEL FOR ACCURATE WIND SPEED PREDICTION

V. Prema[1], Sushmita Sarkar[2], K. Uma Rao[3] and Amrutha Umesh[4]
[1,2,3]*Department of Electrical and Engineering, RV College of Engineering, India*
[4]*Department of Computer Science and Engineering, NMAM Institute of Technology, India*

Abstract

Wind power is credited with large generation capacity among the different types of renewable energy. Increased growth in wind power generation calls for accuracy in wind speed forecasting as it intermittent on various time scales. Due to its stochastic nature, many models have been developed for accurate prediction of wind speed. This paper proposes an accurate prediction model with deep learning using LSTM (Long Short-Term Memory) technique. The model is trained with real time data collected from a wind farm and it leads to a day ahead prediction of wind speed. Another model is developed with recurrent Neural Network and trained with the same data. The Mean Average Percentage Error (MAPE) for both the models are compared. By analyzing the results, the error in prediction of the wind speed is very minimal in Deep Learning using LSTM model. This is due to their nature of selectively remembering patterns for a long duration of time over the traditional neural network.

Keywords:

Wind Speed, Forecast, Deep Learning, Neural Network

1. INTRODUCTION

The demand of energy for the world is humongous and is increasing tremendously, the current annual power consumption being approximately 15 terawatts (TW). The current trends indicate a doubling of this by 2050 and by the end of the century it is expected to triple. Two scenarios may be considered to alleviate this threat – carefully utilize the available energy or find alternative energy sources. Even with focused energy conservation and enhancement of efficiency of the current systems, the fossil fuels which are primary sources of energy are fast depleting and may not be able to cater to this mammoth demand for long, owing to energy hungry technology development and increasing human population on earth. One of the prime objectives and daunting challenges of mankind is to find and employ a clean energy source that could suffice and secure the demand for the future.

Among the available renewable energy sources, wind energy has great potential for electric power conversion and will be able to ensure a sizeable contribution to the electrical energy demand of the planet. The advantage of wind energy is that it is easily scalable. Wind energy has an enormous offering to the scope of renewable power generation. This is evident from the studies that operating 20% rated capacity of 2.5 MW turbines restricted to warm, rural and non-forested areas is capable of exceeding the cumulative global consumption of energy by 5 fold and that of electricity alone by 40 times. Off-shore wind capacity is also colossal. Studies indicate that it could serve the energy demand of Europe seven times over and energy demand of Unites States four times over.

Power generation from wind is highly susceptible to climatic variables viz. geographic allocation, wind speed and its direction, seasonal changes, time of the day, etc. Thus efficient power dispatch from a wind power plant necessitates predictive analysis. A uniform efficiency cannot be warranted for a particular forecasting method across different geographies. Hence, it is critically essential to examine the seasonality and other influencing parameters to determine the best fit model for the given location. Literature provides a good number of options for forecast such as statistical models, intelligent models, etc.

The pattern of wind speed and consequently the pattern of wind power is highly erratic in nature. Due to sudden wind gusts, outliers are more. Thus direct statistical models cannot give accurate predictions. Most of the works in the literature employ hybrid models which combine statistical, intelligent and physical models.

In [1] a new multivariate Least Square - Support Vector Machine (LS-SVM) model is proposed. The results are compared with the standard models like ARIMA, neural network models, etc. SVM is a regression model used to find out the mapping between the predicted values and actual values of wind speed. The data is collected from various wind farms for a duration of 1 year. MAPE is calculated and compared. The lowest MAPE is found to be 10.06%.

In [2], wavelet, Particle Swam Optimization (PSO) and adaptive Network based Fuzzy Inference System (ANFIS) are combined to get a hybrid approach. Wavelet Transform is used to decompose the measured values of wind power. The prediction of this data is performed using ANFIS. PSO technique is used to improve the performance. The accuracy of the model is increased by training the parameters of membership function of ANFIS using PSO. The proposed model is compared with many existing models. The least MAPE obtained is 4.98%.

In [3], Extreme Learning Machine and Seasonal ARIMA methods are used for wind speed prediction. The lowest MAPE obtained is 6%. In [4], Wavelet Neural Network (WNN) is used to predict wind power. WNN is a ANN where the activation function is a wavelet function. A new training strategy is proposed based on Clonal Search Algorithm (CSA). The proposed model is compared with existing models such as Simulated Annealing (SA), Particle Swam Optimization (PSO), CSA and Differential Evolution (DE). The proposed method has an error of 9.7%, which is the lowest. In [5], hybrid model is developed to predict wind speed where, wind speed is decomposed into several sub layers using empirical mode decomposition. Each decomposed series is predicted with neural network optimized by genetic algorithm and mind evolutionary algorithm. The lowest MAPE obtained was 2.5%. In [6], lifting wavelet transform and Support Vector Machine (SVM) are used for model building. Wavelet transform characterizes the original wind speed and SVM improves the prediction accuracy. Each decomposed data is predicted separately and superimposed to obtain the final prediction. The prediction error is found to be 14.9%.

In [7], SVM enhanced Markov Chain model is chosen for wind power forecast. Markov chain is used to capture the normal variations in wind speed, whereas SVM identifies the wind ramp dynamics. [8] Proposes a mixed ARMA model with wind direction as one of the inputs. The relationship between wind direction and wind speed is obtained by K-Mean clustering. It can be observed from the above studies that the methods involving Neural Network, Fuzzy or optimization algorithms give MAPE ranging from 2.5% to 18%. Hence, more accurate Forecasting models for solar and wind power prediction are required with better MAPE.

In [9], two-layer ensemble machine learning technique is used to develop data driven wind forecasting multi-model where first layer generates individual forecasts by using multi machine learning technique and the second layer creates an ensemble of the forecast by suitable algorithm and then generate both probabilistic and deterministic forecasts. Forecasting accuracy improved up to 30% with the implementation of multi model framework with deep feature selection procedure. [10], proposes various combinations of deep learning techniques and ANN algorithms such as auto encoder and LSTM for solar forecasting to show forecast strength compared to physical forecasting model. The experimented result shows that Auto-LSTM is the best performing model with root mean squared error (RMSE) of 0.0713 and also Deep learning algorithms gives superior forecasting results than ANN and any other physical model.

In [11], Long short-time memory recurrent neural network (LSTM-RNN) is used for accurate forecasting the output power of PV system. Evaluation of the proposed method was done using hourly data sets of different sites for a period of one year and the results were compared using three other forecasting models. Five LSTM models were tested for forecasting PV power with various architectures where model 3 based on multiple linear regression (MLR), bagged regression trees (BRT) and neural network (NN) methods proved to be the best with the least error. The author of [12], recommended a shallow and DNN linked with input of different selection algorithm are compared on ultra-short-term wind prediction task for various locations. Proper selection of input variable reduces the NN complexity and ease the DNN training. Although ultra-short-term wind improves the prediction performance and increased efficiency of deep architectures but increases the complexity of prediction model.

This paper proposes a new forecast model for wind speed prediction with deep learning algorithm using LSTM technique.

2. ARTIFICIAL NEURAL NETWORK

Neural Networking is a computational model that is based on the structure and biological functions of neural networks. It is typically best suited when the data used is non-linear. One of the most recognized advantage of ANN is its ability to learn by training data sets. They are interconnected by three layers as shown in Fig.1 with the first layer consisting of input neurons which send data to the second layer which in turn sends data to the output neurons of the third layer. Here the information flow is in one direction. All these connections have a weight associated to it which a Neural Network learns during training and this set of weights controls the cost of the overall model. If the prediction has a minimal error, then the cost will be low. In this traditional feed-forward neural networks (FFNNs), the test cases are independent and are widely used in various domain such as image processing, object recognition and data classification. Another type of ANNs, is Recurrent Neural Networks (RNNs) that are structurally same as FFNNs but allow the connections between the neurons within the same hidden layer. It is capable of mapping all historical input data to the final output by permitting historical input data to be stored in the network internal state.

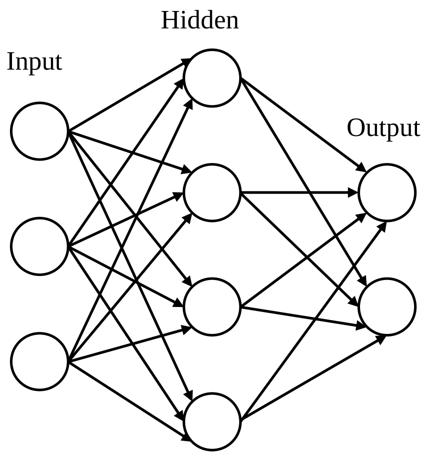

Fig.1. Architecture of ANN

ANN work even if one or a few units fail to react to the network. Many processing and storage resources have to be committed to implement large and effective neural networks. Whilst the brain has hardware that is designed to process signals through a neuron graph, a neural network designer may be forced to fill in millions of database rows with connections by simulating even a most simplified form using Von Neumann Technologies, which can use a huge amount of computer memory and hard disk space. The ANN learns from analyzed data and needs no re-programming. It is also known as blackbox model and gives little insight into the actual effects of the models. The user needs only to feed and monitor the input and wait for the output. ANN is considered to be simple math models for improving existing technology for data analysis.

3. DEEP LEARNING WITH LONG SHORT-TERM MEMORY

A subdivision of Machine learning is deep learning which an end to end process is. Machine learning algorithms need structured data and by understanding these structured data further outputs with more sets of data are produced and if the desired output is not obtained then it needs human intervention. Whereas, deep learning network works on layers of ANN as represented in Fig.2 and learn through their own errors. Hence, they do not require any human intervention. One of the advantages of Deep Learning is its ability to learn and capture high level features in an incremental order from the provided data which eliminates the need of having the expertise in the domain.

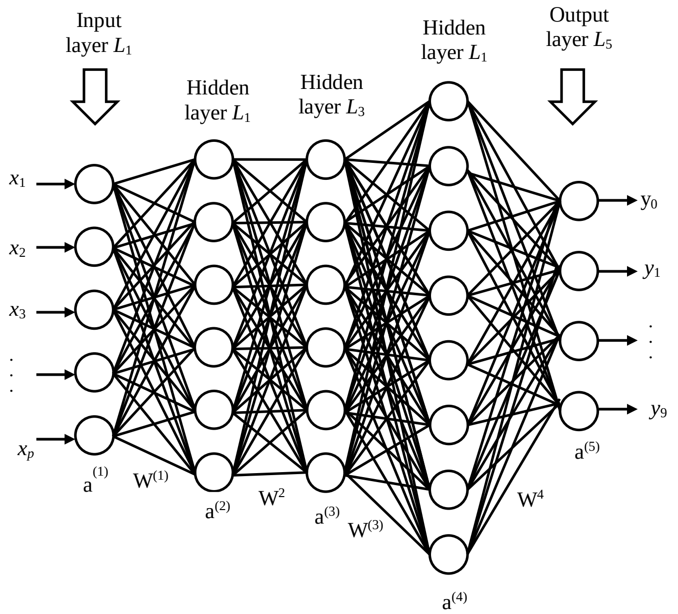

Fig.2. Architecture of deep learning (Source: University of Cincinnati)

Support Vector Regression (SVR) has been used for prediction using times series data. However, this method lacks a structured procedure to determine the parameters crucial to the efficacy of the model. Deep learning techniques are gaining popularity because of their flexible structure. Recurrent Neural Networks (RNN) are a modification of conventional neural networks, wherein the impact of the input on the network hidden layers and output decays exponentially, while cycling around the recurrent networks. This is again modified in Long Short Term Memory (LSTM) by changing the structure of the hidden neurons in RNN.

3.1 LONG SHORT TERM MEMORY MODELS (LSTM)

Long short-term memory is an evolution on a recurrent neural network (RNN). Normal RNN modules take output of last layer over a single Tanh function whereas, LSTMs use feedback loop and gates to remember. LSTMs have four interactive NNs layer in each module and comprises of a cell, an input gate, an output gate and a forget gate where the cell recalls values over self-assertive time intervals and the three gates direct the progression of data into and out of the cell. LSTM can add or remove the data from module state (which is the main chain of data flow) via sigmoid gates. The Fig.3 represents a simple architecture of LSTM cell.

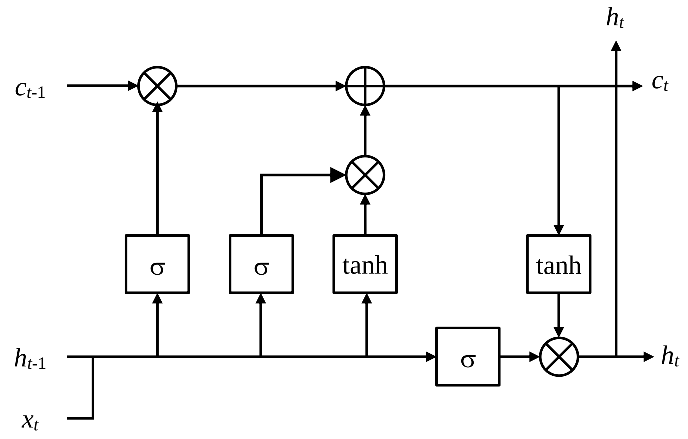

Fig.3. Long Short- Term Memory cell

Long Short Term Memory Model is a type of supervised deep learning that is highly efficient for processing the time series prediction. Here, information is passed into these cell states by a mechanism. In this manner LSTM selectively decides to which data to retain. The data present at a particular cell state has three dependencies:

- Previous cell state.
- Previous hidden state.
- Input at the current time step

4. DATA USED

Two models are developed based on different time series prediction concepts and the results are compared. The data used for the proposed work were collected from a wind farm in the state of Karnataka, India. The recorded data consists of values of air temperature, relative humidity, pressure, wind speed and the wind direction measured for different altitudes for every minute over a period of five months. For analysis, the values at one particular altitude were considered. The choice is made based on the fact that wind velocity is higher in upper strata of atmosphere, thereby having more potential to generate wind energy. The data as such consisted of more than 150 thousand points. Each of them had been measured with an interval of approximately 1 min. It was observed that some points were not exactly spaced at that interval. Thus, the data was averaged over an interval of 10 min to reduce the number of data points as well as to have a uniform time interval. A sample data is show in Fig.4.

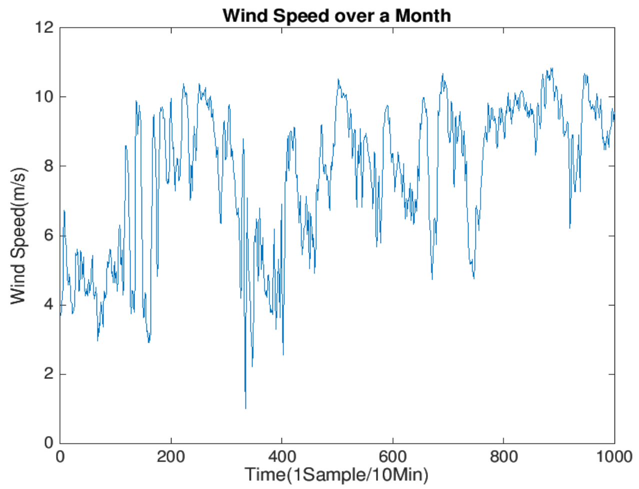

Fig.4.Sample plot of wind speed

5. ERROR MEASURES

It is essential to perform a comparative analysis to determine the accuracy of individual model when different models are employed for the test data. Right choice of the error measures is pivotal in determining the accuracy of the predictive model. Mean Absolute Percentage Error (MAPE) is used to compare performances of different models in this work. MAPE can be defined as given in Eq.(1).

$$MAPE = \frac{1}{N}\sum\frac{|X_t - F_t|}{X_t}\times 100 \quad (1)$$

where

X_t - Input data at time t

F_t - Predicted value at time t

N - Number of samples used in computing the error

MAPE represents the quantum of error relative to the actual data expressed as percentage, thereby directly indicating the accuracy of the model. Considering an error value of 10, MAPE is 1% for an actual value of 10 and 20% for an actual value of 50. This enhances the choice of MAPE as an effective pointer to the accuracy of forecasting techniques.

6. MODEL DEVELOPMENT

Two models are developed in this paper. The first model is using ANN. Neural network models are best choice for predictive models where the data is non-linear. It is evident from Fig.1 that the nature of wind speed is non-linear and erratic. Thus neural network is chosen for the first model. The second model is developed using deep learning LSTM technique.

6.1 DATA PRE-PROCESSING

It was observed for the pattern of wind speed data that there are many outliers due to sudden wind gusts. This can lead to high amount of errors in prediction. Thus a data pre-processing is necessary before the model development. There are many methods mentioned in literature for data smoothening and pre-processing. In case of wind speed, sudden peaks due to wind gusts can be considered as high frequency components. This can be best eliminated by applying wavelet transformation to the original data [13]. Discrete wavelet transform with Daubechies wavelet family is used in this paper for wavelet decomposition. Daubechies wavelets can be decomposed into many levels. However, as the number of levels is increased the computation increases. It was found that a three level decomposition was sufficient and this has been used in the paper.

6.2 ANN MODEL

The wavelet decomposed data is fed to the neural network model. A data duration of 1 year is taken for model development. 80% of data is used as training set and 20% is taken for validation and testing. Non-Linear Autoregressive network with Exogenous input (NARX) network is chosen. It was found that relative humidity and air temperature have a close relationship with wind speed. Thus in this network wind speed is used in the neural network along with humidity and temperature values as external or exogenous inputs.

The network specifications are as follows:

- Delay: 20
- Training algorithm: Scaled Conjugate Gradient
- Inputs that are exogenous: Humidity and Temperature
- Hidden neurons: [5 5] (5 in the first layer and 5 in the second layer)

A training strategy named recursive training is used for training the neural network [15]. With conventional neural network it was observed that as the forecast horizon is increased, the error also increases. Therefore, a new strategy was tried out, wherein a few samples are predicted ahead, and using these the next block of samples are predicted and so on. In this paper, for training, 800 samples are used to predict the next 5 samples. Further, the earliest 5 samples are dropped, the newly predicted 5 samples are added to the time series, to create a new set of 800 samples. These are used to predict the next 5 samples and so on. This is called re-train. This process is continued, till 100 samples are predicted ahead. The Fig.5 shows the block diagram of the strategy used. Humidity and temperature are exogenous parameters, while the wind speed is the parameter to be predicted. 800 samples of each of the three parameters, are fed as inputs to the NARX network to predict the next 5 samples. The process is continued as explained until 100 samples are predicted.

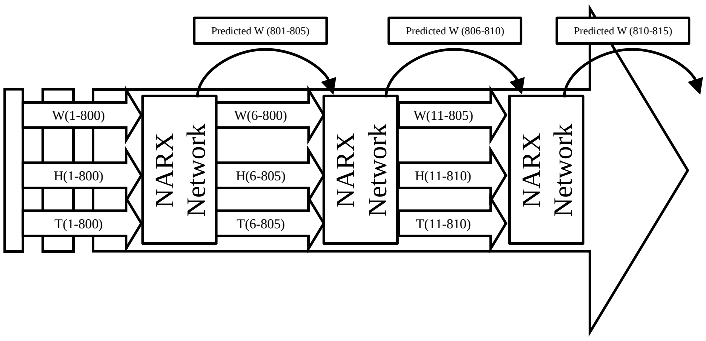

Fig.5. Training by Recursive model

6.3 DEEP LEANING LSTM MODEL

For the deep learning, the data set is partitioned into 90% for training and 10% for testing. The data is standardized to have a zero mean value and unit variance. This helps to prevent the training from diverging. At each time step, the LSTM network is trained to predict the wind speed for the next time step. The number of hidden layers in the LSTM regression network is chosen as 200. The network was trained for 250 epochs. The gradient threshold was set to 1 to restrain the gradients from exploding.

7. RESULTS AND DISCUSSION

Data for a duration of one year has been trained with both ANN and deep learning models. Results are generated for a day ahead prediction of wind speed. MAPE was calculated for each model. The prediction plots for ANN and Deep learning models are shown in Fig.6 and Fig.7 respectively. The comparison of MAPE is shown in Table.1. It can be observed from the table that the MAPE value with Deep learning model is 1.7% which is very less compared the MAPE of ANN model or any model in the literature.

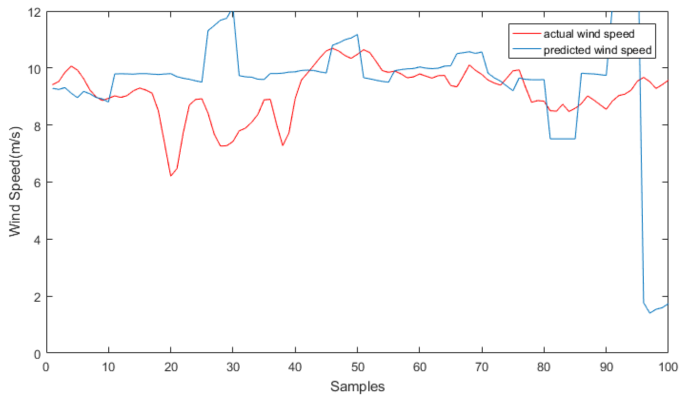

Fig.6. Plot of predicted 1 day ahead wind speed for ANN Model

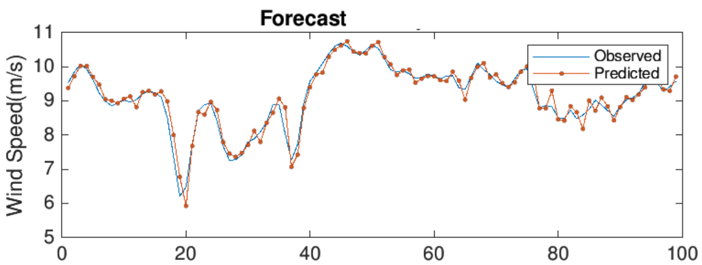

Fig.7. Plot of predicted 1 day ahead wind speed for Deep Learning Model

Table.1. Comparison of MAPE

Model	MAPE (%)
ANN Model	17.25
Deep Learning Model	1.7

8. CONCLUSIONS

This research work presented accurate predictive models which can give a day-ahead prediction for wind speed. Predictive models for wind speed using neural network along with data pre-processing techniques are developed. A new training strategy named recursive training is tested. It is seen that the neural networks which use wavelet decomposition have better prediction performance compared to neural networks without wavelet decomposition. The MAPE for a day-ahead prediction with this model is found to be 17.25%. A novel deep learning LSTM model is proposed in this paper. At each time step, the LSTM network is trained to predict the wind speed for the next time step, thus making the forecast more accurate than other models. The MAPE was found to be 1.7% which is lesser than any model found in literature.

REFERENCES

[1] Qiang Wang, Kin Keung, Lai Dongxiao Niu and Qian Zhang, "A Multivariate Wind Power Forecasting Model Based on LS-SVM", *Proceedings of 5th International Joint Conference on Computational Sciences and Optimization*, pp. 127-131, 2012.

[2] J.P.S. Catalao, H.M.I. Pousinho and V.M.F. Mendes, "Hybrid Wavelet-PSO-ANFIS Approach for short-Term Wind Power Forecasting in Portugal", *IEEE Transactions on Sustainable Energy*, Vol. 2, No. 1, pp 50-59, 2011.

[3] Jianzhou Wang, Jianming Hu, Kailiang Ma and Yixin Zhang, "A Self-Adaptive Hybrid Approach for Wind Speed Forecasting, Renewable Energy", Vol. 78, pp. 374-385, 2015.

[4] Hamed Chitsaz, Nima Amjady and Hamidreza Zareipour, "Wind Power Forecast using Wavelet Neural Network Trained by Improved Clonal Selection Algorithm", *Energy Conversion and Management*, Vol. 89, pp. 588-598, 2015.

[5] Nima Amjady, "New Wind Speed Forecasting Approaches using Fast Ensemble Empirical Model Decomposition, Genetic Algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks", *Renewable Energy*", Vol. 83, pp. 1066-1075, 2015.

[6] Jinbin Wen, Xin Wang, Yihui Zheng, Lixue Li, Lidan Zhou, Gang Yao and Hongtao Chen, "Short-Term Wind Power Forecasting based on Lifting Wavelet Transform and SVM", *Proceedings of IEEE Conference on Power Engineering and Automation*, pp. 1-4, 2012.

[7] Lei Yang, Miao He, Junshan Zhang and V. Vittal, "Support-Vector-Machine Enhanced Markov Model for Short-Term Wind Power Forecast", *IEEE Transactions on Sustainable Energy*, Vol. 6, No. 3, pp. 791-799, 2015.

[8] Ergin Erdem, Jing Shi and Yidong Peng, "Short-Term Forecasting of Wind Speed and Power-A Clustering Approach", *Proceedings of Research Conference on Industrial and Systems Engineering*, pp. 1-8, 2014.

[9] Cong Feng, Mingjian Cui, Bri-Mathias Hodge and Jie Zhang, "A Data-Driven Multi-Model Methodology with Deep Feature Selection for Short-Term Wind Forecasting", *Applied Energy*, Vol. 190, pp. 1245-1257, 2017.

[10] Andre Gensler, Janosch Henze, Bernhard Sick and Nils Raabe, "Deep Learning for Solar Power Forecasting-An

Approach using Autoencoder and LSTM Neural Networks", *Proceedings of IEEE International Conference on Systems, Man and Cybernetrics*, pp. 236-244, 2016.

[11] Mohamed Abdel-Nasser and Karar Mahmoud, "Accurate Photovoltaic Power Forecasting Models using Deep LSTM-RNN", *Neural Computing and Applications*, Vol. 31, No. 7, pp. 2727-2740, 2019.

[12] Mladen Dalto, Jadranko Matusko and Mario Vasak, "Deep Neural Networks for Ultra-Short-Term Wind Forecasting", *Proceedings of IEEE International Conference on Industrial Technology*, pp. 17-19, 2015.

[13] V. Prema, B.S. Jnaneswar, C.A. Badarish, Patil Shreenidhi Ashok, S. Agarwal and K. Uma Rao, "Novel Training Strategies for Wavelet-Neuro Models for Wind Speed Prediction", *Proceedings of IEEE International Conference on Information Technologies, Databases and Information Systems*, pp. 1-4, 2015.

[14] V. Prema and K. Uma Rao "Interactive Graphical User Interface (GUI) for Wind Speed Prediction Using Wavelet and Artificial Neural Network", *Journal of the Institution of Engineers (India): Series B*, Vol. 99, No. 5, pp. 467-477, 2018.

[15] Andre Gensler, Janosch Henze, Bernhard Sick and Nils Raabe, "Deep Learning for Solar Power forecasting-An Approach using Auto Encoder and LSTM Neural Networks", *Proceedings of IEEE International Conference on Systems, Man and Cybernetrics*, pp. 1-8, 2016.