CONFERENCE PROCEEDINGS

of

“3rd International Conference

On

Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things”

(ICETCE-2020)

07th-08th February 2020

Technical Sponsor:IEEE Delhi Section

Conference Record # 48199

Contributors:

Organized By:

Department of Computer Science & Engineering(NBA Accredited)
Department of Information Technology (NBA Accredited)
Swami Keshvanand Institute of Technology, Management & Gramothan, Jagatpura, Jaipur-302017
Telephone: 0141-5160400 (Ext. 286, 284, 213), Fax: 014-2759555
Website: https://icetce.skit.ac.in/2020

TABLE OF CONTENTS

3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE-2020), 07-08 February 2020, (IEEE Conference Record 48199

Short term forecasting based on hourly wind speed data using deep learning algorithms

1st Vikash Kumar Saini
Centre for Energy and Environment
Malaviya National Institute of Technology
Jaipur, India
sainivikash73@gmail.com

2nd Rajesh Kumar, 3rd Akhilesh Mathur
Department of Electrical Engineering
Malaviya National Institute of Technology
Jaipur, India
rkumar.ee@mnit.ac.in,
akhilesh.ee@mnit.ac.in

4th Akash Saxena
Department of Electrical Engineering
SKIT
Jaipur, India
aakash.saxena@hotmail.com

***Abstract*—Application of Internet of Things (IoT) in smart grid is evident in current trends. Smart grid management have greater impact on market economics, security, and distribution of energy. Smart grid is an integration of several components like, wind, solar, cyber-security etc. One of major concern in smart grid is optimal control of wind generation and accurate prediction of wind speed. This paper aims to predict the wind speed with meteorological time series data as input variable using deep learning topology for one-year wind speed data. The dynamic recurrent type network (RNN) integrates and processed with the Extreme Learning-Machine (ELM), nonlinear autoregressive network with exogenous inputs (NARX), and Long short-term memory (LSTM) model. Three models having the same Network's architecture, intermediate layer in architecture have 19 neurons and an activation function. Feature selection method is used for feature extraction from wind data (have four features as wind speed, pressure, humidity, air temperature) and applied to models. Comparative analysis of different models are assessed by performance matrices such as MAPE, MAE, and RMSE.**

***Index Terms*—Deep Learning, Wind Forecasting, RNN, LSTM**

I. Introduction

Innovation in communication structure like IoT leads to modern development in various fields. Specially integration of IoT in smart grid leads to intelligent power system. [1], [2]. In order to improve the reliability of smart grid, a good number of IoT devices are deployed at certain location in power system network. Integration of renewable energy resources with intelligent smart grid leads to further enhancement in efficiency reliability, economy and flexibility of grid. Benefit of renewable energy integration with grid helps in cutting of carbon emission, i.e., makes environmentally friendly. In the renewable energy sources (e.g.,wind,biomass, solar and geothermal energy), wind energy is one of the most prominent and potentially advantageous energy resources. Target set by Indian Government for renewable generation is 175 GW by 2022 [3]. Global installed wind power capacity is rapidly increasing day by day, India have become the fourth largest installed wind capacity in world. Till Oct 2019, total cumulative installation of renewable power generation achieved by India is approx. 85 GW, out of which total wind generation installation is 36.93 GW.

Integration of wind generation with grid is a major challenge because of stochastic nature, intermittent & volatility in wind speed. Forecasting of wind speed helps in overcoming the variability nature of energy production due to fluctuating weather condition (such as wind speed, temperature, pressure, humidity, altitude, latitude). These factors have strong correlation with each other, which introduces the randomness in wind energy that makes prediction of wind speed more difficult that leads in poor performance and accuracy.

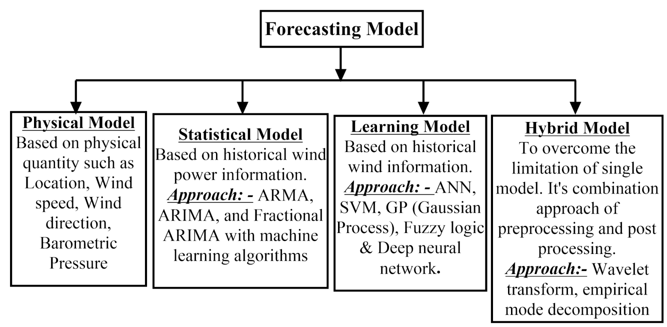

Fig. 1. Categorization of Forecasting Model

In literature, various authors have proposed various methodologies to enhance the ability and accuracy in forecasting of wind speed citeng. These models can be categorized in terms of modeling theory [4] as shown in Fig 1. Likewise, time series method depends on huge amount of data. Parameter estimation, model identification and validation are used for deduction of mathematical model, Thus, this model is used for prediction of wind speed. Most commonly used model in time series is auto regressive and moving average [5], but major drawback of this model is poor performance in terms of prediction accuracy.

Recently, Machine learning techniques are used to prediction of wind power/ wind speed [6]. In literature, several topologies using an artificial neural network (ANN) methods is widely applied for wind energy prediction [7]. ANN have ability of self-learn, self-organize, and self-adaptiveness to

978-1-7281-1683-9/20/$31.00 ©2020

dynamic environment compare to time series method [8], [9]. Some other methods based on fuzzy logic, spatial correlation also reported in literature [10]. But for any machine learning technique to find out the pattern in data and provide an better estimation of prediction is difficult task due to stochastic nature of wind. Thus, in literature, many researcher's interprets this problem as time series forecasting because wind depends on humidity, pressure and temperature for a certain period like a day, month or year. Thus, among the above prediction methods, the neural network method used as time series provides high accuracy and strong adaptive ability to dynamic conditions [11].

In our paper, we have done comparative analysis of deep learning techniques such as Extreme Learning Machine (ELM), Nonlinear Autoregressive Network with Exogenous Input (NARX) and Long Short-Term Memory (LSTM). ELM is a single-layer feedforward neural network algorithm proposed by professor Huang [12]. When data is too large that have non-linear and non-Gaussian degree, traditional loss function optimization is not able to obtain prominent solution. In order to solve this problem, several loss functions like hinge loss, ridge loss etc. used in ELM to enhance its prediction accuracy [13], [14]. LSTM is also provide an alternative solution to this problem. To ensure the accuracy in prediction of wind speed, feature selection by the regression technique is used. Therefore, LSTM model suggested to short-term wind forecasting for less sample size of data. The tools used in our paper is Python Application Programming Interface (API) Keras for simulating all models. All three models are simulated in Python® platform with a Windows based Personal Computer Intel® Core™ i5 CPU having 3.20 GHz processor and 4 GB RAM with 64 bit operating system. Thus, main contributions of this study is listed as follows:

- Short-term wind power forecasting was performed using different deep learning algorithms.
- Effect of Air Temperature, Pressure, Humidity, dew point on prediction of wind speed is studied.
- Effect of varying number of neurons in architecture in learning models is analyzed.
- Comparative analysis of different deep learning method such as ELM, NARX and LSTM based on performance matrices like MAPE, MAE and RMSE is done.

The rest of this paper is organized as follows. Section 2 describes a brief review of suggested models which is used for wind forecasting. Section 3 proposes the wind power prediction scheme based on RNN. The discussion on prediction results and their accuracy in Section 4. Finally, Section 5 concludes with a summary of the main contributions and provides future research directions.

II. Methodology

Prediction of wind energy is a very exciting problem. Thus, accurate and reliable forecasting of wind is a complex task, these is due to large uncertainties present in weather condition. Several deep learning models for prediction is presented in literature for better accuracy in prediction for time series data. In this paper we discuss RNN based deep learning models.

A. *Recurrent Neural Network (RNN)*

Major drawback of feed-forward neural network is that it doesn't have feedback connection. So, when prediction of future steps in sequence is required, feed-forward is not have that forecasting capability because prediction require a memory element in terms of feedback loop. Therefore, Recurrent Neural Network comes in picture. Recurrent Neural Network (RNN) is special type of Artificial Neural Network (ANN) as shown in Fig. (2). Architecture have a feedback loop inside itself means output of prior step is fed to current state. Significant feature have memory element inside its architecture as shown in fig. Memory element stores the information about the data sequence. Parameters of RNN are same for all layers present in architecture because it performs the same task of prediction. So, optimization of parameter value of network is reduced comparable to other neural networks. With increasing the number of layers in architecture, finding the network parameters is complex task due to vanishing and explosion in gradient calculation is further execrated, a major drawback of RNN. Explosion in gradient calculation leads to difficulty in capturing capability of long-term information in network, thus further leads to vanishing of significant information. And also create a complexity in memory for information storage due to exponential rise in gradient multiplication with increase in number of hidden layer. Thus, Long Short-Term Memory units (LSTM) comes in picture to rescue the RNN from explosion of gradient and data vanishing.

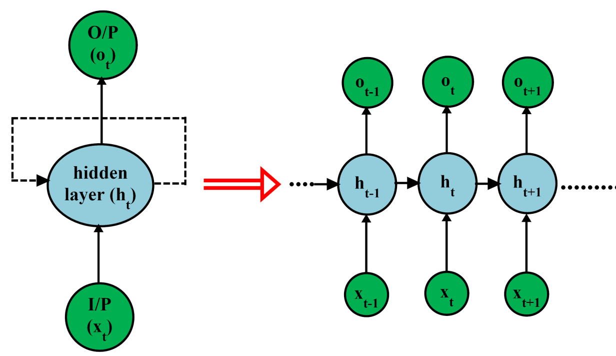

Fig. 2. RNN Architecture

B. *Long Short-Term Memory (LSTM)*

LSTM replaces the conventional neurons architecture present in internal structure of RNN by memory blocks, it enhances the ability to deal with dependencies problems present in RNN. Complete structure of LSTM prediction model with time series multi-variable wind speed data is shown in Fig.3. It shows the key components of LSTM network for memorization of states & updation of information through gates (input, forget & output gates) [15]. Thus, gates perform read, write & reset operation in cell. Therefore, *Input gate* provides the operation for updating the memory cell from input side, *Forget gate* decides which information is needed for storage and discard, & lastly, *Output gate* gives the output based on value present in memory cell and input gate. Updating of cell state & computation of results is described by following equation (1-6),

$$i_t = \sigma(W_i x_t + V_i h_{t-1} + b_i) \tag{1}$$

$$f_t = \sigma(W_f x_t + V_f h_{t-1} + b_f) \tag{2}$$

$$o_t = \sigma(W_o x_t + V_o h_{t-1} + b_o) \tag{3}$$

$$\bar{c_t} = \odot p(W_c x_t + V_c h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t C_{t-1} + i_t \bar{C_t} \tag{5}$$

$$h_t = o_t \odot q c_t \tag{6}$$

where x_t is the input vector, i_t, f_t and o_t represent the different gates (input, forget, and output respectively), c represents the memory cell and h represents the hidden layer. Hadamard product i.e., the element-wise multiplication between two vectors is represented by $\odot$. Here, b is bias vector, whereas W & V denotes the related weighting vectors. σ, p & q are the set of different gates, input & output activation function respectively.

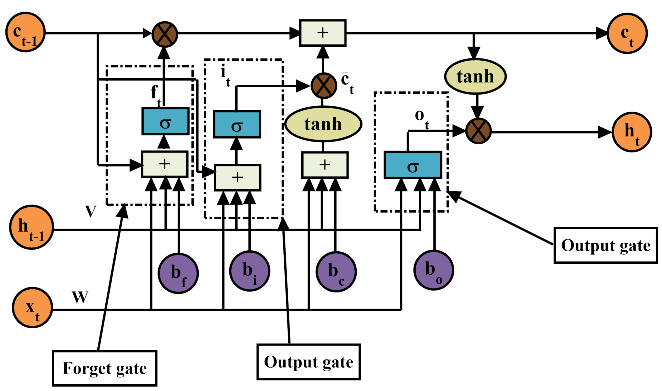

Fig. 3. Structure of LSTM

LSTM overcome the challenge posed by RNN networks such as vanishing of information & exploding of gradient internally. Thus, LSTM provides exciting results for complex sequencial time series data.

C. *Nonlinear Autoregressive exogenous (NARX)*

NARX is a dynamic recurrent network with many layers with feedback connections. It is suitable for modeling of time series data as nonlinear system. Significant property of NARX neural network are :

- Learning of model for time series data in NARX is better than other neural network that are based on gradient descent method.
- Convergence rate of these networks is good [16], [17].

NARX model is trained by either parallel or series-parallel topology. In, parallel topology output is feedback to input of architecture whereas series-parallel topology true input is

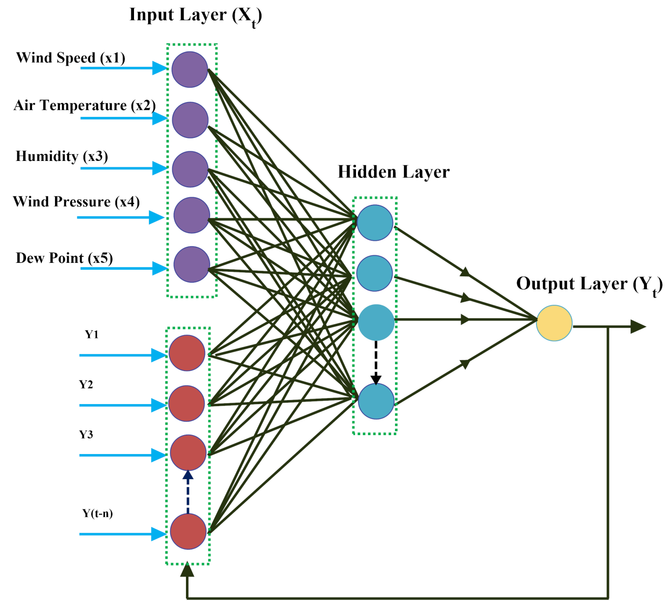

Fig. 4. Architecture of NARX model

applied back instead of estimated output [18]. Architecture of NARX model based on parallel topology is shown in Fig.4 Where input is divided into two parts namely, external input i.e., $w(t)$ and previous output i.e., $s(t)$ at time t. Output of NARX architecture is given by (7) & (8),

$$y(t) = f(y(t-1), y(t-2), y(t-3) ....y(t-n) \tag{7}$$

$$x(t) = f(x(t-1), x(t-2), x(t-3) ....x(t-n_i) \tag{8}$$

Where, n_i is noise delay time. Thus, output of architecture defined by $y(t)$ depends on current & previous state output of neural network i.e., $x(t-i)$ and $y(t-i)$ [19].

D. *Extreme Learning Machine (ELM)*

ELM is single-hidden layer based feedforward neural network [20], important feature of this architecture the total time taken for fine tuning of network parameters, such as hidden layer node, number of neurons in hidden layers, is significantly reduced. Fine tuning of network parameters helps in order to find optimal solution. Neural network weights are obtained by minimizing the square loss function. Architecture of ELM is shown in Fig. 5

Consider N training samples from datasets (x_j, t_j) given by (9) & (10),

$$X_i = [x_{i1}, x_{i2}, ...x_{in}]^T \in R^n \tag{9}$$

$$t_i = [t_{i1}, t_{i2}, ...t_{in}]^T \in R^m \tag{10}$$

If J hidden layers are present in architecture, then expected output is given by (11) & (12),

$$\sum_{j=1}^{J} \beta_j a(w_j x_i + h_j) = o_i j = 1, 2, 3, ...N \tag{11}$$

$$a(x) = 1/1 + \exp(x) \tag{12}$$

where, $a(x)$ is activation function, h_j is the j^{th} bias of hidden layer, $W_j = [w_{j1}, w_{j2}, ..., w_{jn}]^T$ is weight matrix, $\beta_j = [j1, j2, ..., jm]^T$ is j^{th} input/output weights of hidden layer, o_i is the expected output and $W_j \cdot X_i$ is described as inner product of W_j and X_i.

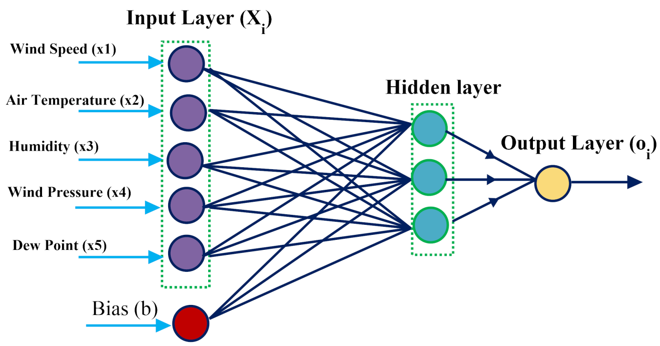

Fig. 5. Architecture of ELM

III. Discussion

A. Procedure

Step:I Data Collection → Collect the metrological wind dataset for the specified location Jodhpur (26.23890^o N X 73.02430^o E) from Dark Sky [21] for the year 2015. The time interval was hourly, total samples are 8760 with five features i.e., wind speed(m/s), air temperature($^o c$), Pressure, Dew point, and Humidity

Step:II Normalization and Standardization→ It is done by MinMaxScaler & StandardScaler objects in scikit-learn package using python library .

Step:III Split the Dataset→ After the normalization of dataset, the data-set is divided into training and testing sets for evaluation.

Step:IV Define the Model→ The Model is defined using Keras in python 3.6 as a sequence of neural layers.

Step:V Compile the Model→ After defining the model, its compilation is done. Compilation is a step that transforms the sequence of different layers into effective series of matrix for execution using GPU or CPU units, depends on the configuration of keras.

Step:VI Fit the Model→ After the compilation of model, the model is fitted on training data-set. It tells that model parameters are well trained on training data-set. The training data-set is specified as matrix of input features and output label as represented by X & y respectively.

Step:VII Evaluate the Model→ Model is validated on training data set using performance indices and classification accuracy.

Step:VIII Make Predictions on the Model→ After training & validation of model on data set, we will make prediction of wind speed using testing data-set and performance is evaluated using MAPE, MAE & RMSE.

Take care of over-fitting & under-fitting of model when model is trained. Fig.6 summarizes the complete procedure of implementation of models as given by above steps.

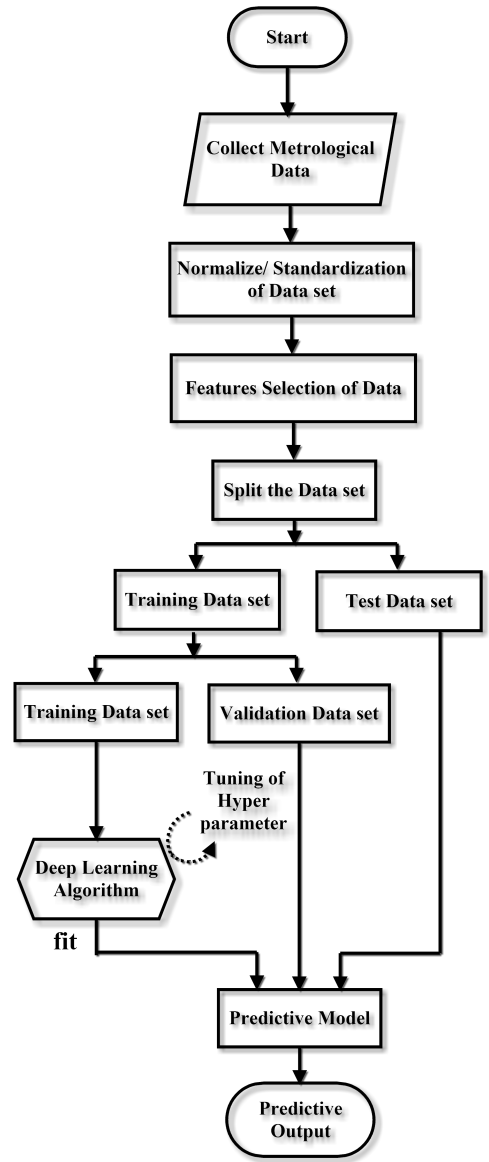

Fig. 6. flow chart

B. Data Processing

There is a lot of uncertainty and inconsistency in the data. Thus, it would be difficult to train the model on whole year's data simultaneously. Therefore, It is better to categorizes the data according to length and then process it. The plot of wind data features is shown in Fig 7. The data categorizes into three group namely, G1: Oct-Dec, G2: July-Dec & G3: Jan-Dec

C. Wind Turbine Specification

The specification of wind turbine are required for the analysis of wind power generation from wind speed. Thus, the technical specification of wind turbine is taken from SUZLON S97-2.1 MW turbine as shown in Table I. As per data given by DarkSky this model is installed at Jodhpur (26.23890^o N X 73.02430^o E). The relationship between output power from turbine and wind turbine speed is given by (13). Due

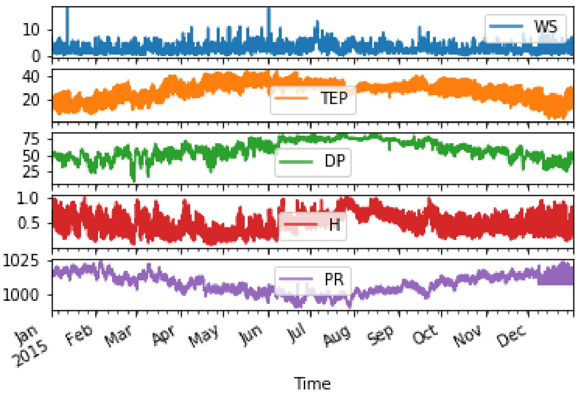

Fig. 7. Features of Wind Speed

TABLE I
SPECIFICATION PARAMETER OF TURBINE [22]

Specifications	S-97
Rated Power	2100 kW
Rotor Diameter	97 m
No. of Blades	3
Swept area	7386 m^2
Air density	1.225 kg/m^3
Rotational speed	Variable 12-15.7 rpm
Cut out wind speed	20 m/s
Cut in wind speed	3.5 m/s
Rated wind speed	11 m/s

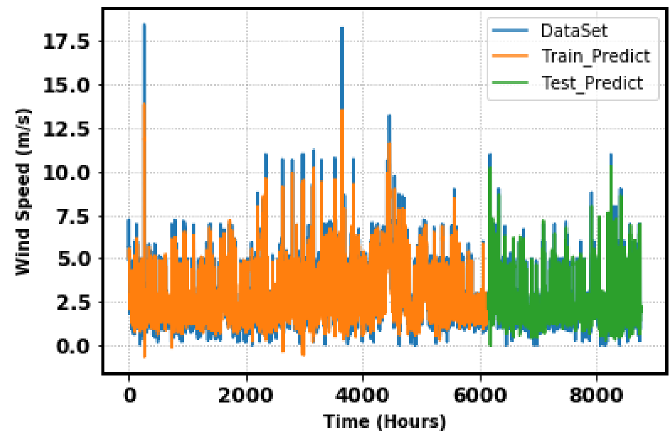

Fig. 8. Actual v/s Predicted Wind Speed

TABLE II
ACCURACY PARAMETER

Data Length	Type of error	ELM	NARX	LSTM
Oct-Dec	MAPE	0.5699	0.3737	0.3017
	MAE	0.3978	1.1624	0.8742
	RMSE	0.9215	0.5657	1.2360
July-Dec	MAPE	0.4304	0.3476	0.4364
	MAE	1.3139	0.9726	0.4464
	RMSE	1.5462	1.3394	0.5505
Jan-Dec	MAPE	0.3945	0.3762	0.2298
	MAE	1.1035	1.0522	0.6769
	RMSE	1.3573	1.4808	1.0594

to uncertainty in speed, the output power of wind turbine is continuously changing.

D. Wind Power Prediction

Using turbine specification, Generated wind power is found out by (13),

$$P = \frac{1}{2}\rho A C_p V^3 \tag{13}$$

Where P is the power in $watts$, V is speed in m/s. ρ is air density in kg/m^3 , A is turbine rotor area in m^2 , C_p is power coefficient. Also, wind turbines cannot operate at maximum limit (C_{pmax} =0.59). Power Coefficient value is unique to each turbine type and it is a function of wind speed. Practical limit is below than Betz Limit with common range of 0.35-0.45 in best designed wind turbines. Fig. 8 shows the estimated wind power speed in m/s for the geographical location under study.

E. Accuracy Assessment

Performance indices such as Root mean square Error (RMSE), Mean Absolute Error (MAE) and Absolute Percentage Error (MAPE) are used for evaluation accuracy assessment for wind speed forecasting. Various Performance indices is given by (14-16),

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(E_t - P_t)^2} \tag{14}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{E_t - P_t}{E_t} \tag{15}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \mid E_t - P_t \mid \tag{16}$$

where, P_t is predicted value, E_t is measured value and N is size of data samples at time horizon. Prediction Error (PE) is defined by (17),

$$PE = P_t - E_t \tag{17}$$

Table II comprises the type of error for different data length for ELM, NARX and LSTM. Out of different methods, LSTM outperform other methods. Prediction accuracy is also carried out by varying the number of neurons present in the hidden layer. Thus variation of MAPE, MAE & MSE with number of neurons in the hidden layer are shown in Fig. 9. It shows that optimal number of neurons is **19** for trade-off between different performance indices.

IV. CONCLUSION

In this paper, Prediction of wind speed using deep learning is implemented. Here, we does not have any knowledge what wind speed is going to be in future and simultaneously other parameter of weather is also dynamic in nature. So, we try to predict the power only by analyzing pattern in the past data using different RNN based models. LSTM analyzed the prior data and tried to get useful insightful knowledge about the patterns in data. And using that knowledge it is used for prediction of wind speed and compared its performance w.r.t NARX and ELM. LSTM outperforms other methods as shown in Table II. As a future scope, hybrid model of deep learning based on neuro-evolution will be developed & combined with Decision Tree/Random Forest and uncertainty in wind features can be accommodated in model by probabilistic modeling and parameters of model can be optimized using bayesian optimization.

ACKNOWLEDGEMENT

Authors would like to thank Mr. Bharat Singh, Research Scholar, Department of Electrical Engineering, MNIT, Jaipur, India, and Mr. Mahipal Bukya, Assistant Professor, Department of Electrical Engineering, Manipal University, Jaipur, India for their helpful suggestion and contribution, which help us to improve the quality of paper.

REFERENCES

[1] Z. Guan, J. Li, L. Wu, Y. Zhang, J. Wu, and X. Du, "Achieving efficient and secure data acquisition for cloud-supported internet of things in smart grid," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1934–1944, 2017.

[2] S. Reddy, R. Kumar, B. Panigrahi, *et al.*, "Cloud energy storage management system with price fluctuations and distributed generation intermittency," in *2018 8th IEEE India International Conference on Power Electronics (IICPE)*, pp. 1–6, IEEE, 2018.

[3] MNRE, "Ministry of new and renewable energy, government of india." Available online: http://mnre.gov.in/.

[4] J. Zhang, M. Cui, B.-M. Hodge, A. Florita, and J. Freedman, "Ramp forecasting performance from improved short-term wind power forecasting over multiple spatial and temporal scales," *Energy*, vol. 122, pp. 528–541, 2017.

[5] Y. Cao, Y. Liu, D. Zhang, W. Wang, and Z. Chen, "Wind power ultra-short-term forecasting method combined with pattern-matching and arma-model," in *2013 IEEE Grenoble Conference*, pp. 1–4, IEEE, 2013.

[6] N. Bokde, A. Feijóo, D. Villanueva, and K. Kulat, "A novel and alternative approach for direct and indirect wind-power prediction methods," *Energies*, vol. 11, no. 11, p. 2923, 2018.

[7] K. P. Senthil, "Improved prediction of wind speed using machine learning," *EAI Endorsed Transactions on Energy Web*, vol. 6, no. 23, 2019.

[8] G. Zhang, L. Zhang, and T. Xie, "Prediction of short-term wind power in wind power plant based on bp-ann," in *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pp. 75–79, IEEE, 2016.

[9] D. Fadare, "The application of artificial neural networks to mapping of wind speed profile for energy application in nigeria," *Applied Energy*, vol. 87, no. 3, pp. 934–942, 2010.

[10] I. G. Damousis, M. C. Alexiadis, J. B. Theocharis, and P. S. Dokopoulos, "A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation," *IEEE Transactions on Energy Conversion*, vol. 19, no. 2, pp. 352–361, 2004.

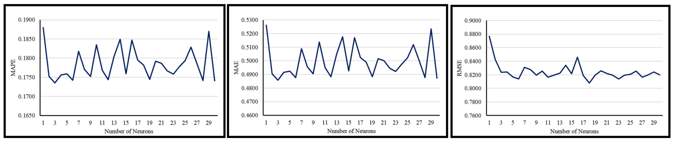

Fig. 9. Effect of Neurons on Performance Indices

[11] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE transactions on neural networks and learning systems*, vol. 25, no. 2, pp. 303–315, 2013.

[12] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[13] W. Liu, P. P. Pokharel, and J. C. Príncipe, "Correntropy: Properties and applications in non-gaussian signal processing," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5286–5298, 2007.

[14] G. Sideratos and N. D. Hatziargyriou, "An advanced statistical method for wind power forecasting," *IEEE Transactions on power systems*, vol. 22, no. 1, pp. 258–265, 2007.

[15] M.-R. Chen, G.-Q. Zeng, K.-D. Lu, and J. Weng, "A two-layer nonlinear combination method for short-term wind speed prediction based on elm, enn and lstm," *IEEE Internet of Things Journal*, 2019.

[16] E. Diaconescu, "The use of narx neural networks to predict chaotic time series," *Wseas Transactions on computer research*, vol. 3, no. 3, pp. 182–191, 2008.

[17] T. Lin, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in narx recurrent neural networks," *IEEE Transactions on Neural Networks*, vol. 7, no. 6, pp. 1329–1338, 1996.

[18] H. Xie, H. Tang, and Y.-H. Liao, "Time series prediction based on narx neural networks: An advanced approach," in *2009 International conference on machine learning and cybernetics*, vol. 3, pp. 1275–1279, IEEE, 2009.

[19] R. A. Khalil, "Comparison of four neural network learning methods based on genetic algorithm for non-linear dynamic systems identification," *AL Rafdain Engineering Journal*, vol. 20, no. 1, pp. 122–132, 2012.

[20] N. Li, F. He, and W. Ma, "Wind power prediction based on extreme learning machine with kernel mean p-power error loss," *Energies*, vol. 12, no. 4, p. 673, 2019.

[21] Available online: https://darksky.net/dev.

[22] V. Reddy, S. M. Verma, K. Verma, and R. Kumar, "Hybrid approach for short term wind power forecasting," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–5, IEEE, 2018.