ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/369888158

Energy Demand Forecasting Using Machine Learning Perspective Bangladesh

Conference Paper · March 2023

DOI: 10.1109/GlobConHT56829.2023.10087679

CITATIONS
3

READS
213

4 authors, including:

Avijit Paul Piyal
BRAC
1 PUBLICATION **3** CITATIONS

SEE PROFILE

Khan Fahad Rahman
BRAC University
1 PUBLICATION **3** CITATIONS

SEE PROFILE

Abu S. M. Mohsin
BRAC University
61 PUBLICATIONS **518** CITATIONS

SEE PROFILE

All content following this page was uploaded by Abu S. M. Mohsin on 03 May 2024.

The user has requested enhancement of the downloaded file.

Energy Demand Forecasting Using Machine Learning Perspective Bangladesh

Avijit Paul Piyal[1]
Department of Electrical and Electronic Engineering
BRAC University
Dhaka, Bangladesh
avijit.paul.piyal@g.bracu.ac.bd

Siam Ahmed[2]
Department of Electrical and Electronic Engineering
BRAC University
Dhaka, Bangladesh
siam.ahmed@g.bracu.ac.bd

Khan Fahad Rahman[3]
Department of Electrical and Electronic Engineering
BRAC University
Dhaka, Bangladesh
khan.fahad.rahman @g.bracu.ac.bd

Abu S. M. Mohsin[4*]
Department of Electrical and Electronic Engineering
BRAC University
Dhaka, Bangladesh
asm.mohsin@bracu.ac.bd

***Abstract*— Bangladesh is a largely populated country with a total area of 1, 47,570 square km and per capita electricity generation of 182kWh, which is one of the world's lowest. Supplying an uninterrupted power supply to this huge population becomes a challenge for the govt. of Bangladesh. Therefore it becomes necessary to use modern energy management tools like machine learning-based load forecasting techniques to make the decision-making action more efficient. Due to the chaotic nature of electric load demand, an artificial neural network (ANN) is preferred for electrical load forecasting purposes. In this study, we explored several machine learning algorithms like Long Short-Term Memory Network (LSTM), Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors (SARIMAX), and Fbprophet on 11 years of power generation data (2003 to 2014) of Bangladesh to forecast the load demand. The findings of this study reveal that LSTM methods outperformed SARIMAX and Fbprophet methods with the least RMSE and MAPE error 150.26 and 0.4821%. The findings of this study will help the government in policy making and the individual consumer to tackle the energy challenges in the near future.**

Keywords*— *Deep learning, LSTM, SARIMAX, Fbprophet, Max power Generation, load forecasting.

I. Introduction

Electricity is one of the most fundamental requirements of human life, with approximately 20 trillion kWh utilized globally daily. [1] The demand for energy in developing countries has increased dramatically in recent years. Bangladesh has a population of about 167 million or 2.18 percent of the world's population. [2] According to the Ministry of Energy, and Mineral Resources, Bangladesh's government supplied power to 70% of the country's 160 million inhabitants in 2015. [3] The population density is 1278 persons per square kilometer, with 36.5 percent of the population living in urban areas. [3] Electricity demand is increasing in Bangladesh as a result of the extensive usage of electricity with the country's fast urbanization. Though Bangladesh's installed power generating capacity expanded fast to 13265MW with captive generation capacity.

However, with a total installed capacity of 13265 MW, produced a maximum of 8122 MW in the fiscal year 2015, compared to a demand of 10,283 MW. [5, 6]. This is insufficient to meet the country's electricity demand. The demand for power is increasing, but the supply is not keeping up. Furthermore, in sector wise domestic, industry, commercial, irrigation, and other sector consumed electric power are respectively 50.07%, 34.47%, 9.09%, 4.58%, and 1.79%.[4] The sector-wise consumption is shown in Figure 1.

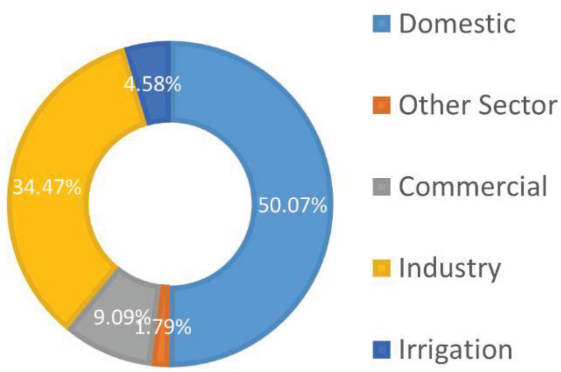

Fig. 1. sector wise electric power consumption [4]

Bangladesh has a higher power generation shortage than demand, which is a barrier to the country's growth and development. To ensure energy security and long-term development, the growth rate of generation must be increased. The main purpose of this study is to help the government in terms of policy making and in order for consumers to handle energy difficulties in the foreseeable future.

This paper is based on several models such as LSTM, SARIMAX, and Fbprophet for forecasting power consumption and comparing their performance. The models are based on actual data which was collected from Bangladesh Power Development Board (BPDB). Furthermore, the LSTM method performed better than the other methods. The LSTM, in particular, is a popular and widely utilized approach. The paper's LSTM model is presented to be decently accurate and has the lowest overall Mean Absolute Percentage Error (MAPE) of 0.4821% and an RMSE of 150.26. Zheng et al. [13] used LSTM to apply load forecasting strategies in smart grids, and their findings suggest that LSTM outperforms the other prediction approaches.

2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT) | 979-8-3503-3211-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/GLOBCONHT56829.2023.10087679

Authorized licensed use limited to: BRAC UNIVERSITY. Downloaded on May 02,2024 at 17:07:55 UTC from IEEE Xplore. Restrictions apply.

II. LITERATURE REVIEW

Electric load forecasting is essential for every country's future power usage planning. Accurate load forecasting is vital in the planning procedures for energy distribution in the smart grid. The government of Bangladesh is finding it more difficult to run and maintain its power system in order to satisfy rising demand and keep up with the developing economy. The majority of Bangladesh load forecasting forecasts have been focused on a regional level. [7] Though some statistics cover the whole country's electrical power demand, some publications focus on daily peak needs. Load forecasting methods may be divided into two types: statistical load forecasting and artificial intelligence-based load forecasting. [4] Traditional statistical approaches have a lot of weaknesses that affect the accuracy of short-term power load forecasting. For example, basic regression functions are ineffective in expressing non-linear and dynamically uncertain prediction relationships.[7] Load forecasting has been approached using a variety of statistical and artificial intelligence approaches.[8] The auto-regressive moving average (ARMA) model was used to predict electricity in the California power market (Nowicka-Zagrajek and Weron 2002). In terms of short-term forecasting accuracy, the autoregressive integrated moving average (ARIMA) method was good. (Hong, Gui, Baran, & Willis, 2010). [9] ARIMA and seasonal ARIMA (SARIMA) methods were used in research (Kim, 2013; Tran, Debusschere, & Bacha, 2012). Their findings suggested that the SARIMA model is preferred. For STLF, simple ARIMA models are sufficient. [10] Md. Rashidul and coworkers proposed (RNN-LSTM) model for forecasting electrical load in Bangladesh's Chattogram region. The suggested LSTM approach outperforms the SVM method in terms of RMSE and MAE value. [15] In another paper Mr. Rahul and coworkers [16], forecasted power demand over a five-year period using the (LSTM-RNN) model. The model was based on real-time power statistics from ISO New England. The forecasts were made every hour. The suggested model was proven to be extremely accurate, with a Mean Absolute Percentage Error.[16] In this study, we would like to explore several load forecasting techniques and find appropriate models for accurate load forecasting.

III. DATASET

A. Data set of demand side

We worked on the power generation dataset of Bangladesh for around 11 years from 2003 to 2014 as we already mentioned that the data was collected from the BPDB website [4] which was freely accessible. The first and last five rows of demand data from the dataset are shown in Table 1 and Table 2.

TABLE I. FIRST FIVE ROWS OF DEMAND

Date	Demand	Month	Year	Time	Day
2003-03-01	12863	3	2003	00:00	Saturday
2003-03-01	12389	3	2003	00:00	Saturday
2003-03-01	12155	3	2003	00:00	Saturday
2003-03-01	12072	3	2003	00:00	Saturday
2003-03-01	12160	3	2003	00:00	Saturday

TABLE II. LAST FIVE ROWS OF DEMAND SIDE

Date	Demand	Month	Year	Time	Day
2014-03-31	16955.0	12	2014	00:00	Wednesday
2014-03-31	16243.0	12	2014	00:00	Wednesday
2014-03-31	15525.0	12	2014	00:00	Wednesday
2014-03-31	14759.0	12	2014	00:00	Wednesday
2014-03-31	14071.0	12	2014	00:00	Wednesday

B. Dataset description of demand side

Table 3 contains the total number of demand data, here it is 103776 where the maximum demand for a day is 27622 MW and the minimum demand for a day is 7794 MW. In addition, the mean demand value of this dataset is 14674.947 MW.

TABLE III. DATA DESCRIPTION OF DEMAND DATA

	Demand	Month	Year
Count	103776.00	103776.000	103776.00
Mean	14674.947493	6.591813	2008.574699
Std	2894.544130	3.420534	3.414726
Min	7794.000	1.00000	2003.0000
25%	12514.0000	4.00000	2006.0000
50%	14773.000	7.00000	2009.0000
75%	16443.000	10.00000	2012.0000
Max	27622.0000	12.00000	2014.0000

C. Pearson correlation Function and heatmap of demand side

We also calculated the Pearson correlation function (Table 4) to identify which has a strong correlation with on-demand data as it measured the strength of the linear relationship between two variables. Pearson correlation shows a value between -1 to 1 where 0 is no correlation, + 1 means a total positive correlation, and -1 means a total negative linear correlation that found it is negatively correlated with the month and positively correlated with temperature and demand. (Figure 2).

Furthermore, a heatmap is a figure that contains two-dimensional matrix data to visualize the numerical data in the form of cells where each cell of the heatmap is colored and the shades of the color represent the relationship of the value with the dataframe.

TABLE IV. PEARSON CORRELATION FUNCTION OF DEMAND DATA

	Demand	Temp	Month
Demand	1.00000	0.208075	-0.013437
Temp	0.208075	1.00000	0.120281
Month	-0.013437	0.120281	1.00000

Authorized licensed use limited to: BRAC UNIVERSITY. Downloaded on May 02,2024 at 17:07:55 UTC from IEEE Xplore. Restrictions apply.

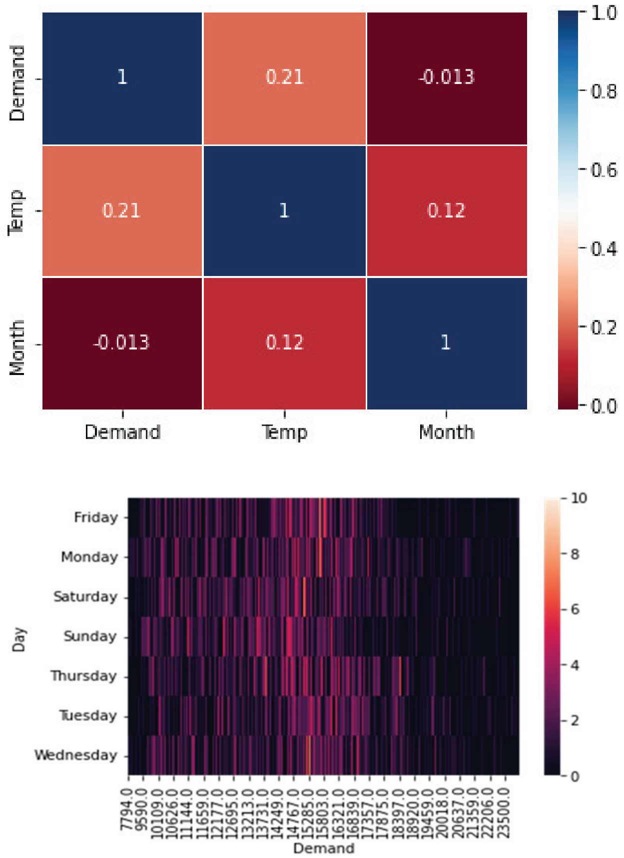

Fig. 2. Heat map of Pearson correlation Function of demand data (a) and b) demand vs. day.

IV. PERFORMANCE MATRICES

To evaluate the performance of the model, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) were determined. The difference between predicted and truth value for each data point was calculated and using the mean and square root of that mean RMSE was computed.

Absolute percent errors were calculated for the available data point and MAPE was calculated by dividing the absolute error by the total number of data point. In the below formulas, x_t is the real value, $\widehat{x_t}$ is the predicted value and t indicates the time.

$$RMSE = \sqrt[2]{\frac{1}{n}\sum_{t=1}^{n} \quad (x_t - \widehat{x_t})^2} \tag{1}$$

$$MAPE = \frac{\sum_{t=1}^{n} \quad |(x_t - \widehat{x_t})|}{n} * 100\% \tag{2}$$

V. LOAD FORECASTING MODEL DEVELOPMENT

A. Training and Test split of load consumption of LSTM

We have categorized the collection of the dataset on a daily basis where we group the demand value by multiplying it with the maximum value of it to get the daily maximum value and stored the maximum of the daily maximum. Later, we split the dataset into two parts called the train side and test side where we used the train side data to train the model and try to predict future values with this train dataset. After that, we compared test data and predicted data to verify the stability of the method we used. Moreover, we maintained a ratio for splitting the dataset which is 65:35. Then we reshaped the train and test dataset according to the requirement of LSTM models.

B. Building the model of load consumption

Then, we imported various libraries to scale our dataset and featured the range. After that, we started converting an array of values into a dataset matrix by using some built-in functions. With the numpy function and the coordination of the array, we get the dataset matrix.

Furthermore, we have taken the time laps for 7 days to train the side dataset and stored these by using the built-in function. Again, we reshaped the input such as samples, time step, and features required for constructing the LSTM models.

C. Training the data of load consumption

We have imported various functions such as sequential, dropout, dense, LSTM, adam, etc. for optimizing the model, making the layer, building the sequence, and many more from different libraries. In addition, these functions were used for building the LSTM model. Sequential regression had been used and applied to build the LSTM model. For adding the first layer, we set LSTM as 50 to build the more stable model. Later, we add an output layer where we use Dense as unit 1. For compiling the (RNN) model, we implement an optimizer called adam and a performance matrix called mean absolute percentage error (MAPE) to calculate the loss. Finally, to train the model, we need to implement an Epoch in the learning algorithm where every sample in the training dataset has a chance to update the internal model once during an epoch. Epoch defines the number of times the entire dataset has to be worked through the learning algorithm. In this model, wan e used epoch of 200 and a batch size of 64.

VI. RESULT ANALYSIS

A. Load Forecasting using LSTM

Besides, we evaluate through the test side data with the taken time steps with the help of a function, we discovered history. Then we do the prediction and check the performance matrices.

Furthermore, we used a function called matplotlib for plotting our final prediction in Figure 3, where we used the X-axis as dates in a year and the Y-axis as load consumption in MW. The orange color indicated the true data and the blue color indicated the predicted demand.

In this model, we calculated RMSE and MAPE to evaluate the performance of the model. We train the model

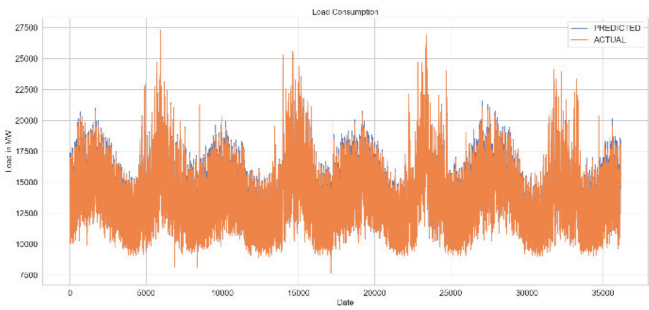

Fig. 3. LSTM prediction for demand side

Authorized licensed use limited to: BRAC UNIVERSITY. Downloaded on May 02,2024 at 17:07:55 UTC from IEEE Xplore. Restrictions apply.

with train side data and predict the future value with these train side data. The LSTM model generated a root mean square error (RMSE) of 150.26 and a mean absolute percentage error (MAPE) of 0.4821 percent.

B. Load Forecasting using SARIMAX.

SARIMAX, or Seasonal Auto-Regressive Integrated Moving Average with eXogenous Factors, is a model that extends the ARIMA class. ARIMA models are made up of two parts: the autoregressive term (AR) and the moving-average term (MA).

SARIMA (p,d,q)*(p,d,q,s)…… .(9) [17]

Here, P = seasonal AR order, D = seasonal differencing, Q = seasonal MA order, and S = length of repeating seasonal pattern, where p = non-seasonal autoregressive (AR) order, d = nonseasonal differencing, q= non-seasonal moving average (MA) order.

For building the SARIMAX model, the combination of p, d and q must be determined which is now set by giving the range (0 to 2) and the format of p, d, q, and seasonal value can be determined.

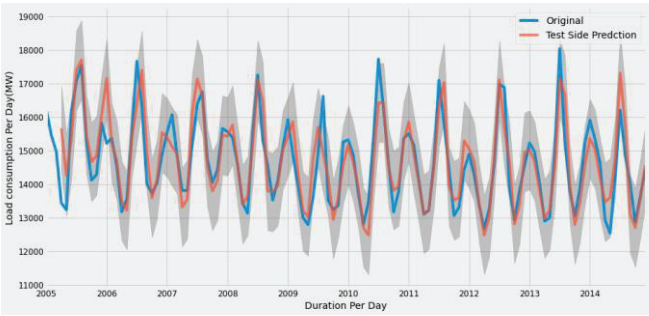

Fig. 4. Original and test side prediction graph using SARIMAX

In this model, we call the ARIMA function and provide five parameters to support the model which are order p, d, q and seasonal_order, df, enforce_stationarity, enforce_invertibility. Therefore, to fit this SARIMAX model, a built-in function has been applied where order is given as (1,1,1) and seasonal order is given as (1,1,1,12) and enforce_stationarity and enforce_invertibility are fixed as "False". After determining all the parameters, a function is called to plot standardized residual, Histogram plus estimated density, Normal Q-Q and correlogram where correlation and Autocorrelation is justified which verified after p value is justified. Finally, with the x-axis set to data and the y-axis set to load consumption, the value of last years' time series is programmed as the Test side, and the prediction is made for this time series data. The blue color indicated the original data and the orange color indicated the test side prediction data. The SARIMAX model (Figure 4) produces a performance matrix with RMSE values of 573.57, and MAPE values of 3.15 percent.

C. Prediction of load consumption using Fbprophet

In addition, for predicting the future value (Figure 5), a function has been called where the period is set as 1 and stored it.

Then, the value of stored data will go through a function to predict and store all this information. Next, we implemented a function to plot the graph and the stored values are constructed by using this function.

Finally, cross-validation comes into action which is required to justify all the information. To check the validation of the dataset a function was imported. Then, a frame has been introduced to pass the network through this validation process where the initial is used as 730 days, the period is used as 180 days and the horizon is used as 365 days. And all these values are stored in the frame.

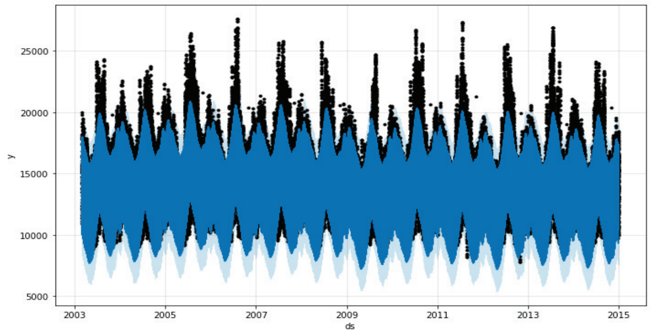

Fig. 5. Prediction plot of load consumption

The RMSE (1852.55699), MAPE (0.091914%), MAE (1425.256563), and MSE (3.431966) results are calculated using the Fbprophet model (Table 6). The comparison table indicates that the LSTM model produces better results compared to the reference paper.

TABLE V. OUR CALCULATED RESULTS ARE COMPARED LITERATURE FOR LOAD CONSUMPTION SIDE.

Reference paper	Model Name	RMSE	nRMSE	MAPE	MPE
Ref 11	WNN+ANN (without clustering)		0.0232	2.1771 %	
Ref 11	WNN+ ANN (with clustering)		0.01920	1.981%	
Ref 8	Linear Regression	847.62			
Ref 8	LSTM	341.40			
Ref 12	ARIMA			5.16%	
Ref 12	ARIMA-SVM			4.15%	
Ref 11	RNN-LSTM (for 200 Epochs)			5.36%	-1.26
Ref 11	RNN-LSTM (for 400 Epochs)			5.27%	-1.17
Ref 11	DNN-W3(for 200 Epochs)			2.64%	1.61
Ref 11	DNN – SAS (for 400 Epochs)			1.84%	0.53
Ref 16	LSTM			6.54%	
Our Approach	**LSTM**	**150.26**		**0.4821**	
Our Approach	SARIMAX	573.57		3.15%	
Our Approach	Fbprophet	1852.5567		0.09191%	

VII. CONCLUSION AND FUTURE WORK

To conclude, in this study we have used LSTM, SARIMAX and Fbprophet to forecast future power demand of Bangladesh. We have compared the result of reviewing

Authorized licensed use limited to: BRAC UNIVERSITY. Downloaded on May 02,2024 at 17:07:55 UTC from IEEE Xplore. Restrictions apply.

three machine learning algorithms and found LSTM shows better precision and less error compared to other techniques. The RMSE values using the LSTM technique is around 150.26 and MAPE is 0.4821% which is much better compared to the available literature. The findings of this study will be helpful for the government in energy management and policymaking and also for the consumer for better energy planning.

ACKNOWLEDGMENT

The authors would like to thank BRAC University, Bangladesh for the research support.

AUTHOR'S CONTRIBUTION

Abu S. M. Mohsin[4] developed the project, verified the simulation, and prepared the manuscript, and Avijit Paul Piyal[1], Siam Ahmed[2], Khan Fahad Rahman[3] contributed to the simulation and manuscript preparation.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author* (asm.mohsin@bracu.ac.bd) upon reasonable request.

REFERENCES

[1] Anik Nath, Sujoy Barua, Nur Mohammad," Electric Power Generation-Mix for Bangladesh and Its Future", 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE) 7-9 February 2019.

[2] Hasan-Al-Shaikh, Md. Asifur Rahman, Ahmed Zubair," Electric Load Forecasting with Hourly Precision Using Long Short-Term Memory Networks". 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February 2019.

[3] Md. Kamrul Hasan, Nur Mohammad, "An Outlook over Electrical Energy Generation and Mixing Policies of Bangladesh to Achieve Sustainable Energy Targets -Vision 2041", DOI: 10.1109/ECACE.2019.8679446.

Availablelink:https://www.researchgate.net/publication/332703318.

[4] M. HUSSAIN, "BANGLADESH ENERGY RESOURCES AND RENEWABLE ENERGY PROSPECTS".

[5] .https://www.bpdb.gov.bd/

[6] Prof. (retd) Muhtasham Hussain, "Solar and Wind Energy Resource Assessment (SWERA) – Bangladesh" Available link : https://www.researchgate.net/publication/282665355

[7] Hasan-Al-Shaikh, Md. Asifur Rahman, Ahmed Zubair," Electric Load Forecasting with Hourly Precision Using Long Short-Term Memory Networks". 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February 2019.

[8] Salah Bouktif , Ali Fiaz , Ali Ouni and Mohamed Adel Serhani," Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches". Published: 22 June 2018

[9] Chao-Ming Huang, Chi-Jen Huang, and Ming-Li Wang ‘’ A Particle Swarm Optimization to Identifying the ARMAX Model for Short-Term Load Forecasting’’ IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 20, NO. 2, MAY 2005.

[10] Faisal Mohammad, Young-Chon Kim "Energy load forecasting model based on deep neural networks for smart grids". Received: 26 March 2019 / Revised: 15 July 2019. Available link: https://doi.org/10.1007/s13198-019-00884-9.

[11] Hamed H.H. Aly „" A proposed intelligent short-term load forecasting hybrid models of ANN, WNN, and KF based on clustering techniques for smart grid"" Available link: https://doi.org/10.1016/j.epsr.2019.106191.

[12] Stefan Hosein and Patrick Hosein" Load Forecasting using Deep Neural Networks". 978-1-5386-2890-4/17/$31.00 ©2017 IEEE.

[13] Jian Zheng, Cencen Xu, Ziang Zhang, and Xiaohua Li, "Electric Load Forecasting in Smart Grid Using Long-Short-Term-Memory based Recurrent Neural Network,". Available link: 10.1109/CISS.2017.7926112. ©2017 IEEE.

[14] Kasun Amarasinghe, Daniel L. Marino and Milos Manic, "Deep Neural Networks for Energy Load Forecasting" Available link: DOI-10.1109/ISIE.2017.8001465. ©2017 IEEE.

[15] Md. Rashidul Islam, Abdullah Al Mamun, Md. Sohel, Md. Lokman Hossain, and Md. Mofij Uddin$ "LSTM-Based Electrical Load Forecasting for Chattogram City of Bangladesh". Available link: https://www.researchgate.net/publication/340279719. DOI: 10.1109/ESCI48226.2020.9167536. (2020)

[16] Rahul Kumar Agrawal and Frankle Muchahary, "Long Term Load Forecasting with Hourly Predictions based on Long-Short-Term-Memory Networks". Available link: https://www.researchgate.net/publication/323712454. DOI: 10.1109/TPEC.2018.8312088. (Conference paper: February, 2018).

[17] J.Brownlee, "A Gentle Introduction to SARIMA for Time Series Forecasting in Python" machine learning mastery.com,Aug 17, 2018.[Online]

Authorized licensed use limited to: BRAC UNIVERSITY. Downloaded on May 02,2024 at 17:07:55 UTC from IEEE Xplore. Restrictions apply.

View publication stats