A Platform Independent Web-Application for Short-Term Electric Power Load Forecasting on a 33/11 kV Substation Using Regression Model

Venkataramana VEERAMSETTY [1], *Gudelli Sushma VAISHNAVI* [2], *Modem Sai Pavan KUMAR* [2], *Prabhu KIRAN* [2], *Nagula SUMANTH* [2], *Potharaboina PRASANNA* [2], *Surender Reddy SALKUTI* [3]

[1]Center for AI and Deep Learning, SR University, 506371 Warangal, India
[2]Department of Electrical and Electronics Engineering, School of Engineering, SR University, Ananthasagar, 506371 Hasanparthy, Telangana, India
[3]Department of Railroad and Electrical Engineering, Woosong University, 171 Dongdaejeon, Dong-gu, 34606 Daejeon, Republic of Korea

dr.vvr.research@gmail.com, gudellisushma81@gmail.com, saipavankumar104@gmail.com, prabhukiran426@gmail.com, sumanthnagula168@gmail.com, prasannapotharaboina44543@gmail.com, surender@wsu.ac.kr

DOI: 10.15598/aeee.v20i4.4561

Article history: Received May 22, 2022; Revised Jun 23, 2022; Accepted Oct 18, 2022; Published Dec 31, 2022.

Abstract. *Short-term electric power load forecasting is a critical and essential task for utilities of the electric power industry for proper energy trading and that enable the independent system operator to operate the network without any technical and economical issues. In this paper, machine learning model such as linear regression model is used to forecast the active power load one hour and one day ahead. Real time active power load data to train and test the machine learning model is collected from a 33/11 kV substation located in Telangana State, India. Based on the simulation results, it is observed that linear regression model can forecast the load with less mean absolute error i.e. 0.042 with training data and 0.045 with testing data in comparison with support vector regressor model for an hour ahead operation. Whereas in the case of the day ahead operation, linear regression model can forecast the load with less mean absolute error i.e. 0.055 with training data and 0.057 with testing data in comparison with support vector regressor model. A platform independent web application is developed to help the operators of the 33/11 kV substation which is located in Godishala, Telangana State, India.*

Keywords

Day ahead forecasting, hourly ahead forecasting, linear regression model, load forecasting, web application.

1. Introduction

Electric power distribution substations take the power from one or more transmission or sub-transmission lines and deliver this power to residential, commercial and industrial customers through multiple feeders. Distribution transformers will step down the primary voltage to the voltage level that customers can use [1]. Electrical power load forecasting is classified as very short-term, short-term, medium-term and long-term based on the length of the prediction horizon [2], [3] and [4]. Short-term load forecasting at the distribution level estimate the active power load on substation in a time horizon range from 30 minutes to 1 week [5]. Accurate load forecasting is essential at the distribution level for proper planning and operation of the electric power distribution system. Load forecasting on the distribution system sends out an alarm in advance to the operator about the overloading of the feeders and substations. Load forecasting helps the distribution substation operator to schedule and dispatch the storage batteries to shave peak load in a smart grid environment [6].

Forecasting distribution-level load is far more challenging than forecasting system-level load, such as Telangana state electric power demand, due to the intricate load characteristics, huge number of nodes, and probable switching actions in distribution systems. Load estimates over a vast region are highly accurate because aggregated load is steady and consistent. The distribution-level load, on the other hand, may be dominated by a few major clients, such as industrial businesses or schools, and the load pattern may not be as regular as that of a vast region. Furthermore, due to re-configurations caused by switching activities, the load may be temporarily moved from one feeder to another, causing significant changes in distribution-level load profiles and affecting the trend in a given time. Short term electricity load forecast in Bulgaria is studied in [7]. In this paper, neural networks, multilayer perceptron network models and radial basis function network models are used to forecast the load and the performance of these models observed in terms of mean absolute percentage error. Three days ahead energy generation and consumption prediction in Estonia are studied [8]. In this paper, recurrent neural networks are used to forecast energy consumption and generation.

Many researchers are working on short-term load forecasting on distribution systems. ANN based methodology is developed in [9] to forecast the load on the 33/11 kV substation near Kakatiya University in Warangal, Telangana State. Short-term load forecasting by considering previous three hours load i.e., L(T-1), L(T-2), L(T-3) and load at same time but previous four days i.e., L(T-24), L(T-48), L(T-72) and L(T-96) as input features to forecast the load at time 'T' i.e., L(T) on an electric power distribution system using various regression models is proposed in [10]. Short-term load forecasting on a electric power distribution system at a particular time 'T' i.e. L(T) by considering previous three hours load i.e. L(T-1), L(T-2), L(T-3), load at same time but previous three days i.e L(T-24), L(T-48), L(T-72) and load at same time but previous three weeks i.e L(T-168), L(T-336), L(T-504) as input features using factor analysis and long-short term memory is proposed in [11], using random forest and the gated recurrent unit is proposed in [12] and using principal component analysis and recurrent neural network is proposed in [13]. Electric power load forecasting distribution level using correlation concept and ANN is proposed in [14] by considering previous two hours load i.e. L(T-1) and L(T-2), load at same time but previous three days i.e. L(T-24), L(T-48) and L(T-72) as input features.

Electric power load forecasting on a medium voltage level based on regression models and ANN is proposed in [15] using time series DSO telemetry data and weather records from the Portuguese Institute of Sea and Atmosphere and applied to the urban area of Evora, one of Portugal's first Smart Cities. A new top-down algorithm based on a similar day type method to compute an accurate short term distribution loads forecast using only SCADA Data from transmission grid substations is proposed in [16]. This study is evaluated on the RBTS test system with real power consumption data to demonstrate its accuracy. Convolutional neural network based load forecasting methodology is proposed in [17].

Electric demand forecasting with Jellyfish Search Extreme Learning Machine, Harris Hawk Extreme Learning Machine and Flower Pollination Extreme Learning Machine is discussed in [18]. Electric power load forecasting using the gated recurrent units with multi-source data is discussed in [19]. Short-term load forecasting using Niche Immunity Lion Algorithm and Convolutional Neural Network is studied in [20]. Electricity demand forecasting using dynamic adaptive entropy-based weighting is discussed in [21]. Demand side management technique by identifying and mitigating the peak load of a building is studied in [22]. Electric power demand forecasting using vector auto-regressive state space model is discussed in [23].

Electric power load forecasting using the random forest model is discussed in [24]. In this study, authors considered wind speed, wind direction, humidity, temperature, air pressure and irradiance as input features. Electric power load forecasting using group method of data handling and Support Vector Regression is discussed in [25]. Electric power load prediction at the building and district levels for day-ahead energy management using Genetic Algorithm (GA) and Artificial Neural Network (ANN) power predictions is discussed in [26]. Short-term electric power load forecasting using feature engineering, Bayesian Optimization algorithms with a Bayesian Neural Network is discussed in [27]. Active power load forecasting using the Sparrow Search Algorithm (ISSA), Cauchy mutation and Opposition-Based Searning (OBL) and the Long- and Short-Term-Memory (LSTM) network is studied in [28].A new hybrid model is proposed in [29] based on CNN, LSTM, CNN_LSTM, and MLP for electric power load forecasting. All these methodologies provide valuable contribution towards load forecasting problem but these studies did not include weather impact, season and day status in load forecasting. All these methodologies provide valuable contribution towards load forecasting problem but these studies did not include weather impact, season and day status in load forecasting.

Main contributions of this paper are as follows:

- New active power load dataset is developed to work on load forecasting problem by collecting the data from 33/11 kV distribution substation in Godishala (Village), Telangana State, India and available in `https://data.mendeley.com/datasets/tj54nv46hj/1`.
- Machine learning model i.e. linear regression model is used to forecast the load on a 33/11 kV distribution substation in Godishala.
- Active power load on the 33/11 kV substation is forecast one hour before based on input features L(T-1), L(T-2), L(T-24), L(T-48), Day, Season, temperature and Humidity.
- Active power load on the 33/11 kV substation is forecast one day before based on input features L(T-24), L(T-48), Day, Season, temperature and Humidity.
- A Web application is developed based on linear regression model to forecast the load on the 33/11 kV distribution substation in Godishala. This application hosted in Heroku cloud [30].
- The developed web application is a platform independent, can be used in any edge computing device like personal computer or mobile, and also can be used in any operating system.

The remaining part of this paper is organized as: Sec. 2. describes dataset and machine learning model, Sec. 3. presents results discussion and Sec. 4. demonstrates the conclusions of the paper.

2. Methodology

This section presents machine learning model used for electric power load forecasting on the 33/11 kV distribution substation in Godishala. This substation has four feeders, First Feeder (F1) supplies load Godishala town, second feeder supplied load to Bommakal, and third feeder supplies load to Godishala rural, and Fourth Feeder (F4) supplies load to Raikal.

2.1. Active Power Load Data Analysis

To train and test the machine learning model, active power load data is required. Hourly data consists of Voltage (V), Current (I) and power factor ($\cos(\phi)$) on the 33/11 kV distribution substation in Godishala is collected over the duration from 01.01.2021 to 31.12.2021. Based on this data hourly active power load is calculated using Eq. (1) and the sample load data is presented Tab. 1. For the data preparation weekend is labelled as '1' and weekday is labelled as '0'. Weather information like temperature and humidity at each hour of the day in Kazipeta region which is near to Godishala substation location is collected from `https://www.wunderground.com/history/daily/in/kazipet/VOWA/date/2021-1-2`. The complete dataset that is used to train and test the linear regression model is available in [31].

$$P = \sqrt(3)VI\cos(\phi). \tag{1}$$

Tab. 1: Sample load data for first 5 hours at 01.01.2021 on 33/11 kV substation in Godishala.

Time	Voltage (kV)	Current (A)	cos (ϕ)	Power (kW)
01-00	11.6	102	0.96	1967
02-00	11.6	102	0.96	1967
03-00	11.6	102	0.96	1967
04-00	11.3	130	0.96	2443
05-00	11.2	148	0.96	2756

2.2. Machine Learning Models

In this paper, three Machine Learning (ML) models called Linear Regression (LR), Xgboost Regressor (XGBR) [32] and Support Vector Regressor (SVR) [33] are used to forecast the load on the 33/11 kV substation for one hour before and one day before. The problem in discussion is a regression problem. Models need to predict the load on substation based on input features like L(T-1), L(T-2), L(T-24), L(T-48), DAY, SEASON, Temperature and Humidity in the case of hourly ahead forecasting, based on input features like L(T-24), L(T-48), DAY, SEASON, Temperature and Humidity in the case of a day ahead forecasting. The performance of each machine learning model for electric power load forecasting on the 33/11 kV substation is observed in terms of MSE as shown in Eq. (6). MSE is a numerical characteristic used for assessing models [34].

1) Linear Regression Model

Linear regression model represents an output variable as a linear function of input variables [35]. In this paper, linear regression model is express the predicting output variable i.e. L(T) as a linear function of input features i.e L(T-1), L(T-2), L(T-24), L(T-48), DAY (D), SEASON (S), Temperature (Temp) and Humidity (H) as shown in Eq. (2) for an Hour

Ahead Load Forecasting (HALF), and L(T) is predicted as a linear function of input features i.e L(T-24), L(T-48), DAY (D), SEASON (S), Temperature (Temp) and Humidity (H) as shown in Eq. (3) for Day Ahead Load Forecasting (DALF). Stochastic Gradient Descent (SGD) algorithm is used to train the linear regression model. During training process, SGD algorithm is updates linear regression model parameters i.e. m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8 and c in such way that the Half Mean Square Error (HMSE) shown in Eq. (4) is minimized. Architecture of linear regression model is presented in Fig. 1 for hour ahead prediction, whereas architecture of linear regression model for day ahead forecasting is presented in Fig. 2 . Step by step procedure to construct regression model with a sample data is explained in Alg. 1 for an hour ahead forecasting. Similarly, for a day ahead forecasting is presented in Alg. 2. Performance of the regression model on short-term load forecasting problem for Godishala substation is measure in terms of error metrics like MSE [36], RMSE [37], [38] and [39] and MAE [40].

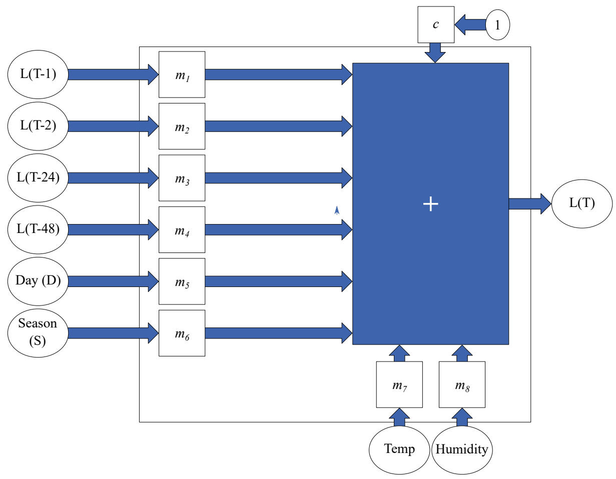


Fig. 1: Linear regression model for hour ahead forecasting.

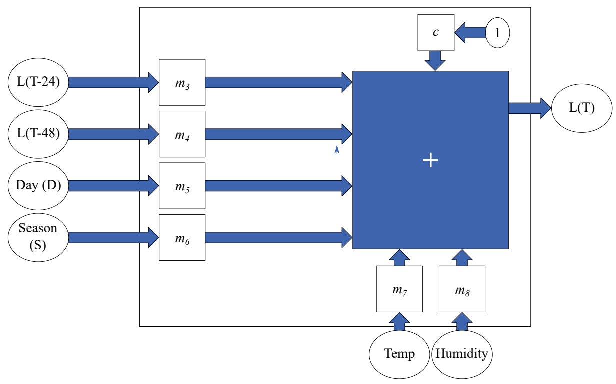


Fig. 2: Linear regression model for day ahead forecasting.

$$\begin{aligned} L(T) = m_1 L(T-1) + m_2 L(T-2) + \\ + m_3 L(T-24) + m_4 L(T-48) + m_5 D + \\ + m_6 S + m_7 Temp + m_8 H + c, \end{aligned} \tag{2}$$

$$\begin{aligned} L(T) = m_3 L(T-24) + m_4 L(T-48) + m_5 D + \\ + m_6 S + m_7 Temp + m_8 H + c, \end{aligned} \tag{3}$$

$$HMSE = \frac{1}{2} \sum_{1}^{n_s} (L^a(T) - L(T))^2. \tag{4}$$

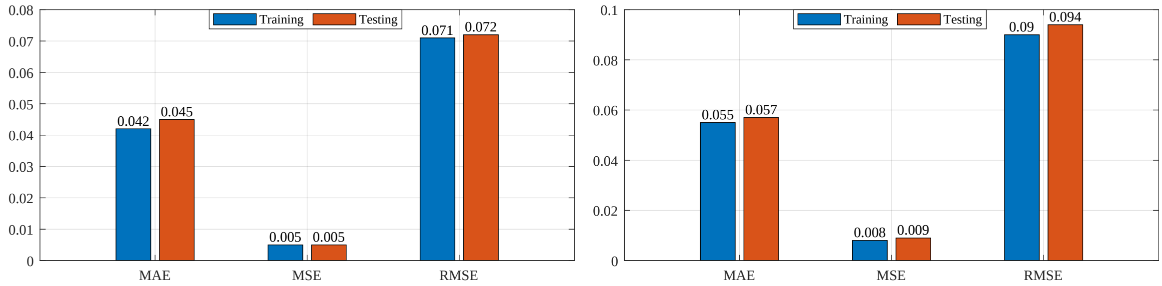
Algorithm 1 Linear Regression Model Training Process using SGD.

1: Read data [L(T-1), L(T-2), L(T-24), L(T-48), D, S, Temp, H] and initialize model parameters [m_1, m_2, Eq. (7)], bias [c], epochs and n_s (number of samples in data)
2: **for** $iteration = 1, 2, \ldots, epochs$ **do**
3: **for** $sample = 1, 2, \ldots, n_s$ **do**
4: Predict the load $L(T)$ using Eq. (2) for HALF and using Eq. (3) for DALF.
5: Update the model parameters using Eq. (5).

$$\begin{aligned} m_1 &= m_1 - \eta \cdot L(T-1) \cdot (L(T) - L^a(T)), \\ m_2 &= m_2 - \eta \cdot L(T-2) \cdot (L(T) - L^a(T)), \\ m_3 &= m_3 - \eta \cdot L(T-24) \cdot (L(T) - L^a(T)), \\ m_4 &= m_4 - \eta \cdot L(T-48) \cdot (L(T) - L^a(T)), \\ m_5 &= m_5 - \eta \cdot D \cdot (L(T) - L^a(T)), \\ m_6 &= m_6 - \eta \cdot S \cdot (L(T) - L^a(T)), \\ m_7 &= m_7 - \eta \cdot Temp \cdot (L(T) - L^a(T)), \\ m_8 &= m_8 - \eta \cdot H \cdot (L(T) - L^a(T)), \\ c &= c - \eta \cdot (L(T) - L^a(T)). \end{aligned} \tag{5}$$

6: **end for**
7: **end for**
8: Read final model parameters [m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8], bias [c]. Calculate mean square error (MSE) of the model based on training and testing data based on Eq. (6).

$$\begin{aligned} Training\ MSE = \frac{1}{n_s} \sum_{1}^{n_s} (L^a(T) - L(T))^2, \\ Testing\ MSE = \frac{1}{n_t} \sum_{1}^{n_t} (L^a(T) - L(T))^2. \end{aligned} \tag{6}$$



(a) Hour ahead load forecasting.

(b) Day ahead load forecasting.

Fig. 3: Training and testing errors of linear regression model.

Tab. 2: Linear regression model parameters for HALF and DALF.

Type of forecasting	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	c
HALF	0.58	-0.098	-0.32	0.15	-0.0002	-0.016	-0.012	-0.069	0.07
DALF	NA	NA	0.61	0.28	-0.0001	-0.02	-0.04	-0.09	0.12

Algorithm 2 Linear Regression Model Training Process using SGD for day ahead forecasting.

Read data [L(T-24), L(T-48), D, S, Temp, H] and initialize model parameters [m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8], bias [c], epochs and n_s (number of samples in data)
for $iteration = 1, 2, \ldots, epochs$ **do**
  for $sample = 1, 2, \ldots, n_s$ **do**
    Predict the load $L(T)$ using Eq. (3) for DALF.
    Update the model parameters using Eq. (5).

$$\begin{aligned} m_3 &= m_3 - \eta \cdot L(T-24) \cdot (L(T) - L^a(T)), \\ m_4 &= m_4 - \eta \cdot L(T-48) \cdot (L(T) - L^a(T)), \\ m_5 &= m_5 - \eta \cdot D \cdot (L(T) - L^a(T)), \\ m_6 &= m_6 - \eta \cdot S \cdot (L(T) - L^a(T)), \\ m_7 &= m_7 - \eta \cdot Temp \cdot (L(T) - L^a(T)), \\ m_8 &= m_8 - \eta \cdot H \cdot (L(T) - L^a(T)), \\ c &= c - \eta \cdot (L(T) - L^a(T)). \end{aligned} \tag{7}$$

  end for
end for
Read final model parameters [m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8], bias [c]. Calculate Mean Square Error (MSE) of the model based on training and testing data based on Eq. (6).

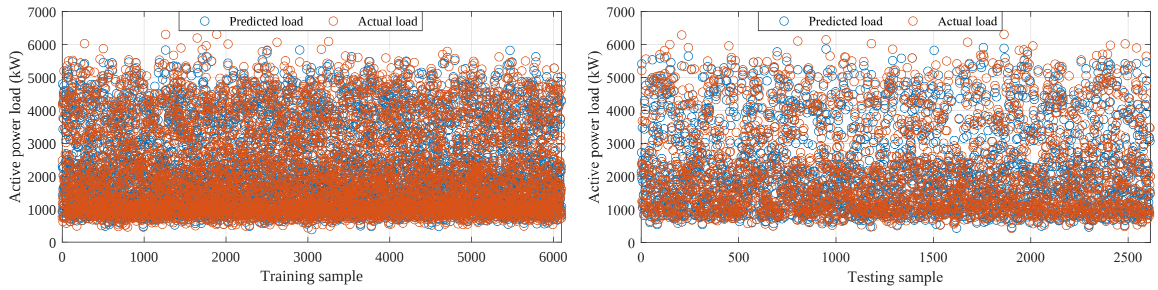
3. Result Analysis

All the machine learning models developed based on data available in [31] using Google Colab. This section presents data analysis, training and testing performance of machine learning models and the web application developed to predict the load. Out of 8712 samples 70 % of samples i.e., 6098 samples are used for training and remaining 30 % of samples i.e., 2614 samples are used for testing. Date processing techniques for observing the data distribution, outliers and data normalization have been used before to train and test the regression model. Stochastic gradient descent optimizer has been used train the regression models.

3.1. Linear Regression Model

The performance of the linear regression model that is developed to forecast the load L(T) based on features L(T-1), L(T-2), L(T-24), L(T-48), day and season status, temperature and humidity is observed based on training and testing accuracy for an HALF. Similarly, for a DALF based on features L(T-24), L(T-48), day and season status, temperature and humidity. The training and testing error metrics of this model is presented in Fig. 3 for both HALF and DALF. From the Fig. 3, it is observed that the model is well fitted without any under or over fitting problems. Linear regression model is predicting load with less error for HALF in comparison with DALF as in later case load is predicting one day earlier i.e 24 hours time horizon. The model parameters of linear regression model for HALF and DALF are presented in Tab. 2.

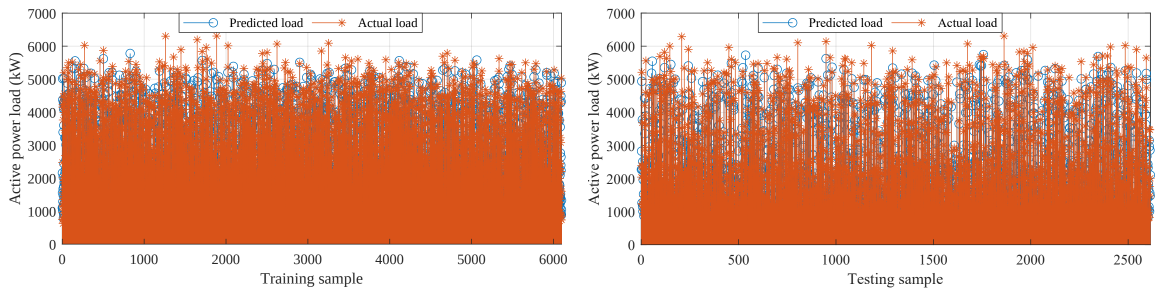
The distribution of predicted load with linear regression model having training and testing MSE 0.005 is compared with actual load samples for training and testing data for HALF is presented in Fig. 4. Similarly, the distribution of predicted load with linear regression model having training MSE 0.008 and testing MSE 0.009 is compared with actual load samples for training and testing data for DALF is presented in Fig. 5. From the Fig. 4 and Fig. 5, it is observed



(a) Hour ahead load forecasting: Predicted Load Vs. Actual Load for training data.

(b) Hour ahead load forecasting: Predicted Load Vs. Actual Load for testing data.

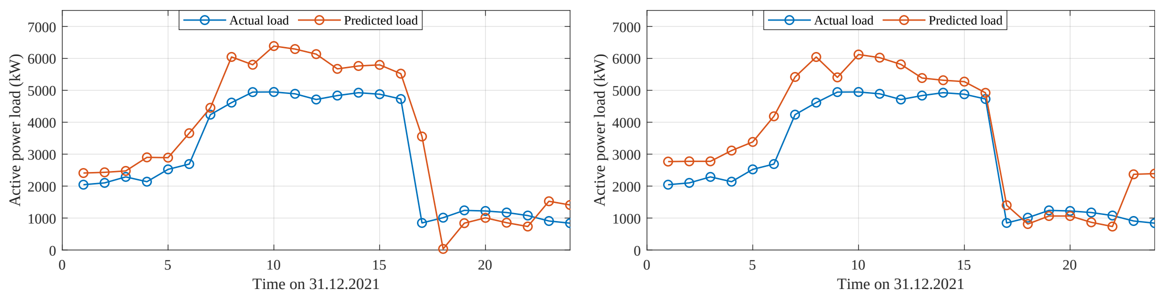
Fig. 4: Distribution of predicted and actual load samples with linear regression model for HALF.



(a) Day ahead load forecasting: Predicted Load Vs. Actual Load for training data.

(b) Day ahead load forecasting: Predicted Load Vs. Actual Load for testing data.

Fig. 5: Distribution of predicted and actual load samples with linear regression model for DALF.



(a) HALF: Predicted Load Vs. Actual Load.

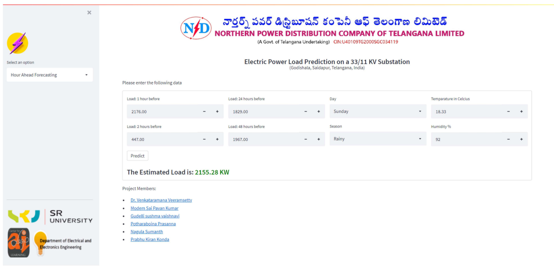
(b) DALF: Predicted Load Vs. Actual Load.

Fig. 6: Distribution of predicted and actual load samples with linear regression model on 31.12.2021.

that most of predicted and actual load samples are overlapping each other.

The predicted load using a linear regression model having training and testing MSE 0.005 is compared with an actual load on 31.12.2021 and presented in Fig. 6. From Fig. 6, it is observed the predicted load using a linear regression model is almost following the actual load curve during night time but has more difference during day time. And also, observed that predicting load is slightly near to actual load curve in case of HALF in comparison with DALF. As in later case the load is predicting one day earlier i.e 24 hours time horizon. At 5PM, the error is approximately 300 % as a sudden change in load due to closing of schools and other working places. Similarly, the error is approximately 100 % at 23rd and 24th hours because of abnormal load patterns in these hours due to new year celebrations.

A web application was developed using an optimal linear regression model to predict load one hour before and also one day before for real time usage as a prototype and shown in Fig. 7. This web application is accessible through the link `https://loadforecasting-godishala-lrm.herokuapp.com/` or through the QR code shown in Fig. 7.



(a) Hour ahead load forecasting.

(b) Day ahead load forecasting.

(c) About application page.

(d) QR code to access the web application.

Fig. 7: Web application to predict active power load on a 33/11 kV substation in Godishala, Telangana State- India, developed using linear regression model.

3.2. Comparative Analysis

The proposed linear regression model to forecast active power load on a 33/11 kV substation one hour before is validated by comparing with XG Boost Regressor and Support Vector Regressor in terms of mean absolute error and presented in Fig. 8. From Fig. 8 it is observed that the proposed model forecasts the load one hour before with MAE value 0.055 with training data and 0.057 with testing data. Whereas SVR, forecasts the load with more MAE value i.e. 0.06 with training and testing data. Whereas XGBR forecasts the load with MAE value 0.055 with training data and 0.057 with testing data same as the LR model but XGBR model complexity is more in comparison with the LR model.

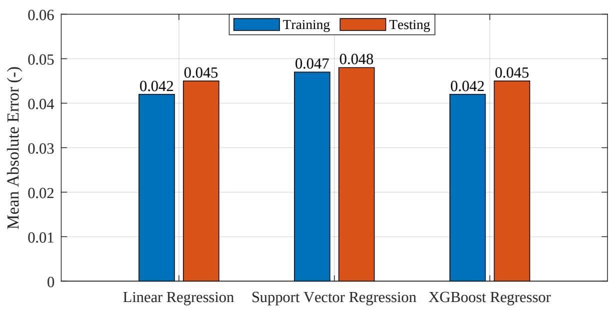


Fig. 8: Comparative analysis - HALF.

The proposed linear regression model to forecast active power load on a 33/11 kV substation one day before is validated by comparing with XG Boost Regressor and Support Vector Regressor in terms of mean absolute error and presented in Fig. 9. From Fig. 9 it is observed that the proposed model forecasts the load one day before with MAE value 0.055 with training data and 0.057 with testing data. Whereas SVR forecasts the load with more MAE value i.e. 0.06 with training and testing data. Whereas XGBR forecasts the load with MAE value 0.055 with training data and 0.057 with testing data same as the LR model but XGBR model complexity is more in comparison with the LR model.

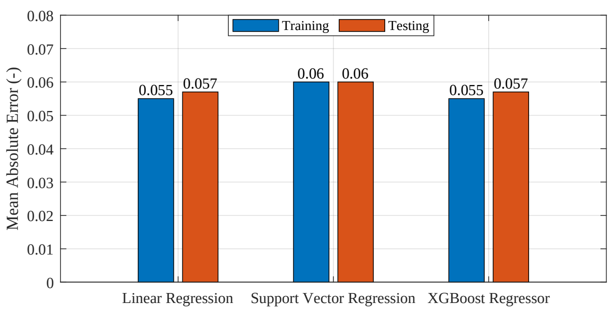


Fig. 9: Comparative analysis - DALF.

4. Conclusions

Electric power load forecasting one hour ahead and one day ahead is required for the utilities to place a bid successfully in hourly ahead energy markets and day ahead energy markets. In this paper, active power on a 33/11 kV substation is predicted one hour before based on load available at last two hours and last two days at the time of prediction, and day status, season status, temperature and humidity. Similarly, the load is predicted one day before based on the load available at last two days at the time of prediction, and day status, season status, temperature and humidity.

In this work, three machine learning models i.e. linear regression model, XG Boost Regressor and Support Vector regressor are developed to predict the active power load on a 33/11 kV substation located at Godishala village in Telangana State, India. Based on the results, it is observed that the linear model predicted the load one hour and one day before with less mean square error in comparison with SVR, and almost with same error in comparison with XGBR but with less model complexity.

This work can be further extended by considering the deep neural networks, sequence models and conventional time series data prediction models. In this paper, temperature and humidity data at the time of prediction are considered from open source website. However, current work can be extended further by integrating temperature and humidity forecasting models with current load forecasting models.

Author Contributions

V.V., G.S.V., M.S.P, N.S. and P.P. developed the theoretical formalism, performed the analytic calculations, data collection and performed numerical simulations. V.V. and S.R.S. contributed to the analysis of the results and the proofreading of this manuscript. V.V. and P.K. developed the web application. All the authors provided critical feedback, helped to shape the research and conduct the analysis and thus contributed to the final version of the manuscript.

References

[1] SUN, X., P. B. LUH, K. W. CHEUNG, W. GUAN, L. D. MICHEL, S. S. VENKATA and M. T. MILLER. An Efficient Approach to Short-Term Load Forecasting at the Distribution Level. *IEEE Transactions on Power Systems*. 2016, vol. 31, iss. 4, pp. 2526–2537. ISSN 1558-0679. DOI: 10.1109/TPWRS.2015.2489679.

[2] HENSELMEYER, S. and M. GRZEGORZEK. Short-Term Load Forecasting Using an Attended Sequential Encoder-Stacked Decoder Model with Online Training. *Applied Sciences*. 2021, vol. 11, iss. 11, pp. 1–15. ISSN 2076-3417. DOI: 10.3390/app11114927.

[3] SHOHAN, M. J. A., M. O. FARUQUE and S. Y. FOO. Forecasting of Electric Load Using a Hybrid LSTM-Neural Prophet Model. *Energies*. 2022, vol. 15, iss. 6, pp. 1–18. ISSN 1996-1073. DOI: 10.3390/en15062158.

[4] GRZESZCZYK, T. A. and M. K. GRZESZCZYK. Justifying Short-Term Load Forecasts Obtained with the Use of Neural Models. *Energies*. 2022, vol. 15, iss. 5, pp. 1–20. ISSN 1996-1073. DOI: 10.3390/en15051852.

[5] GRIGSBY, L. L. *Electric Power Generation, Transmission, and Distribution*. 3rd ed. Boca Raton: CRC press, 2018. ISBN 978-1-315-22242-4.

[6] WILLIS, H. L. *Spatial Electric Load Forecasting*. 2nd ed. Boca Raton: CRC Press, 2002. ISBN 978-0-429-22274-0.

[7] ASENOVA, I. and D. GEORGIEV. Short-Term Load Forecast in Electric Energy System in Bulgaria. *Advances in Electrical and Electronic Engineering*. 2010, vol. 8, iss. 4, pp. 102–106. ISSN 1804-3119.

[8] SHABBIR, N., L. KUTT, M. JAWAD, M. N. IQBAL and P. S. GHAHFAROKI. Forecasting of Energy Consumption and Production Using Recurrent Neural Networks. *Advances in Electrical and Electronic Engineering*. 2020, vol. 18, iss. 3, pp. 190–197. ISSN 1804-3119. DOI: 10.15598/aeee.v18i3.3597.

[9] VEERAMSETTY, V. and R. DESHMUKH. Electric power load forecasting on a 33/11 kV substation using artificial neural networks. *SN Applied Sciences*. 2020, vol. 2, iss. 5, pp. 1–10. ISSN 2523-3971. DOI: 10.1007/s42452-020-2601-y.

[10] VEERAMSETTY, V., A. MOHNOT, G. SINGAL and S. R. SALKUTI. Short Term Active Power Load Prediction on A 33/11 kV Substation Using Regression Models. *Energies*. 2021, vol. 14, iss. 11, pp. 1–21. ISSN 1996-1073. DOI: 10.3390/en14112981.

[11] VEERAMSETTY, V., D. R. CHANDRA and S. R. SALKUTI. Short-term electric power load forecasting using factor analysis and long short-term memory for smart cities. *International Journal of Circuit Theory and Applications*. 2021, vol. 49, iss. 6, pp. 1678–1703. ISSN 1097-007X. DOI: 10.1002/cta.2928.

[12] VEERAMSETTY, V., K. R. REDDY, M. SANTHOSH, A. MOHNOT and G. SINGAL. Short-term electric power load forecasting using random forest and gated recurrent unit. *Electrical*

Engineering. 2022, vol. 104, iss. 1, pp. 307–329. ISSN 1432-0487. DOI: 10.1007/s00202-021-01376-5.

[13] VEERAMSETTY, V., D. R. CHANDRA, F. GRIMACCIA and M. MUSSETTA. Short Term Electric Power Load Forecasting Using Principal Component Analysis and Recurrent Neural Networks. *Forecasting*. 2022, vol. 4, iss. 1, pp. 149–164. ISSN 2571-9394. DOI: 10.3390/forecast4010008.

[14] SALKUTI, S. R. and P. RAY. *Next Generation Smart Grids: Modeling, Control and Optimization*. 1st ed. Singapore: Springer Singapore, 2022. ISBN 978-981-16-7794-6.

[15] CAMARINHA-MATOS, L. M., A. J. FALCAO, N. VAFAEI and S. NAJDI. *Technological Innovation for Cyber-Physical Systems*. 1st ed. Cham: Springer Cham, 2016. ISBN 978-3-319-31165-4.

[16] COURAUD, B. and R. ROCHE. A distribution loads forecast methodology based on transmission grid substations SCADA Data. In: *2014 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*. Kuala Lumpur: IEEE, 2014, pp. 35–40. ISBN 978-1-4799-1300-8. DOI: 10.1109/ISGT-Asia.2014.6873760.

[17] ANDRIOPOULOS, N., A. MAGKLARAS, A. BIRBAS, A. PAPALEXOPOULOS, C. VALOUXIS, S. DASKALAKI, M. BIRBAS, E. HOUSOS and G. P. PAPAIOANNOU. Short Term Electric Load Forecasting Based on Data Transformation and Statistical Machine Learning. *Applied Sciences*. 2020, vol. 11, iss. 1, pp. 1–22. ISSN 2076-3417. DOI: 10.3390/app11010158.

[18] BORIRATRIT, S., C. SRITHAPON, P. FUANGFOO and R. CHATTHAWORN. Metaheuristic Extreme Learning Machine for Improving Performance of Electric Energy Demand Forecasting. *Computers*. 2022, vol. 11, iss. 5, pp. 1–34. ISSN 2073-431X. DOI: 10.3390/computers11050066.

[19] WANG, Y., M. LIU, Z. BAO and S. ZHANG. Short-Term Load Forecasting with Multi-Source Data Using Gated Recurrent Unit Neural Networks. *Energies*. 2018, vol. 11, iss. 5, pp. 1–19. ISSN 1996-1073. DOI: 10.3390/en11051138.

[20] LI, Y., Y. HUANG and M. ZHANG. Short-Term Load Forecasting for Electric Vehicle Charging Station Based on Niche Immunity Lion Algorithm and Convolutional Neural Network. *Energies*. 2018, vol. 11, iss. 5, pp. 1–18. ISSN 1996-1073. DOI: 10.3390/en11051253.

[21] HU, Z., J. MA, L. YANG, X. LI and M. PANG. Decomposition-Based Dynamic Adaptive Combination Forecasting for Monthly Electricity Demand. *Sustainability*. 2019, vol. 11, iss. 5, pp. 1–25. ISSN 2071-1050. DOI: 10.3390/su11051272.

[22] AMOASI ACQUAH, M., D. KODAIRA and S. HAN. Real-Time Demand Side Management Algorithm Using Stochastic Optimization. *Energies*. 2018, vol. 11, iss. 5, pp. 1–14. ISSN 1996-1073. DOI: 10.3390/en11051166.

[23] NAGBE, K., J. CUGLIARI and J. JACQUES. Short-Term Electricity Demand Forecasting Using a Functional State Space Model. *Energies*. 2018, vol. 11, iss. 5, pp. 1–24. ISSN 1996-1073. DOI: 10.3390/en11051120.

[24] KIPTOO, M. K., O. B. ADEWUYI, M. E. LOTFY, T. AMARA, K. V. KONNEH and T. SENJYU. Assessing the Techno-Economic Benefits of Flexible Demand Resources Scheduling for Renewable Energy–Based Smart Microgrid Planning. *Future Internet*. 2019, vol. 11, iss. 10, pp. 1–16. ISSN 1999-5903. DOI: 10.3390/fi11100219.

[25] YU, J., J. H. PARK and S. KIM. A New Input Selection Algorithm Using the Group Method of Data Handling and Bootstrap Method for Support Vector Regression Based Hourly Load Forecasting. *Energies*. 2018, vol. 11, iss. 11, pp. 1–20. ISSN 1996-1073. DOI: 10.3390/en11112870.

[26] KAMPELIS, N., E. TSEKERI, D. KOLOKOTSA, K. KALAITZAKIS, D. ISIDORI and C. CRISTALLI. Development of Demand Response Energy Management Optimization at Building and District Levels Using Genetic Algorithm and Artificial Neural Network Modelling Power Predictions. *Energies*. 2018, vol. 11, iss. 11, pp. 1–22. ISSN 1996-1073. DOI: 10.3390/en11113012.

[27] JIN, X.-B., W.-Z. ZHENG, J.-L. KONG, X.-Y. WANG, Y.-T. BAI, T.-L. SU and S. LIN. Deep-Learning Forecasting Method for Electric Power Load via Attention-Based Encoder-Decoder with Bayesian Optimization. *Energies*. 2021, vol. 14, iss. 6, pp. 1–18. ISSN 1996-1073. DOI: 10.3390/en14061596.

[28] HAN, M., J. ZHONG, P. SANG, H. LIAO and A. TAN. A Combined Model Incorporating Improved SSA and LSTM Algorithms for Short-Term Load Forecasting. *Electronics*. 2022, vol. 11, iss. 12, pp. 1–24. ISSN 2079-9292. DOI: 10.3390/electronics11121835.

[29] TALEB, I., G. GUERARD, F. FAUBERTEAU and N. NGUYEN. A Flexible Deep Learning Method for Energy Forecasting. *Energies*. 2022, vol. 15, iss. 11, pp. 1–16. ISSN 1996-1073. DOI: 10.3390/en15113926.

[30] HANJURA, A. *Heroku Cloud Application Development*. 1st ed. Birmingham: Packt Publishing, 2014. ISBN 978-1-78355-097-5.

[31] VEERAMSETTY, V., G. S. VAISHNAVI, M. S. P. KUMAR, N. SUMANTH and P. PRASANNA. Electric power load dataset. In: *Mendeley Data* [online]. 2022. Available at: https://data.mendeley.com/datasets/tj54nv46hj/1.

[32] CHEN, T. and T. HE. xgboost: eXtreme Gradient Boosting. In: *CRAN* [online]. 2017. Available at: https://cran.microsoft.com/snapshot/2017-12-11/web/packages/xgboost/vignettes/xgboost.pdf.

[33] AWAD, M. and R. KHANNA. *Efficient Learning Machines*. 1st ed. Berkeley: Apress Berkeley, 2015. ISBN 978-1-4302-5990-9.

[34] FALAT, L. and D. MARCEK. Financial Time Series Modelling with Hybrid Model Based on Customized RBF Neural Network Combined With Genetic Algorithm. *Advances in Electrical and Electronic Engineering*. 2014, vol. 12, iss. 4, pp. 307–318. ISSN 1804-3119. DOI: 10.15598/aeee.v12i4.1206.

[35] NOWAKOVA, J. and M. POKORNY. On Multidimensional Linear Modelling Including Real Uncertainty. *Advances in Electrical and Electronic Engineering*. 2014, vol. 12, iss. 5, pp. 511–517. ISSN 1804-3119. DOI: 10.15598/aeee.v12i5.1143.

[36] ALDHYANI, T. H. H. and H. ALKAHTANI. A Bidirectional Long Short-Term Memory Model Algorithm for Predicting COVID-19 in Gulf Countries. *Life*. 2021, vol. 11, iss. 11, pp. 1–26. ISSN 2075-1729. DOI: 10.3390/life11111118.

[37] ZHANG, W., P. WU, Y. PENG and D. LIU. Roll Motion Prediction of Unmanned Surface Vehicle Based on Coupled CNN and LSTM. *Future Internet*. 2019, vol. 11, iss. 11, pp. 1–15. ISSN 1999-5903. DOI: 10.3390/fi11110243.

[38] LU, Y., Y. LI, D. XIE, E. WEI, X. BAO, H. CHEN and X. ZHONG. The Application of Improved Random Forest Algorithm on the Prediction of Electric Vehicle Charging Load. *Energies*. 2018, vol. 11, iss. 11, pp. 1–16. ISSN 1996-1073. DOI: 10.3390/en11113207.

[39] MAITAH, M., K. MALEC, Y. GE, Z. GEBELTOVA, L. SMUTKA, V. BLAZEK, L. PANKOVA, K. MAITAH and J. MACH. Assessment and Prediction of Maize Production Considering Climate Change by Extreme Learning Machine in Czechia. *Agronomy*. 2021, vol. 11, iss. 11, pp. 1–14. ISSN 2073-4395. DOI: 10.3390/agronomy11112344.

[40] LOPEZ-ESPINOZA, E. D., J. ZAVALA-HIDALGO, R. MAHMOOD and O. GOMEZ-RAMOS. Assessing the Impact of Land Use and Land Cover Data Representation on Weather Forecast Quality: A Case Study in Central Mexico. *Atmosphere*. 2020, vol. 11, iss. 11, pp. 1–20. ISSN 2073-4433. DOI: 10.3390/atmos11111242.

About Authors

Venkataramana VEERAMSETTY was born in Srikakulam, India, in 1985. He received his B.Tech from JNTU, Hyderabad in 2006, and M.E. from university college of engineering, Osmania University in 2011, and Ph.D. from National Institute of Technology Waranagl in 2018. His research interests include power system deregulation and application of artificial intelligence in electrical engineering.

Gudelli Sushma VAISHNAVI was born in Warangal, India, in 2001. She received her B.Tech from SR University, Warangal in 2022. Her research interests include application of artificial intelligence in electrical engineering.

Modem Sai Pavan KUMAR was born in Karimnagar, India, in 2001. He received her B.Tech from SR University, Warangal in 2022. His research interests include application of artificial intelligence in electrical engineering.

Prabhu KIRAN was born in Warangal, India, in 2002. He is undergraduate student at SR University, Warangal. His research interests include application of artificial intelligence in electrical engineering.

Nagula SUMANTH was born in Warangal, India, in 2001. He received her B.Tech from SR University, Warangal in 2022. His research interests include application of artificial intelligence in electrical engineering.

Potharaboina PRASANNA was born in Warangal, India, in 2001. She received her B.Tech from SR University, Warangal in 2022. Her research interests include application of artificial intelligence in electrical engineering.

Surender Reddy SALKUTI (corresponding author) received the M.Tech and Ph.D. degrees in electrical engineering from NIT Warangal and IIT Delhi, New Delhi, India, in 2009 and 2013 respectively. He was a Postdoctoral Researcher at Howard University, Washington, DC, USA, from 2013 to 2014. He is currently working as an Associate Professor in the Department of Railroad and Electrical Engineering, Woosong University, Daejeon, Republic of Korea. His current research interests include power system restructuring issues, market-clearing including renewable energy sources, demand response, smart grid development with integration of wind and solar photovoltaic energy sources, battery storage and electric vehicles, and artificial intelligence applications in power systems.

Appendix A

In this section, step by step procedure that is used to train the linear regression model is presented. For this purpose, a sample dataset that is build from few samples on original dataset is shown in Tab. 3.

Assume random linear regression model parameters as shown in Tab. 4.

1.1. Linear regression model training

- Iteration:01 and Sample:01 - [0.12, 0.12, 0, 0, 0.6, 0.73, 0.05].
- Predict the load using Eq. (3).

$$\begin{aligned} L^p(T) &= 1 \cdot 0.12 - 1 \cdot 0.12 + 1 \cdot 0 - 1 \cdot 0 + \\ &+1 \cdot 0.6 - 1 \cdot 0.73 + 1 = 0.87. \end{aligned} \tag{8}$$

- Update model parameters m_3, m_4, m_5, m_6, m_7, m_8 and c using Eq. (5).

$$\begin{aligned} m_3 &= 1 - 0.1 \cdot 0.12 \cdot (0.87 - 0.05) = 0.9901, \\ m_4 &= -1 - 0.1 \cdot 0.12 \cdot (0.87 - 0.05) = -1.0098, \\ m_5 &= 1 - 0.1 \cdot 0 \cdot (0.87 - 0.05) = 1, \\ m_6 &= -1 - 0.1 \cdot 0 \cdot (0.87 - 0.05) = 1, \\ m_7 &= 1 - 0.1 \cdot 0.6 \cdot (0.87 - 0.05) = 0.9508, \\ m_8 &= -1 - 0.1 \cdot 0.73 \cdot (0.87 - 0.05) = -1.0598, \\ c &= 1 - 0.1 \cdot (0.87 - 0.05) = 0.918. \end{aligned} \tag{9}$$

- Iteration:01 and Sample:02 - [0.3, 0.35, 0, 0, 0.55, 0.3, 0.33].
- Predict the load using Eq. (3).

$$\begin{aligned} L^p(T) &= 0.9901 \cdot 0.3 - 1.0098 \cdot 0.35 + 1 \cdot 0 + \\ -1 \cdot 0 &+ 0.9508 \cdot 0.55 - 1.0598 \cdot 0.3 + 0.918 = 1.0665. \end{aligned} \tag{10}$$

- Update model parameters m_3, m_4, m_5, m_6, m_7, m_8 and c using Eq. (5).

$$\begin{aligned} m_3 &= 0.9901 - 0.1 \cdot 0.3 \cdot (1.0665 - 0.33) = 0.968, \\ m_4 &= -1.0098 - 0.1 \cdot 0.35 \cdot (1.0665 - 0.33) = -1.0356, \\ m_5 &= 1 - 0.1 \cdot 0 \cdot (1.0665 - 0.33) = 1, \\ m_6 &= -1 - 0.1 \cdot 0 \cdot (1.0665 - 0.33) = -1, \\ m_7 &= 0.9508 - 0.1 \cdot 0.55 \cdot (1.0665 - 0.33) = 0.9102, \\ m_8 &= -1.0598 - 0.1 \cdot 0.3 \cdot (1.0665 - 0.33) = -1.0818, \\ c &= 0.918 - 0.1 \cdot (1.0665 - 0.33) = 0.8443. \end{aligned} \tag{11}$$

- Iteration:02 and Sample:01 - [0.12, 0.12, 0, 0, 0.6, 0.73, 0.05].
- Predict the load using Eq. (3).

$$\begin{aligned} L^p(T) &= 0.968 \cdot 0.12 - 1.03561 \cdot 0.12 + 1 \cdot 0 - 1 \cdot 0 + \\ &+0.91029 \cdot 0.60 - 1.0818 \cdot 0.73 + 0.84435 = 0.5925. \end{aligned} \tag{12}$$

- Update model parameters m_3, m_4, m_5, m_6, m_7, m_8 and c using Eq. (5).

$$\begin{aligned} m_3 &= 0.968 - 0.1 \cdot 0.12 \cdot (0.5925 - 0.05) = 0.9614, \\ m_4 &= -1.0356 - 0.1 \cdot 0.12 \cdot (0.5925 - 0.05) = -1.0421, \\ m_5 &= 1 - 0.1 \cdot 0 \cdot (0.5925 - 0.05) = 1, \\ m_6 &= -1 - 0.1 \cdot 0 \cdot (0.5925 - 0.05) = -1, \\ m_7 &= 0.9102 - 0.1 \cdot 0.60 \cdot (0.5925 - 0.05) = 0.8777, \\ m_8 &= -1.0818 - 0.1 \cdot 0.73 \cdot (0.5925 - 0.05) = -1.1214, \\ c &= 0.8443 - 0.1 \cdot (0.5925 - 0.05) = 0.7901. \end{aligned} \tag{13}$$

- Iteration:02 and Sample:02 - [0.3, 0.35, 0, 0, 0.55, 0.30, 0.33].
- Predict the load using Eq. (3).

$$\begin{aligned} L^p(T) &= 0.9614 \cdot 0.30 - 1.0421 \cdot 0.35 + 1 \cdot 0 - 1 \cdot 0 + \\ &+0.8777 \cdot 0.55 - 1.1214 \cdot 0.3 + 0.7901 = 0.8601. \end{aligned} \tag{14}$$

Tab. 3: Sample normalized data to build linear regression model.

Data	L(T-24)	L(T-48)	Day	Season	Temperature	Humidity	L(T)
Training	0.12	0.12	0	0	0.6	0.73	0.05
	0.3	0.35	0	0	0.55	0.3	0.33
Testing	0.19	0.19	0	1	0.57	0.31	0.19

Tab. 4: Initial random parameters of linear regression model.

m_3	m_4	m_5	m_6	m_7	m_8	c
1	-1	1	-1	1	-1	1

- Update model parameters m_3, m_4, m_5, m_6, m_7, m_8 and c using Eq. (5).

$$\begin{gathered} m_3 = 0.9614 - 0.1 \cdot 0.3 \cdot (0.8601 - 0.33) = 0.9454, \\ m_4 = -1.0421 - 0.1 \cdot 0.35 \cdot (0.8601 - 0.33) = -1.0606, \\ m_5 = 1 - 0.1 \cdot 0 \cdot (0.8601 - 0.33) = 1, \\ m_6 = -1 - 0.1 \cdot 0 \cdot (0.8601 - 0.33) = -1, \\ m_7 = 0.8777 - 0.1 \cdot 0.55 \cdot (0.8601 - 0.33) = 0.8485, \\ m_8 = -1.1214 - 0.1 \cdot 0.3 \cdot (0.8601 - 0.33) = -1.1373, \\ c = 0.7901 - 0.1 \cdot (0.8601 - 0.33) = 0.737. \end{gathered} \tag{15}$$

At end of the two iterations, final model parameters for the linear regression model are presented in Tab. 5.

Tab. 5: linear regression model parameters.

m_3	m_4	m_5	m_6	m_7	m_8	c
0.9454	-1.0606	1	-1	0.8485	-1.1373	0.737

Load $L^p(T)$ is predicted for both training and testing samples using model parameters shown in Tab. 5 and the predicted values are shown in Tab. 6.

Tab. 6: Predicted values using linear regression model.

Data	L(T)	L^p(T)
Training	0.05	0.4019
	0.33	0.7751
Testing	0.19	-0.1538

Calculate mean square error for training and testing data using Eq. (6), also by using actual and predicted load values shown in Tab. 6 as shown below:

$$\begin{gathered} Training\ MSE = \\ \frac{(0.05 - 0.402)^2 + (0.33 - 0.7751)^2}{2} = 0.1609, \end{gathered} \tag{16}$$

$$Testing\ MSE = \frac{(0.19 + 0.1538)^2}{1} = 0.1156. \tag{17}$$