FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 4 | Pages 1461-1465 | e-ISSN: 2347-470X

Short Term Load Prediction based on LSTM Network for Iraqi Thermal Power Plant

Safa Abdulwahid[1,2], and Mahmoud-Reza Haghifam[3]

[1]*Department of Electrical Engineering, South of Tehran Branch, Islamic Azad University, Tehran, Iran*
[2]*College of Engineering, Al Muthanna University, Al Muthanna, Iraq, safa.abdulwahid@mu.edu.iq*
[3]*Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran, haghifam@modares.ac.ir*

***Correspondence:** Safa Abdulwahid; E-mail: *safa.abdulwahid@mu.edu.iq*

ABSTRACT- Electricity generation must satisfy the demand for electric loads in order to optimize the functioning of the power system. Load prediction can assist power companies in safely and efficiently operating their electrical systems. Load prediction is a process employed by power providers to predict the quantity of power or energy always demanded for balancing supply and demand. Short-term load prediction (STLP) with high accuracy is crucial to the seamless operation of the power system and the improvement of economic benefits. An approach for predicting short-term electrical demand utilizing long short-term memory (LSTM) based on actual data collected from Wasit Thermal Power Plant in Iraq is proposed in this paper. MATLAB software is utilized to implement the data used in this work. The assessment metrics employed were mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE) and the coefficient of determination (R-squared) to assess the precision of load prediction. The findings demonstrate that the LSTM model is highly effective at forecasting the random characteristics of an electrical demand.

Keywords: Load prediction; short term; electrical load; LSTM; neural network; evaluation metrics.

ARTICLE INFORMATION

Author(s): Safa Abdulwahid and Mahmoud-Reza Haghifam;
Received: 09/09/2024; **Accepted**: 09/12/2024; **Published**: 30/12/2024;
e-ISSN: 2347-470X;
Paper Id: IJEER 0909-07;
Citation: 10.37391/IJEER.120440
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-12/ijeer-120440.html

crossref member
CROSSREF.ORG
THE CITATION LINKING BACKBONE

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Electric companies have several economic and technological obstacles that they must overcome in order to deliver reliable and cost-efficient energy to their customers. Scheduling, load flow analysis, planning, and electrical power system management are the most important of these issues. Load prediction is a very promising area of study that has gained significant attention in recent years. Load estimation can be stated as the level of precision in the difference between the real and predicted amounts of future load demand. Predicting load demand can optimize the generation unit start-up costs and prevent the need for additional power facilities [1]. Load prediction is an essential and crucial aspect of electric utility organization and administration. The system operator can efficiently arrange the location of spinning reserves using load prediction, which encompasses timeframes ranging from a few minutes to many months [2]. Four different types are used to categorize real power load forecasting: very short-term load prediction (VSTLP) contains load predictions for the next day [3], short-term load prediction (STLP) deals with load estimations for a few hours to a few days out, medium-term load prediction (MTLP) focuses on predictions for a few weeks to a few months out, and long-term load prediction (LTLP) focuses on load estimations from one year to a few years out [4].

To date, numerous STLF approaches have been developed, and these techniques may often be categorized into two categories. The first category comprises conventional methods like time series analysis [5], auto regressive integrated moving average (ARIMA) model [6], multiple linear regression [7], and so on. The second group includes computational intelligence (CI) methods like expert systems [8], support vector machine (SVM) [9], fuzzy systems [10], artificial neural network (ANN) [11], recurrent neural network (RNN) [12], convolutional neural network (CNN) [13], and a combination of these methods.

This research is focused on short-term load prediction utilizing a long short-term memory (LSTM) network by dividing the input data into two models: all-data and weekday models, and comparing the outcomes of these models based on four assessment metrics.

2. LSTM NETWORK

The LSTM network is a robust configuration of recurrent neural networks (RNN) that is highly effective at analyzing time series data and making predictions [14]. It is an advanced framework that accurately evaluates and understands the relationships between distant elements in a sequence of data. Several applications, including natural language processing and audio identification, frequently use it [15]. The architecture of the single cell of the LSTM network is illustrated in *figure 1*.

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 4 | Pages 1461-1465 | e-ISSN: 2347-470X

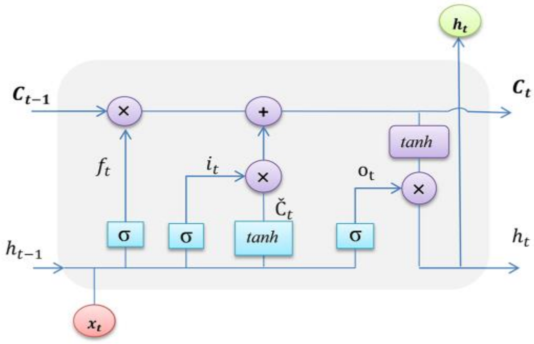

Figure 1. Architecture of single cell of LSTM

A LSTM network consists of four fundamental parts: a cell, an input gate, an output gate, and a forget gate. The cell transmits data over a variety of time periods. The gates facilitate the transfer of data between the cell's input and output. An LSTM network calculates its outputs in the following manner:

$$i_t = \sigma(W_i.[h_{(t-1)}, x_t] + b_i) \quad (1)$$

$$f_t = \sigma(W_f.[h_{(t-1)}, x_t] + b_f) \quad (2)$$

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad (3)$$

where x_t represents the input variable at a certain time; W_i, W_f, W_c, W_o refer to the weight matrices for the gates; f_t, o_t, and C_t refer to the forget gate, output gate, and cell output, respectively. The symbol σ denotes the sigmoid activation function, and h_t is refers to the hidden state at time t. Additionally, b_i, b_f, and b_o stand for the biased values of various gates. The new candidate value vector $\hat{C}_t$ is developed by the hyperbolic tangent function ($tanh$) and also added to the cell state C_t.

$$\widehat{C_t} = \tanh\ (W_c.[h_{(t-1)}, x_t] + b_c) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \widehat{C_t} \quad (5)$$

$$h_t = o_t \times tanh\ (C_t) \quad (6)$$

The $\boldsymbol{C_t}$ and $\boldsymbol{o_t}$ of the hyperbolic tangent function ($\boldsymbol{tanh}$) are utilized to update the $\boldsymbol{h_t}$ state [16].

3. RESEARCH METHODOLOGY

3.1 Input Dataset

This study presents the actual data collected from the Thermal Power Plant of Wasit in Iraq. Hourly load and temperature measurements from August 1 to October 31, 2022 comprise the data. Six electricity-generating units equip this power station. This study uses unit 1 and unit 2 load values of the power plant to predict the future loads for these units. We have suggested two models to handle the acquired data: the all-data model and the weekday model. We execute the suggested models using MATLAB R2023a.

3.2 Reprocessing Dataset

The input dataset passes through different steps before the suggested approaches are implemented. The first step is data cleaning; it is the initial step following data collection when trying to develop an effective and reliable model [17]. In our load data, we have 7 hours in unit 1 and 5 hours in unit 2 equal to zero on August 6-7 because the power plant is shut down during these hours, so we used the correcting data inconsistencies method to replace these data so that the accuracy of the results is not affected. The second step involves feature selection, which refers to the methods used to choose the portion of input features that are most relevant to the expected target value. To enhance the accuracy of electrical demand prediction, the dataset incorporates the hourly load data, the corresponding temperature for each day, the day of the week, and the hour of the day. The third stage divides the data into separate parts. This study uses a dataset consisting of more than 2000 lines, each representing the load values for every day of the week. We divided the data into two models, as mentioned previously:

a) **All-Data Model:** It includes data for all days, from August 1, 2022, to October 31, 2022.
b) **Week-Days Model:** It includes data for the days of the week, excluding holidays (Friday and Saturday).

The fourth step is encoding the dataset, which means encode the day of the week as a numerical variable from 0 to 6 instead of a string variable, and encode the hour of the day as a numerical variable from 0 to 24. The fifth step is to normalize the load data to be between 0 and 1. The mathematical formula for normalization is:

$$Z = \frac{y - \min(y)}{\max(y) - \min(y)} \quad (7)$$

Here, z refers to the normalized value, y refers to the actual value, $min(y)$ is the minimum value of y, and $max(y)$ is the maximum value of x [18]. The last step is converting to matrices, which means storing the training and testing values within matrices, in order to facilitate handling and avoid errors in the training process.

3.3 Hyperparameters Tuning

Tuning is the process of selecting the ideal set of hyperparameters for a learning algorithm. The parameters utilized in this research are:

- *Optimizer:* It is employed during training to update the model weights.
- *Learning rate:* It specifies the step size at each iteration.
- *Learn rate drop factor:* is a multiplicative factor that is applied to the learning rate when a particular number of epochs have passed.
- *Learn rate drop period:* is the number of epochs at which the learning rate is dropped.
- Max. epochs It refers to the maximum number of times the full training dataset is processed by the model.
- *Mini batch size:* is the minimum number of samples utilized in each training iteration.

3.4 Evaluation Metrics

The suggested approach utilizes four indices to evaluate its effectiveness. The mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), and

Research Article | Volume 12, Issue 4 | Pages 1461-1465 | e-ISSN: 2347-470X

the coefficient of determination (R-squared) are four metrics represented by:

$$MAE = \frac{1}{N_o}\sum_{i=1}^{N_o}|\hat{y} - y| \quad (8)$$

$$MAPE = \frac{1}{N_o}\sum_{i=1}^{N_o}\frac{|\hat{y}-y|}{\hat{y}} \quad (9)$$

$$RMSE = \sqrt{\left[\frac{1}{N_o}\sum_{i=1}^{N_o}(\hat{y} - y)^2\right]} \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N_o} y-\hat{y}}{\sum_{i=1}^{N_o} y-\bar{y}} \quad (11)$$

where N_o stand to the number of test samples, $\hat{y}$ is the predicted value, and $\bar{y}$ refers to the mean of $\hat{y}$ value. *Figure 2* illustrates the flowchart of the suggested approach.

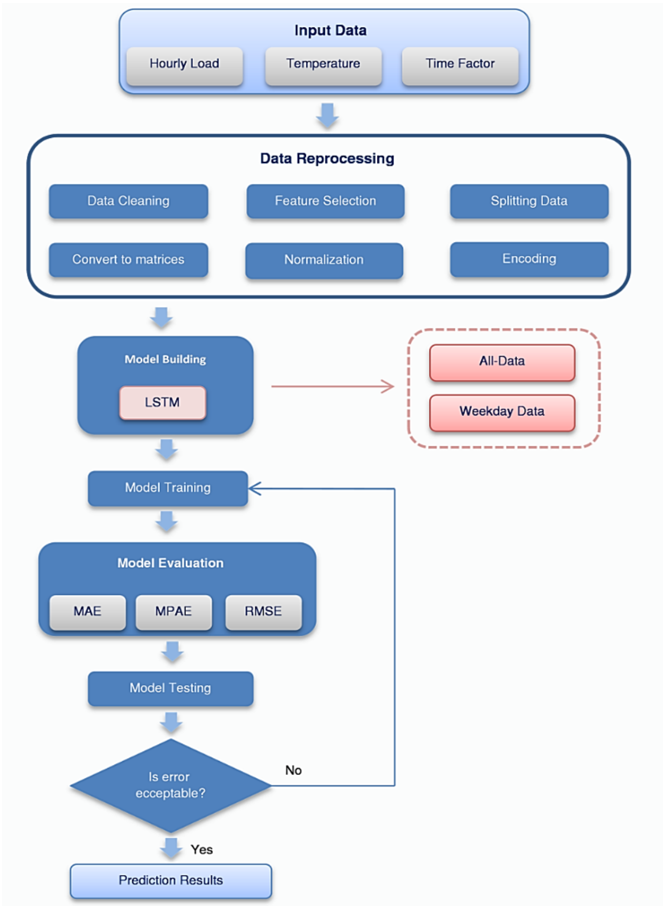

Figure 2. Flowchart of the proposed methodology

4. RESULTS

The dataset of all the suggested models is divided into training sets for training and testing processes to evaluate its performance on new data. The training set comprises 70% of the total data, whereas the test values comprise the remaining 30%. *Table 1* illustrates the number of values after being divided into training and testing sets for the models. All models and cases in this research are trained by the ADAM optimizer for the input data with a 0.01 initial learn rate.

Table 1. The number of samples (rows) of input data used in the two models proposed

Model	No. of samples	
	Training	Testing
All-Data	1545	663
Weekday	1109	476

4.1. All-Data Model

The features used for training the input dataset to predict the load demand using the LSTM approach for Unit 1 and Unit 2 are described in *table 1*. *Figure 3* and *figure 4* display the simulation results for the first model's prediction for Unit1 and Unit 2, respectively.

Table 2. The features used in LSTM approach for training process of all-data model

Feature	Unit 1	Unit 2
LSTM layer	1	1
Hidden cell	280	200
Learn rate drop factor	1	1
Learn rate drop period	2	2
Max. epochs	200	200

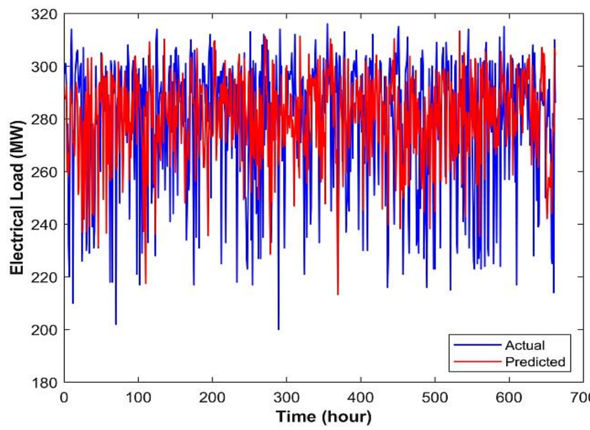

Figure 3. Electrical load prediction results of Unit 1 based on the LSTM approach using all input data (all days)

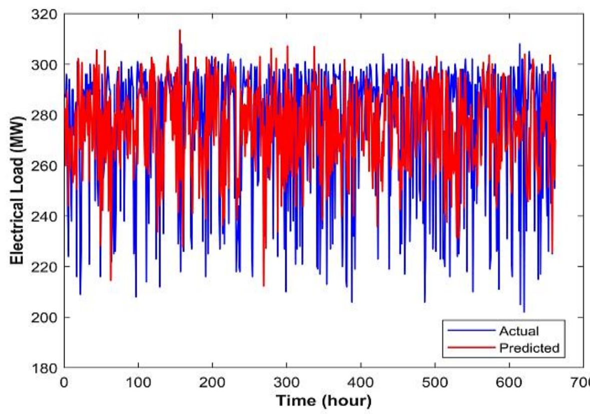

Figure 4. Electrical load prediction results of Unit 2 based on the LSTM approach using all input data (all days)

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 4 | Pages 1461-1465 | e-ISSN: 2347-470X

The proposed model requires a small number of epochs for learning, preventing an exceedingly slow learning process. Additionally, the proposed model provides an accurate prediction for the STLP problem. The performance comparison findings of load prediction between Unit 1 and Unit 2 using the evaluation indices for the all-data model are presented in *table 3*.

Table 3. Evaluation metrics of Unit 1 and Unit 2 for the all-data model

	MAE	MAPE	RMSE	R^2
Unit 1	8.1646	2.7490	0.0123	0.7706
Unit 2	5.2593	1.8262	0.0079	0.8470

4.2. Weekday Model

The features used for training the input dataset to predict the load demand using the LSTM approach for Unit 1 and Unit 2 described in *table 4*. *Figure 5* and *figure 6* illustrate the simulation results for the second model's prediction for Unit1, and Unit 2, respectively.

Table 4. The features used in LSTM approach for training process of weekday model

Feature	Unit 1	Unit 2
LSTM layer	1	1
Hidden cell	350	350
Learn rate drop factor	1	1
Learn rate drop period	2	2
Max. epochs	300	300

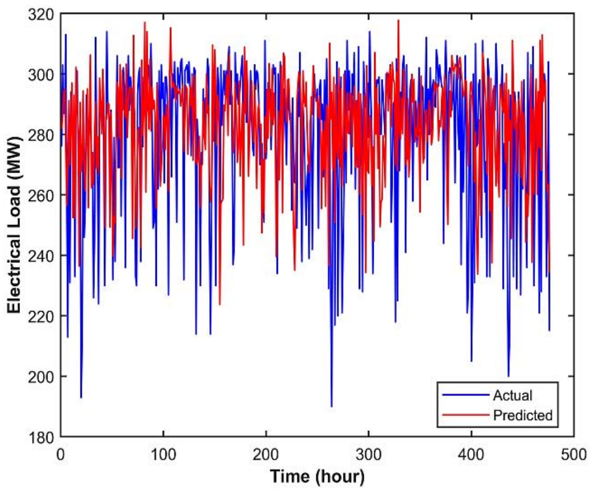

Figure 5. Electrical load prediction results of Unit 1 based on the LSTM approach using only weekday data (Sunday to Thursday)

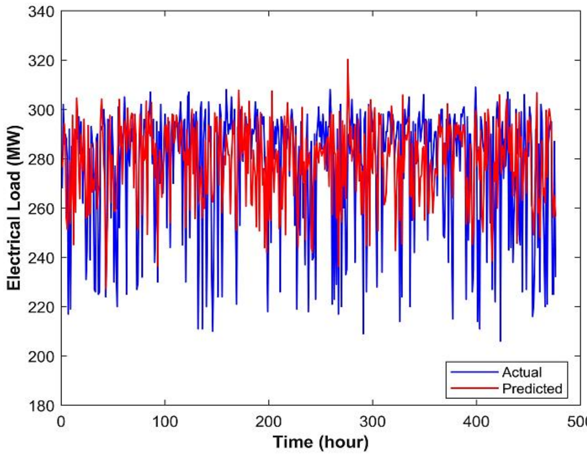

Figure 6. Electrical load prediction results of Unit 2 based on the LSTM approach using only weekday data (Sunday to Thursday)

In this model, the quantity of data is less than in the previous model. Therefore, we need a greater number of hidden cells for the two generating units. Furthermore, increasing the number of maximum epochs resulted in improved prediction accuracy. The performance comparison findings of load prediction between Unit 1 and Unit 2 using the evaluation indices for the weekday model are presented in *table 5*. In this model, the value of these indices increases.

Table 5. Evaluation metrics of Unit 1 and Unit 2 for the weekday model

	MAE	MAPE	RMSE	R^2
Unit 1	8.1418	3.0380	0.0171	0.7281
Unit 2	8.4016	3.1349	0.0177	0.7395

5. CONCLUSION

This paper employed the LSTM approach to predict the short-term loads. We evaluated the efficacy of the proposed models using actual data from the Wasit Thermal Power Plant in Wasit Governorate, Iraq, as a real-world example. We implemented the proposed methodology on two distinct data models: an all-model data model that encompasses all values from the input dataset, and a weekday model that includes the values of the input dataset for weekdays from Sunday to Thursday.

The prediction findings obtained from MATLAB R2023a demonstrated that the LSTM approach for the all-data model has more precise load prediction than the weekday model measured by prediction error (MAE, MAPE, RMSE, and R-squared). The average MAE for the all-data model is 6.71195, and for the weekday model, it is 8.2717. The average MAPE for the all-data model is 2.2876, and for the weekday model, it is 3.08645. The average RMSE for the all-data model is 0.0101, and for the weekday model, it is 0.0174. The average R-squared for the all-data model is 0.8088, and for the weekday model, it is 0.7338.

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 4 | Pages 1461-1465 | e-ISSN: 2347-470X

REFERENCES

[1] Baliyan, A., Gaurav, K., & Mishra, S. K. A review of short term load forecasting using artificial neural network models. *Procedia Computer Science*, 2015, *48*, 121-125. https://doi.org/10.1016/j.procs.2015.04.160

[2] Metaxiotis, K., Kagiannas, A., Askounis, D., & Psarras, J. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. *Energy conversion and Management*, 2003, *44*(9), 1525-1534. https://doi.org/10.1016/S0196-8904(02)00148-6

[3] Walther, J., Spanier, D., Panten, N., & Abele, E. Very short-term load forecasting on factory level–A machine learning approach.. *Procedia CIRP*, 2019, 80, 705-710. https://doi.org/10.1016/j.procir.2019.01.060

[4] Din, G. M. U., & Marnerides, A. K. Short term power load forecasting using deep neural networks. In *2017 International conference on computing, networking and communications (ICNC)*. Silicon Valley, CA, USA, 2017, Jan. https://doi.org/10.1109/ICCNC.2017.7876196

[5] Sheshadri, G. S. Electrical load forecasting using time series analysis. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*, Vijiyapur, India, October 2020. https://doi.org/10.1109/B-HTC50970.2020.9297986

[6] Tarmanini, C., Sarma, N., Gezegin, C., & Ozgonenel, O. Short term load forecasting based on ARIMA and ANN approaches. *Energy Reports* 2023, 9, 550-557. https://doi.org/10.1016/j.egyr.2023.01.060

[7] Dhaval, B., and Deshpande, A. Short-term load forecasting with using multiple linear regression. *International Journal of Electrical and Computer Engineering*, 2020, 10(4), 3911-3917. https://doi.org/10.11591/ijece.v10i4.pp3911-3917

[8] Rao, M. S. S., Soman, S. A., Menezes, B. L., Chawande, P., Dipti, P., and Ghanshyam, T. An expert system approach to short-term load forecasting for Reliance Energy Limited Mumbai. In *IEEE Power India Conference*, New Delhi, India, 10 April 2006. https://doi.org/10.1109/POWERI.2006.1632604

[9] Ahmad, W., Ayub, N., Ali, T., Irfan, M., Awais, M., Shiraz, M., & Glowacz, A. Towards short term electricity load forecasting using improved support vector machine and extreme learning machine. *Energies*, 2020, 13(11), 2907. https://doi.org/10.3390/en13112907

[10] Çevik, H. H., & Çunkaş, M. Short-term load forecasting using fuzzy logic and ANFIS. *Neural Computing and Applications*, 2015, *26*, 1355-1367. https://doi.org/10.1007/s00521-014-1809-4

[11] Arvanitidis, A. I., Bargiotas, D., Daskalopulu, A., Laitsos, V. M., & Tsoukalas, L. H. Enhanced short-term load forecasting using artificial neural networks. *Energies*, 2021, 14(22), 7788. https://doi.org/10.3390/en14227788

[12] Eskandari, H., Imani, M., & Moghaddam, M. P. Convolutional and recurrent neural network based model for short-term load forecasting. *Electric Power Systems Research*, 2021 *195*, 107173. https://doi.org/10.1016/j.epsr.2021.107173

[13] Wu, L., Kong, C., Hao, X., & Chen, W. A short-term load forecasting method based on GRU-CNN hybrid neural network model. *Mathematical problems in engineering*, 2020, (1), 1428104. https://doi.org/10.1155/2020/1428104

[14] Wang, Y., Gan, D., Sun, M., Zhang, N., Lu, Z., & Kang, C. Probabilistic individual load forecasting using pinball loss guided LSTM. *Applied Energy*, 2019, *235*, pp 10-20. https://doi.org/10.1016/j.apenergy.2018.10.078

[15] Jiao, R., Zhang, T., Jiang, Y., & He, H. Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network. *IEEE Access*, 2018, *6*, 59438-59448. https://doi.org/10.1109/ACCESS.2018.2873712

[16] Rafi, S. H., Deeba, S. R., & Hossain, E. A short-term load forecasting method using integrated CNN and LSTM network. *IEEE Access*, 2021, *9*, pp 32436-32448. https://doi.org/10.1109/ACCESS.2021.3060654

[17] Venkatesan, S. Design an intrusion detection system based on feature selection using ML algorithms. *Mathematical Statistician and Engineering Applications*, 2023, *72*(1), pp 702-710.

[18] Velasco, L. C. P., Arnejo, K. A. S., & Macarat, J. S. S.. Performance analysis of artificial neural network models for hour-ahead electric load forecasting. *Procedia Computer Science*, 2022, *197*, 16-24. https://doi.org/10.1016/j.procs.2021.12.113.