Prediction of Electricity Load at PT PLN (Persero) Jayapura Using Recurrent Neural Network and Long Short-Term Memory Models

Dwi Putra Prasatya[1,3], Finny Oktariani[1,2]

[1]Computational Science, Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology, Bandung, Indonesia
[2]Combinatorial Mathematics Research Group, Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology 10, Bandung 40132, Indonesia
[3]PT. PLN (Persero), Jakarta, Indonesia
Email : dwi.putrha@gmail.com

Abstract. The primary challenge in Power System Operation Planning is the uncertainty in predicting electricity load. Inaccurate electricity demand forecasts can lead to issues such as resource wastage, increased operational costs, and supply failure risks. Traditionally, operational planning has relied on estimating load history using Microsoft Excel worksheets, with calculations based on load growth (%) from previous periods. This research aims to improve the accuracy of electricity load prediction for system operation planning at PT. PLN (Persero) Jayapura by utilizing deep learning models, specifically Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM). For model optimization, a grid search method was employed for hyperparameter tuning, ensuring the best performance in load forecasting. The study was conducted at PT. PLN (Persero) Jayapura, using daily electricity load data from January 2020 to August 2024, sourced from the SCADA (Supervisory Control and Data Acquisition) histori server. The results showed that the LSTM model outperformed the traditional RNN. While the RNN model achieved a Mean Absolute Error (MAE) of 1.106, a Root Mean Squared Error (RMSE) of 1.7650, and a Mean Absolute Percentage Error (MAPE) of 0.0142, the LSTM model demonstrated more accurate predictions with a MAE of 1.0047, RMSE of 1.6186, and MAPE of 0.0129. These findings demonstrate the potential of LSTM, enhanced by grid search optimization, for improving load forecasting accuracy and contributing to more reliable power system operation planning.

Keywords: *Deep Learning, Multilayer Perceptron, Long Short-Term Memory, Electricity Load Forecasting, System Operation Planning.*

1. Introduction

Amidst global dynamics and increasingly intense economic competition, Indonesia has successfully maintained consistent economic growth. One of the key factors supporting economic growth is the rapid increase in electricity consumption across all sectors of society. Electricity, as a primary energy source,

Received ________, Revised _________, Accepted for publication __________
 ISSN: xxxx-xxxx, DOI: 10.5614/xxxx

serves as the backbone for the development of industry, business, and daily household life.

To meet the community's electricity needs, PT. PLN (Persero) has undertaken upstream to downstream efforts to ensure system readiness across all regions in Indonesia, based on the Ministry of Energy and Mineral Resources Regulation of the Republic of Indonesia Number 20 of 2020 concerning the Grid Code of the Electric Power System [1].

One of the main challenges in Power System Operation Planning is the uncertainty in electricity load forecasting. The inability to accurately predict electricity demand can lead to various problems, such as resource waste, increased operational costs, and supply failure risks. Currently, PLN's approach to forecasting electricity demand relies on a statistical approach, which involves calculating load growth (%) from the previous period.

Caicedo-Vivas et al. [2] demonstrated LSTM's superiority in short-term load forecasting for a Colombian grid operator, achieving a MAPE of 3.73%, outperforming methods like XGB. LSTM also proved robust in handling irregular data, reinforcing its reliability for dynamic and complex forecasting challenges.

Deep learning is a branch of machine learning that uses multi-layer neural networks to automatically learn representations from raw data. This method has significantly improved performance in various tasks such as image recognition, speech recognition, natural language processing, and drug discovery. Unlike traditional machine learning, which relies on manually designed features, deep learning can uncover complex structures in high-dimensional data through a process called backpropagation, which enables training of complex models with minimal human intervention. In the context of natural language processing, deep learning utilizes neural networks, such as recurrent neural networks (RNN) and long short-term memory (LSTM) networks, to learn distributed word representations. These models can predict the next word in a sequence and perform tasks such as machine translation. The advantage of neural networks over traditional statistical methods is their ability to generalize across semantically related words [3].

This study aims to leverage deep learning, specifically LSTM, to improve electricity load forecasting accuracy. LSTM is a specialized type of recurrent neural network designed to capture patterns in sequential data, such as electricity load data. To optimize the performance of the LSTM model, the study utilizes grid search for hyperparameter tuning, ensuring the selection of the best model configuration for accurate predictions.

2. Recurrent Neural Network (LSTM)

A Recurrent Neural Network (RNN) is a type of neural network specifically designed for processing sequential data, which is particularly useful for forecasting tasks. It maintains a 'state vector' in its hidden units that retains information about the history of all previous elements in the sequence. RNNs process input sequences one element at a time, allowing each output to depend on all prior inputs. This architecture is highly effective for tasks involving sequential inputs, such as time series forecasting, where the prediction of future values relies on the patterns learned from past data [3].

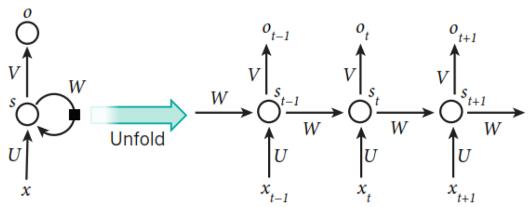

Figure 1. Recurrent Neural Network (RNN) Architecture

The diagram of the architecture illustrates a Recurrent Neural Network (RNN), where computations are performed sequentially over time, with each step depending on the previous input and the updated hidden state. At each time step t, the input x_t is processed and produces an output o_t, which depends on the entire sequence of previous inputs. The hidden state s_t at time t is influenced by the hidden state s_{t-1} from the previous time step, enabling the RNN to retain temporal information. This architecture uses the same parameters (matrices U, V, W) at each time step, known as weight sharing, to process input and update the hidden state. The backpropagation through time (BPTT) algorithm is applied to compute error gradients and update the model's parameters, minimizing errors in predictions or classifications of the output. Additionally, there are RNN variants that generate output sequences, where the output at one step is used as the input for the next.

3. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a type of artificial neural network architecture that falls under the category of Recurrent Neural Networks (RNN). LSTM is specifically designed to address the issues faced by RNNs in retaining long-term information in memory. [4]. LSTM has strong capabilities in predicting cases involving time-series data, one of which is forecasting electricity usage.

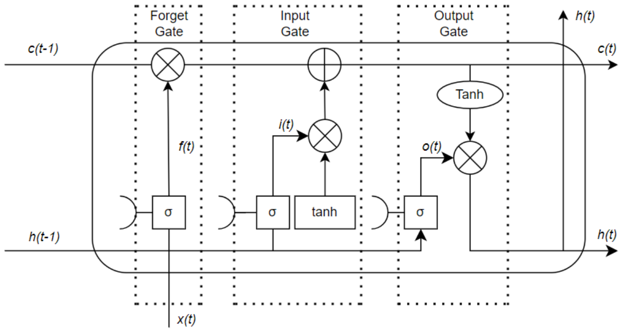

Figure 2. Long Short-Term Memory Architecture

The architecture of the LSTM model consists of three main components known as gates and one sub-component within a layer, or cell state. These components include the forget, input, and output gates, as well as the memory cell candidate, which is an additional part of the input gate. The gates function to control the flow of information by learning which input data in the sequence should be retained or discarded. The final information from the cell state is then passed to the next cell state and hidden node. Based on the architecture shown in Figure 1, the LSTM cell can be mathematically expressed as follows:

$$f_t = \sigma\,(W_f \cdot [h_{t-1}, x_t] + b_f \tag{1}$$

$$i_t = \sigma\,(W_i \cdot [h_{t-1}, x_t] + b_i \tag{2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

$$o_t = \sigma\,(W_o \cdot [h_{t-1}, x_t] + b_0 \tag{5}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

Where σ is the sigmoid function, tanh is the hyperbolic tangent function, and W and b represent the weights and biases learned during training.

In the LSTM model, the Mean Squared Error (MSE) loss function is used to measure the accuracy of the model's predictions by comparing the predicted values with the actual values and calculating the average of the squared differences. With MSE, the model is penalized more heavily for larger errors, encouraging it to make predictions that are closer to the true values. The MSE equation is as follows:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left(Y_i - \hat{Y}_i\right)^2 \tag{7}$$

Where $\mathrm{Y_i}$ represents the actual value, $\mathrm{\hat{Y}_i}$ represents the predicted value and and m is the total number of data points.

4. Evaluation Metrics

The evaluation metrics in electricity load prediction using the LSTM model aim to measure the accuracy and reliability of the model in capturing complex and temporal energy consumption patterns. Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) are the three indicators used in this study to assess the accuracy of the LSTM model applied. [5].

$$\mathrm{MAE} = \frac{1}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} |\mathrm{Y_i} - \mathrm{\hat{Y}_i}| \tag{8}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} \left(\mathrm{Y_i} - \mathrm{\hat{Y}_i}\right)^2} \tag{9}$$

$$\mathrm{MAPE} = \frac{100}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} \left|\frac{\mathrm{Y_i} - \mathrm{\hat{Y}_i}}{\mathrm{Y_i}}\right| \tag{10}$$

In the formulas for MAE, RMSE, and MAPE, the variable $\mathrm{Y_i}$ represents the actual or true value of the i-th data point, while $\mathrm{\hat{Y}_i}$ denotes the predicted value generated by the model for the i-th data point. The variable *m* indicates the total number of data points, or the sample size used in the calculations, representing the number of $\mathrm{Y_i}$ and $\mathrm{\hat{Y}_i}$ pairs in the dataset. These variables help compute the average error between the predicted and actual values, reflecting the model's accuracy.

Mean Absolute Error (MAE) measures the average magnitude of prediction errors while disregarding their direction; it represents the average of the absolute differences between predicted and actual values in the test set, considering all differences equally weighted. Root Mean Squared Error (RMSE) calculates the square root of the average squared prediction errors, giving greater weight to larger errors, making it more sensitive to outliers compared to MAE. Mean Absolute Percentage Error (MAPE) is a relative metric that expresses the average value of the relative errors as a percentage of the actual data.

5. Hyperparameter Tuning

Hyperparameter tuning refers to the process of selecting the most effective set of parameters that control the learning process during model training. These parameters, such as the learning rate, batch size, and optimizer, significantly influence how well the model learns from the data. The goal of tuning is to find the optimal combination that minimizes error and improves model performance. This process involves systematically adjusting these parameters and evaluating the model's performance to identify the configuration that results in the most accurate predictions. Table 1 presents the values used in the hyperparameter tuning process.

Table 1. Hyperparameter Value

Combination	Value
Optimizer	{Adam; RMSprop}
Learning Rate	{0,001; 0,0001}
Batch Size	{8; 16; 32}

6. Data and Model

In this study, the methodology begins with collecting electricity load data from the SCADA historical server at PT PLN (Persero) Jayapura, covering the period from 2020 to August 2024. The collected dataset consists of 1,705 rows and 49 columns, with each feature representing a 30-minute interval over a 24-hour period. The next stage is data preprocessing, which involves normalizing the data to a range between 0 and 1. This normalization is intended to facilitate the model training process.

The dataset is then divided into two parts: 90% of the data is used for training (from January 2020 to December 2023), and the remaining 10% is used for evaluation (from January 2024 to August 2024). In the model training phase, the Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) model architectures are built by specifying the number of units, activation functions, and additional layers as needed, as shown in Table 2.

Table 2. Model Architecture

Model : Sequential			
Layer	**Output Shape**	**Param #**	**Activation**
LSTM	(None, 100)	40400	Internal (Tanh, Sigmoid)
LSTM	(None, 50)	30000	
Dense	(None, 1)	51	

7. Results and Analysis

Both the RNN and LSTM models underwent hyperparameter tuning using grid search to identify the best-performing model. The results of the modeling are presented in Table 3.

Table 3. Hyperparameter Tuning Results for RNN and LSTM Models

RNN					
Optimizer	**Batch Size**	**Learning Rate**	**MAE**	**RMSE**	**MAPE**
Adam	8	0.0001	0.014	0.023	0.021
LSTM					
Optimizer	**Batch Size**	**Learning Rate**	**MAE**	**RMSE**	**MAPE**
RMSprop	32	0.001	0.016	0.025	0.024

The best model obtained was used for testing data. The testing results reveal a downward trend in Loss and Validation Loss, indicating that the RNN and LSTM models successfully learned from the data, as shown in Figure 3. This decline suggests that the model effectively minimized errors over time, achieving improved performance and stability across both training and validation datasets.

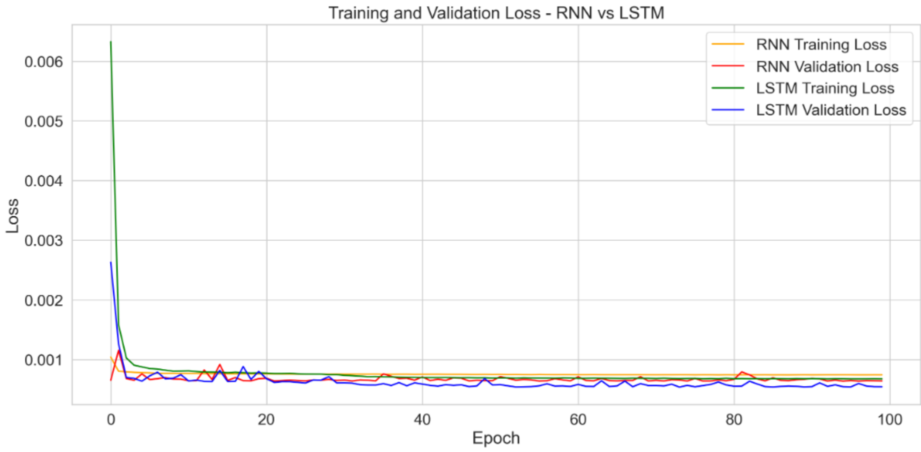

Figure 3. Training Loss & Validation Loss

Initially, the loss decreased rapidly, indicating that the model was learning effectively; however, after several epochs, the loss began to stabilize, signifying that the model had reached a point of convergence or stability in training. This pattern suggests that the model had successfully captured the underlying data patterns and was no longer making significant adjustments, indicating that further training would likely not yield substantial improvement in performance.

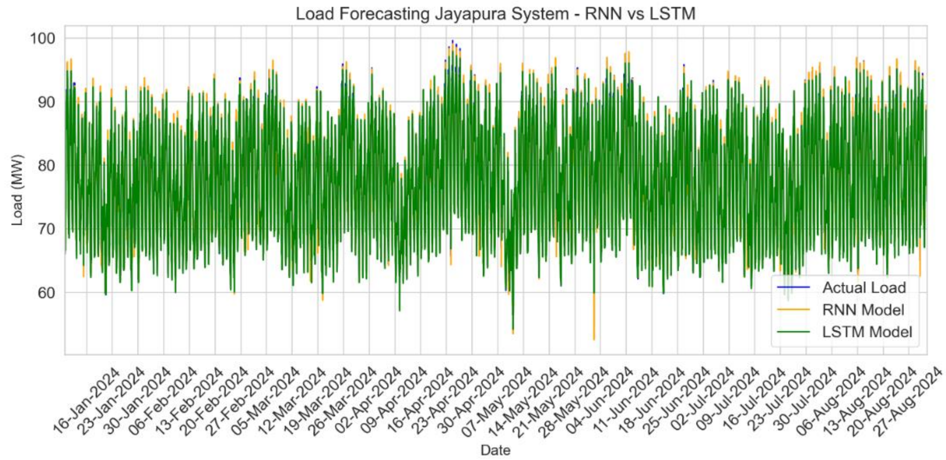

Figure 4. Performance Comparison of RNN and LSTM Models

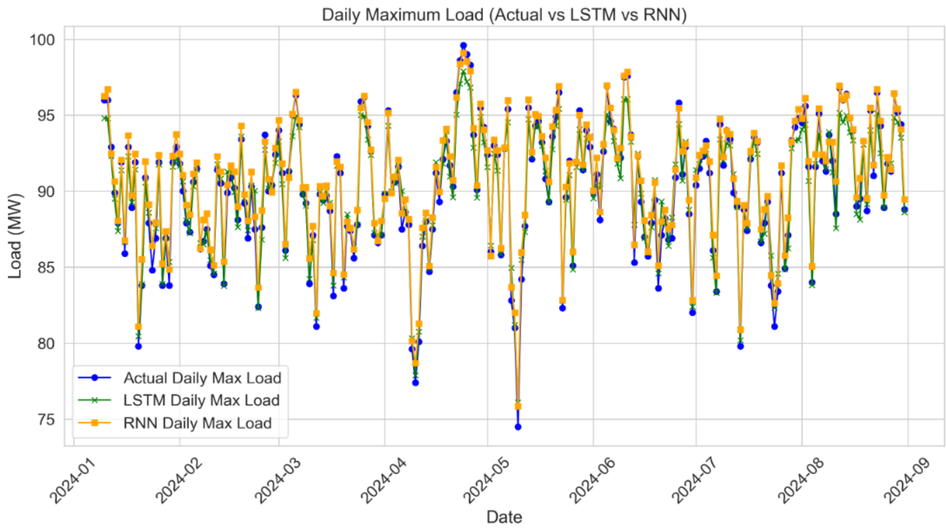

Figure 5. Daily maximum load values

Figure 4 shows the test results comparing the actual electricity load (blue) and the predicted load generated by the RNN model (orange) an LSTM model (green) in megawatts (MW) and Figure 5 shows the daily maximum load values . The X-axis represents the time index of the predicted data, while the Y-axis represents the electricity load values, ranging from approximately 60 to 100 MW. The test results demonstrate the models' ability to closely follow the pattern of actual values, with strong evaluation metrics. The RNN model achieved a Mean Absolute Error (MAE) of 1.106, Root Mean Squared Error (RMSE) of 1.7650, and Mean Absolute Percentage Error (MAPE) of 0.0142. Meanwhile, the LSTM model performed slightly better with a MAE of 1.0047, RMSE of 1.6186, and MAPE of 0.0129..

8. Conclusion

Based on the test results, the Long Short-Term Memory (LSTM) model outperforms the Recurrent Neural Network (RNN) model in terms of accuracy. The LSTM achieved a lower Mean Absolute Error (MAE) of 1.0047, compared to RNN's MAE of 1.106, indicating that the LSTM model is more accurate in predicting electricity load. Similarly, the LSTM showed a lower Root Mean Squared Error (RMSE) of 1.6186, whereas the RNN had a higher RMSE of 1.7650, further demonstrating the superior predictive accuracy of the LSTM. Additionally, the LSTM achieved a lower Mean Absolute Percentage Error (MAPE) of 0.0129 compared to RNN's MAPE of 0.0142, which signifies better performance in minimizing the percentage error between the predicted and actual values. These results suggest that the LSTM model is more optimal for electricity load forecasting, as it is better equipped to capture and model complex temporal dependencies, leading to more reliable and precise predictions..

Acknowledgement

Special thanks to PT. PLN (Persero) for providing support and data information for this study, as well as to the Bandung Institute of Technology, particularly the Computational Science Study Program, for their assistance and support throughout this research.

9. References

[1] Kementerian Energi dan Sumber Daya Mineral Republik Indonesia, “Aturan Perencanaan Dan Pelaksanaan Operasi (Scheduling And Dispatch Code - SDC),” 2020.

[2] J. S. Caicedo-Vivas and W. Alfonso-Morales, “Short-Term Load Forecasting Using an LSTM Neural Network for a Grid Operator,” *Energies (Basel)*, vol. 16, no. 23, Dec. 2023, doi: 10.3390/en16237878.

[3] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” May 27, 2015, *Nature Publishing Group*. doi: 10.1038/nature14539.

[4] T. H. Bao Huy, D. N. Vo, K. P. Nguyen, V. Q. Huynh, M. Q. Huynh, and K. H. Truong, “Short-Term Load Forecasting in Power System Using CNN-LSTM Neural Network,” in *Conference Proceedings - 2023 IEEE Asia Meeting on Environment and Electrical Engineering, EEE-AM 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/EEE-AM58328.2023.10395221.

[5] M. S. Hossain and H. Mahmood, “Short-Term Load Forecasting Using an LSTM Neural Network.”