PARAMETER
JURNAL MATEMATIKA, STATISTIKA DAN TERAPANNYA

e-ISSN. 2829-6303
Vol. 05, No. 2, September 2026
Pages 421-436

Comparison of ANN and LSTM Models for Forecasting Power Usage Effectiveness

Yuliana[1], Teguh Herlambang[2*], Endang Sulistiyani[3], Firman Yudianto[4]

[1,2,3,4]Department of Information Systems, Faculty of Business Economics and Technology Digital, Universitas Nahdlatul Ulama Surabaya, Indonesia
Jl. Raya Jemursari No. 57, Surabaya, 60237, Jawa Timur, Indonesia

E-mail Correspondence Author: *teguh@unusa.ac.id

Abstract

Energy efficiency in data center operations has become a critical issue in the digital era, including for public utilities such as Regional Drinking Water Companies (PDAM). One of the key indicators of data center efficiency is Power Usage Effectiveness (PUE), with PDAM recording a daily average of 2.11. Significant spikes in PUE values during certain months suggest operational dynamics that are not yet fully understood, highlighting the need for a comprehensive analysis of the contributing factors. This study employs Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) algorithms to forecast PUE values. The research process includes data understanding, preprocessing, modeling, evaluation, and forecasting, conducted under three time-based split scenarios: 60%:20%:20%, 40%:20%:40%, and 20%:20%:60%. Among these, the third scenario demonstrated the best performance. The ANN model achieved a Mean Absolute Percentage Error (MAPE) of 10.10%, a Root Mean Square Error (RMSE) of 0.2565, and a derived forecasting accuracy of 89.90%, calculated as 100% MAPE. Meanwhile, the LSTM model recorded a MAPE of 10.62%, RMSE of 0.2600, and derived forecasting accuracy 89.38%. These results fall within the 10%–20% error range, indicating reliable forecasting performance. The predicted PUE values were 2.38 for ANN and 1.85 for LSTM, both categorized as "Average" in terms of energy efficiency, suggesting room for improvement. The highest energy loads were attributed to cooling systems, specifically air conditioning (AC), air conditioning and mechanical ventilation (ACMV), and uninterruptible power supply (UPS) systems. Therefore, future infrastructure optimization efforts should prioritize improvements in these areas.

Keywords: *ANN, data center, energy efficiency, LSTM, PUE*

: https://doi.org/10.30598/parameterv5i2pp421-436

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright © 2026 The Author(s).
Research Ariticle ***Open Acces***
Email: jurnalparameter@gmail.com
Homepage: https://ojs3.unpatti.ac.id/index.php/parameter

Submitted: August 2025
Accepted: September 2026

1. INTRODUCTION

The use of information technology (IT) has become a vital element in supporting modern operations, where data centers play a key role as core infrastructure in the management of modern companies such as the Regional Water Company (PDAM). In measuring data center efficiency, one of the main indicators is Power Usage Effectiveness (PUE) **[1].** PUE is used to measure the ratio between the total energy consumed by the data center and the energy actually used by IT equipment **[2].** PDAM Surya Sembada Surabaya, a technology-based water utility, operates a traditional data center to support its IT and operational processes. The data center operates continuously and relies on IT equipment and related infrastructure, particularly power and cooling systems, which contribute to overall energy consumption. Therefore, energy efficiency is a critical factor in data center operations. The observed data shows that the lowest daily PUE value at PDAM Surya Sembada averages 2.11. Furthermore, strong PUE spikes in certain months indicate fluctuations in the data center's energy efficiency, which could be related to changing operating conditions. To improve energy efficiency and promote sustainable practices in the public sector toward an environmentally friendly data center, a more detailed analysis of the factors influencing PUE is needed **[3]** and promoting sustainable practices in the public sector toward Green Data Centers.

The trend of increasing PUE not only impacts operational costs but also contributes to a surge in carbon dioxide emissions in the environment. For instance, an increase in the PUE ratio can result in an estimated additional 3.5 tons of CO2 emissions annually **[4]**. This has become a major issue due to the high operational costs of data centers resulting from energy inefficiency. Additionally, the absence of predictive modeling to anticipate PUE fluctuations makes it difficult to plan energy savings at PDAM Surya Sembada in Surabaya. This issue is critical because even a 0.1 increase in PUE can add millions of rupiah in operational costs per year **[5]**. To address this inefficiency and align with Environmental, Social, and Governance (ESG) principles mandated by Presidential Regulation No. 22/2017 on the National Energy General Plan (RUEN) **[6]**, immediate strategic initiatives are required at PDAM.

One effective way to reach this efficiency goal is to predict PUE by using data with Machine Learning. Past data can be used to create models that find patterns and things that affect PUE changes. Earlier studies have shown that machine learning methods can work well for predicting PUE. Research that used the -layer Perceptron (MLP) method got an accuracy of 91%, in predicting PUE **[3]**. Showing how neural network models can understand patterns, in PUE data.. The research looked at just one prediction method, which made it hard to see how it compares to other deep learning methods. In another study, the Multi-layer Perceptron (MLP) algorithm was compared with Deep Neural Network (DNN) and Long Short-Term Memory (LSTM), with the LSTM model outperforming the other two models in predicting PUE **[7].** Additionally, PUE has also been predicted using a time series approach with the Long Short-Term Memory (LSTM) and Back-Propagation Neural Network (BPNN) algorithms. From the model evaluation results, it was concluded that the LSTM model was superior with an RMSE of 0.013 and R^2 of 0.89 **[8]**. Until the latest research on the comparison between Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Deep Neural Network (DNN). All models showed maximum performance with low evaluation values in the 80% training data split scenario. Among all these models, LSTM produced the lowest error rate with an MSE matrix of 4.488, MAE of 0.000289, and R2 of 0.9565977 **[9] [10].**

Previous approaches have failed to assess the predictive effects and have not considered feature selection [7]. In the study conducted by Deepika and Prakash [11], predictions were made without first knowing the characteristics of the data. Similarly, in previous research, the process of model integration and hypertuning was not performed in the Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Deep Neural Network (DNN) models [9]. However, the hypertuning process in time series forecasting, particularly the implementation of machine learning approaches, can improve training efficiency and model performance [12]. In handling time series data, appropriate techniques for dividing training and test data are also required, as data splitting must consider the order of values and should not be random [13] [14]. In addition to hypertuning and training/test data splitting, the three studies did not consider factors influencing PUE. Furthermore, among the numerous studies on PUE prediction, none have linked the forecasting results to their efficiency status to provide measurable policy recommendations.

Despite numerous studies on PUE forecasting, comprehensive research that integrates systematic hyperparameter optimization, specific data splitting techniques, and practical interpretations for public sector data centers remains limited. Therefore, this study, titled "A Comparative Study of the Performance of Artificial Neural Networks and Long Short-Term Memory for PUE Forecasting at PDAM Kota Surabaya," aims to develop a PUE forecasting model for the PDAM Surya Sembada Kota Surabaya data center. The novelty and main contribution of this research lie in a holistic framework that combines several approaches: First, a rigorous hypertuning process is conducted to establish the optimal network architecture, specifically identifying a configuration of an input layer with 64 neurons, followed by two hidden layers of 32 and 16 neurons utilizing the ReLU activation function. Second, a correlation matrix method is utilized to identify the most influential features for feature engineering, and the dataset is rigorously divided using a time-series splitting approach to preserve data sequential integrity. The model evaluation stage will use the RMSE, MSE, MAE, and MAPE metrics to ensure accuracy. Finally, unlike previous studies that primarily focus on model evaluation metrics, this research extends its contribution by interpreting the forecasted PUE values to determine the operational efficiency status of the PDAM data center. With this integrated approach, the research is expected to provide measurable policy recommendations and drive the transformation toward a Green Data Center in the public service sector.

2. RESEARCH METHOD

The purpose of this study is to analyze the performance of two algorithms, the Neural Network Algorithm (ANN) and the Long Short-Term Memory (LSTM) Algorithm, for forecasting Power Usage Effectiveness (PUE), identifying, and evaluating simulation results on each training and testing data set to obtain the best model for forecasting so that the efficiency level of the data center at PDAM Surya Sembada Kota Surabaya can be measured. In this study, the Cross-Industry Standard Process for Data Mining (CRISP-DM) method is used as a systematic framework to guide the entire research process [13]. This process begins with Business Understanding, Data Understanding, Data Pre-processing, Feature Engineering, Data Training and Data Testing, Modelling, Model Evaluation, and Data Visualization. The research stages to be conducted will be illustrated in the methodological framework presented in **Figure 1**.

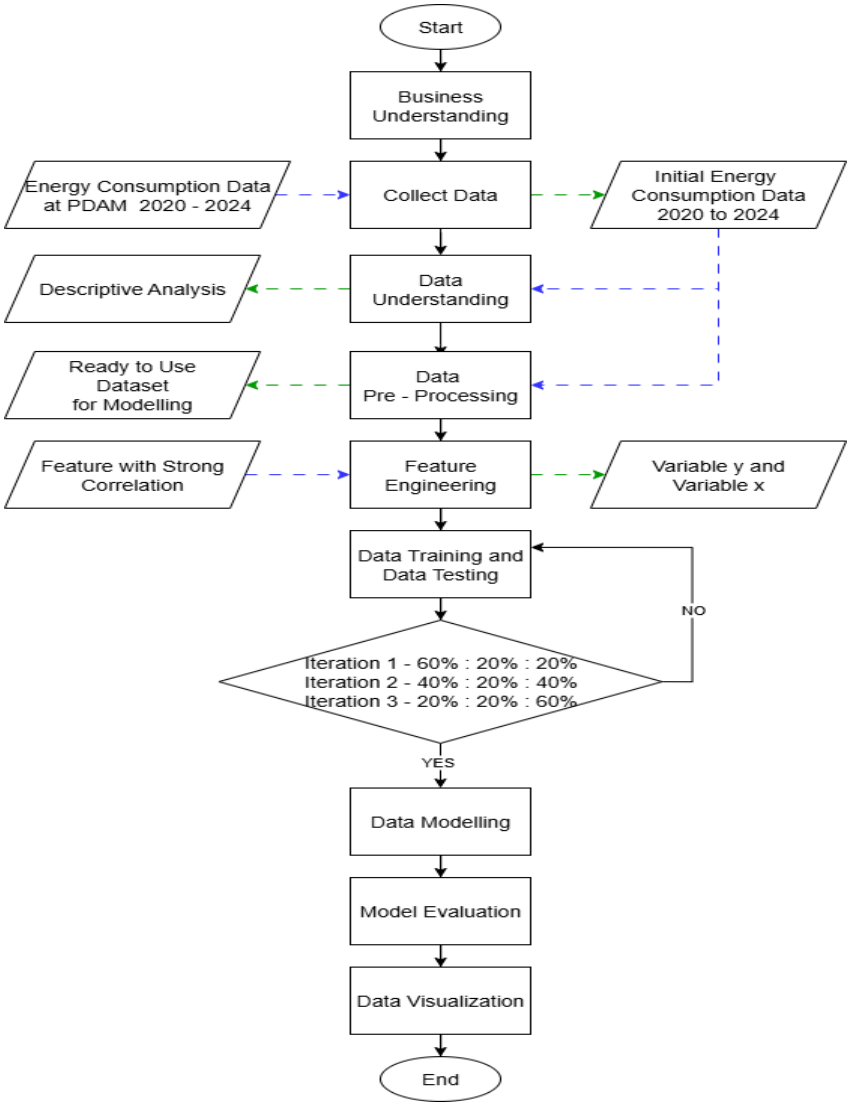


Figure 1. Research Methodology Flowchart

The details of the process in the flowchart above are explained as follows :

1. Business Understanding
   This study began by identifying problems occurring at PDAM Surya Sembada in Surabaya, suboptimal energy efficiency management at the data center. One important indicator used is Power Usage Effectiveness (PUE), but unpredictable fluctuations in PUE values can hinder efficiency evaluation and planning. Inaccurate PUE forecasting risks causing energy waste and unnecessary operational costs. Therefore, an accurate and adaptive forecasting model is needed to predict PUE values.
2. Collect Data
   The next step in this research is collecting relevant data to support the modeling process. The data used is primary data obtained directly through a field survey at PDAM Surya Sembada, Surabaya City. This dataset records energy consumption for the historical period of 2020 to 2024, thus providing a comprehensive foundation for analyzing data center operational efficiency. The collected data consists of 1,679 entries and includes eight main variables: Date, ITE, PUE, Total_Energy, UPS, AC, ACMV, and DCiE. These variables represent various aspects of energy use and performance indicators, which are crucial in developing predictive models and

evaluating performance trends over time. To build and evaluate the forecasting models, this historical dataset (2020–2024) is divided into training and validation/testing periods. Subsequently, the best-performing model is utilized to project the forecasting horizon for the upcoming month, specifically January 2025. The detailed structure of this dataset is presented in

Table 1.

Table 1. Dataset

Date	ITE	PUE	Total_Energy	UPS	AC	ACMV	DCiE
1/1/2020	1062.18	1.83	1941.78	1034.80	437.68	489.81	5.47
1/2/2020	1119.25	1.85	2068.30	997.85	493.49	369.37	5.41
1/3/2020	1182.70	2.16	2559.57	1150.38	583.54	635.98	4.62
1/4/2020	803.87	1.79	1438.95	723.56	571.39	253.87	5.58
1/5/2020	1165.91	2.59	3021.61	1213.36	638.36	682.77	3.85
…	…	…	…	…	…	…	…
12/24/2024	1351.63	2.69	3631.01	1475.39	1075.89	1026.08	3.72
12/25/2024	1098.10	2.36	2588.13	1053.01	840.79	320.32	4.24
12/26/2024	1364.89	2.16	2943.36	1459.56	1008.60	481.85	4.63
12/27/2024	864.00	1.55	1336.85	850.55	387.94	387.57	6.46
12/28/2024	1017.87	1.67	1699.20	896.55	652.97	262.37	5.99

3. Data Understanding
   The collected dataset was then subjected to descriptive analysis, which consisted of understanding data patterns, data anomalies, and data distribution **[14].** This process was carried out to gain an initial understanding of the characteristics of the data that could be used in model development.
4. Data Pre-Processing
   After understanding the data, the next process is to prepare the data before modeling, so that the dataset used can meet the modeling requirements **[15].** This process consists of normalization and feature engineering.
5. Feature Engineering
   Feature engineering aims to select, combine, or engineer attributes (features) that are most relevant to the target to be predicted **[16].** The feature engineering process carried out in this study was greatly influenced by the results of the correlation analysis between features and the target variable, namely PUE.
6. Data Training and Data Testing
   Once the dataset is ready for modeling, it will be divided into training data and testing data with an iterative ratio in accordance with the time-based split method **[17].** This approach aims to comprehensively evaluate model performance, given that the behavior of predictive models can vary depending on the proportion and range of data used in the learning and testing processes.
7. Data Modelling
   At this step, the forecasting process is carried out by building two main models, namely Neural Network (NN) and Long Short-Term Memory (LSTM). Neural Network is used as a baseline model because of its ability to capture nonlinear relationships between variables, despite its limitations in handling complex time series data. As a comparison, LSTM is applied because its architecture is specifically designed to process sequential data and retain long-term information.

a. Neural Network

Artificial neural network is a computational system inspired by the way the human brain recognizes patterns and organizes information. This system consists of three main components: the input layer, hidden layer, and output layer, which are interconnected through weights and adjusted during the training process [18]. The process begins with input entering the system, which is then gradually transformed through each layer until it produces an output aligned with the predefined objective.

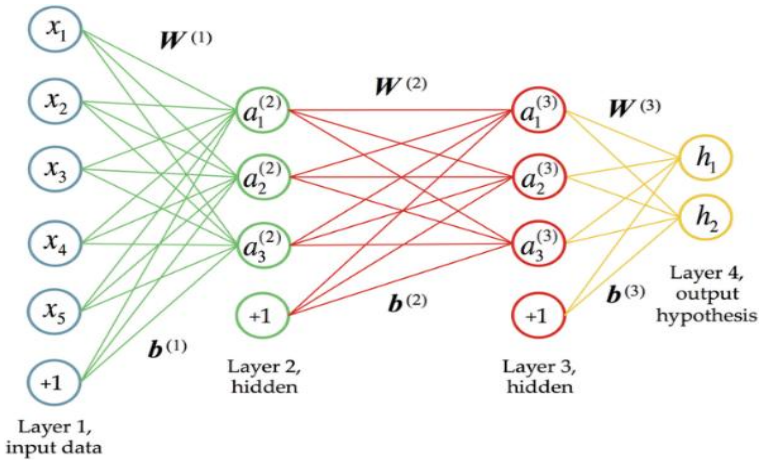


Figure 2. Neural Network Architectures

b. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a special type of Recurrent Neural Network (RNN), which itself is one of the architectures in the more general category of Neural Networks [18]. LSTM is able to retain long-term information through a gate mechanism that controls the flow of information.

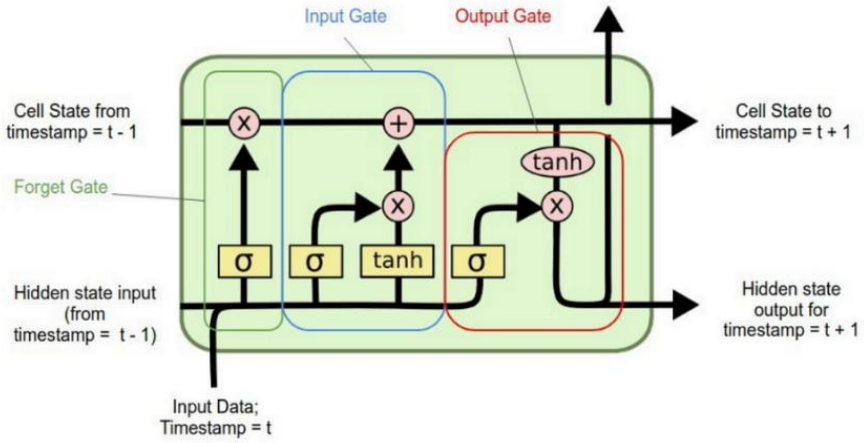


Figure 3. LSTM Architectures

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

Next, the candidate memory value is calculated using the tanh activation function, as shown in **Equation 2**.

$$\tilde{c}_t = \tanh\,(W_c \times [h_{t-1}, X_t] + b_i) \quad (2)$$

The forget gate determines the proportion of information from previous memories that will be retained or discarded, and it is formulated as shown in **Equation 3**.

$$f_t = \sigma\left(W_i \times [h_{t-1}, X_t] + b_i\right) \tag{3}$$

Cell memory is updated by combining information retained from the previous state with new information filtered by the input gate. The update process is formulated as shown in Equation 4

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

Then, the output gate calculates the activation for the next hidden state, as shown in **Equation 5**.

$$o_t = \sigma\left(W_o \times [h_{t-1}, X_t] + b_o\right) \tag{5}$$

Then the output gate generates a new hidden state that is used both for prediction and as input to the next time step. The output gate formula is as follows:

$$h_t = O_t * tanh\left(\tilde{c}_t\right) \tag{6}$$

8. Evaluation Model

The performance of the models was evaluated using three main metrics, namely Root Mean Square Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE), to assess the accuracy level of each model. The accuracy scale of the models is shown in **Table 2.**

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|Actual_i - Predicted_i|}{Actual_i} \text{ x } 100\% \tag{7}$$

Mean Absolute Percentage Error is calculated as the average percentage of the absolute difference between actual and predicted values.

Table 2. MAPE Accuracy Scale **[19]**

Interval	MAPE
< 10%	Highly Accurate Forecast
10% - 20%	Good Forecast
20% - 50%	Acceptable Forecast
> 50%	Unacceptable Forecast

9. Data Visualization

The final results of the PUE modeling process are visualized through line plot graphs to facilitate the interpretation and analysis of model performance. This visualization presents a comparison between actual PUE values and the predicted results of each algorithm, making it easier to observe patterns, trends, and model accuracy levels.

3. RESULTS AND DISCUSSION

This section presents the results of the implementation of each process, starting from data understanding, modeling, data visualization, and model comparison. Specifically for the modeling phase, it is important to state that Power Usage Effectiveness (PUE) is established as the target variable to be forecasted. Meanwhile, the other variables—ITE, Total_Energy, UPS, AC, ACMV, and DCiE—are treated as the initial predictor variables. These predictors subsequently undergo feature selection

based on correlation analysis to determine the most influential inputs for the forecasting models.

3.1 Data Understanding

Before conducting a more in-depth analysis, it is important to perform an analysis using descriptive statistics to explore the characteristics of the data used in this study. These descriptive statistics provide numerical summaries such as the mean, standard deviation, minimum and maximum values, and quartiles that can help identify the distribution of data.

Table 3. Descriptive Statistics of The Dataset

Description	PUE	Total_Energy	AC	ACMV
Count	1679	1679	1679	1679
Mean	2.11	2424.99	725.56	479.00
Std	0.35	592.57	232.91	182.58
Min	1.06	1208.54	263.84	141.86
25%	1.80	1991.49	554.12	339.65
50%	2.11	2361.02	683.80	456.46
75%	2.41	2808.39	870.20	589.82
Max	2.80	4019.18	1566.91	1168.63

As shown in the **Table 3**, the count or number of rows in the dataset is 1,679. The descriptive statistics include the mean or average value, where the PUE variable has an average value of 2.11, with the highest PUE value at 2.8 and the minimum PUE value at 1.06. Additionally, **Figure 4**.2 also indicates that Quartile 1 (Q1) of the PUE variable is 1.80, meaning that 25% of PUE data has values below 1.90. Quartile 2 (Q2) or the median PUE is 2.11, which means that 50% of the data is below that value. And the third quartile (Q3) is 2.41, indicating that 75% of the PUE data is below the value of 2.41.

3.2 Modelling

In this section, we will discuss the simulation of the implementation of the Neural Network (NN) and Long Short-Term Memory (LSTM) algorithms in PUE forecasting with hyperparameter **[20]**. The simulation results will be discussed by comparing the estimation results of the Neural Network (NN) and Long Short-Term Memory (LSTM) with the analytical solution. In this case, the Neural Network (NN) and Long Short-Term Memory (LSTM) models will be compared with the estimation results obtained through three iterations using different training and testing data splits **[17]**, namely:

1. Iteration 1 - 60% : 20% : 20%
2. Iteration 2 - 40% : 20% : 40%
3. Iteration 3 - 20% : 20% : 60%

3.3 Data Visualization

This section presents a comparison graph between the actual values and the estimated values of the Neural Network (NN) and Long Short-Term Memory (LSTM) models for Data Center Power Usage Effectiveness (PUE) at each iteration.

3.3.1 Iteration 1

The **Figure 4** shows the results of the prediction comparison visualized with a scatter plot over 100 epochs in iteration 1 with a data split of 60%:20%:20%. This visualization uses three parameters: actual values, ANN model predictions, and LSTM model predictions. The actual value represents the true value, while the prediction value is the result of the model's prediction. The **Figure 4** shows that the ANN prediction points show a slight deviation from the actual values, especially in areas with sharper fluctuations. Meanwhile, the LSTM prediction distribution appears more regular and consistent. LSTM projects PUE values in the range of 1.90 to 2.20 with higher stability, while ANN estimates values between 2.00 and 2.35, showing sharper fluctuations, especially during peak consumption periods.

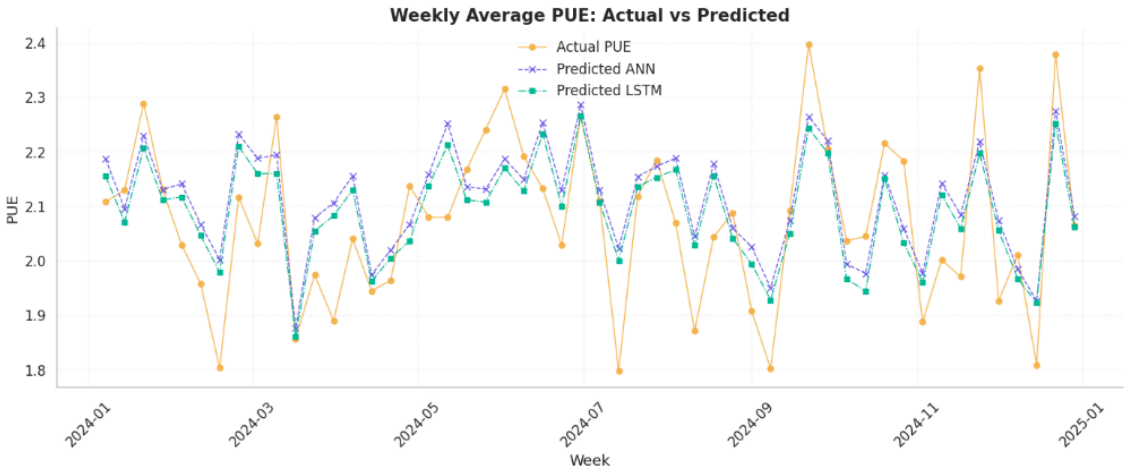


Figure 4. Simulation Result of Iteration 1

3.3.2 Iteration 2

The **Figure 5** shows the results of the prediction comparison visualized with a scatter plot over 100 epochs in iteration 2 with a data split of 40%:20%:40%. This visualization uses three parameters: actual values, ANN model predictions, and LSTM model predictions. The actual value represents the true value, while the prediction value is the result of the model's prediction.

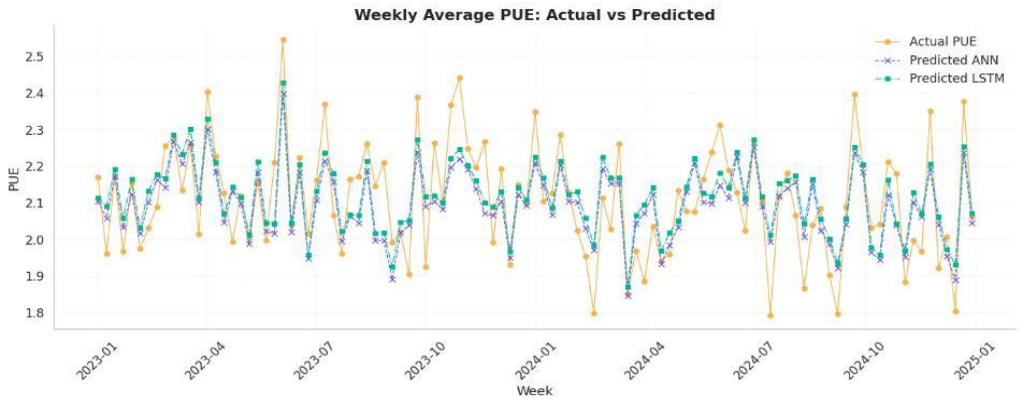


Figure 5. Simulation Result of Iteration 2

The weekly PUE graph for the period from 2023 to early 2025 illustrates the comparison between actual values and prediction results from the ANN and LSTM models. Actual values show PUE variations ranging from 1.85 to 2.60, with sharp fluctuations during certain periods. The LSTM predictions are more stable and consistently follow the weekly trend within the range of 1.90 to 2.20, while the ANN projects PUE values between 2.00 and 2.35 with a sharper response to changes. These differing characteristics indicate that LSTM is superior in maintaining temporal continuity.

3.3.3 Iteration 3

The **Figure 6** shows the results of the prediction comparison visualized with a scatter plot over 100 epochs in iteration 3 with a data split of 60%:20%:20%. This visualization uses three parameters: actual values, ANN model predictions, and LSTM model predictions. The actual value represents the true value, while the prediction value is the result of the model's prediction.

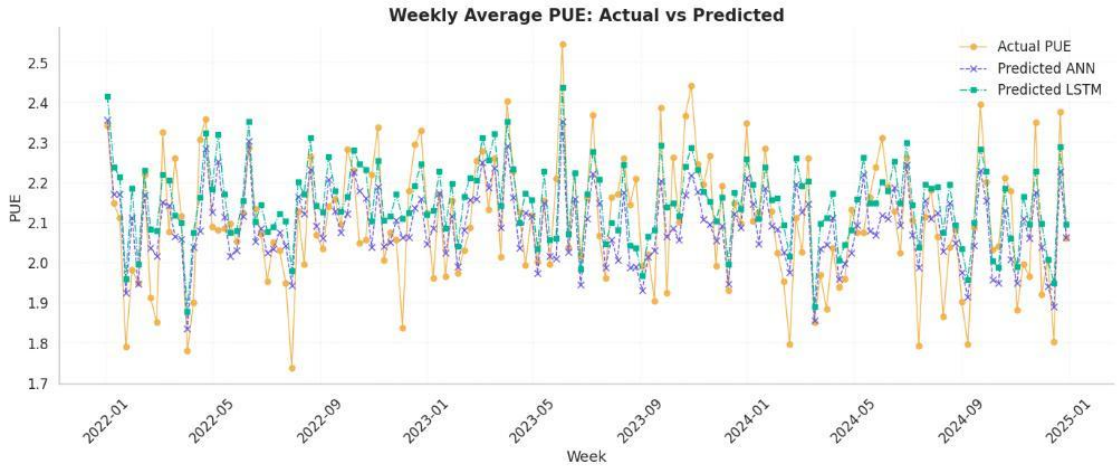


Figure 6. Simulation Result of Iteration 3

The weekly graph for iteration 3, covering the period from early 2022 to early 2025, shows a comparison of the predictive performance of the ANN and LSTM models against actual PUE values over a four-year horizon. Actual PUE values range from 1.7 to 2.5 with fairly sharp fluctuations in some weeks of 2023-05. The LSTM model projects PUE values within a narrower and more stable range of 2.00 to 2.35, maintaining the continuity of the weekly pattern. Meanwhile, the ANN predictions appear more aggressive, moving within a range of approximately 1.90 to 2.20, and responsive to data peaks. When compared to the results of Iteration 1 and 2, Iteration 3 demonstrates a significantly higher level of complexity and dynamic patterns in PUE, where the differences in characteristics between models become more evident.

3.4 Model Comparison

The simulation process was conducted using Python software version 3.12.0. In this case, the algorithms used in the simulation process were the Neural Network (NN) and Long Short-Term Memory (LSTM) models. The simulation model results obtained will be evaluated using the Root Mean Square Error (RMSE) and Mean Square Error (MSE) metrics to determine the error value and accuracy level of the model's performance, as well as the Mean Absolute Percentage Error (MAPE).

Table 4. Model Comparison

Metrics	Scenario	Hyperparameter	ANN	LSTM
RMSE	60% : 20% : 20%	Input : 64 2 Hidden : 32 and 16	0.271	0.2708
	40% : 20% : 40%	Input : 64	0.263	0.2708
	20% : 20% : 60%	2 Hidden : 32 and 16	0.2565	0.26
MSE	60% : 20% : 20%	Input : 64	0.07345	0.073327
	40% : 20% : 40%	2 Hidden : 32 and 16	0.06918	0.073327
	20% : 20% : 60%	Input : 64	0.06581	0.067577
MAPE	60% : 20% : 20%	2 Hidden : 32 and 16	11.19%	11.24%
	40% : 20% : 40%	Input : 64	10.29%	11.24%
	20% : 20% : 60%	2 Hidden : 32 and 16	10.10%	10.62%

3.5 PUE Analytics

After forecasting the PUE value using the ANN and LSTM models, the minimum and maximum PUE values were obtained. In the next stage, the PUE status of the data center will be analyzed based on the Green Grid standard **[2].** The analysis obtained is as follows on

Appendix 1.

Table 5. PUE Efficiency Values

PUE	DCiE	Level of Efficiency
3.0	33%	Very Inefficient

2.5	40%	Inefficient
2.0	50%	Average
1.5	67%	Efficient
1.2	83%	Very Efficient

From this, it can be concluded that the PUE status based on estimates for the next 3 months to 1 year is at an "average" level [21].

4. CONCLUSION

Based on the simulation and analysis results the third scenario showed the forecasting performance. The ANN model achieved a MAPE of 10.10% and an RMSE of 0.2565. The LSTM model had a MAPE of 10.62% and an RMSE of 0.2600. These results show that the ANN model performed better than the LSTM model in the third scenario. The findings suggest that machine learning-based forecasting can help support PUE monitoring and energy-efficiency management at PDAM Surya Sembada Surabaya. However this study has limitations. The prediction performance depends on the data and input variables that are used in the model. These may not fully capture factors that affect data center energy consumption and PUE fluctuations. Also the study only compares ANN and LSTM models. Other modeling approaches have not been tested. Future research should look into adding external variables. These could include conditions, cooling-system parameters and operational or workload-related factors. This would help better represent the variations in PUE. Moreover hybrid forecasting methods that combine ANN or LSTM with optimization or machine learning techniques could be explored. These methods might improve prediction accuracy and model robustness. Such improvements could offer support for proactive energy management and the use of Green Data Center practices, in the public sector.

Acknowledgments

The author would like to thank those who have helped during this research process. In addition, the author would also like to thank the anonymous reviewers for their valuable suggestions and comments. The author would like to extend their deepest gratitude to Dr. Teguh Herlambang, S.Si., M.Si., for the guidance, knowledge, and financial support provided during this research. Additionally, thanks to PDAM Surya Sembada Kota Surabaya as the research partner.

REFERENCES

[1] Laureguialzo, Enrique, "PUE: The Green Grid metric for evaluating the energy efficiency in DC (Data Center). Measurement method using the Power Demand.," 2011.

[2] V. Avelar, S. Electric, D. Azevedo, A. French, and E. Network Power, "PUE™: A Comprehensive Examination Of The Metric Editors," 2012.

[3] A. Kumar Jain, "Predictive Analytics-Driven Management of Data Center Lifecycles," *JETIR2506882 Journal of Emerging Technologies and Innovative Research*, vol. 12, 2025, [Online]. Available: https://learn.microsoft.com/en-us/azure/architecture/data-science-process/lifecycle

[4] O. Ilmiah Guru Besar Institut Teknologi Bandung and P. Kridanto Surendro, "Forum Guru Besar," 2024. [Online]. Available: www.itbpress.id

[5] M. Ilham *et al.*, "Analisa Beban Pendingin Pada Ruang Data Center Di PT BSH," *Jurnal Teknologi Industri*, vol. 12, no. 1, 2023.
[6] "Perpres Nomor 22 Tahun 2017," 2017.
[7] H. A. Ounifi, A. Gherbi, and N. Kara, "Deep machine learning-based power usage effectiveness prediction for sustainable cloud infrastructures," *Sustainable Energy Technologies and Assessments*, vol. 52, Aug. 2022, doi: 10.1016/j.seta.2022.101967.
[8] J. Bi, S. Li, H. Yuan, Z. Zhao, and H. Liu, "Deep Neural Networks for Predicting Task Time Series in Cloud Computing Systems," 2019, [Online]. Available: https://github.com/google/cluster-data
[9] K. Saranraj and C. M. Liu, "Enhancing Power Efficiency Prediction in Small-Scale Data Centers using Attention-Based Deep Neural Networks," in *2023 IEEE 5th International Conference on Architecture, Construction, Environment and Hydraulics, ICACEH 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 44–49. doi: 10.1109/ICACEH59552.2023.10452493.
[10] M. Y. Anshori, V. Asy'Ari, T. Herlambang, and I. W. Farid, "Forecasting occupancy rate using neural network at Hotel R," in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)*, 2023, pp. 347–351. doi: 10.1109/ICAMIMIA60881.2023.10427752.
[11] T. Deepika and P. Prakash, "Power consumption prediction in cloud data center using machine learning," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 2, pp. 1524–1532, 2020, doi: 10.11591/ijece.v10i2.pp1524-1532.
[12] H. Dhake, Y. Kashyap, and P. Kosmopoulos, "Algorithms for Hyperparameter Tuning of LSTMs for Time Series Forecasting," *Remote Sens (Basel)*, vol. 15, no. 8, Apr. 2023, doi: 10.3390/rs15082076.
[13] J. Han, J. Pei, and H. Tong, "Data Mining: Concepts and Techniques," 2023.
[14] N. Khatun, "Applications of Normality Test in Statistical Analysis," *Open J Stat*, vol. 11, no. 01, pp. 113–122, 2021, doi: 10.4236/ojs.2021.111006.
[15] A. Chernikov, C. W. Tan, P. Montero-Manso, and C. Bergmeir, "FRANS: Automatic Feature Extraction for Time Series Forecasting," Sep. 2022, [Online]. Available: http://arxiv.org/abs/2209.07018
[16] N. P. Nur Fauzi, S. Khomsah, and A. D. Putra Wicaksono, "Penerapan Feature Engineering dan Hyperparameter Tuning untuk Meningkatkan Akurasi Model Random Forest pada Klasifikasi Risiko Kredit," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 2, pp. 251–262, Apr. 2025, doi: 10.25126/jtiik.2025128472.
[17] M. Hanasov, M. Wolter, and E. Glende, "Time Series Data Splitting for Short-Term Load Forecasting," 2022.
[18] F. Chollet, *Deep Learning with Python*. 2021.
[19] N. Alifiya, S. Putri, D. E. Ratnawati, and B. Rahayudi, "Prediksi Kebangkrutan Menggunakan Metode Backpropagation (Studi Kasus: Perseroan Terbatas Terdaftar Pada Bursa Efek Indonesia)," vol. 3, no. 3, pp. 2548–964, 2019, [Online]. Available: http://j-ptiik.ub.ac.id
[20] C. Tzagkarakis, P. Charalampidis, S. Roubakis, A. Fragkiadakis, and S. Ioannidis, "Evaluating Short-Term Forecasting of Multiple Time Series in IoT Environments," Sep. 2022, [Online]. Available: http://arxiv.org/abs/2206.07784
[21] M. Uddin, A. Shah, R. Alsaqour, and J. Memon, "Measuring efficiency of tier level data centers to implement green energy efficient data centers," *Middle East J Sci Res*, vol. 15, no. 2, pp. 200–207, 2013, doi: 10.5829/idosi.mejsr.2013.15.2.2356.

Appendix 1. **PUE Analytics Status**

Aspect	Forecast Conditions	Ideal Conditions	Gap
PUE Values	Minimal PUE : a. ANN : 2.38 b. LSTM :1.85 Status*'Average'*	PUE Ideal ≤ 1.5	PUE forecast results for three months to one year based on the minimum PUE indicate an "Average" level or are not yet optimal. The difference between the ANN forecast results and the ideal PUE (2.38 – 1.5 = 0.88) and the LSTM forecast results and the ideal PUE (1.85 – 1.5 = 0.35)
	Maximum PUE : a. ANN : 2.39 b. LSTM :1.86 Status *'Average'*	PUE Ideal ≤ 1.5	The PUE forecast results for three months to one year based on the maximum PUE indicate a level of 'Average' or not yet optimal. The difference between the ANN forecasting results

			and the ideal PUE (2.39 – 1.5 = 0.89) and the LSTM forecasting results and the ideal PUE (1.86 – 1.5 = 0.36)

Appendix 2. **Result of Forecasting PUE Data Center**

Date	Forecast with ANN	Forecast with LSTM
1/1/2025	2.3881686	1.8529968
2/1/2025	2.3890526	1.8536553
3/1/2025	2.3852437	1.851782
4/1/2025	2.390148	1.8546408
5/1/2025	2.3892329	1.8539962
6/1/2025	2.3883448	1.8536282
7/1/2025	2.388386	1.8537447
8/1/2025	2.3914514	1.8551462
9/1/2025	2.3880053	1.853625
10/1/2025	2.39109	1.8550066
11/1/2025	2.3896658	1.8535796
12/1/2025	2.386403	1.852163