

# PREDICTING ELECTRICITY DEMAND USING MACHINE LEARNING AND DEEP LEARNING APPROACHES: A CASE STUDY OF PLN PALANGKARAYA

Rifki Zamzammi - 23522301 (*Author*)

Program Studi Magister Informatika  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung  
E-mail (gmail): 23522314@std.stei.itb.ac.id

**Abstract**— Access to and demand for electricity supply is a primary need supporting societal, business, social, and governmental functions. PLN is committed to improving quality of life by providing reliable electricity. The PLN Unit Layanan Pelanggan (ULP) in Palangkaraya Timur serves 109,074 customers as of December 2023, with an average monthly electricity consumption of 26.26 GWh and a year-on-year customer growth rate of 4.81%, with electricity consumption increasing by 8.21%, surpassing Central Kalimantan's economic growth rate of 4.14%. As a company focused on service quality and revenue enhancement through electricity sales, PLN, particularly PLN Palangkaraya, is expected to meet power supply demands and infrastructure needs through proper planning. This study predicts electricity demand by tariff and power using machine learning and deep learning based on historical electricity sales data from PLN Palangkaraya. The results indicate that the best model is the Random Forest Regressor with an MSE of 7717.28 and  $R^2$  of 0.994, while deep learning models like LSTM, GRU, and RNN show lower performance due to data complexity. This demonstrates that machine learning is more effective in predicting electricity demand in this case.

**Keywords**— *Machine Learning, Deep Learning, Supervised Learning, Electricity Sales, Prediction, Regression*

## I. INTRODUCTION

Access and demand for electricity supply has become one of the primary needs to support the needs of human life both in general, business activities, social, government and others. In accordance with one of PLN's missions to make electricity a medium to improve the quality of people's lives, PLN continues to strive to provide the best service to the needs and demand for electricity supply.

To reach and serve the community directly and more closely, PLN provides Customer Service Unit offices in each PLN operational area. PLN Palangkaraya Unit Layanan Pelanggan (ULP) is one of the PLN service units operating in the

Palangkaraya City area, Central Kalimantan Province. As of December 2023, the ULP Palangkaraya Timur serves 109,074 customers with an average monthly electricity consumption of 26.26 GWh and an average YoY growth in the number of customers of 4.81% and electricity consumption of 8.21% from the previous year, higher than the economic growth rate of Central Kalimantan Province in 2023, which is 4.14%.

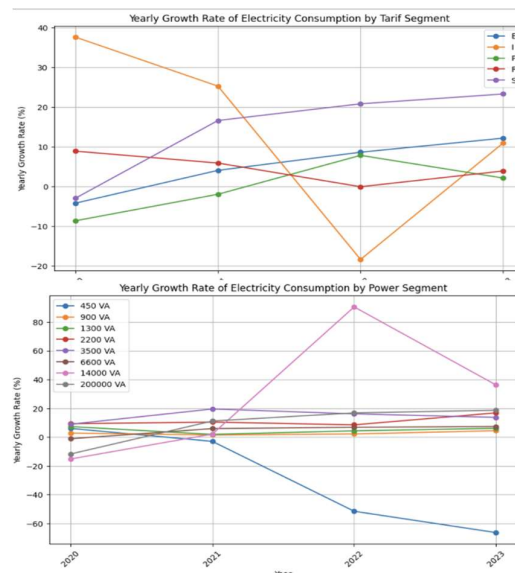


Fig 1 Electricity Consumption Trends by tariff and Power

The increase in the number of requests for access to electricity services, in addition to being one of the company's achievement targets, namely increasing electricity sales, is also a challenge for local units such as the readiness of connection materials according to the request, the condition and availability

of the existing network, as well as other technical problems such as power interruptions and voltage quality.

As a company that focuses on service quality and the need to increase revenue through electricity sales, it is expected that PLN, especially PLN Palangkaraya, can fulfill the supply of electrical power, proposed material needs and good electricity network infrastructure through proper planning.

In line with PLN's efforts to continuously improve and enhance services and infrastructure availability in the electricity sector, this research is aimed at forecasting the fulfillment of electricity demand needs for marketing activities and segmenting customer needs effectively in the Palangkaraya City area using Machine Learning and Deep Learning.

The aim or The benefits of this research are to provide an overview and prediction of electricity demand in Palangkaraya City, provide one of the recommendation options related to the planning and readiness of materials and infrastructure according to the predictions given, Provide insight into the potential sales and revenue of electrical energy in Palangkaraya city.

This research will be limited to the electricity sales report dataset of PLN ULP Palangka Raya Timur, so the results obtained may not be directly applicable to other regions or work units. Historical datasets of customer delta, power sales trend, and connected power delta were collected for five years, from January 2019 to December 2023. 3. The research will focus on predicting electricity demand for various tariffs and power offered by PLN ULP Palangka Raya Timur. However, it will not consider other variables that may affect electricity demand, such as economic factors or policy changes.

## II. LITERATURE REVIEW

### A. Machine Learning

Machine learning is a set of techniques that can handle and predict large amounts of data using learning algorithms. [1] first used the term machine learning. Arthur Samuel said that machine learning is a field of computer science that gives computers the ability to learn without clear programming. So it can be concluded that machine learning itself is a field of study in artificial intelligence (AI) that focuses on developing algorithms and models that allow computers to learn from data and make decisions or predictions without having to be explicitly programmed.

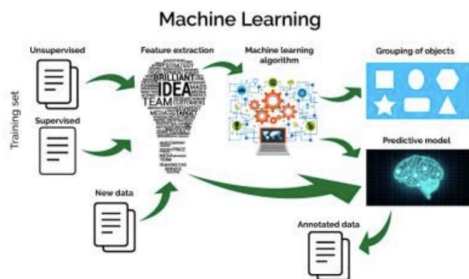


Fig 2 Machine Learning Illustration (Pantech, 2018)

In machine learning, there are scenarios such as : 1. Supervised Learning scenario, learning is done using learning data that is pre-labeled and then a model is created that makes predictions based on that data. 2. Unsupervised Learning Using the unsupervised learning scenario, learning uses unlabeled input learning data. After that, it tries to group the data based on the characteristics encountered. 3. In reinforcement learning scenarios the learning and test phases are mixed. To gather information the learner actively interacts with the environment so as to get a reward for each action of the machine learning.

### B. Deep Learning

Deep Learning is a subset of artificial intelligence that is the development of multiple layer neural networks to provide precision for tasks such as object detection, speech recognition, language translation and others. Deep Learning is a set of algorithms in Machine Learning that attempts learning in multiple levels, corresponding to different levels of abstraction[2], in this case usually using artificial neural networks, the levels in these learned statistical models correspond to different levels of concepts, where higher-level concepts are determined from lower levels, and lower-level concepts can help to define many higher-level concepts [3].

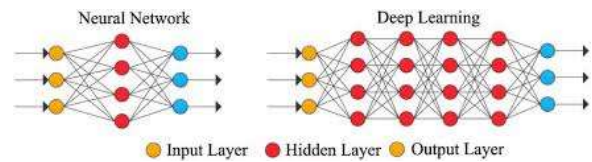


Fig 3 Illustration of Deep Learning Model

The main benefit of Deep Learning is that it can generate new features without the need for human intervention in the process. This shows that the technology can accomplish even complex tasks. Even for complex tasks that require extensive feature engineering.

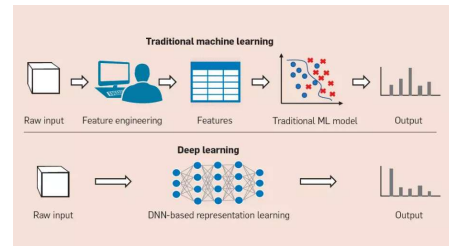


Fig 4 Comparison of ML and DL

Machine learning (ML) and deep learning (DL) are two sub-fields of artificial intelligence (AI) with different approaches and applications. ML uses various algorithms such as linear regression and decision trees to make predictions based on data and often requires manual feature engineering. ML algorithms tend to be simpler and require less computational resources, and work well on relatively small datasets. On the other hand, DL uses deep neural networks with many layers to automatically extract features from very large and complex data, such as facial recognition and text analysis. DL requires high computational resources and very large data volumes. In general, ML is more

suitable for problems with limited data and high interpretability requirements, while DL excels in tasks involving unstructured and complex data with large volumes.

### C. Electric Energy Demand

Electrical energy demand refers to the need to use electrical energy at any given time. It can be divided into two main categories: elastic demand and inelastic demand. Elastic demand is demand that can be reduced or increased based on price, such as electricity use for household appliances. Inelastic demand, due to physical or economic conditions, cannot be reduced or increased based on price, such as electricity use for power generation.

Demand for electricity is a key variable because of its relationship with economic activity and development. Electricity plays an important role in the economy and technology of every nation [4]. The availability of electrical energy is essential for economic development. Electricity energy can be forecasted or estimated based on factors such as economic growth, previous energy consumption, and technological changes [5].

Commonly used forecasting methods include quantitative methods, such as trend analysis, regression analysis, and qualitative methods, such as research analysis.

### D. PLN Tariff and Power Segmentation

In accordance with the provisions of Minister of Energy and Mineral Resources Regulation No. 28/2016 (amended by Minister of Energy and Mineral Resources Regulation No. 18/2017 and Minister of Energy and Mineral Resources Regulation No. 41/2017) concerning Electricity Tariffs provided by PT PLN (Persero), it is conveyed that the Electricity Tariff (TTL) is a tariff that may be imposed by the government to PLN customers.

The TTL provided by PLN is 37 tariff classes from the main sub-tariffs, namely Household (R), Business (B), Social (S), Industry (I), Public (P), Special Services (L) and Traction or Bulk (C). 13 of these divisions follow the Tariff Adjustment mechanism. Meanwhile, PLN's power segmentation is divided into 3 groups according to the connected power, namely 1. Low Voltage (TR) consumers for 220 - 197,000 VA power, which are divided into TR 1 Phase consumers (220 - 5,500 VA) and TR 3 Phase consumers (6,600 - 197,000 VA), 2. Medium Voltage (TM) consumers for power above 200 kVa, 3. High Voltage (TT) consumers for 30,000 kVA power.

## III. RELATED WORK

The Study titled Forecasting Revenue from Electricity Sales of Household Customers using Various Methods [6] This study aims to forecast electricity sales revenue of household customers using various algorithmic methods, such as k-nearest neighbor, random forest, linear regression, gradient increase

and adaboost. The results show that the random forest method has the highest R2 value, which increases the predicted revenue in 2021 by 21.44% of the actual revenue.

Predictive Analysis of Energy Consumption and Electricity Demand Using Machine Learning Techniques [7] Using Machine Learning algorithms, the author addresses the problem of anticipating energy demand prediction. The author compares the performance of three widely used gradient boosting algorithms namely, XGBoost, CatBoost, and LightGBM. MSE, RMSE, R-squared and MAE are used for error calculation and final evaluation.

Electricity Demand Forecasting In Kerala Using Machine Learning Models [8] research data was taken from Kerala State Electricity Board (KSEB). Linear Regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Regression (SVR), K-Nearest Neighbors (KNN), XGBoost, Artificial Neural Network (ANN) based Machine Learning (ML) methods have been applied to observe how these algorithms perform in forecasting electricity. The performance of the discussed ML methods has been analyzed based on various evaluation metrics such as accuracy, MAE, MSE, and RMSE. The results obtained from the analysis show that Random Forest ML approach outperforms other conventional ML approaches in terms of accuracy, MAE, MSE, and RMSE as it has the highest accuracy (82.72%) and the lowest MAE, MSE, and RMSE scores. (0.038, 0.002, and 0.054 respectively) compared to the other conventional ML approaches discussed.

Ultra-short term wholesale electricity price forecasting through Deep Learning [9] The purpose of this research is to address the extent to which Deep Learning can contribute to price threshold forecasting in volatile real-time electricity pricing markets. Additional research in this area is needed post NEM's transition to a 5-minute unified bidding and delivery cycle. Effective short-term forecasting can enable the transition from manual to automated bidding thereby fully optimizing the revenue potential for generators. The algorithm presented in this paper provides ultra-short horizon forecasting for 15-minute forward predictions. Queensland wholesale electricity market prices are used to present the research findings. The results show the Deep Learning model can achieve significant performance improvements.

Electricity Consumption Forecast Based on Empirical Mode Decomposition and Gated Recurrent Unit Hybrid Model [10] In this study, a hybrid empirical mode decomposition (EMD) and gated recurrent unit (GRU) model is proposed to predict users' electricity consumption. First, the original nonstationary electricity consumption time-series data is decomposed into several sets of stationary components through EMD, then each component is predicted through a multi-layer GRU network, and finally the prediction results of each component are combined to obtain the final forecast. The experimental results show that, compared with the direct use of LSTM, the proposed model can effectively reduce the error,

achieve better adjustment effect, and improve the training efficiency to a certain extent.

#### IV. METHODOLOGY

In completing this study, several stages of the research method were carried out which were described in the form of a flow chart related to the research process carried out while describing the research as a whole which can be seen in fig

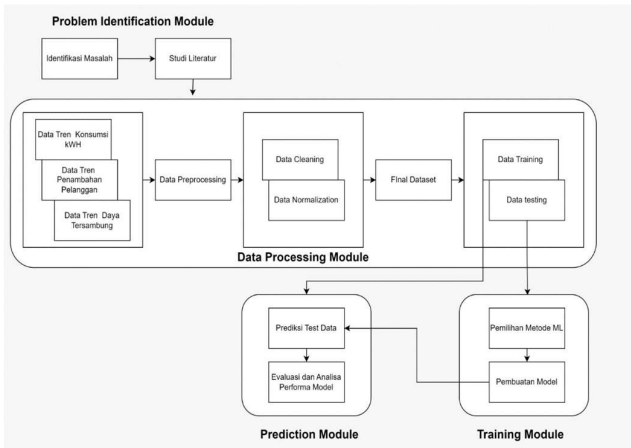


Fig 5 Flowchart of Research Methods

#### A. Data Collection

The data collection stage is the main stage in the research process, which involves the acquisition of information or data needed to answer research questions or test proposed hypotheses. This study uses data for the years 2019 - 2023 collected from available data sources, including historical customer delta data, electricity sales trends, and total connected power sourced from PT PLN (Persero) 309 centralized sales report and monthly weather data from the Palangkaraya BMKG station.



LAPORAN PENJUALAN TEMAGA LISTRIK VERSI PUSAT TOTAL  
BULAN : Desember 2022

PERMUKAAN / PLUS	TARIK	MULAI			JANGKA MATA			PERMUKAAN			JANGKA MATA			Jumlah dan Rata-rata			T.M	Jumlah	Rata-rata
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3			
81.1.200 VA	1.812.3	1.812.4	1.812.5	1.812.6	1.812.7	1.812.8	1.812.9	1.812.10	1.812.11	1.812.12	1.812.13	1.812.14	1.812.15	1.812.16	1.812.17	1.812.18	1.812.19	1.812.20	

Fig 6 Monthly Sales Report Dataset

#### B. Data Processing and Preprocessing

The first step in the system design was to combine data source which historical electricity consumption data from PLN Palangkaraya. This data included information on tariff, power, number of customers, electricity consumption (MWh), connected power (MVA), average temperature, and average rainfall. The data was collected from various internal sources of PLN and an external source, BMKG.

After collection, the data must be cleaned and processed to ensure quality and consistency. These steps include handling missing values, removing duplicates, and normalizing the data. In addition, the data will also be accumulated and calculated per month to ensure that each variable can be used effectively in the prediction model.

The next step is the selection of relevant features. Based on the needs analysis, features such as number of customers, connected power, average temperature, and average rainfall were selected as they have the potential to significantly influence electricity consumption. A dataset is then formed by combining these features together with the target electricity consumption variable.

#### C. Model Building

Machine Learning Linear Regression models used in this research include Polynomial Regression, Random Forest Regressor, and K-Nearest Neighbor as well as Deep Learning such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Recurrent Neural Network (RNN). These models were chosen for their ability to handle time series data and produce accurate predictions.

The dataset is divided into training and testing sets. The model is trained using the training data and validated using the testing data. Once the model is trained, the next step is model evaluation and validation using metrics such as Mean Squared Error (MSE) and coefficient of determination (R<sup>2</sup>). Models are evaluated to ensure that they can predict electricity consumption with high accuracy.

#### V. EXPERIMENTAL RESULTS

In this research test scenario, the Machine Learning approach will be implemented to predict electricity demand per tariff and power at PLN Palangkaraya. The following are the stages of the test scenario conducted in this research.

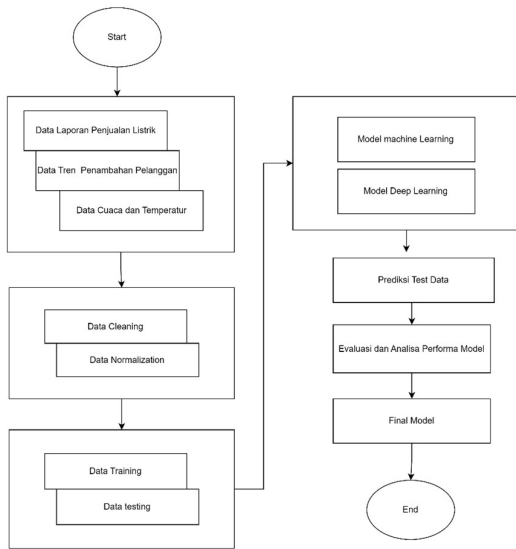


Fig 7 Flowchart of Research Scenario

### A. Test Scenario

The testing process will be carried out with systematic steps, starting from the collection of historical electricity demand data, merging historical electricity consumption datasets and historical weather datasets, data shape transformation to data cleaning to eliminate invalid data, then selecting a Machine Learning model that matches the characteristics of the data and prediction objectives. Furthermore, the data will be divided into training and testing subsets to validate the model performance.

The model training process will be carried out using Machine Learning and Deep Learning algorithms for regression needs such as linear regression, Polynomial regression, Random Forest, kNN as well as Deep Learning models such as LSTM, GRU, and RNN.

After training, the model will be tested using test data to evaluate the prediction performance in estimating electricity demand per tariff and power at PLN Palangkaraya. The test results will be further analyzed to evaluate the accuracy and reliability of the model in producing reliable predictions for electricity distribution planning and management purposes.

### B. Testing Result

The test results show that the implemented Machine Learning approach is able to provide electricity demand predictions per tariff and power with varying degrees of accuracy. Analysis of the model performance also shows that the use of sufficient historical data and the selection of appropriate algorithms contribute significantly to the quality of the resulting predictions. Based on the test results with several methods used, the model performance is as follows:

Tipe Pembelajaran	Model	Parameter	MSE	R <sup>2</sup>
Machine Learning	Linear Regression	StandardScaler(), RidgeCV()	37386.10	0.973
Machine Learning	Polynomial Regression	degree=2, include_bias=False	22148.48	0.984
Machine Learning	Random Forest Regressor	n_estimators=100, random_state=4	7717.28	0.994
Machine Learning	K-Nearest Neighbors	n_neighbors=5	9604.93	0.993
Deep Learning	LSTM	unit=5, activation='relu', input_shape=5	0.1569	-0.0602
Deep Learning	GRU	unit=5, activation='relu', input_shape=5	0.1509	-0.019
Deep Learning	RNN	unit=5, activation='relu', input_shape=5	0.1772	-0.1972

Fig 8 Testing Results Table

Based on the test scenario, visualization is carried out to display actual data and predictions given using the best model, namely Random Forest Regression with visualization of actual and predicted data as follows:

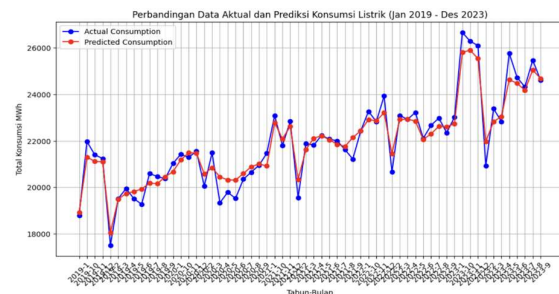
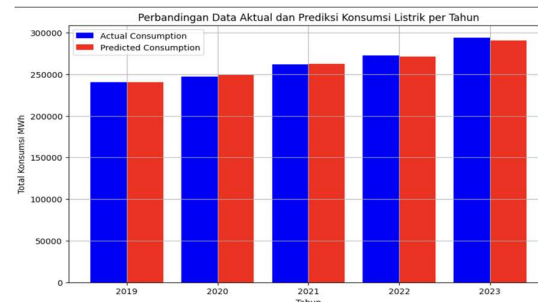


Fig 9 Trend Comparison of Actual and Predicted Yearly and Monthly Electricity Consumption Data

Based on the visualization trends in Figures 9 and 10, the performance of the selected model can provide a fairly good representation of the actual data. From the test data table and visualization results, the Machine Learning model has a better accuracy rate than the Deep Learning model. Based on the source dataset provided, ML can provide good results with less data and does not require a very large amount of data to train the model well. Whereas in Deep Learning the performance will be better if the dataset has higher complexity and many parameters are considered.



### C. Analysis and Evaluation of test results

Analysis and evaluation of test results is an important stage in understanding the performance of Machine Learning and Deep Learning approaches in predicting electricity demand per tariff and power. Evaluation of prediction accuracy is a key aspect in evaluating model performance. Comparing the predicted and actual values of electricity demand per tariff and power gives an idea of how well the model can predict correctly.

The machine learning and deep learning models have been evaluated using mean squared error (MSE) and coefficient of determination ( $R^2$ ). In this case, the deep learning model (LSTM, GRU, RNN) has a very low MSE, but ( $R^2$ ) is negative. In this context, it shows that although the prediction error (MSE) is low, the deep learning model fails to explain the variability in the data well. The best model used in this case is the Random Forest Regressor. This model shows the best combination of low MSE and high  $R^2$ , signifying a good ability to predict and explain the variability in the data.

Checking the consistency of predictions over time is crucial in the context of this research. A model that is consistent in providing predictions that are close to the actual values shows high reliability and can be relied upon for long-term planning purposes.

## VI. CONCLUSION AND FUTURE WORKS

In this research, a prediction model for electricity demand per tariff and power using machine learning and deep learning approaches was developed and evaluated, with a case study of PLN Palangkaraya. The development process involves several important stages, from data collection and processing, feature selection, machine learning model development, to system evaluation and implementation.

The machine learning models used in this study, namely Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Recurrent Neural Network (RNN), have shown good ability in predicting electricity consumption based on historical data. Although each model has its own advantages and disadvantages, the evaluation results show that the GRU model provides the best performance with the lowest Mean Squared Error (MSE) value and the highest coefficient of determination ( $R^2$ ) compared to other models.

Features such as the number of customers, connected power, average temperature, and average rainfall have a significant influence on the prediction of electricity consumption. Proper feature selection is essential to improve the accuracy of the prediction model. Data normalization and aggregation processes also play an important role in ensuring the quality of data used for model training.

Based on those conclusions, as an evaluation for future research it is important to identify the limitations of the approach used and possible improvements to this research. For example, using a variety of appropriate hyperparameters and Grid Search can provide more alternative prediction results and the presence of unexpected external factors or rapid changes in trends can affect the performance of the model.

## REFERENCES

- [1] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," 1959.
- [2] R. Jufro Bangalangi and M. Tumba, "PENERAPAN AI (ARTIFICIAL INTELEGENCE) SEBAGAI ASISTEN PEMBELAJARAN DALAM MENINGKATKAN HASIL BELAJAR MAHASISWA TEKNOLOGI PENDIDIKAN UNIVERSITAS KRISTEN INDONESIA TORAJA," 2022.
- [3] Q. Deng, D. Ramsköld, B. Reinius, and R. Sandberg, "Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells," *Science (1979)*, vol. 343, no. 6167, pp. 193–196, Jan. 2014, doi: 10.1126/SCIENCE.1245316/SUPPL\_FILE/DENG-SM.PDF.
- [4] J. Moral-Carcedo and J. Vicéns-Otero, "Modelling the non-linear response of Spanish electricity demand to temperature variations," *Energy Econ*, vol. 27, no. 3, pp. 477–494, May 2005, doi: 10.1016/J.ENECO.2005.01.003.
- [5] S.-Y. Han, S.-H. Yoo, and S.-J. Kwak, "The role of the four electric power sectors in the Korean national economy: an input-output analysis," *Energy Policy*, vol. 32, no. 13, pp. 1531–1543, 2004, Accessed: Jun. 09, 2024. [Online]. Available: <https://ideas.repec.org/a/eee/enepol/v32y2004i13p1531-1543.html>
- [6] M. Z. Harisandy and D. Purwitasari, "Forecasting Revenue from Electricity Sales of Household Customers using Various Methods," *Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022*, pp. 508–513, 2022, doi: 10.1109/ICITISEE57756.2022.10057941.
- [7] A. Thakur, K. A. Shukla, A. Choudhary, and J. Atrey, "Predictive Analysis of Energy Consumption and Electricity Demand Using Machine Learning Techniques," *International Conference on Smart Systems for Applications in Electrical Sciences, ICSSSES 2023*, 2023, doi: 10.1109/ICSSSES58299.2023.10200636.
- [8] A. Saji, A. Prakash, and M. S. Krishnan, "Electricity Demand Forecasting in Kerala Using Machine Learning Models," *2023 3rd International Conference on Intelligent Technologies, CONIT 2023*, 2023, doi: 10.1109/CONIT59222.2023.10205925.
- [9] A. Alvarez, W. Luo, and S. Fryer, "Ultra-short term wholesale electricity price forecasting through deep learning," *2021 IEEE PES Innovative Smart Grid Technologies - Asia, ISGT Asia 2021*, 2021, doi: 10.1109/ISGTASIA49270.2021.9715585.
- [10] Q. Wang, R. Lin, Y. Zhao, and H. Zou, "Electricity Consumption Forecast Based on Empirical Mode

Decomposition and Gated Recurrent Unit Hybrid Model,” *International Conference on Communication Technology Proceedings, ICCT*, vol. 2020-October, pp. 1670–1674, Oct. 2020, doi: 10.1109/ICCT50939.2020.9295955.

#### PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 10 Juni 2024

Ttd

Rifki Zamzammi 23522301