Research Paper

International Journal of Research and Review
Volume 11; Issue: 6; June 2024
Website: www.ijrrjournal.com
E-ISSN: 2349-9788; P-ISSN: 2454-2237

Utilizing Machine Learning Techniques for Power Estimation in Residential Electricity Consumption

Fadhillah Azmi[1], Amir Saleh[2]

[1]Department of Electrical Engineering, Universitas Medan Area, Medan, Indonesia.
[2]Department of Informatics Engineering, Universitas Prima Indonesia, Medan, Indonesia.

Corresponding Author: Fadhillah Azmi

DOI: https://doi.org/10.52403/ijrr.20240637

ABSTRACT

Electricity consumption analysis and estimation play a vital role in various contexts, ranging from business planning to household financial management. In an era where electrical energy is a primary source of human activity, understanding power usage patterns and estimating them becomes crucial. This research enables an understanding of energy consumption from individual to national scales and supports efficient energy policy planning. Power estimation provides insights into the expenditures that consumers or companies will incur, enabling effective financial planning. Research and development of accurate predictive models using machine learning techniques are essential to providing useful information for the energy industry and the general public. This study uses machine learning to compare the performance of different power prediction models. The decision tree (DT) method, with an MAE of 0.3613, an MSE of 0.4184, and an RMSE of 0.6469, gives the best results. The results show that the model can provide accurate predictions of electricity consumption, potentially becoming a reliable tool in power management and financial management. Its contribution is crucial in the energy industry and daily life, providing insights needed to enhance energy efficiency and budget management. As a result, this research provides relevant and beneficial solutions to energy and financial management for society and industry.

Keywords: *Electricity consumption analysis, Estimation, Machine learning, Power management, financial management*

INTRODUCTION

The need for electricity is an important aspect of daily life because it is used to operate various electrical equipment in the household (1). As the number of electrical appliances owned increases, a good understanding of electricity consumption and costs becomes very important. This will certainly be helpful in managing household finances and reducing the environmental impact of excessive energy use (2). Electrical appliances such as lights, air conditioners, computers, refrigerators, ovens, and more all contribute to daily or monthly electricity consumption (3). Therefore, it is important for household owners to have a clear understanding of how much energy each appliance consumes and how much it will cost based on the usage of the appliance and the applicable electricity tariff.

Additionally, understanding how appliances contribute to overall electricity costs can help household owners plan their electricity use to save on their finances. Therefore, household electrical power analysis and estimation are critical in effectively meeting electricity needs (4). Estimates are basically guesses or predictions about the future occurrence of an event or events (5). In this case, estimates are critical for users to know how much they will incur based on electrical power consumption (6). Power is a measure of energy per unit time, where power provides the level of energy consumption or production, with the unit being a watt (W) (7). In Indonesia, the amount of electrical energy used by each user can be determined by PLN from a tool called a KWh (kilowatt hour) meter. A KWh (kilowatt hour) meter is a tool to measure how much electrical energy is used every hour. The use of electrical power in a building depends on usage; the more equipment used, the greater the power used, and the electricity costs that will be paid will increase.

When performing power usage predictions, long-used techniques have inherent limitations in terms of accuracy and scalability. Today, it is possible to precisely predict power usage using previous data, thanks to improved machine learning techniques (8). Various studies have conducted tests on predicting or estimating electricity loads using different machine learning methods. One of the methods used is KNN, and the results show that the proposed method can accurately predict power usage with an accuracy rate of 90.92% (8). Other research attempts to use linear regression to predict long-term electrical energy needs (case study of Lampung Province, Indonesia). This method can be used to predict electricity needs accurately, with the resulting improvements (9).

In this research, we will analyze household electricity consumption based on data collected by measuring daily electricity consumption. This study employs a machine learning method to estimate daily electrical power. This machine learning approach is very effective in making estimates based on patterns in historical data (10). This method enables us to identify complex relationships between variables that influence electrical power consumption and make more accurate predictions based on these patterns. With this approach, we hope to produce a model that is able to provide better estimates of household electrical power consumption, which in turn will support more effective energy policy planning and resource management.

Based on the previous explanation, this research aims to analyze electrical power consumption and estimate it using machine learning methods. It is hoped that the analysis that will be carried out will provide invaluable insight to household owners in managing their daily electricity usage, as well as help them save electricity costs effectively. Therefore, it is hoped that this research can provide a strong foundation for the application of data analysis practices to managing household electricity consumption. Thus, this research will make a significant contribution to increasing understanding and efficiency in electricity use, as well as promoting sustainable practices for reducing excessive energy consumption.

METHODS

The analysis and estimation of household electrical power usage in this research were carried out using the machine learning method, and the data processing was carried out using the Python application. Figure 1 depicts the general steps for conducting the analysis in this study.

A. Identify household electrical equipment

The data collected includes information about electricity consumption from various household appliances, such as TVs, washing machines, fans, and other kitchen equipment. This data may include measurements of electrical power, usage time, and other characteristics related to energy consumption. The variables measured involve parameters related to electricity usage, such as electrical power (in watts), duration of use (in hours), and applicable electricity tariffs. Data may be collected at specific time intervals, for example, hourly, daily, or monthly, depending on the complexity and objectives of the research.

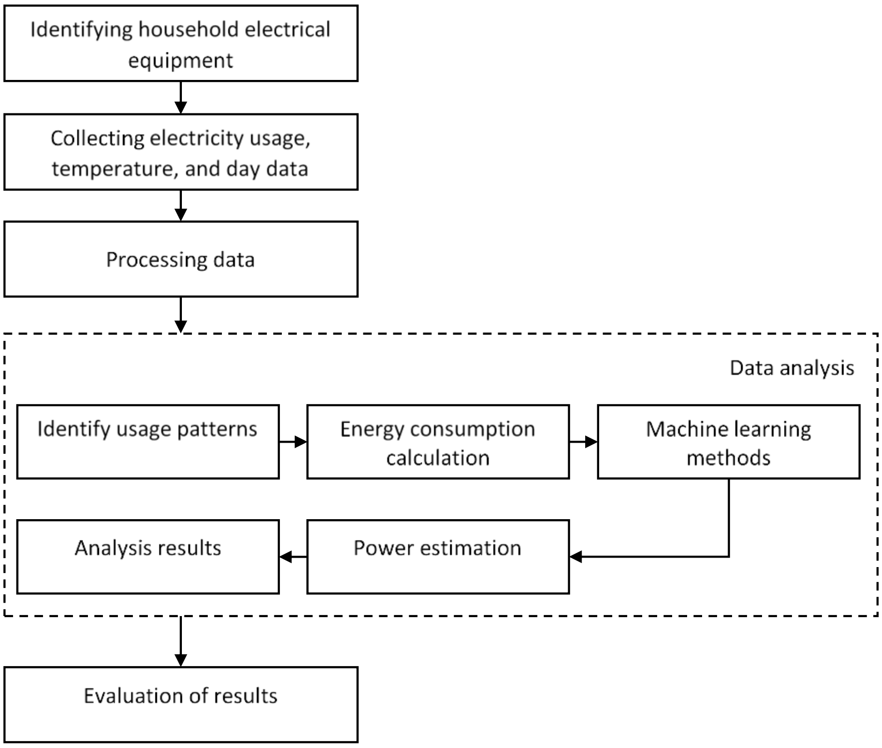

Figure 1. Electrical Power Estimation Using Machine Learning Methods

The research method in Figure 1 can be divided into 4 steps which can be described as follows:

B. Collect electricity usage

The data in this research was obtained by directly collecting data on electrical power consumption every day. Daily data collection on electrical power consumption was carried out in October 2023–February 2024, with a total of 134 data points. This data was obtained from each use of electrical equipment at the research site, and the total amount of power used was calculated to obtain the number of watts, or kWh. Table 1 contains data on daily electric power consumption (in kilowatt-hours, kWh) for analysis and estimation of electric power consumption.

Table 1. Daily Electricity Usage Data

No.	Date	Day	Temperature (^{0}C)	Power Electrical (kWh)
1	1-Oct-23	Sunday	33	9.55
2	2-Oct-23	Monday	34	15.39
3	3-Oct-23	Tuesday	34	14.87
4	4-Oct-23	Wednesday	31	12.58
5	5-Oct-23	Thursday	31	12.55
…	…	…	…	….
…	…	…	…	….
…	…	…	…	….
132	9-Feb-24	Friday	28	9.56
133	10-Feb-24	Saturday	27	8.45
134	11-Feb-24	Sunday	26	7.32

In this research, the features used to predict electrical energy consumption are days, average temperature, and daily power usage. The “day” feature indicates the day of the week, expressed in numbers 0 to 6, representing Sunday to Saturday. The “average temperature” feature displays the daily average air temperature in degrees Celsius. Meanwhile, the “daily power usage” feature displays the amount of electrical energy used by the building in a single day. These features are expected to provide enough information to make accurate predictions about building energy consumption at the initial design stage.

C. Processing data with machine learning

Machine learning (ML) is a branch of artificial intelligence that allows systems to learn from data (11). The main goal is to develop models or algorithms that are able to learn patterns from existing data and use these patterns to make predictions or decisions. This research will use machine learning methods to predict daily electrical power. The description of the method used is as follows:

Linier Regression (LR)

This method is used to find the optimal solution directly, without iterating. This method is suitable for problems with relatively small datasets. Equation 1 below shows the prediction results used (12).

$$\hat{Y} = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \ldots + \theta_n X_n \qquad 1$$

where,

$\hat{Y}$: Predicted value

$\theta_0, \theta_1, \ldots, \theta_n$: Model coefficients

$X_1, X_2, \ldots, X_n$: Input features

Ridge Regression (RR)

Ridge regression is a linear regression method that adds an L2 regularization penalty to the objective function to prevent overfitting. This method is useful when there is multicollinearity in the data. Equation 2 below shows the prediction results used (13).

$$\hat{Y} = \theta_0 + \theta_1 X_1 + \theta_2 X_2 + \ldots + \theta_n X_n + \alpha \sum_{i=1}^{n} \theta_i^2 \qquad 2$$

where,

$\hat{Y}$: Predicted value

$\theta_0, \theta_1, \ldots, \theta_n$: Model coefficients

$X_1, X_2, \ldots, X_n$: Input features

α : Regularization parameter

Decision Tree (DT)

Decision tree methods are prediction models that use a tree structure to make decisions based on the features in the dataset. These features infer rules from which decisions are made.

Random Forest (RF)

Random forest is an ensemble learning method that uses a large number of decision trees to make predictions. Each decision tree is generated independently, and finally, the results are combined to produce more accurate predictions. Predictions are based on the majority of prediction results from each decision tree.

Support Vector Machine (SVM)

SVM is a learning method used for both classification and regression. The main goal is to find the best hyperplane that optimally separates the data into different classes. The prediction results used can be seen in equation 3 below (14).

$$\hat{Y} = sign(\sum_{i=1}^{n} \alpha_i y_i K(X_i, X) + b) \qquad 3$$

where,

α_i : Weight of support vector

y_i : Class label of support vector

$K(X_i, X)$: Kernel function measuring proximity Between two data points

b : Bias

D. Performance Evaluation

Evaluate the research results by comparing the performance of the various methods or models used. Some common evaluation metrics used in research are mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) (15). This metric helps measure how well a model or method can predict the true value.

MAE (Mean Absolute Error)

MAE measures the average of the absolute differences between predicted values and actual values. The lower the MAE value, the better the model is at making predictions. Equation 4 below can be used to determine the MAE value.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \qquad 4$$

where,

n : Number of samples

y_i : True value

$\hat{y}_i$: Predicted value

MSE (Mean Squared Error)

The MSE measures the average of the squares of the differences between predicted values and actual values. The lower the MSE value, the more accurate the model. The following equation 5 can be used to determine the MSE value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad 5$$

where,

n : Number of samples

y_i : True value

$\hat{y}_i$: Predicted value

RMSE (Root Mean Squared Error)

RMSE is the square root of MSE and gives an idea of how close the predicted average is to the actual value. The lower the RMSE value, the better the models performance. Equation 6 can be utilized to determine the RMSE value.

$$RMSE = \sqrt{MSE} \quad 6$$

RESULT & DISCUSSION

In this study, we will evaluate various machine learning methods to estimate power consumption based on the day, temperature, and input electrical power. For the training stage, the 134 data points will be used as historical data. After the data collection stage is carried out, machine learning methods will be used to obtain a prediction model. Machine learning methods are used, such as linear regression, ridge regression, decision trees, random forests, and support vector machines, to identify historical data patterns of electricity usage. After the model is formed, the final step is to predict electricity consumption for the next 60 days using day and temperature parameters. With this technique, it is hoped that an accurate model can be obtained to help understand household electrical energy consumption patterns.

The results of estimating electrical power for 60 days using various machine learning methods such as linear regression, ridge regression, decision trees, random forests, and support vector machines can be seen in Figure 2 below, based on the tests carried out:

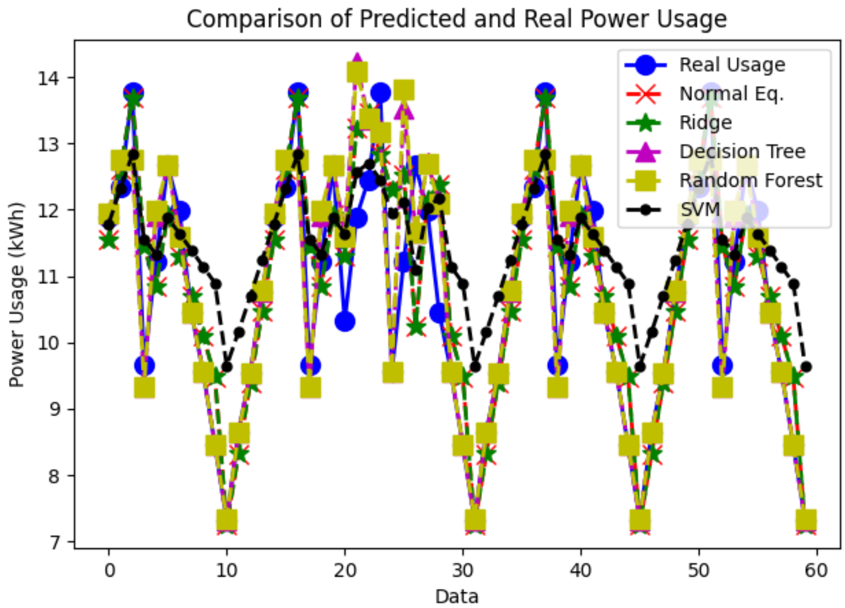

Figure 2. Electric Power Prediction Results Using Machine Learning Methods

In Figure 2, it can be seen that the decision trees (DT) method has a line that is closest to or tangential to the real power line obtained from direct measurements. Apart from looking at the graphic results, we will also look at several other test parameters to find out the best machine learning performance. Table 2 below compares machine learning performance results for estimating electrical power based on the data used.

Table 2. Performance Test Results for Estimating Electrical Power Using Machine Learning Methods

Methods	MAE	MSE	RMSE
Linier Regression (LR)	0.6501	0.7916	0.8897
Ridge Regression (RR)	0.6493	0.7911	0.8894
Decision Tree (DT)	0.3613	0.4184	0.6469
Random Forest (RF)	0.3784	0.4438	0.6662
Support Vector Machine (SVM)	1.0719	1.7770	1.3330

Prediction results from various machine learning methods have been evaluated using three evaluation metrics, namely mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). The LR and RR methods show similar results, with MAE around 0.65 kWh and RMSE around 0.89 kWh. In addition, the DT method provides better results with an MAE of around 0.36 kWh and an RMSE of around 0.65 kWh, indicating its ability to capture data patterns more effectively. Although the RF method provides slightly lower results than the DT method, it still performs adequately, with an MAE of around 0.38 kWh and an RMSE of around 0.67 kWh. On the other hand, the SVM method shows the lowest performance, with MAE around 1.07 kWh and RMSE around 1.33 kWh. Thus, the results of this evaluation show that the DT method is the best choice for predicting power usage in this scenario, followed by RF and RR, while SVM shows the lowest performance.

Evaluation results using MAE, MSE, and RMSE metrics provide a clear picture of the performance of various machine learning methods in predicting power usage. MAE is an average measure of the absolute difference between predicted values and actual values. The lower the MAE value, the better the models performance in making accurate predictions. In this context, the DT method shows the lowest MAE value, indicating its ability to produce predictions that are close to the actual value with an average difference of around 0.36 kWh. Meanwhile, for the LR and RR methods, even though they have slightly higher MAE values (around 0.65 kWh), they still provide relatively accurate predictions.

Next, MSE measures the average of the squared differences between the predicted value and the actual value. The lower the MSE value, the better the model is at reducing overall prediction error. In this case, the MSE results for the DT and RR methods are quite low, indicating their ability to minimize prediction errors. Although RF provides slightly higher results, it is still quite good. Meanwhile, SVM shows the highest MSE value, indicating that this model has a larger prediction error compared to the others.

Finally, RMSE is the square root of MSE, which gives an idea of the magnitude of the prediction error in the same units as the predicted variable. In this case, the RMSE results for the DT method are quite low, indicating a relatively small error rate in the predictions. The RR and RF methods also show good performance, with almost the same RMSE values. However, SVM again shows the worst results, with a much higher RMSE than the others, indicating that this model tends to provide predictions that are far from the true value.

Overall, this analysis shows that the DT method provides the most accurate predictions in the context of power usage prediction. Although the LR and RR methods provide good results, the DT method's ability to capture data patterns makes it a superior choice. The RF method also provides adequate performance, while SVM shows the lowest performance in this case. Therefore, selecting the right machine learning method is critical to obtaining accurate and reliable predictions for power usage.

Utilizing machine learning techniques for power estimation in household electricity consumption has a variety of benefits. First and foremost, this technique can make power consumption estimation more accurate. This accuracy is important for effective energy management and optimal resource allocation. Furthermore, precise estimates enable power companies to plan better and allocate resources more efficiently. They can optimize infrastructure and distribution networks based on predicted consumption patterns, which in turn can result in cost savings. These efficiencies can be passed on to consumers through lower rates or improved service. Additionally, accurate estimates enable consumers to identify areas of high consumption and adopt appropriate energy-saving measures, contributing to overall energy conservation efforts. Machine learning models can also help in predictive maintenance, predicting potential problems in electrical systems based on consumption patterns.

CONCLUSION

In comparing several machine learning methods for predicting daily electrical power usage, it was found that the Decision Tree (DT) technique had the best performance with MAE values of 0.3613, MSE of 0.4184, and RMSE of 0.6469. This shows that the DT method is able to produce predictions that are closest to the true value of all the methods evaluated. Meanwhile, the linear regression (LR) and ridge regression (RR) methods also provide good results, although slightly lower than the DT method. On the other hand, the Random Forest (RF) method also provides quite good results with relatively low prediction errors. However, the SVM method shows the lowest performance in this case, with higher prediction errors compared to other methods. Thus, the results of this research provide a strong basis for applying machine learning to predict daily electrical power usage, with Decision Tree being the most optimal choice based on the evaluation carried out.

Declaration by Authors

Acknowledgement: None

Source of Funding: None

Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

1. Novianto D, Donny M, Munawir M, Sekartaji D. Impact of lifestyle changes on home energy consumption during pandemic COVID-19 in Indonesia. Sustain Cities Soc. 2022;(January).
2. Qureshi M, Arbab MA, Rehman S ur. Deep learning-based forecasting of electricity consumption. Sci Rep [Internet]. 2024;14(1):1–11. Available from: https://doi.org/10.1038/s41598-024-56602-4
3. Napis, Farhan M, Rahmatulloh, Hakim AR, Apriyanto MT. Meningkatkan Kesadaran Masyarakat dalam Budaya Hemat Energi Melalui Penyuluhan Efisiensi Penggunaan Listrik Rumah Tangga. J Pendidik dan Pengabdi Masy [Internet]. 2023;6(2):107–15. Available from: https://doi.org/10.29303/jppm.v6i2.4980
4. Rahman MA, Al Masud TMM, Biswas P. Prediction of Electric Energy Consumption using Recurrent Neural Networks. Int J Smartcare Home. 2021;1(1):23–34.
5. Arungpadang TAR, Hontong FA, Tarigan L, Teknik J, Fakultas M, Universitas T, et al. Analisis Kebutuhan Energi Listrik dengan Jaringan Syaraf Tiruan. J Tekno Mesin. 2021;4(2):84–9.
6. Chung J, Jang B. Accurate prediction of electricity consumption using a hybrid CNN-LSTM model based on multivariable data. PLoS One [Internet]. 2022;17(11 November):1–16. Available from: http://dx.doi.org/10.1371/journal.pone.0278071
7. Adiwiranto MN, Waluyo BC, Sudibya B. Prototipe Sistem Monitoring Konsumsi Energi Listrik Serta Estimasi Biaya Pada Peralatan Rumah Tangga Berbasis Internet of Things. ELECTRON J Ilm Tek Elektro. 2021;2(2):13–22.
8. Bhandarkar M, Suryawanshi A, Deoghare S, Isaac N, Hegu A, Wankhede P. Electricity Billing and Consumption Prediction using Machine Learning. 2023 7th Int Conf Comput Commun Control Autom ICCUBEA 2023. 2023;01048:1–9.

9. Syafruddin M, Hakim L, Despa D. Metode Regresi Linier Untuk Prediksi Kebutuhan Energi Listrik Jangka Panjang (Studi Kasus Provinsi Lampung). J Inform dan Tek Elektro Terap. 2014;2(2).
10. Seyedzadeh S, Rahimian FP, Glesk I, Roper M. Machine learning for estimation of building energy consumption and performance: a review. Vis Eng. 2018;6(1).
11. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci [Internet]. 2021;2(3):1–21. Available from: https://doi.org/10.1007/s42979-021-00592-x
12. Theofani G, Sediyono E. Multiple Linear Regression Analysis on Factors that Influence Employees Work Motivation. SinkrOn. 2022;7(3):791–8.
13. Bastiaan RM, Salaki DT, Hatidja D. Comparing the Performance of Prediction Model of Ridge and Elastic Net in Correlated Dataset. Oper Res Int Conf Ser. 2022;3(1):8–13.
14. Saikin S, Fadli S, Ashari M. Optimization of Support Vector Machine Method Using Feature Selection to Improve Classification Results. JISA(Jurnal Inform dan Sains). 2021;4(1):22–7.
15. Wang W, Lu Y. Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. IOP Conf Ser Mater Sci Eng. 2018;324(1).

How to cite this article: Fadhillah Azmi, Amir Saleh. Utilizing machine learning techniques for power estimation in residential electricity consumption. *International Journal of Research and Review*. 2024; 11(6): 326-333. DOI: *https://doi.org/10.52403/ijrr.20240637*

******