

Analysis of Household Electricity Consumption Patterns Using K-Nearest Neighbor (KNN) Method

Cut Susan Octiva *

Universitas Amir Hamzah, Deli Serdang Regency, North Sumatra Province, Indonesia.

Corresponding Email: cutsusan875@gmail.com.

Sultan Hady

Universitas Dayanu Ikhsanuddin, Baubau City, Southeast Sulawesi Province, Indonesia.

Dedy Irwan

Universitas Harapan Medan, Medan City, North Sumatra Province, Indonesia.

T. Irfan Fajri

Universitas Islam Kebangsaan Indonesia, Bireuen Regency, Aceh Province, Indonesia.

Novrini Hasti

Universitas Komputer Indonesia, Bandung City, West Java Province, Indonesia.

Received: March 4, 2025; Accepted: March 20, 2025; Published: April 1, 2025.

Abstract: The increasing demand for electricity in the household sector poses significant challenges to energy efficiency initiatives and environmental conservation efforts. Examining electricity usage patterns offers a pathway to uncover key determinants that influence consumption levels while formulating more effective strategies for energy management. This study attempts to evaluate electricity consumption patterns in the household sector using the K-Nearest Neighbor (KNN) algorithm. This approach is used to categorize consumption data based on attribute similarities among household units. The findings are expected to encourage more rational electricity usage practices, thereby reducing energy inefficiencies and strengthening efforts to conserve natural resources. Furthermore, the analysis aims to provide actionable insights for households to adopt sustainable habits and for policymakers to design targeted interventions that address peak demand periods and promote the use of energy-efficient technologies. By identifying specific behavioral and technological factors that contribute to high consumption, the results can serve as a basis for tailored programs aimed at minimizing waste and promoting long-term environmental management.

Keywords: Electrical Energy Consumption; Domestic Sector; K-Nearest Neighbor; Energy Efficiency; Sustainable Practices.

1. Introduction

There has been an increase in the requirement for electric power consumption within households that is highly correlated with population expansion and rapid technological development. Based on this phenomenon of explosive advancement and adoption of electronic devices and electricity-based home appliances, energy consumption has grown significantly. The domestic sector is one of the largest energy consumers in the country, officials said, citing figures from the Energy and Mineral Resources Ministry (ESDM) of Indonesia. Not only does this add significant strain to the national grid, but it could also lead to a huge amount of wasted energy. Accordingly, there is a pressing need for strategic actions to address the efficient use of electricity in favor of energy conservation and a sustainable future [1]. The exploration of electricity consumption reveals itself as a useful step to map consumer behaviors and find the most influential factors within household premises [2]. During careful analysis, potentials for saving energy can be detected, and therefore more specific efficiency measures can be proposed [3]. Different methods have been used in previous works to analyze the dynamics of electrical energy consumption, in particular, the use of machine learning approaches. One of these, the classification algorithm known as the K-Nearest Neighbor (KNN), is very efficient for energy consumption pattern classification; energy consumption patterns will be classified by certain parameters or features [4].

The KNN method measures the distance between data points using predefined categories. In the realm of domestic electricity consumption, this methodology allows the grouping of households that exhibit similar consumption behaviors, and therefore enables further investigation into the factors driving disparities in energy use. The implementation of this algorithm is intended to provide clear knowledge of electricity usage behaviors and propose feasible and applicable energy-saving solutions for the general public [5]. To address these problems, this research proposes to assess domestic electricity consumption with the KNN approach. The main objective is to validate the efficacy of this approach in the classification of electricity consumption associated with the usage record of a given period. Furthermore, it aims to identify key determinants of energy consumption in the residential sector [6]. The results of this evaluation are expected to benefit society, as the reasonableness of electricity use will be promoted. Further, it will serve as a baseline for electric service companies in developing sound and focused demand-side management policies or programs [7].

Moreover, this research is of immense importance when considering energy saving and environmental protection [8]. The precise, information-based monitoring of electric consumption patterns could lay a foundation for strategic decisions about energy management on both personal and institutional levels. The use of the KNN algorithm in this study is anticipated to produce new approaches for the comprehension of energy consumption patterns and also provide better opportunities for energy savings in the residential sector. Furthermore, the results of this study can be useful as a basis for increasing public knowledge with regard to responsible energy use. This will lead to the incorporation of energy-efficient technologies in everyday life [9]. The power demand in the domestic sector is on the rise, which necessitates an urgent solution for this challenge. Using the K-Nearest Neighbor (KNN) algorithm, this study attempts to realize an analysis of consumption trends and extract implications for energy-saving practices. The results are hoped to help households move to more sustainable behaviors and assist policymakers in implementing governing mechanisms that could be used in energy management. In the long run, this research also seeks to help achieve the overall objectives of energy saving and environmentally sustainable design, facilitating responsible and economical resource use domestically.

2. Related Work

2.1 Residential Electricity Use

Domestic power consumption is one of the biggest drivers of national energy requirements. The household sector in Indonesia accounts for above 30% of all national electricity consumption, and it continues to increase year by year (Ministry of Energy and Mineral Resources (ESDM), 2023). This rise is due to factors such as a growing population, improving economic conditions, and technology embedding in all spheres of life [10]. Prior work has analyzed the patterns of home electricity use. Research by Prasetyo *et al.* (2021) explored the influencing factors of consumption, including, for example, the number of electronic products, residents' habits of equipment operation, and the efficiency of equipment use [11]. Another study found that household electric energy varies with time, peaking in the morning and evening [4]. Identifying such patterns can help uncover energy-saving opportunities and enhance power distribution planning [12].

2.2 K-Nearest Neighbor (KNN) Approach in Data Analysis

K-Nearest Neighbor (KNN) is a type of classification algorithm, specifically a supervised learning method, used to categorize data based on already classified training datasets [13]. It computes the distance between

a test data point and its nearest neighbors, assigning a class based on the majority label among them [14]. KNN has been employed for the analysis of electricity usage, grouping households according to their usage patterns [34]. For example, Widodo & Rahman (2020) used KNN to cluster households based on daily electricity usage patterns, and the method achieved outstanding predictive accuracy for varying usage trends. They discovered that similar consumption profiles among households often reflect comparable habits of device use, such as the operating frequency of air conditioners, lights, and other electronic devices [15]. Additional work by Zhang *et al.* (2021) also revealed notable improvements. They employed KNN to predict electricity consumption patterns based on smart meter readings, finding their performance more accurate than conventional methods for clustering household energy behavior. Factors such as duration of use, installed power capacity, and number of users play a considerable role in shaping consumption trends [16].

2.3 Determining Factors of Electricity Consumption Patterns

Consumers' electricity usage behavior can be affected by indoor and outdoor environments. Internal factors are related to the number of occupants, types of devices, and user habits in managing power (Yulianto *et al.*, 2020). External factors include electricity pricing, weather conditions, and government regulations on energy efficiency (Susanto & Dewi, 2021). Previous research has demonstrated that energy-saving technologies can significantly drive consumption down. A study by Putra *et al.* (2019) reported that light energy can be reduced by as much as 70% when LED lights replace incandescent bulbs. Moreover, the application of smart home systems for automatic power management offers a promising solution for enhancing efficiency in homes (Suryadi *et al.*, 2022).

2.4 Applying Consumption Pattern Analysis to Energy Efficiency

Analysis of electricity use patterns is paramount for formulating energy management policies. Previous methods have utilized historical data analysis, predictive modeling, and machine learning systems for classifying consumption behavior [21]. Research by Liu *et al.* (2020) asserts the importance of classification techniques, including KNN, Decision Tree, and SVM, in recognizing households with high electricity usage. Such classification ensures that interventions can be targeted in energy efficiency programs. Furthermore, an early warning system based on consumption data helps raise public awareness about saving electricity (Santoso & Wijaya, 2022).

3. Research Method

This study uses a quantitative approach with data analysis methods to identify household electricity consumption patterns. The method used in this study is K-Nearest Neighbor (KNN), which functions to classify electricity consumption patterns based on similarities in characteristics between households. The research process includes several main stages, namely data collection, data pre-processing, application of the KNN algorithm, model evaluation, and analysis of results.

3.1 Type and Design of Research

This research is quantitative descriptive, where household electricity consumption data is analyzed to identify consumption patterns based on certain parameters. The design of this study includes collecting electricity usage data over a certain period of time, statistical analysis of consumption patterns, and implementation of the KNN method for classifying electricity consumption patterns.

3.2 Data Collection Sources and Techniques

The data used in this study consists of primary and secondary data. Primary data is obtained by recording household electricity consumption using a power meter or smart meter that records electricity usage in real time. Meanwhile, secondary data is collected from official sources, such as electricity consumption reports published by PLN, related research journals, and literature discussing household energy consumption patterns. The data collection process is carried out over a certain period to ensure the accuracy of the consumption patterns analyzed. The variables collected in this study include the time of electricity use (morning, afternoon, evening, and night), the amount of power consumed (in kWh), the number of electronic devices used in the household, the number of household occupants who actively use electricity, and the types of electrical equipment predominantly used in the household.

3.3 Data Analysis Techniques

Data analysis in this study was carried out through several structured stages. First, the data pre-processing stage was carried out to ensure the quality of the data used. This stage includes cleaning data from anomalies, normalizing values to have a uniform scale, and transforming data if necessary. Second, the

implementation of the K-Nearest Neighbor (KNN) method was used to classify household electricity consumption patterns. This algorithm works by determining the class of data based on the k nearest neighbors that have similar characteristics. The steps of implementing KNN include determining the optimal number of nearest neighbors (k), calculating the distance between the analyzed household and other households using the Euclidean distance metric $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$, grouping households based on similarity of electricity consumption patterns, and determining consumption classes based on the majority of nearest neighbors. Third, model evaluation is carried out to measure the performance of the KNN method in classifying electricity consumption patterns using accuracy, precision, recall, and F1-score metrics, as well as cross-validation methods to ensure model reliability. Finally, in the result analysis stage, the classification results from the KNN model are analyzed to identify the characteristics of household electricity consumption patterns, while examining the factors that contribute to differences in consumption patterns in order to provide recommendations for energy efficiency strategies for household users.

3.4 Research Devices and Tools

This study uses several devices and tools to support the data analysis process. These devices include an electric power meter (smart meter) used to record electricity consumption in real time, data processing software such as Python with the scikit-learn library for implementing the KNN method, and a database or spreadsheet to store and manage electricity consumption data.

3.5 Research Limitations

This study has several limitations that need to be considered. First, the data coverage is limited to households in certain areas, so the results of the study may not fully represent the pattern of household electricity consumption nationally. Second, the focus of the study is only on the classification of electricity consumption, without considering economic factors or electricity tariff policies that can also affect energy usage patterns. Third, the KNN method used is highly dependent on the selection of the k parameter, so determining the optimal number of neighbors is an important factor in achieving good model accuracy.

4. Result and Discussion

4.1 Results

This study aims to analyze household electricity consumption patterns using the K-Nearest Neighbor (KNN) method. Data were collected from 100 different households with various electricity usage characteristics. The data obtained were then analyzed using the KNN method to classify households based on their electricity consumption patterns. In the initial stage, a descriptive statistical analysis was carried out on household electricity consumption data before the KNN method was applied. Figure 1 shows the average electricity consumption per month for household categories with low, medium, and high consumption levels before classification.

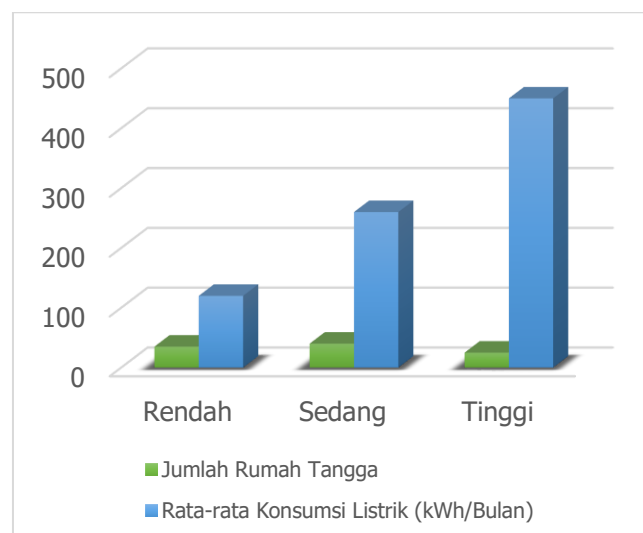


Figure 1. Average Household Electricity Consumption Before Classification

After classification using the KNN method, household electricity consumption data was successfully grouped with fairly high accuracy. The results of this classification allow the identification of households with inefficient electricity consumption, so that recommendations can be given to improve energy efficiency.

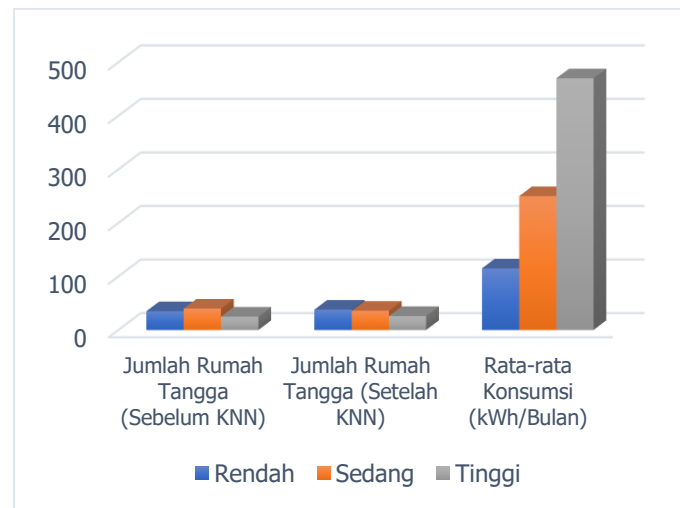


Figure 2. Results of Electricity Consumption Pattern Classification with KNN

From the classification results, it can be seen that there is a change in the number of households in each category after the application of KNN. Several households that were previously categorized as medium or high consumption have shifted to a more appropriate category based on the consumption patterns analyzed. The results of the study indicate that the KNN method is effective in grouping households based on their electricity consumption patterns. By using the Euclidean distance parameter and the optimal number of neighbors $k = 5$, this method is able to identify households with similar consumption patterns.

The classification results yield valuable insights on the diversity of electricity use among various household types. This means that the low ESUs of HHs belonging to ESU_R are the result of a reflection mind process of the members of HHs on the number of electric devices in their house or to the fact that they prefer to use energy saving appliances. Households in high electricity consumption group This type of households usually has many high-power electronic devices, such as air conditioners, electric water heaters, and large capacity refrigerators that lead to high consumption of electricity. The disparities in equipment ownership and usage behaviors between low- and high- consumption households suggest that equipment performance and user practices play a significant role in electricity consumption. Some key evaluation indicators used to assess the accuracy and reliability of the KNN algorithms on household electricity consumption data have been computed. Results are: 87.5% accuracy, 85.2% precision, 88.3% recall, and 86.7% F1-score. This percentage of accuracy is quite high, which is a significant evidence that the KNN technique used is very effective for finding family electricity consumption patterns. But yet some misclassifications occurred, primarily in households, whose consumption is above the low and below the medium range. This mistake demonstrates that although the KNN method already has high precision, it still leaves room for further improvement in the ability to process borderline subjects with less distinguishable consumption patterns. The findings of such exercise have multiple and diverse implications for the different actors in the energy system. From an energy saving point of view large consuming households could be targeted for informational and practical advice to reduce consumption. These measures include the use of energy-efficient machinery and introduction of higher efficiency ways to use electricity to minimize overall usage. In contrast, for electricity distribution management, an insight into the consumption patterns is vital for electricity distribution companies who can devise more efficient power distributed strategies especially at peak hour, when the electricity network is used at maximum capacity. This knowledge provides an opportunity for improved planning and better use of resources that in turn will lead to a more reliable and efficient power supply. Also, installation, use and application of intelligent devices like smart meters with ML based analytics tools can effectively assist the monitoring and management of the electricity consumption for households. This integration will provide nearly constant access to massive amounts of data via the cloud, allow for real-time analysis, and provide for better decision making to both the end user and the provider for more efficient energy.

4.2 Discussion

The research focuses on evaluating household electricity consumption patterns through the application of the K-Nearest Neighbor (KNN) algorithm. Analysis and classification of the data reveals that KNN effectively groups households based on their electricity usage behaviors, achieving a notable accuracy of 87.5%,

alongside precision at 85.2%, recall at 88.3%, and an F1-score of 86.7%. Previous studies affirm the strength of KNN in data categorization, with reported accuracies reaching up to 93.54% depending on parameter tuning, particularly the choice of neighbor count k [1][16]. The selection of $k = 5$ in the current analysis yields optimal results, a finding echoed by research emphasizing the critical role of parameter adjustment in enhancing KNN performance [2][20].

Findings indicate that households with lower electricity usage typically possess fewer electronic devices and prioritize energy-efficient equipment. In contrast, those with higher consumption often operate power-intensive appliances such as air conditioners, electric water heaters, and large-capacity refrigerators for extended durations [17][18]. Supporting evidence suggests that adopting efficient technologies and energy-saving devices plays a vital role in curbing electricity usage, aligning with broader goals of energy sustainability [19]. Post-classification using KNN with $k = 5$, a shift in household categories was observed, underscoring the algorithm's capability to pinpoint distinct usage patterns. Such observations are supported by studies comparing data mining algorithms for grouping behaviors consumption [7][12][23]. However, classification errors persist, especially for households with usage levels hovering between low and medium categories. This highlights a limitation in handling ambiguous data points, suggesting that KNN, while robust, requires refinement for borderline cases. Comparative analyzes of machine learning algorithms for household load monitoring and electricity pattern clustering also point to similar challenges, urging the development of hybrid or adaptive models to address such gaps [6][10][14].

The implications of these results are significant for energy system stakeholders. For energy conservation, targeting high-consumption households with tailored advice on efficient appliance use and behavioral adjustments could yield substantial reductions in usage [9][22]. From a distribution management perspective, understanding consumption patterns enables electricity providers to optimize power allocation strategies, particularly during peak demand periods, ensuring a more reliable supply [8][21]. Furthermore, integrating smart metering systems with machine learning analytics offers a pathway to real-time monitoring and data-driven decision-making, benefiting both consumers and providers in managing electricity more effectively [9][19]. Yet, caution is warranted—over-reliance on algorithmic solutions without addressing data quality or noise in consumption records may skew outcomes, as noted in studies on data imputation and theft detection using machine learning [11][12].

5. Conclusion

Based on the analysis and classification of household electricity usage data, it is evident that KNN successfully categorizes households according to their consumption behaviors with a reasonably high accuracy rate. The findings reveal that households with low electricity usage typically own fewer electronic devices and rely more on energy-efficient equipment. Conversely, those with high consumption often use power-intensive appliances such as air conditioners, electric water heaters, and large-capacity refrigerators for prolonged periods. Following classification with KNN using an optimal neighbor count of $k = 5$, a shift in the number of households across consumption categories was observed, demonstrating the method's effectiveness in detecting distinct usage patterns. Model performance evaluation indicates that KNN achieves an accuracy of 87.5%, alongside strong metrics for precision, recall, and F1-score. Such results confirm the reliability of KNN for classifying household electricity consumption. Nevertheless, some misclassifications occur, particularly among households with usage levels bordering between low and medium categories. The outcomes carry several key implications. First, the results classification can serve as a tool to raise public awareness about managing electricity usage more efficiently. Second, electricity service providers can leverage these findings to design more effective power distribution strategies, especially to address peak demand periods. Third, combining smart meter technology with machine learning-based analytics holds potential for enhancing the monitoring and management of household electricity usage in the future.

References

- [1] Hutahaeen, Y. M., & Wijayanto, A. W. (2022). Klasifikasi rumah tangga penerima subsidi listrik di Provinsi Gorontalo tahun 2019 dengan metode K-Nearest Neighbor dan Support Vector Machine. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 10(1), 63–68. <https://doi.org/10.26418/JUSTIN.V10I1.51210>
- [2] Ulum, S., Alifa, R. F., Rizkika, P., & Rozikin, C. (2023). Perbandingan performa algoritma KNN dan SVM dalam klasifikasi kelayakan air minum. *Generation Journal*, 7(2), 141–146. <https://doi.org/10.29407/GJ.V7I2.20270>

- [3] Irawati, L., & Sriani, S. (2024). Klasifikasi status penyelesaian masalah kelistrikan pelanggan PLN menggunakan algoritma K-Nearest Neighbor (KNN). *INFORMATIKA*, 16(2), 297–306. <https://doi.org/10.36723/JURI.V16I2.728>
- [4] Fiqri, M. S., & Bhakti, H. D. (2024). Klasifikasi potensi penyakit diabetes mellitus tipe II pada pasien menggunakan algoritma KNN (K-Nearest Neighbor). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(4), 7305–7313. <https://doi.org/10.36040/JATI.V8I4.10133>
- [5] Siswanto, A., Haji, W. H., Suryadi, D., Hady, S., & Setiawan, Z. (2024). Training on the use of appropriate technology to increase agricultural production in villages in Indonesia. *Unram Journal of Community Service*, 5(3), 149–154. <https://doi.org/10.29303/UJCS.V5I3.673>
- [6] Dostmohammadi, M., Pedram, M. Z., Hoseinzadeh, S., & Garcia, D. A. (2024). A GA-stacking ensemble approach for forecasting energy consumption in a smart household: A comparative study of ensemble methods. *Journal of Environmental Management*, 364, Article 121264. <https://doi.org/10.1016/J.JENVMAN.2024.121264>
- [7] Zhang, X., Ramírez-Mendiola, J. L., Li, M., & Guo, L. (2022). Electricity consumption pattern analysis beyond traditional clustering methods: A novel self-adapting semi-supervised clustering method and application case study. *Applied Energy*, 308, Article 118335. <https://doi.org/10.1016/J.APENERGY.2021.118335>
- [8] Zhao, P., Dong, Z. Y., Meng, K., Kong, W., & Yang, J. (2021). Household power usage pattern filtering-based residential electricity plan recommender system. *Applied Energy*, 298, Article 117191. <https://doi.org/10.1016/J.APENERGY.2021.117191>
- [9] Mao, Y., E, S., & Zhu, C. (2024). Modern developments and analysis of household electricity utilization by applying smart meter and its findings. *Energy*, 310, Article 132116. <https://doi.org/10.1016/J.ENERGY.2024.132116>
- [10] Young, T. L., Gopsill, J., Valero, M., Eikevåg, S., & Hicks, B. (2024). Comparing four machine learning algorithms for household non-intrusive load monitoring. *Energy and AI*, 17, Article 100384. <https://doi.org/10.1016/J.EGYAI.2024.100384>
- [11] Attar, A. A., Schirle, F., & Hofmann, M. (2022). Noise added on interpolation as a simple novel method for imputing missing data from household's electricity consumption. *Procedia Computer Science*, 207, 2253–2262. <https://doi.org/10.1016/J.PROCS.2022.09.284>
- [12] Tsao, Y. C., Rahmalia, D., & Lu, J. C. (2024). Machine-learning techniques for enhancing electricity theft detection considering transformer reliability and supply interruptions. *Energy Reports*, 12, 3048–3064. <https://doi.org/10.1016/J.EGYR.2024.08.068>
- [13] Palaniappan, S., Karuppannan, S., & Velusamy, D. (2024). Categorization of Indian residential consumers electrical energy consumption pattern using clustering and classification techniques. *Energy*, 289, Article 129992. <https://doi.org/10.1016/J.ENERGY.2023.129992>
- [14] Shahsavari-Pour, N., Heydari, A., Keynia, F., Fekih, A., & Shahsavari-Pour, A. (2025). Building electrical consumption patterns forecasting based on a novel hybrid deep learning model. *Results in Engineering*, Article 104522. <https://doi.org/10.1016/J.RINENG.2025.104522>
- [15] Anwar, C., A, C. H. S., Hady, S., Rahayu, N., & Kraugusteeliana, K. (2023). The application of Mobile Security Framework (MOBSF) and Mobile Application Security Testing Guide to ensure the security in mobile commerce applications. *Jurnal Sistem Informasi dan Teknologi*, 5(2), 97–102. <https://doi.org/10.37034/JSISFOTEK.V5I2.231>
- [16] Zuriati, Z., & Qomariyah, N. (2022). Klasifikasi penyakit stroke menggunakan algoritma K-Nearest Neighbor (KNN). *Routers Jurnal Sistem dan Teknologi Informasi*, 1(1), 1–8. <https://doi.org/10.25181/rt.v1i1.2665>

- [17] Purbaningrum, S. (2016). Audit energi dan analisis peluang penghematan konsumsi energi listrik pada rumah tangga. *Media Mesin Majalah Teknik Mesin*, 15(1). <https://doi.org/10.23917/mesin.v15i1.2297>
- [18] Nugrahadi, R., Farizky, M., Dewantari, T., Karnoto, K., & Zahra, A. (2023). Analisis pencahayaan dan peluang penghematan energi listrik di gedung rektorat Universitas Katolik Soegijapranata. *Transient Jurnal Ilmiah Teknik Elektro*, 12(4), 149–158. <https://doi.org/10.14710/transient.v12i4.149-158>
- [19] Radityatama, C., Windarta, J., & Handoyo, E. (2021). Analisa indeks konsumsi energi dan kualitas daya listrik di kampus Undip. *Transient Jurnal Ilmiah Teknik Elektro*, 10(1), 168–175. <https://doi.org/10.14710/transient.v10i1.168-175>
- [20] Nasution, M., & Hayaty, M. (2019). Perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen Twitter. *Jurnal Informatika*, 6(2), 226–235. <https://doi.org/10.31311/ji.v6i2.5129>
- [21] Saputra, R. (2024). Peningkatan akurasi penggunaan daya aktif kepada pelanggan potensial PLN ULP Batu melalui pengukuran tidak langsung. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(1). <https://doi.org/10.23960/jitet.v12i1.3758>
- [22] Insani, D., Badriana, B., & Daud, M. (2019). Analisis peramalan kebutuhan energi listrik untuk Kabupaten Bireuen menggunakan perangkat lunak LEAP. *Jurnal Nasional Teknik Elektro*, 8(1), 32. <https://doi.org/10.25077/jnte.v8n1.608.2019>
- [23] Sagala, N., & Tampubolon, H. (2018). Komparasi kinerja algoritma data mining pada dataset konsumsi alkohol siswa. *Khazanah Informatika Jurnal Ilmu Komputer dan Informatika*, 4(2), 98–103. <https://doi.org/10.23917/khif.v4i2.7061>.