Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

Utilization of Data Mining in Finding Household Electricity Consumption Patterns in Indonesia

Ferdiansyah Prayoga[1*], Najwa Aulia Meliala[2], Dhea Aulia[3], Seroja Fi Maaris Mar'ah[4], Harly Bima Perkasa Alam[5], Mhd. Furqan[6]

[1,2,3,4,5,6]*Computer Science, Universitas Islam Negeri Sumatera Utara*
ferdiansyahprayoga1133@gmail.com[1*], nazwaauliaa2004@gmail.com[2], dheaauliaa282@gmail.com[3], serojafi884@gmail.com[4], harlybima04@gmail.com[5], mfurqan@uinsu.ac.id[6]

Abstract

This study aims to identify household electricity consumption patterns in Indonesia based on usage time, namely day and night, using data mining techniques, especially the K-Means clustering method. The data used consists of daily electricity consumption in kWh units during the day and night periods. The analysis process begins with data pre-processing to ensure the quality and uniformity of the data scale, then continues with grouping the data into three clusters that represent different electricity consumption characteristics. The clustering results show three main patterns: even consumption between day and night, high consumption at night, and high consumption during the day. These findings reveal the diversity of electricity usage behavior influenced by socio-economic factors and daily habits. By understanding these patterns, the government and electricity providers can design more targeted energy efficiency policies, including the development of time-based electricity tariffs and energy-saving education programs. This approach supports more effective and sustainable energy management in the household sector in Indonesia.

Keywords: *Data Mining, Electricity Consumption, Households, Clustering*

1. Introduction

Household electricity consumption is the use of electrical energy by individuals or families in their daily activities at home. Household electricity consumption includes the use of electronic equipment such as lights, televisions, refrigerators, air conditioners, and other household appliances. Household electricity consumption can vary depending on factors such as the number of occupants, energy usage habits, and the type of equipment used. Understanding household electricity consumption patterns can help improve energy efficiency and resource management [1]. In today's era of globalization, information is an important element in everyday life. With the development of technology, especially the internet, access to information is becoming easier and faster [2]. From the mining process, data transformation can be carried out for the process of predicting electricity needs. [3]. This model allows the identification of the main factors that influence electricity needs and helps predict future needs more accurately. As one of the largest archipelagic countries in the world, Indonesia has its own challenges in meeting the need for electricity for its population, which numbers hundreds of millions of people [4].

The projection of electricity use in Indonesia, which has a population of 268,583,016, is expected to continue to be dominated by the household sector until 2050. The increase in energy consumption in this sector was triggered by the Covid-19 pandemic and large-scale social restrictions in Indonesia, which caused a decrease in energy consumption in the social, business, and industrial sectors. This was also felt by the community in the Bulan Terang Utama Housing Complex, where this situation encouraged sustainable consumer behavior, so that unknowingly it caused energy use, especially electricity, to become less efficient[5]. In essence, this Energy Consumption Intensity is the result of dividing electricity consumption over a certain period of time by the unit area of the building. The IKE unit is KWh/m² per period. The resulting IKE value will determine whether a building is classified as very efficient, efficient, fairly efficient, and wasteful[6]. This IKE value can also be used as a reference in implementing energy efficiency strategies and as an indicator of building energy performance over time.

2. Literature Review

Data mining is also a method used in large-scale data processing, therefore data mining has a very important role in several areas of life including industry, finance, weather, science and technology. In data mining there are also methods that can be used such as classification, clustering, regression, variable selection, and market basket analysis [7].
Electricity is a vital energy for the sustainability of human activities for individuals, community groups and the industrial world. In other words, electrical energy can be used to carry out activities with very great benefits where various equipment to meet life's needs is operated using electrical energy. Data mining is the process of finding interesting patterns or information in selected data using certain techniques or methods. Techniques, methods, or algorithms in data mining vary widely [8].

Human activity in the use of electricity has increased over time. This is because electrical energy has become an important part of the development of human civilization in various fields including economics, technology, social and human culture. The strategy for forecasting electrical energy needs is very much needed. The community's need for electrical energy continues to grow every year. In addition to population growth, economic growth in a region is believed to be one of the factors that influences the increasing consumption of electricity in the region [9]. Data mining is a field of science that studies data processing, by looking for patterns of relationships and summarizing data using certain methods to find new patterns and information. Decision trees are one method or model of data mining that produces a pattern or rule in the form of a decision tree [10].

According to Wahyudi, the main characteristics and objectives of data mining:

1. Data is often buried in very large databases, which sometimes contain data for years. In many cases, data is cleaned and combined into a data warehouse.
2. The data mining environment is generally a client-server architecture or a web-based information system architecture.
3. Various new sophisticated tools, including various sophisticated visualization tools, help to lift the seeds of information buried in corporate files or archive records. To get it will involve polishing and synchronizing data to get the right results [11].

3. Research Method

The research method used in this study is a quantitative method with a data mining approach, especially clustering techniques using the K-Means algorithm [12]. This study utilizes household electricity consumption data consisting of two main variables, namely electricity consumption during the day and at night (in kWh units) [13]. The analysis process begins with the data preprocessing stage, which includes checking the data to avoid missing values and normalization so that the data scale is uniform and can be analyzed optimally [14].

The first step begins with the data preprocessing process to ensure that there are no missing values and that the data is on the appropriate scale. Furthermore, the number of clusters (in this case three clusters) is selected which are considered the most representative in mapping consumption patterns [15]. The K-Means algorithm then groups the data into three groups, which are visualized in the form of a two-dimensional scatter plot, where the x-axis represents daytime consumption and the y-axis nighttime consumption. The colors red, blue, and green indicate different clusters. This approach allows researchers to identify the characteristics of household electricity consumption in a more segmented manner, so that it can be used as a basis for formulating more targeted energy efficiency policies.

With this method, sharper insights are obtained regarding the electricity consumption habits of the community, which can be used by related parties, such as the government and electricity providers, to develop more effective energy management strategies. In addition, the results of this clustering can also be used to develop a fairer electricity tariff structure and design energy-saving education programs

4. Results and Discussion

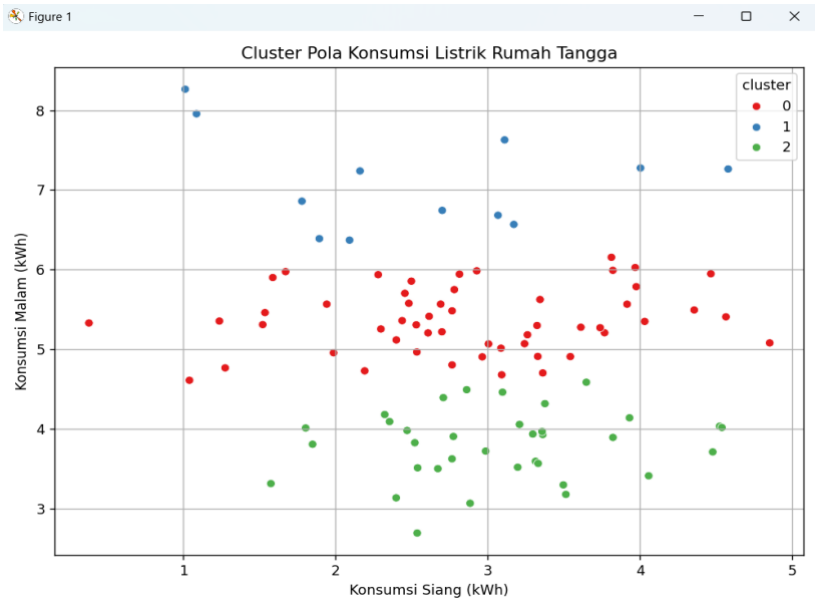

Fig. 1: Visualization of Clustering Results of Household Electricity Consumption Patterns Based on Usage Time

The results of this study aim to identify the characteristics of household electricity consumption in Indonesia based on the time of use, especially between the day and night periods, using the K-Means Clustering analysis approach. Based on the results of the clustering visualization in the image above, the data is divided into three main groups: cluster 0 (red), cluster 1 (blue), and cluster 2 (green). Each point on the graph represents one household unit, with the X-axis coordinates indicating the amount of electricity consumption during the day and the Y-axis indicating consumption at night, both in kilowatt-hours (kWh). Cluster 0 (red) dominates the number of households and describes a relatively even energy consumption pattern between day and night, with usage levels in the moderate range. Meanwhile, cluster 1 (blue) represents a group of households that have high electricity consumption at night but low during the day. This pattern likely reflects the habits of residents who are more active at night, or the use of electronic devices that are predominantly used after sunset. On the other hand, cluster 2 (green) indicates a group of households with low nighttime electricity usage, but their daytime consumption tends to be high, which can be associated with dense domestic activities during the day or the use of electrical equipment such as air conditioners and washing machines at that time.

This difference in consumption patterns indicates the diversity of energy behavior among households that can be influenced by various factors, such as socio-economic background, daily habits, working hours of residents, and the type and efficiency of electrical equipment used. The findings from this cluster analysis are very relevant as an initial basis for designing evidence-based energy management policies, including the development of smart grid systems and the implementation of dynamic electricity tariffs that are adjusted to time-use consumption patterns. This approach can increase the efficiency of energy use nationally and provide incentives for the community to use electricity more wisely according to their needs.

The clustering analysis results show significant differences in electricity usage habits between household groups. Cluster 1 (blue) shows a dominance of electricity usage at night. This most likely reflects households with family members who are more active at night, such as those who work outside the home during the day and only return to using electronic devices after sunset. Meanwhile, Cluster 2 (green) describes a group of households that consume electricity intensively during the day but low at night. This pattern can be associated with household activities that are more concentrated during the day, or a tendency to adopt energy-saving habits at night, for example by limiting the use of air conditioners or other electronic devices while sleeping. In contrast, Cluster 0 (red) represents a group of households with relatively stable consumption patterns between day and night, or is at a balance point between the other two groups. This group reflects moderate energy behavior and is not too focused on one particular time of day.

This finding emphasizes the importance of consumer segmentation of electricity usage times as a basis for formulating more targeted energy efficiency policies. The implementation of schemes such as time-based tariffs (Time-of-Use Tariff) becomes more relevant when the government or electricity service providers understand when peak consumption occurs in each segment of society. For example, households in the night cluster can be the main target in a peak shaving program, while groups with high daytime consumption can be directed to utilize renewable energy, such as solar panels, to support the clean energy transition.

The K-Means Clustering method has proven to be an effective tool for uncovering hidden patterns in household electricity consumption data. Through this approach, households can be objectively classified based on similarities in energy usage patterns, allowing for more accurate load management strategies. This clustering also helps in mapping energy needs spatially and temporally, which is important in infrastructure planning and electricity supply optimization. Thus, the results of this analysis not only serve as statistical information, but also serve as a foundation for designing more adaptive, efficient, and data-driven energy policies.

5. Conclusion

Based on the research results, the use of data mining techniques through the clustering method has shown its effectiveness in identifying household groups based on daily electricity consumption habits. This approach makes a significant contribution to the formulation of more efficient energy policies that are oriented towards community needs. By knowing the different consumption patterns in each household segment, the government together with PLN can design more adaptive energy saving programs, determine a fairer electricity tariff structure, and plan the development of electricity infrastructure that is responsive to real conditions in the field. This information can also be an important basis for long-term decision making to support the sustainability and reliability of the national energy system.

6. Suggestions

Based on the research results, it is recommended that the government and PLN continue to develop the application of data mining techniques, especially the clustering method, in analyzing household electricity consumption patterns in more depth and sustainably. The results of the clustering can be used to design energy saving programs that are more targeted according to the characteristics of each segment of society. In addition, the electricity tariff structure also needs to be reviewed to be fairer and reflect actual consumption patterns. Planning for the development of electricity infrastructure should also be adjusted to the real needs identified through data analysis, so as to create a more responsive and efficient energy system. To support national energy sustainability, collaboration is needed between the government, electricity service providers, academics, and the community in integrating the results of data analysis into strategic long-term policies.

References

[1] S. Lestari, F. Teknik, P. S. Informasi, and U. P. Utama, "Penggunaan Metode K-Means dalam Data Mining untuk Identifikasi Pola Konsumsi Listrik pada Rumah Tangga," 2025.

[2] A. A. Zr, A. Ikhwan, and I. A. Sinaga, "Penerapan Data Mining untuk Rekomendasi Pilihan Lagu di RRI Medan dengan Metode Apriori," vol. 4, no. 2, pp. 37–49, 2024.

[3] R. Yanto Program Studi Sistem Informasi STMIK Bina Nusantara Jaya Lubuklinggau, "Implementasi Data Mining Prediksi Kebutuhan Tenaga Listrik Di Kota Lubuklinggau Data Mining Implementation to Predict Electricity Needs in Lubuklinggau City," vol. 19, no. 2, pp. 197–206, 2020.

[4] Salma, "PLN BUMN atau Swasta? – Mengenal Status PLN sebagai Badan Usaha Milik Negara." [Online]. Available: https://jadibumn.id/pln-bumn-

atau-swasta/
[5] A. Hafiz, F. Amanah, R. Damurti, and P. Harliana, "MENGIDENTIFIKASI POLA KONSUMSI ENERGI RUMAH TANGGA MENGGUNAKAN ALGORITMA GRAF BERBASIS C ++," vol. 8, no. 6, pp. 12057–12061, 2024.
[6] L. M. K. Amali, Y. Mohamad, A. I. Tolago, N. Elysiantobuo, and A. Y. Dako, "Analisis Konsumsi Energi Listrik Menggunakan Metode Internsitas Konsumsi Energi," *Jambura J. Electr. Electron. Eng.*, vol. 6, no. 1, pp. 103–107, 2024, doi: 10.37905/jjeee.v6i1.22567.
[7] U. Ma'rifatin, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng," *Semin. Nas. Inov. Teknol.*, no. 2549–7952, pp. 285–291, 2020.
[8] Y. P. Tanjung, S. R. Sentinuwo, and A. Jacobus, "Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree," *J. Tek. Inform.*, vol. 9, no. 1, 2016, doi: 10.35793/jti.9.1.2016.14141.
[9] Yulia and N. Azwanti, "Data Mining Prediksi Besarnya Penggunaan Listrik Rumah Tangga di Kota Batam Dengan Menggunakan Algoritma C4.5," *Semin. Nas. Ilmu Sos. dan Teknol.*, no. 1, pp. 175–180, 2018.
[10] N. Daamastuti, "Prediksi Penggunaan Listrik Rumah Tangga Menggunakan Metode Decision Tree C4.5 dan CART," *JREC (Journal Electr. Electron.*, vol. 10, no. 1, pp. 31–40, 2022, doi: 10.33558/jrec.v10i1.4443.
[11] A. Kukuh Wahyudi, N. Azizah, and H. Saputro, "Data Mining Klasifikasi Kepribadian Siswa Smp Negeri 5 Jepara Menggunakan Metode Decision Tree Algoritma C4.5," *J. Inf. Syst. Comput.*, vol. 2, no. 2, pp. 8–13, 2022, doi: 10.34001/jister.v2i2.392.
[12] Fajarudin, Wijaya, and Y. Arie, "Penerapan Metode K-Means Clustering dalam Mengelompokan Pelanggan PT . PLN Kota Cirebon," vol. 0577, 2023.
[13] A. N. A. Maulidhia, "Implementation of the K-Means Algorithm on Household Electric Load".
[14] M. K. Sinaga and S. R. Akbar, "Sistem Akuisisi Dan Pelabelan Data Penggunaan Energi Listrik Alat Elektronik Rumah Tangga Menggunakan K-Means Clustering," vol. 1, no. 1, pp. 1–9, 2023.
[15] R. D. Herdiansyah, F. Sains, D. A. N. Teknologi, U. Islam, N. Sultan, and S. Kasim, "Penerapan Algoritma K-Means Clustering Pada Penentuan Tingkat Pemakaian Listrik (Studi Kasus : PLTBg Rantau Sakti)," 2023.