This research focuses on developing a cost-effective non-intrusive load monitoring system (NILM) to identify household appliances using machine learning, specifically the k-nearest neighbors (kNN) algorithm which is not disturbing the existing system. The object of this research is the process of appliance identification based on power consumption characteristics in residential energy monitoring. The main problem to be solved is the lack of accessible, affordable, and efficient tools for monitoring household electricity consumption, as existing solutions are often costly or require complex installations. Existing solutions are expensive or require complicated setup. This research seeks to design a low-cost NILM that can identify household appliances without invasive system while ensuring high accuracy. This study successfully designed and implemented an electrical recording device that integrates machine learning algorithms, achieving an identification accuracy of 83.33 % across six test scenarios involving various household appliances. The findings of this study show that utilizing active power and power factor as classification parameters allows for effective equipment identification. The moderate accuracy of the system indicates that the proposed design is quite promising but can be improved with more advanced algorithms and additional sensor data. The resulting system is cost-effective due to its inexpensive components which are achieved due to the modular design and the use of inexpensive components, such as the Wemos D1 mini and PZEM-004T V3 sensors, which simplify implementation and enhance system scalability. Its built-in LCD provides real-time monitoring without the need for internet connectivity. This research demonstrates the feasibility of a scalable and cost-effective NILM system, which can be further improved with advanced algorithms and additional sensor data for broader applications in smart energy management

Keywords: non-intrusive, monitoring, appliance, identification, kNN, smart, grid, energy, logger, power, consumer

UDC 37.018
DOI 10.15587/1729-4061.2025.316694

NON-INTRUSIVE LOAD MONITORING: A COST-EFFECTIVE APPROACH FOR HOME APPLIANCE IDENTIFICATION UTILIZING MACHINE LEARNING

Levin Halim
Corresponding author
Head of Laboratory
Laboratory of Electronics, Measurement, and Instrumentation*
E-mail: halimlevin@unpar.ac.id
Reyvaldo Barthez
Bachelor of Engineering*
Nico Saputro
Associate Professor*
*Department of Electrical Engineering
Parahyangan Catholic University
Ciumbuleuit str., 94, Bandung, Indonesia, 40141

Received 06.11.2024
Received in revised form 30.12.2024
Accepted 16.01.2025
Published 28.02.2025

How to Cite: *Halim, L., Barthez, R., Saputro, N. (2025). Non-intrusive load monitoring: a cost-effective approach for home appliance identification utilizing machine learning. Eastern-European Journal of Enterprise Technologies, 1 (8 (133)), 46–55. https://doi.org/10.15587/1729-4061.2025.316694*

1. Introduction

The increase in electrical energy consumption in the household sector is unavoidable and will continue to increase [1, 2]. Therefore, the International Electrotechnical Commission states that intelligent electricity use can be crucial in energy problems [3]. Based on data in [4], it is known that the level of electrical energy savings can reach 17 % per year if specific feedback is provided regarding the load on household equipment. In addition, consumer habits are vital in utilizing energy effectively [5]. Consumers tend to change energy usage patterns if they receive feedback regarding detailed data on their energy usage [6, 7]. The solution that can be done is to monitor loads and create an identification system that makes it possible to determine what equipment is used at home [8, 9].

Load monitoring and identification techniques are generally divided into two, namely, the intrusive monitoring process (ILM) and the non-intrusive monitoring process (NILM) [9, 10]. The non-invasive load monitoring technique makes monitoring the overall electrical load possible by using only one sensor at a certain measurement point [11, 12]. The type of equipment used and power consumption can also be identified based on the load characteristics from the NILM results [13, 14]. Another advantage of NILM is the low initial installation costs, which only use one monitoring tool, compared to ILM, which requires one monitoring tool for each household appliance to be monitored [15].

Addressing the ever-increasing electricity consumption in households is critical to ensure sustainable energy management. As global energy demand increases, the efficient use of energy in the housing sector will be one of the essential focuses. Despite advances in energy management systems, many solutions require high costs, complicated installations, or the need to disrupt the installed system. Additionally, as energy costs rise and environmental concerns continue, non-intrusive load monitoring (NILM) has gained attention due to its potential to reduce energy consumption without requiring invasive hardware installations. The NILM system provides real-time feedback on energy usage, encouraging consumers to adopt energy-efficient behaviors. Given the continuous advancements in machine learning algorithms and the increasing need for smart energy solutions, research on developing more accurate, cost-effective, and user-friendly NILM systems will be particularly relevant. Given the continuous advancements in machine learning algorithms and the increasing need for smart energy solutions, research on developing more accurate, cost-effective, and user-friendly NILM systems will be particularly relevant.

2. Literature review and problem statement

Several related studies try to provide solutions for identifying household equipment through non-intrusive load monitoring. In research [15], the data collection process uses the HIOKI PW3390 power analyzer, which has a high level of precision to monitor voltage and current values and uses the kNN algorithm for the household equipment identification process. The characteristics used in research [15] are active power (Watts) and real power in units of voltamperes (VA). In research [16], the data collection process was carried out using a smart meter and the support vector machine (SVM) algorithm to identify household equipment that features the voltage and current values forming a recognition path. The trajectory features in research [16] can identify similar types of electrical equipment, which is insufficient to determine the kind of equipment if the time interval is too short. In research [17], the data collection process also uses a smart meter, and the algorithm used is a deep neural network to classify household equipment. Deep neural networks require a labeling process for machine learning models to be able to identify them.

Another critical aspect of the NILM system is integrating energy-efficient feedback mechanisms tailored to user behavior. Studies have shown that real-time feedback on energy consumption can influence user habits and significantly reduce overall electricity usage [6, 18]. Research [9, 19] has explored the inclusion of personalized dashboards, combining NILM data with predictive analytics to suggest optimal load usage schedules, leading to improved energy efficiency. On the other hand, using cloud-based storage for NILM data, long-term trend analysis can also be performed so that users can make informed decisions regarding equipment replacement or upgrade [20, 21].

On the other hand, previous research has explored various methods of non-intrusive load monitoring (NILM) using multiple algorithms and data collection techniques. The research [22] uses the V-I trajectory method and the support vector machine (SVM) for load identification. Still, it cannot distinguish the equipment when the time interval between uses is too short. Similarly, research [23] uses deep learning models that require extensive labeled datasets, making real-time implementation expensive and complicated. Other research [24] implements the k-Nearest Neighbors (kNN) algorithm with a high-precision power analyzer, but the system cost and hardware dependency limit its practical application in residential environments. These challenges arise from the trade-off between system accuracy, hardware cost, and computing efficiency.

Based on the literature review conducted, it is clear that although significant progress has been made in NILM, several key issues remain unresolved. Existing solutions often rely on expensive hardware that is not accessible to users. Furthermore, many approaches use computationally heavy algorithms that limit scalability and real-time applications. The reliance on large datasets to train machine learning models poses a challenge for real-world implementation. An identified unresolved issue is the lack of a cost-effective and easy-to-use NILM system that balances affordability, accuracy, and scalability while minimizing hardware and computational requirements. This issue directly aligns with the objective of this study to design and implement a low-cost NILM system using the k-Nearest Neighbors (kNN) algorithm, capable of accurately identifying household appliances based on their power consumption characteristics.

3. The aim and objectives of the study

The aim of the study is to design and develop a cost-effective non-intrusive load monitoring (NILM) system capable of accurately identifying household appliances using the k-Nearest Neighbors (kNN) algorithm based on power consumption characteristics.

To achieve this aim, the following objectives are accomplished:

– to develop a low-cost electricity logger using commercially available hardware components, ensuring easy deployment and scalability;

– to design and develop a data acquisition system that collects load characteristics, including active power and power factor, from household appliances;

– to apply the k-Nearest Neighbors (kNN) algorithm to classify and identify appliances based on the acquired load characteristics;

– to evaluate the system's performance through experimental testing, analyzing its identification accuracy, reliability, and cost-effectiveness.

4. Materials and methods

4. 1. Object and hypothesis of the study

The research object is the process of appliance identification based on power consumption characteristics in residential energy monitoring. The hypothesis is that using the k-Nearest Neighbors (kNN) algorithm with active power and power factor as classification features can achieve accurate appliance identification with minimal hardware requirements.

Some of the assumptions made for this study include that the household appliances reviewed can exhibit different load characteristics, particularly in active power parameters and power factors. In addition, it is assumed that the equipment tested under controlled experimental conditions represents a typical residential energy consumption scenario, and the results can be generalized to the wider household environment.

In addition, there are several simplifications adopted in this study to ensure feasibility and cost-effectiveness. First, the system uses only two load characteristics, namely active power and power factor. This is done for ease of classification which will simplify the data acquisition process but there are still significant differences between types of loads. Second, hardware components, such as the PZEM-004T V3 sensor and Wemos D1 Mini, were chosen because of their affordable price, which may result in lower precision compared to high-end alternatives. Finally, the experimental arrangement was carried out under controlled conditions with a limited number of equipment, which did not take into account the variability of the type of load. This simplification, while necessary for initial development, highlights areas for future work, such as integrating additional features, testing more diverse types of equipment, and improving the precision of the system.

4. 2. Electrical design of electricity logger

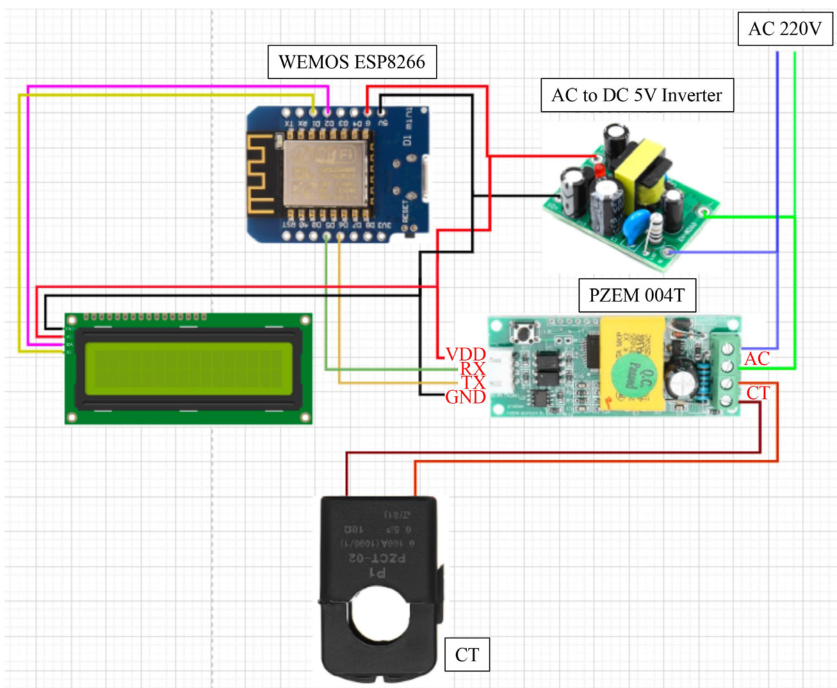
A schematic diagram of the electricity logger is displayed in Fig. 1, *a*. In addition to measuring AC voltage, the PZEM-004T V3 can also measure active power, energy, frequency, and power factor [25]. The P1 PZCT-02, connected to the PZEM 004T V3, measures the current in the cable. The 5V AC to DC converter supplies power to the Wemos D1 mini and LCD with a consistent DC supply. The Wemos D1 mini LCD shows real-time data on electricity usage monitoring. The Wemos D1 Mini, programmed with the Arduino IDE, is set up to capture and transmit data readings through USB or WiFi using the PZEM-004T V3 and P1 PZCT-02 sensors. The LCD is operated through the SDA (Data) and SCL (Clock) pins on D2 and D1, respectively. To link the PZEM-004T V3 to the Wemos D1 Mini, the RX pin is connected to pin D5, and the TX pin is connected to pin D6. On the other hand, the electronic system of this research consists of several components in the electrical design. In simple terms, each working component can be categorized as input, process, or output, as shown in Fig. 1, *b*.

The system efficiently integrates sensors and a control unit for real-time load monitoring. The PZEM-004T V3 and P1 PZCT-02 sensors measure key electrical parameters, while the Wemos D1 Mini manages data processing and transmission. This modular design ensures a cost-effective and scalable solution for residential energy monitoring.

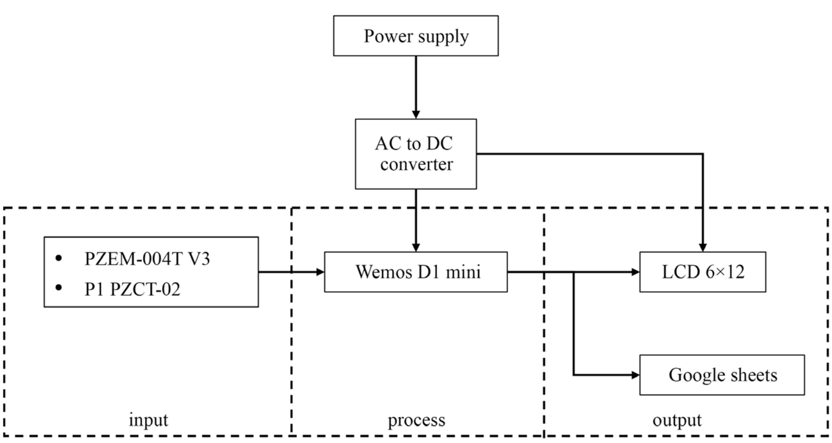
4. 3. System framework

This research builds on existing works, such as those by [17] and [20], which explored deep learning and high-precision hardware for NILM. Unlike these studies, the proposed system uses lightweight algorithms and low-cost sensors to balance performance and affordability.

An electrical load is first connected to the socket, as shown in Fig. 2, *a*. Then, the PZEM-004T V3 module and P1 PZCT-02 sensor on the electricity logger will read the characteristics of the electrical load, such as current, voltage, active power, energy, frequency, and power factor. The reading results will then be sent to Wemos D1 mini as the electricity logger control center. After that, the Wemos D1 mini will display the sensor reading results via a 16×2 LCD and send measurement data using a WiFi connection to Google Sheets. The data can also be sent directly to a computer or laptop using a USB cable. Only after that can the electrical load identification process be carried out using machine learning on a laptop or computer. Fig. 2, *b* shows the electrical load reading process, the data retrieval or acquisition process carried out by the Wemos D1 mini microcontroller, and then the load data identification process.

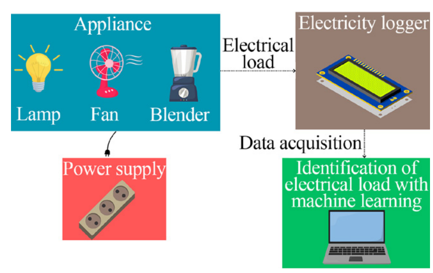


a

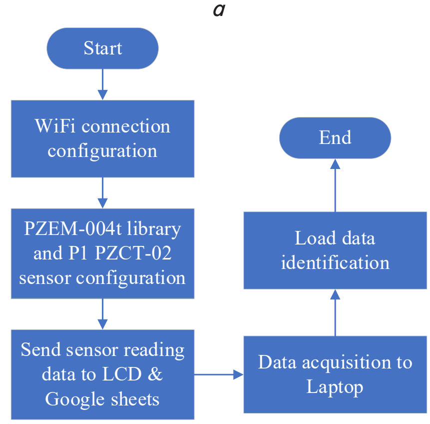


b

Fig. 1. Electrical system design: *a* – schematic; *b* – block diagram



a



b

Fig. 2. System framework: *a* – illustration; *b* – flowchart

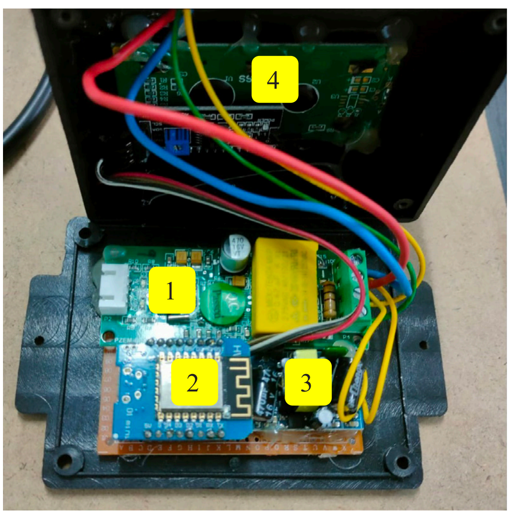
The system framework enables real-time load monitoring using the PZEM-004T V3 and P1 PZCT-02 sensors, managed by the Wemos D1 Mini. Data is displayed on an LCD and transmitted via USB or WiFi for appliance identification using the k-Nearest Neighbors (kNN) algorithm, ensuring accurate and efficient load classification.

Experimental testing method was chosen to simulate realistic residential energy scenarios, to ensure the system can be implemented in real life. Theoretical modeling using kNN was chosen because of its simplicity, computational efficiency, and proven reliability in similar classification.

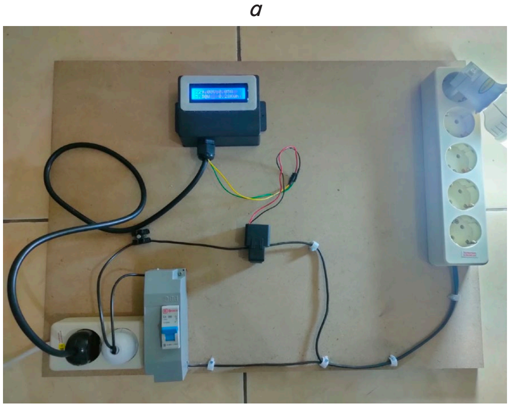
5. Results of research on system performance and appliance identification accuracy

5. 1. Development of a low-cost electricity logger

The logger device is a black beam with a 16×2 LCD screen to display the reading. Fig. 3, *a* shows a Wemos 8266, PZEM003Tv3, LCD i2c 16×2, and AC to DC Converter 5V. The Wemos 8266, indicated with the number "1" functions as a control center that will read and collect data from energy measurements by sensors, and the data can be sent using WiFi or a USB cable. PZEM003Tv3, indicated by the number "2", measures energy consumption with the PZCT sensor next to the electricity logger. As noted in the number "3", the 5V AC to DC Converter will change the AC voltage to DC to provide power to the Wemos 8266 and other components. Meanwhile, the 16×2 i2c LCD, indicated by the number "4" functions to display reading results from data originating from the sensor and recorded by Wemos 8266.



a



b

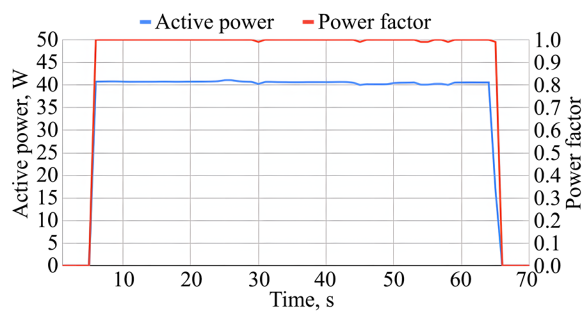
Fig. 3. Testing kit: *a* – logger; *b* – implementation

On the other hand, the base of the testing kit uses an MDF wooden board measuring 400×550×12 mm, as shown in Fig. 3, *b*. Several components other than the electricity logger are installed on the model: a two-socket extension socket, a five-socket extension socket, and a Mini Circuit Breaker. The two-socket extension socket functions as a grid electricity source to turn on the electricity logger and channel electricity to the five-socket extension socket filled by various electronic equipment as an electrical load. Mini Circuit Breaker protects in the event of overload and short circuit.

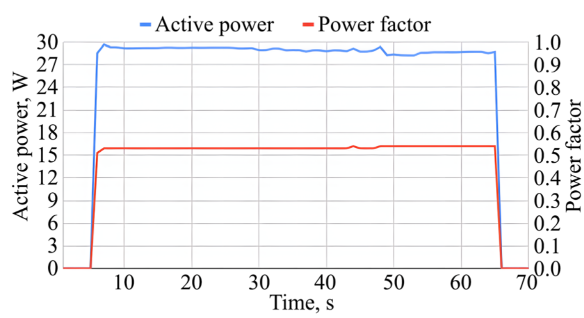
5. 2. Data acquisition system for load characteristics

Load data is collected by connecting the load to a 5-socket extension socket. Data retrieval is carried out in stages, starting from connecting only one load per data retrieval, then combining two types per data retrieval, to combining three types per data retrieval. The data retrieval process is carried out at intervals every second, and Coolterm software is used to communicate with serial devices via USB.

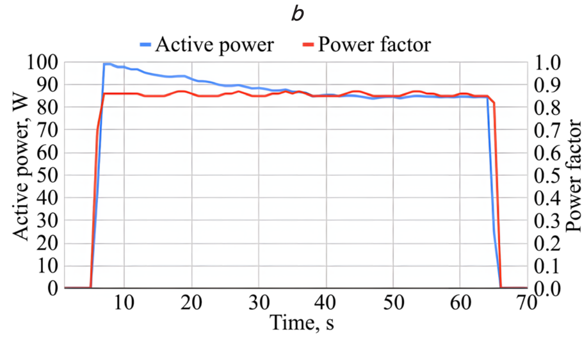
For one type of load scenario, Fig. 4, *a* shows the lamp load characteristics when data is collected for 70 seconds with one-second intervals per data. The lamp is on for 60 seconds, from five to 65 seconds. The load characteristic of the lamp is the power value, which is 40.24–41.08 Watts on the left vertical axis. Meanwhile, the Power Factor value is 0.99–1, which can be seen on the right vertical axis.



a



b



c

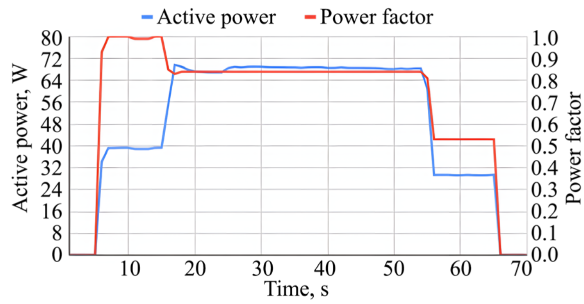
Fig. 4. One type of load scenario characteristics: *a* – lamp; *b* – fan; *c* – blender

Fig. 4, *b* presents a graph of the load characteristics of Fan 1 when data was collected for 70 seconds with an interval of one second per data. A fan with a speed I is on for 60 seconds, from five seconds to seconds 65. The load characteristics of the fan are the power value, which is 28.23–29.68 Watts on the left vertical axis. Meanwhile, the Power Factor value is 0.51–0.54, which can be seen on the right vertical axis.

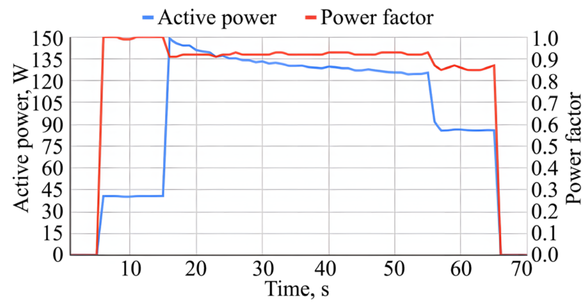
Fig. 4, *c* illustrates the overlapping power factor measurements for of the load characteristics of the blender, highlighting their similar operational characteristics. While these curves may appear visually dense, they are essential for demonstrating the system's ability to classify appliances with closely related load profiles. When data was collected for 70 seconds with an interval of one second per data. The blender is on for 60 seconds, from 5 to 65 seconds. The load characteristic of the blender is the power value, which is 83.87–98.99 Watts on the left vertical axis. Meanwhile, the Power Factor value is 0.82–0.87, which can be seen on the right vertical axis.

For two types of load scenario, Fig. 5, *a* is displays the active power consumption trends of multiple appliances under simultaneous operation. The overlapping curves emphasize the challenges inherent in distinguishing similar appliances, which the proposed kNN algorithm addresses effectively. The data collection process was carried out for 70 seconds with one-second intervals. As seen, the change in power consumption from 0 Watts to 34.3 Watts at second 6 and moving to 38.81–39.28 Watts until second 15 due to the lamp load being used. Then, the power consumption increases at 16 seconds to 55.11 Watts, and until 55 seconds, the power value moves between 66.93–69.65 Watts due to the lamp load and fan load. At 56 seconds, the power consumption is reduced to 29.32 Watts because the lamp is turned off. Power consumption at seconds 56 to 65 moves between the values 29.2–29.37 Watts. At 66 seconds, the fan is turned off so that power consumption becomes zero. The power factor value of the combination of lamp load and aquarium pump is 0.53–1. At 6 seconds, the power factor value becomes 0.93 and moves between 0.99–1 until 15 seconds due to the lamp load. Then, at seconds 16 to 55, the power factor is 0.81–0.85 due to the load of the lamp and fan. At 56 seconds, the lamp is turned off, and the power factor value drops and remains stable at 0.53 for 10 seconds. After 10 seconds, the fan is turned off, and data collection is complete.

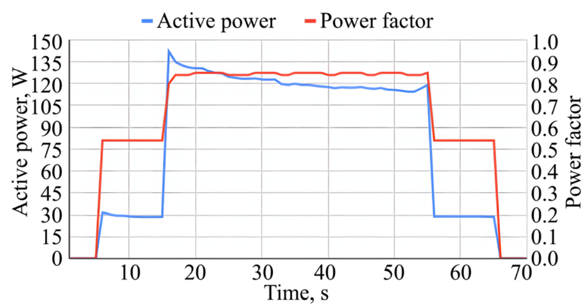
Fig. 5, *b* presents a graph of the characteristics of the combination of lamp and blender loads. The data collection process was carried out for 70 seconds with one-second intervals. The change in power consumption after the lamp is turned on, from 0 Watt to 40.62 Watt at second 6 and moving up and down at 40.17–40.67 Watt until the second 15. Power consumption at second 16 reaches 149 Watt immediately after the blender is turned on, then until second 55, and the power value moves between 124.46–145.95 Watts due to the load of the lamp and blender. At 56 seconds, the lamp is turned off, causing power consumption to decrease to 91.63 Watts. The blender was turned off at 65 seconds, so the power was zero. The power factor value of the combined lamp and blender load is 0.85–1. At 6 seconds, the lamp turns on, making the power factor value 1, and moves to 0.99–1 until 15 seconds. At 16 to 55 seconds, the blender is turned on simultaneously as the lamp is on. The power factor value changes to 0.71 at 16 seconds; from 17 to 55 seconds, the value is 0.59–0.71. At 56 seconds, the power factor value drops to 0.43–0.46 for 10 seconds because the lamp is turned off. After 10 seconds, the blender is turned off, and data collection is complete.



a



b



c

Fig. 5. Two types of load scenario characteristics: *a* – lamp+fan; *b* – lamp+blender; *c* – fan+blender

Fig. 5, *c* presents a graph of the characteristics of the combination of fan and blender loads. The data collection process was carried out for 70 seconds with one-second intervals. The power value changes from 31.53 Watts at 6 seconds to 28.52–30.43 Watts at 15 seconds because the fan is on. Power consumption at 16 seconds reached 142.09 Watts immediately after the blender was turned on, then until 55 seconds, the power value moved between 114.44–134.83 Watts due to the load of the fan and blender. At 56 seconds, the blender is turned off, causing power consumption to decrease to 28.81 Watts, and the value changes between 28.69 and 28.85 Watts until 65 seconds. At 66 seconds, the fan is turned off, and the power value becomes 0. The power factor value of the fan load combination and blender is 0.54–0.85. At 6 seconds, the power factor value becomes 0.54, which is stable at that number when the fan is turned on for 15 seconds. Then, from 16 seconds to 55 seconds, the blender is turned on at the same time as the fan is turned on, and the power factor value changes to 0, 8 at 16 seconds. Then, from 17 to 55 seconds, the value moves to 0.8–0.85. At 56 seconds, the blender is turned off, causing the power factor value to drop to 0.54 for 10 seconds. After that, at 66 seconds, the fan is turned off, and data collection is complete.

For three types of load scenario, Fig. 6 is a graph of the characteristics of a combination of three types of loads, measured for 90 seconds with a data recording interval of one second.

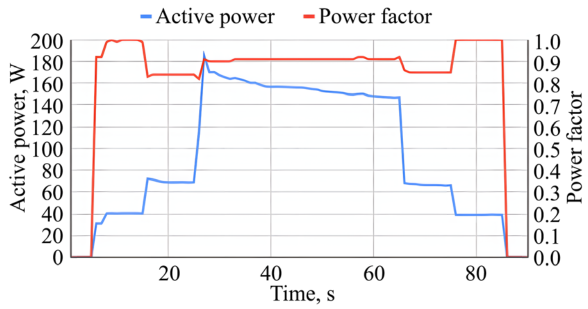


Fig. 6. Chart of combination characteristics of three types of loads

Table 2

Example of the dataset used

Active power (W)	Power factor	Type	Active power (W)	Power factor	Type
40.53	1	L	69.04	0.84	LK
40.03	0.99	L	127.24	0.92	LB
28.53	0.51	KA	128.63	0.93	LB
29.68	0.53	KA	142.09	0.8	KB
84.45	0.85	B	132.52	0.84	KB
84.45	0.85	B	185.22	0.91	LKB
69.65	0.83	LK	170.31	0.9	LKB

The power value increases at the second 5 to 31.04 Watts and moves to 40.27–40.5 Watts until the second 15 because the lamp is turned on. Then, at 16 seconds, the fan is turned on at the same time as the lamp, causing the power value to increase and move to 68.78–72.27 Watts. At 26 seconds, the three loads are turned on simultaneously (lamp, fan, and blender) so that at 27 seconds, the power jumps to 185.22 Watts. In seconds 26 to 65, the power is at 146.68–185.22 Watts. The blender shuts down at 66 seconds, dropping the power value to 65.91–68.35. Next, at 76 seconds, the remaining lamp is on until 86 seconds, with the power value recorded at 39.06–39.4 Watts.

On the other hand, the combination of three types of loads (lamp, fan, and blender) has a power factor value of 0.82–1. However, the power factor value is 0.91–0.92 when using the three loads simultaneously.

5. 3. Application of the k-nearest neighbors (kNN) algorithm

Identifying load types using machine learning, precisely the k-nearest neighbors method, requires several processes that must be carried out until the system can determine the type of load. Starting from creating a dataset so that the learning process can be carried out, it is necessary to choose the correct k value in k-nearest neighbors so that the average identification error is low. Then, the identification process is carried out based on the k value. Based on the existing load combinations, Table 1 describes the name labels of the combinations of household equipment used in this research.

Table 1

List of load labels and load combinations

Load name label	Load combination
L	Lamp
KA	Fan
B	Blender
LK	Lamp+fan
LB	Lamp+blender
KB	Fan+blender
LKB	Lamp+fan+blender

The dataset is created from load combination data taken previously, explained previously. The data in the form of a collection of active power and power factor values is then labeled as in Table 2 so that the machine learning algorithm used can understand the results that should be given. The machine learning algorithm will learn to identify loads with active power characteristics and power factors. The type column in Table 2 labels the equipment used, as seen in Table 1.

The need for an appropriate k value can be determined using the k vs Error Rate plot, as seen in Fig. 7, *a* which shows that $k=1$ to $k=12$ has an error rate of 0.0, which can be called the optimal k value. Furthermore, at values $k=13$ to $k=30$, the error rate increases as the value of k used increases. The graph in Fig. 7, *a* was obtained through a testing process using a separate dataset utilizing Python language.

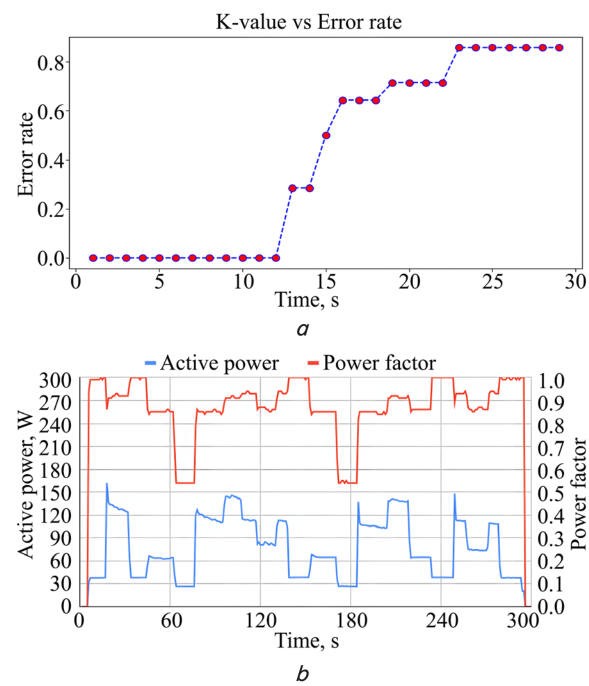


Fig. 7. Determining k-value: *a* – error value compared to the number of k used; *b* – dataset

Next, a load identification process is carried out, which begins with taking five minutes of electrical load combination data, which will be used as test data. The electrical load consists of a lamp, fan, and blender. After the test data has been taken, six electrical loads are selected in seconds over the five minutes in Table 3. On the other hand, Fig. 7, *b* is a graph of the test data load for this research.

The selected data tests are at the 25th second, 68th second, 104th second, 125th second, 264th second, and 290th second. The load characteristic values from the selected test data can be seen in Table 3.

The load characteristic data is then put into machine learning for the load identification process. The characteristics entered are each data's active power and power factor, starting from the 25th second of data until the sixth data.

This determines the k value and defines the test data according to the characteristics in Table 3. The data entered is the active power value first, followed by the power factor value.

Table 3

Data selected from datatest

Time (s)	Voltage (V)	Current (A)	Active power (W)	Power factor
25	221	0.63	129.66	0.93
68	223	0.22	26.98	0.54
104	222.6	0.71	143.91	0.91
125	222.4	0.43	82.24	0.86
264	221.7	0.39	74.07	0.86
290	221.3	0.17	37.68	0.99

After the test data is elaborated, the machine learning algorithm will measure and look for the five closest to each test data because the k value used is five. After the calculation process in the program, each data test value was identified as a combination of certain types of household equipment load.

5. 4. System performance evaluation

The results of identifying load characteristics can also be understood by looking at the Decision Boundary Matrix in Fig. 8. Each type of load combination is represented by a different color with a circle shape. Meanwhile, the test data is in the form of a red triangle. The position of each data that it is necessary to identify can be seen in the figure. The data is located and adjacent to different data groups. The existing identification results are compared with the data observations carried out manually during testing. The results of this comparison can be seen in Table 4.

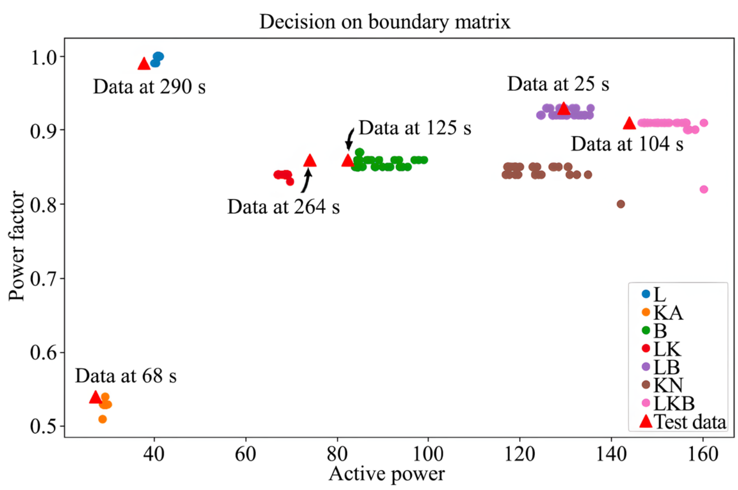


Fig. 8. Decision boundary matrix overall identification system testing

Based on Table 4, the results of machine learning predictions can be compared with the results of manual observations from the data in Table 2. Testing on the 25th second of data has resulted in lamp+blender loads, the same between machine learning algorithm predictions and observations made manually. After that, testing the data for the 68th second had results in the form of fan load, where the prediction results by machine learning and observation showed similar answers. For data at the 104th second, machine learning prediction and observation results identify the same load, namely lamp+fan+blender. The data at the 125th second has the same identification results when predicted by machine learning and is matched with the observation results in blender load. The load identification results at the 264th second were expected as a lamp+fan load by the machine learning algorithm, while based on observations made, the load being used was a blender. The load identification results at the 290th-second show that a lamp load was successfully predicted by machine learning and following the observation results.

Table 4

Confusion matrix from load identification results

Results	Observation results							
Machine learning result	Load	L	KA	B	LK	LB	KB	LKB
	L	1	0	0	0	0	0	0
	KA	0	1	0	0	0	0	0
	B	0	0	1	0	0	0	0
	LK	0	0	1	0	0	0	0
	LB	0	0	0	0	1	0	0
	KB	0	0	0	0	0	0	0
	LKB	0	0	0	0	0	0	1
Accuracy	83.33 %							

The error in identifying the load in the 264th second of data, a type of blender load, occurs because the blender's power usage characteristics tend to increase when first used and slowly decrease. So, the blender load data identified at the 264th second has an active power value close to the lamp+fan load combination characteristics. The nearest neighbor of the 264th-second data is the lamp+fan load type, so the data at the 264th second is identified as lamp+fan. The k-nearest neighbors algorithm successfully identified household loads when testing the identification system. Five load types were correctly identified in the six identifications, and one was incorrectly identified. So, the accuracy value of the identification system testing is 83.33 %.

6. Discussion of results on system performance and appliance identification

This research contributes to the development of NILM technology by demonstrating cost-effective solutions using available components and lightweight machine learning algorithms. The use of new parameters of active power and power factor as classification features offers a practical balance between accuracy and affordability.

The development of NILM, as illustrated in Fig. 1, *a* integrates the PZEM-004T V3 sensor, Wemos D1 Mini microcontroller, and LCD screen. The successful integration of these components allows for reasonably accurate monitoring of time loads while maintaining cost efficiency. The system framework in Fig. 2, *b* shows how the data from sensor readings to machine learning-based load identification shows a well-structured hardware-software connection.

A graph of load characteristics supports the effectiveness of system data acquisition. Fig. 4 shows the active power and power factor readings for the lights, fans, and blenders, indicating consistent and precise load monitoring. Similarly,

Fig. 5, 6 present a dynamic response when multiple equipment is operating simultaneously, confirming the system's ability to track changes over time.

The k-nearest neighbors (kNN) algorithm's application is explained by determining the optimal k value in Fig. 7, *a* which shows that the value between $k=1$ and $k=12$ minimizes the error rate. The optimal selection of k-values ensures accurate classification, supported by the sample dataset in Table 2.

The overall performance evaluation of the system is reflected in the equipment identification results shown in Table 4. With an identification accuracy of 83.33 %, the confusion matrix showed the correct classification in five of the six test cases. The error of classifying the blender as a lamp-fan combination at the 264th second occurs due to the characteristics of the overlapping loads, which are visually explained by the decision limit matrix in Fig. 8.

The proposed NILM system provides several features that offer significant advantages compared to existing solutions. This study's cost-effective design utilizes affordable components such as the Wemos D1 Mini and PZEM-004T V3 sensors, unlike systems that use high-precision devices such as the HIOKI PW3390 power analyzer [15]. In addition, the system in the study simplifies data collection and processing by using only two load characteristics, namely active power and power factor, which are then processed using the k-Nearest Neighbors (kNN) algorithm, which reduces the need for labeled datasets and extensive computing resources [17].

The system's modular structure increases scalability, thus opening up options for integration with other smart devices and advanced machine learning models. In the future, these systems can function independently or be integrated into cloud-based platforms, which ensures flexible deployment [21]. Furthermore, the electricity logger is equipped with a built-in LCD screen for real-time monitoring, which provides immediate energy consumption feedback without internet connectivity [9].

The kNN-based classification algorithm achieves a competitive identification accuracy of 83.33 %, showing reliable performance despite minimal hardware and limited features. This differs from systems that use support vector machines (SVMs), which struggle with overlapping load characteristics [16].

The proposed system effectively addresses the problems identified before, precisely the trade-off between system cost, complexity, and monitoring accuracy in non-intrusive load monitoring (NILM). The selection of hardware in this study successfully minimized the cost of the system while maintaining adequate measurement precision, as demonstrated by an identification accuracy of 83.33 % during experimental testing (Table 4).

The problematic reliance on complex algorithms and extensive labeled datasets found in previous studies [17] was solved by applying the k-Nearest Neighbors (kNN) algorithm. This lightweight and efficient machine learning model requires only two features-active power and power factor-that simplify data collection and processing, as seen in Fig. 7, *a* where the optimal k-value can be specified to minimize the error rate. The real-time data visualization the internal LCD provides ensures immediate feedback, eliminating reliance on external cloud-based systems [9]. This feature further reduces implementation complexity and operational costs.

Experimental results, such as the load identification charts in Fig. 4–6, show that the system can monitor a wide range of equipment, even in scenarios with up to three different load types. The system successfully classifies equipment based on its unique load characteristics, as shown by the decision limit matrix in Fig. 8.

Thus, the study effectively provides a scalable, cost-effective, and accurate NILM solution. This comprehensive approach can solve the unmet need for a practical and affordable load identification system.

The proposed NILM system also has limitations and shortcomings that must be considered when implementing it. The system relies on only two parameters-active power and power factor-that limit its ability to distinguish between equipment with the same power consumption pattern. This obstacle was evident in misclassifying the blender as a lamp-fan combination during testing (Table 4). Expanding the feature set by incorporating additional parameters such as harmonic distortion or reactive power may improve classification accuracy.

In addition, the PZEM-004T V3 and P1 PZCT-02 sensors may not be able to capture high-frequency transient data, reducing accuracy for equipment with rapidly changing loads. Future designs may incorporate higher-resolution sensors to improve system precision. On the other hand, one main shortcoming is its reliance on specific hardware components, such as the PZEM-004T V3 and Wemos D1 Mini. If these components are unavailable, the system design will require substantial modifications, limiting long-term sustainability.

The system currently operates as a standalone device with limited integration capabilities. While real-time monitoring can be done through LCD screens, advanced data storage and analytics require manual data transfer. Therefore, implementing a cloud-based storage platform and real-time analytics can improve scalability and usability. The system's user interface is limited to a basic LCD screen with minimal data visualization capabilities. There is no companion mobile or web app for remote monitoring, detailed analysis, or system control, which may limit its appeal to modern consumers who are used to feature-rich smart home solutions.

Other limitations are related to environmental factors, such as the stability of the power supply and WiFi connectivity. Voltage fluctuations or unstable wireless connections can interfere with data acquisition and system operation. Therefore, it is recommended that the system be implemented in an environment with a stable power infrastructure.

On the other hand, the performance of the k-nearest Neighbors (kNN) algorithm relies heavily on the selection of optimal k-values and having a well-labeled training dataset. This dependency can limit the system's adaptability in a dynamic home environment where equipment configurations change frequently. More advanced algorithms, such as neural networks or decision trees, can provide better performance.

The system also does not have a built-in fault tolerance mechanism. There is no provision for error detection, self-correction, or system fallback in case of hardware failure or data transmission issues. In addition, experimental settings are limited to controlled environments with a small set of household appliances, so the system's performance in diverse and complex real-world settings has not been tested.

This research can be further developed by addressing its significant shortcomings and expanding the scope of its technology. One is to improve the system's data processing capabilities by incorporating more sophisticated machine learning models, such as neural networks or decision trees, that can improve classification accuracy, especially in scenarios with characteristics with different types of loads.

Another potential improvement lies in expanding the parameter beside active power and power factors. Adding harmonic distortion, current waveforms, and transient response data can improve tool differentiation. This will help reduce false positives and improve the system's overall reliability.

The system's hardware architecture can also be improved by integrating higher-resolution sensors. These improvements will reduce reliance on low-cost sensors with limited precision and improve data quality for better machine learning performance.

Additionally, by enabling real-time data synchronization with the cloud platform, users can access detailed energy consumption reports, predictive analytics, and load management recommendations via mobile and web apps. It will transform the system from a self-recorder to a fully integrated smart home energy management solution.

On the other hand, conducting more extensive experimental testing in various residential environments will validate the system's performance under real-world conditions. It will provide valuable insights into its adaptability to various electrical infrastructures and load profiles, strengthening its commercial and industrial potential. The proposed NILM system has been tested in simulated residential settings, demonstrating its practical application for real-time appliance monitoring. Its cost-effective design and ease of deployment make it suitable for widespread use in energy management programs, particularly for households aiming to reduce energy consumption and costs.

7. Conclusions

1. For development of low-cost electricity non-intrusive load monitoring system was using commercially available components such as the Wemos D1 Mini and the PZEM-004T V3. Developed logger enables the accurate collection of active power and power factor data from household appliances, supporting a scalable and cost-effective monitoring system.

2. The proposed system can effectively collect the load characteristics of multiple household appliances, demonstrating consistent measurement performance. The data acquisition results show stable power factor readings and accurate active power values, supporting real-time load monitoring and precise power usage tracking.

3. Implementing the k-nearest neighbors (kNN) algorithm proves that the classification of household appliances is successfully carried out based on active power parameters and power factors. The optimal selection of the k-value can minimize classification errors, achieving an overall identification accuracy of 83.33 %.

4. The performance evaluation of the system confirms its reliability and scalability for household energy monitoring. The experiment results demonstrated an identification accuracy of 83.33 % across six test scenarios involving various household appliance combinations. This consistent performance highlights the system's capability to effectively classify appliances based on active power and power factor measurements.

Conflict of interest

The authors declare that they have no conflict of interest concerning this research, whether financial, personal, authorship, or otherwise, that could affect the research and its results presented in this paper.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence

The authors confirm they did not use artificial intelligence technologies when creating the current work.

Acknowledgments

The authors thank the Center of Control, Automations, and Systems Engineering at Parahyangan Catholic University for providing the resources and facilities necessary to conduct this research.

References

1. Khan, I. (2019). Household factors and electrical peak demand: a review for further assessment. Advances in Building Energy Research, 15 (4), 409–441. https://doi.org/10.1080/17512549.2019.1575770
2. Zuki, N. A. M., Othman, R. N. F. K. R., Shukor, F. A. A., Ahmad, S. R. C. (2023). Analysis of linear motor with symmetrical EMF vector for household elevator application. International Journal of Power Electronics and Drive Systems (IJPEDS), 14 (1), 51. https://doi.org/10.11591/ijpeds.v14.i1.pp51-59
3. Divina, F., García Torres, M., Goméz Vela, F. A., Vázquez Noguera, J. L. (2019). A Comparative Study of Time Series Forecasting Methods for Short Term Electric Energy Consumption Prediction in Smart Buildings. Energies, 12 (10), 1934. https://doi.org/10.3390/en12101934
4. Marangoni, G., Tavoni, M. (2021). Real-time feedback on electricity consumption: evidence from a field experiment in Italy. Energy Efficiency, 14 (1). https://doi.org/10.1007/s12053-020-09922-z
5. Hussein, H. I., Abdullah, A. N., Jafar, A. S. J. (2023). A novel online monitoring system of frequency oscillations based intelligence phasor measurement units. International Journal of Power Electronics and Drive Systems (IJPEDS), 14 (3), 1589. https://doi.org/10.11591/ijpeds.v14.i3.pp1589-1596
6. Eirinaki, M., Varlamis, I., Dahihande, J., Jaiswal, A., Pagar, A. A., Thakare, A. (2022). Real-time recommendations for energy-efficient appliance usage in households. Frontiers in Big Data, 5. https://doi.org/10.3389/fdata.2022.972206

7. Mohammed, N., A. Danapalasingam, K., Majed, A. (2018). Design, Control and Monitoring of an Offline Mobile Battery Energy Storage System for a Typical Malaysian Household Load Using PLC. International Journal of Power Electronics and Drive Systems (IJPEDS), 9 (1), 180. https://doi.org/10.11591/ijpeds.v9.i1.pp180-188
8. Tundis, A., Faizan, A., Mühlhäuser, M. (2019). A Feature-Based Model for the Identification of Electrical Devices in Smart Environments. Sensors, 19 (11), 2611. https://doi.org/10.3390/s19112611
9. Welikala, S., Dinesh, C., Ekanayake, M. P. B., Godaliyadda, R. I., Ekanayake, J. (2019). Incorporating Appliance Usage Patterns for Non-Intrusive Load Monitoring and Load Forecasting. IEEE Transactions on Smart Grid, 10 (1), 448–461. https://doi.org/10.1109/tsg.2017.2743760
10. Ghosh, S., Chatterjee, A., Chatterjee, D. (2019). Improved non-intrusive identification technique of electrical appliances for a smart residential system. IET Generation, Transmission & Distribution, 13 (5), 695–702. https://doi.org/10.1049/iet-gtd.2018.5475
11. Wójcik, A., Łukaszewski, R., Kowalik, R., Winiecki, W. (2019). Nonintrusive Appliance Load Monitoring: An Overview, Laboratory Test Results and Research Directions. Sensors, 19 (16), 3621. https://doi.org/10.3390/s19163621
12. Cannas, B., Carcangiu, S., Carta, D., Fanni, A., Muscas, C. (2021). Selection of Features Based on Electric Power Quantities for Non-Intrusive Load Monitoring. Applied Sciences, 11 (2), 533. https://doi.org/10.3390/app11020533
13. Garcia, F. D., Souza, W. A., Diniz, I. S., Marafão, F. P. (2020). NILM-based approach for energy efficiency assessment of household appliances. Energy Informatics, 3 (1). https://doi.org/10.1186/s42162-020-00131-7
14. Schirmer, P. A., Mporas, I. (2023). Non-Intrusive Load Monitoring: A Review. IEEE Transactions on Smart Grid, 14 (1), 769–784. https://doi.org/10.1109/tsg.2022.3189598
15. Cimen, H., Cetinkaya, N., Vasquez, J. C., Guerrero, J. M. (2021). A Microgrid Energy Management System Based on Non-Intrusive Load Monitoring via Multitask Learning. IEEE Transactions on Smart Grid, 12 (2), 977–987. https://doi.org/10.1109/tsg.2020.3027491
16. Wang, A. L., Chen, B. X., Wang, C. G., Hua, D. (2018). Non-intrusive load monitoring algorithm based on features of V–I trajectory. Electric Power Systems Research, 157, 134–144. https://doi.org/10.1016/j.epsr.2017.12.012
17. Laouali, I., Ruano, A., Ruano, M. da G., Bennani, S. D., Fadili, H. E. (2022). Non-Intrusive Load Monitoring of Household Devices Using a Hybrid Deep Learning Model through Convex Hull-Based Data Selection. Energies, 15 (3), 1215. https://doi.org/10.3390/en15031215
18. Mataloto, B., Ferreira, J. C., Resende, R. P. (2023). Long Term Energy Savings Through User Behavior Modeling in Smart Homes. IEEE Access, 11, 44544–44558. https://doi.org/10.1109/access.2023.3272888
19. Aboulian, A., Green, D. H., Switzer, J. F., Kane, T. J., Bredariol, G. V., Lindahl, P. et al. (2019). NILM Dashboard: A Power System Monitor for Electromechanical Equipment Diagnostics. IEEE Transactions on Industrial Informatics, 15 (3), 1405–1414. https://doi.org/10.1109/tii.2018.2843770
20. Chen, Y.-Y., Chen, M.-H., Chang, C.-M., Chang, F.-S., Lin, Y.-H. (2021). A Smart Home Energy Management System Using Two-Stage Non-Intrusive Appliance Load Monitoring over Fog-Cloud Analytics Based on Tridium's Niagara Framework for Residential Demand-Side Management. Sensors, 21 (8), 2883. https://doi.org/10.3390/s21082883
21. Chen, C., Geng, G., Yu, H., Liu, Z., Jiang, Q. (2023). An End-Cloud Collaborated Framework for Transferable Non-Intrusive Load Monitoring. IEEE Transactions on Cloud Computing, 11 (2), 1157–1169. https://doi.org/10.1109/tcc.2021.3132929
22. Zhang, R., Wang, Y., Song, Y. (2022). Nonintrusive Load Monitoring Method Based on Color Encoding and Improved Twin Support Vector Machine. Frontiers in Energy Research, 10. https://doi.org/10.3389/fenrg.2022.906458
23. Chen, S., Zhao, B., Zhong, M., Luan, W., Yu, Y. (2023). Nonintrusive Load Monitoring Based on Self-Supervised Learning. IEEE Transactions on Instrumentation and Measurement, 72, 1–13. https://doi.org/10.1109/tim.2023.3246504
24. Li, W., Kong, D., Wu, J. (2017). A Novel Hybrid Model Based on Extreme Learning Machine, k-Nearest Neighbor Regression and Wavelet Denoising Applied to Short-Term Electric Load Forecasting. Energies, 10 (5), 694. https://doi.org/10.3390/en10050694
25. Hasan, I. J., Waheib, B. M., Jalil Salih, N. A., Abdulkhaleq, N. I. (2021). A global system for mobile communications-based electrical power consumption for a non-contact smart billing system. International Journal of Electrical and Computer Engineering (IJECE), 11 (6), 4659. https://doi.org/10.11591/ijece.v11i6.pp4659-4666