Advanced Engineering Informatics 52 (2022) 101556

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Full length article

Selection of features from power theories to compose NILM datasets

Wesley A. Souza [a,*], Augusto M.S. Alonso [b], Thais B. Bosco [c], Fernando D. Garcia [c], Flavio A.S. Gonçalves [c], Fernando P. Marafão [c]

[a] Department of Electrical Engineering, Federal University of Technology - Parana (UTFPR), Cornélio Procópio, PR, Brazil
[b] School of Electrical and Computer Engineering, University of Campinas (UNICAMP), Campinas, SP, Brazil
[c] Institute of Science and Technology, São Paulo State University (UNESP), Sorocaba, SP, Brazil

ARTICLE INFO

Keywords:
Load disaggregation
Nonintrusive load monitoring
Features quality
Smart meters
Electric consumption management

ABSTRACT

The load disaggregation concept is gaining attention due to the increasing need for optimized energy utilization and detailed characterization of electricity consumption profiles, especially through Nonintrusive Load Monitoring (NILM) approaches. This occurs since knowledge about individualized consumption per appliance allows to create strategies striving for energy savings, improvement of energy efficiency, and creating energy awareness to consumers. Moreover, by using feature extraction to devise energy disaggregation, one can achieve accurate identification of electric appliances. However, even though several literature works propose distinct features to be utilized, no consensus exists in the literature about the most appropriate set of features that ensure high accuracy on load disaggregation. Thus, beyond presenting a critical analysis of some significant features often selected in the literature, this paper proposes identifying the most relevant ones considering collinearity and machine learning algorithms. The results show that high-performance metrics can be achieved with fewer features than usually adopted in the literature. Moreover, it is demonstrated that the Conservative Power Theory can offer the most representative features for appliance identification, leading to efficient power consumption disaggregation.

1. Introduction

Electric power systems are undergoing infrastructure updates led by the smart grid trend. Among such changes, one fundamental transition is related to the substitution of electromechanical electricity meters for digital ones. The motivation for employing digital metering devices lies on their improved measurement accuracy, in addition to their superior processing, storage, and data communication capabilities [1]. Such digital devices, so-called smart meters, are capable of transmitting and receiving data for information, monitoring, management, and grid control. Consequently, they can complementarily support modern functionalities, beyond the basic energy consumption measurement, boosting energy services, and improving the relationship between consumers and utilities [2,3].

Particularly, the automatic and detailed information about electricity consumption, when attained by smart meters, devises visualization of power profiles via in-home displays (IHD), web pages, or mobile applications [1,4,5]. In fact, from a residential consumer's point of view, the visualization of power consumption profiles provides means to achieve more efficient management of electricity usage, such as in home energy management systems (HEMS) [6]. This allows a homeowner to improve usage habits through consumption awareness [7]. Nevertheless, adequate understanding of electricity consumption is cumbersome, especially when it comes to identifying the behavior of individual appliances. Hence, based on that, electrical load disaggregation techniques are seen as attractive alternatives to reach such an identification [8,9].

The electrical load disaggregation concept is explored in the literature according to three configurations: Intrusive Load Monitoring (ILM), Semi Intrusive Load Monitoring (sILM), and Nonintrusive Load Monitoring (NILM) [9,10]. ILM monitors energy usage via smart plugs that are attached to each appliance or device that needs monitoring. sILM uses smart devices at the electric circuit breakers to capture a subset of the electric energy usage. Finally, NILM is described by the principle of identifying loads based on the analysis of either macroscopic or microscopic parameters.

Particularly focusing on the NILM concept (i.e., microscopic aspect), operational load features are inferred from current and voltage waveforms, which are electrical measurements acquired at a high sampling rate. On the other hand, for the latter (i.e., macroscopic aspect), load features are analyzed by data attained at low sampling rates, typically based on measurements from the point of common coupling (PCC)

* Corresponding author.
E-mail address: wesleyangelino@utfpr.edu.br (W.A. Souza).

https://doi.org/10.1016/j.aei.2022.101556
Received 20 October 2021; Received in revised form 13 January 2022; Accepted 9 February 2022
Available online 24 February 2022
1474-0346/

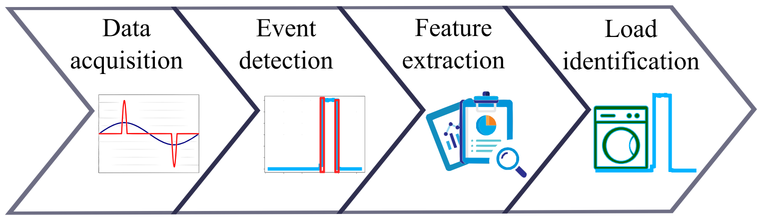

Fig. 1. Basic diagram of NILM approach stages.

of the residential electrical installation. Such a connection point is the electric branch in which a residence has access to the distribution grid [11]. Thus, in summary, NILM is a computational problem that dissociates aggregated electrical consumption into comprehensible household appliance electric behaviors.

Although interesting, due to each household installation and appliance's peculiarities, NILM is a non-trivial task. This occurs since a method for signal and event analysis must be constructed to characterize and identify appliances [12,13]. Several works in the literature propose the use of a variety of features for this task, presenting different performance capabilities. For instance, research proposals have presented the analysis of harmonic decomposition, power theories, statistical indicators, and effective values of electrical signals, among others [13,14].

Based on the importance of properly selecting the features for NILM approaches, this paper presents a proposal for a more effective selection of the most representative features employed on the characterization and classification of household appliances. By using collinearity analysis, the proposed approach is able to reduce the quantity of features considered for the NILM, while minimizing the dataset features to only the most relevant ones. In this paper, it is suggested for future research on NILM which features best represent the analyzed appliances, while performing load identification. Finally, another contribution in this paper is the extraction of features from the Conservative Power Theory (CPT) [15] and the IEEE 1459–2010 standard [16], since that is nonexistent in state-of-the-art studies.

This paper is divided into five sections. Section 2 presents an overview of the literature with the most relevant features from NILM studies and power theories. Section 3 presents the collinearity approach for the selection of the most representative features. Section 4 presents the results and discussions, and Section 5 brings the conclusions.

2. State of art

The pioneering work of Hart [17] indicated four stages of the NILM approach: signal acquisition, event detection, feature extraction, and load identification, as shown in Fig. 1. All of them are described as follows.

Data acquisition is the stage in which electrical signals are converted into digital values for processing and calculating electrical quantities. **Event detection** corresponds to the load state identification (when the load is turned on or off). The **feature extraction** stage converts the collected signals (current and voltage waveforms) to a set of features (e.g., active power, reactive power, current harmonics, among others) used to create parameters for the load recognition. Finally, the **load identification** stage is performed through pattern recognition algorithms to identify the load that has been turned on or off.

After almost two decades of little interest in NILM by the academic community, the evolution of acquisition devices and artificial intelligence techniques brought a new reality to this research field, and NILM approaches began to be improved. In this context, machine learning methods are seen as attractive tools to perform load identification, as depicted in Fig. 2. To take advantage of such intelligent algorithms, it is necessary a dataset with information about load characteristics, since it helps on the creation of a classification model for a proper load identification. In addition, the quality of the adopted features becomes a determining item for adequate accuracy in load identification.

Moreover, training supervised models with datasets comprising many features frequently requires a high computational cost and processing time, which prevents their application in dynamic scenarios or embedded solutions. Thus, dimensionality reduction or feature selection techniques are required, both to remove unnecessary attributes and to improve the load identification model.

Thus, many research efforts [9,18–32] have focused on feature extraction to describe and distinguish one appliance from the others, while focusing on features automatically extracted by algorithms and selected using calculations from the physical or quantitative analysis of load phenomena. Features from standards and electrical studies assist in performing the load disaggregation and expanding functionalities to the meter, such as improvements in energy efficiency and possible malfunctions of electrical appliances [33].

Some works from the literature use Principal Component Analysis (PCA) [34,35], artificial neural networks, and deep learning for the automatic feature identification capable of distinguishing appliances [36–41], creating features by means of waveform differentiation between appliances. Although such researches present quite significant results, the extracted features do not necessarily correspond to the appliances' physical phenomena.

Hence, this paper focuses on features that represent the electrical and physical phenomena of the appliances, providing means for their future use in systems related to fault detection, power quality interventions, and energy efficiency improvement. As a result of this perspective, using features from power theories can devise a flexible alternative to integrate load disaggregation and a smart energy manager, shedding light on a new generation of smart meters [42,43]. The following subsection discusses the power theory attributes used in the literature.

2.1. Features from the NILM literature

Sadeghianpourhamami et al. [13] and Souza et al. [14] presented an extensive list of relevant works and raised features that can be used for the feature extraction towards an accurate appliance classification. In both works, the analysis of performance metrics was studied by using several features from the literature. In Sadeghianpourhamami et al. [13], the accuracy is above 90%, and in Souza et al. [14], the accuracy is above 98%. In the comprehensive literature search of both works, the principal features are:

- Power components: The active (i.e., resulting in useful work), reactive and apparent powers are used in several studies for event detection and load identification. Such features are the most used features in the literature, having as examples the studies carried out in [18–22];
- Voltage and current effective values (RMS): correspond to the square root of the mean of the squared value of the voltage or current signal during a complete cycle and widely used in the literature, having as examples the studies done in works [37,44];

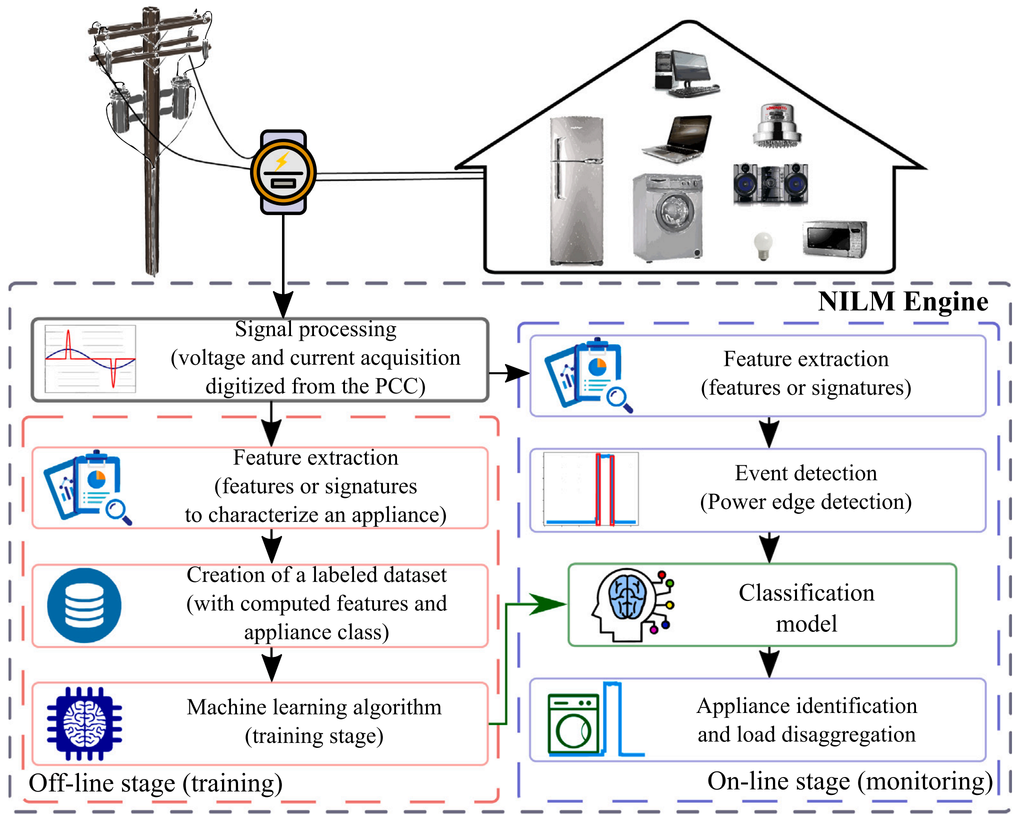

Fig. 2. Diagram of the NILM approach with consideration of intelligent algorithms.

- Power factor: corresponds to the ratio of active power over apparent power. This factor can be used to indicate the possible angle displacement between the voltage and current signals. Examples are found in studies, such as [23,24,31];
- Odd harmonic current components: the effective values of the harmonic components, according to the multiple periodic signals of the fundamental frequency extracted through the Fourier series. For example, the 3rd-order current harmonic corresponds to the 180 Hz harmonic current's effective value if considered the fundamental grid frequency of 60 Hz. The studies made in [25,26,45] are examples of the use of these features for the NILM system;
- Harmonic distortion of voltage and current: it is the ratio given by the sum of the non-fundamental harmonic components over the fundamental harmonic component (harmonic order equivalent to (1). This quantity provides means to quantify the harmonic content in a periodic signal. The works presented in [27,46,47] are examples that use harmonic distortion;
- VI Trajectory: corresponds to the graphical information of the voltage value by the current in a cycle. Such information provides a graphical analysis of the characteristic value of the current drawn by a load. Works like [48–50] use the VI trajectory as information for feature extraction;
- Load behavior information: is the characteristic information or statistics of power or current of loads, such as the maximum value and average value. For example, there are studies made in [27,37,51].

Recent works in the literature employ neural networks to attain derived values (i.e., features) intending to bring informative content and being non-redundant, facilitating the subsequent learning and generalization steps, also leading to better human interpretations. Additionally, such researches relate feature extraction to dimensionality reduction. As examples, Le et al. [39,41,52] presented the outstanding accuracy of 95.04% using the Hilbert Transform Long Short-Term Memory (HT-LSTM), the average accuracy of 92.96% using the Gated Recurrent Unit (GRU), and the accuracy of 92.04% using the Generative Adversarial Network (GAN).

Many other features can be found in the literature, but the above-listed concepts are the ones that best indicate the physical characteristics of appliances, which is the premise of this paper. Recently, it has been explored in the literature the use of power theories for feature extraction applied to load identification, as highlighted in the next section.

2.2. Features from power theories

Since 1865, starting with Maxwell [53], the study of feature extractions based on electrical circuit observations has been advancing in a historical context. Later on, modern power theories have been created to further understand the behavior of loads and electric power systems [23,54].

Electrical quantities calculated by power theories can quantify: (i) the angle deviation between voltages and currents, caused by elements with reactive properties; (ii) the power consumed by non-linear elements of loads; (iii) the useful power that generates work; (iv) the total power that considers the useful and non-useful work; (v) the levels of load unbalance; (vi) power quality indexes; and many others [15,16,55,56].

The study of power theories also supported the development of the IEEE standard 1459–2010 [16], which allows one to define electrical quantities under sinusoidal, non-sinusoidal, balanced and unbalanced power grid conditions. Such a standard presents broad definitions of well-defined concepts in the literature to indicate the active, reactive, non-active, fundamental, apparent, and distortion power parcels. In addition, this standard was created to be a reference for designing measurement instruments to quantify electricity and power [16] consumption.

Table 1
Comparison of feature selection methods [59–61].

Feature selection methods			
	Filter	Wrapper	Intrinsic/Embedded
Meaning	It uses statistical techniques to evaluate the relationship between each input variable and the target variable and these scores are used as the basis to choose (filter) those input variables that will be used in the model.	It creates many models with different input features and select those that result in the best performing model according to a performance metric.	It performs feature selection automatically as part of learning the model.
Advantages	The computational cost is lower than the other methods, fastest running time, lower risk of overfitting, ability of good generalization, the independence of the classifier can also be an advantage.	Interacts with the classifier for feature selection, considers feature dependencies, better generalization than the filter methods, model feature dependencies.	Less computationally intensive as compared with the wrapper methods, lower risk of overfitting, model feature dependencies.
Disadvantages	No interaction with classification model for feature selection, mostly ignores feature dependencies and considers each feature separately in case of univariate techniques, which may lead to low computational performance to wrapper or embedded.	High computational cost, longer running time, no guarantee of optimality of the solution, higher risk of overfitting when is compared with the filter and embedded.	May be problematic when use in small set of features, the classifier dependent selection can be a disadvantage.
Algorithms	Correlation/Colinearity, Chi-Square Test, ANOVA, information gain, ReliefF	Forward selection, Backward elimination, Stepwise selection, Sequential Search, Genetic algorithm, Recursive feature elimination.	LASSO, Elastic Net, Ridge Regression, Random Forest, Spatial SVM, Memetic, Decision tree

One of the most recently-developed power theories is the so-called Conservative Power Theory (CPT), which was proposed by Tenti et al. [15]. The CPT is an approach that provides decomposition of power signals in the time domain, being valid for single and poly-phase systems (i.e., with or without neutral conductor). Another advantage is that the CPT is applied to sinusoidal or non-sinusoidal conditions [54]. The CPT offers an orthogonal decomposition of current and power in terms directly related to the electrical characteristics of the load under analysis at the PCC. Moreover, the CPT power decomposition results in characterization of power components that interpret existence of active, reactive, unbalance, and void (i.e., related to nonlinearities) behaviors. Studies carried out by Souza et al. [31,57] present interesting results while using the CPT, as a mathematical basis, for the feature extraction during load disaggregation.

2.3. Feature selection

Feature selection is a process that reduces the number of features to decrease the computational cost of modeling and, when applicable, improve the model's performance. Methods related to statistical-based feature selection evaluate the relationship among input variables and use statistics to select the ones with the most robust relationship with the target variable. These methods can be fast and effective, although statistical measures depend on the data type of both the input and output variables. Moreover, attribute selection reduces the size of the dataset and supports data interpretation, reducing computational requirements, and improving model performance [58]. In general, feature selection may be divided into the wrapper, filter, and intrinsic methods, having their specifications and methodology explained in Table 1.

Regarding NILM approaches, Akarslan and Doğan [62] used the Total Harmonic Distortion (THD), the angles and amplitudes of the 1st, 3rd, 5th, 7th, and 9th of the current signal as features, and used the ReliefF method to determine the importance of each feature. In such a research, the authors obtained an accuracy of 97.5% using the Elman neural networks as classifiers and using the five most relevant features: the current harmonic amplitude of the 1st, 3rd, 5th, 7th, and the angle of the 3rd harmonic. Valencia-Duque, Álvarez Meza and Orozco-Gutiérrez [63] also used the ReliefF, which reduced the initial 17 features to only 4 and reached an accuracy of 98.95%.

Sadeghianpourhamami et al. [13] use a typical feature extraction from the literature based on recursive feature elimination (RFE). The authors described the presence of high-correlated features and, due to that reason, the RFE is said to not be a good approach. Later on, the authors used the most relevant features and the initial number of 78 features was reduced to 58, 30, 20, and 10 elements. The best accuracy of 93.2% was obtained with the use of the 20 most important features.

The inclusion and reinforcement of power theories as a useful tool for identifying loads is another important issue. Houdi et al. [64] performed the reduction of features deduced from voltage and current measurements, targeting applications in a digital system using a low-cost processor with a low storage capacity. From ninety features derived from the IEEE 1459–2010 [16] standard, RMS voltage, RMS current, and the apparent, reactive, distortion, and active powers, the authors compared the success rate of the whole feature set with: (i) the heuristic search, (ii) Inertia Ratio Maximization using Features Space Projection (IRMFSP), (iii) Principal Component Analysis (PCA), and (iv) the use of only the P and Q features. The best result obtained was the heuristic search considering 36 features and an accuracy of 99.4%, whereas the PCA presented 16 remaining features and 97.4% of accuracy. Cannas et al. [6] also extracted features from IEEE 1459–2010 and applied the neighborhood component analysis (NCA) as feature selection. In such a work, the authors used the 8 most relevant features and reached an f-score of 94.0%.

In feature selection, the goal is to choose a subset of features that provides better distinctive performance than when using all of them. Moreover, it is worth highlighting that feature selection also should be distinguished from the feature extraction process. This latter creates new features from transformations of the original features into low-dimensional data. For instance, in the NILM literature, some research efforts use PCA [31,64], Linear Discriminant Analysis (LDA) [65], Autoencoder [66], and t-distributed stochastic neighbor embedding (t-SNE) [67].

Finally, the literature related to attribute extraction does not compare power theories to determine which one best represents a load. On the other hand, this paper complementarily incorporates such a contribution. Furthermore, with regards to feature selection, the NILM literature does not cover the correlation-based selection method. Therefore, Section 3 presents a proposed feature selection method based on collinearity, as well as the appointed features listed as best indicators for the desired NILM task.

3. Collinearity-based feature selection approach

Feature selection is the identification and selection of the most useful characteristics in a dataset [68]. It is a critical step in the machine learning process because unnecessary features can: (i) increase computational time in the stages of training and classification; (ii) decrease the model's interpretability; and (iii) decrease the generalization of the machine learning algorithm [69]. Feature reduction is considered

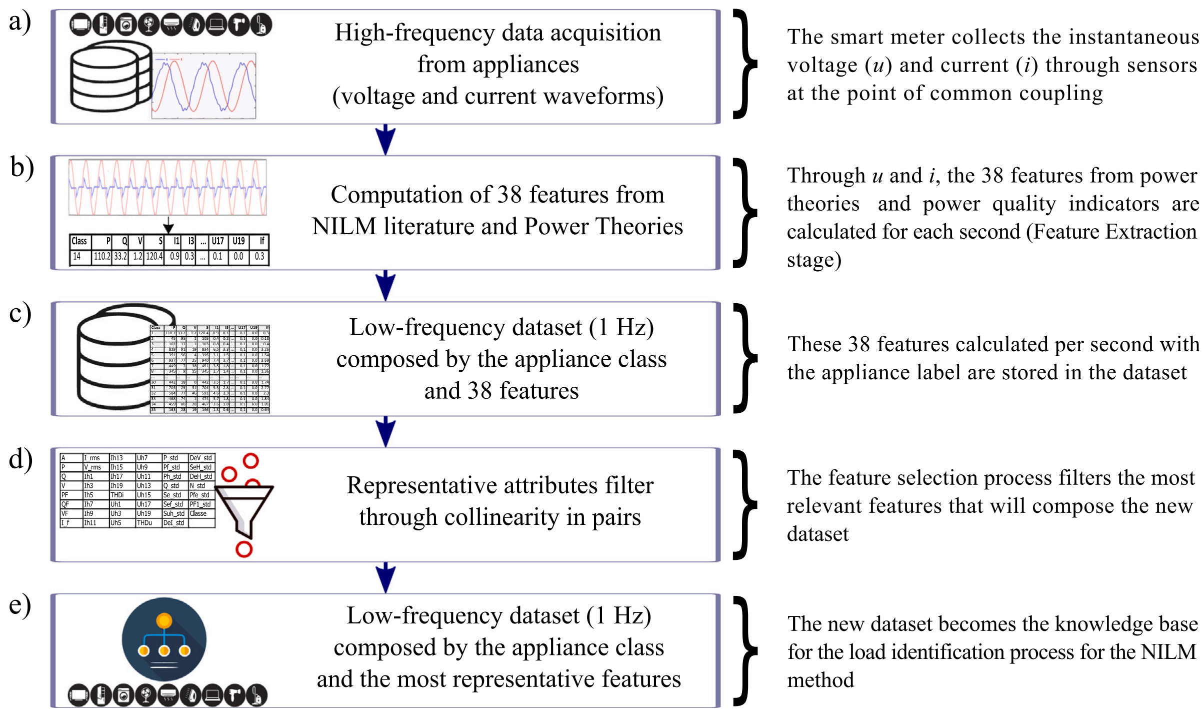

Fig. 3. Diagram of the proposed feature selection approach.

an essential tool to decrease the computational burden caused by large data aggregation (e.g., like BigData) [70].

Taking into account the categories of features presented in Section 2, the most relevant ones for the load identification are selected according to Fig. 3. Therein, the five steps of the proposed collinearity-based method are described for the process of feature selection. Complimentary explanation of each of these steps is given as follows.

(a) **High-frequency data acquisition:** In this step, the raw data (voltage and current waveforms of a given load) is collected through a meter or from a high-frequency NILM dataset [71]. At this stage, the instantaneous voltage (u) and current (i) are gathered through sensors at the PCC;
(b) **Feature extraction — Computation of 38 features from the NILM literature:** 38 features presented in Table 2 are extracted once at every second of raw data from the previous step. The importance of using power theories and PQ indicators is that the features can perform the load analysis and characterization. The power terms can describe the physical phenomena of the appliance. The harmonic components help in designing a fingerprint of the interaction between source and load: the voltage harmonic quantities can describe the source interaction, and the current harmonic quantities can describe load power consumption interaction.
(c) **Low-frequency dataset (1 Hz) composed of the appliance class and 38 features:** The calculated values of the 38 features and the appliance identification (class) are calculated once at each second. The values of these features, combined with the true label of the measured appliance, represent one sample of a dataset derived from the high-frequency data acquisition from the Process (a).
(d) **Representative feature filter considering collinearity in pairs:** With the dataset created in the previous step, the features' relevance in the classification process is analyzed, removing those that do not interfere in the performance metrics of the machine learning algorithms. The method used for the feature reduction in this study is called collinearity, which is a filter method of feature selection, and it is used in different areas of applications [72–74].
In statistics, collinearity is the phenomenon in which one feature is highly correlated to another. From an analytical perspective, when a high correlation between pairs of features exists, there is a behavior approximately linear between them, as shown in Fig. 4.
Considering the samples plotted pair-wise as depicted in Fig. 4, the pairs of features (I_Q, Q) and (I_V, V) are highly correlated, and it is possible to note the linear behavior and dependence between them, and one of such features can be removed. If there is a high correlation between the two features, there is similar information and the possibility of data redundancy. Another phenomenon that happens is overfitting, which corresponds to the optimal adjustment of the classification model to the training data, but the classification model becomes ineffective when new data are classified [75]. Features with a correlation greater than 0.8 are considered highly dependent [76]. Removing one feature of highly correlated features reduces training and classification time and could increase the machine learning methods' assertiveness [70]. In this approach, the parameter to select the features is the value of the correlation and the removed feature from the correlated pair is the feature with more correlation with other features.
(e) **Low-frequency dataset (1 Hz) composed of appliance identification and the most representative features:** A dataset derived from step (c) is the ending process of the proposed method, keeping the appliance identification and containing only the most representative features. The computational time in the execution of all NILM steps is reduced in this derived dataset and consequently decreasing the meter's processing charge. Another point is the reduction of the dataset size, making it smaller and more representative.

Based on the above-mentioned steps, Section 4 presents the experimental procedure and results of the collinearity-based approach. In addition, it is highlighted that two public NILM Datasets are used: NILMbr [14] and PLAID (Plug-Level Appliance Identification Dataset) [71].

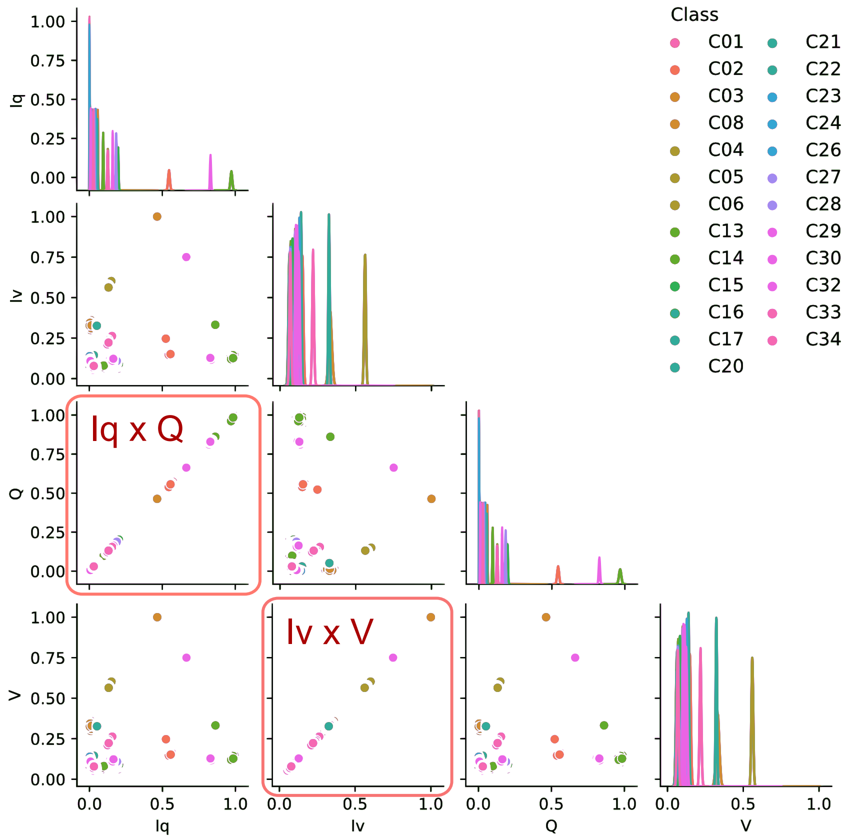

Fig. 4. Collinearity between features example.

Table 2
Features used in relevant studies from NILM literature and power theories.

Features	From	Description	References
A		Apparent power	[18,24]
P		Active power	[17,20,22]
Q		Reactive power	[21,22]
V	CPT	Void power	[31,57]
I_a	CPT	Apparent current	[15]
I_p	CPT	Active current	[15]
I_q	CPT	Reactive current	[15]
I_v	CPT	Void current	[15]
U_{RMS}		RMS Voltage	[44]
I_{RMS}		RMS Current	[37,44]
$PF(\lambda)$		Power factor	[23,24,31]
$QF(\lambda_Q)$	CPT	Reactivity factor	[31]
$VF(\lambda_D)$	CPT	Distortion factor	[31]
$I_{1,3,5,\ldots,19}$		Current harmonic components	[25,26,45]
THD_U		Voltage total harmonic distortion	[27]
THD_I		Current total harmonic distortion	[27,46,47]
P_1	std	Fundamental active power	[16]
P_h	std	Harmonic active power	[16]
Q_1	std	Fundamental reactive power	[16]
S_1	std	Apparent power	[16]
S_h	std	Harmonic apparent power	[16]
S_n	std	Nonfundamental apparent power	[16]
D_I	std	Current distortion power	[16]
D_V	std	Voltage distortion power	[16]
D_h	std	Harmonic distortion power	[16]
N	std	Nonactive power	[16]
PF_1	std	Fundamental power factor	[16]

4. Results and discussion

In this paper, the main goal of the feature selection process is to point the most representative features and guarantee the prediction power of the classifier algorithms for the NILM approach.

4.1. Experimental procedure

Each step of Fig. 3 is executed in the experimental procedure. Initially, a high-frequency dataset is required, as shown in Fig. 3(a). It is possible to find several bases in the literature, such as NILMbr [14], PLAID [71], COOLL [77], LIT-Dataset [78], REDD [79], BLUED [80], UK-DALE [81], WHITED [82] and EnerTalk [83]. Two high-frequency datasets were selected – NILMbr and PLAID – for the assessment of the results. Both datasets were converted into `csv` files containing acquisition data of voltage and current. The NILMbr dataset contains the sampling rate of 15.36 kHz and data gathered from 25 appliances, while the PLAID dataset contains the sampling rate of 30 kHz and data gathered from 11 appliances, as presented in Table 3.

The dataset is handled by a feature extractor algorithm developed in `Python` 3 language. The 38 features (shown in Table 2) are calculated for the amount of data corresponding to one second of high-frequency samples (15,360 and 30,000 samples for NILMbr and PLAID, respectively). Such features with the load identification become an instance inserted into the low-frequency dataset. The low-frequency dataset allows the performance analysis of supervised classification algorithms to be used in the load identification of NILM approaches.

After removing inconsistent data (*e.g.*, read errors and obvious outliers), is was normalized the features of the low-frequency dataset using the Z-score normalization [84,85]. To check if the set of features offers a good computational representation of the loads, we have used them to train three well-known machine learning algorithms: k-nearest neighbors (k-NN) [86], decision tree (DT) [87], and random forests (RF) [88]. To evaluate the results, the traditional 10-fold cross-validation [89,90] was used. It is reinforced that a methodological comparison among the k-NN, DT and RF algorithms is out of scope in this paper, knowing that this can be found in literature [91]. Herein such different algorithms are only used as means to achieve an unbiased assessment of features for NILM applications.

Table 3
Appliances represented in each dataset.

Household appliance	NILMbr	PLAID
Incandescent lamp	X	X
Air conditioner	X	X
Refrigerator	X	X
Microwave	X	X
CRT TV	X	
LCD TV	X	
Electrical shower	X	
Iron	X	
Washing machine	X	X
Hairdryer	X	X
Fluorescent lamp	X	X
Mix lamp	X	
ASD Dryer	X	
ASD Fridge	X	X
Blender	X	
Bread maker	X	
Desktop PC	X	
Food processor	X	
Freezer	X	
Furnace	X	
Garage door	X	
Laptop	X	X
Regular dryer	X	
Regular fridge	X	X
Vacuum	X	X
Fan		X
Heater		X

The accuracy and macro f-measure are considered for the evaluation of the classification algorithms performance. Accuracy is a metric used to predict the correctness of a classifier model. The accuracy is equal to the number of correctly predicted data points out of all the data points used in the test.

F-measure is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but f-measure is usually more useful than accuracy, especially if it has an uneven class distribution.

Additionally, the average training and testing time were collected in results with a 10-fold cross-validation. The tests were performed on a computer with Ubuntu Linux OS, i7-8750H processor, 16 GB of RAM. The experiments were implemented using `Python 3` with the `sklearn` library [92].

The classification algorithms are performed on the 38 features dataset and for every derived dataset created after the collinearity analysis. This method aims to evaluate if the feature elimination does not interfere adversely in the classification algorithms' performance. In the collinearity analysis, the correlation between the features is calculated and compared pairwise. For each pair with the correlation value greater than the cutoff, one of the features is removed.

Sections 4.2 and 4.3 show the collinearity approach results for the NILMbr and PLAID datasets.

4.2. NILMbr dataset

A programmable meter has been used to implement the NILM system presented in [57]. The voltage and current waveforms of each 25 appliances were collected for 120 min of operation, with a sampling frequency of 15.36 kHz (256 samples per cycle). Subsequently, the 38 features were extracted for each second. A dataset with 191,308 samples, 38 features, and the class was created in the final, and it is available in [https://bit.ly/334na2g].

After data treatment (normalization, missing data verification and duplicate instances removal), the performance indices of three traditional classification methods were analyzed. The cross-validation technique was used through the traditional k-fold stratified validation, with $k = 10$.

Table 4
Results achieved by the classifiers using the 38 features from the literature.

Measure	DT	RF ($n = 20$)	k-NN ($k = 3$)
Accuracy	0.8463	0.9228	**0.9311**
F-measure (macro)	0.8032	0.8998	**0.9339**
t_{train} (s)	4.0588	5.8875	–
t_{test} (s)	**0.1143**	0.2214	11.3982

Table 5
Results achieved by the classifiers using the 11 features filtered by the collinearity to represent the appliances.

Measure	DT	RF ($n = 10$)	k-NN ($k = 3$)
Accuracy	0.9375	0.9464	**0.9537**
F-measure (macro)	0.9233	0.9380	**0.9561**
t_{train} (s)	0.9313	1.1129	–
t_{test} (s)	**0.0553**	0.1023	2.2388

Table 6
Results achieved by the classifiers using the 9 features filtered by the collinearity to represent the appliances.

Measure	DT	RF ($n = 10$)	k-NN ($k = 1$)
Accuracy	0.9573	**0.9641**	0.9635
F-measure (macro)	0.9433	**0.9644**	0.9639
t_{train}(s)	0.5409	0.7340	–
t_{test}(s)	**0.0183**	0.0645	0.7355

The scoring methods were analyzed: k-NN with grid search for best parametrization $k = 1, 3, 5, 7$; RF with grid search $n = 10, 20, 30, \dots, 90, 100$ and DT. The accuracy values, macro f-measure, training, and testing times of the classification algorithms applied in the dataset with the 38 features are presented in Table 4. The best performance was the k-NN with $k = 3$, with an accuracy of 93.11% and the macro f-measure of 93.39%, but the test time is 50 times greater than the second best in accuracy, the RF.

Subsequently, the collinearity is applied to the dataset. The correlation between pairs of features is presented in Fig. 5. Several pairs of features with high correlation values can be identified, such as: active current (I_p) with apparent current (I_a), apparent power (A), active power (P), effective current (I_{RMS}), fundamental active power (P_1) and fundamental apparent power (S_1). For such a case, only one feature has been maintained.

After removing features with collinearity above 0.8, there are 11 remaining features, as shown in Fig. 6. The classification performance metrics were calculated with the remaining 11 features. Realizing grid search for the best parameters, the k-NN, RF and DT algorithms were executed, having results depicted in Table 5.

For the 11 features dataset, the accuracy and macro f-measure have been improved as well as the training and testing times are reduced. The improvement of the metrics' value is caused by the phenomenon known as "Curse of Dimensionality" [68,93], which can occur in the data analysis in multi-dimensional spaces (features). This phenomenon indicates that adding characteristics does not necessarily improve the performance of a classification algorithm. Therefore, it is reinforced the necessity to identify the quality of features to describe the data accurately and improve the classification algorithms' results.

After the substantial feature reduction, the collinearity analysis with the value of 0.5 was performed to reduce the number of features without interfering in the performance metrics. Two features have been removed and 9 remain, as shown in Fig. 7. The classification algorithms were performed with a grid search for the best parameters, with the results presented in Table 6.

The performance metrics have been improved, and the best performance classification algorithm is the RF, before it was the k-NN. The accuracy values in this scenario are close to the best values in the NILM literature that ranges from 70% to 98% [94].

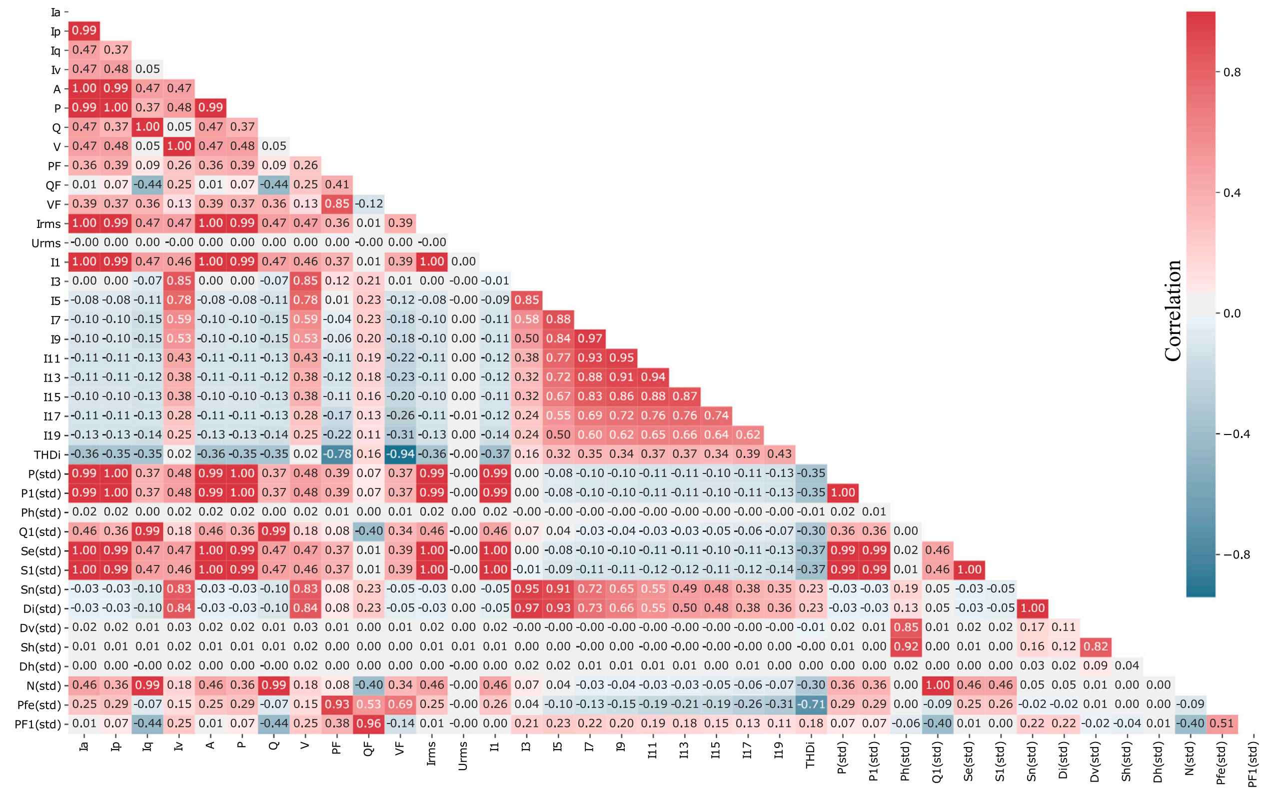

Fig. 5. Correlation matrix of the 38 features dataset.

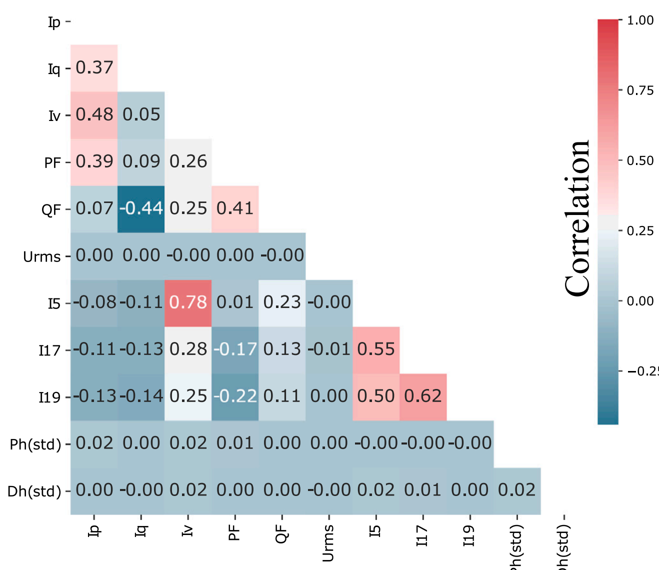

Fig. 6. Correlation matrix of the 11 remaining features with a correlation lower than 0.8.

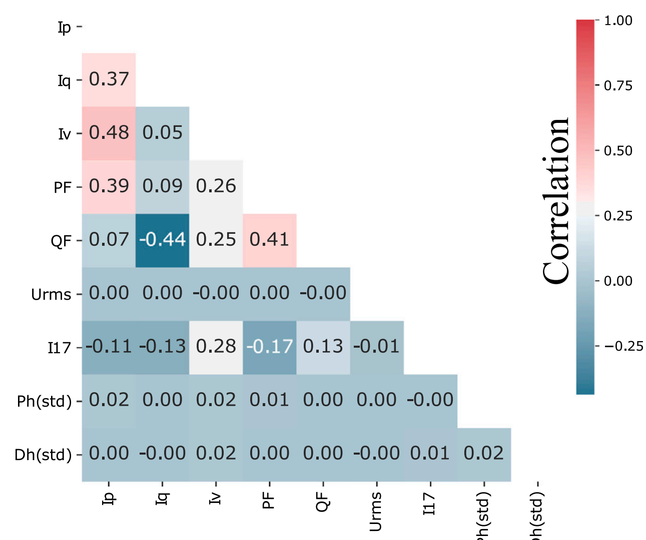

Fig. 7. Correlation matrix of the 9 remaining features with a correlation lower than 0.5.

The relevance of each feature was analyzed using the RF method, as shown in Fig. 8. The features U_{RMS}, P_h, D_h e I_{17} have low relevance in the ranking, being therefore removed in the last step. The correlation between the remaining features is shown in Fig. 9. The classification algorithms were performed again, with the results presented in Table 7. There was an improvement in the performance metrics for the RF and k-NN algorithms and a worsening for the DT algorithm. The remaining features for the NILMbr dataset are I_p, I_q, I_v, PF, and QF, all from CPT. Consequently, CPT is an excellent tool for feature extraction to the characterization of equipment in NILM systems.

Table 7
Results achieved by the classifiers using the 4 final features to represent the appliances.

Measure	DT	RF	k-NN
Accuracy	0.9545	0.9782	**0.9874**
F-measure (macro)	0.9417	0.9782	**0.9874**
t_{train}(s)	0.4234	0.5349	–
t_{test}(s)	**0.0134**	0.0486	0.5709

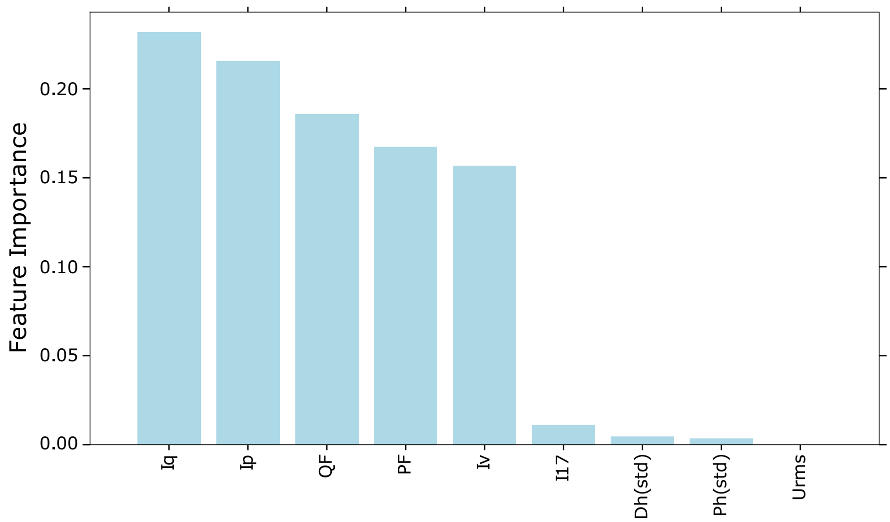

Fig. 8. Relevance of features for the RF algorithm.

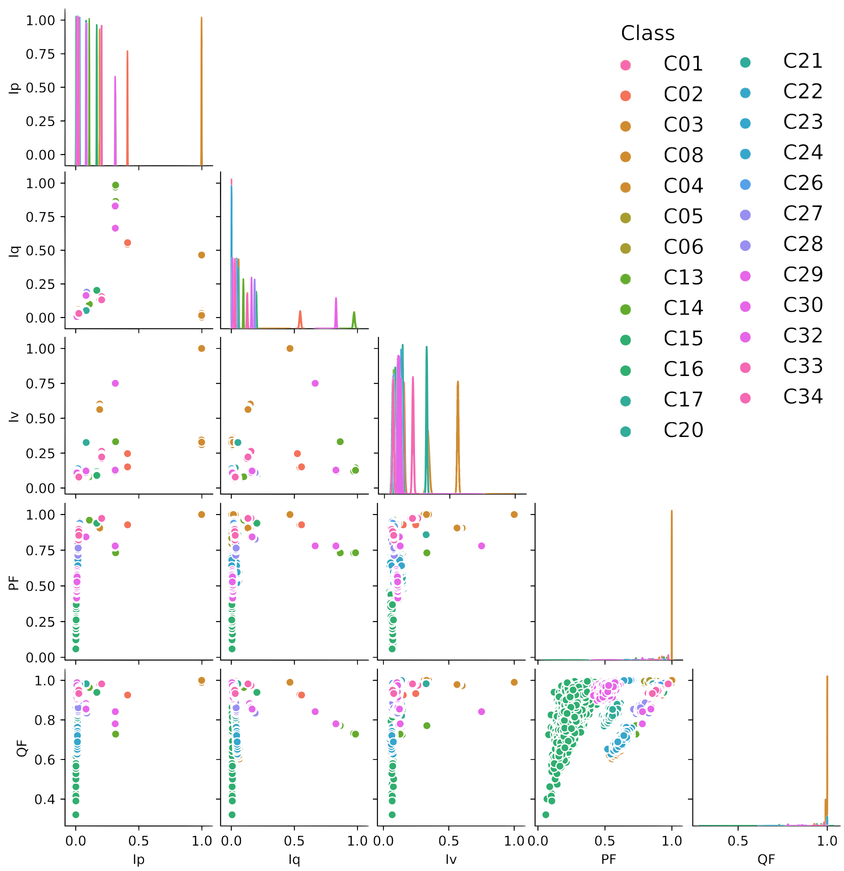

Fig. 9. Collinearity of the 5 remaining features of the NILMbr dataset.

In the next Section, the PLAID dataset was analyzed to check if the selected features are the same in the collinearity approach presented on the NILMbr dataset.

4.3. PLAID dataset

The same procedure was done to collect the 38 features presented in Section 4.2 and Fig. 3. After data treatment, the performance indices of RF, k-NN, and DT were analyzed, with k-fold stratified used with $k = 10$. After removing features with collinearity greater than 0.5, the remaining features were I_p, I_q, I_v, and PF. Fig. 10 presents the correlation of the remaining features to the PLAID dataset. Fig. 11 shows the collinearity between the selected features, and Table 8 presents the evolution of performance measures for each feature removal process. For the PLAID dataset, the four remaining features are I_p, I_q, I_v, and PF from the CPT. It reinforces the CPT as a mathematical framework that allows a proper feature extraction for the load classification.

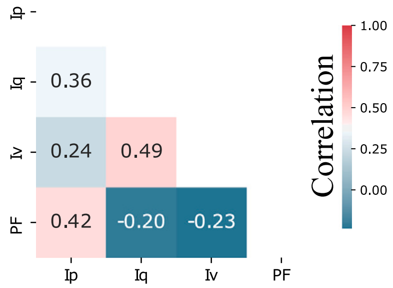

Fig. 10. Correlation matrix of the 4 remaining features with a correlation lower than 0.5.

Table 8
Results achieved for the PLAID dataset using the 4 final features to represent the appliances.

Number of features	Measure	DT	RF	k-NN
38	Accuracy	0.9155	**0.9235**	0.9246
	F-measure	0.9008	**0.9179**	0.9168
	t_{train}(s)	2.5272	3.4119	–
	t_{test}(s)	0.1037	**0.0140**	10.4131
10 ($corr(x,y) < 0.8$)	Accuracy	0.9325	0.9707	**0.9859**
	F-measure	0.9275	0.9596	**0.9672**
	t_{train}(s)	1.1898	0.8836	–
	t_{test}(s)	**0.0070**	0.0509	1.5678
4 ($corr(x,y) < 0.5$)	Accuracy	0.9330	0.9756	**0.9865**
	F-measure	0.9278	0.9659	**0.9695**
	t_{train}(s)	0.9795	4.1669	–
	t_{test}(s)	**0.0069**	0.2044	0.5510

4.4. Comparison with others feature selection approaches

The reached accuracy of 98.74% presented in Sections 4.2 and 4.3 are aligned with the best results found in the literature [13,31,39]. However, it is necessary to verify if the feature selection is adequate for the NILM datasets and, therefore, three other well-known algorithms from the literature were used as comparison benchmarks:

- **Wrapper method — Recursive feature elimination**: Given an external estimator that assigns weights to features, the recursive feature elimination (RFE) aims to select features by recursively considering smaller and smaller sets of features [95]. First, the estimator is trained on the initial set of features, and the importance of each feature is obtained by means of any specific attribute. Then, the least important features are pruned from the current set of features. Finally, that procedure is recursively repeated on the pruned set until the desired number of selected features is eventually reached [96].
- **Filter method — Select k-Best**: The Select k-Best (SKB) is a univariate feature selection that selects the best features based on univariate statistical tests. It can be seen as a preprocessing step for an estimator. The transform method Select k-Best generates a direct search function that takes the k features with the most outstanding individual evaluations [97]. The SKB is handled herein according to the χ^2 (chi-square) test, which measures the dependence between stochastic variables. Hence, by using this function one can eliminate the features that are the most likely to be class independent, therefore, being irrelevant for the classification.
- **Intrinsic method — Tree-based estimator**: Tree-based estimators (TBE) are meta-transformers that can be used alongside any estimator that assigns importance to each feature through a specific attribute or fitting. The features are considered unimportant and removed if the corresponding importance of the feature values is below the provided threshold parameter. Apart from specifying the threshold numerically, there are built-in heuristics for finding a threshold using a string argument [98]. The tree-based estimator can compute impurity-based feature importances, discarding irrelevant features. In this benchmark, the Extra Trees Classifier is used for the estimator.

Table 9 depicts the selected attributes by the collinearity method, as well as by the RFE, SKB, and TBE methods. For the SKB, k (remaining features) was parametrized to 11, 9, and 5. For RFE, the features selected were configured to 11, 9, and 5. Finally, for TBE, the number of trees estimator was set to 50, 30, and 10.

The features related to the current terms from the CPT remained in the SKB, TBE, and collinearity approaches. The CPT power terms (which result from the effective value of the current times voltage) appear in the RFE method. Therefore, the CPT power decomposition terms are relevant for all approaches and techniques. Among the selected features, the 5th, 7th, 17th, and 19th harmonic currents are relevant terms, as well as the effective current and the IEEE std 1459 power parcels. However, when reducing the number of attributes, the CPT parcels remain, while the IEEE std 1459 components are elected to be removed from the selected attributes group. Thus, the collinearity process determined significant attributes for selection, and the CPT is ensured as a suitable power theory to be used as an attribute extractor for NILM methods.

To further present comparative discussions among the methods, the k-NN with grid search (for best parametrization $k = 1, 3, 5, 7$), the RF with grid search ($n = 10, 20, 30, \dots, 90, 100$), and the DT approaches are implemented for each dataset from the feature selection depicted in Table 9. Additionally, the accuracy values are presented in Table 10.

According to Table 10, the TBE with 10 estimators gives the best results. Thus, the suggestion is to use the features I_a, I_p, I_q, I_v, QF, VF, I_5, I_7, and S_1 for the composition of the NILMbr dataset. Another statement relates to the fact that, from the 9 features within the best result, 6 are CPT-derived and only 1 is derived from the IEEE 1459–2010. Moreover, the last aspect noticed is that the collinearity method presents results similar to the conventional techniques from the literature, even though it is devised by a simple filter method.

4.5. Results analysis and discussion

The collinearity presented in this article had the CPT as an extractor of characteristics that can better represent an appliance for both datasets. According to Tenti et al. [15], the CPT helps in load characterization, in terms that represent resistive, inductive/capacitive and non-linear characteristics of loads and is an approach that is applied in sinusoidal or nonsinusoidal conditions in the power grid, thus becoming a mathematical framework for NILM approaches.

Considering only the CPT features, there was more than 98.6% accuracy for both datasets, and the time for training and classification was considerably reduced because there were less data to be performed by the classification algorithms. The overall average accuracy of 98.6% for NILMbr and PLAID 1 datasets is aligned with the best results available in the NILM literature, as presented in Table 11. It is a remarkable performance because with few features it is possible to reach high accuracy for the load disaggregation.

5. Conclusions

This paper presented the feature engineering process for NILM approaches, offering relevant features that can perform accurate load identification. To demonstrate the particularities and contribution of the collinearity-based approach, it was performed the process of choosing 38 features found in the NILM literature, standards, and power theories. A low-frequency dataset with such 38 features was built derived from high-frequency data (voltage and current waveforms of

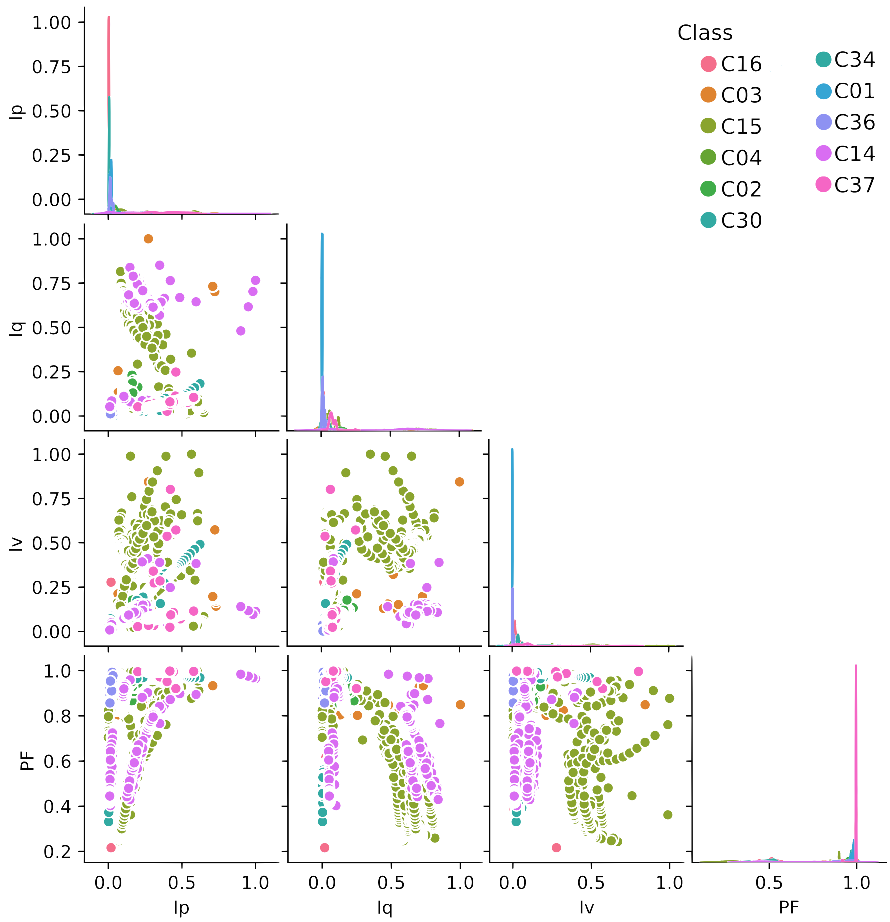

Fig. 11. Collinearity of the 4 remaining features of the PLAID dataset.

Table 9
Comparison of the feature selection approaches and the remaining features for the NILMbr dataset.

Approach	Parameters	Number of features	Remaining features
SKB	k = 11	11	I_a, I_p, I_v, I_q, I_{RMS}, I_5, P_1, Q_1, S_1, D_h, PF_1
	k = 9	9	I_a, I_p, I_v, I_q, I_{RMS}, I_5, S_1, D_h, PF_1
	k = 5	5	I_a, I_p, I_v, I_q, I_5
RFE	Features = 11	11	P, Q, V, I_7, I_{13}, P_1, P_h, Q_1, S_1, D_h
	Features = 9	9	P, Q, V, I_7, I_{13}, P_1, P_h, S_1
	Features = 5	5	P, Q, V, I_7, S_1
TBE	Estimators = 50	12	I_a, I_p, I_q, I_v, A, P, Q, PF, I_{RMS}, Q_1, S_1, PF_1
	Estimators = 30	11	I_a, I_p, I_q, I_v, P, Q, QF, VF, I_5, I_7, P_1
	Estimators = 10	9	I_a, I_p, I_q, I_v, QF, VF, I_5, I_7, S_1
Collinearity	Correlation < 80	11	I_p, I_q, I_v, PF, QF, U_{RMS}, I_5, I_{17}, I_{19}, P_h, D_h
	Correlation < 50	9	I_p, I_q, I_v, PF, QF, U_{RMS}, I_{17}, P_h, D_h
Collinearity + features relevance	–	4	I_p, I_q, I_v, PF

electrical appliances from NILMbr and PLAID datasets), with the features being calculated for every second. Regarding the low-frequency base, the performances of three traditional classifiers were verified: DT, RF, and k-NN.

The collinearity-based process discussed in this paper provided the means to significantly reduce the number of features. As a result, the classification algorithms presented better performance and reduced computational time. Five features have remained for the NILMbr dataset: I_p, I_q, I_v, PF, and QF and the best classifier was the k-NN, with an accuracy of 98.74% and macro f-measure of 98.74%. For the PLAID dataset, four features have been selected: I_p, I_q, I_v, and PF with $k-$NN as the best classifier, with an accuracy of 98.65% and a macro f-measure of 96.95% and such a results are in line with the NILM efforts in the literature.

As presented in the results section, the collinearity method was compared with the RFE, SKB, and TBE approaches, and all presented similar feature selections. Moreover, in the feature selection methods, the features from the CPT presented the highest relevance. In the comparison of the feature selection methods, the best results are from the TBE with Extra-trees and 10 estimators, in this way, the features I_a, I_p, I_q, I_v, QF, VF, I_5, I_7, S_1 represents the best set of features for the NILMbr dataset.

Therefore, it was possible to obtain high values of performance metrics from the classification algorithms with few features. For the NILM approach, the feature reduction helps by decreasing the computational

Table 10
Comparison of the feature selection approaches and the accuracy for the NILMbr dataset.

Approach	Parameter	Accuracy		
		DT	RF (n = 30)	k-NN (k = 3)
SKB	k = 11	0.9401	0.9502	0.9529
	k = 9	0.9671	0.9689	0.9681
	k = 5	0.9651	0.9808	0.9883
RFE	Features = 11	0.9519	0.9591	0.9627
	Features = 9	0.9603	0.9614	0.9703
	Features = 5	0.9667	0.9873	0.9899
TBE	Estimators = 50	0.9613	0.9672	0.9706
	Estimators = 30	0.9708	0.9757	0.9787
	Estimators = 10	**0.9731**	**0.9883**	**0.9901**
Collinearity	Correlation < 0.80	0.9375	0.9464	0.9537
	Correlation < 0.50	0.9573	0.9641	0.9635
Collinearity + features relevance	–	0.9545	0.9782	0.9874

Table 11
Comparison of the results obtained with the ones available in the literature.

Dataset	Reference	Accuracy %
PLAID 1	Sadeghianpourhamami et al. [13]	93.2
	Hoyo-Montano et al. [30]	90.0
	Le et al. [40]	93.2
	Le et al. [41]	95.0
	The collinearity approach	**98.6**
NILMbr	Souza et al. [57]	95.0
	The collinearity approach	**98.7**
	TBE (10 estimators)	**99.0**

burden in the feature extraction and the appliance classification processes — two of the essential steps for the proper load disaggregation. Thus, the results in this paper demonstrate that the CPT can be used as an efficient means to select the most relevant features for the load identification process in the considered residential scenario.

The feature reduction process based on the presented technique allows the implementation of modern power management systems and equipment (smart meters, HEMS, EMS, among others). Moreover, such a technique is also suitable to be incorporated into the entire model of the NILM systems. At last, this paper demonstrates that based on the data computed from power theories, it is possible to analyze the load's physical characteristics, allowing to perform an accurate NILM. Consequently, the NILM process can be more efficiently used on strategies related to energy efficiency improvement and reduction of electricity consumption.

CRediT authorship contribution statement

Wesley A. Souza: Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Augusto M.S. Alonso:** Methodology, Validation, Writing - review & editing. **Thais B. Bosco:** Investigation, Software, Writing - original draft, Data curation. **Fernando D. Garcia:** Methodology, Investigation and review. **Flavio A.S. Gonçalves:** Funding acquisition, Project administration, Writing - review & editing. **Fernando P. Marafão:** Formal analysis, Resources, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors are grateful to the São Paulo Research Foundation (FAPESP) for the grant 2016/08645-9.

References

[1] F.D. Garcia, F.P. Marafao, W.A.d. Souza, L.C.P.d. Silva, Power metering: History and future trends, in: IEEE Green Technologies Conference, 2017, pp. 26–33, http://dx.doi.org/10.1109/GreenTech.2017.10.
[2] Y. Kabalci, A survey on smart metering and smart grid communication, Renew. Sustain. Energy Rev. 57 (2016) 302–318, http://dx.doi.org/10.1016/j.rser.2015.12.114.
[3] S.S.S.R. Depuru, L. Wang, V. Devabhaktuni, N. Gudi, Smart meters for power grid; Challenges, issues, advantages and status, in: IEEE/PES Power Systems Conference and Exposition, 2011, pp. 1–7, http://dx.doi.org/10.1109/PSCE.2011.5772451.
[4] J. Leiva, A. Palacios, J.A. Aguado, Smart metering trends, implications and necessities: A policy review, Renew. Sustain. Energy Rev. 55 (2016) 227–233, http://dx.doi.org/10.1016/j.rser.2015.11.002.
[5] R. Ford, M. Pritoni, A. Sanguinetti, B. Karlin, Categories and functionality of smart home technology for energy management, Build. Environ. 123 (2017) 543–554, http://dx.doi.org/10.1016/j.buildenv.2017.07.020.
[6] B. Cannas, S. Carcangiu, D. Carta, A. Fanni, C. Muscas, Selection of features based on electric power quantities for non-intrusive load monitoring, Appl. Sci. 11 (2) (2021) http://dx.doi.org/10.3390/app11020533.
[7] M. Baldini, A. Trivella, J.W. Wente, The impact of socioeconomic and behavioural factors for purchasing energy efficient household appliances: A case study for Denmark, Energy Policy 120 (2018) 503–513, http://dx.doi.org/10.1016/j.enpol.2018.05.048.
[8] F.D. Garcia, W.A. Souza, F.P. Marafão, Embedded NILM as home energy management system: A heterogeneous computing approach, IEEE Latin Am. Trans. 18 (02) (2020) 360–367, http://dx.doi.org/10.1109/TLA.2020.9085291.
[9] S. Giri, M. Bergés, An error correction framework for sequences resulting from known state-transition models in Non-Intrusive Load Monitoring, Adv. Eng. Inform. 32 (2017) 152–162, http://dx.doi.org/10.1016/j.aei.2017.01.006.
[10] Q. Liu, K.M. Kamoto, X. Liu, M. Sun, N. Linge, Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models, IEEE Trans. Consum. Electron. 65 (1) (2019) 28–37, http://dx.doi.org/10.1109/TCE.2019.2891160.
[11] L. Li, L. Yang, H. Chen, M. Li, C. Zhang, Multi-objective evolutionary algorithms applied to non-intrusive load monitoring, Electr. Power Syst. Res. 177 (2019) 105961, http://dx.doi.org/10.1016/j.epsr.2019.105961.
[12] S. Welikala, C. Dinesh, M.P.B. Ekanayake, R.I. Godaliyadda, J. Ekanayake, Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting, IEEE Trans. Smart Grid 10 (1) (2019) 448–461, http://dx.doi.org/10.1109/TSG.2017.2743760.
[13] N. Sadeghianpourhamami, J. Ruyssinck, D. Deschrijver, T. Dhaene, C. Develder, Comprehensive feature selection for appliance classification in NILM, Energy Build. 151 (2017) 98–106, http://dx.doi.org/10.1016/j.enbuild.2017.06.042.
[14] W.A. Souza, T.A. Almeida, Data characterization for electrical load disaggregation using supervised learning, in: Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional, SBC, 2019, pp. 226–237, http://dx.doi.org/10.5753/eniac.2019.9286.
[15] P. Tenti, H.K.M. Paredes, P. Mattavelli, Conservative power theory, a framework to approach control and accountability issues in smart microgrids, IEEE Trans. Power Electron. 26 (3) (2011) 664–673, http://dx.doi.org/10.1109/TPEL.2010.2093153.
[16] IEEE, IEEE Standard Definitions for the Measurement of Electric Power Quantities Under Sinusoidal, Nonsinusoidal, Balanced, or Unbalanced Conditions, IEEE Std 1459-2010 (Revision of IEEE Std 1459-2000), 2010, pp. 1–50, http://dx.doi.org/10.1109/IEEESTD.2010.5439063.
[17] G.W. Hart, Nonintrusive appliance load monitoring, Proc. IEEE 80 (12) (1992) 1870–1891, http://dx.doi.org/10.1109/5.192069.
[18] S. Makonin, F. Popowich, B. Gill, The cognitive power meter: Looking beyond the smart meter, in: IEEE Canadian Conference on Electrical and Computer Engineering, 2013, pp. 1–5, http://dx.doi.org/10.1109/CCECE.2013.6567686.
[19] S. Drenker, A. Kader, Nonintrusive monitoring of electric loads, IEEE Comput. Appl. Power 12 (4) (1999) 47–51, http://dx.doi.org/10.1109/67.795138.
[20] P. Xiao, S. Cheng, Neural network for NILM based on operational state change classification, 2019, arXiv arXiv:1902.02675.
[21] M. Marceau, R. Zmeureanu, Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings, Energy Convers. Manage. 41 (13) (2000) 1389–1403, http://dx.doi.org/10.1016/S0196-8904(99)00173-9.
[22] M. Baranski, J. Voss, Genetic algorithm for pattern detection in NIALM systems, in: IEEE International Conference on Systems, Man and Cybernetics, Vol. 4, 2004, pp. 3462–3468, http://dx.doi.org/10.1109/ICSMC.2004.1400878.
[23] D.F. Teshome, T.D. Huang, K. Lian, Distinctive load feature extraction based on fryze's time-domain power theory, IEEE Power Energy Technol. Syst. J. 3 (2) (2016) 60–70, http://dx.doi.org/10.1109/JPETS.2016.2559507.
[24] M.B. Figueiredo, A. de Almeida, B. Ribeiro, An experimental study on electrical signature identification of non-intrusive load monitoring (NILM) systems, in: Adaptive and Natural Computing Algorithms, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 31–40, http://dx.doi.org/10.1007/978-3-642-20267-4_4.

[25] S.R. Shaw, S.B. Leeb, L.K. Norford, R.W. Cox, Nonintrusive load monitoring and diagnostics in power systems, IEEE Trans. Instrum. Meas. 57 (7) (2008) 1445–1454, http://dx.doi.org/10.1109/TIM.2008.917179.

[26] A.S. Bouhouras, P.A. Gkaidatzis, E. Panagiotou, N. Poulakis, G.C. Christoforidis, A NILM algorithm with enhanced disaggregation scheme under harmonic current vectors, Energy Build. 183 (2019) 392–407, http://dx.doi.org/10.1016/j.enbuild.2018.11.013.

[27] S. Lin, L. Zhao, F. Li, Q. Liu, D. Li, Y. Fu, A nonintrusive load identification method for residential applications based on quadratic programming, Electr. Power Syst. Res. 133 (2016) 241–248, http://dx.doi.org/10.1016/j.epsr.2015.12.014.

[28] P. Held, S. Mauch, A. Saleh, D.O. Abdeslam, D. Benyoucef, Frequency invariant transformation of periodic signals (FIT-PS) for classification in NILM, IEEE Trans. Smart Grid (2018) 1–8, http://dx.doi.org/10.1109/TSG.2018.2886849.

[29] J.M. Gillis, S.M. Alshareef, W.G. Morsi, Nonintrusive load monitoring using wavelet design and machine learning, IEEE Trans. Smart Grid 7 (1) (2016) 320–328, http://dx.doi.org/10.1109/TSG.2015.2428706.

[30] J.A. Hoyo-Montano, N. Leon-Ortega, G. Valencia-Palomo, R.A. Galaz-Bustamante, D.F. Espejel-Blanco, M.G. Vazquez-Palma, Non-intrusive electric load identification using wavelet transform, Ingen. Investig. 38 (2018) 42–51, http://dx.doi.org/10.15446/ing.investig.v38n2.70550.

[31] W.A. Souza, F.P. Marafão, E.V. Liberado, M.G. Simões, L.C.P. Silva, A NILM dataset for cognitive meters based on conservative power theory and pattern recognition techniques, J. Control Autom. Electr. Syst. 29 (6) (2018) 742–755, http://dx.doi.org/10.1007/s40313-018-0417-4.

[32] F. Jazizadeh, B. Becerik-Gerber, M. Berges, L. Soibelman, An unsupervised hierarchical clustering based heuristic algorithm for facilitated training of electricity consumption disaggregation systems, Adv. Eng. Inform. 28 (4) (2014) 311–326, http://dx.doi.org/10.1016/j.aei.2014.09.004.

[33] I. Opris, L. Caracasian, The relation between smart meters and electricity consumers, in: International Conference on Environment and Electrical Engineering, 2013, pp. 325–329, http://dx.doi.org/10.1109/EEEIC.2013.6549536.

[34] R. Machlev, D. Tolkachov, Y. Levron, Y. Beck, Dimension reduction for NILM classification based on principle component analysis, Electr. Power Syst. Res. 187 (2020) 106459, http://dx.doi.org/10.1016/j.epsr.2020.106459.

[35] A. Moradzadeh, O. Sadeghian, K. Pourhossein, B. Mohammadi-Ivatloo, A. Anvari-Moghaddam, Improving residential load disaggregation for sustainable development of energy via principal component analysis, Sustainability 12 (8) (2020) 3158, http://dx.doi.org/10.3390/su12083158.

[36] Y. Lin, M. Tsai, A novel feature extraction method for the development of nonintrusive load monitoring system based on BP-ANN, in: International Symposium on Computer, Communication, Control and Automation, Vol. 2, 2010, pp. 215–218, http://dx.doi.org/10.1109/3CA.2010.5533571.

[37] J. Kelly, W. Knottenbelt, Neural NILM: Deep neural networks applied to energy disaggregation, in: ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, 2015, pp. 55–64, http://dx.doi.org/10.1145/2821650.2821672.

[38] H. Yang, H. Chang, C. Lin, Design a neural network for features selection in non-intrusive monitoring of industrial electrical loads, in: International Conference on Computer Supported Cooperative Work in Design, 2007, pp. 1022–1027, http://dx.doi.org/10.1109/CSCWD.2007.4281579.

[39] T.-T.-H. Le, J. Kim, H. Kim, Classification performance using gated recurrent unit recurrent neural network on energy disaggregation, in: International Conference on Machine Learning and Cybernetics, Vol. 1, 2016, pp. 105–110, http://dx.doi.org/10.1109/ICMLC.2016.7860885.

[40] T.-T.-H. Le, H. Kang, H. Kim, Household appliance classification using lower odd-numbered harmonics and the bagging decision tree, IEEE Access 8 (2020) 55937–55952, http://dx.doi.org/10.1109/ACCESS.2020.2981969.

[41] T.-T.-H. Le, S. Heo, H. Kim, Toward load identification based on the Hilbert transform and sequence to sequence long short-term memory, IEEE Trans. Smart Grid 12 (4) (2021) 3252–3264, http://dx.doi.org/10.1109/TSG.2021.3066570.

[42] X. Wang, Z. Liu, H. Zhang, Y. Zhao, J. Shi, H. Ding, A review on virtual power plant concept, application and challenges, in: IEEE Innovative Smart Grid Technologies - Asia, 2019, pp. 4328–4333, http://dx.doi.org/10.1109/ISGT-Asia.2019.8881433.

[43] Z. Yi, Y. Xu, W. Gu, W. Wu, A multi-time-scale economic scheduling strategy for virtual power plant based on deferrable loads aggregation and disaggregation, IEEE Trans. Sustain. Energy 11 (3) (2020) 1332–1346, http://dx.doi.org/10.1109/TSTE.2019.2924936.

[44] A.G. Ruzzelli, C. Nicolas, A. Schoofs, G.M.P. O'Hare, Real-time recognition and profiling of appliances through a single electricity sensor, in: IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, 2010, pp. 1–9, http://dx.doi.org/10.1109/SECON.2010.5508244.

[45] F. Sultanem, Using appliance signatures for monitoring residential loads at meter panel level, IEEE Trans. Power Deliv. 6 (4) (1991) 1380–1385, http://dx.doi.org/10.1109/61.97667.

[46] M. Dong, P.C.M. Meira, W. Xu, W. Freitas, An event window based load monitoring technique for smart meters, IEEE Trans. Smart Grid 3 (2) (2012) 787–796, http://dx.doi.org/10.1109/TSG.2012.2185522.

[47] H.-H. Chang, Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses, Energies 5 (11) (2012) 4569–4589, http://dx.doi.org/10.3390/en5114569.

[48] L.D. Baets, J. Ruyssinck, C. Develder, T. Dhaene, D. Deschrijver, Appliance classification using VI trajectories and convolutional neural networks, Energy Build. 158 (2018) 32–36, http://dx.doi.org/10.1016/j.enbuild.2017.09.087.

[49] L. Du, D. He, R.G. Harley, T.G. Habetler, Electric load classification by binary voltage–current trajectory mapping, IEEE Trans. Smart Grid 7 (1) (2016) 358–365, http://dx.doi.org/10.1109/TSG.2015.2442225.

[50] T. Hassan, F. Javed, N. Arshad, An empirical investigation of V-I trajectory based load signatures for non-intrusive load monitoring, IEEE Trans. Smart Grid 5 (2) (2014) 870–878, http://dx.doi.org/10.1109/TSG.2013.2271282.

[51] M. Dong, P.C.M. Meira, W. Xu, C.Y. Chung, Non-intrusive signature extraction for major residential loads, IEEE Trans. Smart Grid 4 (3) (2013) 1421–1430, http://dx.doi.org/10.1109/TSG.2013.2245926.

[52] A. Mukaroh, T.-T.-H. Le, H. Kim, Background load denoising across complex load based on generative adversarial network to enhance load identification, Sensors 20 (19) (2020) http://dx.doi.org/10.3390/s20195674.

[53] J.C. Maxwell, VIII. A dynamical theory of the electromagnetic field, Philos. Trans. R. Soc. Lond. 1 (155) (1865) 459–512.

[54] A.C. Moreira, L.C. da Silva, H.K. Paredes, Power quality study and analysis of different arc welding machines, J. Control Autom. Electr. Syst. 29 (2) (2018) 163–176, http://dx.doi.org/10.1007/s40313-017-0363-6.

[55] T. Hong, F. de León, Lissajous curve methods for the identification of nonlinear circuits: Calculation of a physical consistent reactive power, IEEE Trans. Circuits Syst. I. Regul. Pap. 62 (12) (2015) 2874–2885, http://dx.doi.org/10.1109/TCSI.2015.2495780.

[56] H. Kirkham, A. Emanuel, M. Albu, D. Laverty, Resolving the reactive power question, in: IEEE International Instrumentation and Measurement Technology Conference, 2019, pp. 1–6, http://dx.doi.org/10.1109/I2MTC.2019.8826915.

[57] W.A. Souza, F.D. Garcia, F.P. Marafao, L.C.P. da Silva, M.G. Simoes, Load disaggregation using microscopic power features and pattern recognition, Energies 12 (14) (2019) 2641, http://dx.doi.org/10.3390/en12142641.

[58] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40 (1) (2014) 16–28, http://dx.doi.org/10.1016/j.compeleceng.2013.11.024.

[59] S.V. Schooten, R. Harel, S. Ercan, E.D. Groot, Applying feature selection methods on fMRI data, in: Student Project Report, 2014, pp. 1–21.

[60] S. Biswas, M. Bordoloi, B. Purkayastha, Review on feature selection and classification using neuro-fuzzy approaches, Int. J. Appl. Evol. Comput. (IJAEC) 7 (4) (2016) 28–44, http://dx.doi.org/10.4018/IJAEC.2016100102.

[61] V. Prachayasittikul, A. Worachartcheewan, W. Shoombuatong, N. Songtawee, S. Simeon, V. Prachayasittikul, C. Nantasenamat, Computer-aided drug design of bioactive natural products, Curr. Top. Med. Chem. 15 (18) (2015) 1780–1800, http://dx.doi.org/10.2174/1568026615666150506151101.

[62] E. Akarslan, R. Doğan, A novel approach based on a feature selection procedure for residential load identification, Sustain. Energy Grids Netw. 27 (2021) 100488, http://dx.doi.org/10.1016/j.segan.2021.100488.

[63] A. Valencia-Duque, A.A. Meza, A. Orozco-Gutiérrez, Automatic identification of power quality events using a machine learning approach, Sci. Tech. 24 (2) (2019) 183–189.

[64] S. Houidi, F. Auger, H. Ben Attia Sethom, D. Fourer, L. Miègeville, Relevant feature selection for home appliances recognition, in: Electrimacs 2017, Toulouse, France, 2017, pp. 1–7.

[65] C. Chalmers, P. Fergus, C.A.C. Montanez, S. Sikdar, F. Ball, B. Kendall, Detecting activities of daily living and routine behaviours in dementia patients living alone using smart meter load disaggregation, IEEE Trans. Emerg. Top. Comput. (2020) http://dx.doi.org/10.1109/TETC.2020.2993177.

[66] H. Chen, Y.-H. Wang, C.-H. Fan, A convolutional autoencoder-based approach with batch normalization for energy disaggregation, J. Supercomput. 77 (3) (2021) 2961–2978, http://dx.doi.org/10.1007/s11227-020-03375-y.

[67] S. Ghosh, D.K. Panda, S. Das, D. Chatterjee, Cross-correlation based classification of electrical appliances for non-intrusive load monitoring, in: International Conference on Sustainable Energy and Future Electric Transportation, IEEE, 2021, pp. 1–6, http://dx.doi.org/10.1109/SeFet48154.2021.9375687.

[68] B.S. Wade, S.H. Joshi, B.A. Gutman, P.M. Thompson, Machine learning on high dimensional shape data from subcortical brain surfaces: A comparison of feature selection and classification methods, Pattern Recognit. 63 (2017) 731–739, http://dx.doi.org/10.1016/j.patcog.2016.09.034.

[69] O. Abedinia, N. Amjady, H. Zareipour, A new feature selection technique for load and price forecast of electrical power systems, IEEE Trans. Power Syst. 32 (1) (2017) 62–74, http://dx.doi.org/10.1109/TPWRS.2016.2556620.

[70] H. Midi, S. Sarkar, S. Rana, Collinearity diagnostics of binary logistic regression model, J. Interdiscip. Math. 13 (3) (2010) 253–267, http://dx.doi.org/10.1080/09720502.2010.10700699.

[71] J. Gao, E.C. Kara, S. Giri, M. Bergés, A feasibility study of automated plug-load identification from high-frequency measurements, in: IEEE Global Conference on Signal and Information Processing, 2015, pp. 220–224, http://dx.doi.org/10.1109/GlobalSIP.2015.7418189.

[72] I.M. Nasir, M.A. Khan, M. Yasmin, J.H. Shah, M. Gabryel, R. Scherer, R. Damaševičius, Pearson correlation-based feature selection for document classification using balanced training, Sensors 20 (23) (2020) http://dx.doi.org/10.3390/s20236793.
[73] E.C. Blessie, E. Karthikeyan, Sigmis: A feature selection algorithm using correlation based method, J. Algorithms Comput. Technol. 6 (3) (2012) 385–394, http://dx.doi.org/10.1260/1748-3018.6.3.385.
[74] Y. Liu, Y. Mu, K. Chen, Y. Li, J. Guo, Daily activity feature selection in smart homes based on pearson correlation coefficient, Neural Process. Lett. 51 (2) (2020) 1771–1787, http://dx.doi.org/10.1007/s11063-019-10185-8.
[75] D.M. Hawkins, The problem of overfitting, J. Chem. Inf. Comput. Sci. 44 (1) (2004) 1–12, http://dx.doi.org/10.1021/ci0342472.
[76] N. Meghanathan, Assortativity analysis of real-world network graphs based on centrality metrics, Comput. Inf. Sci. 9 (3) (2016) 7–25, http://dx.doi.org/10.5539/cis.v9n3p7.
[77] T. Picon, M. Nait Meziane, P. Ravier, G. Lamarque, C. Novello, J.-C. Le Bunetel, Y. Raingeaud, COOLL: Controlled On/Off loads library, a public dataset of high-sampled electrical signals for appliance identification, 2016, arXiv:1611.05803.
[78] D.P.B. Renaux, F. Pottker, H.C. Ancelmo, A.E. Lazzaretti, C.R.E. Lima, R.R. Linhares, E. Oroski, L.d.S. Nolasco, L.T. Lima, B.M. Mulinari, J.R.L.d. Silva, J.S. Omori, R.B.d. Santos, A dataset for non-intrusive load monitoring: Design and implementation, Energies 13 (20) (2020) http://dx.doi.org/10.3390/en13205371.
[79] J.Z. Kolter, M.J. Johnson, REDD: A public data set for energy disaggregation research, in: Proceedings of the Workshop on Data Mining Applications in Sustainability, 2011, pp. 59–62.
[80] K. Anderson, A. Ocneanu, D.R. Carlson, A. Rowe, M. Bergés, BLUED : A fully labeled public dataset for event-based non-intrusive load monitoring research, in: Proceedings of the KDD Workshop on Data Mining Applications in Sustainability, 2012, pp. 1–8.
[81] J. Kelly, W. Knottenbelt, The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes, Sci. Data 2 (1) (2015) 150007, http://dx.doi.org/10.1038/sdata.2015.7.
[82] M. Kahl, A.U. Haq, T. Kriechbaumer, H.-A. Jacobsen, WHITED-A worldwide household and industry transient energy data set, in: Proceeding of the European Workshop on Non-Intrusive Load Monitoring, 2016, pp. 1–5.
[83] C. Shin, E. Lee, J. Han, J. Yim, W. Rhee, H. Lee, The ENERTALK dataset, 15 Hz electricity consumption data from 22 houses in Korea, Sci. Data 6 (1) (2019) 193, http://dx.doi.org/10.1038/s41597-019-0212-5.
[84] G.W. Milligan, M.C. Cooper, A study of standardization of variables in cluster analysis, J. Classification 5 (2) (1988) 181–204, http://dx.doi.org/10.1007/BF01897163.
[85] A. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, Pattern Recognit. 38 (12) (2005) 2270–2285, http://dx.doi.org/10.1016/j.patcog.2005.01.012.
[86] T.M. Cover, P.E. Hart, et al., Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27, http://dx.doi.org/10.1109/TIT.1967.1053964.
[87] J.R. Quinlan, C4. 5: Programs for Machine Learning, Elsevier, 2014.
[88] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, http://dx.doi.org/10.1023/A:1010933404324.
[89] J. Shao, Linear model selection by cross-validation, J. Amer. Statist. Assoc. 88 (422) (1993) 486–494, http://dx.doi.org/10.2307/2290328.
[90] R. Kohavi, et al., A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Proceedings of the International Joint Conference on Artificial Intelligence, Vol. 14, Montreal, Canada, 1995, pp. 1137–1145. http://dx.doi.org/10.5555/1643031.1643047.
[91] P. Thanh Noi, M. Kappas, Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery, Sensors 18 (1) (2018) http://dx.doi.org/10.3390/s18010018.
[92] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[93] P. Indyk, R. Motwani, Approximate nearest neighbors: Towards removing the curse of dimensionality, in: Proceedings of the Annual ACM Symposium on Theory of Computing, 1998, pp. 604–613, http://dx.doi.org/10.1145/276698.276876.
[94] F.D. Garcia, W.A. Souza, I.S. Diniz, F.P. Marafão, NILM-based approach for energy efficiency assessment of household appliances, Energy Inform. 3 (1) (2020) 10, http://dx.doi.org/10.1186/s42162-020-00131-7.
[95] Z. Li, W. Xie, T. Liu, Efficient feature selection and classification for microarray data, PLoS One 13 (8) (2018) e0202167, http://dx.doi.org/10.1371/journal.pone.0202167.
[96] L. Kerkeni, Y. Serrestou, M. Mbarki, K. Raoof, M.A. Mahjoub, C. Cleder, Automatic speech emotion recognition using machine learning, in: A. Cano (Ed.), Social Media and Machine Learning, IntechOpen, Rijeka, 2020, http://dx.doi.org/10.5772/intechopen.84856, Ch. 2.
[97] M. Mera-Gaona, U. Neumann, R. Vargas-Canas, D.M. López, Evaluating the impact of multivariate imputation by MICE in feature selection, Plos One 16 (7) (2021) e0254720, http://dx.doi.org/10.1371/journal.pone.0254720.
[98] X. Wu, W. Zheng, M. Pu, J. Chen, D. Mu, Invalid bug reports complicate the software aging situation, Softw. Qual. J. (2020) 1–26, http://dx.doi.org/10.1007/s11219-019-09481-2.