Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

Residential Electricity Customers Classification Using Multilayer Perceptron Neural Network

P. Asghari* and A. Zakariazadeh*(C.A.)

Abstract: This paper proposes a novel approach to analyzing and managing electricity consumption using a clustering algorithm and a high-accuracy classifier for smart meter data. The proposed method utilizes a multilayer perceptron neural network classifier optimized by an Imperialist Competitive Algorithm (ICA) called ICA-optimized MLP, and a CD Index based on Fuzzy c-means to optimally determine representative load curves. A case study involving a real dataset of residential smart meters is conducted to validate the effectiveness of the proposed method, and the results demonstrate that the ICA-optimized MLP method achieves an accuracy of 98.62%, outperforming other classification methods. This approach has the potential to improve energy efficiency and reduce costs in the power system, making it a promising solution for analyzing and managing electricity consumption.

Keywords: Smart Meter; Fuzzy C-Means; MLP Neural Network; ICA Algorithm; Residential Electricity Customers.

Nomenclature

Acronym	
ANN	Artificial Neural Network
CD	Name of an index determining the Distance between two Clusters
FCM	Fuzzy c-means
ICA	Imperialist Competitive Algorithm
RPROP	Resilient back-propagation
MLP	Multilayer perceptron
Variables and parameters	
x_0	Un-preprocessed data in a user's load profile
x	Normalized data in a user's load profile
x_{min}	Minimum value in a user's load profile
x_{max}	Maximum value in a user's load profile
dis_{max}	Maximum distance between the centers of clusters
dis_{min}	Minimum distance between the centers of clusters
$dis(.)$	Euclidean distance function
$c^{(q)}$	Center of qth cluster
$c^{(l)}$	Center of lth cluster
q	qth cluster
l	lth cluster
NC	Number of clusters
$\|.\|$	Norm-2
NO	Number of objects
x_j	jth object
U	Partition matrix
μ_{lj}	Membership degree of object x_j in cluster $C^{(l)}$
C_l	lth cluster
J_c	Objective function of FCM clustering algorithm
M	Fuzzifier parameter
NOC_l	Number of objects in cluster C_l
r	Number of clusters
h_k	kth input node in hidden layer
n	nth node in the hidden layer
m	mth node in the input layer
i	Number of nodes in hidden nodes
w_{mn}	Connection weight between mth input node and nth hidden node
X_i	ith node in the input layer
θ_n	The bias of nth node in the hidden layer
H_k	kth node in the hidden layer
Sigmoid(.)	Transfer function
Y_j	jth node in the output layer
y_j	jth input node in the output layer

Iranian Journal of Electrical & Electronic Engineering, 2023.
Paper first received 08 Apr 2023 and accepted 19 Jul 2023.
* The author is with the Department of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran.
E-mails: pardis.asghari@mazust.ac.ir, zakaria@mazust.ac.ir.
Corresponding Author: A. Zakariazadeh

a	ath node in the input layer
b	bth node in the output layer
w_{ab}	Connection weight between ath input node and bth output node
$\acute{\theta}_b$	The bias of bth node in the output layer
j	Number of nodes in the output layer
TDS	Training Data Set
$\vec{o}^{(p)}$	pth desired output pattern
$\vec{x}$	Input pattern
$\vec{w}$	Weight vector
e	Error
E	Cost Function
Δ_{ij}	Weight changes
$\eta^{+}\&\eta^{-}$	Update Factor
α	A random variable by a uniform distribution
β	A number greater than 1
S	The distance between colony and imperialist
$U(.)$	Uniform distribution
$\delta(.)$	Derivative Function
$F_1(.)$	Griwank Function
$F_2(.)$	Rastrigin Function
$F_3(.)$	Rosenbrock Function
D	Dimension
χ	Optimal solution

1 Introduction

PROBLEMS of energy shortage, environmental pollution, and increasing demand for electricity have been challenged in the power grid. The smart grid can discover solutions for traditional power system failures through external energy sources for instance wind, solar, and other clean energy systems. It can be made a more sustainable, reliable, effective, and secure system for customers [1-2]. Smart meters have been known as the tools of Advanced Metering Infrastructure (AMI) through the improvement of smart grids and it has been very popular around the world [3]. Massive load data describing power consumption patterns has been generated because smart meters measure every 15-minute, 30-minute, or 1-hour time intervals. For example, in China, the annual volume of electricity measurement data gathered from smart meters will reach 117 Terabytes [4]. Also, when the data is collected from 1 million smart meters every 15 minutes, 35.04 billion records are achieved with a volume of 2920 terabytes of data [5]. On the other hand, the traditional approach isn't suitable for analyzing a large volume of electricity consumption data. So, data mining methods would be a proper strategy to extract hidden information from the big data of smart meter readings. Typical load data mining techniques consist of association rules, classification, prediction, and clustering [6]. In [7], for example, an architecture for big data analytics of smart meters aiming at performing data storage, query, analysis, and visualization tasks on large smart meter data sets has been presented.

In [8], a medium voltage customer characterization method supported by knowledge discovery in the database has been presented. The method identified typical load profiles of medium voltage consumers. Also, an automatic classification of new consumers based on a rule set has been developed. In [9], a clustering method aiming at grouping different types of factories based on their electricity consumption patterns has been proposed. In order to estimate the number of clusters, data imagining has been implemented and the k-means algorithm was used on pre-processing data to create separation. In [10], a finite mixture of model-based clustering has been presented to better understand the peak demand and customers' behavior. Ten distinct groups were offered based on customers' demand and their behavior. Finally, an existing bootstrap technique has been employed to show that the clustering is trustworthy. In [11], a developed k-means algorithm to cluster residential load profiles was proposed. The method defined the link between the clustering ratio and the division time window. In [12], a method to estimate load profiles based on the k-means algorithm has been presented. The algorithm estimated lost and future measurements by employing earlier cluster centers and distance functions such as Canberra, Manhattan, Euclidean, and Pearson correlation distances. In [13], a classification method for household usage and activities by using residential smart meter data has been developed. Common household electrical utilization from a characteristic household smart meter measurement can be recognized by applying a deep learning methodology based on multilayer, feedforward neural networks. Clustering algorithms to recognize customer clusters for analysis in a business environment have been presented in [14]. In [15], a clustering method focusing on improving load forecasting has been presented. Before performing load forecasting, customers with similar load consumption patterns have been categorized into the same group. By using time-series energy consumption data, a method for categorizing buildings and conducting energy benchmarking has been presented in [16]. The framework was based on 3-step k-means clustering and was tested on a dataset of 81 buildings in Singapore.

Feature engineering and data-driven classification models are at the forefront of the analysis of large temporal sensor data from the built environment [17]. Several studies have extended the understanding of classification methods for electricity consumers. In [18], a multi-label classification method for non-intrusive load monitoring has been presented. By modifying the popular sparse representation-based classification technique, the method has solved multi-label classification problems. In [19], five pre-screening analysis methods for those who tend to carry out energy savings through home automation devices have been proposed. The methods used interval energy consumption data aiming at better characterizing building energy consumption for residential customers. In [20], a model to evaluate the importance of several variables on commercial office building electricity usage has been presented. To provide demand impact profiles, smart meter data, as well as information on installed equipment, has been used. In [21], a method for identifying customers with photovoltaic units using net energy consumption data of smart meters has been presented. The method has implemented k-means and neural network-based techniques on real consumption data.

In [22], a classification approach based on the k-means algorithm for daily load profiles of nonresidential buildings has been presented. Also, the results have been validated using three other methods time-series features extraction, dynamic time warping, and raw time series. In [23], a characteristics identification technique based on federated learning for electricity consumers has been presented. To this end, features have been extracted from smart meter data using principal component analysis where an artificial neural network has been implemented to link between electricity consumption patterns and other social characteristics of customers. Four multi-label classification algorithms including k-nearest neighbors, radial basis function networks, backpropagation networks, and support vector machines for household appliance disaggregation and non-intrusive load monitoring have been investigated in [24]. The results showed that multi-label k-nearest neighbors outperformed existing methods. In [25], the K-means algorithm has been applied to energy meter data of homes aiming at grouping similar heating daily profiles. Then, logistic regression and random forest as classification algorithms have been implemented to predict the heating consumption level.

The classification techniques have been widely implemented in power system problems. The deep neural network to identify and classify power system events by using measurements of phasor measurement units [26], binary classification method to identify the characteristics of residential customers in a time of use tariff [27], and multi-layered deep neural network for real-time event classification in power grids with renewable generation [28] are some of the applications.

Although various classification methods have been used for smart meter data so far, achieving a reliable algorithm with high accuracy is still of particular importance. One way to increase the accuracy of classification algorithms is to combine them with optimization methods. In this paper, using the imperialist competitive optimization algorithm, the multilayer perceptron neural network classification method is upgraded and its accuracy in implementation on smart meter data is increased. To the best of our knowledge, no smart meter data classification approach using a multilayer-perceptron neural network has been reported in the literature. Accordingly, the contributions of the present paper are summarized as follows:

- The application of the CD index-based fuzzy c-means clustering algorithm has been assessed on electricity consumption data.
- An optimized multilayer perceptron neural network has been applied to classify residential customers.
- Resilient back-propagation (RPROP) algorithm has been used to train MLP which has fast convergence and minimal storage requirements.
- Imperialist competitive algorithm (ICA) has been used to optimally determine the number of hidden layers and the number of neurons in each hidden layer of a Multilayer Perceptron (MLP) neural network
- As the application of the proposed method, assigning new or unknown customers to its related cluster method has been investigated.

The rest of the paper is organized as follows: Section 2 formulates the proposed method for electrical customers' characterization and classification. Some simulation results are described in Section 3, and finally concluding remarks and future works are presented in Section 4.

2 Proposed Methodology

The electricity consumption data has been periodically collected from smart meters. The data has been gathered in the meter data management system [29]. To conduct electricity customers' characterization, the combination of unsupervised (clustering technique) and supervised (classification) learning methods is required. As the quality of raw data directly affects the quality of final decisions, an initial data pre-processing step is essential. Fig. 1 shows the procedure of the proposed methodology and the details are described in the following subsections.

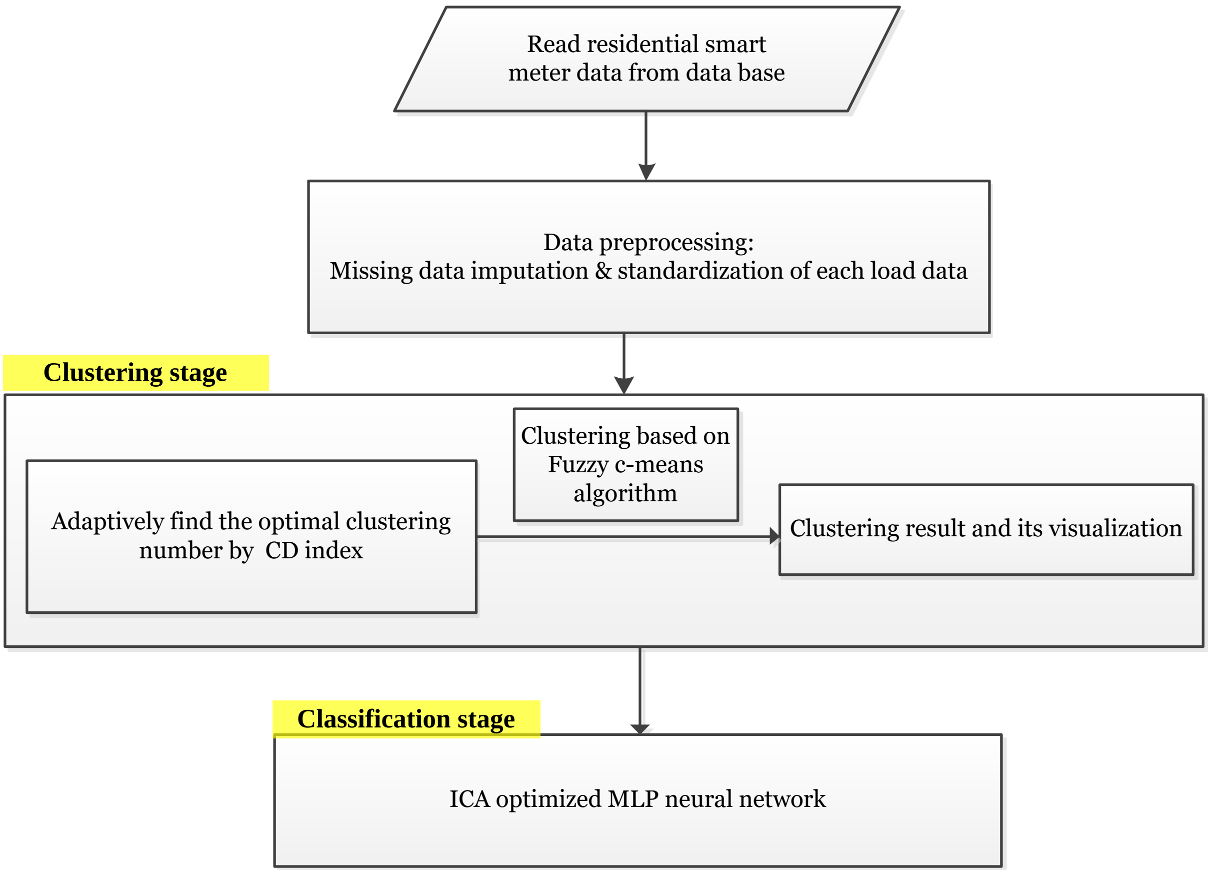

Fig. 1 Flowchart of the proposed method.

2.1 Load Profile Normalization

Due to analyzing and checking for missing values in all data, the pre-processing level is requested. As a result, in this paper, min-max normalization [30] has been employed. The formulation of this method is given as follows:

$$x = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where x and x_0 are un-preprocessed and normalized data in a user's load profile, respectively; x_{max} is the maximum value in a user's load profile; and x_{min} is the minimum value in a user's load profile. By using this approach, the value of all normalized data will be linearly between [0, 1]. In this way, smart meter raw data has been pre-processed and is ready to use.

2.2 Fuzzy C-Means (FCM) Clustering Algorithm

Clustering is an unsupervised technique used for data mining to discover significant patterns within datasets [31]. Currently, there is no definitive answer as to which clustering method is superior. However, by analyzing a particular database, we can determine which clustering algorithm produces favorable outcomes for our purposes. In this paper, Fuzzy c-means (FCM) methodology has been used to group load profiles into several clusters. In FCM, each data point of the data set can be related to more than one cluster [32]. The mathematical description of FCM is as follows. Let's assume that $x_1, x_2, \dots, x_n$ as a dataset, $c_1, c_2, \dots, c_{NC}$ as centers of clusters, $U = [\mu_{lj}]_{NC\times NO}$ $(l = 1, \dots, NC)$ $j = (1, \dots., NO)$ is the partition matrix, where μ_{lj} represents the membership degree of object x_j in cluster c_l. In fuzzy clustering, an object belongs to the cluster when the object has the maximum value of membership degree. The membership degree, μ_{lj} is given as follows:

$$\mu_{lj} \in [0,1] \tag{2}$$

$$0 < \sum_{j=1}^{NO} \mu_{lj} < NO \qquad \forall l = 1, \dots, NC \tag{3}$$

$$\sum_{l=1}^{NC} \mu_{lj} = 1 \qquad \forall j = 1, \dots, NO \tag{4}$$

$$\sum_{l=1}^{NC} \sum_{j=1}^{NO} \mu_{lj} = NO \tag{5}$$

The objective function of the FCM clustering algorithm is defined by:

$$J_c = \sum_{l=1}^{NC} \sum_{j=1}^{NO} \mu_{lj}^m \left\| x_j - v_l \right\|^2 \tag{6}$$

$$c_l = 1/NOC_l \sum_{x_j \in C_l} x_j \tag{7}$$

where c_l and NOC_l represent the center of cluster C_l and the number of objects in cluster C_l respectively. M is the fuzzifier parameter. $\|.\|$ points to the Euclidean distance between object x_j and cluster center c_l.

FCM updates the membership degree μ_{lj} and center of cluster C_l iteratively as follows:

$$\mu_{lj} = 1 \Big/ \sum_{r=1}^{NC} \left(d_{lj}/d_{rj} \right)^{\frac{2}{m-1}} \tag{8}$$

$$c_l = \sum_{j=1}^{NO} \mu_{lj}^m x_j \Big/ \sum_{j=1}^{NO} \mu_{lj}^m \tag{9}$$

The M as the fuzzifier parameter can be optimally determined by a Genetic Algorithm [33] and a basic structure that illustrates how a genetic algorithm can be used to optimize the fuzzifier parameter in a fuzzy clustering algorithm is presented as follows:

1. Define the fitness function: The fitness function should evaluate the quality of the clustering results based on a measure.
2. Initialize the population: Create an initial population of candidate solutions (values of M) randomly.
3. Evaluate fitness: Evaluate the fitness of each candidate solution using the fitness function.
4. Selection: Select the fittest individuals to be parents for the next generation.
5. Crossover and mutation: Breed the selected individuals to produce offspring with desirable traits. The offspring inherit traits from their parents, and mutation is introduced to maintain diversity in the population.
6. Evaluate the fitness of offspring: Evaluate the fitness of the offspring using the fitness function.
7. Selection: Select the fittest individuals (parents and offspring) to be parents for the next generation.
8. Repeat steps 5-7: Repeat steps 5-7 until convergence is reached (e.g., the fitness function no longer improves or a maximum number of iterations is reached). Return the best solution: The best solution is the candidate solution with the highest fitness value.

2.2.1 CD Index

To address the challenge of manually specifying the number of clusters in FCM, the CD index has been introduced as a means of determining the optimal number of clusters. The CD index is used to identify the most appropriate number of clusters for FCM, taking into account the characteristics of the dataset being analyzed. The CD index measures the distance between the two clusters [34]. CD index is computed as given as follows:

$$CD = \frac{dis_{max}}{dis_{min}} \sum_{q=1}^{NC} \left(\sum_{l=1}^{NC} dis\left(c^{(q)}, c^{(l)}\right) \right)^{-1} \tag{10}$$

where dis_{max} and dis_{min} point to the maximum and minimum distances between the cluster centers, respectively. A high value of this index represents the best partitions.

The flowchart of the fuzzy c-means clustering algorithm based on the CD index has been presented in Fig. 2.

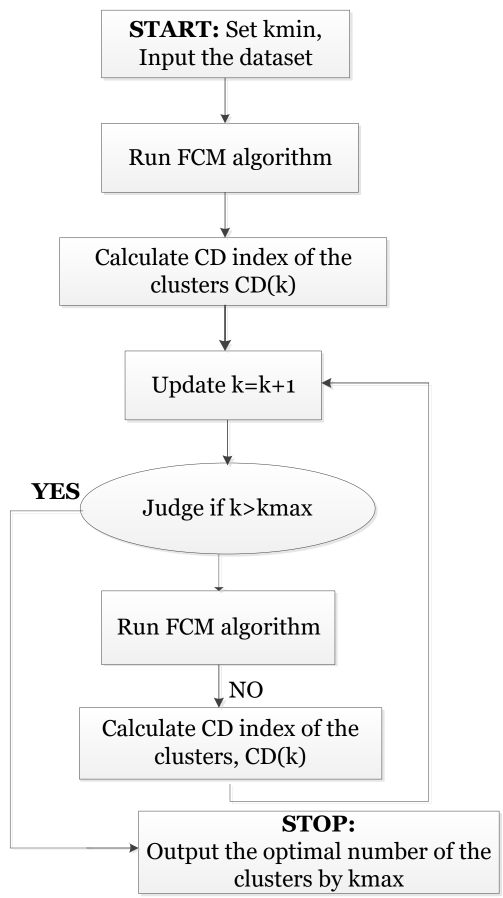

Fig. 2 Flowchart of the Fuzzy c-means clustering based on CD index.

2.3 The proposed classification algorithm

Classification is a supervised data mining technique that is implemented to train a target function aiming at mapping an attribute set to a predefined class label. One application of classification methods is to model smart meter data for assigning new or unknown consumers into existing customer classes. In this paper, the Multi-

Layer Perceptron (MLP) neural network [35] is recommended as a classifier in which the parameter corresponding to the MLP is optimally determined by the Imperialist Competitive Algorithm (ICA) [36]. Here, the classifier, optimization algorithm, and proposed classification algorithm are briefly described, respectively.

2.3.1 The proposed classification algorithm

Multi-Layer Perceptron (MLP) is one of the feed-forward neural networks that have one or more hidden layers. MLP neural network method consists of an input layer, one or more hidden layers, and an output layer. For better understanding, the MLP neural network concept with one hidden layer is shown in Fig. 3. By following these steps, the output of MLP is demonstrated.

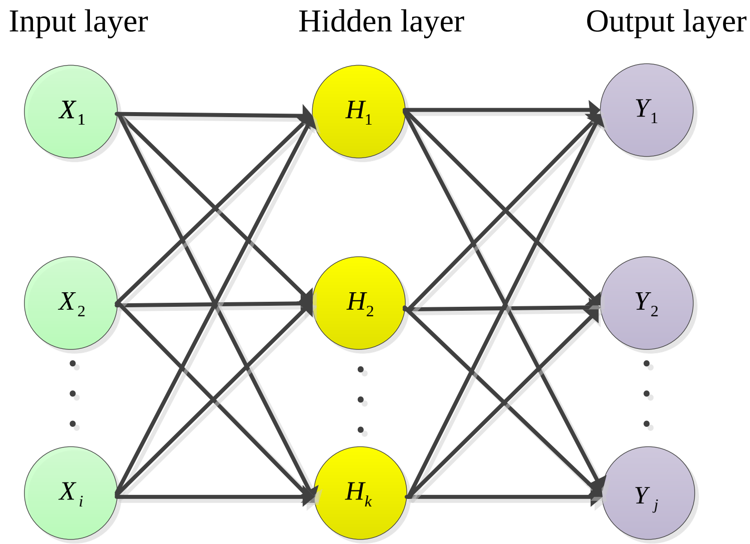

Fig. 3 The structure of a multilayer perceptron with one hidden layer.

Step1: calculate the sum of weights corresponding to input nodes by (11)

$$h_n = \sum_{m=1}^{i}(w_{mn}X_m) - \theta_n\,, \quad n = 1,2,\dots,k \tag{11}$$

Step2: compute the outputs of the hidden layer by (12)

$$H_n = sigmoid(h_n) = \frac{1}{1+exp(-h_n)} \quad n = 1,2,\dots,k \tag{12}$$

Step3: obtain the final output by (13) and (14)

$$y_b = \sum_{a=1}^{i} w_{ab}H_n - \acute{\theta}_b \qquad b = 1,2,\dots,j \tag{13}$$

$$Y_b = sigmoid(y_b) = \frac{1}{1+\exp(-y_b)} \qquad b = 1,2,\dots,j \tag{14}$$

In this paper, MLP is trained by the resilient back-propagation algorithm (RPROP) to develop the MLP performance and calculate the best value for weights and biases [37]. In the following section, this training algorithm is described.

2.3.2 Imperialist Competitive Algorithm (ICA)

(11)

The Imperialist competitive algorithm is called an evolutionary algorithm. ICA has been inspired by imperialist competition which was first introduced by Atashpaz and Lacus in 2007. This optimization algorithm has been applied to solve the problems optimally. ICA is started by a random initial population and each population is called a country. The objective function of the initial population is computed. The most powerful countries are assigned as imperialists and the other countries are selected as the colony of these imperialists. Then, to get more colonies, a competition between imperialists is took place. The more powerful imperialists had more chances to own more colonies. At that point, an empire is made by one imperialist with its colonies. After the colonies between the imperialists are divided, these colonies are moved to the imperialist by α units and arrive at their new sites. This movement is defined by (15).

$$\alpha \approx U(0, \beta \times S) \tag{15}$$

where α, β, and S represent a random variable by uniform distribution, a number greater than 1, and the distance between the colony and imperialist respectively. After the movement, if one of the colonies owned more power than its related imperialist, their sites were exchanged. Based on the total objective function value of each empire, the competition among empires is started. In this competition, the powerful empire tries to capture possessions and the weak one loses them. The winner of the competition goes to the most powerful empire. The failure of the competition goes to the empire that lost all its colonies. The pseudocode of the Imperialist Competitive Algorithm (ICA) has been presented in Algorithm I.

Algorithm I: Imperialist Competitive Algorithm (ICA)

Step1: Select some random points on the function and initialize the empires.

Step2: Move the colonies toward their relevant imperialist (Assimilating).

Step3: If there is a colony in an empire that has a lower cost than that of the imperialist, exchange the position of the imperialist and the colony.

Step4: Compute the total cost of an empire (Related to the power of both the imperialist and its colonies).

Step5: Pick up the weakest colony from the weakest empire and give it to the empire that has the most likelihood to possess it (Imperialist competition).

Step6: Eliminate the powerless empires.

Step7: If there is just one empire, else, go to step 2.

In this paper, the Imperialist Competitive Algorithm (ICA) is applied to optimize the number of hidden layers and the number of neurons in each hidden layer of a Multilayer Perceptron (MLP) neural network. The proposed MLP neural network has been trained using the resilient backpropagation algorithm. It is worth noting that in previous applications of MLP neural networks, the number of hidden layers and the number of neurons in each hidden layer were either selected manually or randomly. Therefore, the proposed approach of using the ICA algorithm to optimize the architecture of the MLP neural network can lead to better performance and more efficient use of resources.

The process of using ICA to optimize the number of hidden layers and neurons in an MLP network involves the following steps:

1. Define the optimization problem: In this case, the objective is to find the optimal number of hidden layers and neurons in each layer of an MLP network that can minimize the error between the actual outputs and the desired outputs.
2. Initialize the population: The initial population of solutions is generated randomly, representing different combinations of the number of hidden layers and neurons.
3. Evaluate the fitness: The fitness function is defined to evaluate the performance of each solution in the population. In this case, the fitness function could be the mean squared error (MSE) between the actual and desired outputs of the MLP network.
4. Determine the imperialists: The solutions with the highest fitness values are selected as the imperialists. These solutions will control the territories and lead to the evolution of the population.
5. Determine the colonies: The solutions with lower fitness values are selected as the colonies. These solutions will compete with each other to improve their fitness.
6. Perform competition: The colonies will compete with each other to overthrow the imperialists and take control of the territories. The competition is based on a probabilistic rule that takes into account the fitness of each colony and the distance to the nearest imperialist.
7. Perform assimilation: The winners of the competition will assimilate the characteristics of the imperialists and improve their fitness.
8. Perform revolution: A small percentage of the colonies will undergo a revolution, which means that they will change their characteristics randomly to introduce diversity in the population.
9. Repeat the process: Steps 3 to 8 are repeated until a stopping criterion is met, such as reaching a maximum number of iterations or a specific level of fitness.
10. Select the best solution: The final solution with the highest fitness value is selected as the optimal combination of the number of hidden layers and neurons in the MLP network.

In summary, the Imperialist Competitive Algorithm (ICA) can be used to optimize the number of hidden layers and neurons in a Multilayer Perceptron (MLP) neural network by generating a population of solutions, evaluating their fitness, and performing a competition and assimilation process to improve the solutions' fitness until an optimal solution is found.

2.3.3 Benchmark Functions

The Griewank, Rastrigin, and Rosenbrock functions as three classical benchmark functions have been implemented to evaluate the performance of the proposed ICA with other optimization algorithms [38].

The Griewank function as the first function (F_1) is given in (16). At its global minimum$(0,0,\ldots,0)$, the value of $F_1$is 0. The F_1 function is used to evaluate convergence and the computational efficiency of the optimization algorithms. As shown in (16) D and χ_d are the problem domain size and the optimal solution, respectively.

$$F_1(\vec{\chi}) = \frac{1}{4000}\left(\sum_{d=1}^{D}(\chi_d{}^2)\right) - \left(\prod_{i=1}^{D}\cos\left(\frac{\chi_d}{\sqrt{d}}\right)\right) + 1 \tag{16}$$

The Rastrigin function as the second function (F_2) is defined in (17). The value of this function is 0 at its global minimum $(0,0,\ldots,0)$. F_2 has several local minima. it is a classic case of a non-linear multimodal function, but sites of the minima are commonly distributed.

$$F_2(\vec{\chi}) = \sum_{d=1}^{D}(\chi_d{}^2 - 10\cos(2\pi\chi_d) + 10) \tag{17}$$

The Rosenbrock function as the third function (F_3) is given by (18). The value of F_3 is 0 at its global minimum $(1,1,\ldots,1)$. F_2 is continually used to evaluate the efficiency of the optimization algorithms.

$$F_3(\vec{\chi}) = \sum_{d=1}^{D}100(\chi_d{}^2 - \chi_{d+1}{}^2)^2 + (1-\chi_d)^2 \tag{18}$$

3 Numerical Study

3.1 Description of Smart Meter Data

To evaluate the performance of the developed data mining algorithm, the modified data corresponding to 2524 residential smart meters were collected [39]. The smart meter data providing load profiles represent daily power consumption. In the next step, the bad data or missing values such as reading errors or outliers have been removed from the database. Therefore, 2160 customers with reliable smart meter data are considered as input data to assess the efficiency of the proposed method. Then, the load profile of each customer has been normalized. Also, the normalized total daily load profile has been shown in Fig. 4.

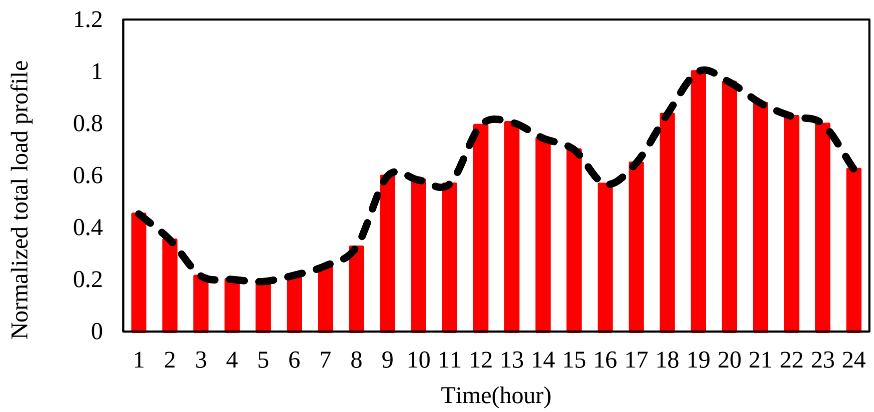

Fig. 4 Load profiles of all residential customers.

3.2 Clustering Results

A CD index based on the Fuzzy c-means clustering algorithm is utilized to cluster the residential smart meter data. In the FCM clustering algorithm, the number of clusters must be specified in advance and given as input to this clustering algorithm. So, for evaluating the proper value for cluster numbers, the corresponding CD index should be calculated and compared with the ones for other cluster numbers. In this paper, the range of cluster numbers for evaluating the CD index is between 2 and 20. For example, the number of clusters is first given two as input to the Fuzzy c-means algorithm. Then, the algorithm divides the data into two clusters and then, the index CD for two clusters is calculated. Next, the number of clusters is increased one by one and the clustering process is performed again and the corresponding CD index is calculated. This continues until the number of clusters reaches 20. Finally, the number of clusters for which the value of the CD index is the highest of the others is selected as the optimal number of clusters for the studied data. CD index for different cluster numbers k has been presented in Fig. 5. As shown, the best clustering number is 7, which achieves the maximum CD Index of 6.654.

Fig. 5 CD index for different cluster number k.

Therefore, the residential customers are clustered into seven clusters. Also, that the optimal value of M as a fuzzifier parameter for our dataset was 4.53 which is optimally determined by the Genetic Algorithm. The seven centroids or representative load profiles named C1 to C7 are illustrated in Fig. 6.

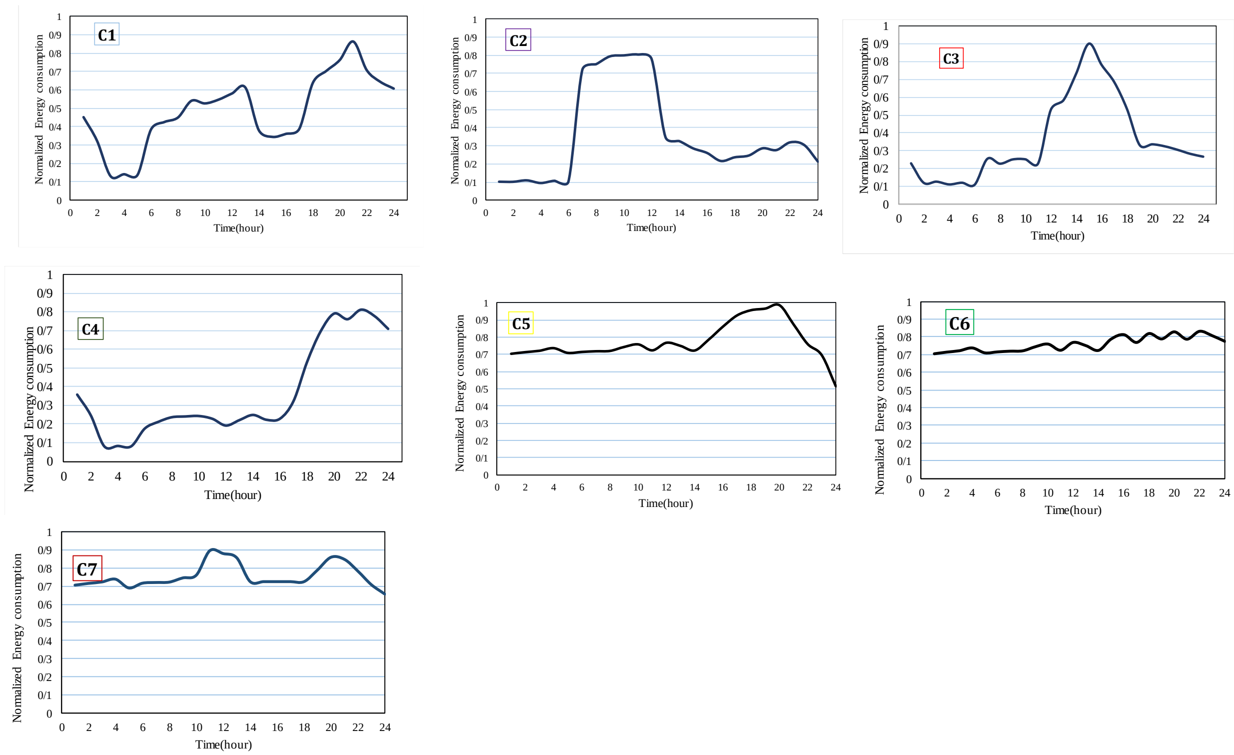

Fig. 6 Results of Fuzzy c-means clustering.

3.3 Result of Classification

For the residential smart meter data, investigating an accurate classifier is the finishing stage. In this paper, the Imperialist Competitive Algorithm (ICA) optimizes the number of hidden layers and the number of neurons in each hidden layer of a Multilayer Perceptron (MLP) neural network. The proposed MLP neural network has been trained using the resilient backpropagation algorithm. In Table 1, the ICA-optimized MLP neural network has been run 10 times to optimally determine the number of hidden layers and the number of neurons in each hidden layer for the MLP neural network. Also, for each run, the fitness function and complexity are calculated. Table 1 presents several columns, each containing information on the specific run of the algorithm. The "Run" column lists the number of the run, while the "Number of Hidden Layers" column displays the optimal number of hidden layers found by the algorithm. The "Number of Neurons per Hidden Layer" column corresponds to the optimal number of neurons per layer. The "Fitness Value" column shows the performance of each optimized MLP network in terms of fitness, with lower values being better. Finally, the "Complexity" column indicates the complexity of each optimized MLP network, which is calculated as the sum of the number of neurons in each hidden layer plus the number of input and output neurons. each run of the ICA algorithm produced a different optimal number of hidden layers and neurons per layer. The fitness value column shows that the algorithm was successful in finding MLP networks with low error rates, with all networks having a fitness value less than 0.003. Based on the results presented in the table, it can be observed that the optimal configuration of the neural network is achieved in the 9th run, which corresponds to the lowest values of both the fitness function and complexity. Specifically, the optimal number of hidden layers is found to be 1, while the optimal number of neurons per hidden layer is determined to be 8. These findings suggest that a neural network with a single hidden layer and 8 neurons in that layer is the most effective configuration for the given dataset and problem being addressed.

Table 1 MLP neural network's hidden layer and neuron selection based on Fitness function and Complexity.

Run	Number of Hidden Layers	Number of Neurons per Hidden Layer	Fitness Function	Complexity
1	2	10,5	0.002	46
2	3	8,6,4	0.0015	49
3	1	15	0.0022	46
4	4	12,8,6,4	0.0013	61
5	2	20,10	0.0018	61
6	1	12	0.0017	43
7	3	10,15,6	0.0012	62
8	2	18,6	0.0019	55
9	1	8	0.0016	39
10	2	15,12	0.0021	58

MLP architecture and training parameters are shown in Table. 2. In order to evaluate the accuracy of the proposed method, the result of the MLP neural network based on ICA optimization has been compared to the ones provided by the traditional MLP neural network. For this study, 60% of the data has been used for training the classifier and the rest for testing. In both algorithms, the number of neurons and accuracy have been compared in Table 3. In this section, the accuracy formula based on false positive (FP), false negative (FN), true positive (TP), and true negative (TN) can be expressed as:

$$accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{19}$$

The higher value of accuracy indicates better performance. Table 3 presents the results of the proposed method, which show that the optimal number of neurons in a hidden layer is 12. In contrast, the traditional MLP approach involves three hidden layers with 20, 15, and 8 neurons per layer, respectively. The accuracy of the traditional MLP algorithm and the ICA-optimized MLP algorithm are reported as 88.45% and 97.84%, respectively. These findings demonstrate that by optimizing the parameters of the MLP neural network algorithm, the accuracy of this classifier can be significantly improved.

Table 2 MLP architecture and training parameters.

training algorithm	Resilient back-propagation
The initial weights and basis	Random
Activation function (RP)	Tangent-sigmoid Linear

Table 3 The comparison of the MLP neural network optimized by ICA and Traditional MLP.

Classifier	Number of Hidden Layers	Number of Neurons per Hidden Layer	Accuracy (%)

Traditional MLP	20, 15,8	3	88.45
ICA optimized MLP	**12**	**1**	**97.84**

The distribution of errors is determined by the confusion matrix [40]. Diagonal elements and off-diagonal present the performance of the classifier and errors, respectively. As shown by the confusion matrix in Table 4, the classifier has correctly detected 330 out of 334 customers belonging to cluster one. According to the samples from cluster two, 210 out of 212 customers have been correctly selected. The classifier has incorrectly selected 4 out of 84 customers from cluster three. The proposed classifier has correctly detected 100 out of 102 customers belonging to cluster 4. The presented classifier has incorrectly selected 2 out of 50 customers from Cluster Five and 1 out of 60 customers from Cluster Six. From cluster 7, 5 out of 6 customers have been correctly selected. As a result, the proposed classification technique has been able to approximate **98.49%** of the data truthfully. To show the robustness of the ICA algorithm, the proposed method has been also optimized by the Grasshopper Optimization Algorithm (GOA) [41], Ant Colony Algorithm (AC) [42], and Particle Swarm Optimization (PSO) [43] and the results are analyzed. The overall accuracy and recall of the classification technique optimized by three algorithms (ICA, GOA, AC, and PSO) are listed in Tables 5–8. To this end, each algorithm has been run 10 times. Then, Recall and accuracy were calculated based on functions F1, F2, and F3. To conduct a comparison between the results provided by the three algorithms, the average values of functions have been calculated. The results evidenced that the ICA algorithm achieves the highest values in both recall and accuracy

Table 4 The proposed method Confusion Matrix.

True Label	Estimated Labels							
	1	2	3	4	5	6	7	total
1	330	1	-	-	-	-	1	332
2	1	210	-	-	1	-	-	212
3	2	-	80	1	1	-	-	84
4	1	-	-	100	-	-	-	101
5	-	1	-	-	50	-	-	51
6	-	-	2	1	-	60	-	63
7	-	-	2	-	-	1	5	8
Total	334	212	84	102	52	61	6	851

Table 5 The overall accuracy and Recall using the ICA algorithm.

Run	Recall			Accuracy		
	F1	F2	F3	F1	F2	F3
1	100	100	100	100	100	100
2	100	100	100	97	100	100
3	100	100	100	100	100	88
4	100	100	100	96	100	96
5	100	100	100	96	100	99.12
6	92.45	100	100	100	100	100
7	100	100	100	95.78	98.58	96.8
8	100	91.35	100	100	100	92
9	100	100	100	94.12	100	100
10	100	100	90.48	96.87	99	100
Ave.	99.142			98.175		

Table 6 The overall accuracy and Recall using the GOA algorithm.

Run	Recall			Accuracy		
	F1	F2	F3	F1	F2	F3
1	100	85.54	100	100	98.41	78.88
2	100	72.48	94.23	92.15	75.84	100
3	100	100	100	100	92.74	100
4	81.12	100	88.37	100	87.88	92.12
5	100	81.17	100	96.14	85.98	100
6	100	100	100	100	72.88	100
7	100	100	100	97.87	95.05	94.24
8	100	72.45	99.15	100	92.87	100
9	100	100	100	88.78	89.14	92.98
10	100	88.47	100	100	98	89.98
Ave.	95.43			93.731		

Table 7 The overall accuracy and Recall using the PSO algorithm.

Run	Recall			Accuracy		
	F1	F2	F3	F1	F2	F3
1	100	100	100	100	100	88.14
2	87.21	100	76.38	91.78	92.87	94.25
3	100	100	89.71	92.35	100	100
4	100	96.45	100	100	100	76.25
5	98.45	100	100	99.98	89.74	100
6	79.78	100	100	89.97	100	100
7	89.34	100	95.45	100	97.84	94.14
8	100	100	100	91.25	92.47	85.68
9	100	87.78	100	94.47	100	100
10	100	100	100	100	100	100
Ave.	96.685			95.706		

Table 8 The overall accuracy and Recall using the AG algorithm.

Run	Recall			Accuracy		
	F1	F2	F3	F1	F2	F3
1	100	100	78.19	94.15	100	78.88
2	100	80.34	100	100	79.25	87.18
3	84.23	100	92.46	100	100	93.47
4	100	91.86	89.18	91.45	80.37	88.17
5	76.14	100	100	99.28	100	100
6	84.17	80.79	81.12	100	100	100
7	100	77.71	76.12	100	82.27	73.58
8	84.23	100	100	93.34	92	79.64
9	100	89.61	85.41	100	100	78.88
10	75.24	100	100	76.14	100	100
Ave.	90.893			89.915		

The performance of four different machine learning algorithms, including Support Vector Machines (SVM) [44], Decision Trees [45], K-Nearest Neighbors (KNN) [46], and a proposed method are compared in Table 9. The comparison is based on two metrics, namely run and accuracy. The experiments were conducted with an iteration of 100. The results show that the proposed method performed better than the other algorithms in terms of both accuracy and runtime. The proposed method achieved the highest accuracy of 98.62% with the shortest runtime of 2 seconds. In comparison, SVM achieved an accuracy of 90.18% but required a longer runtime of 25 seconds, while Decision Trees achieved an accuracy of 95.36% with a runtime of 3 seconds. KNN achieved an accuracy of 92.77% with a runtime of 20 seconds.

Therefore, the results prove that the ICA-optimized MLP neural network outperforms the other methods. So, the proposed method is the most efficient classification technique for analyzing energy consumption data collected from smart meters.

Table 9 Comparison of classification accuracy among different algorithms.

Classifiers	run	Accuracy
SVM	25	90.18
KNN	20	92.77
Decision tree	3	95.36
The proposed method	**2**	**98.62**

As a result, ICA optimized MLP neural network has been implemented to classify the customers and discover more appropriate information from load profiles. The ICA-optimized MLP as a developed classification tool assigned new or unknown customers to the appropriate clusters by recognizing the customer's load profile pattern in training and testing processes. The achieved accuracy is 98.62%, which statistically indicates the customer's load profile pattern of more than 98.62% power system network can be accurately predicted. As a result, utilities can better understand their customer base and tailor their services to meet the specific needs of each group. For example, they can offer targeted energy efficiency programs and other services to each cluster, reducing energy consumption and lowering costs.

4 Conclusion

In this study, a new method for analyzing and managing electricity consumption using a clustering algorithm and a high-accuracy classifier for smart meter data is proposed. The method employs a novel ICA-optimized MLP classifier and a CD Index based on Fuzzy c-means to determine representative load curves effectively. The proposed approach is validated through a case study using a real dataset of residential smart meters, demonstrating an outstanding accuracy rate of 98.62%, surpassing that of other classification methods. The results suggest that this approach could be a promising solution for improving energy efficiency and reducing costs in the power system. Overall, the proposed method presents a practical and effective approach to analyzing and managing electricity consumption, with potential benefits for both consumers and utility providers.

As a future direction, it would be interesting to explore the potential of implementing GMDH (Group Method of Data Handling) for classification purposes in smart meter data analysis. For example, the Polynomial neural network is commonly associated with constructing the GMDH model. An application of the polynomial neural network as a model GMDH for the prediction of wind speed and solar radiation has been investigated in [47]. In addition, the probabilistic approach is applied to the model to generate the related scenarios for uncertain parameters. By incorporating GMDH, it may be possible to further enhance the accuracy and efficiency of the classification process. Additionally, it would be valuable to investigate the use of other clustering algorithms and classification techniques to further advance smart meter data analysis.

Reference

[1] G. B. Gharehpetian, M. S. Naderi, H. Modaghegh, and A. Zakariazadeh, "Iranian smart grid: road map and metering program," in Application of Smart Grid Technologies, Academic Press, 2018, pp. 13-60.

[2] T. Agheb, I. Ahmadi, and A. Zakariazadeh, "Optimum sizing and placement of wind turbines in distribution networks considering correlation of load demand and wind power," Iranian Journal of Electrical & Electronic Engineering, vol. 17, no. 3, pp. 253-261, Sep. 2021.

[3] A. Cooper, "Electric company smart meter deployments: foundation for a smart grid," Inst. Electr. Innov., Washington, DC, USA, Tech. Rep., 2016.

[4] S. Lu, G. Lin, H. Liu, C. Ye, H. Que, and Y. Ding, "A weekly load data mining approach based on hidden Markov model," IEEE Access, vol. 7, pp. 34609-34619, Mar. 2019.

[5] S. Sagiroglu, R. Terzi, Y. Canbay, and I. Colak, "Big data issues in smart grid systems," in 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), Birmingham, UK, 2016, pp. 1007-1012.

[6] Y. Wang, Q. Chen, C. Kang, M. Zhang, K. Wang, and Y. Zhao, "Load profiling and its application to demand response: A review," Tsinghua Sci. Technol., vol. 20, no. 2, pp. 117-129, Apr. 2015.

[7] T. Wilcox, N. Jin, P. Flach, and J. Thumim, "A Big Data platform for smart meter data analytics," Computers in Industry, vol. 105, pp. 250-259, May 2019.

[8] S. Ramos, J. M. Duarte, F. J. Duarte, and Z. Vale, "A data-mining-based methodology to

support MV electricity customers' characterization," Energy and Buildings, vol. 91, pp. 16-25, Feb. 2015.
[9] I. Khan, J. Z. Huang, M. A. Masud, and Q. Jiang, "Segmentation of factories on electricity consumption behaviors using load profile data," IEEE Access, vol. 4, pp. 8394-8406, Dec. 2016.
[10] S. Haben, C. Singleton, and P. Grindrod, "Analysis and clustering of residential customers energy behavioral demand using smart meter data," IEEE Transactions on Smart Grid, vol. 7, no. 1, pp. 136-144, Jan. 2016.
[11] A. Al-Wakeel and J. Wu, "K-means based cluster analysis of residential smart meter measurements," Energy Procedia, vol. 88, pp. 754-760, Apr. 2016.
[12] A. Al-Wakeel, J. Wu, and N. Jenkins, "K-means based load estimation of domestic smart meter measurements," Applied Energy, vol. 194, pp. 333-342, Mar. 2017.
[13] M. A. Devlin, B. P. Hayes. "Non-intrusive load monitoring and classification of activities of daily living using residential smart meter data." IEEE Transactions on Consumer Electronics 65, no. 3 (2019): 339-348.
[14] C. Flath, D. Nicolay, T. Conte, C. Dinther, L. Filipova-Neumann. "Cluster analysis of smart metering data." Business & Information Systems Engineering 4, no. 1 (2012): 31-39.
[15] F. L. Quilumba, W. Lee, H. Huang, D. Y. Wang, R. L. Szabados. "Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities." IEEE Transactions on Smart Grid 6, no. 2 (2014): 911-918.
[16] S. Zhan, Z. Liu, A. Chong, D. Yan. "Building categorization revisited: A clustering-based approach to using smart meter data for building energy benchmarking." Applied Energy 269 (2020): 114920.
[17] C. Miller. "What's in the box?! Towards explainable machine learning applied to non-residential building smart meter classification." Energy and Buildings 199 (2019): 523-536.
[18] S. Singh, A. Majumdar. "Non-intrusive Load Monitoring via Multi-label Sparse Representation based Classification." IEEE Transactions on Smart Grid (2019).
[19] S. Oh, J. S. Haberl, J.Baltazar. "Analysis methods for characterizing energy saving opportunities from home automation devices using smart meter data." Energy and Buildings (2020): 109955.
[20] C. Roach. "Estimating electricity impact profiles for building characteristics using smart meter data and mixed models." Energy and Buildings 211 (2020): 109686.
[21] D. L. Donaldson, D. Jayaweera. "Effective solar prosumer identification using net smart meter data." International Journal of Electrical Power & Energy Systems 118 (2020): 105823.
[22] M. Bourdeau, P. Basset, S. Beauchêne, D. Da Silva, T. Guiot, D. Werner, and E. Nefzaoui, "Classification of daily electric load profiles of non-residential buildings," Energy and Buildings, vol. 233, p. 110670, Mar. 2021.
[23] Wang, Y., Bennani, I.L., Liu, X., Sun, M. and Zhou, Y. "Electricity Consumer Characteristics Identification: A Federated Learning Approach." IEEE Transactions on Smart Grid (2021).
[24] Li, D. and Dick, S. "Non-intrusive load monitoring using multi-label classification methods." Electrical Engineering 103, no. 1 (2021): 607-619.
[25] Sala, J., Li, R. and Christensen, M.H. "Clustering and classification of energy meter data: A comparison analysis of data from individual homes and the aggregated data from multiple homes." In Building Simulation, vol. 14, no. 1, pp. 103-117. Tsinghua University Press, 2021.
[26] J. Shi, B. Foggo, and N. Yu, "Power system event identification based on deep neural network with information loading," IEEE Transactions on Power Systems, 2021, doi: 10.1109/TPWRS.2021.3099079.
[27] Y. Kiguchi, M. Weeks, and R. Arakawa, "Predicting winners and losers under time-of-use tariffs using smart meter data," Energy, vol. 236, p. 121438, Apr. 2021.
[28] R. Yadav, S. Raj, and A. K. Pradhan, "Real-time event classification in power system with renewables using kernel density estimation and deep neural network," IEEE Transactions on Smart Grid, vol. 10, no. 6, pp. 6849-6859, Nov. 2019.
[29] G. Gharehpetian, M. Salay Naderi, H. Modaghegh, A. Zakariazadeh. "Iranian smart grid: road map and metering program." In Application of Smart Grid Technologies, pp. 13-60. Academic Press, 2018.
[30] S. G. O. P. A. L. Patro and K. K. Sahu, "Normalization: A preprocessing stage," arXiv preprint arXiv:1503.06462, Mar. 2015.
[31] L. Cao, Z. Zhao, and D. Wang, "Clustering algorithms," in Target Recognition and Tracking for Millimeter Wave Radar in Intelligent Transportation, Singapore: Springer Nature Singapore, 2023, pp. 97-122.
[32] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm,"

Computers & Geosciences, vol. 10, no. 2-3, pp. 191-203, Apr. 1984.
[33] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," Multimedia Tools and Applications, vol. 80, pp. 8091-8126, Apr. 2021.
[34] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "On clustering validation techniques," J. Intell. Inf. Syst., vol. 17, no. 2, pp. 107145, Dec. 2001.
[35] J. Zhang, C. Li, Y. Yin, J. Zhang, and M. Grzegorzek, "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer," Artificial Intelligence Review, vol. 56, no. 2, pp. 1013-1070, May 2023.
[36] J. Cai, H. Yang, T. Lai, and K. Xu, "A new approach for optimal chiller loading using an improved imperialist competitive algorithm," Energy and Buildings, vol. 253, p. 112835, Mar. 2023.
[37] M. P. S. Bhatia and P. Chandra, "Effect of weight initialization on training of sigmoidal FFANN for back propagation algorithms," in Proceedings of ICETIT 2019, Cham, Switzerland: Springer, 2020, pp. 462-471.
[38] M. Jamil and X.-S. Yang, "A literature survey of benchmark functions for global optimisation problems," International Journal of Mathematical Modelling and Numerical Optimisation, vol. 4, no. 2, pp. 150-194, Jan. 2013. [39] Commission for Energy Regulation, “Electricity Smart Metering Customer Behaviour Trials (CBT) Findings Report - CER11080a,” Dublin, 2011.
[40] A. Luque, A. Carrasco, A. Martín, and A. de Las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," Pattern Recognition, vol. 91, pp. 216-231, Aug. 2019.
[41] S. Z. Mirjalili, S. Mirjalili, S. Saremi, H. Faris, and I. Aljarah, "Grasshopper optimization algorithm for multi-objective optimization problems," Applied Intelligence, vol. 48, no. 4, pp. 805-820, Apr. 2018.
[42] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," IEEE Computational Intelligence Magazine, vol. 1, no. 4, pp. 28-39, Nov. 2006.
[43] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proceedings of ICNN'95-International Conference on Neural Networks, vol. 4, Perth, Australia, Nov. 1995, pp. 1942-1948.
[44] M. E. Mavroforakis and S. Theodoridis, "A geometric approach to support vector machine (SVM) classification," IEEE Transactions on Neural Networks, vol. 17, no. 3, pp. 671-682, May 2006.
[45] Safavian, S. Rasoul, and David Landgrebe. "A survey of decision tree classifier methodology." IEEE transactions on systems, man, and cybernetics 21, no. 3 (1991): 660-674.
[46] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in OTM Confederated International Conferences "On the Move to Meaningful Internet Systems", Berlin, Heidelberg, Germany, 2003, pp. 986-996.
[47] A. Azizi, H. Karimi, and S. Jadid, "Daily operation of multi-energy systems based on stochastic optimization considering prediction of renewable energy generation," IET Renewable Power Generation, vol. 16, no. 2, pp. 245-260, Feb. 2022.


P. Asghari received the B.Sc. and M.Sc. (Hons.) degrees in electrical engineering from the Babol Noshirvani University of Technology in 2017 and 2020, respectively. Since 2020, she has been collaborating as a researcher at the University of Science and Technology of Mazandaran. Her main research interests are machine learning applications in power systems, demand response, electric vehicles, and renewable energy.


A. Zakariazadeh received a B.Sc. (Hons.) degree in electrical engineering from the University of Mazandaran, Babol, in 2007 and the M.Sc. (Hons.) and the Ph.D. degrees in electrical engineering from Iran University of Science and Technology, Tehran in 2010 and 2014, respectively. Since 2016, he has been an Assistant Professor at the University of Science and Technology of Mazandaran, Behshahr, Iran. His research interests include smart grids, demand response, Electric vehicles, and stochastic optimization.