doi: 10.17586/2226-1494-2023-23-2-374-381

An intelligent shell game optimization based energy consumption analytics model for smart metering data

Ramalingam Saravanan[1]✉, Arulnanthisivam Swaminathan[2], Sankaralingam Balaji[3]

[1,3] Sri Manakula Vinayagar Engineering College, Puducherry, 605107, India
[2] Panimalar Engineering College, Chennai, 600123, India

[1] saravanan@smvec.ac.in✉, https://orcid.org/0000-0003-3503-1133
[2] swamisivam19@gmail.com, https://orcid.org/0000-0001-7672-1339
[3] balaji@smvec.ac.in, https://orcid.org/0000-0002-9013-9801

Abstract
Smart metering is a hot research topic and has gained significant attention since the electromechanical metering is not reliable and requires more energy and time. All the existing methods are focused only on how to deal with data rather than how to do efficiently. Prediction of electricity consumption is essential to gain intelligence to the smart gird. Precise electricity prediction allows a service provided in resource planning and also controlling actions for the demand and supply balancing. The users are beneficial from the smart metering solution by effective interpretation of their energy utilization, and labelling them to efficiently handle the utilization cost. With this motivation, the paper presents intelligent energy consumption analytics using smart metering data (ECA-SMD) model to determine the utilization of energy. The presented ECA-SMD model involves three major processes namely data pre-processing, feature extraction, classification, and parameter optimization. The presented ECA-SMD model uses Extreme Learning Machine (ELM) based classification to determine the optimum class labels. Besides, shell game optimization (SGO) algorithm is applied for tuning the parameters involved in the ELM and boosts the classification efficiency. The efficacy of the ECA-SMD model is validated using an extensive set of smart metering data and the results are investigated based on accuracy and mean square error (MSE). The proposed model exhibited supremacy with the maximum accuracy of 65.917 % and minimum MSE of 0.096.

Keywords
electricity consumption, predictive model, data analytics, smart metering, machine learning

For citation: Saravanan R., Swaminathan A., Balaji S. An intelligent shell game optimization based energy consumption analytics model for smart metering data. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 2, pp. 374–381. doi: 10.17586/2226-1494-2023-23-2-374-381

УДК 004.94

Модель аналитики энергопотребления на основе интеллектуальной оболочки Game Optimization для данных интеллектуального учета

Рамалингам Сараванан[1]✉, Арулнантишивам Сваминатан[2], Санкаралингам Баладжи[3]

[1,3] Инженерный колледж Шри Манакулы Винаягар Пудучерри, Пудучерри, 605107, Индия
[2] Панималарский инженерный колледж, Ченнаи, 600123, Индия

[1] saravanan@smvec.ac.in✉, https://orcid.org/0000-0003-3503-1133
[2] swamisivam19@gmail.com, https://orcid.org/0000-0001-7672-1339
[3] balaji@smvec.ac.in, https://orcid.org/0000-0002-9013-9801

Аннотация
Интеллектуальные измерения привлекают к себе все большое внимание из-за ненадежности электромеханических измерений, больших затрат труда и времени. Существующие методы прогнозирования сосредоточены на работе с данными и не уделяют должного внимания полученным результатам. Точное прогнозирование потребления электроэнергии позволяет предоставлять услуги по планированию ресурсов, контролю

действия по балансированию спроса и предложения. Пользователи получают выгоду при применении интеллектуального учета за счет эффективной интерпретации результатов использования энергии и благодаря экономичному) управлению затратами на электроэнергию. В работе представлена интеллектуальная аналитика энергопотребления с применением модели данных интеллектуального учета ECA-SMD для определения использования энергии. Модель включает предварительную обработку данных, извлечение признаков, классификацию и оптимизацию параметров. Использована классификация на основе машин экстремального обучения (Extreme Learning Machine, ELM) для определения оптимальных меток классов. Применен алгоритм оптимизации Shell Game Optimization для настройки параметров, участвующих в ELM и повышения эффективности классификации. Работоспособность модели ECA-SMD проверена с использованием обширного набора данных интеллектуальных измерений. Предложенная модель показала максимальную точность 65,9 % и среднеквадратичное отклонение 0,096.

Ключевые слова
потребление электроэнергии, прогнозирующая модель, анализ данных, интеллектуальный учет, машинное обучение

Ссылка для цитирования: Сараванан Р., Сваминатан А., Баладжи С. Модель аналитики энергопотребления на основе интеллектуальной оболочки Game Optimization для данных интеллектуального учета // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 2. С. 374–381 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-2-374-381

Introduction

Conventional power grids are being replaced by smart grids which include the use of solar energy and wind energy, and the smart metring is essential to collect the data efficiently. In previous years, smart meters have been rapidly utilized all over the world. At the end of 2018, nearby 86 (UK) and 11 (US) million smart meters have been installed by small and large suppliers. The major aim of smart metering in residential sectors is to inspire the user for consuming lesser energy with increased awareness regarding their consumption level [1]. The smart meter provides data on cost and sum of energy utilization in real world for both consumers and suppliers. This data with incentive programs and demand response will assist them to reduce their energy utilization on peak times and schedule its appliances based on electricity prices [2]. High resolution data created by smart meters, alternatively, give suppliers various managing tasks like power loss detection and power quality monitoring. Furthermore, it unlocks several doors of chances in electricity load analyses like load predicting with higher accurateness at low aggregation levels [3, 4]. The important benefits of the smart metering are automated meter reading, dynamic pricing updates, and early alert of blackouts, efficient energy usage and savings. The load control of smart meters helps the consumers and distributors to disable the meters when the price gets higher which saves the energy when there is scarcity. The data generated from the smart meters are helpful in market demands, abruption of changes, load forecasting through data analytics.

Electrical load predicting is the forecast of the load demand that an electricity user would have later [5]. Load prediction assists suppliers to balance demand and supply also in ensuring the consistency of power grids during power insufficiency. Load predictions are also significant to electricity traders for balancing their electricity sales and purchase [6]. Load predicting is implemented in extensive time-horizon aims at distinct targets: short term load prediction (seconds to one day in advance) for adjusting demand and supply; medium term load forecast (one day to one year in advance) for planning maintenance and outage; and long term load prediction (over a year in advance) for planning the growth of power framework. The predicting process becomes very difficult for low aggregation levels, for example at building level, since several fluctuating factors affect a building energy consumption with differing degrees, such as building properties, weather variables, Ventilating, Heating, and Air Conditioning (HAVC) services and utilization behavior of occupiers [7, 8]. Additionally, several researchers have profited from smart metering information for developing innovative modules to load predicting at separate building levels. The approaches for forecasting building energy consumption are commonly categorized into 2 classes: data driven and engineering (physical) methods. Engineering approach utilizes scientific equations for presenting the physical modules and thermal efficacy of buildings. Nowever, they require higher details regarding distinct variables of the buildings that aren't often presented. Furthermore, a higher level of skill is needed for performing elaborate and expensive computations. Alternatively, data driven methods does not require this complete information regarding the inspired building and rather learns from historical/real world data. This approach is categorized into 2 classes: Artificial Intellegence (AI) based and statistical methods [9, 10].

Statistical approaches utilize historical data as a goal to correlate energy consumption with significant parameters as input. Thus, a large number of historical data with higher quality performs a major part in the efficiency of statistical modules. Conventional linear statistical modules, like Conditional demand analysis, Autoregressive Integrated Moving Average, Gaussian mixture models (GMM), Auto Autoregressive Moving Average, and Regression models, have endured the standard for time sequence forecast with an extensive utilization in several applications [11]. Though it is easier for utilizing statistical methods, the fundamental assumptions of this module are depending upon time sequence that is deliberated linear and thus follows a particularly known distribution of statistics. Several Machine Learning (ML) modules were established to conquer these restrictions. The modules are depending upon Support Vector Machines (SVM), and Classification and Regression Trees, which are between the effective ML methods utilized in time sequence predicting and energy application.

In previous years, several scientists have examined the application of AI based methods in predicting challenges. Amongst AI based methods, Artificial Neural Networks (ANNs) with distinct structures have been extensively utilized in the load predicting field [12]. ANN is equivalent to statistical approaches that utilize historical data for building a module. However, hidden layer structure and learning capability offer numerous benefits on statistical and traditional ML methods for predicting time sequence. They are deliberated data driven and self-adaptive approaches that could take subtle and functional patterns via a trained procedure on historical records of data, when the fundamental relationships among input and output parameters are unknown/complex. However, the neural networks with shallow structures have the drawbacks of considering entire inputs and outputs that are autonomous of one another, while handling consecutive data [13, 14].

This paper presents intelligent Energy Consumption Analytics using Smart Metering Data (ECA-SMD) model to determine the utilization of energy. The presented ECA-SMD model involves three major processes, namely, data pre-processing, feature extraction, classification, and parameter optimization. The presented ECA-SMD model uses Extreme Learning Machine (ELM) based classification to determine the optimum class labels. Besides, Shell Game Optimization (SGO) algorithm is applied for tuning the parameters involved in the ELM and it boosts the classification efficiency. The efficacy of the ECA-SMD model is validated using an extensive set of smart metering data and the results are investigated based on accuracy and Mean Square Error (MSE).

The Proposed ECA-SMD Model

The ECA-SMD model encompasses different processes as shown in Fig. 1, such as data pre-processing, feature extraction, ELM based prediction, and SGO based parameter optimization. Initially, the smart metering data is pre-processed to enhance the quality of the data. Followed by, the features in the pre-processed data are extracted and are then fed into the ELM model to predict the electricity utilization. In order to further improve the predictive performance of the ELM model, the SGO algorithm is applied to optimally alter the parameters involved in the ELM model.

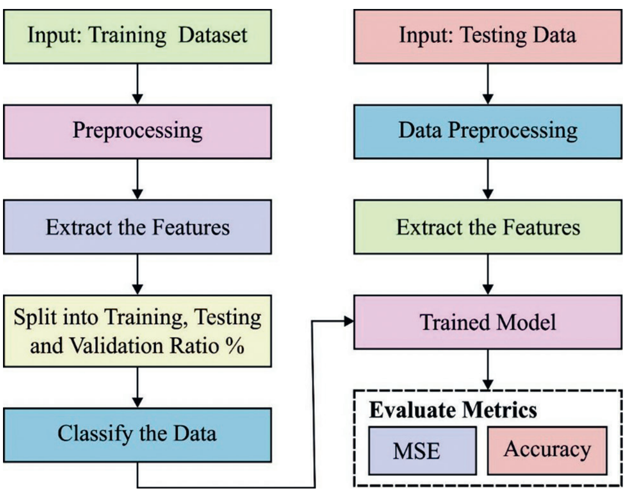

Fig. 1. Working process of ECA-SMD model

Data Pre-processing

The pre-processing phase is the early procedure that is assumed to be major procedure of the energy consumption analyses where the dataset is loaded. The dataset used in the work is smart meter dataset acquired from Kaggle[1] which contains the readings of 5567 households. The dataset contains hourly household energy consumption that is expected and is kept in the local database to compare the real measured data. The Data pre-processing phase is consisting of several procedures, such as data integration, data transformation, data cleaning, and data reduction. Data cleansing is a procedure of finding the missing variables and it fills them with precise values. The processed data is summarized and executed standardization function for reducing the redundant values in the datasets. The data kept in the datasets should be dependent on one another with essential logic between these data. With the removal of redundant data, the data reduction procedure is executed thus to raise the speediness of the data process while relating the trained dataset with the test data.

Feature Extraction

It is the next procedure of defining the accuracy of matching the test data to the trained data. The measured value of voltage, global active reactive powers, power intensity was assumed to be a test value and is related to standard deviation. The low standard deviation produces higher level of accuracy in the feature extraction stage. The time sequence data of the reactive powers and global active have been related to the extracted features. The relation coefficient $C(t)$ between the test and trained data is defined by

$$C(t) = \frac{\left|\sum_{n=1}^{\infty}(x_n - \bar{x}_n\right|}{\sqrt{\left|\sum_{n=1}^{\infty}(x_n - \bar{x}_n\right|}}, \quad (1)$$

where x_n is the training data and $\bar{x}_n$ is testing data.

Eq. (1) represents that the relation coefficient produces the corresponding factor relating the test and trained values of global reactive power, power intensity global active power. The Standard Deviation (SD) for the test and trained values has been as

$$SD(\sigma) = \sqrt{\frac{\Sigma(x_i - \mu)^2}{N}}, \quad (2)$$

where x_i is a value in the data set.

Eq. (2) denotes the scientific form of defining the SD of the prediction value, and it is based on overall electrical appliances in the test residential building (N) and mean of electrical appliances in the target residential building (μ). The approach used in the feature extraction process is Pearson correlation coefficient-based feature selection method. With the help of ranking function, the features

[1] Smart meter data from London area. Available at: https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london (accessed 18.10.2022).

are ranked based on the correlation among the features. Then the optimal features are selected by eliminating the irrelevant features in the dataset based on the rank obtained by the features. The selection of the relevant feature purely depends on the heat consumption and the outside temperatures of the current and previous 21 days.

ELM based Prediction

In order to perform prediction, mathematical model that accepts input and output should be modeled for energy consumption. It is necessary to study the relationship between the input parameters that affect consumption and output values providing the consumption. Next to the feature extraction process, the Extreme Learning Machine (ELM) model is applied to determine the predictive value of the input data. Huang et al. [15] present a novel Neural Network (NN) method named ELM. The ELM topology is a generalized Single hidden Layer Feedforward Network (SLFN) where the input layer weight is set arbitrarily, and the hidden layer weight should be changed. Therefore, in computation terms, ELM is a light weight method. As the trained method isn't a gradient descent-based one, it enables the utilization of diverse activation functions like sinusoid, sigmoid, Gaussian, logistic, identity, Rectifier Linear Unit, Radial Basis Function between another's.

They assumed a regular SLFN with m output node, n input node, and L hidden node. Every hidden node has a similar activation function h. They assume a time sequence $\{x_t\}_{t\in\mathbb{Z}}$ and, moreover, N random pairs (x_j, t_j), whereas

$$\mathbf{x}_j = [x_{j,1}, x_{j,2}, \ldots, x_{j,n}]^T := [x_j, x_{j+1}, \ldots, x_{j+n-1}]^T \in \mathbb{R}^n$$

denotes input vector and where the equivalent output targeted vector t_j for $j = 1, \ldots, N$ is denoted as $\mathbf{t}_j = [t_{j,1}, t_{j,m}]^T \in \mathbb{R}^m$ and also the scalar output of i-th hidden node is denoted as:

$$h(\mathbf{w}_i\mathbf{x}_j + b_i) \in \mathbb{R},\ i = 1, \ldots, L, j = 1, \ldots, N. \quad (3)$$

Eq. (3) denotes the hidden node L, where $\mathbf{w}_i = [w_{i,1}, \ldots, w_{i,n}]^T \in \mathbb{R}^n$ denotes weight vector related to the connection among n input nodes of input layer and i-th hidden node with b_i as bias of the hidden node that is:

$\mathbf{w}_{i,v}$ = *weight associated to the connection between the v-th input layer node and i-th inner layer node*. (4)

Eq. (4) denotes weight vector $\mathbf{w}$, where $v = 1, 2, \ldots, n$ and $i = 1, \ldots, L$. The k-th element of output of the SLFN for input $\mathbf{x}_j$ is assumed as [16]:

$$\mathbf{o}_\mu(x_j) = \sum_{i=1}^{L} \beta_{i,\mu} h(\mathbf{w}_i\mathbf{x}_j + b_i), \quad (5)$$

where $\mu = 1, \ldots, m$, $j = 1, \ldots, N$, $b_i \in \mathbb{R}$ indicates bias of i-th hidden node, and

$\beta_{i,\mu}$ = *weight associated to the connection between the i-th input layer node and* μ*-th inner layer node*. (6)

The output of j-th input vector by the structure of ELM module denoted in eq. (5) is written in matrix form and it is represented as:

$$\overbrace{\begin{bmatrix} o_1(x_1) & \ldots & o_m(x_1) \\ \vdots & \ddots & \vdots \\ o_1(x_N) & \ldots & o_m(x_N) \end{bmatrix}}^{\mathbf{O}\equiv\mathrm{O}(x_j)} =$$

$$= \underbrace{\begin{bmatrix} h(w_1|x_1 + b_1) & \ldots & h(w_L x_1 + b_L) \\ \vdots & \ddots & \vdots \\ h(w_1|x_N + b_1) & & h(w_L x_N + b_L) \end{bmatrix}}_{\mathbf{H}\equiv\mathrm{H}(w_i, x_j, b_i)} \times \quad (7)$$

$$\times \underbrace{\begin{bmatrix} \beta_{1,1} & \ldots & \beta_{1,m} \\ \vdots & \ddots & \vdots \\ \beta_{L,1} & \ldots & \beta_{L,m} \end{bmatrix}}_{\mathbf{B}}.$$

The matrix $\mathbf{H}$ represents *hidden layer output matrix* of the NN (as shown in eq. (7)), matrix $\mathbf{B}$ is explained in eq. (7) also. For training, SLFN should detect, for a presented set of vectors $x_1, \ldots, x_N$, the certain vector weights $w_i = = [w_{1,i}, \ldots, w_{n,i}]^T$ and $\beta_\mu = [\beta_{1,\mu}, \ldots, \beta_{L,\mu}]^T$. The scalar bias b_i, resolves the succeeding minimization problem:

$$\min_{w_i, b_i \beta_{i,\mu}} ||\mathbf{HB} - \mathbf{T}||^2 =$$

$$= \min_{w_i, b_i \beta_{i,\mu}} \sum_{\mu=1}^{m} \sum_{j=1}^{N} \Big(\sum_{i=1}^{L} (\beta_{i,\mu} h(\mathbf{w}_i\mathbf{x}_j + b_i) - t_{i,\mu})^2, \quad (8)$$

where ||.|| denotes typical Frobenius matrix norm, $\mathbf{T} = [t_{i,\mu}]N\times m$ is targeted matrix. Afterward, the variables $\mathbf{w}_i$, $\beta_{i,\mu}$ and b_i, that mentioned in (8), are established in the trained stage, they remain fixed for entire run of *new* vectors $\mathbf{x}_j$. The succeeding results ensure that in the case $L = N$ the problem (8) is precisely resolved "with likelihood one".

Parameter Optimization

Finally, the parameters involved in the ELM model are optimally selected using the SGO algorithm in order to improve the efficiency of the ELM model. The shell game is inspired for inventing a novel optimization method named SGO. Thus, the succeeding assumptions are deliberated:

— in this game, an individual is deliberated as game operator;
— the 3 shells and 1 ball are presented to the operator;
— every player has only 2 chances for guessing the accurate shell.

It can be mathematically defined as follows. Here, a group of N individuals is considered as the game player [17]. In equation below, the location 'd' of player 'I' is given by x_i^d.

$$\mathbf{X}_i = (x_i^1, \ldots, x_i^d, \ldots, x_i^n). \quad (9)$$

In the eq. (9), $\mathbf{X}_i$ denotes arbitrary value for problem parameter. Depending upon $\mathbf{X}_i$, the value of Fitness Function (FF) is calculated for every player.

Once estimating the FF values for every player, the game operators select 3 shells in which all the shells are interrelated to the location of an optimum player and 2 other shells are selected arbitrarily using

$$game's\ operator: \begin{cases} shell_1 = ball = X_{best} \\ shell_2 = X_{k_1} \\ shell_3 = X_{k_2} \end{cases}, \quad (10)$$

where, X_{best} denotes location of minimum (in minimization problem) or maximum (in maximization problem) of fitness; X_{k_1} and X_{k_2} indicate position of 2 members of the population; k_1 and k_2 denote arbitrary numbers among 1 to N that is selected arbitrarily. Once estimating the FF and finding the shell for every player, accuracy and intelligence of the player must be calculated in this phase. Every player guesses the shell depending on whether each player is determined based on fitness accuracy and intelligence normalized value using.

$$AI_i = \frac{fit_i - fit(X_{worst})}{\sum_{j=1}^{N}[(fit_j - fit(X_{worst})]}, \quad (11)$$

where AI_i denotes accuracy and intelligence of player i and X_{worst} represents location of minimum (in maximization problem) or maximum (in minimization problem) of fitness.

Here, the player is prepared for guessing the ball. Assuming that the game is played with 3 shells and every player has 2 opportunities, there are 3 states of guess for every player. In initial state, the initial guess might be right and the position of the ball would be identified. In next state, the player later an incorrect guess in the initial choice might guess the ball position in the next time. Lastly, in the third state, both guesses of players might be incorrect and therefore the player was ineffective to identify the ball position. The guess vector detailed by $\mathbf{G}_v$ is inspired using for every player the following equation:

$$\mathbf{G}_v(x) = \begin{cases} state1: [1\ 0\ 0],\ at\ first \\ state2: \begin{cases} [0.5\ 0.5\ 0] \\ 0.5\ 0\ 0.5 \end{cases},\ at\ second. \\ state3: [0\ 0.5\ 0.5],\ else \end{cases} \quad (12)$$

The probability of choosing one of the states for shell selection is simulated by

$$state = \begin{cases} state1: ifAI_i > r_{g1} \\ state2: ifAI_i > r_{g2} \\ state3 = else \end{cases}, \quad (13)$$

where r_{g1} denotes probability of right guess at the initial choice and r_{g2} represents probability of accurate guess at the next time. Lastly, $\mathbf{X}_i$ vector that is considered as the position of every member of population is upgraded based on equations

$$dx^d_{i,ball} = r_1(ball - x_i{}^d)state(1,1) \quad (14)$$

$$dx^d_{i,shell_2} = r_2(shell_2{}^d - x_i{}^d)sign(fit_i - fit_{shell_2})state(1,2) \quad (15)$$

$$dx^d_{i,shell_3} = r_3(shell_3{}^d - x_i{}^d)sign(fit_i - fit_{shell_3})state(1,3) \quad (16)$$

$$x_i{}^d = x_i{}^d + dx^d_{i,ball} + dx^d_{i,shell_2} + dx^d_{i,shell_3}, \quad (17)$$

where r_i indicates arbitrary value in the range of zero and one, $dx^d_{i,ball}$, $dx^d_{i,shell_2}$, and $dx^d_{i,shell_3}$ denotes displacement of dimension 'd' of player 'i' according to $shell_1$, $shell_2$, and $shell_3$.

The steps of SGO is generalized by:

Step 1: Arbitrary creation of early population by eq. (9)

Step 2: Evaluating the fitness value of agents

Step 3: Choice of i-th member

Step 4: Choosing 3 shells by eq. (10)

Step 5: Estimation of accuracy and intelligence (AI) by eq. (11)

Step 6: Determining the state of guess by eqs. (12) and (13)

Step 7: Choice of d-th dimension of i-th member

Step 8: Evaluating $dx^d_{i,ball}$, $dx^d_{i,shell_2}$, and $dx^d_{i,shell_3}$ using eqs. (14)–(16)

Step 9: Upgrading position of d-th dimension of i-th member by eq. (17)

Step 10: When each dimension of i-th member are upgraded, go to Step 11, otherwise return Step 7

Step 11: When each member is upgraded, go to Step 12, otherwise return Step 3

Step 12: When the end criteria are recognized, go to Step 13, otherwise return Step 2

Step 13: Print the optimum solution.

Performance Validation

This section validates the performance of the ECA-SMD model with other existing methods such as ANN and SVM. The results are examined in terms of MSE and accuracy. For improved predictive results, the value of accuracy should be maximum and MSE value should be minimum. Table 1 provides the comparative results analysis of the ECA-SMD model in terms of accuracy and MSE. The dataset is processed and aggregated in to hourly data. We used 3 months data of households for prediction. The forecasting is done using the proposed model and the performance is evaluated by comparing the proposed model with other models. The performance metrics used in the work is MSE, and its accuracy is shown in Table 1.

An accuracy analysis of the ECA-SMD model is made with the existing methods in Fig. 2. The figure showcased that the ANN model had shown poor performance and resulted in a lower accuracy value. At the same time, the SVM model demonstrated slightly enhanced outcome over the ANN but not higher than the proposed ECA-SMD model. For instance, at T_1 Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 73 %, whereas the ANN and SVM models have achieved lower accuracy of 67 % and 70 %, respectively. In addition, at T_4 Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 71 %, whereas the ANN and SVM models have achieved lower accuracy of 69 % and 57 %, respectively. Also, at T_8 Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 71 %, whereas the ANN and SVM models have achieved lower accuracy of 66 % and 61 %, respectively. Additionally, at T_{12} Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 59 %, whereas the ANN and SVM models have achieved lower accuracy of 57 % and 52 %, respectively. Besides, at T_{16} Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 63 %, whereas the ANN and SVM models have achieved lower

Table 1. Result Analysis of Existing with Proposed ECA-SMD Model in terms of MSE and Accuracy

Hour	ANN		SVM		ECA-SMD	
	Accuracy, %	MSE	Accuracy, %	MSE	Accuracy, %	MSE
T_1	67	0.10	70	0.08	73	0.06
T_2	59	0.11	59	0.12	62	0.10
T_3	66	0.10	58	0.12	71	0.09
T_4	69	0.09	57	0.12	71	0.08
T_5	66	0.10	58	0.12	68	0.09
T_6	63	0.10	58	0.11	68	0.09
T_7	60	0.11	59	0.11	61	0.10
T_8	66	0.10	61	0.11	71	0.09
T_9	66	0.10	60	0.12	67	0.10
T_{10}	61	0.11	59	0.12	64	0.10
T_{11}	62	0.11	57	0.12	66	0.11
T_{12}	57	0.12	52	0.12	59	0.10
T_{13}	59	0.12	55	0.12	62	0.12
T_{14}	57	0.12	60	0.11	65	0.10
T_{15}	61	0.12	56	0.13	63	0.10
T_{16}	62	0.12	60	0.12	63	0.10
T_{17}	60	0.11	59	0.12	65	0.10
T_{18}	67	0.10	60	0.11	68	0.09
T_{19}	62	0.12	68	0.11	70	0.11
T_{20}	59	0.12	62	0.11	65	0.11
T_{21}	66	0.11	63	0.10	68	0.09
T_{22}	58	0.12	61	0.10	65	0.08
T_{23}	57	0.12	60	0.11	62	0.10
T_{24}	61	0.11	61	0.13	65	0.10
Avg.	**62.130**	**0.110**	**59.708**	**0.114**	**65.917**	**0.096**

accuracy of 62 % and 60 %, respectively. Moreover, at T_{20} Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 65 %, whereas the ANN and SVM models have achieved lower accuracy of 59 % and 62 %, respectively. Furthermore, at T_{24} Hour, the proposed ECA-SMD model has attained effective outcome with higher accuracy of 65 %, whereas the ANN and SVM models have achieved lower accuracy of 61 % and 61 %, respectively.

A brief MSE analysis of the ECA-SMD model takes place with the existing techniques in Fig. 3. The figure depicted that the ANN and SVM models have failed to outperform the proposed ECA-SMD model which has achieved least MSE values. For instance, at T_1 Hour,

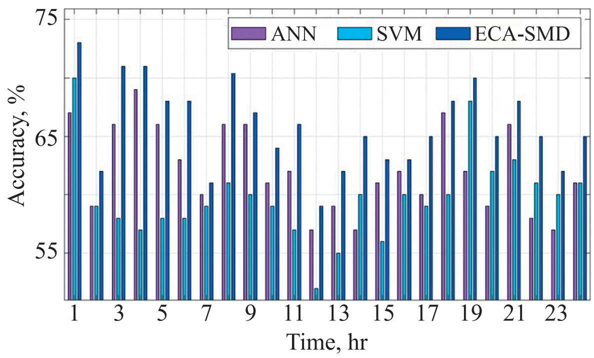

Fig. 2. Result analysis of ECA-SMD model in terms of accuracy

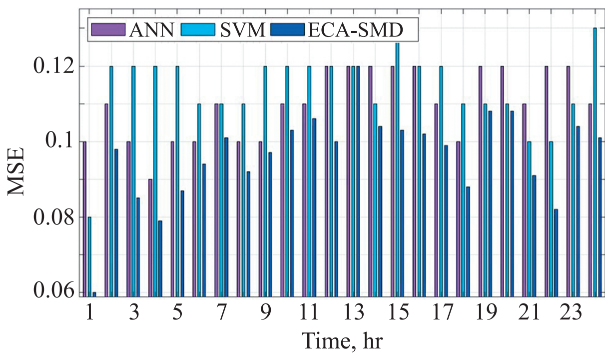

Fig. 3. Result analysis of ECA-SMD model interms of MSE

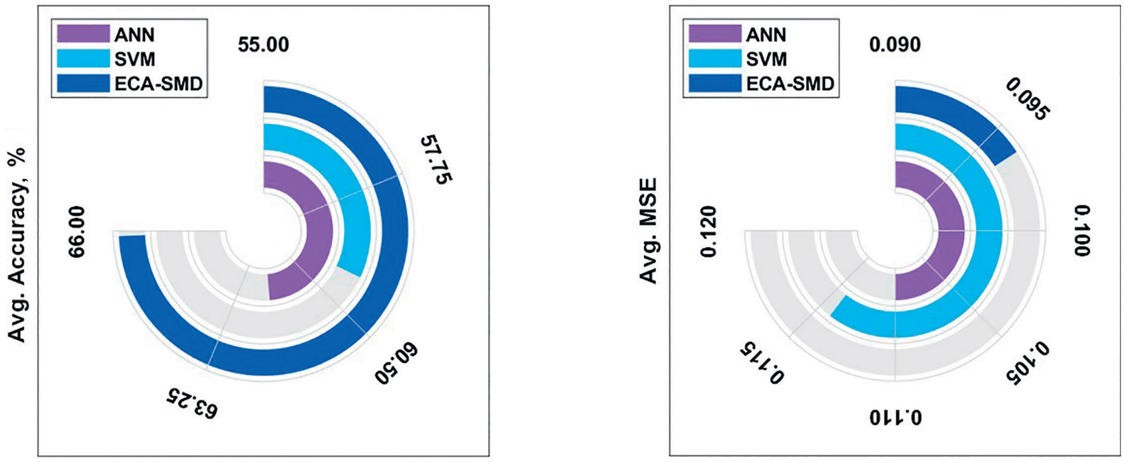

Fig. 4. Average accuracy analysis (*a*) and MSE (*b*) of ECA-SMD model

a minimal MSE of 0.06 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.10 and 0.08 have been gained by the ANN and SVM models. In the meantime, at T_4 Hour, a minimal MSE of 0.08 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.09 and 0.12 have been gained by the ANN and SVM models. At the same time, at T_8 Hour, a minimal MSE of 0.09 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.10 and 0.11 have been gained by the ANN and SVM models. Meanwhile, at T_{12} Hour, a minimal MSE of 0.10 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.12 and 0.12 have been gained by the ANN and SVM models. In line with, at T_{16} Hour, a minimal MSE of 0.10 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.12 and 0.12 have been gained by the ANN and SVM models. Along with that, at T_{20} Hour, a minimal MSE of 0.11 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.12 and 0.11 have been gained by the ANN and SVM models. Simultaneously, at T_{24} Hour, a minimal MSE of 0.10 has been accomplished by the ECA-SMD model, whereas a maximum MSE of 0.11 and 0.13 have been gained by the ANN and SVM models.

Fig. 4, *a* portrays the average accuracy analysis of the ECA-SMD model with the existing ANN and SVM models. From the figure, it is depicted that the SVM model has achieved least performance over the other methods with the reduced average accuracy of 59.708 %, whereas the ANN model has demonstrated slightly enhanced performance with an average accuracy of 62.13 %. But the proposed ECA-SMD model has resulted in a maximum average accuracy of 65.917 %.

Fig. 4, *b* depicts the average MSE analysis of the ECA-SMD model with the existing ANN and SVM models. From the figure, it can be seen that the SVM model has realized minimum performance over the other methods with the increased average MSE of 0.114, whereas the ANN model has established somewhat improved outcomes with an average MSE of 0.110. However, the proposed ECA-SMD model has offered superior performance with the lowest average MSE of 0.096.

Conclusion

The ECA-SMD model encompasses different processes, such as data pre-processing, feature extraction, ELM based prediction, and SGO based parameter optimization. Initially, the smart metering data is pre-processed to enhance the quality of the data. Followed by, the features in the pre-processed data are extracted and are then fed into the ELM model to predict the electricity utilization. In order to further improve the predictive performance of the ELM model, the SSO algorithm is applied to optimally alter the parameters involved in the ELM model. The efficacy of the ECA-SMD model is validated using an extensive set of smart metering data, and the results are investigated based on accuracy and MSE. The proposed model exhibited supremacy with a maximum accuracy of 65.917 % and minimum MSE of 0.096. In future, the proposed ECA-SMD model can be extended to the utilization of deep learning architectures.

References

1. Mehdipour Pirbazari A., Farmanbar M., Chakravorty A., Rong C. Short-term load forecasting using smart meter data: A generalization analysis. *Processes*, 2020, vol. 8, no. 4, pp. 484. https://doi.org/10.3390/pr8040484
2. Uthayakumar J., Metawa N., Shankar K., Lakshmanaprabu S.K. Intelligent hybrid model for financial crisis prediction using machine learning techniques. *Information Systems and e-Business Management*, 2020, no. 4, pp. 617–645. https://doi.org/10.1007/s10257-018-0388-9

Литература

1. Mehdipour Pirbazari A., Farmanbar M., Chakravorty A., Rong C. Short-term load forecasting using smart meter data: A generalization analysis // Processes. 2020. V. 8. N 4. P. 484. https://doi.org/10.3390/pr8040484
2. Uthayakumar J., Metawa N., Shankar K., Lakshmanaprabu S.K. Intelligent hybrid model for financial crisis prediction using machine learning techniques // Information Systems and e-Business Management. 2020. N 4. P. 617–645. https://doi.org/10.1007/s10257-018-0388-9

3. Wang Y., Chen Q., Hong T., Kang C. Review of smart meter data analytics: applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, 2019, vol. 10, no. 3, pp. 3125–3148. https://doi.org/10.1109/tsg.2018.2818167
4. Farmanbar M., Parham K., Arild Ø., Rong C. A widespread review of smart grids towards smart cities. *Energies*, 2019, vol. 12, no. 23, pp. 4484. https://doi.org/10.3390/en12234484
5. Rajakumar R., Sivanandakumar D., Uthayakumar J., Vengattaraman T., Dinesh K. Optimal parameter tuning for PID controller using accelerated grey wolf optimisation in smart sensor environments. *Electronic Government, an International Journal*, 2020, vol. 16, no. 1-2, pp. 170–189. https://doi.org/10.1504/eg.2020.105237
6. Ryu S., Noh J., Kim H. Deep neural network based demand side short term load forecasting. *Energies*, 2016, vol. 10, no. 1, pp. 3. https://doi.org/10.3390/en10010003
7. Mocanu E., Nguyen P.H., Gibescu M., Kling W.L. Comparison of machine learning methods for estimating energy consumption in buildings. *Proc. of the 2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2014, pp. 1–6. https://doi.org/10.1109/pmaps.2014.6960635
8. Pirbazari A.M., Chakravorty A., Rong C. Evaluating feature selection methods for short-term load forecasting. *Proc. of the 2019 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 2019, pp. 1–8. https://doi.org/10.1109/bigcomp.2019.8679188
9. Zhao H., Magoulès F. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 2012, vol. 16, no. 6, pp. 3586–3592. https://doi.org/10.1016/j.rser.2012.02.049
10. Pérez-Lombard L., Ortiz J., Pout C. A review on buildings energy consumption information. *Energy and Buildings*, 2008, vol. 40, no. 3, pp. 394–398. https://doi.org/10.1016/j.enbuild.2007.03.007
11. Hyndman R., Koehler A., Ord K., Snyder R. *Forecasting with Exponential Smoothing*. Berlin/Heidelberg, Germany, Springer, 2008, 362 p. https://doi.org/10.1007/978-3-540-71918-2
12. Ahmad A.S., Hassan M.Y., Abdullah M.P., Rahman H.A., Hussin F., Abdullah H., Saidur R. A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews*, 2014, vol. 33, pp. 102–109. https://doi.org/10.1016/j.rser.2014.01.069
13. Khashei M., Bijari M. An artificial neural network (p,d,q) model for timeseries forecasting. *Expert Systems with Applications*, 2010, vol. 37, no. 1, pp. 479–489. https://doi.org/10.1016/j.eswa.2009.05.044
14. Murugan B.S., Elhoseny M., Shankar K., Uthayakumar J. Region-based scalable smart system for anomaly detection in pedestrian walkways. *Computers & Electrical Engineering*, 2019, vol. 75, pp. 146–160. https://doi.org/10.1016/j.compeleceng.2019.02.017
15. Huang G.B., Zhu Q.Y., Siew C.K. Extreme learning machine: a new learning scheme of feedforward neural networks. *Proc. of the 2004 IEEE International Joint Conference on Neural Networks. V. 2*, 2004, pp. 985–990. https://doi.org/10.1109/ijcnn.2004.1380068
16. Fernández C., Salinas L., Torres C.E. A meta extreme learning machine method for forecasting financial time series. *Applied Intelligence*, 2019, vol. 49, no. 2, pp. 532–554. https://doi.org/10.1007/s10489-018-1282-3
17. Dehghani M., Montazeri Z., Malik O.P., Givi H., Guerrero J.M. Shell game optimization: A novel game-based algorithm. *International Journal of Intelligent Engineering and Systems*, 2020, vol. 13, no. 3, pp. 246–255. https://doi.org/10.22266/ijies2020.0630.23

3. Wang Y., Chen Q., Hong T., Kang C. Review of smart meter data analytics: applications, methodologies, and challenges // IEEE Transactions on Smart Grid. 2019. V. 10. N 3. P. 3125–3148. https://doi.org/10.1109/tsg.2018.2818167
4. Farmanbar M., Parham K., Arild Ø., Rong C. A widespread review of smart grids towards smart cities // Energies. 2019. V. 12. N 23. P. 4484. https://doi.org/10.3390/en12234484
5. Rajakumar R., Sivanandakumar D., Uthayakumar J., Vengattaraman T., Dinesh K. Optimal parameter tuning for PID controller using accelerated grey wolf optimisation in smart sensor environments // Electronic Government, an International Journal. 2020. V. 16. N 1-2. P. 170–189. https://doi.org/10.1504/eg.2020.105237
6. Ryu S., Noh J., Kim H. Deep neural network based demand side short term load forecasting // Energies. 2016. V. 10. N 1. P. 3. https://doi.org/10.3390/en10010003
7. Mocanu E., Nguyen P.H., Gibescu M., Kling W.L. Comparison of machine learning methods for estimating energy consumption in buildings // Proc. of the 2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). 2014. P. 1–6. https://doi.org/10.1109/pmaps.2014.6960635
8. Pirbazari A.M., Chakravorty A., Rong C. Evaluating feature selection methods for short-term load forecasting // Proc. of the 2019 IEEE International Conference on Big Data and Smart Computing (BigComp). 2019. P. 1–8. https://doi.org/10.1109/bigcomp.2019.8679188
9. Zhao H., Magoulès F. A review on the prediction of building energy consumption // Renewable and Sustainable Energy Reviews. 2012. V. 16. N 6. P. 3586–3592. https://doi.org/10.1016/j.rser.2012.02.049
10. Pérez-Lombard L., Ortiz J., Pout C. A review on buildings energy consumption information // Energy and Buildings. 2008. V. 40. N 3. P. 394–398. https://doi.org/10.1016/j.enbuild.2007.03.007
11. Hyndman R., Koehler A., Ord K., Snyder R. Forecasting with Exponential Smoothing. Berlin/Heidelberg, Germany: Springer, 2008. 362 p. https://doi.org/10.1007/978-3-540-71918-2
12. Ahmad A.S., Hassan M.Y., Abdullah M.P., Rahman H.A., Hussin F., Abdullah H., Saidur R. A review on applications of ANN and SVM for building electrical energy consumption forecasting // Renewable and Sustainable Energy Reviews. 2014. V. 33. P. 102–109. https://doi.org/10.1016/j.rser.2014.01.069
13. Khashei M., Bijari M. An artificial neural network (p,d,q) model for timeseries forecasting // Expert Systems with Applications. 2010. V. 37. N 1. P. 479–489. https://doi.org/10.1016/j.eswa.2009.05.044
14. Murugan B.S., Elhoseny M., Shankar K., Uthayakumar J. Region-based scalable smart system for anomaly detection in pedestrian walkways // Computers & Electrical Engineering. 2019. V. 75. P. 146–160. https://doi.org/10.1016/j.compeleceng.2019.02.017
15. Huang G.B., Zhu Q.Y., Siew C.K. Extreme learning machine: a new learning scheme of feedforward neural networks // Proc. of the 2004 IEEE International Joint Conference on Neural Networks. V. 2. 2004. P. 985–990. https://doi.org/10.1109/ijcnn.2004.1380068
16. Fernández C., Salinas L., Torres C.E. A meta extreme learning machine method for forecasting financial time series // Applied Intelligence. 2019. V. 49. N 2. P. 532–554. https://doi.org/10.1007/s10489-018-1282-3
17. Dehghani M., Montazeri Z., Malik O.P., Givi H., Guerrero J.M. Shell game optimization: A novel game-based algorithm // International Journal of Intelligent Engineering and Systems. 2020. V. 13. N 3. P. 246–255. https://doi.org/10.22266/ijies2020.0630.23

Authors

Ramalingam Saravanan — PhD, Associate Professor, Sri Manakula Vinayagar Engineering College, Puducherry, 605107, India, SC 57211236494, https://orcid.org/0000-0003-3503-1133, saravanan@smvec.ac.in

Arulnanthisivam Swaminathan — PhD, Associate Professor, Panimalar Engineering College, Chennai, 600123, India, SC 55624126400, https://orcid.org/0000-0001-7672-1339, swaminathan@panimalar.ac.in

Sankaralingam Balaji — PhD, Associate Professor, Sri Manakula Vinayagar Engineering College, Puducherry, 605107, India, SC 55310821000, https://orcid.org/0000-0002-9013-9801, balaji@smvec.ac.in

Авторы

Сараванан Рамалингам — PhD, доцент, Инженерный колледж Шри Манакулы Винаягар Пудучерри, Пудучерри, 605107, Индия, SC 57211236494, https://orcid.org/0000-0003-3503-1133, saravanan@smvec.ac.in

Сваминатан Арулнантишивам — PhD, доцент, Панималарский инженерный колледж, Ченнаи, 600123, Индия, SC 55624126400, https://orcid.org/0000-0001-7672-1339, swamisivam19@gmail.com

Баладжи Санкаралингам — PhD, доцент, Инженерный колледж Шри Манакулы Винаягар Пудучерри, Пудучерри, 605107, Индия, SC 55310821000, https://orcid.org/0000-0002-9013-9801, balaji@smvec.ac.in

Received 14.06.2022
Approved after reviewing 10.02.2023
Accepted 31.03.2023

Статья поступила в редакцию 14.06.2022
Одобрена после рецензирования 10.02.2023
Принята к печати 31.03.2023