

Multi Objective Fuzzy Network Based Big Data Analytics System for Energy Consumption Prediction in Smart Grid

¹Praveen Kumar R, ²Ravi Kumar, ³Gokulakrishnan S, ⁴Devi Kosuru S N V J, ⁵Neeraj Shrivastava and ⁶Jyothi P

¹Department of Electronics and Communication Engineering, SRM Easwari Engineering College, Chennai, Tamil Nadu, India.

²Department of Electronics and Communication Engineering, Jaypee University of Engineering and Technology, Raghogarh, Guna, Mathya Pradesh, India.

³Department of Computer Science and Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India.

⁴Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

⁵Engineering Trainer and Assessor, Dubai, United Arab Emirates.

⁶Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India.

¹rpjcspraveen@gmail.com, ²ravi.kumar6@gmail.com, ³s.gokulakrishnaan@gmail.com, ⁴jyotsnakosuru@gmail.com, ⁵njcool21@gmail.com, ⁶jyothi_p@vnrvjiet.in

Correspondence should be addressed to Praveen Kumar R : rpjcspraveen@gmail.com

Article Info

Journal of Machine and Computing (<https://anapub.co.ke/journals/jmc/jmc.html>)

Doi : <https://doi.org/10.53759/7669/jmc202505207>

Received 18 March 2025; Revised from 21 June 2025; Accepted 15 September 2025.

Available online 05 October 2025.

©2025 The Authors. Published by AnaPub Publications.

This is an open access article under the CC BY-NC-ND license. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Abstract – In the modern era, smart grids are generates vast amount of heterogeneous data, which are making the challenges for effective energy management systems. Consequently, the prediction of energy consumption is important for reducing the operational cost, ensuring grid stability and supporting maintainable energy usage. Moreover, the study of energy usage in buildings is popular due to its significant environmental and financial implications. Smart metering can improve building energy efficiency and management. However, the existing prediction strategies are failed to manage the large scale energy data, dynamic behaviour of grid data and uncertainty due to the diversity of family home types and the non-linear trends. Therefore, a specialized models that consider both temporal and spatial factors are required. This study aims to address these issues through a Multi-objective Golden Eagle Optimization-based (MGEO) hybrid deep dynamic convolutional fuzzy network (HDDCFN) method. Multi-objective optimization is used to optimize the hyperparameters of the algorithm and prediction model. Regarding historical energy usage and weather conditions, the suggested method incorporates a combination of linear and non-linear predictions. Here, fuzzy logic is used to manage the inaccurate and uncertainty data while applying the Multi-objective Golden Eagle Optimization model. Here, large range of data were processed with big data analytics and provides adaptive and reliable prediction outcomes. Moreover, the developed approach improves the energy distribution, load forecasting and demand response performance. The simulation was done using MATLAB and the developed system has demonstrates lower Mean Absolute Error (MAE) of 0.24, decreases the Mean Square Error (MSE) of 0.27 and enhance the R^2 of 0.92 while comparing the traditional models.

Keywords – Artificial Intelligence, Energy Prediction, Energy Time Series, Smart Meters, Smart Grid, Optimization.

I. INTRODUCTION

Greater numbers of individuals are choosing to live in cities rather than rural regions, [1] perhaps to take advantage of the cutting-edge amenities of smart cities [2]. Residential building energy consumption has skyrocketed in recent years due to technology breakthroughs, increased urbanization, more time spent indoors, and so on [3]. Smart System (SS) is a technically sophisticated electrical grid; its development is critical for dependable and economic power for smart city users [4]. Physical infrastructure is converted into digital infrastructure, and traditional analog systems are replaced by enhanced digital and power electronics [5]. This technology improves the existing grid's efficiency and reliability by minimizing outages and giving the grid self-healing or auto-restore capabilities [6]. When an outage occurs, electricity is instantly

diverted, and power is restored to the affected region. Furthermore, because SG is technologically advanced, it includes multiple energy-measuring technologies, such as smart meters and advanced metering infrastructure (AMI) [7].

Smart meters allow bidirectional connections between consumers and utility providers, so, the final results has generates the wide range of data with various velocity and veracity across the abundant computation environment [8]. Here, the generated data is based on time-series concepts and consists of energy use assessments of component equipment over a timespan [9].

Moreover, In addition, the emergence of big data technology has further revolutionized the ability of valuable providers for understanding the energy optimization, demand prediction, energy consumption and other patterns etc. Consequently, data driven and decision making can enable to handle the high-volume data and data streams [10][11].

Researchers have used big data to create dependable and effective energy management systems [12]. Integrating data analytics with a scalable selection process allows for consistent the predictions of demand and supply of energy [13], as reported for the smart city of Singapore [14]. Researchers used SG to show how tools can be used to handle big data strategies [15]. Even though accomplishing the objectives of these initiatives has been somewhat successful, long-lasting outcomes still need to be achieved [16]. This lack of outcomes is due to the difficulty in understanding each consumer's habits individually and tailoring tactics that consider the advantages versus drawbacks of changing behavior in accordance with advised energy-saving schemes [17]. Additionally, there is a dynamic interaction between human behavior and the factors influencing energy consumption patterns. Therefore, it is crucial to create models that can analyze energy time-series data from smart meters and intelligently infer and anticipate energy consumption [18].

The body of knowledge used to predict electric usage is expanding daily. To create consumption-forecasting models that are effective and more widely applicable, therefore various researchers are utilized an advanced data analytic system such as machine learning (ML), AI and deep learning (DL) [19]. The short-term load forecasting (STLF) [20] models that have been previously developed statistical approaches, including autoregressive moving average (ARMA) [21], artificial neural network (ANN) [22], autoregressive integrated moving average (ARIMA) [23], convolutional neural network (CNN) [24] and support vector regression (SVR) [25]. The capacity to extend is the primary metric for measuring model performance in neural networks, and it is significantly influenced by the structure used. For example, large learning mistakes will occur if there are too few nodes, whereas overfitting may occur if there are too many nodes. The generalization capacity of the neural network system is directly impacted by overfitting and overtraining, which might result from an inappropriate structure. Large datasets are needed to train statistical techniques, and feature collinear factors, making it challenging to forecast dependent variables that alter the statistical significance of regression models due to the connection between possible predictor variables. As a result, traditional models can only function reasonably when applied to linear situations and they are unsuitable for capturing complicated electric consumption patterns.

This study aims to propose models based on data analysis of energy with time series, prediction and recognition of various energy consumption series which are closely correlated to customer behaviour and anticipated comfort. This research's main contribution can be summarised as follows:

Consider the energy-efficient framework for a big data analytics-based smart grid data metering system. To define the detailed description based on residential buildings, a linear component of dataset-based time-series model was captured in terms of historical energy consumption and meteorological condition [26]. The result of the linear stage is prediction performance, residuals, and forecasted energy consumption. The historical energy consumption, linear stage prediction performance, forecasted energy consumption, and other meteorological conditions such as temperature, humidity, apparent temperature, visibility, pressure, wind bearing, wind speed, cloud cover, precipitation intensity, precipitation probability, dew point, etc., are stored in the database. Data preparation is the initial stage. Cleaning and performing transformations on the original dataset are the goals of this step. The Multi-Objective Golden Eagle Optimization-based (MGE) [27] Hybrid Deep Dynamic Convolutional Fuzzy Network (HDDCFN) [28] [29] method is proposed to predict the energy consumption of the system. The multi-objective optimization optimizes the hyperparameter of the algorithm and prediction model. Finally, the performance measures are validated for the overall system. The remainder of this essay is structured as follows:

The available research on predicting electric demand and consumption profile clustering is reviewed in Section 2. The proposed technique for this study is described in Section 3. The results and findings of the experiment are covered in Section 4. The conclusion of the research is mentioned in Section 5, which offers suggestions for the future.

II. RELATED WORK

The Following Is a Summary of Some Recent Literature That Is Relevant to This Study

Responsible, efficient, and ecologically conscious energy consumption practices are increasingly essential for a dependable contemporary power system. Singh et al. (2018) [30] provided an intelligent data mining methodology to analyze, predict, and display energy time-series data and reveal diverse temporal energy usage patterns. Based on energy time-series data, they suggested unsupervised data clustering, frequent pattern mining analysis, and Bayesian network forecasting for energy consumption predictions. The proposed model's short-term hourly energy consumption prediction accuracy was 81.89%, while long-term (day, month, week, and season) energy consumption prediction accuracy was 75.88%, 74.74%, 79.23%, and 72.81%, respectively.

Izidio et al. (2021) [31] suggested an Evolutionary Hybrid System (EvoHyS) that integrates statistical and machine learning approaches through error series modelling for consumption electricity usage of a smart grid. EvoHyS is made up of the Seasonal ARIMA (SARIMA) framework, ML, and a supplementary ML model that combines both linear and non-linear predictions. Also, Genetic Algorithm (GA) has utilized for selecting features and hyperparameter adjustment to increase its accuracy under the presence of EvoHyS uses a, An experimental analysis was carried out to validate the presented EvoHyS approach and the outcomes are determined in terms of standard and well-known measures like Mean Square Error (MSE).

Kumar et al. (2022) suggested a smart energy consumption analysis and alert system [32] to keep the user informed and to prevent excessive energy consumption. It tries to create a circuit that enables the user to manage energy consumption without incurring additional costs by enhancing intelligent electrical energy-saving techniques. According to Summit Energy, a smart energy management system (SEMS) would assist with cost reduction. In addition, residential, commercial, and manufacturing usage has been employed in the Internet of Things (IoT) and big data initiatives to effectively accomplish electricity.

To forecast short-term power consumption, Khan et al. (2021) [33] introduced a temporal and spatial-based ensemble forecasting model. Moreover, to analyse the electric utilization level of the higher buildings a clustering procedure with k-means models are integrated. Also, Long short with gated recurrent memory networks are combined to perform the function with an effective manner.

By achieving high results with the least MAPE value is 4.2 and 4.5 at higher building and floor area forecasting results, correspondingly, the results obtained by their suggested ensemble method demonstrates the developed model has better performance while comparing the electrical consumption of other sequential behavior approaches.

To identify the pattern of urban energy consumption, Niu et al. (2021) [34] suggested a unique spatiotemporal visual analytic technique to find possible energy-saving opportunities, plan energy supply, and boost energy efficiency. Additionally, they suggested a feasible flow approach for modelling energy demand change patterns and integrated it into a tool for pattern research. The outcomes show the value of the suggested strategy, and the tool may be included in smart energy systems for a deeper comprehension of user energy consumption habits and preferences.

A single complete LSTM built on a sequential network was proposed by Fatema et al. (2021) [35] as a method for forecasting power consumption and pricing from massive datasets related to smart cities. The efficiency of the suggested approach is verified using actual energy market information from the Australian Energy Market Operator (AEMO). According to the validation results, the developed model is a superior choice for predicting the demand and price for energy with a tolerable lower error.

Ünal et al. (2021) [36] created a sophisticated pre-processing method in conjunction with a hybrid sequential learning integrated energy forecasting model that incorporates a CNN and bidirectional long short-term memory (BiLSTM) into a single framework for precise energy consumption forecasting methods. The prediction step is when the innovative hybrid deep learning technique based on data feature decoding and coding is implemented.

To choose the most effective hybrid design for energy consumption prediction, Haq et al. (2021) [37] offered a comparative examination of a range of deep characteristics using multiple sequential learning models. Convolutional long short-term memory (ConvLSTM) combined with BiLSTM produces the most incredible performance. When applied to publicly accessible datasets, such as the residential and Korean commercial building datasets, the suggested energy consumption-forecasting model beats the existing algorithms.

Due to the ineffective parameter setup, traditional ML algorithms are prone to become trapped in local minima. For instance, neural networks have intrinsic complexity and poor generalization, which causes overfitting problems.

III. PROPOSED METHODOLOGY

This article proposed a novel hybrid big data analytics-based energy-efficient framework for the smart grid to minimize energy consumption. The suggested model is shown in **Fig 1**, along with its many phases, including data pre-processing, training and testing, prediction model, and prediction optimization. The first stage involves the linear energy consumption data based on historical energy consumption and meteorological condition data. Furthermore, the linear time-series model is developed for the prediction performance of the energy consumption and residuals. Then, the linear and time-series model data is given to the non-linear part of energy consumption system. Next, the database is created in a storage system. Consequently, the raw data preparation and processing is executed, including millions of records of energy time-series data consumption. In this step, the unwanted and missing data are cleared and corrected to support accurate predictions.

This article presents a novel MGE-based HDDCFN prediction model to detect energy consumption accurately. The data are split into test and training phases. The training phases are executed based on the proposed hybrid deep learning model with tuning the hyperparameter and the parameters are optimized by MGE method. Moreover, the prediction model performance is optimized by the MGE algorithm. Finally, the performance measures are estimated to validate our proposed research's effectiveness.

Data Pre-Processing

Data preparation changes the raw data into a usable and effective format. Numerous arbitrary variables, including demographic, financial, cultural, climatic circumstances, and so on, can change the performance of the demand for and the

cost of electricity. Because original data quality might vary, pre-processing the data is important to increase the precision of model predictions. Data for this investigation are pre-processed using normalization. Normalization is a common technique used in data pre-processing, which transforms the numerical data into a standard format. This helps to enhance the prediction accuracy of the machine learning models. In the proposed work, the normalization was performed before data training combining all the training data features into a similar scale. Scaling the data involves dividing each feature value by the standard deviation of that feature, which results in a new standard deviation of one for each feature. Here, the normalization is estimated based on the ratio of difference between the mean from the original data to the standard deviation. Thus, it enables the proposed model to learn effectively and enhance the accuracy of the energy consumption predictions.

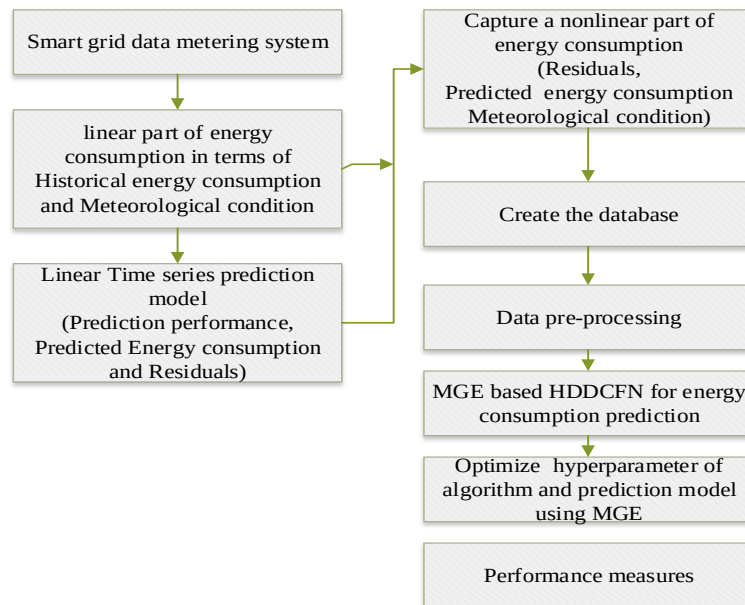


Fig 1. Proposed Methodology.

Proposed Algorithm

Fig 2 depicts the new structure of the MGE-HDDCNN method that uses Deep CNN to operate a fuzzy system based on the TSK model. In contrast to conventional neural networks, the name “DCNN” denotes three things: 1) there are more than four layers; 2) bias is not taken into account; and 3) the weights could be a function rather than a true constant.

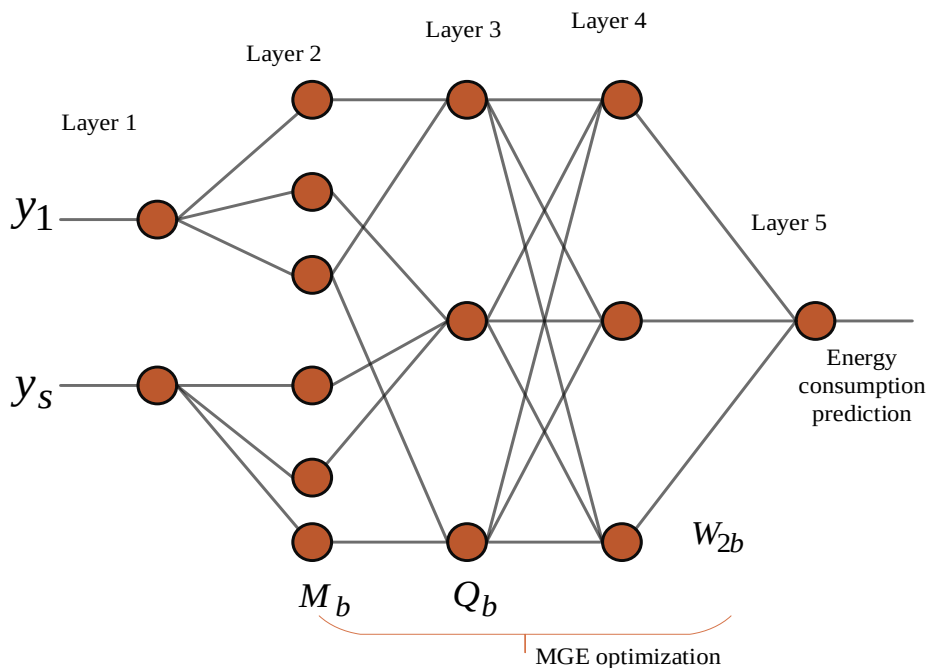


Fig 2. Structure of the MGE-HDDCNN Method.

Each unit in layer 1 indicates a linguistic variable as input, which is articulated using equation (1).

$$y = [y_1, y_2, \dots, y_{s-1}, y_s]^T \quad (1)$$

The second layer, known as the fuzzy layer, distorts the linguistic features of the first layer. The second layer's elements, which are DCNN neurons and the Gauss membership function, are often used, and the number of nodes in this layer corresponds to the number of fuzzy rules. The fuzzy rule's T norm is the neuron's output, which is defined in equation (2):

$$\mu_b = e \left| -\frac{\sum_{a=1}^s |y - d_{ab}|}{\sigma_{ab}^2} \right| \quad (a = 1, 2, \dots, s; b = 1, 2, \dots, v) \quad (2)$$

where d_{ab} and σ_{ab} indicate the center and breadth of the b^{th} DCNN neuron of y as the Gauss membership function. The sum of all the input linguistic variables is s , and the sum of all the membership fuzzy rules is v . The normalization layer, which is the third layer has the same number of nodes as the fuzzy layer. Each unit in the third layer denoted a suitable IF rule for the fuzzy system. The neurons' output expression for the b^{th} rule for Q_b output is expressed as in equation (3), and (4).

$$\mu_b = e \left| -\frac{\sum_{a=1}^s (y - d_{ab})^2}{\sigma_b^2} \right| \quad (3)$$

$$\mu_b = e \left| -\frac{\sum_{a=1}^s (y - d_{ab})^2}{\sigma_b^2} \right| \quad b = 1, 2, \dots, v \quad (4)$$

where the center of the b^{th} DCNN unit is denoted as $Y = (y_1, \dots, y_s)$ and d_b . Each node in this layer also serves as a DCNN unit, as can be seen from equation (2). The normalization layer is represented as third layer which has the same number of nodes with fuzzy layer. Here, normalized unit is considered as M . Therefore, the M is equal to the rule Q units. These are the neurons' output M_b expressions in equation (5), and (6).

$$\beta_b = \frac{\mu_b}{\sum_{l=1}^v \mu_l} \quad (5)$$

$$= \frac{e \left[-\|Y - d_b\|^2 \right]}{\sum_{l=1}^v e \left[-\frac{\|Y - d_b\|^2}{\sigma_l^2} \right]} \quad b = 1, 2, \dots, v \quad (6)$$

Each node in the fifth layer of this structure represents an output variable as the sum of incoming signals. The deblurred output layer is the fifth layer. The weight of all input signals from the preceding layer is expressed in equation (7), and (8).

$$= \frac{e \left[-\|Y - d_b\|^2 \right]}{\sum_{l=1}^v e \left[-\frac{\|Y - d_b\|^2}{\sigma_l^2} \right]} \quad b = 1, 2, \dots, v \quad (7)$$

$$x(Y) = \sum_{l=1}^v W_{2l} \cdot \beta_l = \frac{\sum_{l=1}^v W_{2l} \cdot e \left[-\|Y - d_b\|^2 \right]}{\sum_{l=1}^v e \left[-\frac{\|Y - d_b\|^2}{\sigma_l^2} \right]} \quad (8)$$

Where the output value is denoted as x and the rule weights are denoted as W_{2l} , and it is expressed in equation (9).

$$W_{2b} = r_{b0} + r_{b1}y_1 + \dots + r_{bs}y_s \quad b = 1, 2, \dots, v \quad (9)$$

When the weights are regarded as actual constants, it is represented in equation (10).

$$W_{2b} = d_b \quad b = 1, 2, \dots, v \quad (10)$$

GE Optimization

The combined Golden Eagle Optimization algorithm is the suggested GE algorithm. Here, the parameter selection is done by applying the hunting and encircling fitness function as the HDDCFN algorithm. Initialize the HDDCFN algorithm's settings, as well as its prime number and other parameters. The parameters were chosen at random and are expressed as in equation (11).

$$K_j = K_0, K_2, \dots, K_N \quad (11)$$

where $K_j = 1, 2, \dots, N$ and the overall parameter quantity is denoted as N . The suggested HDDCFN algorithm's primary focus is the best fitness selection. The fitness defines the function, which finds the optimal parameter to of each HDDCFN process with the best solution. The objective function is assessed using equation (12) to evaluate the fitness value of each parameter.

$$\text{Fitness value of key}(F_K) = f(\text{optimal key}) = f(K_0, K_2, \dots, K_N) \quad (12)$$

Here, F_K is represented as optimal key factors. Moreover, the selected fitness function was contributed to the finest parameter function with minimal fitness function coefficients. In an hyperparameter tuning the parameters are selected based on the iteration and which are mapped in terms of HDDCFN's flock data so far. Every time the data was trained and recall the finest results under the iteration of hyperparameter tuning.

In each iteration, a target parameter is selected by each search agent from the store of the entire flock. The tuning of each offensive and transit vector is then established with respect to the selected parameter.

Finally, the memory function was updated based on the parameter section approaches and performs the validation phase. If the current position is updated based on the memory function, which is finest than the previous location of the updated memory function. The parameter selection process is essential in a medical transaction. The parameter can be selected simply by letting each transaction choose it from its memory. To help weights move across the network more efficiently, a random one-to-one mapping technique is proposed, to randomly updates the current iteration while choosing the group components. Here, the parameter is randomly selected so the selected parameter is private or public mode to connect the memory on the single function. The selected parameter is then subjected to the strike and transit techniques for forecasting the performance.

Exploitation Phase

Here, the current location is starting point of vector coefficients which has terminate and transact the parameter location from that parameter exploration phase with respect to the memory models. The equation (13) is used to determine the parameter exploitation vector for weights tuning a .

$$\bar{K}_a \equiv \bar{Y}_{k_n} - \bar{Y}_a \quad (13)$$

Where, $\bar{K}_a$ is denoted as parameter exploitation of weighting vectors of each $\bar{Y}_{k_n}$ is the best parameter stayed so far by bias k_n , and $\bar{Y}_a$ is the present iteration of weights a . Parameter of the vector coefficients has highlighted in exploitation phase to leads the RDDCNN population based on the best-frequented sites.

Exploitation Phase

The parameter exploration is computed based on the exploitation phase. In some case, the parameter exploration is situated in n -directions the tangent hyperplane is located within the round. Therefore, the tangent hyperplane is estimated at first level of the exploration using equation (14),

$$h_1 k_1 + h_2 k_2 + \dots + h_m k_m = r \Rightarrow \sum_{i=1}^m h_i k_i = r \quad (14)$$

From the above estimation, ordinary vectors and decision parameter vectors are represented as $h_1, h_2, \dots, h_f$ and $k_1, k_2, \dots, k_f$ respectively with the i^{th} nodes.

Updating Phase

The hyperparameter tuning phase consist parameter exploration and exploitation to represents the data transaction. Therefore, iterative based data transaction of g^{th} vector step is mentioned in equation. (15)

$$\Delta K_i = r_1 \left(c_{K_u}^0 + \frac{g}{G} |c_{K_u}^G - c_{K_u}^0| \right) \frac{\bar{K}_u}{\|\bar{K}_u\|} + r_2 \left(c_{K_r}^0 - \frac{g}{G} |c_{K_r}^G - c_{K_r}^0| \right) \frac{\bar{K}_r}{\|\bar{K}_r\|} \quad (15)$$

Where, randomly selected range is $[0, 1]$ which is represented as initial and final public parameter values based on exploitation inclination r_1 and r_2 , $c_{K_u}^0$ and $c_{K_u}^G$ with g^{th} iteration. Then, initial and final private parameter values based on exploitation inclination $c_{K_r}^0$ and $c_{K_r}^G$ with g^{th} iteration. Euclidean distance of exploitation and exploration vectors are denoted as $\|\bar{K}_u\|$ and $\|\bar{K}_r\|$ with respect to the g^{th} iteration which are present iteration.

Termination Phase

Here, the each transaction has optimal parameter position which is $g+1$ iteration is analysed by additionally mentioned the g^{th} iterative step vectors. Moreover, termination phase is implemented in in equation. (16),

$$k^{g+1} = f^g + \Delta K_i \quad (16)$$

From the termination phase new positions are generated and that are stored in the memory. Moreover, the previous selected positions are avoided of the newly analysed is better for further processing. Consequently, feature values are moved to the other location to pick the random feature point based on the population of strike vectors. For this loop to continue, any one of the termination requirements must be met. **Fig 3** shows a workflow of the proposed methodology.

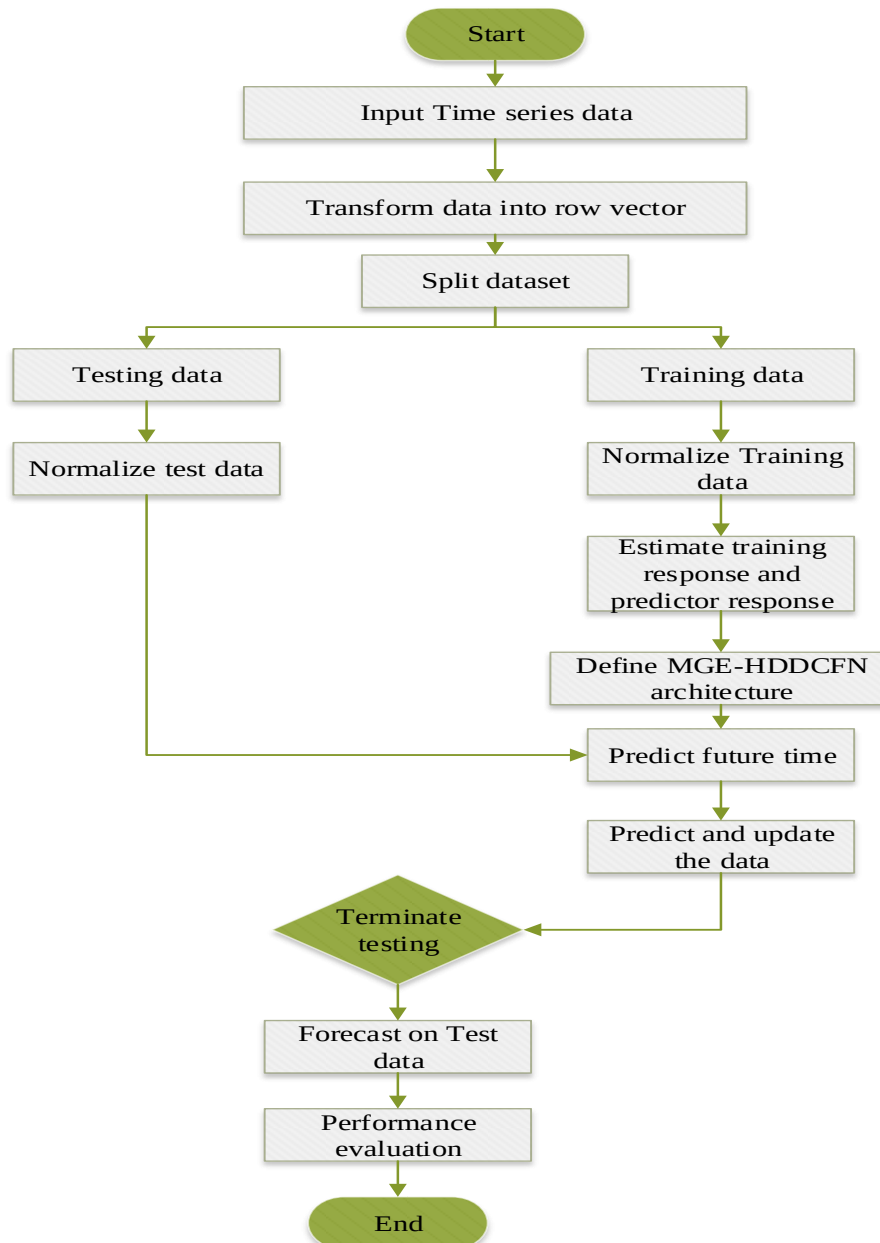


Fig 3. Flowchart of the Proposed Methodology.

IV. RESULT AND DISCUSSION

The experimental design of the suggested research project is described in this section. MATLAB is employed as a general-purpose language in this research study to carry out a number of tests.

Dataset Description

Since smart metering applications are run in actual buildings, getting time-series data is difficult and subject to constraints. The proposed hybrid system is assessed using information gathered by the smart grid infrastructure on the overall hourly energy usage between January 1, 2015, and January 1, 2019, which totals 35.1k points.

Moreover, the dataset contains four years of electricity price, generation and usages, weather related data these are collected from Spain country. then, to estimate the consumption and generation statistics ENTSOE based Transmission Service Operator (TSO) public gateway was utilized. Consequently, Red Electric Espaa providing defrayal assessing on Spanish TSO. Climate data for the five major cities in Spain was obtained from the Open Weather API as part of a personal initiative and made available here. For the implementation, the collected datasets was split based on training and testing. Finally, the performance are validated on entire data for testing the everyday of the weekends.

Performance Metrics

Five error metrics are used to evaluate the models: the correlation coefficient (R), the determination coefficient, the mean absolute error (MAE), the mean square error (MSE), and the root mean square error (RMSE). The actual (Y_t) and prediction series ($\hat{Y}_t$) have linear relationships, which are measured by the correlation coefficient (R) in equation (17); values of R that are closer to 1 suggest more accurate forecasts.

$$R = \frac{\sum(Y_t - \bar{Y})(\hat{Y}_t - \bar{\hat{Y}})}{\sqrt{\sum(Y_t - \bar{Y})^2 \sum(\hat{Y}_t - \bar{\hat{Y}})^2}} \quad (17)$$

Where $\bar{Y}$ and $\bar{\hat{Y}}$ defines the mean of actual and predicted values. Then variation between the actual and predicted classes are fitting the regression line to the collected data which is termed as coefficient of determination (R^2).

The variation that can be explained by fitting a regression line to the data is known as the coefficient of determination (R^2). This is a fundamental result of regression analysis. and this is better fit for calculating the statistical analysis of the model. Then, quantified determination coefficient, which is denoted as numerical value. The squared value of the correlation coefficient is (R^2) and it is expressed in equation (18). Complex statistics suggest a lesser fault change, which is from 0 to 1.

$$R^2 = \frac{\sum_{b=1}^n (Y_t - \bar{Y})(\hat{Y}_t - \bar{\hat{Y}})}{\sqrt{\sum_{b=1}^n (Y_t - \bar{Y})^2 \sum_{b=1}^n (\hat{Y}_t - \bar{\hat{Y}})^2}} \quad (18)$$

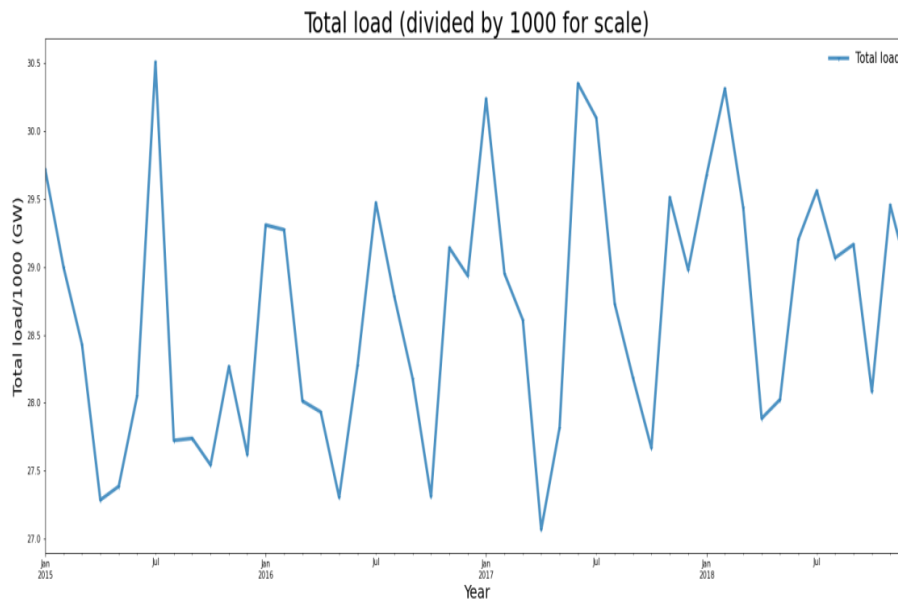


Fig 4. Total Hourly Load Curve.

where n denotes the number of test datasets. This parameter provides the residual error based on MSE, and affects the comparative valuation based on the imitation outcomes under the dimensionless measured metrics by evaluating the grade of correspondence among the real as well as expected values. The root mean square error calculation (RMSE) is expressed in equation (19).

$$RMSE = \sqrt{\frac{\sum_{b=1}^n (Y_t - \bar{Y})^2}{n}} \quad (19)$$

The Mean Squared Error (MSE) is a statistic for evaluating the total squared deviations between the actual and estimated values, or how close a fitted line is to sample points. The MSE is formulated in equation (20).

$$MSE = \frac{1}{n} \sum_{b=1}^n (Y_t - \bar{Y})^2 \quad (20)$$

The mean absolute error (MAE) in equation (21) is determined as the error rate ratio between the mean absolute observed and estimated energy consumption.

$$MAE = \frac{\sum_{b=1}^n |y_t - \hat{y}|}{n} \quad (21)$$

Performance Analysis

The total hourly load curve is illustrated in **Fig 4**. By identifying the hourly demand, each sector's matching hourly load curve is rebuilt of a chosen typical end consumer who represents consumption. The observation shows the overall load curve of the energy efficiency SG system.

The hourly load and apparent temperature are illustrated in **Fig 5**. The hourly load graphs of the changing climatic temperature visually represented the region's consumption patterns. These curves show that households and services account for most of the country's electricity use. The rise in consumption at the beginning of the day begins at approximately 7°C after the commencement of work hours in the various consumption categories. The different activities reach their daily average of around 22.5°C, where they stabilize and stay essentially constant until the end of the day. During the base year, the highest measured temperature was 27.5°C in the summer; the lowest was 7°C in the winter.

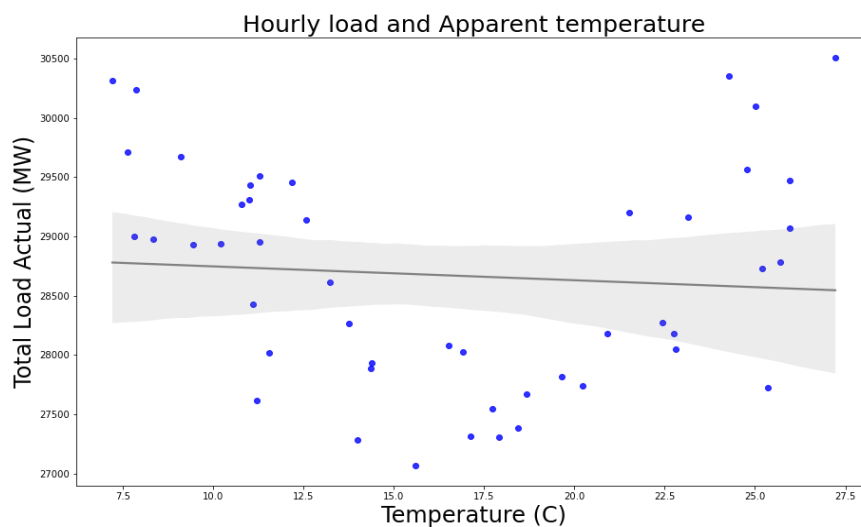


Fig 5. Hourly Load and Apparent Temperature.

Heat, oil, and gas cover more than 90% of heating space and water costs. The electricity use for heating is restricted to a few service segments and high-income homes. Its use grows correspondingly during extremely cold nights and, as a result, it helps to build the year's highest peak in December.

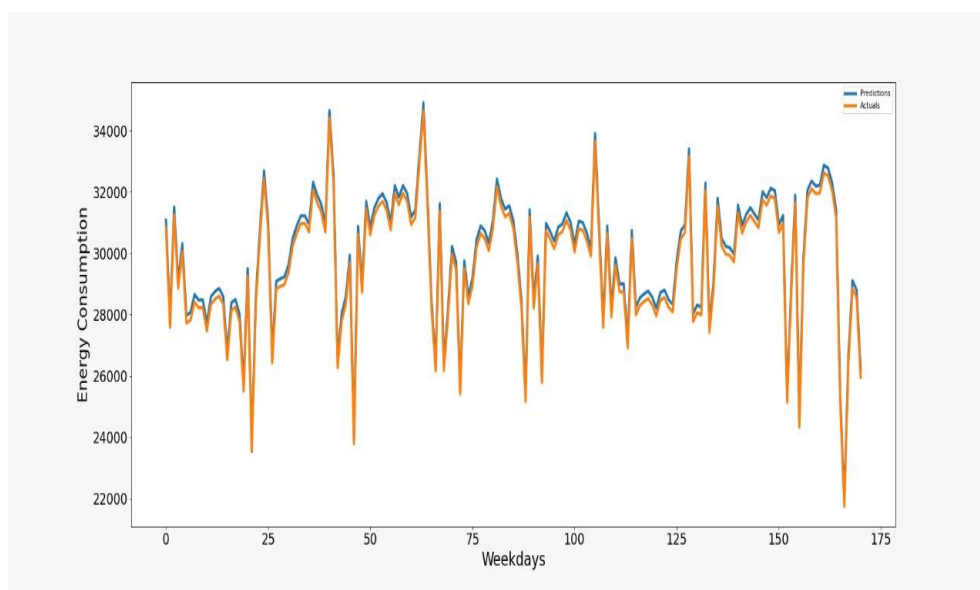


Fig 6. Load Forecast for Weekdays.

Fig 6 and 7 show hourly power demand trends for weekdays and weekends during all seasons of 2015-2019, respectively. **Fig 6** shows that the predicted and the actual observed values are mostly consistent. The blue line in the illustration also suggests the anticipated values, while the yellow line depicts the test/observed data. The projected values are in perfect agreement with the data on observed power demand figures. One of the most popular methods for project load in the SG is weekly projection, which is mostly carried out by the utility for week-ahead programming of the generating units. Given the size of the data, it is clear from the statistics that the energy is in units of 200 MWh. Another finding is that there is a very small error between real and forecast energy consumption.

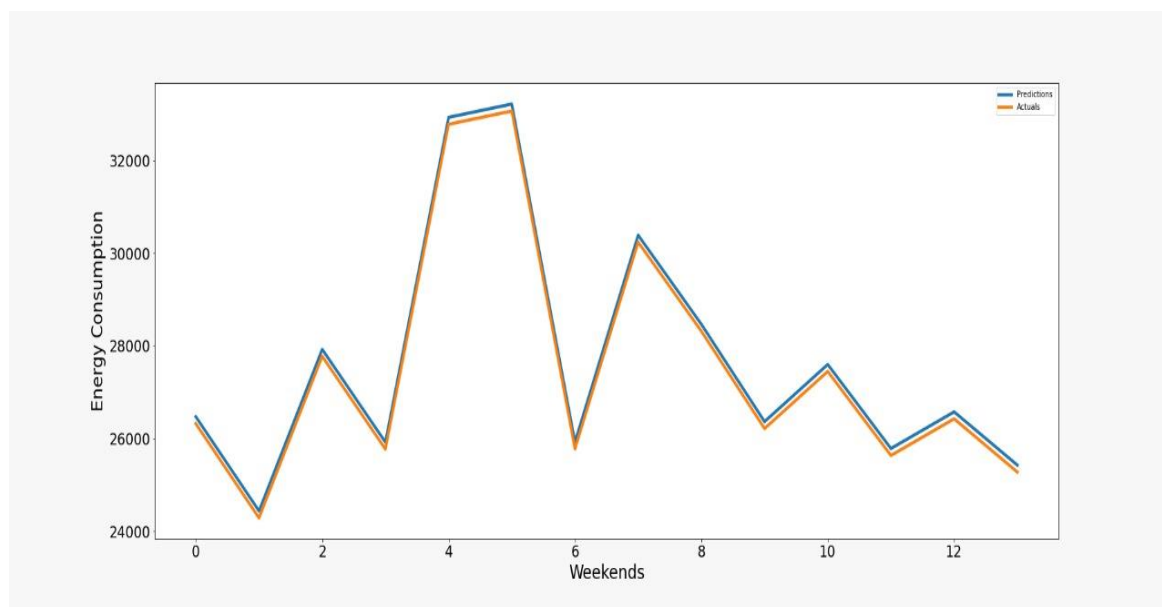


Fig 7. Load Forecast for Weekend Days.

The observation demonstrates a sizeable difference in power usage between weekdays and weekends. The demand for and consumption of power are significantly lower on the weekends due to fewer activities than on weekdays throughout the year. This forecast will result in significant utility cost savings and improved use of the production facilities. These simulation results, along with the RMSE, MSE, R, R^2 , and MAE errors in **Table 1**, explain the performance of the suggested model. Once created, the forecasting model was trained and evaluated on unbiased data. When comparing the observed data and projected values created during the training and testing procedure, the MAE and RMSE values were utilized to assess the success of the scenarios.

Table 1. Performance Measures of a Proposed Approach Under Weekdays and Weekend Approach

Metrics	Weekday	Weekend
MAE	0.15	0.24
MSE	0.2	0.27
RMSE	0.43	0.76
Correlation (R)	0.842	0.861
Coefficient of Determination (R^2)	0.932	0.926

R^2 is a correlation coefficient that measures how well the prediction models performed. The correlation between the actual and expected values is greater the higher the R^2 value. For example, on weekdays and weekends, the R^2 value for power demand is over 0.9, showing a significant connection between actual and expected values.

Comparative Analysis

The attained outcomes of the developed algorithm was validated with respect to the conventional method, like EvoHyS [31], LSTM-GRU [33], and LSTM [35] in terms of different metrics in **Table 2**.

The suggested model has the most incredible R^2 value of 0.932 for weekdays and 0.926 for weekends when compared to EvoHyS [31], LSTM-GRU [33], and LSTM [35]. In this instance, there is a strong connection between actual and anticipated numbers. The ideal outcomes for each model are shown in **Table 2**. The results show that the suggested MGE-based HDDCFN model produces the best outcomes. The RMSE index of the suggested MGE-based HDDCFN model has been improved compared to EvoHyS [31] in the case of power demand, LSTM-GRU [33], and LSTM [35]. Similar

improvements have been made to the MAE over other approaches. When comparing the performance of MSE to that of EvoHyS [31], LSTM-GRU [33], and LSTM [35] for electricity prediction, each approach performs better.

Table 2. Comparison of Performance Measures Using Various Approaches

Parameter	Algorithm	MAE	MSE	RMSE	R	R ²
Weekday	EvoHyS [31]	4.025	1.205	12.05	0.7652	0.91
	LSTM-GRU [33]	5.444	0.310	2.669	-	-
	LSTM [35]	50.96	-	67.35	-	0.62
	Proposed	0.15	0.2	0.43	0.842	0.932
Weekend	EvoHyS [31]	3.052	1.053	23.056	0.682	0.905
	LSTM-GRU [33]	3.769	0.340	2.695	-	-
	LSTM [35]	31.5	-	31.5	-	0.59
	Proposed	0.24	0.27	0.76	0.861	0.926

V. CONCLUSION

The accurate study of power usage has been regarded as necessary because of the rise in electricity use in residential and commercial structures. This research presented an intelligent multi-objective hybrid deep learning framework for an energy-saving decision support system. Specifically, a cutting-edge, multi-objective metaheuristic forecast system inspired by nature was developed behind the analytics layer to allow precise prediction and energy consumption optimization. A future study may focus on creating a single prediction model that operates well under various seasonal and climatic situations. In this study, the simulation tool was MATLAB, and the comparative metrics were RMSE, MAE, MSE, R, and R² for weekdays and weekends. The usefulness of the suggested strategy with forecasting in various test cases and granularity is demonstrated by simulation results. Furthermore, the numerical findings indicate that the proposed forecasting model outperforms existing techniques in terms of MAE, MSE, and RMSE. Applying this approach to additional big data sets to create city-level energy consumption estimates and extending this study to investigate more granular geographical and temporal dimensions to obtain improved forecast accuracy are two potential future work directions.

CRedit Author Statement

The authors confirm contribution to the paper as follows:

Conceptualization: Praveen Kumar R, Ravi Kumar, Gokulakrishnan S, Devi Kosuru S N V J, Neeraj Shrivastava and Jyothi P; **Methodology:** Praveen Kumar R, Ravi Kumar and Gokulakrishnan S; **Software:** Devi Kosuru S N V J, Neeraj Shrivastava and Jyothi P; **Data Curation:** Neeraj Shrivastava and Jyothi P; **Writing- Original Draft Preparation:** Praveen Kumar R, Ravi Kumar, Gokulakrishnan S, Devi Kosuru S N V J, Neeraj Shrivastava and Jyothi P; **Visualization:** Devi Kosuru S N V J, Neeraj Shrivastava and Jyothi P; **Supervision:** Devi Kosuru S N V J, Neeraj Shrivastava and Jyothi P; **Validation:** Praveen Kumar R, Ravi Kumar and Gokulakrishnan S; **Writing- Reviewing and Editing:** Praveen Kumar R, Ravi Kumar, Gokulakrishnan S, Devi Kosuru S N V J, Neeraj Shrivastava and Jyothi P; All authors reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

- [1]. V. Vanitha and E. Vallimurugan, "A hybrid approach for optimal energy management system of internet of things enabled residential buildings in smart grid," *International Journal of Energy Research*, vol. 46, no. 9, pp. 12530–12548, May 2022, doi: 10.1002/er.8024.
- [2]. T. M. Skjølsvold, I. M. Henriksen, and M. Ryghaug, "Beyond the car: how electric vehicles may enable new forms of material politics at the intersection of the smart grid and smart city," *Urban Geography*, pp. 1–18, Apr. 2022, doi: 10.1080/02723638.2022.2044692.
- [3]. A. O. Ali, M. R. Elmarghany, M. M. Abdelsalam, M. N. Sabry, and A. M. Hamed, "Closed-loop home energy management system with renewable energy sources in a smart grid: A comprehensive review," *Journal of Energy Storage*, vol. 50, p. 104609, Jun. 2022, doi: 10.1016/j.est.2022.104609.
- [4]. D. K. Behera, M. Das, S. Swetanisha, and J. Nayak, "XGBoost regression model-based electricity tariff plan recommendation in smart grid environment," *International Journal of Innovative Computing and Applications*, vol. 13, no. 2, p. 79, 2022, doi: 10.1504/ijica.2022.123223.

- [5]. V. F. Pires, I. Colak, and F. Kurokawa, “Smart Grid as a Key Tool for the Future of Electrical Distribution Networks,” *Energies*, vol. 15, no. 9, p. 3281, Apr. 2022, doi: 10.3390/en15093281.
- [6]. D. Kumari and A. Sharma, “A Study of Smart Grid Application and Impact of Renewable Resources When Integrated with Smart Grid,” *Electronic Systems and Intelligent Computing*, pp. 245–258, 2022, doi: 10.1007/978-981-16-9488-2_22.
- [7]. J. F. Martins, A. G. Pronto, V. Delgado-Gomes, and M. Sanduleac, “Smart Meters and Advanced Metering Infrastructure,” *Pathways to a Smarter Power System*, pp. 89–114, 2019, doi: 10.1016/b978-0-08-102592-5.00004-1.
- [8]. N. K. Singh and V. Mahajan, “End-User Privacy Protection Scheme from cyber intrusion in smart grid advanced metering infrastructure,” *International Journal of Critical Infrastructure Protection*, vol. 34, p. 100410, Sep. 2021, doi: 10.1016/j.ijcip.2021.100410.
- [9]. J. Chan, R. Ip, K. W. Cheng, and K. S. P. Chan, “Advanced Metering Infrastructure Deployment and Challenges,” *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*, pp. 435–439, Mar. 2019, doi: 10.1109/gtdasia.2019.8715927.
- [10]. J. S. Guerrero-Prado, W. Alfonso-Morales, and E. F. Caicedo-Bravo, “A Data Analytics/Big Data Framework for Advanced Metering Infrastructure Data,” *Sensors*, vol. 21, no. 16, p. 5650, Aug. 2021, doi: 10.3390/s21165650.
- [11]. Md. M. Rahman, M. B. Pasupuleti, and H. P. Adusumalli, “Advanced Metering Infrastructure Data: Overviews for the Big Data Framework,” *ABC Research Alert*, vol. 7, no. 3, pp. 159–168, Dec. 2019, doi: 10.18034/ra.v7i3.602.
- [12]. M. H. Rashid, “AMI Smart Meter Big Data Analytics for Time Series of Electricity Consumption,” *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, pp. 1771–1776, Aug. 2018, doi: 10.1109/trustcom/bigdata.2018.00267.
- [13]. 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, Feb. 2019, doi: 10.1016/j.compind.2018.12.010.
- [14]. G. F. Huseien and K. W. Shah, “A review on 5G technology for smart energy management and smart buildings in Singapore,” *Energy and AI*, vol. 7, p. 100116, Jan. 2022, doi: 10.1016/j.egyai.2021.100116.
- [15]. P. D. Diamantoulakis, V. M. Kapinas, and G. K. Karagiannidis, “Big Data Analytics for Dynamic Energy Management in Smart Grids,” *Big Data Research*, vol. 2, no. 3, pp. 94–101, Sep. 2015, doi: 10.1016/j.bdr.2015.03.003.
- [16]. M. Ghorbanian, S. H. Dolatabadi, and P. Siano, “Big Data Issues in Smart Grids: A Survey,” *IEEE Systems Journal*, vol. 13, no. 4, pp. 4158–4168, Dec. 2019, doi: 10.1109/jsyst.2019.2931879.
- [17]. D. Alahakoon and X. Yu, “Advanced analytics for harnessing the power of smart meter big data,” *2013 IEEE International Workshop on Intelligent Energy Systems (IWIES)*, pp. 40–45, Nov. 2013, doi: 10.1109/iwies.2013.6698559.
- [18]. T. Zufferey, A. Ulbig, S. Koch, and G. Hug, “Forecasting of Smart Meter Time Series Based on Neural Networks,” *Data Analytics for Renewable Energy Integration*, pp. 10–21, 2017, doi: 10.1007/978-3-319-50947-1_2.
- [19]. T. Kotsiopoulos, P. Sarigiannidis, D. Ioannidis, and D. Tzovaras, “Machine Learning and Deep Learning in smart manufacturing: The Smart Grid paradigm,” *Computer Science Review*, vol. 40, p. 100341, May 2021, doi: 10.1016/j.cosrev.2020.100341.
- [20]. B. Hayes, J. Gruber, and M. Prodanovic, “Short-Term Load Forecasting at the local level using smart meter data,” *2015 IEEE Eindhoven PowerTech*, pp. 1–6, Jun. 2015, doi: 10.1109/ptc.2015.7232358.
- [21]. R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor, “Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy,” *Applied Energy*, vol. 123, pp. 168–178, Jun. 2014, doi: 10.1016/j.apenergy.2014.02.057.
- [22]. G. A. Alkawsu et al., “A hybrid SEM-neural network method for identifying acceptance factors of the smart meters in Malaysia: Challenges perspective,” *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 227–240, Feb. 2021, doi: 10.1016/j.aej.2020.07.002.
- [23]. D. Alberg and M. Last, “Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms,” *Vietnam Journal of Computer Science*, vol. 5, no. 3–4, pp. 241–249, Jun. 2018, doi: 10.1007/s40595-018-0119-7.
- [24]. J. Barzola-Monteses, M. Guerrero, F. Parrales-Bravo, and M. Espinoza-Andaluz, “Forecasting Energy Consumption in Residential Department Using Convolutional Neural Networks,” *Information and Communication Technologies*, pp. 18–30, 2021, doi: 10.1007/978-3-030-89941-7_2.
- [25]. A. Jahanshahi et al., “An Auto Regressive Integrated Moving Average (ARIMA) Model for prediction of energy consumption by household sector in Euro area,” *AIMS Energy*, vol. 7, no. 2, pp. 151–164, 2019, doi: 10.3934/energy.2019.2.151.
- [26]. J.-S. Chou and N.-T. Ngo, “Smart grid data analytics framework for increasing energy savings in residential buildings,” *Automation in Construction*, vol. 72, pp. 247–257, Dec. 2016, doi: 10.1016/j.autcon.2016.01.002.
- [27]. A. Mohammadi-Balani, M. Dehghan Nayeri, A. Azar, and M. Taghizadeh-Yazdi, “Golden eagle optimizer: A nature-inspired metaheuristic algorithm,” *Computers & Industrial Engineering*, vol. 152, p. 107050, Feb. 2021, doi: 10.1016/j.cie.2020.107050.
- [28]. J. Wang, Z. Ma, and X. Chen, “Generalized Dynamic Fuzzy NN Model Based on Multiple Fading Factors SCKF and its Application in Integrated Navigation,” *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3680–3693, Feb. 2021, doi: 10.1109/jsen.2020.3022934.
- [29]. M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu, “Towards Better Analysis of Deep Convolutional Neural Networks,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 91–100, Jan. 2017, doi: 10.1109/tvcg.2016.2598831.
- [30]. S. Singh and A. Yassine, “Big Data Mining of Energy Time Series for Behavioral Analytics and Energy Consumption Forecasting,” *Energies*, vol. 11, no. 2, p. 452, Feb. 2018, doi: 10.3390/en11020452.
- [31]. D. M. F. Izidio, P. S. G. de Mattos Neto, L. Barbosa, J. F. L. de Oliveira, M. H. da N. Marinho, and G. F. Rissi, “Evolutionary Hybrid System for Energy Consumption Forecasting for Smart Meters,” *Energies*, vol. 14, no. 7, p. 1794, Mar. 2021, doi: 10.3390/en14071794.
- [32]. A. Kumar, S. A. Alghamdi, A. Mehbodniya, M. anul haq, J. L. Webber, and S. Navruzbeq Shavkatovich, “Smart power consumption management and alert system using IoT on big data,” *Sustainable Energy Technologies and Assessments*, vol. 53, p. 102555, Oct. 2022, doi: 10.1016/j.seta.2022.102555.
- [33]. A.-N. Khan, N. Iqbal, A. Rizwan, R. Ahmad, and D.-H. Kim, “An Ensemble Energy Consumption Forecasting Model Based on Spatial-Temporal Clustering Analysis in Residential Buildings,” *Energies*, vol. 14, no. 11, p. 3020, May 2021, doi: 10.3390/en14113020.
- [34]. Z. Niu, J. Wu, X. Liu, L. Huang, and P. S. Nielsen, “Understanding energy demand behaviors through spatio-temporal smart meter data analysis,” *Energy*, vol. 226, p. 120493, Jul. 2021, doi: 10.1016/j.energy.2021.120493.

- [35]. I. Fatema, X. Kong, and G. Fang, “Electricity demand and price forecasting model for sustainable smart grid using comprehensive long short term memory,” *International Journal of Sustainable Engineering*, vol. 14, no. 6, pp. 1714–1732, Jul. 2021, doi: 10.1080/19397038.2021.1951882.
- [36]. F. Ünal, A. Almalaq, and S. Ekici, “A Novel Load Forecasting Approach Based on Smart Meter Data Using Advance Preprocessing and Hybrid Deep Learning,” *Applied Sciences*, vol. 11, no. 6, p. 2742, Mar. 2021, doi: 10.3390/app11062742.
- [37]. I. U. Haq et al., “Sequential Learning-Based Energy Consumption Prediction Model for Residential and Commercial Sectors,” *Mathematics*, vol. 9, no. 6, p. 605, Mar. 2021, doi: 10.3390/math9060605.