An Analysis Of Energy Demand In IoT Integrated Smart Grid Based On Time And Sector Using Machine Learning

Jitendra MANAGRE[1], *Namit GUPTA*[1]

[1]Electrical and Electronics Engineering Department, Shri Vaishnav Institute of Technology and Science, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Ujjain Road, village- Baroli, Indore- 453111, Madhya Pradesh, India

jitendra.managre@gmail.com, namitg@hotmail.com

DOI: 10.15598/aeee.v21i4.5291

Article history: Received Jun 23, 2023; Revised Oct 09, 2023; Accepted Oct 19, 2023; Published Dec 31, 2023.

Abstract. *Smart Grids (SG) encompass the utilization of large-scale data, advanced communication infrastructure, and enhanced efficiency in the management of electricity demand, distribution, and productivity through the application of machine learning techniques. The utilization of machine learning facilitates the creation and implementation of proactive and automated decision-making methods for smart grids. In this paper, we provide an experimental study to understand the power demands of consumers (domestic and commercial) in SGs. The power demand source is considered a smart plug reading dataset. This dataset is large dataset and consists of more than 850 user plug readings. From the dataset, we have extracted two different user data. Additionally, their hourly, daily, weekly, and monthly power demand is analysed individually. Next, these power demand patterns are utilized as a time series problem and the data is transformed into 5 neighbour problems to predict the next hour, day, week, and month power demand. To learn from the transformed data, Artificial Neural Network (ANN) and Linear Regression (LR) ML algorithms are used. According to the conducted experiments, we found that ANN provides more accurate prediction than LR Additionally, we observe that the prediction of hourly demand is more accurate than the prediction of daily, weekly, and monthly demand. Additionally, the prediction of each kind of pattern needs an individually refined model for performing with better accuracy*

Keywords

Artificial Neural Network (ANN), Accuracy improvement, Demand Side Management (DSM), Energy Management, Machine Learning (ML), Smart Grids (SGs).

1. Introduction.

The smart grid monitors and performs the various duties with proper management, like electricity, parking signals on traffic, and alert signals on road, earthquake detection, and weather conditions. A smart grid is the interconnected network of components like sensors, transmission and distribution lines, energy storage devices like smart meters, substations, and transformers etc. [1]. It contributes to energy management. The SGs with IoT integration are more capable and accurate because of communication among, sensors, and command centres for rapid response. It is less expensive than the traditional grid [2]. It can detect faults and failures [3]. The smart grid is advantageous for balancing the difference between energy consumption and production by considering power generation, management, and distribution. Additionally [4], ensuring smart energy use, clean energy usage, and low cost, fulfil increasing energy needs and provide faster solutions to regional problems.

Demand side management supports SG functionality for infrastructure development, decentralized energy resource management, and electric cars [5]. Addition-

Tab. 1: Abbreviations

Abbreviations	Abbreviations	Abbreviations	Abbreviations
SGs	Smart Grids	ML	Machine Learning
ANN	Artificial Neural Network	RNN	Recurrent Neural Network
CNN	Convolutional Neural Network	DES	Distributed Energy System
DSM	Demand-Side Management	DR	Demand Response
DL	Deep Learning	LR	Linear Regression
SM	Smart Meter	AMI	Automatic Meter Infrastructure
IoT	Internet of Things	EMM	Energy Management Model
LECs	Lower Energy Consumers	HECs	High Energy Consumers
ToU	Time of Usage	MAE	Mean Absolute Error
HEM	Home Energy Management	NILM	Non-Intrusive load management
HMM	Hidden Markov Model	LSTM	Long Short-Term Memory
GHG	Green House Gas	STLF	Short Term Load forecasting
CLARA	Clustering LArge Applications	SVM	Support Vector Machine
RES	Renewable Energy Source	ELM	Energy Learning Machine
MOAHA	Multi Objective Artificial Hummingbird	DCNN	Deep Convolution Neural Network
SVM	Support Vector Machines	HAN	Home Area Network
GWO	Grey Wolf Optimizer	WDO	Wind-Driven Optimization
WNN	Wavelet Numerical Network	NWP	Numerical Weather Prediction
WPP	Wind Power Prediction	GRU	Gated Recurrent Unit
EED	Electric Energy Demand	EES	Electric Energy Supply

ally, the aim is to minimize energy costs, carbon emission, and improve sustainability [6]. It is also useful for smart billing by using Smart Meters (SMs) [7]. It might result in consumer awareness and help to improve supply chain of power [8]. The consumers can regulate the amount of energy used. The Demand Side Management (DSM) is in relation to the intelligent power grid. The DSM will notify the server for load schedule and load minimization technique [9]. To achieve effective DSM, the SG needs accurate and trustworthy information about power use.

Components of smart grid are described in the given Figure 1. The Smart Grid technology is comparatively new and requires continuous development and improvement. This technology is utilizing various software and hardware technologies to improvise their functioning [10]. By using the implemented technologies, a significant amount of data has been generated. The analysis of this data may provide several benefits to understand the behavior of power utilization, trend, and patterns of demand [11]. Additionally sudden changes and seasonal changes may help to manage the power demand and supply. The analysis of the data is also essential to make balance between the power generation and their better utilization by the different industrial and domestic purpose [12].

In this paper, we first involve a review of the modern trends and development in SG based on ML. Next, we addressed several identified research areas based on review. Further, we present a study of Electric Energy Demand (EED) understanding using smart plug reading dataset. The ML techniques are applied on smart metering dataset and future demand has tried to predict. The following task has been undertaken:

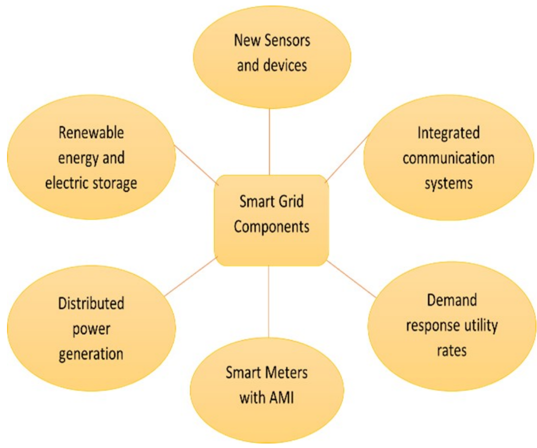


Fig. 1: Smart grid components and its features.

- Analyze SG data to recover the seasonal impact on demand and supply of EED in the domestic sector and industrial sectors.
- Identifying the variations in Electric Energy Supply (EES) for a 24-hour time cycle.
- Identify the weekly pattern of EED and EES.

Further, the concept of time series data analysis is utilized, and ML methods are applied to anticipate the power required for different time frequencies. Finally, the performance of the employed techniques is measured. Two ML algorithms are used for this paper namely artificial neural network and linear regression. They are as follows:

1) Artificial Neural Network (ANN):

An Artificial Neural Network is a data processing approach that is derived from the principles of biological nerve systems. The artificial neural network is organized in the form of a network, consisting of interconnected units referred to as neurons. The configuration is contingent upon the specific application, such as recognition, classification, or prediction. The acquisition of knowledge regarding Artificial Neural Networks has necessitated modifications to the synaptic connections. The Artificial Neural Network is a sophisticated, non-linear, and parallel computational system. The fundamental component of a neural network is the neuron, which is comprised of a certain number of inputs denoted as $x(n)$. Each input is multiplied by a corresponding connection weight, written as $w(n)$. The sum of the product of the input and weight is sent through a transfer function, also known as an activation function, to produce the output.

2) Linear Regression (LR):

Linear regression is a statistical technique used to model the relationship between a dependent variable and one or more independent variables. It is a widely used method in various fields, including economics. Linear regression is a widely recognized approach employed in the resolution of diverse statistical and machine learning challenges. The logistic regression model is commonly employed for predictive modeling purposes. The LR algorithm is primarily focused on lowering the error to achieve accurate predictions. The concept of linear regression is explored as a theoretical framework for comprehending the correlation between input and output variables. The linear regression model is predicated on the assumption of a linear association between the input variable (x) and the output variable (y). The value of y can be determined by performing a linear combination of the input variable x. When a single method is present, it is commonly known as simple LR. When there are numerous variables, it is referred to as multiple linear regression. The most often used strategy for training the linear regression model is known as Ordinary Least Squares. The linear equation is a mathematical model that utilizes a collection of input values (x) in order to make predictions about the corresponding output values (y). The equation assigns a coefficient, represented by Beta (B), to each input, which serves as a scale factor. The coefficient, often known as the intercept, imparts a degree of freedom to the line.

$$y = B_0 + B_1 x. \tag{1}$$

Overall, the paper presents original contributions in terms of applying machine learning algorithm to power demand prediction, analyzing demand patterns, comparing algorithm performance, and addressing the challenges associated with different prediction horizons and demand types within Smart Grids.

2. Related Work

In this section we will discuss the issues and their solutions which have been given in introduction section by recent researchers.

2.1. Review Highlights

The collected review relevant to the smart grid based on internet of things (IoT) and Machine learning (ML) technique is highlighted into Table 2.

2.2. Review Summary

Based on the recently performed study, we have identified the following research opportunities to improve the performance of existing smart grid systems.

- Here is a need to pay attention to the Home Energy Management (HEM) system. This system can deal with services to automate power supply according to demand. Current algorithms for the operations of HEMs require performance improvement and optimization.
- Power usage has increased, which raises concerns about an imbalance between supply and demand.
- Securing SG and realizing its applicability are difficult tasks. Additionally, Real-time monitoring requires high-speed connectivity, and accelerating the transition to smart grids is crucial.
- Due to the rise in the number of smart gadgets in homes, there is an increasing requirement for electrical energy. Thus, we must create methods for controlling energy demand.
- Operations demand ML-based load anticipation in IoT framework. Sending, handling, and processing a lot of data is difficult.
- To reduce CO2 emissions, greenhouse gas emissions, global pollution, combat climate change, and provide energy security, we must address the reliability of renewable energy sources.
- Need to create an Energy Management Model (EMM) that uses ML algorithms to manage renewable energy under varying conditions or needs.

- Data and energy have different values, but we must pay attention to both. Additionally, we must control the electricity supply in accordance with consumers like smart cities, homes, and businesses.

- The covert cyber deception attack is a recent threat. Such attacks are being carried out by hackers, who are undetectable. As a result, affected devices may provide a challenging situation.

- Concern with reducing and modifying energy consumption patterns to reduce peak loads, smooth out load distribution, and reduce carbon emissions.

3. Proposed Work

The ML technique is crucial for the SG. The SG applications depend on effectiveness of communication and the accuracy of the prediction algorithms. Some of these applications may include automation, self-diagnose, error finding and tolerance, and maintenance. To fulfil the applications power demand and supply we need an efficient Electric Energy Management (EEM) system. The data analysis is an essential part of EEM system [29]. Figure 2 shows the flow chart of proposed system which presents a study of Electric Energy Demand (EED) using smart plug reading dataset. The ML techniques are also applied on analysed data and future workload have tried to predict. The following task has been undertaken in this paper.

1) Analyze SG data to recover the seasonal impact on demand and supply of EED in the domestic sector and industrial sectors.

2) Identifying the variations in Electric Energy Supply (EES) for a 24-hour time cycle.

3) Identify the weekly pattern of EED and EES.

1) Dataset

The main objective of this analysis is to recover cycle feature for preparing a prediction system. To study the seasonal influence in EED and EEDS we utilize a Plug Readings of Home Area Network (HAN) dataset. That dataset is publicly available at [29]. The dataset is provided by the Australian Government Department of Climate Change, Energy, the Environment, and Water [30]. The dataset was created on 09-09-2015 and last updated on 11-04-2022. The dataset consists of CUSTOMER_ID (Cid), READING_TIME (date), PLUG_NAME (PName), READING_VALUE (Reading), CALENDAR_KEY (CKey) and RECORD_COUNT (Rcount) attributes. The raw samples of the dataset are demonstrated in Table 3.

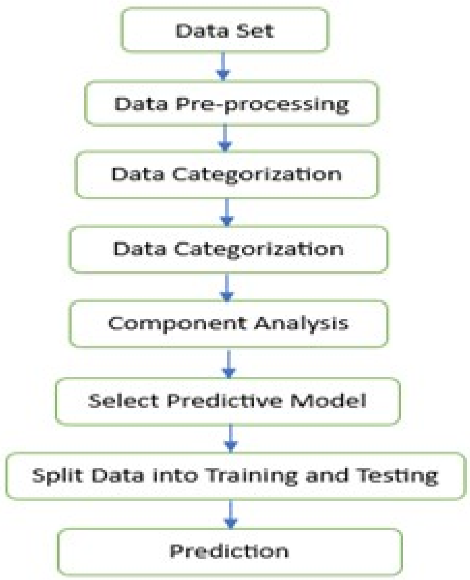


Fig. 2: Flow chart of proposed system.

2) Data Preprocessing

The aim of preprocessing is to prepare the data for utilization in experiments. Therefore, the following steps are followed for preprocessing of data:

1) Attributes, CKey and Cid are used for the same purpose for identifying the consumer uniquely. Therefore, we eliminated Ckey from the initial dataset.

2) Attribute PName describes the product or device used. This attribute is not very appropriate for our study therefore we eliminate this attribute.

3) Next the date attribute is converted into an index. Additionally, the data has been sorted according to the date.

The preprocessed dataset is demonstrated in Table 4.

3) Data Categorization

The dataset contains more than 800 consumer's data. Overall, the dataset has 10828120 available instances or readings. Figure 3 shows the total consumers in dataset and demonstrated on X axis. Y axis shows the total readings available for each client. For the high-quality resolution only 60000 available instances are showing in below Figure. Here, we categorize our dataset into two groups based on the entries available on dataset.

Tab. 2: Review Highlights

Ref.	Work contributed	Method	Results
[13]	Presented an IEMS for managing energy consumption and meeting demand in a microgrid. The IEMS makes sure the hydrogen-battery storage system runs as smoothly as possible.	To solve IEMS, the multi-objective artificial hummingbird optimizer, also known as MOAHA, is used.	The optimized IEMS results in low greenhouse gas (GHG) emissions.
[14]	Work carried out with the purpose of forecasting the amount of energy and load consumption. To accomplish this objective, a basic and uncomplicated long-term memory deep learning model for electrical load has been created and calibrated for multivariate time-series forecasting.	The LSTM model is implemented, and then the six swarm intelligence metaheuristics used to fine-tune the LSTM architecture are described and discussed.	Overall, the results show that FA gave the best results out of all the heuristics that were used, even though the differences between the results were very small.
[15]	A short-term power load forecasting model that is based on multi-factor analysis and a Long-Short Term Memory (LSTM) neural network has been proposed with the goal of increasing the accuracy of short-term power load forecasting and taking into full consideration the influence of weather factors on power load.	Using the sliding window method, the initial time series data are deconstructed and then recreated. LSTM is then used to set up the model for making predictions.	The forecasting model described in this study has a 7.41% average absolute percentage error and 380.67 MW average absolute value error, which is better than the other models listed in the paper.
[16]	This work proposes a self-supervised learning-based nonintrusive-load monitoring strategy to maximise model training using unlabelled monitoring data.	The load identification model uses self-supervised learning to recreate data.	The algorithm suggested in this research outperforms advanced machine learning and deep learning algorithms, proving its efficacy.
[17]	In this paper, proposed a NILM framework based on the Modified Factorial Hidden Markov Model (MFHMM) to model the interdependencies between appliances' operational states and between appliances' differential operational states, all while accounting for differences in their power consumption profiles over time.	The Hidden Markov Model is used to model each appliance as a separate load (HMM).	Each segment's NILM problem is solved, and then the acquired solution is tweaked in light of the aggregated load point's voltage profile.
[18]	Presented a stochastic transition likelihood function (STLF) approach using ensemble hidden Markov models (e-HMM) to learn the dynamic properties of industrial customers' spending patterns in linked multivariate time series and consequently enhance prediction accuracy.	Time series, machine learning, and deep learning models have been used to solve STLF problems.	To lower the overall prediction error rate, we use the Bagging ensemble learning algorithm architecture.
[19]	Presented a two-stage method for increasing STLF's precision by using the outputs of classic algorithms as reference points. By breaking the initial forecast deviation down into its component modes, the desired DR deviation sequence may be built.	Dynamic mode decomposition (DMD) is utilised to acquire the deviation sequences generated by DR. The Hankel matrix is developed to ease the operation.	The accuracy of the final forecast is enhanced by superimposing the findings of conventional algorithms with the discovered deviation sequence.
[20]	This paper proposes a deep convolutional neural network (DCNN)-based NILM framework for profiling the on/off states of all residential appliances and associated power consumption.	The NILM can directly assess appliance-level electricity consumption using a load trajectory and DCNN architecture.	NILM improves demand side control and energy efficiency.

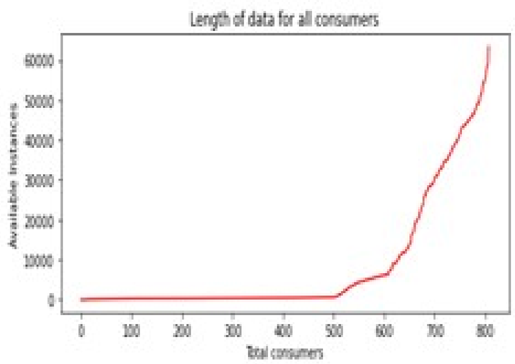


Fig. 3: Shows consumers and contributed data in the dataset.

Based on the number of entries the data is categorized in two categories: High Power consumer (Industrial), Industrial data involve the consumers who have entries higher than 10000 instances is considered as indutrial or high-power consumer group. And next is Low power usage consumer (Domestic), those consumers has less entry than 10000 instances. Based on obtained data based on their category the both group of users is given in Figure 4.

Figure 4(a) shows the total consumers in industrial usage and Figure 4(b) shows the consumers with user of low power usages. The available instances are suddenly increasing from 500-600 in Figure 4(b) because of sudden increasing of appliances used and increased power demand.

Review Highlights

Ref.	Work contributed	Method	Results
[21]	The study makes several important contributions to the field, including research on the economic and technological effects of renewable energy, grid integration, and electricity rates. However, the major objective is to provide a demand-response (DR) model that maximises the benefits to energy retailers, in this case the microgrid clients. Different customers' utility and elasticity during peak and off-peak times are analysed by DR models.	A revolutionary intelligent algorithm minimises microgrid system cost and analyses results with and without DR software.	Numerical data showed that a DR-based energy management microgrid system reduced overall generating costs and pollution compared to the literature.
[22]	Proposed a heterogeneous ensemble method for making short- and intermediate-term load predictions. The ensemble forecaster is formed by a two-level hierarchy of machine learning-based and classical approaches, with the first-stage forecasters' output used as input in the second stage.	ML forecasters include artificial neural networks and support vector regression, whereas classical forecasters include Holt's exponential smoothing and multiple linear regression.	Practical system prediction accuracy improves manifolds. The proposed model outperforms ensemble-based models.
[23]	The proposed computational strategy offers a significant amount of potential to improve the effectiveness of smart grids.	RFECV is used to eliminate the weakest feature and select the highest scoring feature.	Linear regression performs the best in all metrics.
[24]	Presented a combined forecasting model, GWO WNN-VMD-LSTM-Q-learning, based on the integration of NWP and wind power time series.	The Q-learning technique is utilised to superimpose the prediction results based on an ideal weight and obtain the final WPP findings.	The simulation results show that this model has good prediction accuracy and a far bigger predictive impact than traditional models that use time series forecasts.
[25]	A long short-term memory (LSTM) recurrent neural network-based framework was proposed, which is the most recent and one of the most prominent techniques of deep learning. This was done in order to address this challenging problem.	The proposed system is evaluated on publicly available household smart metre data and compared to benchmarks, including load forecasting state-of-the-arts.	The proposed LSTM approach does a better job of predicting short-term load for individual residential households than the other algorithms listed.
[26]	In order to predict energy use and determine peak demand, this research suggests using a random forest supervised learning model.	In order to improve analysis and predictions, the vast smart metre dataset collected during the year is fed into the random forest classifier method.	As compared to other methods, this one excels in precision, consistency, and broad applicability. This paper also explores the previous models and evaluates how they've fared in the past.
[27]	The primary objective of this study is to develop models for anticipating electrical load based on the electricity provider's actual load data.	LSTM, GRU, and RNN deep learning algorithms predict electrical demands (RNN).	According to the findings, the GRU model was successful in attaining an R-squared value of 90.228%, a Mean Square Error (MSE) value of 0.00215, and a Mean Absolute Error (MAE) value of 0.03266.

Next from each group an individual consumer is selected for detailed exploration of data. There are a smaller number of users in industrial users, but the amount of EE is higher enough. Here we denote them as C_1 and C_2. The consumer C_1 is a domestic user, which have a total of 9948 instances. The consumer C_2 is selected as commercial usage group. This user has a total of 105520 instances. The same number of informative instances have been taken for performing training from both the user groups i.e., 4776 instances of data samples.

4) Data preparation

The dataset of user C_1 is denoted here as D_1 and for user C_2 is D_2. Both the datasets D_1 and D_2 it contains consumer id therefore we removed the attribute Cid. The dataset has the issue of duplicate index. Thus, we resample the data according to the hourly frequency. During the re-sampling the sum of duplicate index is used to get the total hourly load. The hourly load for both the consumers is demonstrated in Figure 5.

Figure 5(a) shows the total hourly demand of consumer C_1 and Figure 5(b) shows the total demand of consumer C_2. According to Figure 5(a) the demand is increasing with the time but in domestic demand sudden picks are appeared. On the other hand, for industrial user it increases up to a particular limit and further increasing with the time.

Next the data is re-sampled based on total daily EED. After re-sampling the plot is demonstrated in Figure 6(a) and Figure 6(b). According to Figure 6(a) we can see the increasing demand pattern of the user but due to some special cases we can see sudden downfalls from the regular increasing patterns. The industrial pattern shows the smoother upper limit of power

Tab. 3: Shows raw samples of data set.

S. No.	Cid	Date	Pname	Rea ding	Ckey	Rco unt
1	1001 4678	2013-08-19, 14:48:40	Micro wave	0	281420	1
2	1001 4678	2013-08-19, 14:48:40	TV	0	281420	1
3	1001 4678	2013-08-19, 14:48:41	Dishwasher	0	281420	1
4	1001 4678	2013-08-19, 14:48:41	Kettle	0	281420	1
5	1001 4678	2013-08-19, 14:48:41	Washing ma-chine	.002	281420	1

Tab. 4: Shows data after pre-processing

S.No.	Date	Cid	Reading	Rcount
1	2013-08-19, 14:48:40	10014678	0	1
2	2013-08-19, 14:48:40	10014678	0	1
3	2013-08-19, 14:48:41	10014678	0	1
4	2013-08-19, 14:48:41	10014678	0	1
5	2013-08-19, 14:48:41	10014678	.002	1

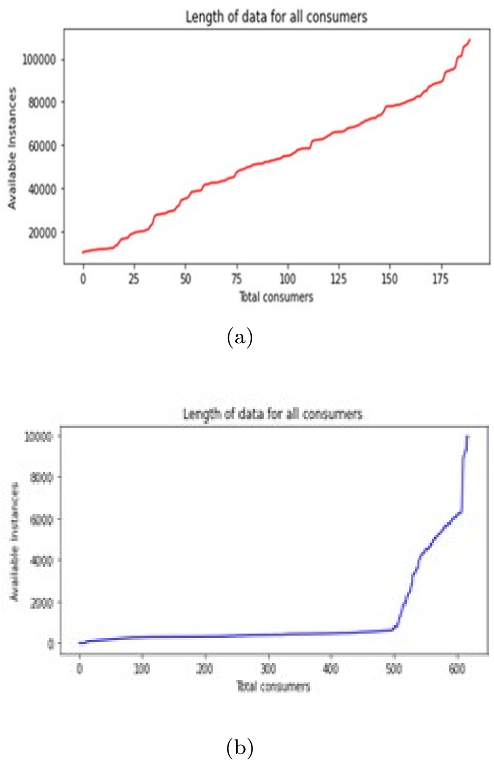


Fig. 4: Shows (a) the industrial consumers (b) the domestic consumers.

demand, but we can also find a significant among of downfalls in demand.

Next, we resample the dataset according to the total weekly EED pattern. Figure 7 provides a plot between weeks and total weekly EED. X axis shows the dates and Y axis contains total weekly EED. The weekly EED is demonstrated in Figure 7(a) for C_1 and Figure 7(b) shows the EED for consumer C_2. According to EED of consumer C_1 we find a regular and increasing demand pattern but reduced during some specific weeks. On the other hand, for consumer C_2 in reduces in specific patterns of week.

Similarly, the monthly total EED is measured and reported in Figure 7(c) and Figure 7(d). Figure 7(c) shows the demand on domestic consumer C_1 and Figure 7(d) shows the EED of consumer C_2. According to both the user's monthly EED we found smooth increasing patterns with the time. Thus, with the time the EED is increasing.

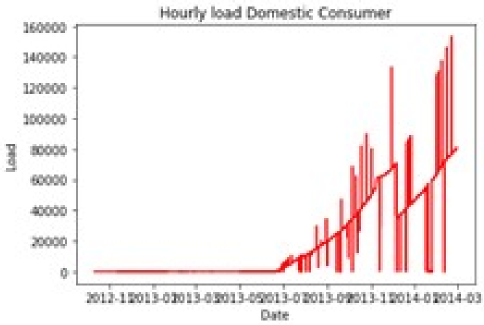
5) Seasonality analysis

The EED data is a time series problem. This time series is needed to solve and understand for building a prediction model. Therefore, time series data decomposition is essential to consider. Using this analysis, we know how the pattern of historical data varies with the time. The time series have four components.

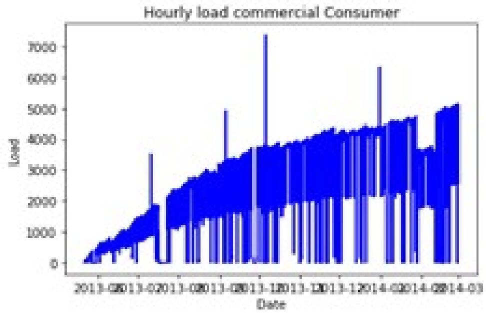
These components are:

- **Level-L** is the base value of the given time series. That provides an average value of consumers total EED with respect to time. The level is denoted as L.
- **Trend-T** is the type and rate in change of the values with respect to time. Mostly it provides an increase or decreases line with time.
- **Seasonality-S** is a cyclic event repeated in a specific time interval in each time series. That shows a type of wave form in increasing or decreasing manner. Therefore, it utilizes two parameters first is the entire EED pattern and a period on which we need to estimate a cyclic effect. The seasonality is denoted as S.
- **Noise or residual-R** is random variations in time series patterns. It is denoted as R.

The combination of the components is the reason of formation of time series. In a case, where seasonality and trend are the main part of the time series. Additionally, that has the major influence on the prediction value. The predicted values may also be different from

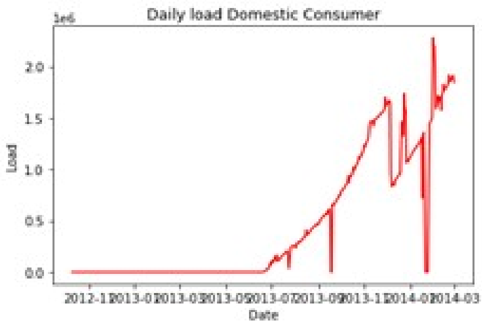


(a)

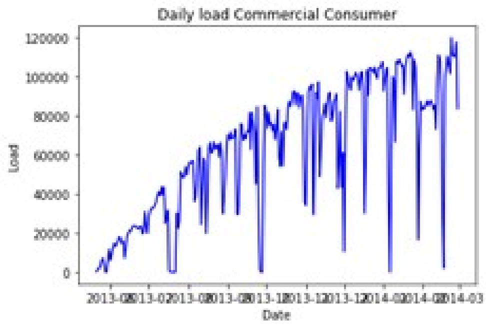


(b)

Fig. 5: Shows the total hourly demand for (a) consumers C_1 and (b) consumer C_2.

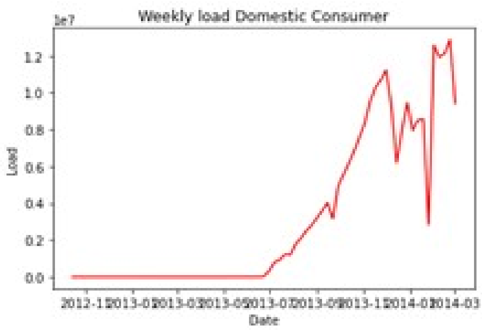


(a)



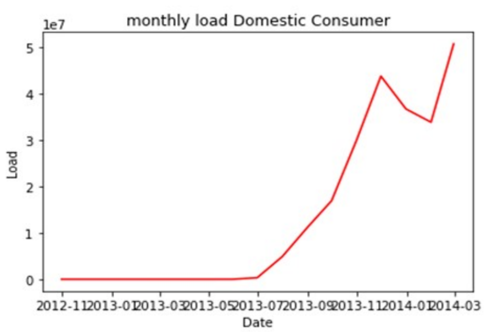
(b)

Fig. 6: Shows the total daily demand (a) consumers C_1 and (b) consumer C_2.

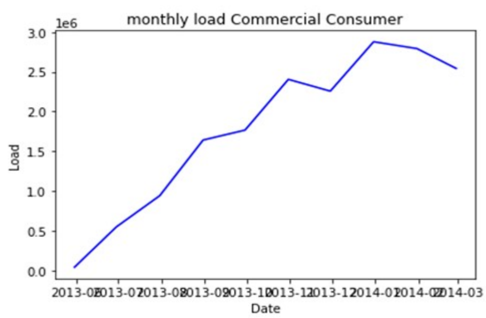


(a)

(b)



(c)



(d)

Fig. 7: Shows the total weekly load for consumers (a) C_1 and (b) consumer C_2, additionally monthly total EED is given by (c) for consumer C_1 and (d) for consumer C_2.

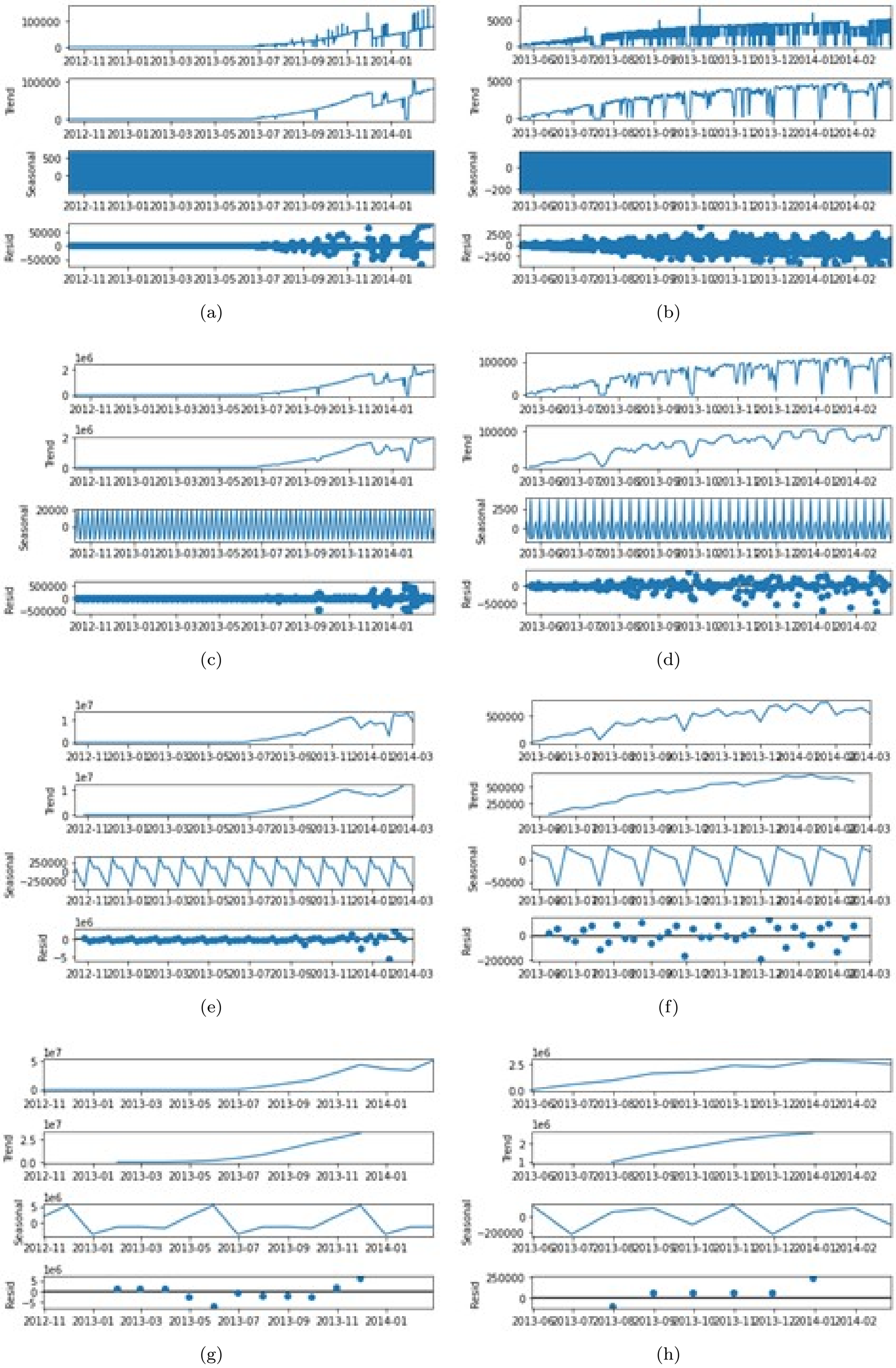


Fig. 8: Shows the visual seasonal time series decomposition of EED for variation of (a) consumer C_1 and (b) for consumer C_2 for hourly EED in 24 hour cycle, (c) and (d) shows the EED pattern for daily total demand and in 7 days period, (e) and (f) shows pattern for total weekly load in 5 week cycle, and (g) and (h) pattern for monthly load in a cycle of six months.

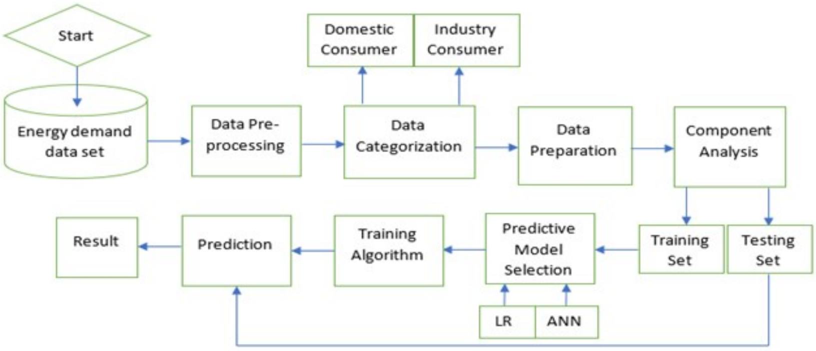


Fig. 9: Shows the proposed model for performing predictive data analysis for per hour EED prediction.

the time series, which is analyzed during training. The time series components can be performed in two types i.e., Additive or Multiplicative.

In additive decomposition the time series components are describing the relationship among data by adding them together. The additive time series is increasing or decreasing uniformly throughout the time. The additive time series can be defined as:

$$y(t) = L + T + S + R. \tag{2}$$

On the other hand, if the components are multiplicative, then it is called the multiplicative time series. This time series is demonstrating an exponential growth or decrement in their trend pattern. That can be given as:

$$y(t) = L \times T \times S \times R. \tag{3}$$

In this presented work the additive time series decomposition is performed. Figure 8 shows the visual time series decomposition. In these Figures the first sub diagram shows the level, second sub plot shows the trend, third provide seasonal cyclic pattern and last sub plot shows the residual. Figure 8(a) and 8(b) shows the total hourly demand-based data components for a period of 24-hour cycle. According to the visual pattern analysis the data shows increasing trend, but seasonality is not much clear. Additionally, the residual is too random with increasing time. But when the data is re-sampled in terms of daily total demand and measured the cyclic effect in period of 7 days is given in Figure 8(c) and 8(d). Next Figure 8(e) and 8(f) shows the component patterns of total weekly EED and seasonality is estimated at 5-week period. Similarly, Figures 8(g) and 8(h) demonstrate, the monthly total EED and cyclic relationship is estimated in six-month duration. According to the obtained visual pattern we found that the hourly demand prediction is complex as compared to other type of demand prediction. As the time duration is increased in data the seasonality is enhanced. Therefore, prediction becomes easier as compared to hourly demand prediction. Next, we are trying to predict the EED for both the consumer. Therefore, we have utilized two popular machine learning techniques for prediction namely Artificial Neural Network (ANN) and Linear Regression (LR). Before utilizing these ML algorithms, we transform the data from a linear manner to two-dimensional vectors. The vector organizes data into the following manner in Table 5.

Tab. 5: Shows data after pre-processing

Training Vector.	Predictable
$t_1, t_2, \ldots, t_{24}$	t_{25}
$t_2, t_3, \ldots, t_{25}$	t_{26}
$t_2, t_3, \ldots, t_{26}$	t_{27}

Then, we utilized a split function to split data into training (70%) and validation ratios (30%). The ML algorithms are applied for both kind of consumers and the predictions has been carried out. After training of the ML algorithms, the prediction on validation data is performed. The configuration of artificial neural network is given in Table 6.

Tab. 6: Configuration of artificial neural network.

Parameter	Values
Model	Sequential
Layer 1	Dense Number of neurons 128 Input size 24 Activation ReLu
Layer 2	Dense Number of neurons 64 Activation ReLu
Layer 3	Dense Number of neurons 32 Activation ReLu
Layer 4	Dense Number of neurons 1 Activation Soft-Max
Loss	Mean squared error
Optimizer	Adam

The complete data model for energy demand prediction is demonstrated in Figure 9. The proposed model first includes a historical energy demand dataset. The dataset is preprocessed in the next step. Preprocessing is aimed at removing unwanted or less relevant attributes from the dataset. Next, the preprocessed data is used for categorizing the EED into two user groups based on the demand. These groups are industrial and domestic consumers. Next, the individual

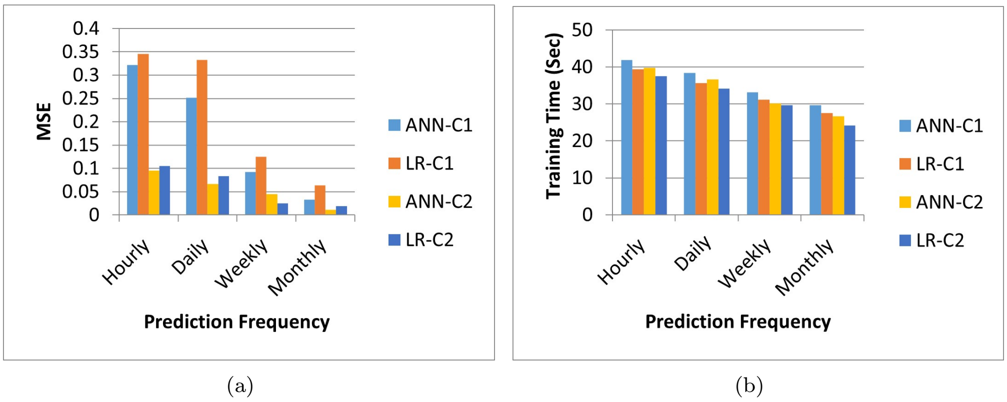


Fig. 10: Performance of the ML models algorithms for EED prediction on hourly, daily, weekly, and monthly frequency for domestic user and commercial user in terms of (a) MSE and (b) Training time.

consumers from each group were selected and the individual client's data is transformed into time series problem. Further, to understand the trend and seasonal impact of data in prediction. The decomposition of the formulated time series has been performed. Based on the seasonal effect of the data variation the machine learning models have been selected. Here we have selected two popular ML algorithms which are frequently used for continuous value prediction problems namely ANN and LR. Next the training of the ML algorithms has been carried out and the prediction has been performed. Based on the experiment conducted with the validation data the performance of both the algorithms for both the consumers has been evaluated in terms of Mean Square Error (MSE) and training time. The detailed evaluation of the performance experiments has been discussed in the next section.

4. Result Analysis

The EED prediction is performed in this experiment for both consumers' group i.e., domestic, and commercial. The experiments include the data according to the frequency of hourly, daily, weekly, and monthly formats. Additionally, we trained the two machine learning models for predicting the future EED. To measure the performance MSE and Training time is considered. Mean squared error (MSE) measures the error in predictions. It shows the average squared difference between the actual and predicted values. When a model has no error, the MSE equals zero. That can be measured using:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2, \qquad (4)$$

Where, y_i' is predicted values and y_i is actual value and n is the total samples for prediction in equation (4).

The MSE of different experimental scenarios is given in Figure 10(a). Figure 10(a) describes the error in predicted values with respect to actual EED values. The prediction is made for hourly, daily, weekly, and monthly demand. In this diagram, the X axis shows the frequency of demand and Y axis shows MSE. Based on the experimental results the ANN model is superior as compared to the LR algorithm. The prediction of hourly patterns is more complex than other frequencies. On the other hand, we also found that the EED prediction for industrial consumption is more accurate than for domestic usage patterns.

The second parameter is training time. The training time is the amount of time required to train an ML algorithm. The training time of the experimental scenarios is given in Figure 10(b). According to the results we have found commercial usage pattern training for both ANN and LR i.e. C2 requires less amount of time than C1. And clearly it has been observed that the ANN is consuming more time than LR in both customer group. Based on both the evaluation parameters, we found the ANN is the most suitable and accurate prediction method, but it has some limitations in training time as consuming more time than LR.

5. Conclusion

The smart grids are incorporating machine learning methods and cutting-edge communication infrastructure for making it more useful and profitable to balance demand and supply. However, currently, SG needs to make a balance between the several methods used to produce electricity and the various customer needs. This paper is focused on to understand the power demand of the different classes of users i.e., domestic, and commercial. Therefore, first, we provide a highlight of recent literature about smart grids based on communication technology and ML. Here, we found that proactive techniques are valuable for scheduling the power demand and supply. In this context, a large smart plug dataset is considered for experimental analysis. Additionally, two types of consumer patterns are obtained from the dataset. Using these two kinds of consumer patterns we perform the visual analysis for identifying the trends of hourly power demand, daily power demand, weekly power demand, and monthly power demand. Further, to design a proactive technique, we have conducted experiments with two popular ML techniques namely ANN and LR algorithms. Additionally, after training of these models, we predict the power demand for hourly, daily, weekly, and monthly demands. The experimentation demonstrates that hourly trend prediction is more complex than other kinds of demand predictions. In addition, both machine learning algorithm's comparative result shows that the ANN is providing a more accurate prediction than LR. And as for training time perspective LR requires less time. But by using proper tuning methods, the training time for ANN may be decreased for better results in ANN model. Furthermore, we found a common model that is not able to predict all types of trends accurately because of large data set and random variation in data. Therefore, to accurately predict trends of individual power demand, we need separate machine learning models and their configurations. The effective execution of the suggested methodology holds the potential to exert a substantial influence on the worldwide energy panorama and establish a pathway towards a sustainable future.

Acknowledgment

Author is thankful to Dr. Namit Gupta, Dean-Faculty of Engineering and Architecture, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India for his continuous guidance and support. Also, thankful to Dr. Navita Khatri for her initial guidance.

Author Contributions

Author, Jitendra Managre proposed the theory of work based on review of literature and done the analysis of modern trends developed in smart grids. Then on the concept of time series data, simulation part has been done in python using two machine learning algorithms, ANN and LR. Then according to the time and sector results has been analysed. The whole work has been done in guidance of Dr. Namit Gupta. He has contributed major in methodology and data analysis.

References

[1] GHASEMPOUR, A. Internet of Things in Smart Grid: Architecture, Applications, Services, Key Technologies, and Challenges. *MDPI Inventions*, Vol. 4, pp. 1-22, 2019, DOI: 10.3390/inventions4010022

[2] DIVAKAR, P. P., G.V. LAKSHMI, L. DEVI. Applications of Internet of things on Smart Grid. *IOP Conf. Series: Materials Science and Engineering*, Vol. 906, pp.1-3, 2020, DOI: 10.1088/1757-899X/906/1/012015.

[3] SALEEM, Y., N. CRESPI, M. H. REHMANI, R. COPELAND. Internet of Things-Aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions. *IEEE Open Access Journal*, Vol. 7, pp.62962-63003, 2019., DOI: 10.1109/ACCESS.2019.2913984.

[4] MOTLAGH, N. H., M. MOHAMMADREZAEI, J. HUNT, B. ZAKERI. Internet of Things (IoT) and the Energy Sector, *Energies*, Vol. 13, pp.1-27, 2020., DOI: 10.3390/en13020494.

[5] MANOJ KUMAR, N., A. A. CHAND, M. MALVONI, K. A. PRASAD, K. A. MAMUN, F.R. ISLAM and S. S. CHOPRA. Distributed Energy Resources and the Application of AI, IoT, and Blockchain in Smart Grids, *Energies*, 2020, Volume 13, page-5739, 2020, DOI: 10.3390/en13215739.

[6] AURANGZEB, K., S. ASLAM, S. M. MOHSIN, and M. ALHUSSEIN. A Fair Pricing Mechanism in Smart Grids for Low Energy Consumption Users, *IEEE Open Access*, 2021, DOI: 10.1109/ACCESS.2021.3056035.

[7] AL-TURJMAN, F., M. ABUJUBBEH. IoT-enabled smart grid via SM: An overview, *Future Generation Computer Systems*, Volume 96, Page 579–590, 2019, DOI: 10.1016/j.future.2019.02.012.

[8] DEWANGAN, F., A. Y. ABDELAZIZ, AND M. BISWAL. Load Forecasting Models in Smart Grid Using Smart Meter Information: A Review, *Energies*, vol. 16, pp. 1404, 2023, DOI: 10.3390/en16031404.

[9] ABDULLAH, A. A., and T. M. HASSAN. Smart grid (SG) properties and challenges: an overview, *Discover Energy* (2022) 2:8, DOI: 10.1007/s43937-022-00013-x.

[10] ALAHAKOON, D. and X. YU. Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey, *IEEE transactions on indus-*

trial informatics, vol. 12, no. 1, February 2016, DOI: 10.1109/TII.2015.2414355.

[11] GAN, L., P. JIANG, B. LEV and X. ZHOU. Balancing of supply and demand of renewable energy power system: A review and bibliometric analysis, *Sustainable Futures*, Vol. 2, pp.100013, 2020, DOI: 10.1016/j.sftr.2020.100013.

[12] DONG, J. Smart grid, Smart science, available at: `https://blog.frontiersin.org/2018/07/11/smart-grids-new-energy-research\protect\discretionary{\char\hyphenchar\font}{}{}specialty-section.`

[13] YOUSRI,D., H.E.Z. FARAG, H. ZEINELDIN, E.F. EL-SAADANY. Integrated model for optimal energy management and demand response of microgrids considering hybrid hydrogen-battery storage systems, *Energy Conversion and Management*, Vol. 280, pp.116809, 2023, DOI: 10.1016/j.enconman.2023.116809.

[14] BACANIN, N., C. STOEAN, M. ZIVKOVIC, M. RAKIC, R. STRULAK-WÓJCIKIEWICZ and R. STOEAN. On the Benefits of Using Metaheuristics in the Hyperparameter Tuning of Deep Learning Models for Energy Load Forecasting, *Energies*, 2023, Vol. 16, pp.1434, 2023, DOI: 10.3390/en16031434.

[15] ZENG, B., Y. QIU, X. YANG, W. CHEN, Y. XIE, Y. WANG and P. JIANG, Research on short-term power load forecasting method based on multi-factor feature analysis and LSTM, *Journal of Physics: Conference Series-ICSCAS-2022*, Vol. 2425, pp.012068, 2023, DOI: 10.1088/1742-6596/2425/1/012068.

[16] TAN, Z., B. LIU, Y. XU, S. FENG, and ZHAO. Application of Self-supervised Learning in Non-intrusive Load Monitoring", *Journal of Physics: Conference Series-ICSCAS-2022*, Vol. 2425, pp. 012037,2023, DOI: 10.1088/1742-6596/2425/1/012037.

[17] KUMAR, P., and A.R. ABHYANKAR. A Time Efficient Factorial Hidden Markov Model Based Approach for Non-Intrusive Load Monitoring," *IEEE Transactions on Smart Grid*, 2023, DOI: 10.1109/TSG.2023.3245019.

[18] WANG, Y., Y. KONG, X. TANG, X. CHEN, Y. XU, J. CHEN, S. SUN. Short-Term Industrial Load Forecasting Based on Ensemble Hidden Markov Model, *IEEE Access*, Vol.8, pp.160858 – 160870, 2020, DOI: 10.1109/ACCESS.2020.3020799.

[19] KONG, X., Z. WANG, F. XIAO, L. BAI. Power load forecasting method based on demand response deviation correction, International *Journal of Electrical Power & Energy Systems*, Vol. 148, pp. 109013, 2023, DOI: 10.1016/j.ijepes.2023.109013.

[20] YIN, H., K. ZHOU, and S. YANG. Non-Intrusive Load Monitoring by Load Trajectory and Multi feature Based on DCNN, *IEEE Transactions on Industrial Informatics*, pp. 1 – 12, 2023, DOI: 10.1109/TII.2023.3240924.

[21] DEY, B., S. MISRA and F. P. GARCIA MARQUEZ. Microgrid system energy management with demand response program for clean and economical operation, *Applied Energy*, Volume 334, pp. 120717, 2023, DOI: 10.1016/j.apenergy.2023.120717.

[22] RAI, S., and M. DE. Load forecasting using Two-level Heterogeneous Ensemble Method for Smart Metered Distribution System, *International Journal of Science and Technology*, Scientia Iranica, 2023, DOI: 10.24200/SCI.2023.59765.6410.

[23] QADIR, Z., S. IMRAN KHAN, E. KHALAJI, H. S. MUNAWAR, F. AL-TURJMAN, M.A. PARVEZ MAHMUD, A. Z. KOUZANI and K. LE. Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids, *Energy Reports*, Vol. 7, pp. 8465-8475, 2021, DOI: 10.1016/j.egyr.2021.01.018.

[24] ZENG, L., X. LAN, and S. WANG. Short-term wind power prediction based on the combination of numerical weather forecast and time series, *Journal of Renewable and Sustainable Energy*, 2022, DOI: 10.1063/5.0123759.

[25] KONG, W., Z. Y. DONG, Y. JIA, D. J. HILL, Y. XU, and Y. ZHANG. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network, *IEEE Transactions on Smart Grid*, Vol. 10, No. 1, 2019, DOI: 0.1109/TSG.2017.2753802.

[26] GEETHA, R., K. RAMYADEVI, and M. BALASUBRAMANIAN. Prediction of domestic power peak demand and consumption using supervised machine learning with smart meter dataset, *Multimedia Tools and Applications*, Volume 80, Issue 13, pp 19675–19693, ACM digital library, 2021, DOI: 10.1007/s11042-021-10696-4.

[27] ABUMOHSEN, M., A. YOUSEF OWDA AND MAJDI OWDA. Electrical Load Forecasting Using LSTM, GRU, and RNN Algorithms, *Energies*, 2023, Volume 16, PP-2283. 2023, DOI: 10.3390/en16052283.

[28] MANAGRE, J., and N. KHATRI. A Review on IoT and ML Enabled Smart Grid for Futurestic and Sustainable Energy Management, *International Conference for Advancement in Technology (ICONAT)*, march-2022, DOI: 10.1109/ICONAT53423.2022.9725932.

[29] Home Area Network Plug Readings, available at: https://search.data.gov.au/dataset/ds-dga-4e21dea3-9b87-4610-94c7-15a8a\77907ef/distribution/dist-dga-c7901c40-28e2-4601-bd3b-\38156428671a/details?q=sgsc

[30] Available at: https://www.dcceew.gov.au/

About Authors

Jitendra MANAGRE (corresponding author) received his B.E. in Electronics and Instrumentation Engineering from Institute of Engineering and Technology, Devi Ahilya Vishwavidyalaya, Indore, India in 2008 and done M.Tech. with specialization in Instrumentation from School of Instrumentation, Devi Ahilya Vishwavidyalaya, Indore, India in 2011. He is currently pursuing his Ph.D. at Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India. His research interest includes Process control automation, Python with machine learning and Smart Grids.

Namit GUPTA received his Ph.D. degree in Electronics Engineering from Devi Ahilya Vishwavidyalaya, Indore, India. He is currently working as the Dean- Faculty of Engineering and Architecture, and Director at Shri Vaishnav Institute of Technology and Science, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India. His research areas are nano electronics, 5G communications, networking, quantum dot automata, digital design logic, IoT and smart grid automation.