Building Energy Time Series Data Mining for Behavior Analytics and Forecasting Energy consumption

Balachander.K[1*], and Paulraj.D[2]
[1]Department of Computer Science and Engineering, Velammal Institute of Technology
Panchetti, Tamilnadu, TiruvallurDist,India
[e-mail:kbalachander.cse@gmail.com]
[2]Department of Computer Science and Engineering, R.M.K College of Engineering and Technology
Puduvoyal, Tamilnadu, TiruvallurDist,India
[e-mail:kingrajpaul@gmail.com]
[*]Correspondingauthor:Balachander.K

Received January 10, 2021; revised March 17, 2021; accepted May 12, 2021; published June 30, 2021

Abstract

The significant aim of this research has always been to evaluate the mechanism for efficient and inherently aware usage of vitality in-home devices, thus improving the information of smart metering systems with regard to the usage of selected homes and the time of use. Advances in information processing are commonly used to quantify gigantic building activity data steps to boost the activity efficiency of the building energy systems. Here, some smart data mining models are offered to measure, and predict the time series for energy in order to expose different ephemeral principles for using energy. Such considerations illustrate the use of machines in relation to time, such as day hour, time of day, week, month and year relationships within a family unit, which are key components in gathering and separating the effect of consumers behaviors in the use of energy and their pattern of energy prediction. It is necessary to determine the multiple relations through the usage of different appliances from simultaneous information flows. In comparison, specific relations among interval-based instances where multiple appliances use continue for certain duration are difficult to determine. In order to resolve these difficulties, an unsupervised energy time-series data clustering and a frequent pattern mining study as well as a deep learning technique for estimating energy use were presented. A broad test using true data sets that are rich in smart meter data were conducted. The exact results of the appliance designs that were recognized by the proposed model were filled out by Deep Convolutional Neural Networks (CNN) and Recurrent Neural Networks (LSTM and GRU) at each stage, with consolidated accuracy of 94.79%, 97.99%, 99.61%, for 25%, 50%, and 75%, respectively.

Keywords: Behavioral Analytics, Big Data Mining, Clustering Analysis, CNN, Energy Consumption, Energy Prediction, LSTM.

http://doi.org/10.3837/tiis.2021.06.001 **ISSN : 1976-7277**

1. Introduction

The measure of the consumption of household energy consumption is a substantial proportion of the usage of maximum energy around the world. The amount of use of household energy for maximum use of energy in certain European and American countries is approximately 30 % [1-2]. The rapid development of the economy and populations in India and China in the last decades has led to strong demand for energy, which has played a key role in the use of household energy consumption. The consumption of household energy thus contributed to real natural problems. For example, nearly 38 % of US carbon flows have their origin in the immediate usage of family units in the United States [3]. The consumption of energy in various households indicates large variability since different factors usually affect their use of energy. Fortunately, the consumption of energy has an enormous potential saving.

The time series estimates are one of the most important examinations of transient information and the expectation of future problems faced by data experts from money and finance into the creation of the board or communications broadcast. A measure is expected for certain events of the future. The current, med-term and long-term times are often named for forecasting problems. Short-term estimates include forecasting periods for a several periods (days, weeks, months) in the expectations. The medium-term calculations range from 1 to 2 years and long-term prediction problems will stretch for many years [4]. The details on time series data can be represented as a set of continuous series of expectations of an excitement element. Most approximate problems indicate utilization of these data, which is typically accomplished through methods for conventional empirical appliances. The enormous numbers of tests which define the time series data in various regions are of exceptional importance for data mining procedures [5].

The aim of the research is to aggressively consider earlier observations that discriminate between data structures that better reflect the internal context used in the system and quantify the hidden value of the data generation process [6]. The ARIMA is a remarkable scientific technique, generally used in time series, which demonstrates precisely the prediction of short-term and transient systems. Based on the time series, the Autoregressive (AR) method offers a sufficient illustration of such data generation portion. The assumption of these approaches is that knowledge regarding time series data is immediately subject to those prior readings from related periods. However, the expected performance of these techniques fails for long and complicated time series because the systems imply linear relations and fixed time-set properties. The electrical load profile since a large metropolis is based on dynamic cyclical and periodic indications defined by the mechanical architecture, human activities, and temperature effects. From the previous history, Nonlinear systems do better than linear representations, such as moving average (MA) model and autoregressive integrated moving average (ARIMA) model [7-9]. Overall, a few methodologies have been proposed to direct electrical load using consumer data and machine learning methods to increase the expectation of accuracy. However, it requires better load assessment models are still high [10].

In this way, deep learning methods are considered by various researchers late on and have shown considerable improvement in multiple domains such as acoustic presentation, image recognition, and natural language training. These fundamental systemic changes provide the potential for deliberation which enables complex non-nonlinear models [11-12]. In general, the recurrent neural network (RNN) is a deep learning method that is precisely designed to function overtime series and LSTM is a variation of RNNs [13]. This enables load sustaining which is propagated in layers forward and backward. Another variety of standard RNNs gated recurrent networks (GRU) which overcome the problem of the evaporated path like the LSTMs to show long term arrangements. LSTM and GRUs are innovative strategies to illustrate successive details by decoding logical data from previous data sources since they can understand complex non-linear instances and ultimately to distinguish the related fields [14-15]. Regardless of their popularity, the use of these deep learning models is typically rare and affects areas relevant to the Computer in general. One of the key parties which suggested an LSTM-based deep learning model-dependent time series data analysis is to determine short and medium electric loads. Although the suggested LSTM model decision has been more successful than the optional machine learning solution [16-17], the absence in feasibility has been exposed and thus a large-scale non-test application does not proceed. Therefore, it ignores complicated electric load qualities, related to periodicity, frequency of data, models, stages, auxiliary breaks, and timing impacts, which are seen by a time series. Specifically, perplexing examples of electrical loads are neglected every day, every week, every month, and last year as contributions to the previous LSTM model [18]. It finds only one category of loads in the past as the approach to forecasting electrical loads in the short and medium-term. The recently proposed model correctly overlooks critical domain details which may impact the validity and vitality that are determined [19]. **Table 1** shows three categories of studies related to the prediction of power consumption.

A load prediction model based on improved LSTM is proposed, where the periodicity of electrical load is taken into account while using multiple sequences as major data sources. So as to consolidate the previous measurements of the electrical load used by LSTM are integrated, a k-means cluster analysis is developed in data mining operation, for recurrence is established and the time lags related to the Multifunctional LSTM are realized. The two deep neural networks, LSTM (Long short-term memory) and GRU (Gated recurrent unit), are developed for comparison and validation. The results of these experiments indicate that LSTM and GRU multi-sequence models predict the other possible methodologies used. Our proposed system is follows

(1) Normalize data and use models based on time series representations to measure standard energy consumption for each consumer;

(2) The K-means clustering is used to divide data into several groups based on the similarity to increase the accuracy of the prediction.

(3) Consumptions within clusters clustering and aggregation. The deep learning prediction model is learned for each cluster and the forecasts for the following time are completed;

(4) Prognosis is paired with actual intake data and compared. Next, a day-ahead forecast for the obtained cluster representations using the prediction methods were created.

Table 1. Related works on electric energy consumption prediction

Category	Reff. No.	Year	Method	Advantages	Disadvantages	Description
Statistical modeling						
	7	2017	ARIMA	Forecasting non-stationarity short term datas	No automatic updating	Analysis of linear correlation structure based prediction performance according to time resolution
	13	2020	ARIMA			Energy, Finance and Medicine long-term and multivariate time series prediction
	12	2018	SARIMA			Household energy consumption prediction analysis using single and ensemble technique
			Linear regression	Lower time complexity	Sensitive to outliers	
	10	2018	FCM	Decision making for uncertain information	Cognitive uncertainties	Fuzzy based energy load prediction and exception calculation
			PSO	Memory storage	local minimization	PSO based Long-Term energy load prediction
	9	2017	Entropy based	Less complexity	only short term data	Multiple temporal scales based ECG feature extraction
Spatial information modeling						
	10	2018	ANN	Unsupervised Learning	Overfitting	ANN-based learning technique for STLF
	12	2018	ANN			Household energy consumption prediction analysis using single and ensemble technique
	17	2018	MLP			Predicting long-term time series commercial and residential building electricity consumption
			SVM	Structural Risk Minimization	Parameter uncertainty	
	13	2020	SVM			Energy, Finance and Medicine long-term and multivariate time series prediction
	7	2017	SVM	Structural Risk Minimization	Parameter uncertainty	Classification technique based on statistic learning
Temporal information modeling						
	13	2020	LSTM	Non-linear function of more accurate	Training time is much longer	Energy, Finance and Medicine long-term and multivariate time series prediction
			GRU	Simple linear operation	Add external information	
			RNN	Extract temporal features	Continuous time dependency	
	17	2018	RNN			Predicting long-term time series commercial and residential building electricity consumption
	16	2019	CNN-LSTM	Extract complex features	Complex design	Household energy consumption prediction and feature extraction

2. The Forecasting Framework of Proposed Approaches

The proposed approach consist of three main developments are found in the methodologies presented in this article. The model consists of three stages namely data collection, data mining, and prediction. In order to calculate the relation between the factors and the appliances' energy consumption, the Pearson coefficient is used, and the key highlights of appliances energy are divided by the Pearson coefficients. Secondly, the K- means clustering technique is often used to separate details in a few clusters which concentrate on the similarity of features. Finally, in each cluster deep learning networks are trained, and the household energy is obtained by a combination of the outcomes of the various LSTM and GRU networks in the preferred cluster.

Step 1: In a database for additional evaluation, rough data is collected from each house containing a huge amount of vitality records of each unit gadget in relation to time arrangement details.

Step 2: Rapidly, successive mining and grouping trends are carried out in this stage. Frequent prototypes are persistent configurations or collections of systems frequently presented in a dataset. For example, light and fan are seen continuously, they are seen as a continuous sample. The goal is to expose the relationships between appliances and consider the time of use in terms of hours and duration for the day, weeks, months and year. Such observed information throughout time series data allows it possible for appliance associations to be found by clustering of devices over time. Cluster analyses are the method of forming groups where cluster members have similarities and discrepancies between clusters and members of other clusters.

The detection of repeated trends and the study of clusters are widely recognized in broad datasets as a costly. Similarly, data creation is a continuous process in real-world scenario, where new records are created and the old records becomes redundant over time, developing new classes of clusters. By this technique, the high resolution energy consumption time series data are produced. Accordingly, an incremental and systematic upgrade strategy is necessary, taking ongoing data changes into consideration and updated knowledge is used to preserve current regular trends and clusters. This goal is accomplished by systematic incremental processing of data and removing the need to repeat mining the entire database at frequent intervals. For a broad database it is possible to use the sample development method for the regular pattern mining, [38, 39] while the cluster analysis can be performed using the k-means analyses clustering [40]. Established frequent patterns and clusters are replaced with new data during each continuous mining activity and new found patterns and clusters are applied gradually to the permanent data base. This methodology minimizes memory overhead and increases performance for real-world applications to just a fraction of the entire database at each iterating phase.

Step 3: The design continuously and honestly emphasizes basic knowledge about the laws of interaction of computers from ordinary appliances. The probabilistic relationship of devices with a different appliance can be defined by usual systems and groups formed. A prediction approach focused on deep recurrent neural networks (LSTM and GRU) to handle the complex use and viability of devices both for a short term and long-term assessment was used.

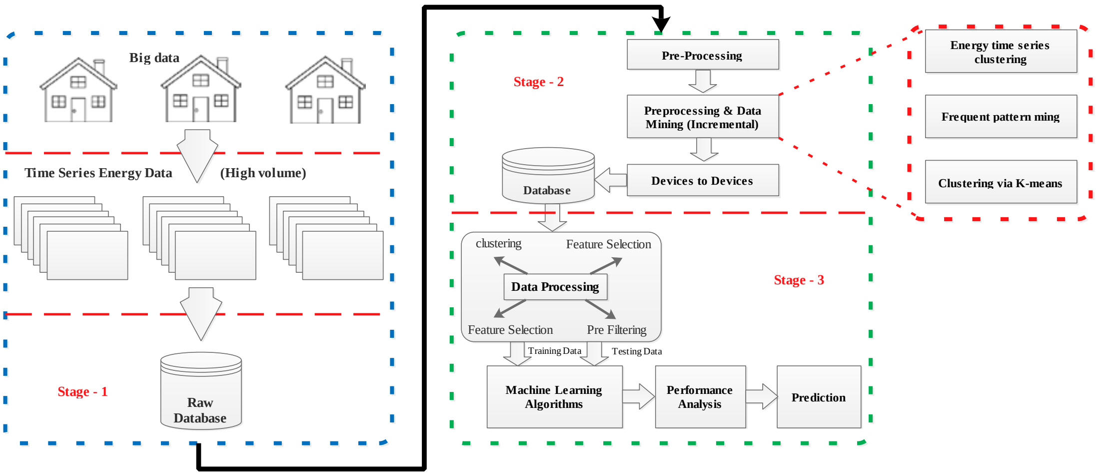


Fig. 1. Proposed Model Flow

Fig. 1 displays a proposed model that illustrates probable similarities between appliance usage and activities. The individual operation of appliances is not only the isolation of human movement designs but also the association of appliances; i.e., examples of movements that are mixed, e.g. washing appliances after preparation or looking at the TV. The basic concept of every model relies on the design pattern growth or frequent pattern development approach, using a depth-first divide-and-conquer technique strategy [20-22]. This operation was not usually being carried out, but it was detached, which is obviously not applicable to well-being applications involving fast dynamic response.

2.1 Data Preparation

The raw energy time-series data is a time-scale, converted into energy usage or load data with a resolution of 1 minute. It is then converted for the next step of the data mining operation into 30-minute time-resolution data sources. Therefore, the data reduction is 24 x 2 = 48 readings per day per unit, while the period of use, average charge, and energy consumption are reported for each active device. The source information for regular pattern mining and cluster analysis comprises all appliances reported active during this 30-minute time period. The 30-minute time resolution is evaluated and it fits ideally because it accurately measures appliance-time and appliance – appliance. This analysis was performed using three datasets (two actual and one synthetic). The first actual data collection includes over 5 houses with a timely settlement of over 400 million raw energy usage documents. During the pre-processing period, it was reduced to just over 20 million without lack of accuracy or exactness. Similarly, from more than 21 million raw records, the second actual dataset AMPds2 was reduced to four million records [23-24], originally 1-minute duration A simulated dataset of more than 1.2 million records for the preliminary assessment have also developed for our model. A collection of source data containing four appliances for one home ready to search for regular trends and clusters can be found in **Tables 1 and 2**.

Finally, the data are recorded continuously; they can also have applications to operate electrical systems in real time and to control them. All these factors stress the need to use new power consumption clustering methods. The clustering method based on K-means with mezoids as a centroid option to classify consumers into groups (clusters) was used. The optimum number of clusters K is determined with the Silhouette coefficient for every representation of the dataset.

2.2 Frequent Pattern Mining

Associations of appliances and appliance-time are important behavioral features of consumer energy usage and can define peak load / economic hours of energy use. These correlations further describe the behavioral patterns of the respective residents and their predicted comfort [25]. With the large amount of data continuous from smart meters collected, it is also of a great interest not only for utilities and energy suppliers, but also for customers to draw up such regular trends and decision-making clusters such as energy cost containment, managing demand responses and energy efficiency strategies. The repeating pattern can be observed in recurrent pattern mining carried out through the input data and provided in **Table 2**: i.e. a pattern containing continuously occurring item sets [26]. Let $\Gamma = \{I_1, I_2, \cdots, I_k\}$ be a set of items consists of k objects, k-item set (I_k). Let DB mean a database of the transaction where each transaction Y is such that $Y \subseteq \Gamma$ and $Y \neq \phi$. The transaction database is seen in **Table 1**. The item set support amount has been specified as its appearance frequency; i.e., number of transactions comprising the item set. Let's set two objects or patterns in $X(X \subseteq Y)$ and $\Upsilon(\Upsilon \subseteq Y)$. X and Υ are called regular patterns because they are more or less equal to their respective s_X and s_Υ support. The minimum assistance threshold is a pre-determined *minsup*. Frequent patterns or typical patterns are processed for the production of association laws, extracted by the mining process. The rules of association, with $\{X \Rightarrow \Upsilon\}$, are formed through support- confidence structures, where $s_{X \Rightarrow \Upsilon}$ support is the proportion of transactions $(X \cup \Upsilon)$ in the database and that can be referred as $P(X \cup \Upsilon)$ probabilities. In the transaction database DB, consists of X which includes Υ; i.e., conditional probability $P(X / \Upsilon)$, confidence $c_{X \Rightarrow \Upsilon}$, as seen in equation, can also be expressed as the ratio of transactions [27-28]. The description of support and confidence is described in equations (1) and (2), in both:

$$support(X \Rightarrow \Upsilon) = s_{X \Rightarrow \Upsilon} = support(X \cup \Upsilon) \tag{1}$$

$$confidence(X \Rightarrow \Upsilon) = \frac{support(X \cup \Upsilon)}{support(X)} \tag{2}$$

Table 2. Source Database for Frequent Pattern Mining

Date	Start time	End Time	Active appliances
2/18/2017	6:00	6:15	1, 3, 4,7
2/18/2017	6:15	6:30	1, 3, 4, 7
2/18/2017	6:30	6:45	1,3,4,11
2/18/2017	6:45	7:00	1,3,4
2/18/2017	7:00	7:15	4
2/18/2017	7:15	7:30	4,12
2/18/2017	7:30	7:45	4
2/18/2017	7:45	8:00	4,13
2/18/2017	8:00	8:15	1,4,15
2/18/2017	8:15	8:30	1,4,14
2/18/2017	8:30	8:45	1,2,3,4
2/18/2017	8:45	9:00	2,3,4
2/18/2017	9:00	9:15	4
2/18/2017	9:15	9:30	4
2/18/2017	9:30	9:45	4
2/18/2017	9:45	10:00	3,4
---	---	---	---

2.3 Cluster Analysis

Appliance-time relationship comprehension can enable the strategic study of consumption energy activity in addition to the introduction of inter-appliance associations. Appliance-time associations can be defined in terms of the hour of the day (00:00–23:59), time of day (Morning/Afternoon /Night), weekday, week of the month, week of the year, and month of the year. Appliance to time comprehension can support critical analysis of consumer energy consumption behavior, in addition to exposing the inter-appliance. Based on the hour of the day, the week, the month, the year, the Appliance-time associations are being specified. This can be calculated in terms of the hour of the day [29]. Looking at device-to-time relations with active appliances to create a class or a cluster for the device may be seen as adding a correctly adjacent time stamp. The defined clusters identify the relationships of equipment-to-time and the size of the clusters, determines the potential strength of the clusters according to the membership count [30]. The discovery of appliances-to-time correlations can also be translated as clustered devices in time-interval groups; where each cluster is one unit, with time-marks as cluster members are shown in **Fig. 2**.

The evaluation of clusters is a method for creating data point batches based on results, but not externally i.e., the unsupervised classification. The whole extracted data describes the interaction between data points and helps to classify data points in order to achieve the data points are similar to each other, yet are far from members of all other clusters. “better” and “faraway” being association measures which define the near relationship between cluster members [31]. Thus, the clustering evaluations performed on Input data provided in **Tables 3 and 4** will produce clusters or groups that describe normal appliance associations in time, whereas the degree of association is specified by support or strength. This appliance to times does not only calculate the full load or hours of energy usage but do also shows the behavioral characteristics of the residents and customers' energy use.

In order to derive device time correlations, one of the most commonly used partial clustering concepts was expanded. The cluster prototype is determined by its centroid, and is the mean of all the data points of the cluster. In this method, clusters consist of groups without overlapping, i.e., each data point belongs to only one group or cluster. In addition, mining was carried out gradually.

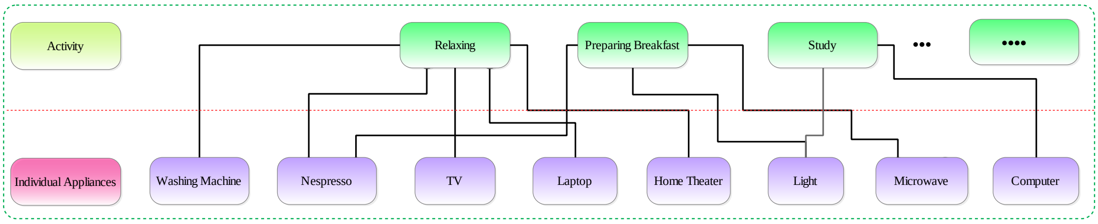


Fig. 2. Possible Clustering between the Appliance

Table 3. Clustering Source Database-I

Appliance	Hour of the day	Time of the day
TV (2)	11:30 - 12:20	Morning and Afternoon
Washing machine (5)	10:15 - 11:10	Morning
Water pump (7)	10:00 - 10:25	Morning
Music Systems (11)	14:25 - 16:15	Afternoon and Evening

Table 4. Clustering Source Database-II

Appliance	Days	week	Month
TV (2)	267	38	All
Fan (3)	186	26	All
Washing machine (5)	98	14	All
AC (6)	46	6	3,4,5
Water pump (7)	136	19	all

An initial perspective on the clustering of k-means centroid [32-33] have been considered. The data clusters of the data dataset DB, which has n data points on the Euclidean, are allocated by partial clusters to a k number of clusters, $C_1, C_2, \cdots, C_k$, which have centroids $c_1, c_2, \cdots, c_k$, so that the data are $C_i \subset DB$, $C_i \cap C_j = \phi$ and $c_i \neq c_j$ to $(1 \leq i, j \leq k)$. The Equation (4) represents the objective function based on the Euclidean distance to calculate the consistency between data points that exhibit the quality of cluster. The objective function is defined as squared error sum (SSE), defined in Equation (3), and the SSE is reduced to a minimum with the k-means clustering algorithm.

$$SSE = \sum_{i=1}^{k} \sum_{d \in C_i} distance(d, C_i)^2 \tag{3}$$

$$distance(x, y) = \sqrt{\sum_{j} (x_j - y_j)^2} \tag{4}$$

The properties of the individual object where x_j and y_j are, and j ranges from 1 to n. The k-means start from database by selecting k data points, where $k \leq n$ and k clusters are generated with centroids. Then, a cluster with a less Euclidean distance (c_i) from its centroid, distance (d, c_i) will be allocated to any of the remaining d data points in DB. Only the updated cluster centroids are evaluate by computing the cluster center used for clusters after a new data object has been allocated, along with the k-means algorithm modifies the cluster structure regularly, thereby minimizing intra-cluster inequalities through reallocating the data points before clusters will be balanced [34]. During the measurement, any update is not possible by the estimation in sum of the squared error (SSE) from its centroid covering all cluster data points.

2.4 K-Means Clustering Algorithm

Original centroids are spontaneously calculated by using K-means. The following steps are:

(1) Evaluate the value of k, where k is the number of clusters that are needed.

(2) Original centroids are determined. The centroid is randomly allocated from current results, and the number of clusters equals the original centroid number.

(3) Measure the distance to a centroid by using the Euclidean Distance Function, in the nearest centroid of each data point.

(4) Minimum distance clustering of results. If it is the nearest from its cluster base, a data point would be part of a cluster.

(5) Identify new centroid data centers for each cluster based on average results.

(6) Repeat step 3.

(7) Stop if the cluster assignment does not alter results.

Fig. 3 illustrates the flow chart of the K-means algorithm. In this research, the data used were a collection of electricity building data, and the input vector form had time-specific characteristics. For forecasts, data from time series are typically used. For building appliance energy prediction used time-series data [35]. The clustering of time series is split into two groups: first the model or feature, and the raw data. In a feature clustering process, the raw data are compiled, converted, or converted using feature elimination or parametric models, e.g., dynamic regression, ARIMA, machine learning, and deep learning models to promote clustering [36-37]. The clustering raw database relates clearly to time-series vectors before the clustering step without any spatial transformations. A raw data-based clustering for the process was employed.

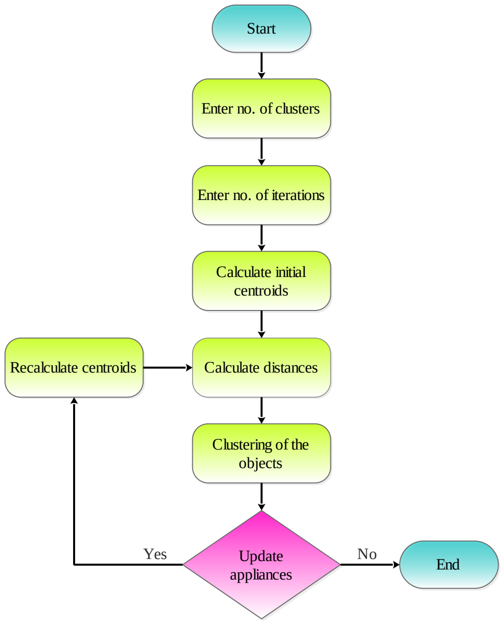


Fig. 3. K-means Clustering Algorithm Flow Diagram

2.4.1. Optimal k-means: Determining k using *Silhouette coefficient*

The cluster quality is assessed by analyzing intra-cluster cohesion and inter-cluster separation of data points. A coefficient of silhouettes which is based on the Euclidean length is exploited to assess an optimum number of clusters. The silhouette coefficient indicates "How well clusters are constructed," measuring the degree of similarity and differences. As defined in equations (5) to (9), the silhouette coefficient can be calculated.

• Compute average distance a_j to all other data points of d_j in cluster C_i

$$a_j = average\left\{distance\left(d_j, d_i\right)\right\} \tag{5}$$

where, $d_i = (d_1, d_2, \cdots, d_n);\ d_i \neq d_j$

• Compute Average distance of d_j to all other data points in clusters C_i, having $i \neq j$; Determine b_j= *minimum*(b_j) across all the clusters except C_i.

$$b_j = average\left\{distance\left(d_j, d_{i/C_x}\right)\right\} \tag{6}$$

where, $d_i = (d_1, d_2, \cdots, d_n)$ and $C_x = (C1, C2, \cdots, Cn); C_x \neq C_i$

• Compute *Silhouette coefficient* for d_j

$$S_{d_j} = \frac{(b_j - a_j)}{maximum(b_j, a_j)} \tag{7}$$

• Compute *Silhouette coefficient* for cluster C_i

$$S_{C_i} = average(S_{d_j}) \; for \; j = d_1 \cdots d_n \tag{8}$$

• Compute *Silhouette coefficient* for clustering, having *k* clusters

$$S_k = average(S_{C_i}) \; for \; i = 1 \cdots k \tag{9}$$

The coefficient in silhouette may range from -1 to 1. In cases where the negative value shows that the average distance of a data point from a data point to a data point in cluster C*i* (a*i*) is greater than the average d_i to a cluster of data points other than C_i (b_i) and there are better clusters. Overall cluster quality can be measured by calculating the Silhouette average coefficient by computing the silhouette average for all cluster member data points (Silhouette width) in equation 9. In the final analysis, the quality analysis process for formed clusters is repeated, where n is the unique database/dataset data point set, while the silhouette (silhouette width) is computed, and k is chosen with the highest Silhouette width.

2.5 Forecasting Methods

The sequential information where the output depends not only on the present inputs but also on the previous inputs are used by a Recurrent neural network. It is called recurring with each variable throughout the data sequence, the data is similarly processed. Owing to the internal memories, RNNs may recall essential inputs and thus desired data for time series. However, the RNNs model stops learning when the values of the gradient become too small because of the absence of the gradients problem. The RNN setup here on input sequence, where the input is x_t and the hidden state of s_t is at step *t*, and is the network memory, is seen in **Fig. 4**. The problem of vanishing gradients affects regular RNNs, since the gradients are becoming smaller over time as the system step back. In contrast to the subsequent layers in the system, neurons in previous layers thus grow very slowly. This complexity of gradient propagation is analyzed by LSTMs and GRUs [38-39]. The input, forget and output portals that decide when new information is introduced to the cell state, extracted from the memory and output port to decide how it is extracted from the memory were implemented. The information flow and gates set inside the LSTM cells are seen in **Fig. 5**.

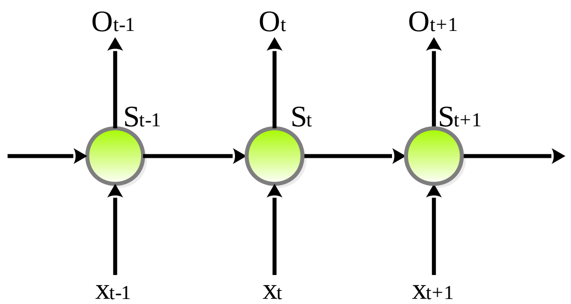


Fig. 4. Unfolding during the RNN network estimation

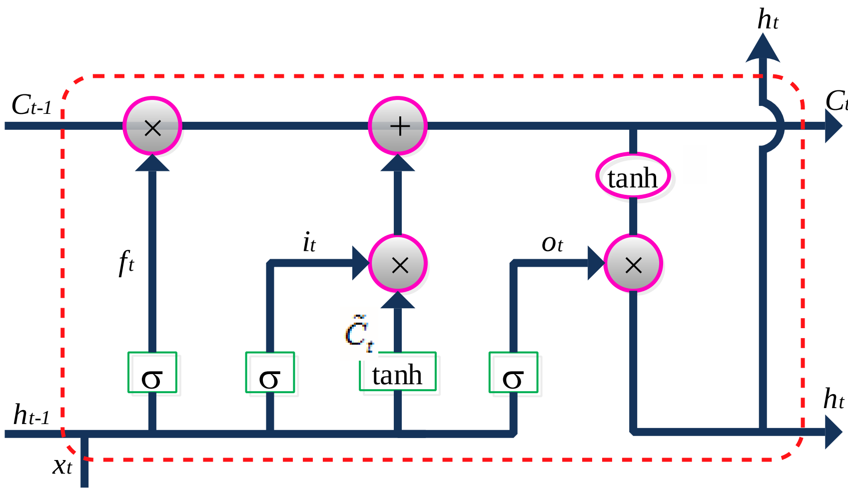


Fig. 5. Data flow in the RNN LSTM block

LSTM cell gates minimize the probability that gradients may disappear and support the learning for long-term dependency [40]. Such gating function allows a great deal of control about what the LSTM cell recalls, forgets and maintains its internal cell memory successfully over time. Using a recurring function as defined in equation (10), the input sequence of the LSTM model is {X1, X2, ... ,Xn}.

$$h_t = f(h_{t-1}, x_t) \tag{10}$$

Where xt is the input and *ht* is the hidden state. In order to solve the problem of the loss of blast, gates are placed into the recurrence function f. LSTM cell states are determined accordingly

$$i_t = \sigma\left(W_i \bullet [h_{t-1}, x_t] + b_i\right) \tag{11}$$

$$f_t = \sigma\left(W_f \bullet [h_{t-1}, x_t] + b_f\right) \tag{12}$$

$$o_t = \sigma\left(W_o \bullet [h_{t-1}, x_t] + b_o\right) \tag{13}$$

$$\tilde{C} = \tanh\left(W_C \bullet [h_{t-1}, x_t] + b_C\right) \tag{14}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{15}$$

$$h_t = o_t \odot \tanh(C_t) \tag{16}$$

where i_t,is the input gates, f_t, is the forgot and o_t is the output gates. W_s represents weights and b_s represents the bias of the LSTM unit parameters and the existing cell status is labeled as C_t and new cell status of the claimant are labeled as C_e. Three sigmoid functions, i_t, f_t, and *ot* gates, are expressed in equations (11) to (13). Owing to the *xt* input and the previous h_{t-1} output, the three gates obstruct the signal or pass it over. The signals are blocked, specifically unless the signals are blocked, unless the gate is 0. The forget gate f_t calculates the previous output h_{t-1} which are allowed to move through the gate. The input gate it to change the cell state determines the data and the output gate *ot* specifies the cell status of the output. Equation 15 transfers the current cell state C_t with the existing cell state C_{t-1}. A hyperbolic-tangent function defined by equations 14 and 16 respectively calculates the new candidate values $\tilde{C}$ of a memory cell as well as the output of the current LSTM block ht. Each stage is immediately passed to the next cell by the two states $\tilde{C}$ and h_t. The weights W's and biases b's are established while minimizing the variations between the LSTM outputs and the actual training samples GRU's configuration is similar to an LSTM cell, but it only has two gates, namely update, reset, and gates. GRU is a popular version [39]. As seen in **Fig. 6**, the model is easier and is therefore computationally quicker than typical LSTM models. Just as with LSTM, it is easier to train as fewer calculations are needed for updating the hidden state due to the simple internal structure. In dynamic modeling scenarios, specially qualified GRU can perform exceptionally well.

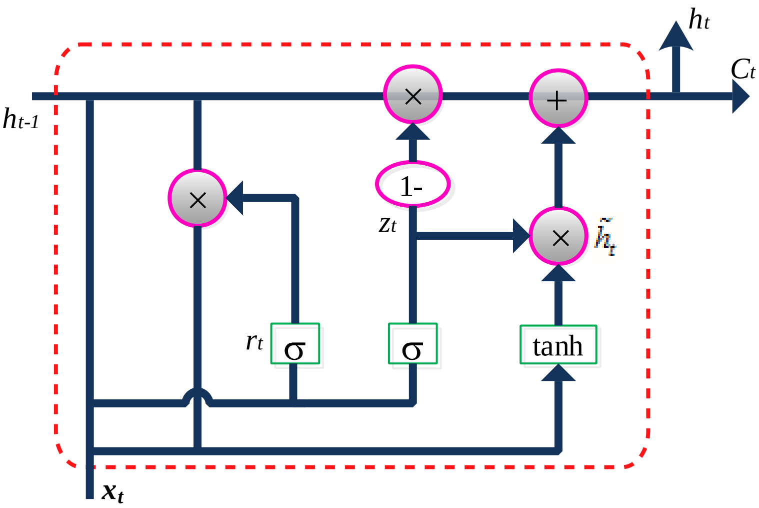


Fig. 6. Information flow in GRU block

Update gate z, reset gate rt and cell states ht and $\tilde{h}t$ for GRU are computed using the following equations (17) to (20):

$$z_t = \sigma\left(W_z \cdot \left[h_{t-1}, x_t\right]\right) \tag{17}$$

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1}, x_t\right]\right) \tag{18}$$

$$\tilde{h}_t = \tanh\left(W \cdot \left[r_t \odot rh_{t-1}, x_t\right]\right) \tag{19}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{20}$$

2.6 Performance Metrics for Evaluation

The four types of measurable standards are used to statistical evolution of time series models' precision in prediction. Root Mean Square Error (RMSE), accuracy, R-square (R^2) is the relevant factor test and is a regression model that discusses to the amount of the statistical difference/variance of an indigenous required attribute, Mean Absolute Error (MAE), and a mean of the value of total estimates/mean of the total estimates for the naive model is known as Relative Mean Absolute Error (RMAE) [41]. In accordance with equation (21) to (24), the four criteria for performing tests used in this investigation can be identified.

$$R - Squared = \frac{n\left(\sum ab\right) - \left(\sum a\right)\left(\sum b\right)}{\sqrt{\left[n\sum a^2 - \left(\sum a\right)^2\right]\left[n\sum b^2 - \left(\sum b\right)^2\right]}} \tag{21}$$

$$RMSE = \frac{\sqrt{\sum_{t=1}^{n}\left(x_{ft} - x_{rt}\right)^2}}{\sqrt{n}} \tag{22}$$

$$MAE = \frac{\sum_{t=1}^{n}\left|x_{ft} - x_{rt}\right|}{\sqrt{n}} \tag{23}$$

$$RMAE = \frac{\sum_{t=1}^{n}\left|x_{ft} - x_{rt}\right|}{\sum_{t=1}^{n}\left|x_{rt}\right|} \tag{24}$$

3. Results Analysis and Discussion

The findings of the processing of 25% of data from the independent data collection for both houses 4 and 5 are seen in **Table 5**. The appliances such as Lights, TV (Television), Fans, Refrigerator, washing machine, Air conditioner, Water pump (Laptop, Smartphone, Music systems, and so on) have been taken into consideration. Through these affiliations, the community will produce daily family behavioral applications that are high- and low-powered (kW) appliances; however, due to the extremely extensive usage of house 4, they have tremendous energy impressions. Huge appliances, including washing machines and water pumps with periodically reduced energy measurements, do not usually cause higher power costs. Machines are often unreasonably used by customers because of their behavior. They are also blamed for this. The actual energy use and long-term energy supply against true energy use in Home 4 are seen in **Fig. 7**, with expanding usage on Sundays and during the late spring season, and in **Fig. 7**.

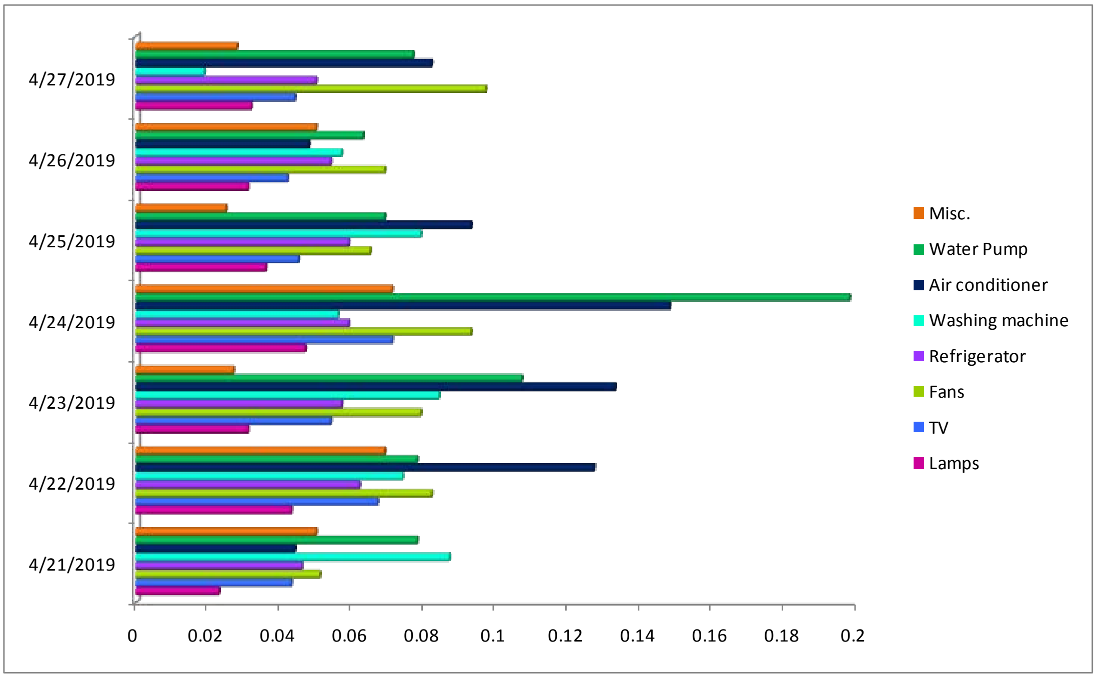


Fig. 7. Energy utilization of appliance-day during the week

Table 5. Appliance - day during the week

Date	4/21/2019	4/22/2019	4/23/2019	4/24/2019	4/25/2019	4/26/2019	4/27/2019
Lamps (kW)	0.023	0.043	0.031	0.047	0.036	0.031	0.032
TV (kW)	0.043	0.067	0.054	0.071	0.045	0.042	0.044
Fans (kW)	0.051	0.082	0.079	0.093	0.065	0.069	0.097
Refrigerator(kW)	0.046	0.062	0.057	0.059	0.059	0.054	0.05
Washing machine (kW)	0.087	0.074	0.084	0.056	0.079	0.057	0.019
Air conditioner (kW)	0.044	0.127	0.133	0.148	0.093	0.048	0.082
Water Pump (kW)	0.078	0.078	0.107	0.198	0.069	0.063	0.077
Misc. (kW)	0.05	0.069	0.027	0.071	0.025	0.05	0.028

The changing effect of time, days and months on appliance use on the basis of appliance behavior was found. **Fig. 8** shows the associations of appliance-time found. The cluster 1 with the highest concentration of clusters 2 appliances, such as water pump or washing machine between 11:30 and 13:00, and between 18:00 and 21:00, which are expanded in one single day was found. Cluster 1 has the lowest energy consumption in TV, music, and diversity products. The same frequency throughout the day with large usage concentration of Cluster1&2 devices increased energy consumption during the day. The number of clusters generated during the data mining process is listed in **Fig. 9** and further supports the invention of associations of appliance time that replicate the behavior of the appliances over a period of time.

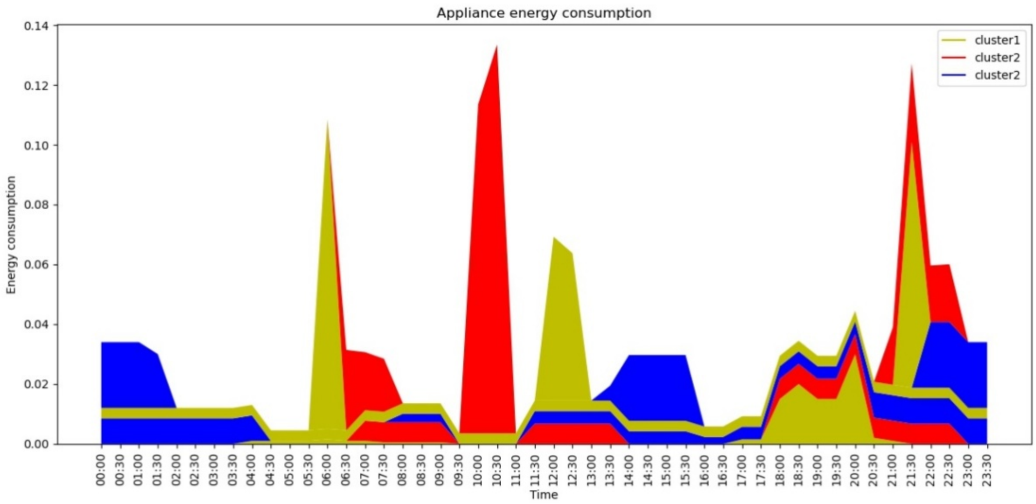


Fig. 8. Appliance-time associations @ hour of day

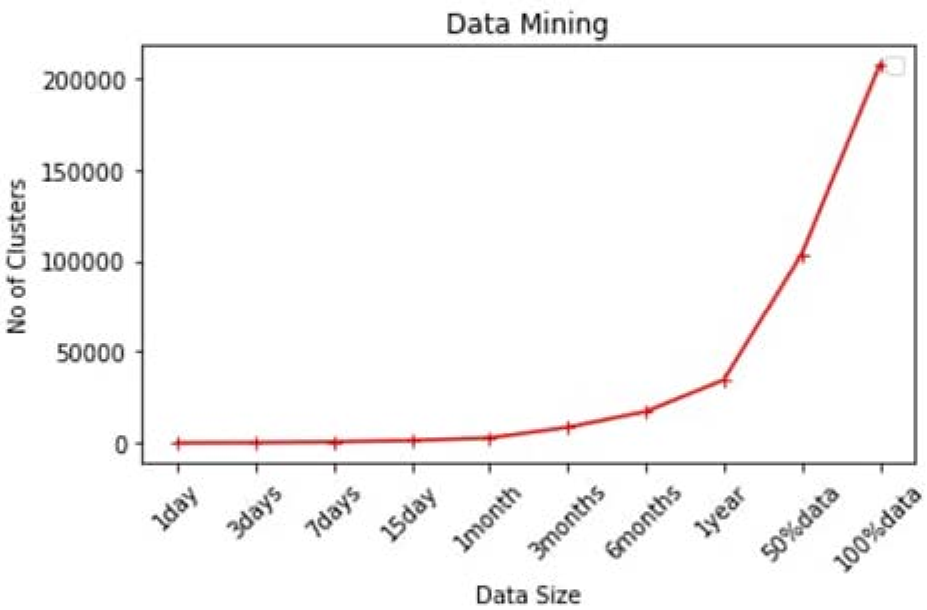


Fig. 9. Number of clusters discovered vs dataset mining.

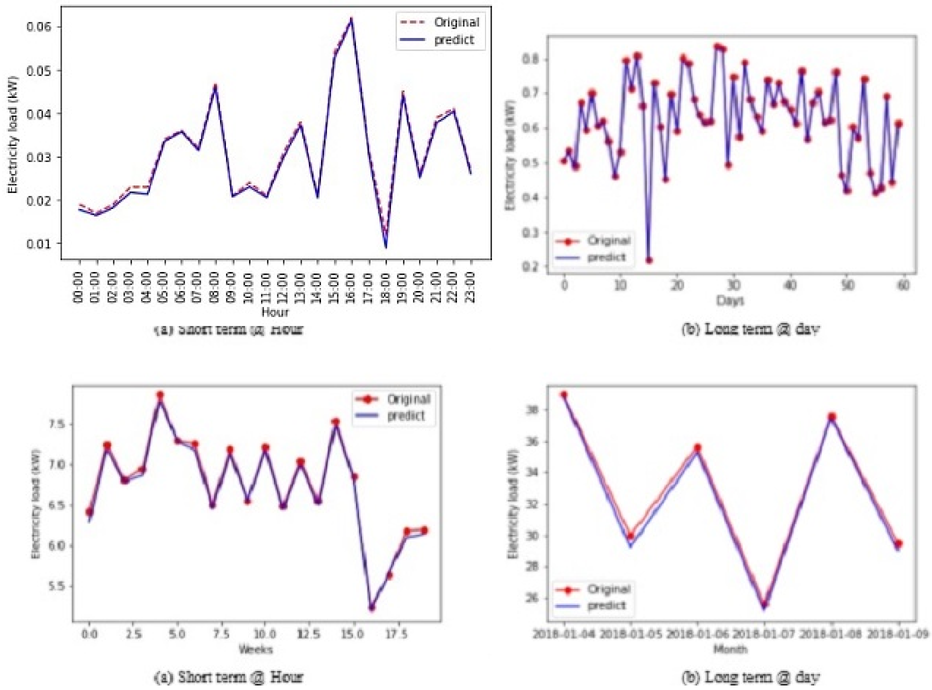
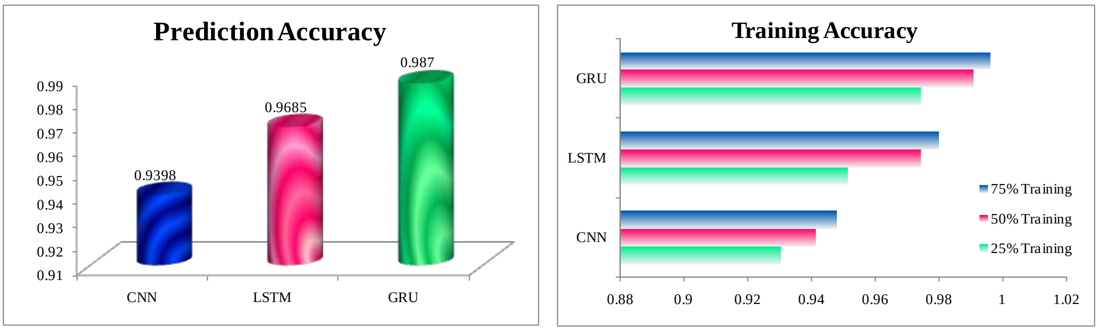


Fig. 10. Home 4 energy consumption prediction vs. actual energy

The energy prediction of house_4 in short and long term is presented and compared in **Fig. 10**. As well as the average averages of 25 %, 50% and 75% of incremental data mining approaches, **Table 6** displays the short- and long-term prediction accuracy achieved. Our suggested model will at any stage in a period be above CNN, LSTM and GRU achieve a combined precision in 97.42% (25%), 99.07% (50%) of all training, 99.61% (75%) of all. In **Fig. 11 (a)**, The gradual mining can identify varieties that are driven by the residential behavioral characteristics and encourage highly skilled efficient use of resources at various speeds. This is a general overview of the proposed model versus CNN, LSTM, and GRU. **Fig. 11 (b)** provides a correspondence of training data set 25%,50% and 75% between the proposed model and CNN, Sequential and LSTM.

Table 6. Prediction model accuracy

Training	CNN	LSTM	GRU
25% Data Training Accuracy			
	0.9302	0.9515	0.9742
50% Data Training Accuracy			
	0.9413	0.9743	0.9907
75% Data Training Accuracy			
	0.9479	0.9799	0.9961
Overall Training Accuracy			
	0.9398	0.9685	0.987



(a) Overall Prediction Accuray (b) Prediction accuracy @ data set 25, 50 and 75%

Fig. 11. Comparison of prediction accuracy

Table 7. Prediction accuracy @ household level

Long Term		R-Squared	RMSE	MAE	medAE	Accuracy
25 % Data as Training Data						
CNN	Home 1	0.488394	0.404559	0.330852	0.27485	0.9362
	Home 2	0.527916	0.37819	0.311093	0.266756	0.9241
LSTM	Home 1	0.215022	0.501121	0.421199	0.394857	0.9549
	Home 2	0.092536	0.524343	0.442557	0.420795	0.9481
GRU	Home 1	0.103503	0.535535	0.45982	0.473143	0.9745
	Home 2	0.301156	0.460141	0.38043	0.342626	0.9738
50 % Data as Training Data						
CNN	Home 1	0.265177	0.494367	0.411117	0.376486	0.9417
	Home 2	0.265528	0.490907	0.410766	0.368873	0.9408
LSTM	Home 1	0.112879	0.543187	0.462083	0.453552	0.9673
	Home 2	0.094203	0.545164	0.461476	0.4664	0.9812
GRU	Home 1	0.098269	0.547642	0.470408	0.454947	0.9887
	Home 2	0.124243	0.536048	0.451231	0.430627	0.9926
75 % Data as Training Data						
CNN	Home 1	0.237057	0.497204	0.415144	0.378056	0.9482
	Home 2	0.182011	0.503956	0.418736	0.385433	0.9476
LSTM	Home 1	0.17623	0.516644	0.433257	0.401592	0.9769
	Home 2	0.072268	0.536699	0.443885	0.428286	0.9828
GRU	Home 1	0.16701	0.519528	0.436487	0.410718	0.9985
	Home 2	0.174758	0.506186	0.41792	0.380161	0.9937

In **Table 7**, the various parameters of accuracy as R-squared, RMSE, MAE, and MedAE in each model, and compared at the current family level. Therefore, to approximate the house energy consumption excepted, the findings were applied with various system requirements. The output value reached 98.71% exactness 99.16%, 97.52%, and 96.37%, time @hour, long term @day, long-term @week, and long-term @month energy usage predictions separately. In addition, in this analysis, Jian Qi Wang, Federico Divina, Salah Bouktif and Jui-Sheng Chou, compared with the proposed LSTM and GRU models, indicates error accuracy parameters in **Table 8**. These assumptions by residents impact the option of energy usage examples that are translated into energy usage straightforwardly. Certain criteria depend on the way of life, the product of the occupant's behavioral features as well as the comparison of individual choice, are critical to the design of a successful energy environment which encourage consumers to engage in the energy use behavior of different occupants.

Table 8. Performance analysis

Reference	Year	Model	R-Squared	RMSE	MAE	medAE	Accuracy
Jian Qi Wang [41]	2020	LSTM	-	1.8118	1.4215	10.815	-
Federico Divina[36]	2019	NN	-	0.54	0.46	-	-
		Ensemble	-	0.65	0.53	-	-
		XGBoost	-	0.55	0.47	-	-
Salah Bouktif[6]	2019	ANN	-	725.89	559.63	-	-
		Random Forest	-	527.25	368.91	-	-
		XGBoost	-	440.16	311.43	-	-
Jui-Sheng Chou [12]	2018	ANN	0.556	0.092	0.057	36.83	-
		Ensemble	0.607	0.094	0.049	42.31	-
Proposed		LSTM	0.124249	0.526672	0.4386	0.414	96.85
		GRU	0.170884	0.512857	0.3954	0.037	98.7

4. Conclusions and Future Work

This research describes the influence of consumers' behavior and their particular preferences on reasons to use energy that can be taken from appliances time associated with time series for energy. These examples can promote realistic, cost-saving plans for purchasers, gracefully balance vitality and ask for advances in reservation and assignment, plan vitality purchasing arrangements and maintenance schedules, and more thoroughly coordinate the skilled formation and essential arrangements of the system. This paper also introduced incremental frequent mining and estimation models by k-means clustering. Qualitative and quantitative research using time periods like 15 minutes, 30 minutes, 1 hour, and 12 hours have revealed results that actively support our approach to constant mining. The proposed model is evaluated using genuine data sets for valuable energy scheduling. It was also noticed that the CNN, LSTM, and GRU networks were provided a deep learning ability. This approach to deep learning can be used to forecast the potential use of resources with greater detailed comparison and various models. The results show that the model RNN-GRU efficiently and steadily predicts residential electricity demand and shows the good performance compared with current standards. The variable that affects the forecast is also explained.

1) To estimate electricity, use in real residential homes, a network of RNN-GRUs of a stable rating of accuracy 98.7 % was proposed.
2) In all cases of minute, hourly, daily and weekly unit resolution, our model forecasts complex electricity consumption with the highest level of performance compared to other methods.
3) Analyzing the method proposed and finding water pump, washing machine and air conditioner variables that exhibit the greatest impact on the forecast model.

The performance metrics like RMSE and MAE respectively 0.512857 and 0.3954, the proposed model of RNN-GRU had a significant prediction limit on the other technique. Per year between 2017 and 2020 the expected efficiency of power usage rises by 9.27 percent. In this research, the residential buildings' energy use in building 4 is projected to be reasonable between 8.2 and 61.8 kW. In future, to learn and develop from significant numbers from the time series of energy big data from individual homes was proposed. It is also intended to develop in the future and learn without ceasing about big data from the timing of energy from individual homes. This enables providers to perform an online/nonstop vitality assessment and directly attract consumers after changes in uses are known by wonderful industries, which preserve their energy.

Further, the proposed model is evaluated and is trying to make it realistic by bringing in real-time big data mining of time series from multiple houses. This will assist the energy providers to develop real time online energy prediction solutions to manage better for the dynamic smart grid Techniques.

References

[1] A.A. Alola, F.V. Bekun, andS.A. Sarkodie, "Dynamic impact of trade policy, economic growth, fertility rate, renewable and non-renewable energy consumption on ecological footprint in Europe," *Science of the Total Environment*, 685, pp. 702-709, 2019. Article (CrossRef Link)

[2] K. Zhou, and S. Yang, "Understanding household energy consumption behavior: The contribution of energy big data analytics," *Renewable and Sustainable Energy Reviews*, 56, pp.810-819, 2016. Article (CrossRef Link)

[3] Q. Ding, W. Cai, C. Wang, and M. Sanwal, "The relationships between household consumption activities and energy consumption in china—an input-output analysis from the lifestyle perspective," *Applied Energy*, 207, pp.520-532, 2017. Article (CrossRef Link)

[4] F. Ziel, and R. Steinert, "Probabilistic mid-and long-term electricity price forecasting," *Renewable and Sustainable Energy Reviews*, 94, pp.251-266, 2018. Article (CrossRef Link)

[5] S. Singh, and A. Yassine, "Big data mining of energy time series for behavioral analytics and energy consumption forecasting," *Energies*, 11(2), pp.452, 2018. Article (CrossRef Link)

[6] S. Bouktif, A. Fiaz, A. Ouni, and M.A.Serhani, "Single and multi-sequence deep learning models for short- and medium-term electric load forecasting," *Energies*, 12(1), pp.149, 2019. Article (CrossRef Link)

[7] P. Wang, P, H. Zhang, Z. Qin, and G. Zhang, "A novel hybrid-Garch model based on ARIMA and SVM for PM2. 5 concentrations forecasting," *Atmospheric Pollution Research*, 8(5), pp.850-860, 2017. Article (CrossRef Link)

[8] C. Chatfield, and H. Xing, *The analysis of time series: an introduction with R*, CRC press, 2019.

[9] L. Faes, A. Porta, M. Javorka, and G.Nollo, "Efficient computation of multiscale entropy over short biomedical time series based on linear state-space models," *Complexity*, 2017. Article (CrossRef Link)

[10] S.N. Fallah, R.C. Deo, M. Shojafar, M. Conti, and S.Shamshirband, "Computational intelligence approaches for energy load forecasting in smart energy management grids: state of the art, future challenges, and research directions," *Energies*, 11(3), pp.596, 2018. Article (CrossRef Link)

[11] J. Gu, Z. Wang, J. Kuen, L. Ma, A.Shahroudy, B. Shuai, andT. Chen, "Recent advances in convolutional neural networks," *Pattern Recognition*, 77, pp.354-377, 2018. Article (CrossRef Link)

[12] Jui-Sheng Chou, and Duc-Son Tran, "Forecasting Energy Consumption Time Series using Machine Learning Techniques based on Usage Patterns of Residential Householders," *Energy*, 165, pp. 709-726, 2018. Article (CrossRef Link)

[13] Y. Liu, C. Gong, L. Yang, and Y. Chen, “DSTP-RNN: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction,” *Expert Systems with Applications*, 143, pp.113082, 2020. Article (CrossRef Link)

[14] L. Mou, P. Ghamisi, and X. Zhu, “Deep recurrent neural networks for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, 55(7), pp.3639-3655, 2017. Article (CrossRef Link)

[15] Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, “Recurrent neural networks for multivariate time series with missing values,” *Scientific Reports*, 8(1), pp.1-12, 2018. Article (CrossRef Link)

[16] T.Y. Kim, and S.B. Cho, “Predicting residential energy consumption using CNN-LSTM neural networks,” *Energy*, 182, pp.72-81, 2019. Article (CrossRef Link)

[17] A. Rahman, V. Srikumar, and A.D. Smith, “Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks,” *Applied Energy*, 212, pp.372-385, 2018. Article (CrossRef Link)

[18] M. Kraus, and S.Feuerriegel, “Forecasting remaining useful life: Interpretable deep learning approach via variational Bayesian inferences,” *Decision Support Systems*, 125, pp.113100, 2019. Article (CrossRef Link)

[19] V.K. Verma, and B. Chandra, “An application of theory of planned behavior to predict young Indian consumers' green hotel visit intention,” *Journal of Cleaner Production*, 172, pp.1152-1162, 2018. Article (CrossRef Link)

[20] Y. Wang, W. Liao, and Y. Chang, “Gated recurrent unit network-based short-term photovoltaic forecasting,” *Energies*, 11(8), pp.2163, 2018. Article (CrossRef Link)

[21] C. Ma, J.H. Menke, J. Dasenbrock, M. Braun, M. Haslbeck, and K.H. Schmid, “Evaluation of energy losses in low voltage distribution grids with high penetration of distributed generation,” *Applied Energy*, 256, pp.113907, 2019. Article (CrossRef Link)

[22] A. Yassine, S. Singh, and A. Alamri, “Mining human activity patterns from smart home big data for health care applications,” *IEEE Access*, 5, pp.13131-13141, 2017. Article (CrossRef Link)

[23] P. Huber, P. Schmieder, M. Gerber, and A. Rumsch, “Poster abstract: Is the run-time of domestic appliances predictable?,” *Computer Science-Research and Development*, 33(1-2), pp.241-243, 2018. Article (CrossRef Link)

[24] A. Harell, S. Makonin, and I.V. Bajic, “A recurrent neural network for multisensory non-intrusive load monitoring on a Raspberry Pi,” in *Proc. of IEEE MMSP*, 18, August 2018.

[25] M. Fahim, K. Fraz, and A. Sillitti, “TSI: Time Series to Imaging based Model for Detecting Anomalous Energy Consumption in Smart Buildings,” *Information Sciences*, 523, pp.1-13, 2020. Article (CrossRef Link)

[26] L.G.B. Ruiz, M.C. Pegalajar, R. Arcucci, and M. Molina-Solana, “A Time-Series Clustering Methodology for Knowledge Extraction in Energy Consumption Data,” *Expert Systems with Applications*, 160, pp.113731, 2020. Article (CrossRef Link)

[27] C.W. Song, H. Jung, and K. Chung, “Development of a medical big-data mining process using topic modeling,” *Cluster Computing*, 22(1), pp.1949-1958, 2019. Article (CrossRef Link)

[28] J. Kong, J. Han, J. Ding, H. Xia, and X. Han, “Analysis of students’ learning and psychological features by contrast frequent patterns mining on academic performance,” *Neural Computing and Applications*, 32(1), pp.205-211, 2020. Article (CrossRef Link)

[29] G. Bode, T. Schreiber, M. Baranski, and D. Müller, “A time series clustering approach for Building Automation and Control Systems,” *Applied Energy*, 238, pp.1337-1345, 2019. Article (CrossRef Link)

[30] S. Singh, and A. Yassine, “Mining energy consumption behavior patterns for households in smart grid,” *IEEE Transactions on Emerging Topics in Computing*, 7(3), pp.404-419, 2019. Article (CrossRef Link)

[31] T.T. Nguyen, P.Krishnakumari, S.C. Calvert, H.L. Vu, and H. Van Lint, “Feature extraction and clustering analysis of highway congestion,” *Transportation Research Part C: Emerging Technologies*, 100, pp.238-258, 2019. Article (CrossRef Link)

[32] Y. Zhou, H. Wu, Q. Luo, and M. Abdel-Baset, "Automatic data clustering using nature-inspired symbiotic organism search algorithm," *Knowledge-Based Systems*, 163, pp.546-557, 2019. Article (CrossRef Link)
[33] E.L. Lydia, P.K. Kumar, K. Shankar, S.K. Lakshmanaprabu, R.M. Vidhyavathi, and A. Maseleno, "Charismatic document clustering through novel K-Means non-negative matrix factorization (KNMF) algorithm using key phrase extraction," *International Journal of Parallel Programming*, 48(3), pp.496-514, 2020. Article (CrossRef Link)
[34] C. Yuan, and H. Yang, "Research on K-value selection method of K-means clustering algorithm," *J—Multidisciplinary Scientific Journal*, 2(2), pp.226-235, 2019. Article (CrossRef Link)
[35] F. Divina, M. García Torres, F.A. Goméz Vela, and J.L.VázquezNoguera, "A comparative study of time series forecasting methods for short term electric energy consumption prediction in smart buildings," *Energies*, 12(10), pp.1934, 2019. Article (CrossRef Link)
[36] H.I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.A. Muller, "Deep learning for time series classification: a review," *Data Mining and Knowledge Discovery*, 33(4), pp.917-963, 2019. Article (CrossRef Link)
[37] A.K. Ozcanli, F. Yaprakdal, and M.Baysal, "Deep learning methods and applications for electrical power systems: A comprehensive review," *International Journal of Energy Research*, 44(9), pp. 7136-7157, 2020. Article (CrossRef Link)
[38] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," *IEEE Access*, 7, pp. 53040-53065, 2019. Article (CrossRef Link)
[39] D. Gilboa, B. Chang, M. Chen, G. Yang, S.S. Schoenholz, E.H. Chi, and J. Pennington, "Dynamical isometry and a mean field theory of LSTMs and GRUs," *arXiv preprint arXiv:1901.08987*, 2019.
[40] J.Q. Wang, Y. Du, andJ. Wang, "LSTM based long-term energy consumption prediction with periodicity," *Energy*, 197, pp.117197, 2020. Article (CrossRef Link)
[41] S. Wang, J. Lian, Y. Peng, B. Hu, and H. Chen, "Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China," *Agricultural Water Management*, 221, pp.220-230, 2019. Article (CrossRef Link)

Nomenclature

CNN - Convolutional Neural Networks
RNN - Recurrent Neural Networks
LSTM - *Long Short Term Memory*
GRU - *Gated Recurrent Unit*
AR - Autoregressive
MA - Moving Average
ARIMA - AutoRegressive Integrated Moving Average
SSE - squared error sum
RMSE - Root Mean Square Error
MAE - Mean Absolute Error
RMAE - Relative Mean Absolute Error
Υ - Patterns
$c_{X \Rightarrow \Upsilon}$ - Confidence
C - centroids
d - distances
a_j - average distance
S_{d_j} - *Silhouette coefficient*
C_i - **C**luster

X - Input sequences
h_t - hidden state
W_s- Weight
b_s- Bias
C_t - current cell state
$\tilde{C}$ - New candidate values
z_t - Update gate
r_t - Reset gate
h_t - Cell states
kW – ***Kilowatt***

K.Balachander, B.E, M.E, is an Associate Professor in the Department of Computer Science and Engineering, since July 2013. He obtained his B.E (ECE) from Madurai Kamaraj University and M.E (Computer Science and Engineering) from Satyabama University, Chennai. He has been in the teaching profession for the past 7 years and has handled UG programs. His areas of interest include Data Mining, Big Data and Machine Learning. He has published 4 papers in refereed International Journals.

Dr.D.Paulraj, B.E, M.E, Ph.D., is a Professor and Head in the Department of Computer Science and Engineering, since November 2012. He obtained his B.E (CSE) from Bangalore University and M.E (Computer Science and Engineering) from Anna University, Chennai. He received his PhD from Anna University, Chennai. He has been in the teaching profession for the past 20 years and has handled UG programs. His areas of interest include Big Data and Machine Learning. He is currently guiding 5 research scholars and 5 Scholars have successfully completed their doctoral degree under his guidance. He has published 26 papers in refereed International Journals.