International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-10, Issue-5, September 2022
https://doi.org/10.55524/ijircst.2022.10.5.10
Article ID IRP1310, Pages 73--79
www.ijircst.org

Predicting Electricity Usage of Houses Incorporating Smart Meters

Musaib Amin[1], Dharminder Kumar[2], and Dr. Satish Saini[3]

[1]M. Tech Scholar, Department of Electrical Engineering, RIMT University, Mandi Gobingarh, Punjab, India
[2]Assistant Professor, Department of Electrical Engineering, RIMT University, Mandi Gobingarh, Punjab, India
[3]Professor, Department of Electrical Engineering, RIMT University, Mandi Gobindgarh, Punjab, India

Correspondence should be addressed to Musaib Amin; rathermusaib123@gmail.com

ABSTRACT- The need for smart grid user-side management is becoming more critical as smart power distribution networks evolve. This study suggests a short-term power load forecasting model based on K-means and Clustering to increase the precision of short-term electric load forecasting for individual customers. Local comparable daily data are used as characteristics for users with poor correlation at adjacent times. Based on the -means cluster analysis technique, a clustering module was created. The clustering module was validated using home load profiles based on smart meters. Utilizing daily and segmented load profiles from individual and collective smart meter data, many case studies were put into practice. The smallest segmentation time window yields the lowest clustering ratio, according to simulation findings described in terms of the correlation between the clustering ratio and the time window. Results also indicate that for strongly coupled load profiles, a limited number of clusters is advised. Last but not least, the feather vectors are sent into the BP Neural Network, which forecasts the short-term load. According to experimental findings, the suggested clustering algorithm matches the features of customers' power consumption pattern. The accuracy of clustering-based load forecasting is higher than that of non-clustering load forecasting when using the same forecasting methodology. This model offers superior predicting accuracy when compared to conventional BP, RBF, and GRNN Neural Network

KEYWORDS- K means, Clustering, Smart Meters, Distributed Networks

I. INTRODUCTION

The quantity of data that may be used to represent the energy usage of residential, small commercial, and small industrial clients has increased thanks to smart meters that take readings every 15 minutes, 30 minutes, or an hour. The operation, planning, and management of distribution networks are significantly enhanced by the conversion of smart metering data into useful information. The data of smart meters are utilized to derive such significant information as load profile classes [1] using statistical, engineering, time-series, and cluster analysis techniques.

In the literature, [1]–[7], the application of the -means cluster analysis approach to group load profiles of residential consumers was described. The prior study's objective was to cluster daily load profiles. However, this study examines how daily and segmented load profiles of unique and combined residential consumers are clustered. Profiles with a period duration of less than or equal to 24 hours are referred to as segmented load profiles. The suggested -means based clustering module's outputs may be applied to load estimation, whereby current and upcoming smart meter data are projected, or to enhance Time of Use (Tou) tariff design. This study presents the suggested -means based clustering module and analyse the simulation outcomes. Based on a thorough examination of the simulation findings, the ideal time frame for segmenting load profiles is established.

A. *Cluster Analysis Methods*

The definition of clustering is the gathering of related things. In order to make profiles inside a cluster comparable to one another, a given set of load profiles is divided into numerous clusters. The load profiles allocated to the various clusters are also as diverse as feasible. When there are more output clusters than there are input load profiles, this is referred to as clustering.

The variety of current clustering applications led to the development of several cluster analysis techniques. Data mining, pattern recognition, and clustering-based estimation are a few examples of applications for cluster analysis approaches. Hierarchical and partitional clustering methods are two primary categories for cluster analysis techniques. Through a series of layered partitions, hierarchical algorithms [8] divide a given dataset of load profiles into the necessary number of clusters. A hierarchy of partitions is created as a result, leading to the ultimate cluster (s).

On the other hand, partitional approaches seek to classify load profiles into several clusters by maximizing an objective function. The objective function that is minimized is the intra-cluster sum of squared distances. The necessary number of clusters must be predefined or known in advance when using partitional clustering. Because each cluster is represented by a matching centre, partitional cluster analysis methods are also known as centre-based approaches. A cluster's centre is frequently thought of as a concise summation of all the load profiles found there. Large and high-dimensional datasets can be clustered effectively using partitional approaches. As a result, partitional approaches are favoured for clustering residential customers' daily load patterns [9], [10].

B. Proposed Cluster Analysis Module

One of the most used partitional cluster analysis techniques is the -means approach [12]. By reducing the intra-cluster sum of squares, an iterative approach that organizes load profiles into clusters made up of half-hourly observations is used in this method.

The suggested k-means based clustering module's inputs comprise residential customers' load profiles and the maximum number of clusters. There can never be more clusters than there are input load profiles. Each load profile will have its own cluster if the allowed number of clusters is achieved. In this instance, load profiles act as the respective hubs of their clusters. Equation is used to compute the average Euclidean distance between the load profiles and cluster centre at each iteration of the -means (2). As a result, each load profile is assigned to the cluster The needed number of clusters is determined using the mean value of the root-mean-square errors (RMSE) between load profiles and their respective cluster centres. Up until the mean RMSE is below a predetermined error level, the number of clusters is incrementally increased. In equation, the error threshold is defined.

The number of clusters, cluster centre, and the assignment of load profiles to specific clusters are all examples of the clustering module's outputs. The average values of all load profiles assigned to this particular cluster, computed at each half-hourly time step, are used to define a cluster centre.

II. OBJECTIVES

The Python 2.7 clustering module was created using Pycluster [12], an open-source cluster analysis program. Performance of the suggested module was evaluated using residential load profiles based on real smart meter observations. The Irish Smart Metering [13] Customer Behaviour Trials provided the load profiles (CBT). Each smart meter's daily load profiles are made up of 48 half-hourly readings. The first measurement, taken at 12:30 a.m., shows the typical active power usage between 12:00 a.m. and 12:30 a.m. The average amount of active power used between 11:30 p.m. and 12:00 a.m. was the final measurement made at 12:00 a.m.

In the current study, load profiles from 100 household smart meters that were gathered between July 20 and August 9, 2009, were employed. These were separated into profiles for the training phase and the test period.

III. LITERATURE REVIEW

Research into identification of specific appliances such as Non-Intrusive Load Monitoring (NILM), data collection systems and protocols, smart meter control and development, data privacy and tariff development are beyond the scope of this paper. Only papers published in English are included in this review to maintain reproducibility, fully acknowledging the quality of non-English research literature. Regardless of the number of reviewers working on a review it is advisable to develop a formal protocol with evaluation criteria for inclusion, exclusion and quality appraisal to ensure consistency across the reviewers and papers. For this paper a protocol was developed for evaluating and extracting data

There are several ways to search for literature. Popular and feasible strategies are to visit multipurpose search-engines like Google, Bing etc. or visit the academic publishers' online resources and identify journals of interest, but as many journals are cross disciplinary it is not a simple task to identify relevant journals

IV. METHODOLOGY

Fig. 1 displays the load curves of five randomly selected users in a Nanjing neighbourhood in 2018. Distinct users are represented by different colour load curves. In the diagram, the load value is represented by the vertical axis, and the collection of the load value is shown by the horizontal axis. Each user provides 336 data points, which are gathered every 30 minutes.

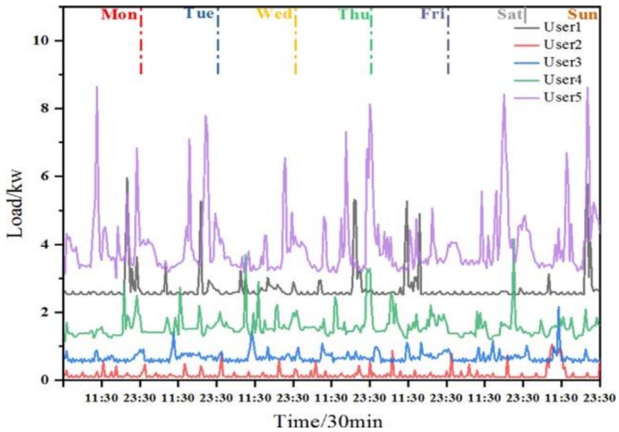

Figure 1: A weekly load curve of 5 users. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Users 2 and 3 use more power at night, as shown in Fig. 1, and their load curves are cyclical and regular with little volatility. In addition to having comparable load values at the same time on various days (daily correlation), the curves also exhibit a gentle load curve over a brief period of time (adjacent time correlation), which is typical for residential users. Users 1, 4, and 5's load curves periodically show notable peaks in the mornings and afternoons. This indicates that, like many corporate customers, their load numbers are comparable at the same time on many days (daily correlation). K-means method is employed in this study to analyse user clusters.

A common clustering algorithm is the K-means algorithm. The basic goal of the method is to continuously cluster the data into various groups until the objective function is maximized.

The precise steps are as follows.

(1) Choose K appropriate points as the initial clustering centres

(2) Calculate the value of d:

$$d = \sqrt{\sum_{i=1}(x_i - y_i)^2} \qquad (1)$$

Xi and Yi in formula (1) stand for the value of the ith variable in the sample x and the ith variable in the clustering centre y, respectively. D denotes the separation between the centre points and the other points.

(3) Classify all the points according to their nearest centre points, thus dividing all the sample points into K clusters.
(4) Calculate the centre of mass of the K clusters and update them as new clustering centre.
(5) Repeat step (2), (3) and (4), keep iterating until the clustering centres stop shifting.

The K-means technique was employed in this study to divide the users into clusters A and B; class A users' electricity consumption behaviour has both daily correlation and adjacent temporal correlation, but class B users' behaviour just has daily correlation.

A. User Side Load Forecast Based on FCM– BP

- *FCM Algorithm*

This study used the FCM (Fuzzy C-Means) to select local similar days. The procedure of FCM is as follows:

(1) Select c appropriate clusters and a weighted index m. Generate an initial cluster centre matrix V^0 randomly. The number of iteration is 0(l = 0).

(2) Calculate each cluster centre V_i^{l+1}

$$V_i^{l+1} = \frac{\sum_{k=1}^{n} (u_{ik}^l)^m x_k}{\sum_{k=1}^{n} (u_{ik}^l)^m} \quad (i = 1,2,...,c) \quad (2)$$

In formula (2), x_k represents the kth element, u_{ik} represents the kth element's membership towards the ith cluster.

(3) Update the membership matrix and calculate the value of the object function.

$$u_{ik}^{l+1} = \left[\sum_{j=1}^{c} \left(\frac{d_{ik}^l}{d_{jk}^l} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (i = 1,2,...,c; k = 1,2,...,n) \quad (3)$$

$$J^l(U^l, V^l) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^l (d_{ik}^l)^2 \quad (4)$$

In the formula above, U represents the membership matrix, d_{ik} represents the distance between the kth sample point and the ith cluster centre

(4) The given threshold $\varepsilon > 0$, if the object function satisfies the condition $\left| J^l - J^{l+1} \right| \le \varepsilon$, then the iteration is stopped. Otherwise, let $l = l + 1$, iterate again and return to step (2).

(5) When the iteration is finished, observe the membership of all the samples of matrix U and classify them into the cluster with the maximal membership.

- *Selection of Local Similar Days Based on FCM*

Users' historical load curves must first be classified in order to choose comparable days for them. The load may be split into four categories based on its characteristics: typical weekday load, typical weekend load, typical holiday load, and big holiday load (National Day or Chinese Spring Festival). The FCM algorithm cannot determine the similarity between the target vector and each sample vector, so in this study, a model is proposed that can reasonably determine the similarity between the target vector and each sample vector by amplitude and trend in order to find the local similar days of the load. The particular rule is as follows: The following definition describes the "similarity difference" between the target vector x0 and the sample vector xk::

$$D_k = \sum_{i=1} \left| u_{i,0} - u_{i,k} \right| \quad (5)$$

In formula (5), D_k is the similarity difference between the target vector x_0 and the sample vector x_k; $\mu_{i,0}$ indicates the membership degree to which target vector x_0 belongs to the ith clustering prototype; and $\mu_{i,k}$ indicates the membership degree to which target vector x_k belongs to the ith clustering prototype. The smaller the D_k, the higher the trend similarity between the target vector x_0 and the sample vector x_k.

The value of the target sequence is the first d load points of the (t+1) times to be predicted. This study select m historical loads with high similarity from recent n historical days as a local similar sequence $L_{i,t}$, where:
$L_{0,t} = \{l_{0,t-d+1}, l_{0,t-d+2}, ..., l_{0,t}\}$, $l_{0,t-d+1}$ represents the historical load at (t−d+1) times on the unpredicted day.
$L_{i,t} = \{l_{i,t-d+1}, l_{i,t-d+2}, ..., l_{i,t}\}, i = 1,2,...,m$, $l_{i,t-d+1}$ represents the historical load at (t−d+1) times on Day i.
The selection of local similar days is divided into the following four steps:

Step 1: the first d points of (t+1) time on n historical days before the unpredicted day are selected as historical load sequence set $L_{k,t}$ (k = 1, 2, ...n), and the first d points of (t+1) time on unpredicted day are selected as the target sequence $L_{0,t}$.

Step 2: the historical load sequence set $L_{k,t}$ (k = 1, 2, ...n) and the target sequence $L_{0,t}$ are selected as samples, and the FCM Algorithm is used to implement fuzzy clustering.

Step 3: the similarity difference between the target sequence $L_{0,t}$ and various historical load sequences $L_{k,t}$ of the same cluster can be obtained by using the Formula (5).

Step 4: get the value of the similarity difference, and select m historical days having the smallest similarity difference as local similar days.

- *FCM– BP Load Prediction Model*

BP Neural Network uses a three-layer network, including Input layer, Hidden layer and Output layer. Its structure is shown in Figure . 2.

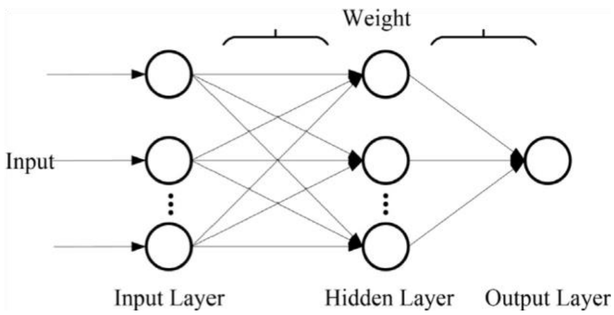

Figure 2: Structure of BP neural network

Neurons will process the input vector after receiving it from the input layer, creating the necessary weights, and transmitting them to the hidden layer. The neurons in the hidden layer then produce weights and send them to the output layer. The output layer will finally produce the output vector. The created output vector will continually review itself and compare deviance from the predicted vector until the perfect neural network is established [5].

Due to the load curves' daily correlation and adjacent temporal correlation properties, the K-means method may be used to divide users into the A and B clusters. While the load curves of Cluster B users only include daily correlation features, the load curves of Cluster A users contain both daily correlation and adjacent time correlation. Selecting comparable days is done using the FCM algorithm. Users of Cluster A can apply the approach outlined in Section 3.2 to determine the local comparable sequence of (d+1) occurrences of the unpredicted day. The preceding d points of the historical load value for L0,t and the specified local comparable sequence are used as the user information set.. For Cluster B users, FCM algorithm is used to calculate the local similar sequence with the minimum similarity difference as the user information set. Finally, the sets are used as input data for BP Neural Network, which is utilized to forecast the short-term load. The specific process is shown in Figure. 3.

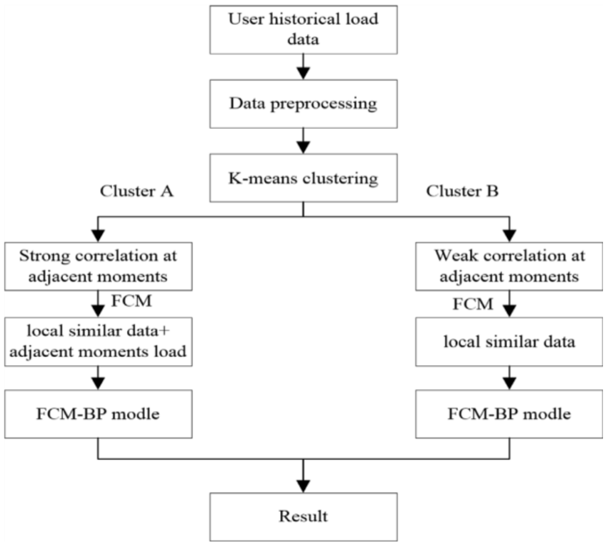

Figure 3: Flow chart of short-term load forecasting model based on K-means and FCM–BP

By processing the local comparable days' load sample data using Formula (6), it is possible to determine the load rate-of-change for each group.

$$\delta_{j,t+1} = \frac{L_{j,t+1} - L_{j,t}}{L_{j,t}} \qquad (6)$$

Both the load value at adjacent periods and the load rate-of-change are used as BP training samples for users in Cluster A, but only the load rate-of-change is used for users in Cluster B. Finally, Formula may be used to calculate the load data at the unexpected time (7).

$$L_{k,t+1} = (1 + \bar{\delta}_{k,t+1}) L_{k,t} \qquad (7)$$

where h is the load rate-of-change at (t+1) times of the unpredicted day, which is calculated by Formula (8):

$$\bar{\delta}_{k,t+1} = f(\delta_{j,t+1}, \delta_{j-1,t+1}, \ldots \delta_{j-n,t+1}) \qquad (8)$$

B. Example simulation

1) Data Description and Pre-Processing

The load statistics of 200 users from a region of Nanjing in 2018 were the data set used in this investigation. Every 30 minutes, one data point is taken, and a daily data set of 48 load points is available. The samples that were gathered for this investigation are first prepped. This study uses the horizontal smoothing approach to handle the distorted or missing abnormal data, replacing the abnormal value with the average value of the first 16 neighbouring data points of the abnormal value. Normalize all load data following the analysis of the aberrant data. The following is the normalizing formula:

$$x_i' = \frac{x_{max} - x_i}{x_{max} - x_{min}} \qquad (9)$$

In formula (9), x_i represents the sample data, x_i' represents the normalized sample data, x_{max} represents the maximum value in the sample data, and x_{min} represents the minimum value in the sample data.

2) Evaluation Index

In order to effectively evaluate the accuracy of the model and compare it with other algorithms, MAPE (Mean Absolute Percentage Error) is adopted as the model evaluation index. The formula is as follows:

$$MAPE = \frac{1}{N} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| * 100\% \qquad (10)$$

In the formula, n represents the number of points to be predicted, y_i represents the real load data of the ith point to be predicted, and $\hat{y}_i$ represents the predicted load data of the ith point

V. SYSTEM IMPLEMENTATION

A. Data Collection

In this study, smart meter data from the dataset was utilised [11]. An anonymized dataset of smart meter data from the "Electricity Smart Metering Customer Behaviour Trials" was made available to the public by the Commission for Energy Regulation [1]. The residences and businesses in this dataset, which includes more than 5000 throughout Ireland, are sampled every 30 minutes.

The information was gathered between July 14, 2009, and December 31, 2010. This data was divided into a training, validation, and testing set for cross-validation purposes. The first 9 months of data were divided into three sets: the training set, which was used to train the models; the validation set, which was used to tweak the hyperparameters; and the test set, which contained the remaining 6 months. These divides were designed to give the models a chance to understand the periodicity present in a year by balancing the training data with the test data.

We used a sample of 709 distinct Irish residences because these algorithms require lengthy training periods. However, we think our method would work well in practice given how rarely it is necessary to train these models.

Four daily load patterns for residential customers are shown in Figure 1. As can be observed, customers 1 and 2 have load profiles that are comparable, whereas customers

3 and 4 have quite distinct load profiles. This shows that, even though individual power usage varies, comparable consumers may be grouped together based on their load patterns.

B. Clustering

We suggest that to increase the accuracy of the models, comparable customer load profiles should be clustered together, along with the power usage of each cluster. This is because aggregated clustered loads reduce the stochasticity of the system's demand, improving the models' capacity for prediction.

Hierarchical clustering was used to evaluate the Euclidean distance and wavelet metrics [21]. However, k-means proved to be the most reliable and effective algorithm, and as a result, it was selected for use in this article [12].

Each user's nine-month power use in the training set was condensed into a single averaged 24-hour load profile in order to cluster the data (48 data points per customer). This was accomplished by averaging the electricity used over the day's half-hour intervals (eg. taking the mean of every 12-12:30pm point in the training set). The distinction between weekends and weekdays for clustering was not considered. The data was then adjusted such that families of various sizes that had comparable use habits were grouped together.

The test set was utilized for cross-validation in order to determine the ideal number of clusters (k). As a consequence, we were able to evaluate the outcomes of each model and choose the one with the highest mean absolute percentage error (MAPE) and mean absolute scaled error (MASE). Because the error did not vary much past seven clusters, k in this research was varied between 1 and 7. For each cluster, many models were fitted and used to forecast the testing results.

The most accurate division was selected after 1000 iterations of the k-means method were completed to overcome local minima [18].Figure 4 show daily load profiles for four different customers on the 22nd of July 2009

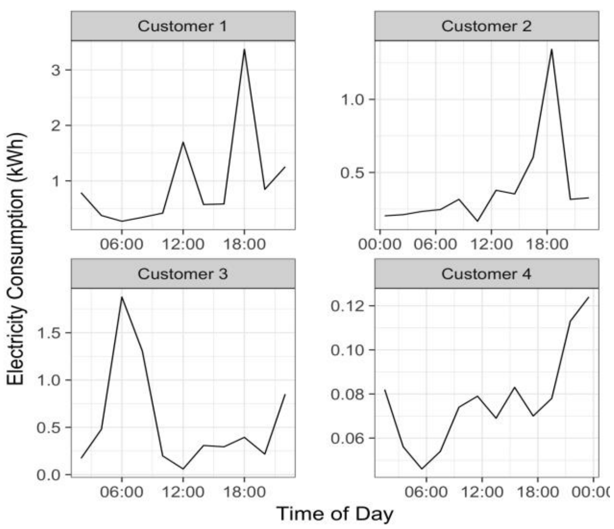

Figure 4: Figure showing daily load profiles for four different customers on the 22nd of July 2009

The total amount of power used by each cluster is determined after assigning each client to its corresponding cluster. The power used at each time interval per cluster is added to achieve this. The system is partially loaded as a result. Each of the various component system loads is fed into a model, and the output forecasts are then combined to provide the forecast for the entire system. The MAPE and MASE for each model are then calculated to assess this forecast.

C. Feature Selection

1) Calendar Attributes

Because of the power consumption's daily, weekly, and yearly periodicity, daily calendar properties were thought to be crucial for effectively modelling the issue. Included on the calendar are the hour of the day, day of the week, month, and public holidays.

Due to the difference in power usage displayed on these days, public holidays were employed as characteristics for the model.

In the findings section, we assess the performance improvement brought on by the modelling of calendar features.

2) Time Series Data.

Lagged data inputs were used to simulate the time-series component. The preceding three hours, the comparable three hours from the day before, and the similar three hours from the week before were all used to accomplish this.

Neural networks with long short-term memory retain values across indefinite time scales. In order to test the hypothesis that the long short-term memory network can recall short-term memory over a long period of time, five delayed inputs from the preceding two and a half hours were utilized as features.

3) Data Representation.

The entry of the lagged data was done using numerical representations. Public holidays were handled with just one binary. One hot encoding is a technique that separates categorical variables from ordinal data. The day of the week and month of the year were encoded using one hot encoding.

D. Experiments

The experiments performed to create the models are examined in this section. Each model's validation set was subjected to cross-validation in order to fine-tune the hyperparameters. The variance of the findings was then investigated by creating each model five times for each cluster.

1) Support Vector Regression

The SVR model's hyperparameters were fine-tuned and its kernel was chosen using the validation set. The radial basis function (RBF), the linear kernel, and the polynomial kernel were all compared [4, 31]. They were picked because of their acceptance, assistance, and computation speed.

In Table 1, Prediction Accuracy Based on Type of Kernel. The final model was chosen because it delivered the best results, the linear kernel.

2) Random Forest.

Utilizing the validation set, the number of variables sampled as candidates at each split was adjusted.

23 was the ideal number for this. The 21 delayed inputs, which are essential to understanding the underlying structure of the time series, are said to be the reason why the number 23 was shown to be optimal.

Table 1: Prediction Accuracy Based on Type of Kernel

Kernel Type	Kernel Parameters	RMSE
Linear	No values	0.02102779
RBF	C=2, γ = 0.016	0.02444950
Polynomial	C=2, d = 2,r = 2	0.03145719

3) Artificial Neural Network.

The first step when creating an Artificial Neural Network is to design the architecture. In our case, the number of input neurons is set to 43 only one output neuron is required, due to the fact that we are only forecasting one step (30 minutes) ahead.
To design the number of hidden layers the Levenberg-Marquardt technique was used. An optimal architecture with three hidden layers was obtained. The first layer contained two neurons, the second contained five, and the third contained four.

4) LSTM.

The Levenberg-Marquardt techniques was once again used to select number of layers and number of memory units. Using this technique, the optimum number of layers was found to be 2 with 50 memory units each.

VI. SIMULATION AND RESULTS ACORN_GROUPED

The following notebook will explore/experiment with K-mean feature clustering on weather data. This was a small test project of mine that I wanted to upload. Table 2 shows The std and TOU values on LC Lid
Outside conditions are one of the leading variables responsible for energy consumption. However, they are often large in number and increase dimensions for the predictive ML models. To solve this issue with a minimal loss of information, I have decided to experiment a bit with K-mean clustering Table 3 shows Energy consumption on Normalised and LC lid values

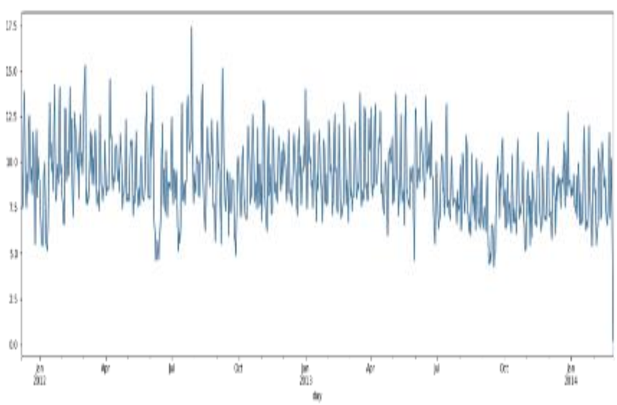

Figure 5: Electricity consumption prediction of smart meters for the years

Table 2: The std and TOU values on LC Lid

	LCLid	stdorToU	Acorn	Acorn_grouped	file
0	MAC005492	ToU	ACORN-	ACORN-	block_0
1	MAC001074	ToU	ACORN-	ACORN-	block_0
2	MAC000002	Std	ACORN-A	Affluent	block_0
3	MAC003613	Std	ACORN-A	Affluent	block_0
4	MAC003597	Std	ACORN-A	Affluent	block_0

Table 3: Energy consumption on Normalised and LC lid values

Acorn_grouped	energy_sum	LCLid	normalised
Adversity	5.373101e+06	633152	8.486273
Affluent	8.709586e+06	758448	11.483432
Comfortable	5.332296e+06	531420	10.034052

Figure 6 Shows Showing Affluent , Adversity and comfortable values of the models
Figure 7 shows the load values of the model

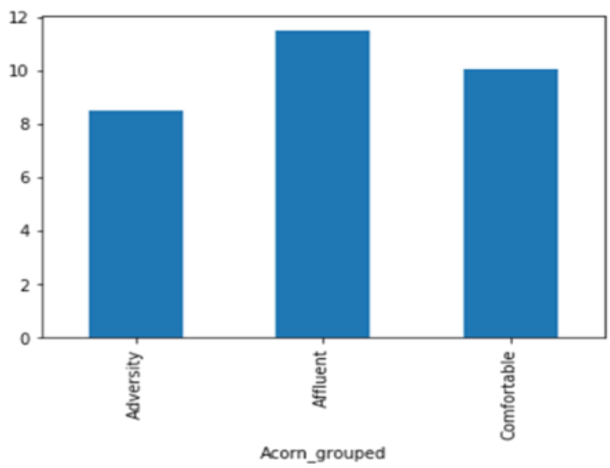

Figure 6: Showing Affluent, Adversity and comfortable values comparison of the models

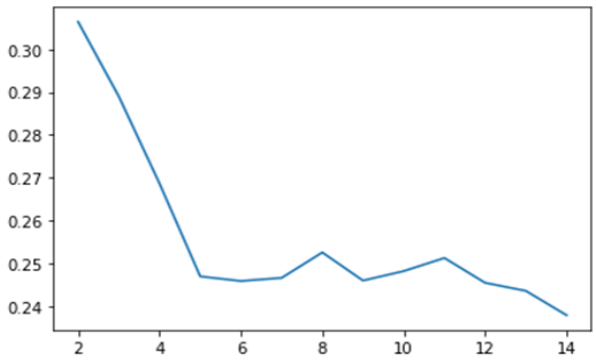

Figure 7: Showing load with time

VII. CONCLUSION

. This study examined the features of user power use in order to better correctly estimate the user-level load. Users were then clustered into two groups using the K-means algorithm. After clustering, the user's distinctive sequence is extracted using FCM. Input of the user information set and forecasting by the BP neural network come to a close. The experimental findings show that the suggested technique may significantly increase the prediction accuracy of the model, as shown by the fact that the prediction accuracy of the model in this study is greater than that of RBF, GRNN, BP neural network, and Unclustered FCM-BP model. . It is challenging to increase the model accuracy further since the existing model only takes the time-series correlation of the load into account and incorporates few external parameters. Future work will involve mining the vast user load data available in the smart grid, optimizing the clustering method, and further examining the impact of economic and climatic variables on the precision of the load forecasting model in order to create a user-level high-resolution and high-precision load forecasting model.

REFERENCES

[1] The Irish Electricity Smart Metering Customer Behaviour Trials. (2011). 21-January-2018].

[2] Hirotugu Akaike. A New Look at the Statistical Model Identification. IEEE Trans. Automat. Control 19, 6 (1974), 716–723.

[3] M.S. Al-Musaylh, R.C. Deo, J.F. Adamowski, and Y. Li. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia. Advanced Engineering Informatics 35, November 2017 (2018), 1–16.

[4] Yin-Wen Chang, Cho-Jui Hsieh, Kai-Wei Chang, Michael Ringgaard, and ChihJen Lin. 2010. Training and Testing Low-degree Polynomial Data Mappings via Linear SVM. Journal of Machine Learning Research 11 (2010), 1471–1490.

[5] C.S. Chen, Y.M. Tzeng, and J.C. Hwang. 1996. The application of artificial neural networks to substation load forecasting. Electric Power Systems Research 38, 1996 (1996), 153–160

[6] Cédric Clastres. Smart grids: Another step towards competition, energy security and climate change objectives. Energy Policy 39, 9 (2011), 5399–5408.

[7] Soma Shekara Sreenadh Reddy Depuru, Lingfeng Wang, and Vijay Devabhaktuni. Smart meters for power grid: Challenges, issues, advantages and status. Renewable and Sustainable Energy Reviews 15, 6 (2011), 2736–2742

[8] T S Dillon, S Sestito, and S Leung. Short term load forecasting using an adaptive neural network. Electrical Power & Energy Systems (1991), 186–192.

[9] Harris Drucker, Chris J C Burges, Linda Kaufman, Alex Smola, and Vladimir Vapnik. 1997. Support vector regression machines. Advances in Neural Information Processing Dystems 1 (1997), 155–161.

[10] Abdollah Kavousi Fard and Mohammad-Reza Akbari-Zadeh. 2014. A hybrid method based on wavelet, ANN and ARIMA model for short-term load forecasting. Journal of Experimental & Theoretical Artificial Intelligence 26, 2 (2014), 167–182.

[11] Commission for Energy Regulation (CER). 2012. CER Smart Metering Project Electricity Customer Behaviour Trial, 2009-2010 (2012).www.ucd.ie/issda/ CER electricity