Turkish Online Journal of Qualitative Inquiry (TOJQI)

Volume 12, Issue 5, June 2021: 1247– 1256

Research Article

A Dynamic and Unique Approach for Energy Consumption Prediction In High Raised Buildings

[1]Ms.R. Prabavathi, [2]Dr.J. ShinyDuela, [3] Mr. S.Senthil Pandi, [4]Ms.Saranya

Abstract

Prediction of power consumption is a crucial chore in vitality conservation. Because of backing vector-basedregression has a good efficiency in coping with non-linear knowledge regression problems, lately it frequently was used to foretell constructing vitality consumption. Primarily according to the historic knowledge its concludedthat the connection between lighting power consumption and its influencing elements is non-linear. The prediction of vitality consumption is a vital activity for power buying and selling companies. The prediction ought to be as correct as doable because the accuracy of the prediction interprets straight into the company's profit. Electrical hundreds and vitality consumption forecasting are a number of the most vital duties in energy scheme operation and planning. However, in some cases, we have to resolve this drawback within the absence of dependable and enough historic data. To develop a correct prediction mannequin of the lighting power consumption, a support vector regression is used along with a radial basis function. The forecast outcomes point out that the prediction accuracy of support vector regression is larger than neural networks. The prediction model can forecast the constructing hourly power consumption and check the impression of workplace constructing vitality administration plans.

Keywords: LSTM(long-short term memory),non-linear data regression problems Prediction, support vector regression, radial basis function, neural networks

[1]Department of Information Technology
[2,3,4]Department of Computer Science and Engineering
[1]Sri Sairam Institute of Technology, West Tambaram, Chennai-44
[2]SRM Institute of Science and Technology, Ramapuram Campus, Chennai-89
[3]Mohamed Sathak A J College of Engneering, Siruseri Chennai
[4]Kongu Engineering College, Erode
Email:,[1]prabavathi.raman@gmail.com,[2]shiny.duela@gmail.com[3]mailtosenthil.ks@gmail.com
[4]mohanasaranya.cse@kongu.edu

I INTRODUCTION

One of the main subjects of energy systems is the use of energy. Following the oil shortage of the 1970s, energy consumption came under consideration [1]. It is also seen that worldwide consumption of energy is steadily increasing [2]. Every country is therefore attempting to use as little energy as possible in its own country in various ways, from manufacturing to farming, from production processes to vehicles [3]. Since energy comes from two different sources, such as fossil fuels and renewable sources, It takes too much time to keep track of their energy consumption and nuclear resources [4].

Species in diverse regions. However, we can estimate the sum of energy by doing so, which is consumed in numerous fields and attempted to create arrangements, specialized for a particular need and location. For any of the above energy types, calculating the utilization of energy is useful for decision-making and strategy creators. They will be able to tell how much energy is used for their process or job by learning how much energy consider any improvements in them in order to minimize the volume of energy consumption. Predicting potential energy demand will also allow us to realize, both in the short and long term, which form of energy is mostly used and attempted to shift the pattern, as is the case with fossil fuels in recent years and now, we have green energies. Different variables affect the amount of electricity consumed in various regions. Factors like water, wind, temperature, for example. With several variables, the prediction of energy consumption is a difficult issue[5].short-term load forecasting (STLF) is relevant for energy scheme management and scheduling.Short-horizon load forecasts (from minutes to days) are important for power companies to assemble decisions that are related to the planning of electrical energy production and transmissions, reminiscent of unit dedication, dispatch of technology, hydro scheduling, organization of hydrothermal, allocation of spinning reserves, trade, and assessment of low circulation.Electrical energy markets additionally require exact load forecasts as a result of the load is the fundamental driver of electrical energy prices. The forecast accuracy interprets to the monetary efficiency of vitality corporations (suppliers, scheme operators) and different market individuals stand monetary establishments working in power markets. Neural networks are broadly utilized in STLF as a result of their flexibility which may replicate course of variability in an unsure dynamic atmosphere and complicated relationships between variables. Climate conditions (temperature, wind speed, cloud cover, humidity, precipitation), time, demographics, economy, electrical energy rates, and various components related to geographical conditions, customer styles, and behaviors are the key drivers of the load capacity organization.Often the relationships between explanatory variables and the load of the energy scheme are uncertain and inconsistent over time.Our work focus on univariate forecasting methods, in which only historical selection of load time is used to forecast the long-term values of the load of the energy entity.The sequence of load time involves a pattern, a number of seasonal variations and a stochastic irregular portion, the position of the hourly electrical load of the Polish energy system, annual, weekly and every day cycles may well be observed.In STLF, we focus on the profile of each day which depends on the day of the 7-day period and the season of the year.Moreover, it might alter over the years. The blare degree in a load time collection is dependent upon the entity measurement and buyer structure. A development amplitude of the annually cycle, weekly and every day profiles, and blast depth may alter a lot from dataset to dataset. The present entity supplies options which provides vital consumption prediction downside of residential buildings. First, to eliminate the noise, outliers and lack of values from the data, it is essential to preprocess the information that is acquired from the databases, which integrates the characteristics of residential buildings, environment information and the consumption of vitality.The sensors used to create measurements for power consumption which are multi-sourced and asynchronous, and during the long-term operation phase, the measurement and management agency can experience community fluctuations or community interruptions, leading to some missing and abnormal values.After the characterization, the options of all of the information sorts are divided into the identical standard. For the reason that function dimension will likely be elevated later developing the fundamental characteristic vector, the coaching course of for the standard classifier is time-consuming, and they are immaterial options which have minute or no affect will have an effect on the experimental results. High-dimensional attributes of function and ample characteristic information describe the information we used in this already.A single function engineering algorithm is notprecise and therefore a function engineering algorithm combines RF-PCA-SVD with random forests (RFs) with key element assessment (PCA) and single cost decomposition (SVD).

Feature choice is vital in function engineering, which may straight decide the outcomes of mannequin training. Random forests (RFs) research algorithm frameworks as an ensemble.The choice tree is managed in RFs since the minimal unit and the nodes are picked randomly from the function house because the nodes split up. A bagging algorithm is used to put together a call tree for a number of coaching sets. RFs are solely relevant to the classification drawback as a worldwide characteristic choice method. RFs undertake the strategy of quantifying the significance of options to pick out the options with the biggest number of information.

II RELATED WORK

Entity Complexities related to modeling of residential electrical energy consumption. In residential buildings, electrical energy consumption accounts for more than 30% of the overall electrical energy consumption and is thus a significant source of CO_2 emissions.The increasing desire to minimize carbon emissions and boost the energy efficiency of residential buildings is becoming imperative in the event of a right mannequin for electrical energy consumption. Even though countless vitality fashions for residential buildings are already accessible to forecast power consumption. It's easy as a,b,c to estimate their energy consumption and High-resolution residential electrical energy [6].

Optimization in Clustering evaluation of Residential electrical energy Consumption habits is used to enhance the effectivity of electrical energy consumption conduct analysis, cluster evaluation is needed [7]. A clustering optimization technique is proposed which aims to optimize the disadvantage of clustering evaluation of residential electrical energy consumption under the context of massive distribution network information.Firstly, the customer's every day electrical energy consumption curve is chosen primarily based on the adjusted cosine similarity and the standard buyer electrical energy consumption curve is fashioned by the Monte Carlo methodology to try not to chooseincorrect electrical energy consumption curve which ends up in unpredictable results. It then uses the enhanced fuzzy C-means (FCM) clustering algorithm based primarily on optimized density peaks that can effectively select cluster facilities and cluster information on electrical energy consumption.This technique overwhelms the issue of the usual FCM clustering algorithm which has random preliminary clustering facilities resulting in issues resembling being trapped into an area optimum.

In order to enhance the effectivity of electrical energy consumption conduct analysis, cluster evaluation is required vitality consumption predictions for residential buildings play a crucial function within the vitality administration and management system, as the availability and demand of power expertise dynamic and seasonal changes. on this paper, month-to-month electrical energy consumption scores are exactly categorized based mostly on open information in a complete region, which incorporates over 16 000 residential buildings. First, knowledge mining methods are used to find and summarize the electrical energy utilization patterns hidden within the data. Second, the particle swarm optimization-K-means algorithm is utilized to the clustering analysis. The benefits had been to foretell Lagrange multiplier and own the best confidence [8].

Wrapper function choice based mostly on a number of logistic regression fashions for determinants evaluation of residential electrical energy consumption supplies helpful steering for the residential buyer to attain the discount of electrical energy consumption Growing data on the determinants of the residential electrical energy consumption stage (RECL) would encourage improved vitality performance and discounted electrical energy consumption.As a result of numerousness, complexity and a number of correlations amongst influence components (IFs) of RECL, characteristic choice is an important step to make sure the precision and stability of an explanatory model. However, the present linear characteristic deciding on fashions similar to stepwise regression is extra prone to take away the determinants which have a strong-nonlinear organization with RECL, which is attributable to the restricted skill to merely seize the linear relationship between factors. To address this problem, the choice of wrapper characteristics that combines a genetic algorithm (GA) and a number of logistic regression (MLR) based mostly classifiers is suggested on this money to look for the optimum subset of functions (FS) [9].

III. SYSTEM MODEL

Not just some day-by-day consumption profiles appear repeatedly in electrical energy load data, but also some specific combos of profiles occur periodically even though each person has distinctive profile patterns and their repeating periods.This proposed scheme takes advantage of customers' day-to-day consumption profiles and their periodic consumption patterns.To prepare a LSTM to memorize the patterns, the consumption load knowledge is reworked right into a sequence of pre-defined profile.

The consumption load data is remodeled into a pre-defined profile sequence to train and train an LSTM to memorize the patterns.In the course of the study process, embedding vectors are realized based on the appearances of the profiles within the sequence, which hint at the characteristics and relationships of the profiles.

The LSTM recognizes the prevailing sample within the profile sequence and generates expected future consumption based primarily on the discovered embedding vectors which is depicted in figure1.

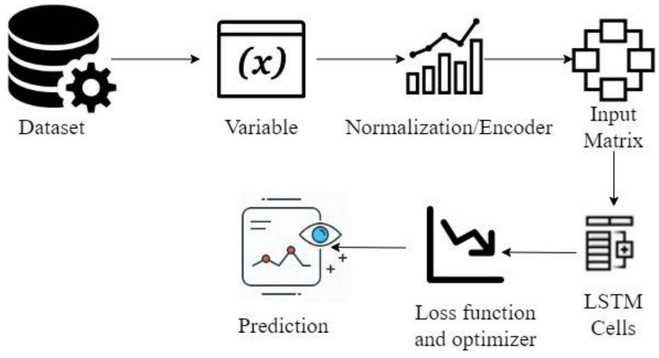

Fig.1 System Model

By using daily foundation consumption profile sequences with an LSTM network, this technique proposes a short-term electrical energy load sample forecast methodology for predicting load quality.The profile signifies that the kind of day by day electrical energy load shapes. The streaming load knowledge is segmented into each day basis, and every day consumption sample are reworked right into a consultant profile. As a substitute for the 60 second or hour-based load knowledge, the LSTM takes the profile sequences for its entry and performs a month-ahead consumption sample profile prediction.Therefore, the proposed methodology focuses on predicting the patterns already in each day load profile sequences as opposed to predicting the load quantity at a particular level of time.

IV PROPOSED SYSTEM

A. Data Analysis

Exploratory information evaluation is an information evaluation technique to investigate knowledge and discover the inherent law based mostly on the precise distribution of data. Exploratory Data Analysis (EDA) utilizing visible strategies to find the construction contained within the data. visible knowledge evaluation strategies in use in a wide selection will be traced backed to countless centuries ago, it's as a result of human eyes and brains possess the robust structural capacity to detect that occupy such a crucial place in knowledge exploring. And visible evaluation is to play quite a lot of human fashions within the processing capability of a particular option to show data.

Analysts at all times do Exploratory knowledge evaluation for information fist, then are sure to choose the mode of construction amount or stochastic quantity; Exploratory knowledge evaluation can also illustrate the sudden deviation which the frequent mannequin cannot. The important thing level of Exploratory knowledge evaluation isn't solely versatile applies to the info construction but in addition a versatile response to the revealed mode of the later analysis-step.

B. Data Preprocessing

Preprocessing is a significant errand and basic advance in Text mining, Natural Language Processing (NLP) and data recovery (IR). In the region of Text Mining, information preprocessing utilized for separating fascinating and non-trifling and information from unstructured content information. Data Retrieval (IR) is simply a matter of selecting which records in an assortment should be retrieved to satisfy the data requirement of a client.The data requirement of the customer is spoken to by an inquiry or profile and includes at least one pursuit conditions, in addition to certain additional data, such as the heaviness of the words.Thus, the recovery choice is made by contrasting the conditions of the question and the list terms (significant words or expressions) showing up in the record itself. The choice might be twofold (recover/reject), or it might include assessing the level of significance that the archive needs to question. Lamentably, the words that show up in records and in inquiries frequently have numerous auxiliary variations. So before the data recovery from the reports, the information preprocessing strategies are applied to the objective informational collection to diminish the size of the informational index which will expand the adequacy of the IR System.

The primary thought behind lemmatization is to diminish sparsity, as various curved types of a similar lemma may happen rarely (or not in any way) during preparing. In any case, this may come at the expense of disregarding significant syntactic subtleties. The importance of these multiword articulations is regularly scarcely recognizable from their individual tokens. Therefore, treating multi-words as single units may prompt better preparing of a given model. Along these lines, word installing toolboxs, for example, Word2vec propose measurable methodologies for extricating these multi-words or legitimately incorporate multi-words alongside single words in their pre-prepared implanting spaces

C. Feature Extraction

Feature selection can be defined on the basis of the importance of characteristics as a cycle of selecting the subset from the first feature set[11,12].There are three classifications of feature selection techniques: coverings, channels and implanted strategies. The extraction of features can be described as a period of separating a bunch of new features from the collection of features that are generated in the selection stage of the feature.

The feature selection technique of Ambiguity Measure (AM) will relegate a higher score to the features that appear consistently in only a single classification. AM score is determined for each feature. This strategy relegates score near 1 if the feature is unambiguous; else, it allocates score near 0. There is one edge and depending on this edge, the AM score features are divided below that limit and the AM score features above that edge are used for the learning level.The data increase of a feature is the count of the distinction of entropy whether it shows up in the content. The bigger data gain, the more noteworthy commitment the attributes to the content. Qualities with high data increase will be chosen as feature.

D. Prediction

A recurrent neural network is the Long Short-Term Memory Network, or LSTM Network, which is prepared using Backpropagation Over Time and defeats the problem of evaporating inclination.All things considered, it can very well be used to build large recurring networks that can then be used to solve problematic grouping problems in AI and achieve best in class outcomes.LSTM networks have memory impediments that are linked by layers instead of neurons.A square has segments that make it more astute than an old-style neuron and a memory for ongoing arrangements. A square contains entryways that deal with the square's state and yield. A square operates on an info grouping and the sigmoid initiation units are used for each entryway within a square to monitor whether they are set off, rolling out the state enhancement and data expansion going through the square dependent.

The neural network[13,14,15] with Long Short-Term Memory is a type of recurrent neural network (RNN).RNNs utilize past time functions to advise the later ones. For instance, to characterize what sort of function is going on in a film, the model requirements to utilize data about past functions. RNNs function admirably if the issue requires just ongoing data to play out the current assignment. On the off chance that the issue requires long term conditions, RNN would battle to show it. The LSTM was intended to learn long term conditions. It recollects the data for long periods.

V RESULTS AND DISCUSSION

For utilization of the long-term load gauging, an introduced model has been utilized to conjecture load interest for a time of a long time from 2011 to 2015 on the ISO New England dataset. Notice that the calculations were made with an hourly target, which is one of the key features of the model proposed. This has permitted the model to accomplish high precision and the determined burden information for a time of five years can serve a few advantages to a service organization for long term arranging and speculations request to additionally comprehend the essentialness of hourly expectations in long term figures, the month to month MAPE estimations of the year 2011. As portrayed in the plot, for lion's share of the months the MAPE esteem is underneath 5% and the greatest incentive in the year viable doesn't surpass 8%. For the exact year, season-wise mistake esteems have been recorded in Table 1. The watched mistake sizes show exceptionally high precision considering that heap has been determined for a time of five years on an hourly premise.

The MSE is used to measure the square of errors ie, it gives the difference between the predicted values and actual values.

$$\text{MSE} = \frac{1}{n}\sum (y - \hat{y})^2 \quad (1)$$

MAE(Mean square error) measures the magnitude of predicted errors which gives clear interpretations

$$\text{MAE} = \frac{1}{n}\sum_1^n |(y - \hat{y})| \quad (2)$$

MAPE (Mean absolute percentage error) is the measure of prediction accuracy for time series

$$\text{MAPE} = \frac{1}{n}\sum_1^n \left| \frac{(y - \hat{y})}{y} \right| \quad (3)$$

The MAPE is calculated for different season. The Support vector machine with LSTM is the best approach used in calculating the energy consumption on high raised buildings. For MSE and RMSE, Support vector regression approach isbetter than Linear Regression, Bi-LSTM and CNN-LSTM. The values got for EECP-CBL approach is 0.298 and 0.546 of MSE and RMSE while CNN-LSTM, LSTM and Linear Regression have obtained the following values (0.355, 0.596), (0.515, 0.717) and (0.425, 0.652) of MSE and RMSE respectively [10].Table 1 shows the following details of the performance measures of different methods for hourly dataset and comparative analysis of various models with training and predicting time.

Table 1: Performance measures of the models for hour-baseddataset [10]

S.No	Models	MSE	MAE	MAPE	Training time(s)	Predicting time (s)
1	LR(Linear Regression)	0.427	0.502	83.74	692.12	2.88
2	LSTM	0.515	0.526	44.37	2281.50	5.95

3	CNN-LSTM	0.355	0.332	32.83	820.72s	2.41
4	EECP-CBL	0.289	0.392	50.09	1296.34	1.87

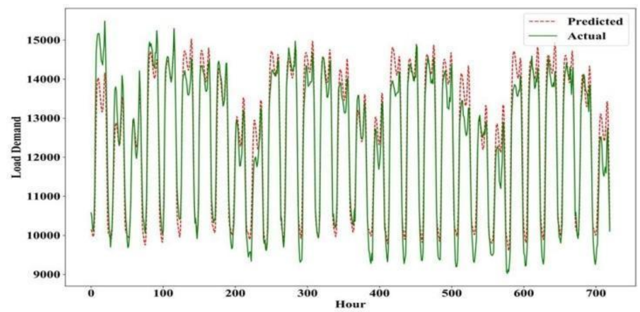

Fig.2: Predicted consumption vs Actual consumption

The table 2 shows the measures of the hourly dataset season wise MAPE.

Table 2. Season-Wise MAE and MAPE

Season	**MAE**	**MAPE**
Summer	170.75	6.12
Fall	120.25	3.48
Winter	140.86	4.65
Spring	119.37	3.37

To completely legitimize the exhibition of any AI model, a plot of anticipated qualities versus genuine qualities helps in giving an away from of the precision of a model. In such manner, the anticipated versus real burden bends for different timeframes have been plotted in Figure 2. In particular, Figure 2 portrays the varieties in the year 2011. From the period somewhere in the range of 2018 and 2019, a month, a week and a day have been selected indiscriminately and the presentation conspiracies have been outlined. Generally, the plots show how the anticipated qualities intently follow the real burden interest and the high precision of the proposed model for long haul load gauging. It should be stated that the model was prepared using a 12 GB NVIDIA Graphics Processing Unit (GPU) and a calculation time of approximately 30 minutes was required.

VI CONCLUSION

The proposed LSTM-Regression based model is prepared and tried on a benchmark dataset that contained power utilization information for various types of structures in America with a one-hour goal. So as to assess the proposed model moderately, MLP, RF, and SVM are likewise made and tried on the equivalent dataset. The week-ahead estimating results had indicated that the proposed LSTM-based model had the option to figure building power utilization better than the near models in nine of a year. a transient power load forecast technique based on a LSTM regression network is proposed. This technique uses successions of consistent schedule load profile, temperaturedata , and mugginess data as contribution to make installing vectors that mean the inherent qualities and connections of the profiles. Moreover, for future work, we will attempt to play out the model by utilizing an alternate sort of dataset. Diverse information investigation is required for various sorts of clients (ex: home, public office, and so on) In a similar case, we will center to improve the exhibition of the model to give better exploratory outcomes later on.

REFERENCES

1. Oluklulu, Ç., A research on the photovoltaic modules that are being used actively in utilizing solar energy, sizing of the modules and architectural using means of the modules. MSc, Gazi University, Ankara, Turkey, 2001.

2. Pérez-Lombard, L., J. Ortiz, and C. Pout, A review on buildings energy consumption information. Energy and buildings, 2008. 40(3): p. 394-398.

3. Barak, S. and S.S. Sadegh, Forecasting energy consumption using ensemble ARIMA– ANFIS hybrid algorithm. International Journal of Electrical Power & Energy Systems, 2016. 82: p. 92-104.

4. Demirbas, A., Potential applications of renewable energy sources, biomass combustion problems in boiler power systems and combustion related environmental issues. Progress in energy and combustion science, 2005. 31(2): p. 171-192.

5. Zhao, H.-x. and F. Magoulès, A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 2012. 16(6): p. 3586- 3592.

6.M. Naveed Iqbal, Lauri Kütt, Argo Rosin,Complexities associated with modeling of residential electricity consumption, IEEE 59th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)

7.Huayou Wang, Yunlian Sun, Binghan Xu, Yiqin Chen, Liangpeng Ye, Shimin Yi, Jinlei Hu, XinlinXie, Optimization of Clustering Analysis of Residential Electricity Consumption Behavior, 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA),pp:418-423

8. Huiling Cai, Shoupeng Shen, Qingcheng Lin, XuefengLi,Hui Xiao, Predicting the Energy Consumption of Residential Buildings for Regional Electricity Supply-Side and Demand-Side Management, 2019, IEEE Access Vol:7, pp: 30386- 30397

9. YiliYu , Bo Wang , Zheng Wang , Fei Wang , Liming Liu,Wrapper feature selection based multiple logistic regression model for determinants analysis of residential electricity consumption, 2017, 2017 Asian Conference on Energy, Power and Transportation Electrification (ACEPT)

10. Huiling Cai, Shoupeng Shen, Qingcheng Lin, XuefengLi,Hui Xiao, Predicting the Energy Consumption of Residential Buildings for Regional Electricity Supply-Side and Demand-Side Management, 2019, IEEE Access Vol:7, pp: 30386- 30397

11. G. Renith, A Senthilselvi , “Accuracy Improvement in Diabetic Retinopathy Detection using DLIA”, Journal of Advanced Research in Dynamical and Control Systems, titled Volume 12, Issue 7, July 2020

12. V. Surya, A Senthilselvi , “A Qualitative Analysis of the Machine Learning Methods in Food Adultery: A Focus on Milk Adulteration Detection”, Journal of Advanced Research in Dynamical and Control Systems, titled Volume 12, Issue 4, May 2020

13. Mohammed Thaha .M, Pradeep mohankumar .K Murugan .B.S, Dhanasekar .S, Vijay Karthick , P,Senthilselvi.A “Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images”,Journal of Medical Systems, Volume 43 No 9 July 2019,

14. A. Senthilselvi , V. Sellam , Saad Ali Alahmari , Sivaram Rajeyyagari ,” Accuracy enhancement in mobile phone recycling process using machine learning technique and MEPH process”, Environmental Technology & Innovation,volume20,November2020.

15. R Nivetha, A Senthilselvi, “Hybrid Feature Matching for Image Forgery Detection”,International Journal of Engineering Science and Computing, volume 7, issue 3, p. 5075 – 5080, 2017