International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Energy Consumption Analysis And Forecasting

Dr. J.B.Jona , Mr. S.A.Gunasekaran , Rohinth S, Mohindran S R, Nareshkumar G

Dept. of Computing Applications, Coimbatore Institute of Technology Coimbatore,India, 641014

ABSTRACT:

Driven by technological advances, there's a rise in electricity-based equipment and this leads to excessive energy consumption (EC) and demand for power on a daily basis. to boost power management and collaboration between electricity utilized in a building and therefore the good grid, the international organization should be predicted. forecasting techniques used for prediction of the energy accurately are restricted thanks to challenges like dynamic behavior of residents and climatic conditions. So, to overcome such challenges we have a tendency to proposed a deep learning primarily based methodology. The projected methodology uses a hybrid model consisting of CNN and Bi-LSTM for predicting EC. The performance of the projected methodology is tested victimization publicly offered real dataset. check results show that the projected methodology is in a position to predict the consumption with terribly little error. The projected methodology helps in management for producing optimum amount of power.

Keywords—Machine learning, deep learning, RNN, LSTM, Bi-LSTM, GRU, Energy consumption

I. Introduction

Demand for energy in today's world is rising chop-chop due to the economic and increase. According to World Energy Outlook 2017, energy demand is predicted to grow at a compound annual growth rate (CAGR) of 1 Chronicles for the 2016-2040 amount. The residential sector conjointly represents twenty seventh of global energy consumption, and contains a vital impact on overall energy consumption [1]. Electric energy should be consumed at an equivalent time because it is generated within the power station thanks to its physical characteristics. Therefore, correct power demand also is needed for stable power provide. In order to keep up stable power provide i.e., energy should be consumed at same rate because it is created, it is important to accurately predict energy consumption ahead [2].Thekey interest for forecasting energy consumption isn't solely the short-run (sub-hourly, hourly) blackout bar, however conjointly long-run (monthly, annual) designing and investments. This prediction conjointly

helps in translating energy into price, therefore helps residents in estimating their bills and creating choices based on the predictions.Nowadays the researchers are using strategies that are based on deep learning because of their marvelous ends up in fields of computer vision and forecasting. [3] proposed a hybrid approach by integrating genetic-algorithms and Long Short-Term Memory (LSTM) for predicting energy. Li et al. [4] used a hybrid combination of Convolutional Neural Network (CNN) and LSTM for short-run prediction of energy consumption.In [5], the authors used hybrid combination of CNN and Multilayer Bidirectional LSTM (MBLSTM) for energy consumption prediction. during this paper, we have a tendency to propose a methodology to forecast energy consumption using deep learning-based strategies. The proposed methodology for prediction of energy consumption uses LSTM, bidirectional LSTM (Bi-LSTM), CNNLSTM and CNN-Bi-LSTM.[6] used RNN, gated recurrent unit (GRU), and LSTM models for electricity load prediction in Turkey and extensively bated the error. to check the performance of the projected methodology, an individual unit wattage consumption information set was used that is available on UCI machine learning repository [7] and is sampled at one-minute intervals. In our check result, it can be over that the CNN GRU model outperforms as compared to alternative models and confirms correct prediction results.

II. Background Study

A. Course Of Action

Although the primary approach to building a stable neural network should be done manually, the objective of this work is the automation of most of the included tasks so as to form solutions for individual requirements. the subsequent steps represent the logical order for working manually, but they are doing not necessarily represent the order inside automated processes as they'll part be done simultaneously by an equivalent or totally different computer systems and servers.Before making a neural network and process information, the scenario of the desired application should be well thought of. That includes: (1) the sort of output, i.e., logical or numeric; (2) interpretation of the output by a sub-system or user; (3) needed and offered data; (4) the requirement of making or preprocessing (artificial) data; (5) the spread of possible outputs and (6) assuring a stable answer.

B. Manual Approach

The second step once making the scenario is to gather all the necessary available data from the central server. once re-organizing the info, they're pre-processed into a format that the neural network will compute. it's completely necessary to make sure that not only the input data used for coaching the network but also all possible information which will occur in check and real data are inside the limit of [-1 to 1] or [0 to 1] respectively.After all data dimensions are known, the neural network architecture is built. The networks that have proven suitable for the scenario at hand consist mainly of input units within the number of dimensions within the input vector and one output unit moreover as many hidden units in multiple layers. Step four implies the creation of a collection of training data as well as one or multiple data sets for testing the quality of the training process. creating the training data set requires a representative variety of realistic data. without that, the neural network may not be able to build up abstraction levels and to reproduce them correctly under real conditions. For that reason, it's also necessary to examine the mapping by representing the check data sets to the trained neural network. Neural networks possess the feature that, in some cases, they are ready to generate cheap outputs with deviant or perhaps missing information, thus, show some form of stability. If, even once handling the data selection, training and test processes, the neural network is not able to reproduce reasonable output or does not show the desired behavior, an iteration of the scenario mapping and the following steps is necessary. In some cases, it should be necessary to reconsider the whole scenario, whereas adding new (artificial) data dimensions or altering existing data (e.g., in the scale) will also show the desired results.

About The Data

	A	B	C	D	E	F	G	H	I	J	K
1	Date	Time	Global_active_power	Global_reactive_power	Voltage	Global_intensity	Sub_meterin	Sub_meterin	Sub_meterin	DateTime	
2	01/01/07	0:00:00	2.58	0.136	241.97	10.6	0	0	0	01/01/07 0:00	
3	01/01/07	0:01:00	2.552	0.1	241.75	10.4	0	0	0	01/01/07 0:01	
4	01/01/07	0:02:00	2.55	0.1	241.64	10.4	0	0	0	01/01/07 0:02	
5	01/01/07	0:03:00	2.55	0.1	241.71	10.4	0	0	0	01/01/07 0:03	
6	01/01/07	0:04:00	2.554	0.1	241.98	10.4	0	0	0	01/01/07 0:04	
7	01/01/07	0:05:00	2.55	0.1	241.83	10.4	0	0	0	01/01/07 0:05	
8	01/01/07	0:06:00	2.534	0.096	241.07	10.4	0	0	0	01/01/07 0:06	
9	01/01/07	0:07:00	2.484	0	241.29	10.2	0	0	0	01/01/07 0:07	
10	01/01/07	0:08:00	2.468	0	241.23	10.2	0	0	0	01/01/07 0:08	
11	01/01/07	0:09:00	2.486	0	242.18	10.2	0	0	0	01/01/07 0:09	
12	01/01/07	0:10:00	2.492	0	242.46	10.2	0	0	0	01/01/07 0:10	
13	01/01/07	0:11:00	2.5	0	242.88	10.2	0	0	0	01/01/07 0:11	
14	01/01/07	0:12:00	2.494	0	242.57	10.2	0	0	0	01/01/07 0:12	
15	01/01/07	0:13:00	2.492	0	242.41	10.2	0	0	0	01/01/07 0:13	
16	01/01/07	0:14:00	2.48	0	241.81	10.2	0	0	0	01/01/07 0:14	
17	01/01/07	0:15:00	2.478	0	241.73	10.2	0	0	0	01/01/07 0:15	
18	01/01/07	0:16:00	2.47	0	241.29	10.2	0	0	0	01/01/07 0:16	
19	01/01/07	0:17:00	2.466	0	241.11	10.2	0	0	0	01/01/07 0:17	
20	01/01/07	0:18:00	2.456	0	240.59	10.2	0	0	0	01/01/07 0:18	
21	01/01/07	0:19:00	2.46	0	240.83	10.2	0	0	0	01/01/07 0:19	
22	01/01/07	0:20:00	2.544	0.092	240.9	10.6	0	0	0	01/01/07 0:20	
23	01/01/07	0:21:00	2.55	0.116	241.15	10.4	0	1	0	01/01/07 0:21	
24	01/01/07	0:22:00	2.554	0.118	241.55	10.6	0	1	0	01/01/07 0:22	
25	01/01/07	0:23:00	2.65	0.218	241.67	11	0	2	0	01/01/07 0:23	
26	01/01/07	0:24:00	2.682	0.258	242.45	11	0	1	0	01/01/07 0:24	
27	01/01/07	0:25:00	2.66	0.252	241.6	11	0	1	0	01/01/07 0:25	
28	01/01/07	0:26:00	2.65	0.25	241.14	11	0	2	0	01/01/07 0:26	
29	01/01/07	0:27:00	2.654	0.25	241.38	11	0	1	0	01/01/07 0:27	
30	01/01/07	0:28:00	2.642	0.248	240.9	11	0	2	0	01/01/07 0:28	
31	01/01/07	0:29:00	2.644	0.25	241.15	11	0	1	0	01/01/07 0:29	
32	01/01/07	0:30:00	2.648	0.252	241.32	11	0	1	0	01/01/07 0:30	
33	01/01/07	0:31:00	2.646	0.252	241.35	11	0	2	0	01/01/07 0:31	
34	01/01/07	0:32:00	2.644	0.252	241.38	11	0	1	0	01/01/07 0:32	
35	01/01/07	0:33:00	2.656	0.256	242.03	11	0	1	0	01/01/07 0:33	

2007_energy_consumption

This archive contains 521679 measurements gathered in a house located in Sceaux (7km of Paris, France) in the year 2007

- (global_active_power*1000/60 - sub_metering_1 - sub_metering_2 - sub_metering_3) represents the active energy consumed every minute (in watt hour) in the household by electrical equipment not measured in sub-meterings 1, 2 and 3.
- The dataset contains some missing values in the measurements (nearly 1,25% of the rows). All calendar timestamps are present in the dataset.

Approach Flowchart

1. Anomaly Detection

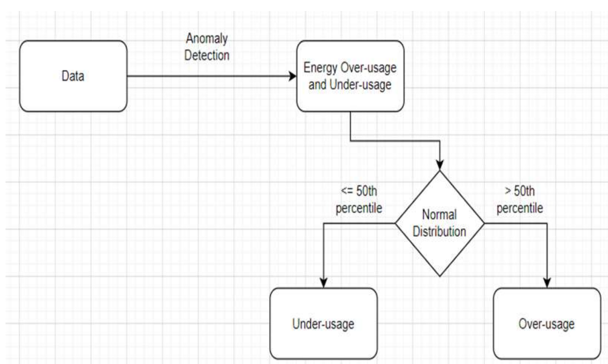

Figure.2.1

2. Future Prediction

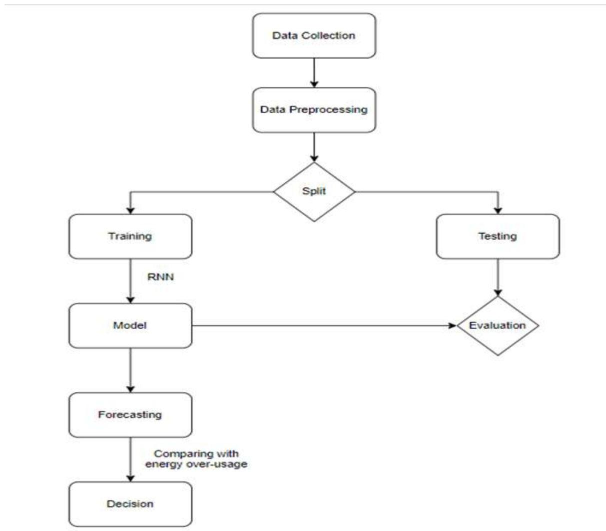

Figure.2.2

III. Proposed Methodology

To forecast the energy consumption of households, we propose a framework based on different training models. In the proposed methodology, 'individual household electric power consumption data set' is used which is available on UCI machine learning repository[20].

Data Preprocessing

Dataset contains a lot of missing values and to deal with missing values we replaced the missing values with the mean value of that column. Label encoding is done to convert the labels into the numeric form so as to convert them into the machine-readable form.

Feature Selection and Training

In feature selection, we have transformed our dataset into features and labels by using a sliding window approach. We have used previous values of the global active power (GAP) column as features and future value as label.

Basic features extracted:

- Month
- Day
- Quarter
- Hour
- Day_of_the_week
- Is_weekend or not

The train-test split is done which is a technique for evaluating the performance of an algorithm. A dataset is divided into two subsets as part of the technique.

- Train Dataset: Used to fit the model.
- Test Dataset: Used to evaluate the fit model.

The objective is to estimate the performance of the model on new data: data not used to train the model.

Standardization shifts the distribution to have a mean of zero and a standard deviation of one by removing the mean (called centering) and dividing by the standard deviation.

The MinMaxScaler estimator will fit on the training data set and the same estimator will be used to transform both the training and the test data set.

z-score

Figure.3.1

The normal distribution is a symmetric probability distribution centered on the mean, indicating that data around the mean occur more frequently than data far from it. In graph form, normal distribution will appear as a bell curve shown in Fig.3.1.

The Probability Distribution Function(PDF) of the Normal Distribution is used and compared with the PDF of energy usage, and classified the usage as "Normal" or "Anomaly" with respect to the PDF of the 40th percentile.

IV. METHODOLOGY

Logistic Regression using Keras

Logistic regression is supervised learning, but contrary to its name, it is not a regression, but a classification method. It is assumed that the data may be split (classified) by a line or an n-dimensional plane,i.e. it is alinear model. To put it another way, the classification is done by calculating the value of the first-degree polynomial of the form:

$$y = w1*x1+w2*x2+\ldots+wn*xn \quad (1)$$

where x is the input parameter, w is the weight assigned to this parameter, and n is the number of input parameters. The next step in logistic regression is to pass the obtained y result through a logistic function to get a value in the range (e.g. sigmoid or hyperbolic tangent) (0, 1). After obtaining this value, we can classify the input data to group A or group B on the basis of a simple rule: if y > = 0.5 then class A, otherwise class B.

LSTM

An LSTM [8] a type of RNN that is way ahead of these conventional networks which are having these feed- forward mechanisms, it provides a better solution for short- term memory problems.

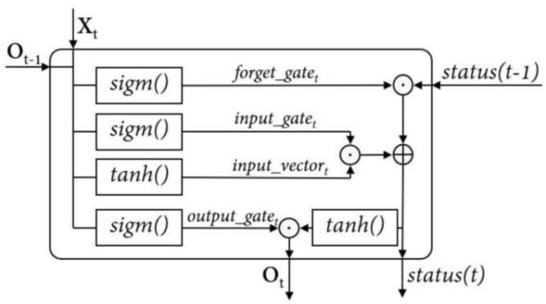

Fig. 4.1 LSTM Architecture

These LSTM networks differ from traditional RNNs as LSTM have an extra gate present called "forget gate" which helps LSTM to overcome issues of RNNs. For computing the hidden states (at) present in LSTM unit we require values of input gate (it), forget gate (ft) and output gate (ot) and all of them are described in equations given below:

ft =σ(wf.[at−1,xt]+bf)

it = σ(wi.[at−1, xt] + bi)

ot= σ(wo.[at−1, xt] + bo)

$\tilde{c}$=tanh(w.[a ,x]+b) t ct−1tc

c=f.c +i.$\tilde{c}$ t t t−1 t t

at = tanh(ct) (2)

where ft is the output from the forget gate, it is the output from the input gate, ot is the the output from the output gate and ct gives the cell state and wf ,wi and wo are weight matrices and bf , bi and bo are the biases values which corresponds to the ft gate, it gate and or gate of the LSTM unit respectively which are being learned by the model during the training process. Here σ denotes the activation function and {.} symbol here implies for element-wise multiplication and xt represents the input.

Bi-LSTM

A bidirectional LSTM, often known as a biLSTM, is a sequence processing model that consists of two LSTMs: one that takes input in one direction and the other that takes it in the opposite direction. BiLSTMs effectively improve the quantity of data available to the network, increasing the algorithm's context.

The BiLSTM connects two hidden layers with opposite orientations to produce the same output that considers both past and future states at the same time.. In many cases, having simultaneous access to these two states helps the model perform better.. The main principle on which the architecture of a BiLSTM is based, consists in splitting each neuron into two directions: one corresponds to the future and reflects the forward states, while the second corresponds to the past, highlighting the backward states [9]

When it comes to the training of BiLSTMs, similar algorithms to those utilised in the case of ordinary LSTMs are applied, for example the ADAM, SGDM, and RMSProp training algorithms. However, because the input and output layers cannot be updated concurrently when using the back-propagation approach for training, a number of additional operations are necessary.

GRU

GRUs are very similar to Long Short Term Memory(LSTM). Just like LSTM, GRU uses gates to control the flow of information. They are relatively new as compared to LSTM. This is the reason they offer some improvement over LSTM and have simpler architecture.

Fig. 4.2 GRU Architecture

A GRU holds a simpler architecture than the LSTM unit; two intuitive advantages of GRU over LSTM are the absence of cell status and the reduction of gate number, which reduces the required amount of computation. Additional details on the mechanisms of the two special units can be achieved at [10]

Several commonly selected GRU based RNN structures are tested in [11] to predict power consumption. In [12], GRU based RNNs are adopted to forecast household power demand on one second resolution.

V. Results and Discussion

In this paper, the programming is done using python, keras-API and TensorFlow at backend. The results have been achieved on the GPU environment of Google Collaboratory.

A. Analysis Of Dataset.

The dataset that we have used in this paper is 'individual household electric power consumption data set' which is available on UCI machine learning repository [20], this dataset contains information on energy consumed between 2006 and 2010 on a house located in France.

Dataset consists of 2072259 instances out of which 1.25% values are missing, and these values are treated in the data processing step. This database contains energy consumption at a sample rate of one minute for a four-year period. In this database, the overall energy consumed is given by the global active power (GAP) column. The amount of energy used is also provided by sub metering, and is collected by implanted sensors

B. Performance Metrics

The performance of the proposed methodology is tested using the specific performance metrics and achieved the lowest error for predicting energy consumption. The optimizer used during the training of different models is Adam. As our problem is regression based so for evaluation of result, we have used performance metrics i.e. MSE, MAE and MAPE. Let yi denotes actual value, yˆi denotes predicted value and N denotes the total sample. Equations (3) to (5) represent MSE, MAE and MAPE respectively.

1) Mean Square Error: Mean square error (MSE) can be defined as the mean of square error between actual and predicted value.

$$MSE = 1/N\,(\Sigma\,(y_i - \hat{y}_i)^2)$$

2) Mean Absolute Error: Mean absolute error (MAE) can be defined as the mean of absolute error between actual and predicted value.

$$MAE = (1/N)\,\Sigma\,|y_i - \hat{y}_i|$$

3) Mean absolute percentage error: Mean absolute percentage error (MAPE) can be defined as the percentage of mean of relative absolute error between actual and predicted value.

$$MAPE = (100\% / N)\,\Sigma\,|(y_i - \hat{y}_i) / y_i|$$

C. Test Results

To check the performance of the proposed methodology, we
have taken the GAP column in the dataset. Then, we converted our dataset into two parts by splitting it in a ratio of 80:20. The First 80% part is the

training set and accounts for 1660146 observations and the remaining 20% part is known as the test set having 412113 observations. Different Models in the proposed methodology are trained using the training dataset and then predict the values of the test dataset. In the next step, compared the predicted and actual values using different performance metrics. In the proposed methodology, the deep learning models used for comparison are LSTM, Bi-LSTM and GRU.

We have trained our model for 100 epochs with early stopping at patience of 2. Table I shows the test results for the different models used in this paper. The MSE is 0.043, 0.042, 0.038, for LSTM, Bi-LSTM, GRU, respectively. From Table I, it is concluded that GRU outperformed all the other models in terms of different metrics and provides least MSE, MAE and MAPE value i.e., 0.038, 0.072 and 8.23 respectively.

We have also checked the performance of different models by plotting the actual and predicted values. Fig. [5.1,5.2,5.3,5.4] illustrate the plot of actual and predicted values obtained from each of the models for a portion of the test set. Out of all the observations, only 60 observations are plotted to show the model performance. From all the plot it can be seen that the predicted value is close to the actual value. In the case of GRU curve fitting is best as compared to other models. Here the GAP is given in kilowatt and the time interval is one minute.

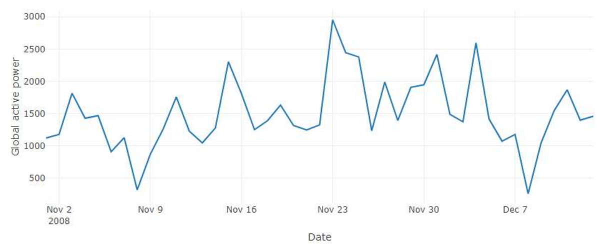

Fig. 5.1 Actual Values of energy consumption

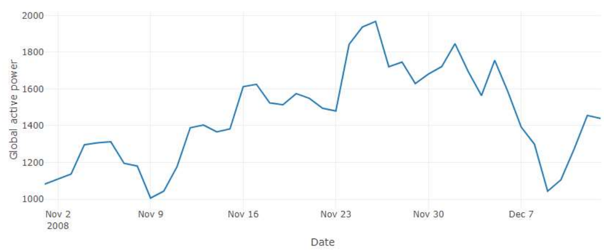

Fig. 5.2 Predicted Values for LSTM method

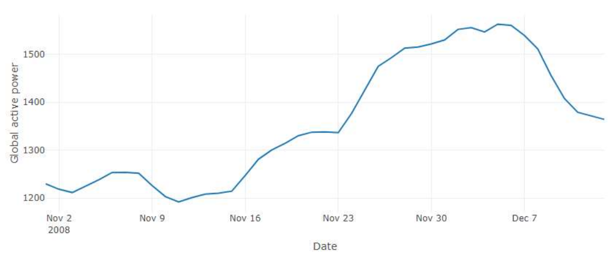

Fig. 5.3 Predicted Values for Bi-LSTM method

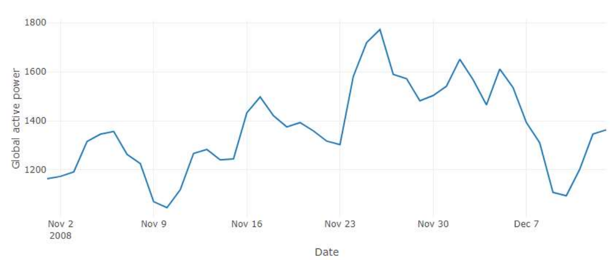

Fig. 5.4 Predicted Values for GRU method

VI. CONCLUSION:

Management of energy usage in smart grids is a daunting task due to noise disturbances and many other factors. Prediction of the energy consumption in households is critical in order to improve the energy consumption and collaboration between smart grids and residential buildings. In this paper, we have proposed a deep learning-based methodology for prediction of energy consumption in households. We have used LSTM, Bi-LSTM and GRU models, for prediction of energy consumption.

We have checked the performance of the proposed methodology on publicly available dataset on household power consumption. Test results show that GRU model provides least MAPE, MSE and MAE and outperforms as compared to other models.

REFERENCES:

[1] P. Nejat, F. Jomehzadeh, M. M. Taheri, M. Gohari, and M. Z. A. Majid, "A global review of energy consumption, co2 emissions and policy in the residential sector (with an overview of the top ten co2 emitting countries)," Renewable and sustainable energy reviews, vol. 43, pp. 843– 862, 2015.

[2] H. Ibrahim, A. Ilinca, and J. Perron, "Energy storage systemscharac- teristics and comparisons," *Renewable and sustainable energy reviews*, vol. 12, no. 5, pp. 1221–1250, 2008.

[3] A. Almalaq and J. J. Zhang, "Evolutionary deep learning-based energy consumption prediction for buildings," *IEEE Access*, vol. 7, pp. 1520–1531, 2018.

[4] T.-Y. Kim and S.-B. Cho, "Predicting residential energy consumption using cnn-lstm neural networks," *Energy*, vol. 182, pp. 72–81, 2019.

[5] F.U.M.Ullah,A.Ullah,I.U.Haq,S.Rho,andS.W.Baik,"Short-term prediction of residential power energy consumption via cnn and multi- layer bi-directional lstm networks," IEEE Access, vol. 8, pp. 123 369– 123 380, 2019.

[6] Tokgoz, A.; Unal, G. A RNN Based Time Series Approach for Forecasting Turkish Electricity Load. In Proceedings of the 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2–5 May 2018; pp. 1–4.

[7] G. Hebrail and A. Berard, "Individual household electric power con- sumption data set," *E´. d. France, Ed., ed: UCI Machine Learning Repository*, 2012.

[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[9] Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. IEEE Trans. Signal Process. 1997, 45, 2673–2681. [CrossRef]

[10] Sak H., Senior A., Beaufays F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. *arXiv.* 2014. 1402.1128

[11] Kumar S., Hussain L., Banarjee S., Reza M. Energy Load Forecasting using Deep Learning Approach-LSTM and GRU in Spark Cluster; Proceedings of the 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT); IIEST; Howrah, India. 12–13 January 2018.

[12] Heghedus C., Chakravorty A., Rong C. Energy Load Forecasting Using Deep Learning; Proceedings of the 2018 IEEE International Conference on Energy Internet (ICEI); Beijing, China. 21–25 May 2018; pp. 146–151.