JOIV : Int. J. Inform. Visualization, 7(4) - December 2023 2436-2444

INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

A Prediction Model of Power Consumption in Smart City Using Hybrid Deep Learning Algorithm

Salam Abdulkhaleq Noaman [a,*], Ali Mohammed Saleh Ahmed [a], Aseel Dawod Salman [a]

[a] *Department of Computer Science, University of Diyala Iraq, Baqubah, Diyala, 32001, Iraq*

Corresponding author: [*]salam2015noaman@gmail.com

***Abstract*—A smart city utilizes vast data collected through electronic methods, such as sensors and cameras, to improve daily life by managing resources and providing services. Moving towards a smart grid is a step in realizing this concept. The proliferation of smart grids and the concomitant progress made in the development of measuring infrastructure have garnered considerable interest in short-term power consumption forecasting. In reality, predicting future power demands has shown to be a crucial factor in preventing energy waste and developing successful power management techniques. In addition, historical time series data on energy consumption may be considered necessary to derive all relevant knowledge and estimate future use. This research paper aims to construct and compare with original deep learning algorithms for forecasting power consumption over time. The proposed model, LSTM-GRU-PPCM, combines the Long -Short-Term -Memory (LSTM) and Gated- Recurrent- Unit (GRU) Prediction Power Consumption Model. Power consumption data will be utilized as the time series dataset, and predictions will be generated using the developed model. This research avoids consumption peaks by using the proposed LSTM-GRU-PPCM neural network to forecast future load demand. In order to conduct a thorough assessment of the method, a series of experiments were carried out using actual power consumption data from various cities in India. The experiment results show that the LSTM-GRU-PPCM model improves the original LSTM forecasting algorithms evaluated by Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) for various time series. The proposed model achieved a minimum error prediction of MAE=0.004 and RMSE=0.032, which are excellent values compared to the original LSTM. Significant implications for power quality management and equipment maintenance may be expected from the LSTM-GRU-PPCM approach, as its forecasts will allow for proactive decision-making and lead to load shedding when power consumption exceeds the allowed level.**

***Keywords*—Power consumption prediction; LSTM; GRU; Mean Absolute Error (MAE); Root Mean Squared Error (RMSE).**

Manuscript received 1 Jun. 2023; revised 7 Aug. 2023; accepted 20 Oct. 2023. Date of publication 31 Dec. 2023.

I. INTRODUCTION

A smart city uses IT (information technology), AI (artificial intelligence), and IoT (Internet of Things) to gather, analyze, and combine critical information from fundamental systems in current cities in order to improve facilities and make them stronger [1]. A smart city is a concept made a reality by calculating a sizable quantity of data acquired by sensors, cameras, and other electronic technologies to deliver services, manage resources, and address issues encountered in everyday life [2]. Smart cities may meet the city's services, public safety, everyday needs, and commercial and industrial operations swiftly, effectively, and intelligently [3]. All systems are based on power or energy. The smart grid is a kind of resilient electrical system that adjusts to the customer's various demands and is inherently self-healing. As a result, it serves as one of the solid foundations for implementing the smart city idea [4].

The design of smart grids, the growth of the electrical market, and the sustainability of power all depend on accurate predictions of power use [5]. Providers and consumers may benefit from accurate projections and helpful advice for improved power management, grid security, and load control. Due to the advancement of machine/deep learning-based techniques, several scholars have begun to concentrate on these tactics [6], [7]. A unique deep learning method for processing time series predictions, the LSTM neural network, has recently shown encouraging prediction results [8], [9]. The gated unit mechanism used in this model improves the conventional recurrent neural network (RNN) by preventing the vanishing gradient issue. There are still several issues with the LSTM neural network model, including poor learning efficiency and gradient disappearance due to the rise of the hidden layer's number. Many academics and practitioners are searching for ways to enhance the LSTM model [10], [11]. The methods used in the field of electricity to forecast short-term power may be separated into four categories: very-short, short, medium, and long-term [12]. The terms "very-short and short-term predictions" refer to forecasts with time horizons between a few minutes to a few days. While every long-term prediction refers to annual or multiple years in advance forecasts, medium-term load prediction comprises developing projections for a week up to a year in advance. Although each of these prediction words has a specific use, for this study, our attention is focused on short- and medium-term predictions [13] and [14]. The major objective of this research is to increase the predictability of short-term power use in residential cities, which would decrease client's consumption and have positive economic effects. The testing portion demonstrates how useful the proposed model is, ensuring the hybrid prediction model's superior performance to the original model. The potential of deep learning algorithms in tackling various difficulties in power grid operations, controls, and monitoring has been examined in a number of research papers. The contributions of other researchers to this subject are mentioned in the section that follows. Realizing the smart city vision via using smart grids is providing new possibilities for study into the real-time employment of intelligent systems, notably in grid stability prediction.

Convolutional neural networks (CNN) were compared to other machine learning models for their efficacy in predicting energy demand at the level of a specific building by Amarasinghe et al. The CNN model produced the best results with an RMSE value of around 0.677 [15]. Yan et al. [16] developed a hybrid deep-learning neural network architecture using CNN and LSTM neural networks to forecast home power demand. The CNN-LSTM architecture has the best MAE and RMSE, 0.0058 and 0.0112. The UK-DALE project provided five real-world residential power consumption statistics.

Wang et al. [17] used LSTM neural networks to estimate electricity demand and detect grid faults. The model detected 9,784 anomaly points from 271 users and 601 windows from 231 users. 134 of 231 users stole energy. The model has 0.58 recall and 0.71 accuracy. An LSTM neural network model was put out by Khafaf et al. [18] to predict the energy consumption of groups of energy consumers three days in advance. The findings demonstrate that a 3-day ahead demand estimate can be made correctly with a mean absolute percent error of 3.15 percent. In [19], authors used the LSTM neural network in their research to forecast the long-term energy usage of a cooling system. Their research showed that the suggested LSTM technique produced a root–mean–square–error 19.7% lower than that of the reference feedforward neural network.

Khan et al. [20] developed a two-step method for forecasting the load on residential structures. Raw data on power utilization must first be cleaned up to successfully train the (CNN) and multilayer-bidirectional (MB-GRU). The suggested model included two datasets, although the appliances load prediction (AEP) dataset produced the best error prediction results (RMSE (2%), MAE (1%). An LSTM network-based attention mechanism-based module for predicting power consumption was suggested by Lin et al. [21]. The experimental data sets utilized came from China Southern Power Grid's power supply station. According to the results, the suggested model's prediction accuracy rose by 6.5%. Somu et al. [22] developed A hybrid LSTM, convolutional neural network, and clustering using k-means model to estimate building power usage. The LSTM layer's 5100 neurons were a fixed number. A strategy for forecasting the demand for day-ahead plug-in energy using LSTM neural networks was reported by Markvoic et al. [23]. To estimate the amount of power utilized in French cities, Mahjoub et al. [6] employed recurrent neural network models with Long -Short-Term -Memory, Gated- Recurrent- Units, and Drop-Gated Recurrent Units. Concerning prediction accuracy and speed, the Drop-GRU neural network model surpasses the LSTM and GRU models. The Drop-GRU model has an average of 0.07195 RMSE and 0.0526 MAE. An LSTM model was suggested by Zhou et al. [24]to forecast the energy use of air-conditioning systems.

The main area of concern is the comparative effectiveness of recently developed procedures versus more established ones. A hybrid model combining LSTM and GRU is presented to handle this issue in order to address problems with energy consumption estimation, peak energy identification, and load reduction. These expectations are motivated by the desire to maintain a harmonious equilibrium between energy suppliers and their customers and to guarantee the stability and dependability of the electrical system in smart cities.

The suggested hybrid model seeks to produce accurate forecasts by successfully identifying future patterns in time series data. The model's GRU component imitates the dynamic changes seen in historical energy consumption data, making it easier to spot emerging characteristics. In order to reduce the possibility of overfitting, the LSTM technique was chosen because of its capacity to maintain and improve the fundamental characteristics of the data throughout time. The article offers a thorough methodology, including data processing and hybrid model training to obtain energy consumption time series data.

II. Materials and Methods

Two common deep-learning neural network subtypes, LSTM and GRU, have attracted a lot of interest lately. This research aims to develop a hybrid algorithm that outperforms conventional techniques to solve single-city power consumption forecasts' high volatility and unpredictability. The proposed approach pre-processes the data using suggested methods for dealing with time series data types and utilizes GRU output to train the LSTM model, producing more reliable and accurate forecasts.

A. Dataset Description

The "Power Consumption in India dataset" (2019-2020) was utilized in this study. It provides a time series of data on power consumption for the 17 months from January 2, 2019, to May 23, 2020. Each column corresponds to an Indian state, and each row is indexed by date. Each data point in the dataset represents the amount of energy used in Mega Units (MU) by the specified state on the specified date. The Power -System -Operation –Corporation- Limited (POSOCO), an Indian government-owned corporation that falls under the Ministry of Power, gathered the data. 43 attributes from the dataset reflect Indian cities and time series [25].

B. Long Short-Term Memory

The long- and short-term memory (LSTM) model is one distinctive variant of the recurrent neural network (RNN) that provides feedback at each neuron. Past neuron inputs influence RNN output in addition to the current neuron input and weight [26]. Thus, conceptually speaking, analyzing time series data frequently uses the RNN structure.

However, difficulties with exploding and disappearing gradients arise when working, including a lengthy and associated sequence of data samples [27], which subsequently serves as the pivotal point for introducing the LSTM model [9]. The RNN model's problem with disappearing gradients is resolved by the internal loops of the LSTM, which preserve relevant information and discard trash. The LSTM model's flowchart has four crucial components: the cell's condition, the input and forget gates, and the output gate, Figure 1 [9], [28]. The input, forget, and output gates manage the update, maintenance, and deletion of data contained in cell status. The method of forward computing is shown by:

$$f\_t = \sigma(w\_f \cdot [h\_(t-1), x\_t] + b\_f), \tag{1}$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i), \tag{2}$$

$$\tilde{c}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c), \tag{3}$$

$$c_t = f_t \cdot c_{t-1} + i_t . \tilde{c}_t, \tag{4}$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o), \tag{5}$$

$$h_t = \tanh(c_t) \tag{6}$$

Where, respectively, c_t , c_{t-1},and $\tilde{c}_t$ stand for the value of the present cell state, the cell status value for the latest time period and the recent modification to the current cell status value. Forget, input, and output gate, respectively, are represented by the notations f_t,i_t, and o_t. Equations (4) and (6) may be used to determine the output value h_t depending on the values of $\tilde{c}_t$ and c_{t-1} values. Following the back-propagation through time (BPTT) process, all weights, w_f, w_i, w_c,and w_0, are adjusted depending on the variance between the produced and real values [16].

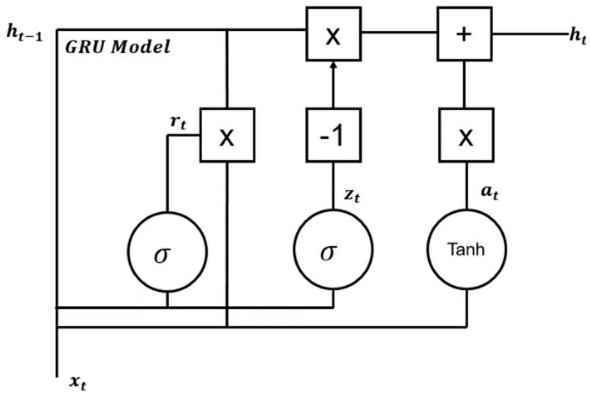

Fig. 1 An example line graph with colors that stand out both on-screen and in a hardcopy of black and white

C. Gated Recurrent Unit

Among the most popular types of modified RNN, GRU employs a special "gated recurrent neural network" constructed on an improved LSTM [29]. Except associating the input gate and forget gate into a single update gate, the GRU's internal design is identical to that of the LSTM [30]. One of the two gates in this architecture is the update gate; it regulates the scope and keeps historical data current. The other gate, which serves as the reset gate, makes the determination of whether past knowledge and the current state should be connected [6]. Figure (2) depicts the basic construction of a GRU unit.

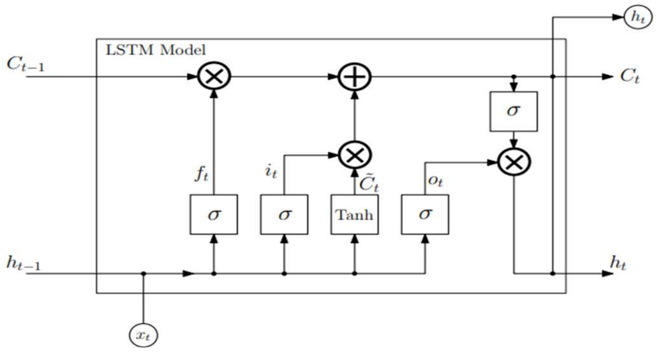

Fig. 2 The GRU Model's Internal Structure [6]

$$z_t = \sigma(w_z \cdot [h_{t-1}, x_t] + b_z), \tag{7}$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_r), \tag{8}$$

$$a_t = \tanh(r_t \cdot w_a \cdot [h_{t-1}, x_t] + b_a), \tag{9}$$

$$h_t = (1 - z_t) \cdot a_t + z_t \cdot h_t - 1, \tag{10}$$

Where h_t is produced by the present layer at time *t* and x_t is the training set's vector input at period *t*. The update and reset gates are identified with z_t and r_t , consecutively. The activation of the candidate a_t is at hand [31].

C. The Proposed Power Consumption Prediction Model

A methodology based on hybrid LSTM + GRU models is put forward to estimate power usage. The proposed work's technique is described using a block diagram in Figure 3, which depicts the many phases used to build the predictive model. The proposed LTSM-GRU-PPCM consists of several key stages, which will be elaborated in detail as follows:

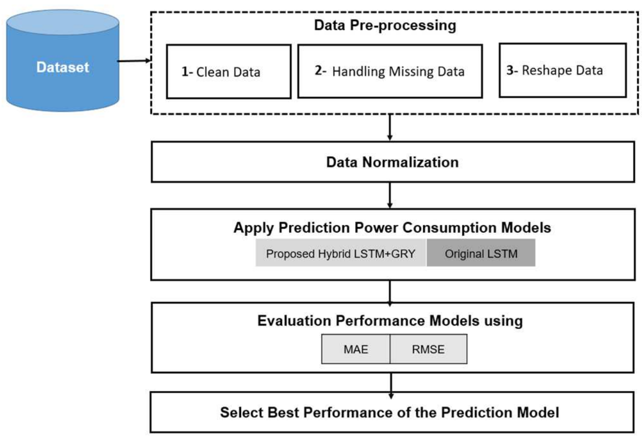

Fig. 3 The Long-Short-Term Memory and Gated-Recurrent-Unit Prediction Power Consumption Model's (LSTM-GRU-PPCM) Suggested Block

D. Data Pre-processing

Given the time series nature of the input data, it requires a distinct approach that caters to its unique properties. As such, the proposed hybrid LSTM-GRU-based power consumption prediction model (LTSM-GRU-PPCM) employs a tripartite methodology to tackle this task effectively.

1) Cleaning Power Consumption Data: Data cleansing is an important step in preparing the data accurately for deep learning prediction models. Our proposed strategy employs three distinct methods due to the unique handling requirements of time series data. To improve the model's performance, irrelevant characteristics must be found and eliminated during the early stage. As a result, the suggested model carefully chooses a particular set of 17 features from the initial collection of 43 attributes. The second step involves removing variables with a high proportion of missing values to enhance the overall quality of the dataset. Finally, features with spaces, % signs, and other unusual characters are renamed to facilitate feature name-calling from the dataset.

2) Handling Missing Data: The proposed methodology scrutinizes the presence of absent values and analyses the remedial actions taken by computing the meaning of the available data for each element. Suppose we have a dataset with an attribute X that has been recorded for a set of entities, and the set of N observed values or observations for X is denoted by 1, $x2$, … xN. Various measures can be used to identify the central tendency of the observations for attribute X, including midrange, mode, median, and mean. The mean of this set of values can be calculated using Equation (11) [32]:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{x_1 + x_2 + \cdots + x_N}{N} \tag{11}$$

3) Reshape Raw Data: The input data is transformed into a 3D matrix of dimensions [samples, time steps, features] using the "Reshape Raw Data" technique to optimize the pre-processing of input data for prediction models. The model must estimate features based on past values to predict time-series data. The time series to supervised function takes samples at time t-n, t-n-1, ... , t-1 as input and samples at time t, t+1,..., t+n as output. Data on daily electricity use from Indian cities makes up the dataset used in this particular case. As a result, the data from the previous day's power use serves as the model's input, and the output is the data from the current day's power usage. This arrangement establishes a time step of one and the input matrix shape of [samples, 1, features].

E. Data Normalization

The consumption profiles display various varieties of data, making the usage of reduced-focused standardization necessary [6]. Z-score normalization is implemented for standardization purposes, which can be expressed mathematically using the following Equations (12), (13), and (14) [33]and [34]:

$$Z - score = \frac{x - x_{mean}}{x_\sigma} \tag{12}$$

$$x_\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - x_{mean})^2} \tag{13}$$

$$x_{mean} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{14}$$

The objective of data standardization is to rescale variable values to a range of -1 to 1 while preserving inter-Val distances [35].

F. Hybrid Deep-Learning Models

In the field of deep learning, the models of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are currently extensively used in sequence prediction tasks, including power consumption prediction. Both models are designed to handle the problem of vanishing gradients in traditional recurrent neural networks by selectively updating and forgetting information.

LSTM and GRU models with two hidden layers are combined to benefit from each model's advantages. GRU has fewer parameters and are generally faster to train than LSTM models. However, LSTM has been shown to be more effective in capturing long-term dependencies in sequences. By combining the two models in a hybrid approach with two hidden layers, we can leverage the benefits of both models and achieve more accurate power consumption predictions.

In the hybrid LSTM+GRU model, the input data contains i of data (I_i) where the number of i =1,2,..,1000 nodes .The model is composed of two hidden layers, the first of which has 1000 GRU nodes and the second of which has 1000 LSTM nodes. We used 100 epochs for training and batch size 64 for the training dataset to prevent overfitting. The model output contains eight dense nodes (D_i) connected in fully connected layers where each node represents one class, as shown in Figure 4. This architecture is intended to give precise predictions for each city in the dataset and efficiently capture the complicated temporal correlations found in the data.

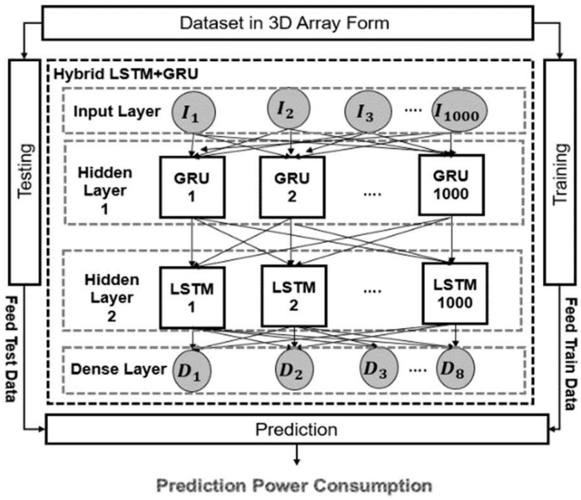

Fig. 4 The Framework of the proposed hybrid lstm-gru prediction model

In the first hidden layer, we use GRU with 1000 nodes to capture the "short-term dependencies" and patterns in the data. The second hidden layer receives the outputs from the first hidden layer, which uses LSTM with 1000 nodes to capture the "long-term dependencies" and patterns in the data. This hybrid approach with two hidden layers collects and integrates short-term and long-term dependency in power usage data. This improves forecast precision and accuracy. The summary of the suggested hybrid model is illustrated in Table (1).

TABLE I
SUMMARY OF THE PROPOSED HYBRID LSTM+GRU MODEL

Layer Type	Output Shape	Parameters
gru(GRU)	(None,1,1000)	4036000
lstm(LSTM)	(None,1,1000)	6006000
dense(Dense)	(None,8)	8008
Total Parameters =10,050,008		
Trainable Parameters =10,050,008		
Non-Trainable Parameters=0		

G. Evaluation of the Prediction's Performance Model for Power Consumption

The data source's reliability, prediction methods, and the testing portion circumstances, among other variables, can influence the accuracy of a forecast. To evaluate the quality of predictions, it is necessary to utilize generic metrics. The most crucial indicator is the accuracy metric, which provides immediate insight into the performance of a forecasting model. Two measures, commonly employed by many models to compare predictions with actual results, have been selected for assessment in this research: MAE (Mean Absolute Error) and RMSE (Root Mean Square Error) [36] and [37]. The equations (15) and (16) define these measures, as outlined in [38][38] and [39]:

$$MAE = \frac{\sum_{i=1}^{n} |x_i - \bar{x}_i|}{N} \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{n} (x_i - \bar{x}_i)^2} \tag{16}$$

where x_i is the actual value; $\bar{x}_i$is the outcome of the forecast for x_i and N is the whole samples of test.

III. RESULTS AND DISCUSSION

The computer system used for the tests in this research had to meet certain hardware and software specifications. A Core i5-CPU 3320M processor running at 2.60 GHz and 8GB of RAM were included in the system, which ran Windows 10 as its operating system. "The Power consumption in India (2019-2020)" dataset was used to put the suggested approach into practice. Dedicated interfaces were created and visually shown in Figures 5 and 6 to make training and testing the model easier. The creation of these user-friendly and clear interfaces streamlined the implementation and performance assessment of the suggested approach.

As depicted in Figure (5), the initial phase for executing the proposed model entails launching the training interface and selecting the application of the hybrid LSTM+GRU model on the training data from the India power consumption dataset, which contains data on 17 distinct cities. The training interface provides comprehensive data analysis for each city, enabling the user to comprehend the consumption behavior of power and its correlation with various factors such as people's lifestyles, weather conditions, and other pertinent parameters.

Testing the power consumption prediction model comes after the training phase. This is accomplished by feeding the model examples from the testing dataset that can be predicted to have high or low power usage. Figure (6) illustrates the testing phase of the planned hybrid LSTM+GRU model. The testing phase is crucial in determining the effectiveness of the model and its ability to generalize to unseen data.

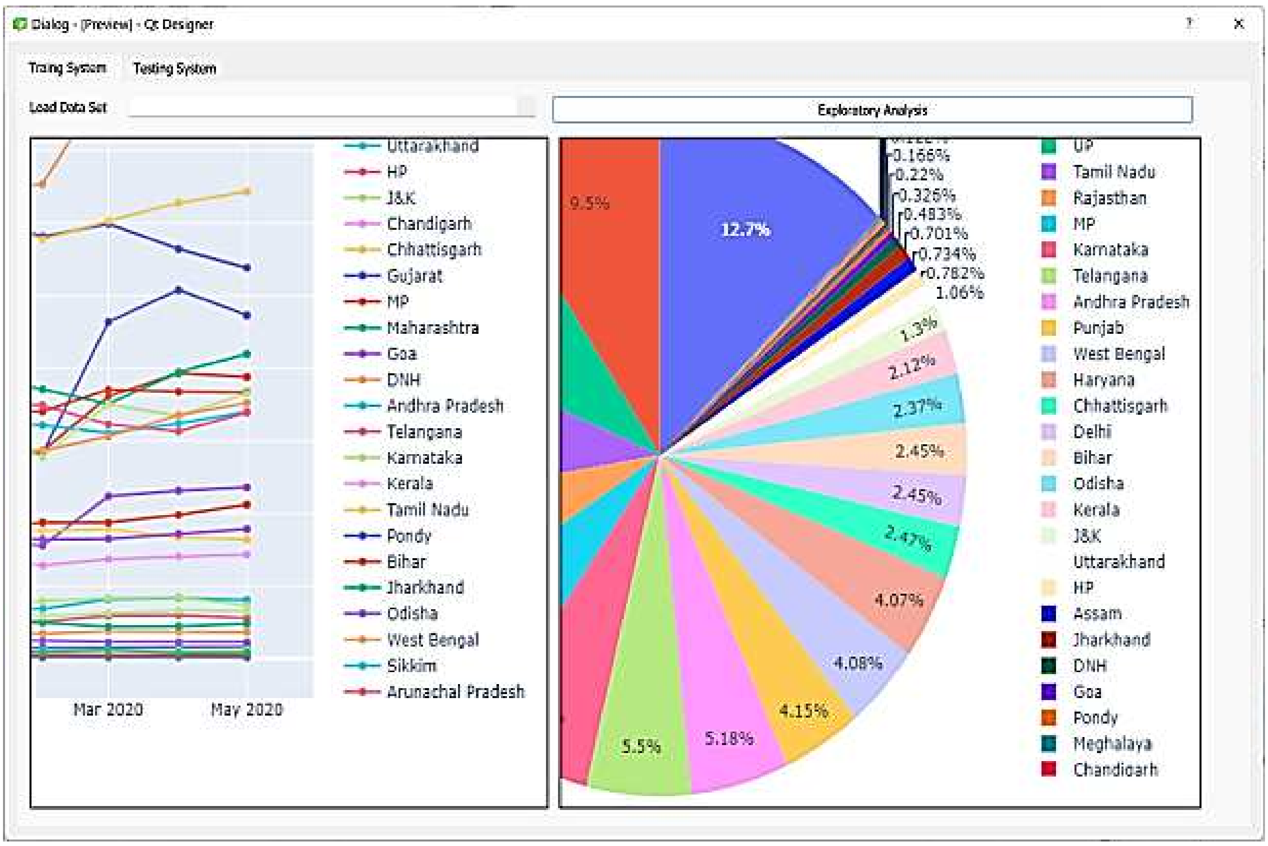

Fig. 5 Implementation training interface of the proposed LSTM-GRU-PPCM model

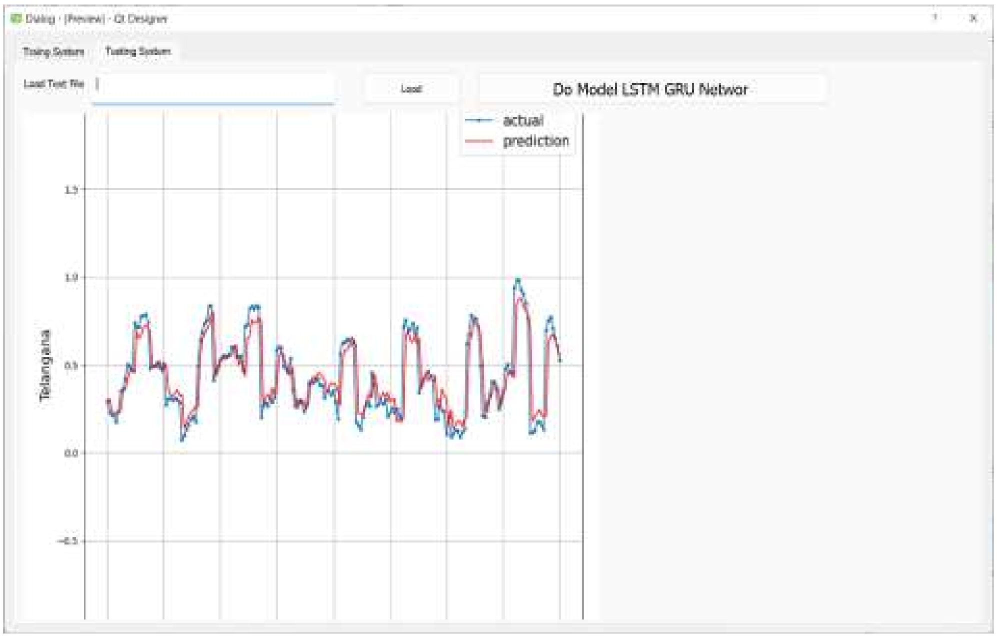

Fig. 6 Implementation testing interface of the proposed LSTM-GRU-PPCM model

The model suggested in this research can analyze power consumption forecast data independently for each of the 17 Indian cities over a period of 200-time steps. This is seen in Figure 7, where the horizontal axis displays the time series, the vertical axis depicts the power consumption for each city, the blue line indicates the actual consumption, and the red line indicates the expected consumption derived from the suggested model. The close closeness of the actual and anticipated values in the figure shows that the suggested model successfully predicts power consumption correctly and reduces the margin of error between the actual and expected values.

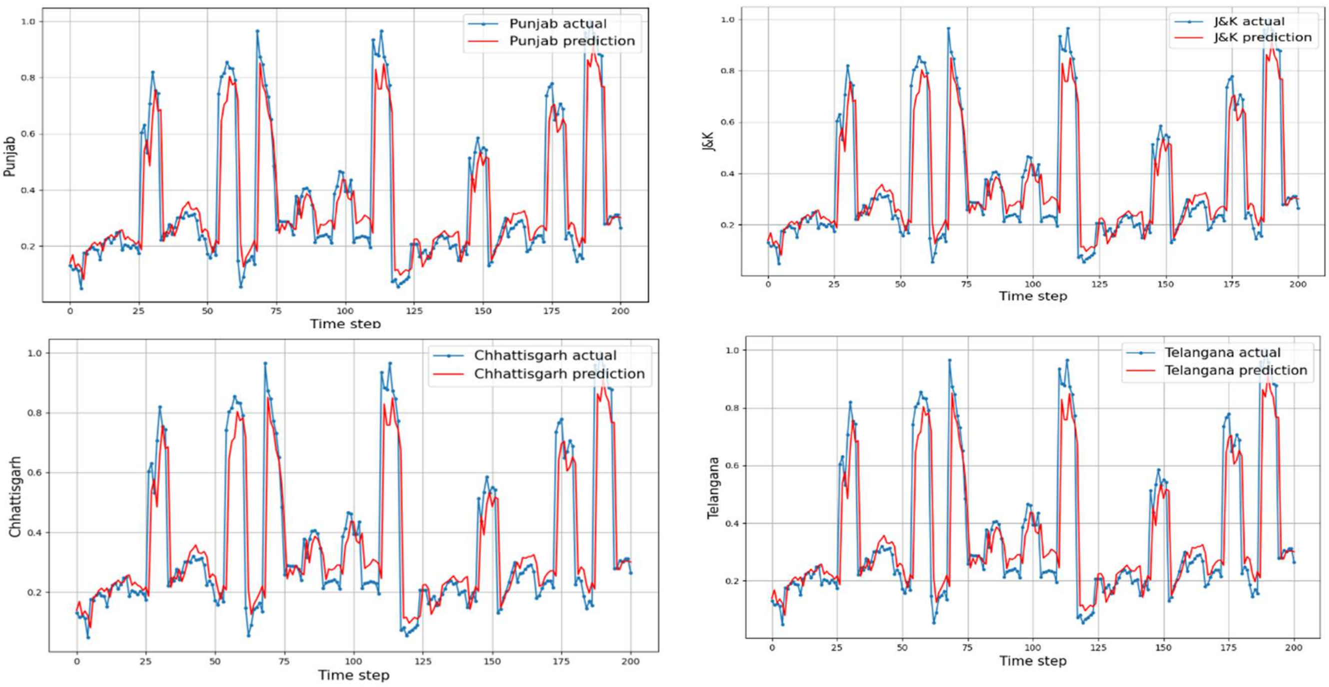

Fig. 7 Samples of the prediction results of 200 day power consumption

The proposed system reorders cities in the database by their energy consumption, from lowest to highest, as shown in Figure 8. This helps identify high- and low-energy-consuming cities, providing insights for future plans and policies to manage energy consumption effectively, especially in high-consumption cities.

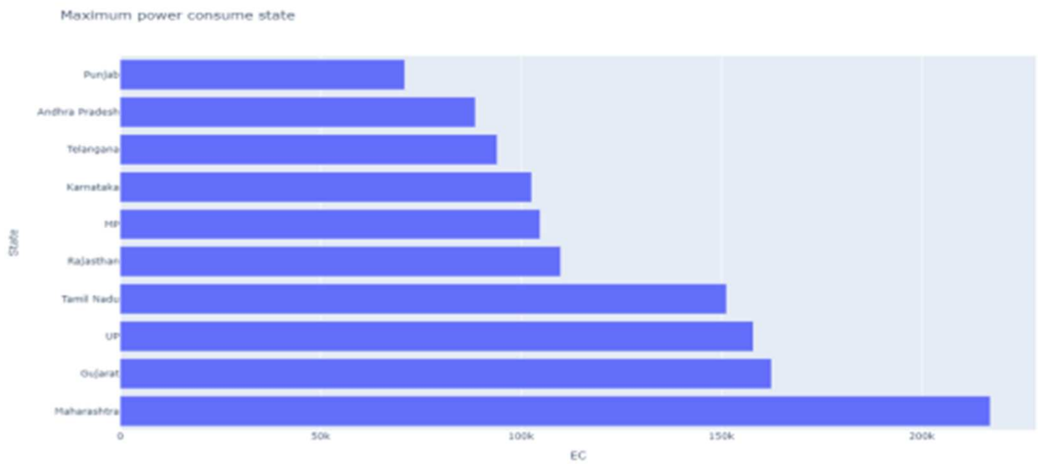

Fig. 8 Rearrange cities in the database based on their energy consumption

Figure (9) depicts a heat map of the 17 cities, allowing for the analysis of the correlation between them to determine which cities exhibit similar power consumption patterns and which are dissimilar. This analysis provides valuable insights into the clustering of cities in terms of their power consumption behavior.

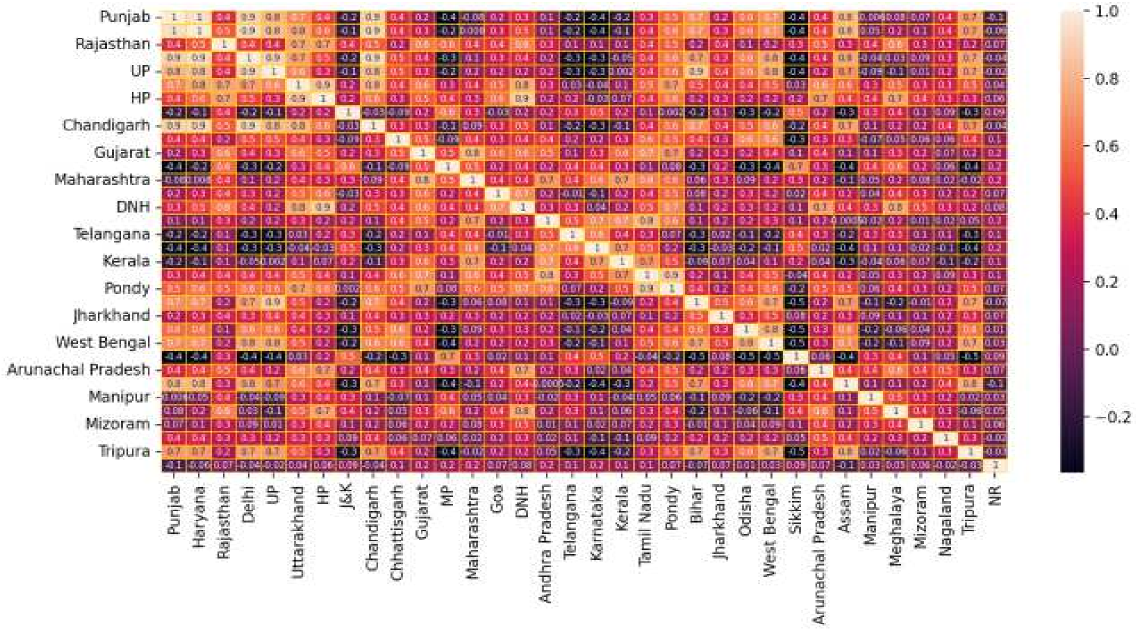

Fig. 9 Heatmap of the Power Consumption for 17 Cities

In this study, the performance of the proposed Hybrid LSTM+GRU model and the original LSTM model is primarily analyzed using MSE and RMSE, as defined in Equations (15) and (16). The MAE and RMSE values are used to compare the hybrid and original models, as illustrated in Figure (10).

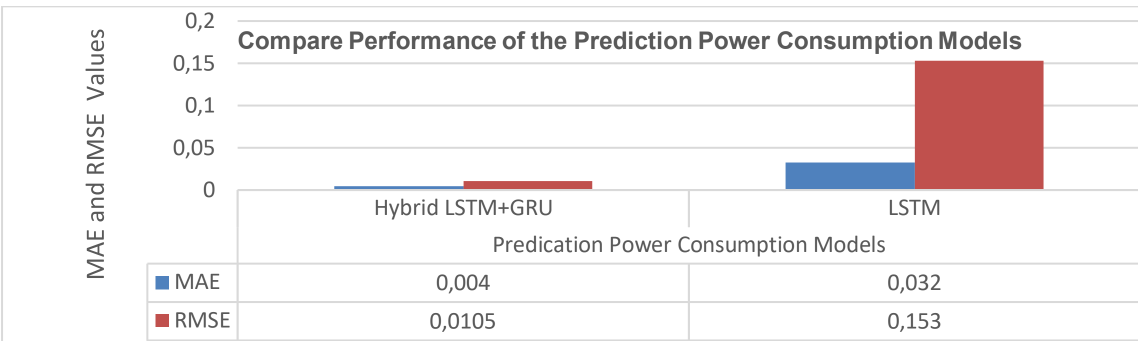

Fig. 10 Comparison Performance between Hybrid and Original Models

The proposed model of hybrid LSTM+GRU has demonstrated superior efficiency compared to the original LSTM model, with a significantly lower prediction error. The hybrid model achieved an impressive MAE value of around 0.004 and an RMSE value of approximately 0.0105, whereas the original LSTM model had an MAE value of about 0.032 and an RMSE value of approximately 0.57. These results highlight the potential of the hybrid model in accurately predicting power consumption and its potential for practical applications in energy management and conservation.

Regarding the prediction technique, dataset name, and MAE and RMSE values. Comparisons were made between the recommended model's performance with several relevant studies, as shown in Table (2). The comparison shows that the LSTM-GRU-PPCM had the lowest MAE and RMSE error value.

TABLE II
COMPARED PERFORMANCE OF THE PROPOSED LSTM-GRU-PPCM WITH SOME RELATED WORKS

Number of reference	**Prediction approach**	**Prediction Method**	**Dataset name**	**Results**	
				MAE	**RMSE**
[15]	Deep Learning	CNN	Individual household electric power consumption dataset	Non	0.732
[16]	Deep Learning	CNN+LSTM	Five real-world household power consumption datasets	0.0058	0.0112
[20]	Deep Learning	CNN	Appliances load prediction (AEP) dataset	0.01	0.02
[6]	Deep Learning	LSTM, GRU and Drop-GRU	A section of Péronne City's electricity usage	0.0526	0.7195
Our model	Deep Learning	Hybrid LSTM+GRU	Power consumption in India with 17 cities	0.004	0.01

IV. CONCLUSION

The prediction model of power consumption in smart cities using a hybrid deep learning algorithm is a research project that aims to develop a highly accurate and efficient model for estimating and managing power consumption in smart cities. The model successfully predicts power consumption over a specified period using a hybrid deep learning method combining the advantages of (LSTM) and (GRU). The suggested model first processed the input data and then conducted data feature normalization before creating the proper network structure to enhance ability prediction. The suggested hybrid model is then compared against other models. A collection of Indian power consumption data from 17 cities was used to develop and test the suggested LSTM-GRU-PPCM, and the results showed that the hybrid LSTM+GRU was more effective than the original LSTM. With the help of prior consumption data, the hybrid LSTM+GRU model is well suited for estimating power consumption over a certain period, essential for estimating consumption peaks and choosing the best course of action for load shedding in advance. Future studies will try to increase the accuracy and speed of hybrid models by considering extraneous elements like weather and holiday information. The goal is to identify high-consumption hotspots that go beyond the permitted limit and safeguard the power grid in smart cities.

REFERENCES

[1] N. A. Jasim, H. TH, and S. A. L. Rikabi, "Design and Implementation of Smart City Applications Based on the Internet of Things.," *Int. J. Interact. Mob. Technol.*, vol. 15, no. 13, 2021.

[2] S. Je and J. Huh, "Estimation of future power consumption level in smart grid: Application of fuzzy logic and genetic algorithm on big data platform," *Int. J. Commun. Syst.*, vol. 34, no. 2, p. e4056, 2021.

[3] S. Tiwari *et al.*, "Machine learning-based model for prediction of power consumption in smart grid-smart way towards smart city," *Expert Syst.*, vol. 39, no. 5, p. e12832, 2022.

[4] Z. Wang, M. Ogbodo, H. Huang, C. Qiu, M. Hisada, and A. Ben Abdallah, "AEBIS: AI-enabled blockchain-based electric vehicle integration system for power management in smart grid platform," *IEEE Access*, vol. 8, pp. 226409–226421, 2020.

[5] M. Zekić-Sušac, S. Mitrović, and A. Has, "Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities," *Int. J. Inf. Manage.*, vol. 58, p. 102074, 2021.

[6] S. Mahjoub, L. Chrifi-Alaoui, B. Marhic, and L. Delahoche, "Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks," *Sensors*, vol. 22, no. 11, p. 4062, 2022.

[7] M. Humayun, M. S. Alsaqer, and N. Jhanjhi, "Energy optimization for smart cities using iot," *Appl. Artif. Intell.*, vol. 36, no. 1, p. 2037255, 2022.

[8] S. Mahjoub, S. Labdai, L. Chrifi-Alaoui, B. Marhic, and L. Delahoche, "Short-Term Occupancy Forecasting for a Smart Home Using Optimized Weight Updates Based on GA and PSO Algorithms for an LSTM Network," *Energies*, vol. 16, no. 4, p. 1641, 2023.

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

[10] K. Chen, "APSO-LSTM: an improved LSTM neural network model based on APSO algorithm," in *Journal of Physics: Conference Series*, IOP Publishing, 2020, p. 12151.

[11] M. Sajjad *et al.*, "A novel CNN-GRU-based hybrid approach for short-term residential load forecasting," *IEEE Access*, vol. 8, pp. 143759–143768, 2020.

[12] M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption—A systematic review," *Eng. Appl. Artif. Intell.*, vol. 115, p. 105287, 2022.

[13] M. G. M. Almihat, M. T. E. Kahn, K. Aboalez, and A. M. Almaktoof, "Energy and Sustainable Development in Smart Cities: An Overview," *Smart Cities*, vol. 5, no. 4, pp. 1389–1408, 2022.

[14] Y. Kaluarachchi, "Implementing data-driven smart city applications for future cities," *Smart Cities*, vol. 5, no. 2, pp. 455–474, 2022.

[15] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *2017 IEEE 26th international symposium on industrial electronics (ISIE)*, IEEE, 2017, pp. 1483–1488.

[16] K. Yan, X. Wang, Y. Du, N. Jin, H. Huang, and H. Zhou, "Multi-step short-term power consumption forecasting with a hybrid deep learning strategy," *Energies*, vol. 11, no. 11, p. 3089, 2018.

[17] X. Wang, T. Zhao, H. Liu, and R. He, "Power consumption predicting and anomaly detection based on long short-term memory neural network," in *2019 IEEE 4th international conference on cloud computing and big data analysis (ICCCBDA)*, *IEEE*, 2019, pp. 487–491.

[18] N. Al Khafaf, M. Jalili, and P. Sokolowski, "Application of deep learning long short-term memory in energy demand forecasting," in *International conference on engineering applications of neural networks*, Springer, 2019, pp. 31–42.

[19] J. Q. Wang, Y. Du, and J. Wang, "LSTM based long-term energy consumption prediction with periodicity," *energy*, vol. 197, p. 117197, 2020.

[20] Z. A. Khan, A. Ullah, W. Ullah, S. Rho, M. Lee, and S. W. Baik, "Electrical energy prediction in residential buildings for short-term horizons using hybrid deep learning strategy," *Appl. Sci.*, vol. 10, no. 23, p. 8634, 2020.

[21] Z. Lin, L. Cheng, and G. Huang, "Electricity consumption prediction based on LSTM with attention mechanism," *IEEJ Trans. Electr. Electron. Eng.*, vol. 15, no. 4, pp. 556–562, 2020.

[22] N. Somu, G. R. MR, and K. Ramamritham, "A deep learning framework for building energy consumption forecast," *Renew. Sustain. Energy Rev.*, vol. 137, p. 110591, 2021.

[23] R. Markovic, E. Azar, M. K. Annaqeeb, J. Frisch, and C. van Treeck, "Day-ahead prediction of plug-in loads using a long short-term memory neural network," *Energy Build.*, vol. 234, p. 110667, 2021.

[24] C. Zhou, Z. Fang, X. Xu, X. Zhang, Y. Ding, and X. Jiang, "Using long short-term memory networks to predict energy consumption of air-conditioning systems," *Sustain. Cities Soc.*, vol. 55, p. 102000, 2020.

[25] B. Nettasinghe, S. Chatterjee, R. Tipireddy, and M. M. Halappanavar, "Extending Conformal Prediction to Hidden Markov Models with Exact Validity via de Finetti's Theorem for Markov Chains," in *International Conference on Machine Learning*, PMLR, 2023, pp. 25890–25903.

[26] M. Bilgili, N. Arslan, A. Şekertekin, and A. Yaşar, "Application of long short-term memory (LSTM) neural network based on deeplearning for electricity energy consumption forecasting," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 30, no. 1, pp. 140–157, 2022.

[27] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *International conference on machine learning*, PMLR, 2015, pp. 2342–2350.

[28] L. Peng, L. Wang, D. Xia, and Q. Gao, "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory," *energy*, vol. 238, p. 121756, 2022.

[29] H. Ikhlasse, D. Benjamin, C. Vincent, and M. Hicham, "Multimodal cloud resources utilization forecasting using a Bidirectional Gated Recurrent Unit predictor based on a power efficient Stacked denoising Autoencoders," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 11565–11577, 2022.

[30] X. Li, X. Ma, F. Xiao, C. Xiao, F. Wang, and S. Zhang, "Time-series production forecasting method based on the integration of Bidirectional Gated Recurrent Unit (Bi-GRU) network and Sparrow Search Algorithm (SSA)," *J. Pet. Sci. Eng.*, vol. 208, p. 109309, 2022.

[31] S. Jung, J. Moon, S. Park, and E. Hwang, "An attention-based multilayer GRU model for multistep-ahead short-term load forecasting," *Sensors*, vol. 21, no. 5, p. 1639, 2021.

[32] S. Agarwal, "Data mining: Data mining concepts and techniques," in *2013 international conference on machine intelligence and research advancement*, *IEEE*, 2013, pp. 203–207.

[33] F. Soldan, A. Maldarella, G. Paludetto, E. Bionda, F. Belloni, and S. Grillo, "Characterization of electric consumers through an automated clustering pipeline," in *2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, IEEE, 2022, pp. 1–5.

[34] R. Chiosa, M. S. Piscitelli, C. Fan, and A. Capozzoli, "Towards a self-tuned data analytics-based process for an automatic context-aware detection and diagnosis of anomalies in building energy consumption timeseries," *Energy Build.*, vol. 270, p. 112302, 2022.

[35] Q. Yuan, Y. Pi, L. Kou, F. Zhang, Y. Li, and Z. Zhang, "Multi-source data processing and fusion method for power distribution internet of things based on edge intelligence," *Front. Energy Res.*, vol. 10, p. 891867, 2022.

[36] I. Ullah, K. Liu, T. Yamamoto, R. E. Al Mamlook, and A. Jamal, "A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability," *Energy Environ.*, vol. 33, no. 8, pp. 1583–1612, 2022.

[37] I. Atik, "A New CNN-Based Method for Short-Term Forecasting of Electrical Energy Consumption in the Covid-19 Period: The Case of Turkey," *IEEE Access*, vol. 10, pp. 22586–22598, 2022.

[38] Y. Li, Z. Zhu, D. Kong, H. Han, and Y. Zhao, "EA-LSTM: Evolutionary attention-based LSTM for time series prediction," *Knowledge-Based Syst.*, vol. 181, p. 104785, 2019.

[39] F. Sommer and M. Stuke, "An efficient and fast method to calculate integral experimental correlation coefficients–S2Cor," *Ann. Nucl. Energy*, vol. 157, p. 108209, 2021.