EXPLORING TIME SERIES ANALYSIS OF RESIDENTIAL ELECTRICAL POWER CONSUMPTION

Maheshwari[1], Chaithanya[2], D. Poojitha[2], D. Pranaya[2] , N. Deepika[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Science Engineering
[1,2]Malla Reddy Engineering College for Women, Maisammaguda, Dhulapally, Kompally, Secunderabad-500100, Telangana, India

ABSTRACT

Residential electricity consumption has become increasingly important as our global population grows and urbanization continues. The rising demand for electricity in households makes it crucial to understand and analyze the patterns of residential power usage. This analysis is essential for energy providers, policymakers, and individuals alike, as it allows us to optimize energy usage, improve efficiency, and make informed decisions about how we consume energy. In the past, analyzing time series data of residential power consumption involved using basic statistical methods and manually examining the data. But now, we aim to explore and analyze this data in a more comprehensive and sophisticated manner. This project focuses on delving into the time series data of residential electricity consumption to gain valuable insights. The exploration of time series analysis in residential electrical power consumption is vital for several reasons. Firstly, it helps us become more energy-efficient by identifying opportunities to adopt energy-saving practices and utilize new technologies based on consumption patterns. Secondly, it aids in load management and forecasting, ensuring that utility companies can efficiently handle supply and demand to prevent blackouts and brownouts. Moreover, by predicting peak demand, we can optimize energy generation and distribution, reducing our reliance on expensive peak-load power plants. Time series analysis also plays a significant role in developing better demand response strategies. This encourages consumers to adjust their electricity usage during peak hours, helping to balance the grid and ensure stable energy distribution. As we transition to renewable energy sources, exploring time series data becomes even more crucial. It enables us to align our energy consumption with the intermittent nature of renewable energy generation, fostering more sustainable practices. Additionally, accurate billing and efficient tariff design can be achieved by gaining a deeper understanding of consumption patterns. This benefits both consumers and energy providers, as it allows for fair and cost-effective billing.

Keywords: Electrical Power Consumption, Time Series Analysis,

1. INTRODUCTION

The historical narrative of residential electrical power consumption is a captivating one, characterized by transformations in how electricity is harnessed and utilized within homes [1]. It includes the advent of electric appliances, the expansion of electrical grids, and major technological breakthroughs that have made power readily accessible to households worldwide [2]. Additionally, regulatory changes and shifts toward sustainable energy sources have played pivotal roles in shaping the landscape of residential power consumption over the years [3, 4]. To comprehend the intricacies of residential electrical power consumption, it is essential to trace its historical development. From the early days of electrification when electricity began powering households to the present era of advanced metering systems and smart grids, the evolution of this domain has been marked by significant milestones. Key concepts, such as time series data, kilowatt-hours (kWh), load profiles, and metering systems, form the foundational knowledge necessary for this exploration [5]. The exploration of residential electrical power consumption through time series analysis is a critical endeavor in the context of modern society's increasing dependence on electricity. As we rely on electrical energy for various aspects of our daily lives, understanding and analyzing how power is consumed in residential settings have become paramount [6]. This research embarks on a journey to delve deeper into the patterns, trends, and dynamics of residential electrical power consumption, with the ultimate goal of uncovering insights that can drive energy conservation, cost reduction, and environmental sustainability [7].

Therefore, this research underscores the pressing need for a thorough examination of residential electrical power consumption. In today's data-driven world, making informed decisions is paramount, and the availability of comprehensive data on power usage in households offers immense potential. Moreover, the need for enhanced energy efficiency is becoming increasingly evident, with residential areas being a focal point for reducing energy waste. Additionally, environmental concerns regarding excessive energy consumption and its implications for climate change highlight the urgency of this endeavor.

2. LITERATURE SURVEY

The research on the prediction of building energy consumption began in the 1970s, when an energy crisis forced countries to start thinking about ways to cut their energy consumption and carbon emissions. The early-developed models of building energy consumption prediction relied on the use of simplified calculation methods that were empirical models based on extensive engineering practices, allowing the estimates to be performed at the early stages of building design to guide the relevant design work. However, it was recognized that simplified calculation methods were not able to adequately capture the dynamicity and complexity of the environment. To tackle this problem, scholars in the mid-1980s started to adopt statistical methods for predicting building energy consumption. Since then, significant progress has been made in the field of building energy consumption prediction. Nowadays, the three most popular methods for predicting energy consumption in buildings include engineering simplification, physical modeling, and ML-based methods.

In [10], Kim, et al. explored the use of Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) neural networks to predict residential energy consumption patterns. The research leverages time series data to train the model and utilizes the CNN-LSTM architecture for its ability to capture both spatial and temporal dependencies within the data. Results indicate promising predictive capabilities, offering valuable insights for energy management and conservation efforts in residential settings.

In [11], Lago, et al. investigated deep learning approaches for forecasting spot electricity prices, aiming to improve accuracy and reliability compared to traditional algorithms. The study employs empirical comparisons to assess the performance of these approaches. Findings demonstrate the potential of deep learning in enhancing spot price forecasting in the energy sector.Fan, et al. [12] studied a comprehensive statistical analysis of the driving factors affecting residential energy demand in the Greater Sydney region of Australia. The research identifies and quantifies the key factors influencing energy consumption patterns, contributing valuable insights for energy policy and demand-side management strategies. However, the findings are specific to the Greater Sydney region, and the applicability of the results to other geographical areas may require further research and validation.

3. PROPOSED SYSTEM

Overview

This project revolves around the exploration and analysis of residential electrical power consumption patterns using a specialized type of neural network called Long Short-Term Memory (LSTM). It begins by diligently handling the data, which is initially loaded from a file. Data preprocessing is carried out meticulously, which involves addressing missing values and augmenting the dataset with additional features related to time. Subsequently, the project delves into data transformation, where the actual power consumption values are rescaled to fit within a normalized range. This scaling is essential for the subsequent modeling phase, ensuring that the neural network can effectively learn from the data.

The heart of the project lies in the development of the LSTM neural network. This neural network architecture is specifically tailored for handling sequential data, making it ideal for time series analysis. The model is meticulously crafted, trained on historical power consumption data, and fine-tuned to predict future consumption patterns based on past trends and observations. To gauge the model's effectiveness, thorough evaluation is conducted. Performance metrics, such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), are employed to assess the model's accuracy in predicting power consumption. This evaluation is carried out on both the training and testing datasets, ensuring a comprehensive understanding of the model's capabilities.

Finally, the project offers a visual representation of the model's performance. It showcases the training loss over epochs, providing insights into the learning process. Additionally, the project generates visualizations that compare the model's predictions with actual power consumption data for a specific time frame. These visualizations allow stakeholders to grasp how effectively the model captures the intricate patterns of residential power consumption. In essence, this project aims to harness the power of LSTM neural networks to gain valuable insights into residential power consumption patterns. By doing so, it opens doors to more accurate forecasting, benefiting both energy providers and consumers by optimizing energy usage and enhancing efficiency.

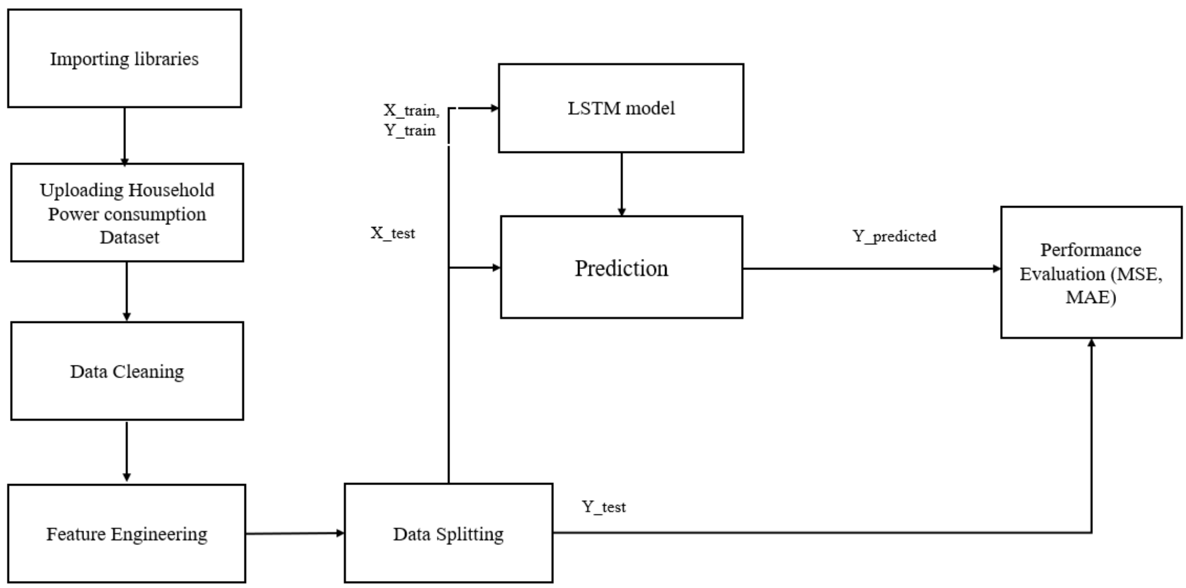

Figure .1: Overall design of proposed methodology.

Data Preprocessing

Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So, for this, we use data pre-processing task. A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

- Getting the dataset
- Importing libraries
- Importing datasets
- Finding Missing Data
- Encoding Categorical Data
- Splitting dataset into training and test set

LSTM Network

Deep learning is a new research direction in the field of artificial intelligence. It is developed on the basis of shallow neural networks with the improvement of computer hardware levels and the explosive growth of the current data volume. Deep learning and shallow neural network structure are both layered. Each layer will process the data input to the model and combine low-level features into potential high-level features by learning data rules. Compared with shallow models, deep learning can express complex high dimensionality such as high-variable functions and find the true relationships within the original data better. In the 1980s, artificial neural network back propagation algorithm was born. This method can automatically learn data rules from a large amount of training data without manual intervention. At present, deep learning is the most concerned research direction in the field of artificial intelligence, which completely subverts the shallow model in traditional machine, proposes a deep learning network model, and elevates it to a new height from theory to application. CNN (convolutional neural network) and RNN are two types of classical deep learning network structures now.

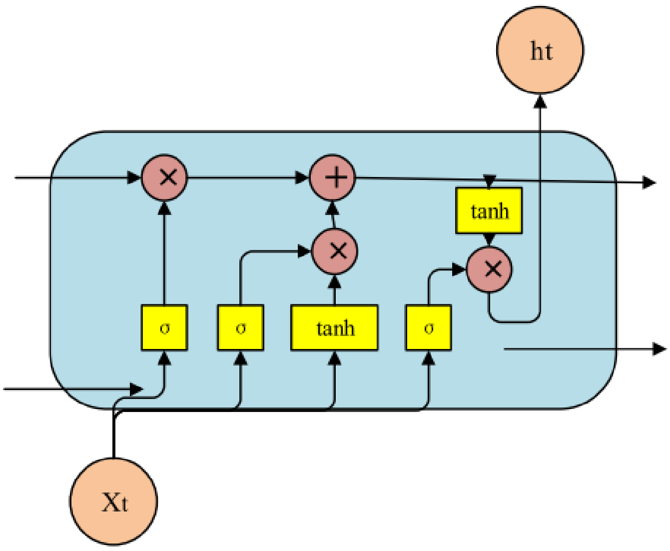

Fig.2: LSTM model structure.

Because there are connections between neurons in the RNN layer, the network can learn the change law of sequence data before and after, and the internal sequence rules of data is easy to be mined. Thus RNN is widely used in the field of sequence data processing such as speech recognition and machine translation. However, this structure also has some problems. When data is transmitted backward, the problem of gradient disappearance or gradient explosion is unavoidable, which limits its processing of long-term dependencies. The LSTM network changes the way of gradient transmission during backpropagation by adding multiple special computing nodes in the hidden layer of RNN, which effectively slows the problem of gradient disappearance or gradient explosion. Its model structure is shown in figure 4.

Where h_{t-1} represents the output of the previous cell, and x_t represents the input of the current cell. σ represents the sigmoid function. The difference between LSTM and RNN is that it adds a ‘‘processor’’ to the algorithm to determine the usefulness of the information. The structure of this processor is called a cell. Three gates are placed in a cell, which are called *Input gate*, *Forget gate*, and *Output gate*. A piece of information enters the LSTM network, and it can be judged whether it is useful according to the rules. Only the information that meets the algorithm authentication will be left, and the non-conforming information will be forgotten through the *Forget gate*.

Advantages

Long Short-Term Memory (LSTM) models offer several advantages in for time series analysis of residential electrical power consumption:

- Sequential Data Handling: LSTMs are specifically designed to handle sequential data, making them well-suited for time series analysis. In the project, LSTMs can capture the temporal dependencies and patterns in power consumption data more effectively compared to traditional machine learning models.
- Long-Term Dependencies: LSTMs are capable of capturing long-term dependencies in time series data. They can learn from historical data over extended periods, which is crucial for understanding seasonal or yearly patterns in residential power consumption. This capability helps in making more accurate predictions.
- Variable Sequence Length: LSTMs can accommodate variable-length sequences, making them adaptable to datasets where the number of time steps may vary. In the project, this flexibility allows the model to consider different look-back periods when making predictions.
- Feature Engineering: LSTMs do not require extensive feature engineering. They can automatically learn relevant features from the sequential data, reducing the need for manual feature selection and engineering, which can be time-consuming.
- Handling Noisy Data: LSTMs can handle noisy time series data effectively. They are robust to outliers and missing values, which is essential when working with real-world data that may have irregularities or data gaps.
- Parallel Processing: LSTMs can be trained efficiently on modern hardware with parallel processing capabilities, enabling faster training times. This is valuable when dealing with large datasets or complex models.
- Model Interpretability: While LSTMs are often considered "black-box" models, efforts can be made to interpret their internal workings. For instance, techniques such as attention mechanisms can be added to LSTMs to provide insights into which parts of the input data are most influential in making predictions.
- Transfer Learning: Pre-trained LSTM models can be leveraged for time series analysis. Transfer learning allows the use of models trained on similar data or domains, potentially reducing the amount of data required for training and improving model performance.
- Real-Time Predictions: LSTMs can be used for real-time predictions, making them valuable for applications where timely forecasting of power consumption is critical. This can aid in proactive load management and energy optimization.

4. RESULTS AND DISCUSSION

Results description

Figure 3 depicts a representation of the dataset used for the analysis of residential electrical power consumption. It shows all the feature attributes such as the data described in dataset description. Figure 4 shows how the date and time information from the dataset are combined into a single format. For time series analysis, it's common to represent the timestamp in a consistent format, such as yyyy-mm-dd hh:mm:ss. This figure illustrates the transformation of the original date and time data into this standardized format

	Date	Time	Global_active_power	Global_reactive_power	Voltage	Global_intensity	Sub_metering_1	Sub_metering_2	Sub_metering_3
0	16/12/2006	17:24:00	4.216	0.418	234.84	18.4	0.0	1.0	17.0
1	16/12/2006	17:25:00	5.360	0.436	233.63	23.0	0.0	1.0	16.0
2	16/12/2006	17:26:00	5.374	0.498	233.29	23.0	0.0	2.0	17.0
3	16/12/2006	17:27:00	5.388	0.502	233.74	23.0	0.0	1.0	17.0
4	16/12/2006	17:28:00	3.666	0.528	235.68	15.8	0.0	1.0	17.0

Figure 3: Illustration of the sample dataset used for exploring time series analysis of residential electrical power consumption.

	date_time	Global_active_power	year	quarter	month	day	weekday
0	2006-12-16 17:24:00	4.216	2006	4	12	16	0
1	2006-12-16 17:25:00	5.360	2006	4	12	16	0
2	2006-12-16 17:26:00	5.374	2006	4	12	16	0
3	2006-12-16 17:27:00	5.388	2006	4	12	16	0
4	2006-12-16 17:28:00	3.666	2006	4	12	16	0
...	...	...	...	...	...	...	...
5190	2006-12-20 07:54:00	2.532	2006	4	12	20	1
5191	2006-12-20 07:55:00	2.522	2006	4	12	20	1
5192	2006-12-20 07:56:00	2.832	2006	4	12	20	1
5193	2006-12-20 07:57:00	3.050	2006	4	12	20	1
5194	2006-12-20 07:58:00	2.982	2006	4	12	20	1

5195 rows × 7 columns

Figure 4: Representing the combining the date and time in one format.

```
Epoch 1/50
4125/4125 [==============================] - 1s 258us/sample - loss: 0.0629 - val_loss: 0.0109
Epoch 2/50
4125/4125 [==============================] - 0s 17us/sample - loss: 0.0239 - val_loss: 0.0048
Epoch 3/50
4125/4125 [==============================] - 0s 13us/sample - loss: 0.0117 - val_loss: 0.0057
Epoch 4/50
4125/4125 [==============================] - 0s 13us/sample - loss: 0.0143 - val_loss: 0.0068
Epoch 5/50
4125/4125 [==============================] - 0s 14us/sample - loss: 0.0154 - val_loss: 0.0058
Epoch 6/50
4125/4125 [==============================] - 0s 14us/sample - loss: 0.0124 - val_loss: 0.0044
Epoch 7/50
4125/4125 [==============================] - 0s 13us/sample - loss: 0.0101 - val_loss: 0.0039
Epoch 8/50
4125/4125 [==============================] - 0s 14us/sample - loss: 0.0100 - val_loss: 0.0039
Epoch 9/50
4125/4125 [==============================] - 0s 13us/sample - loss: 0.0103 - val_loss: 0.0038
Epoch 10/50
```

.

Table 1 provides error metrics for the LSTM algorithm, which are commonly used to evaluate the performance of predictive models for time series analysis. The table has two rows: one for the training data and one for the test data. The error metrics typically include:

- MAE (Mean Absolute Error): This measures the average absolute difference between the predicted values and the actual values. A lower MAE indicates better model performance.
- RMSE (Root Mean Square Error): This calculates the square root of the average of the squared differences between predicted and actual values. RMSE is another measure of prediction accuracy, with lower values indicating better performance.

Table 1: Error Metrics for LSTM algorithm.

Metric	MAE	RMSE
Train data	0.3127457734676216	0.542399620555129
Test data	0.1620324041018154	0.31970784968084065

Figure 6 displays two-line plots. One plot shows the loss (error) of the proposed LSTM model on the training dataset for each epoch, indicating how the training loss changes over time. The other plot shows the loss on the test dataset for each epoch, providing insights into how well the model generalizes to unseen data.

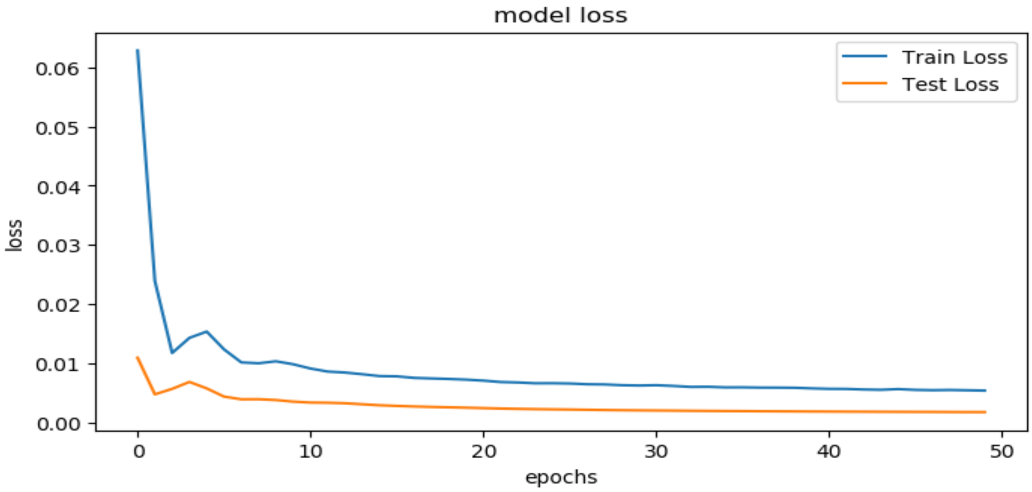

Figure 5: Loss performance of proposed LSTM model for each epoch on train and test dataset.

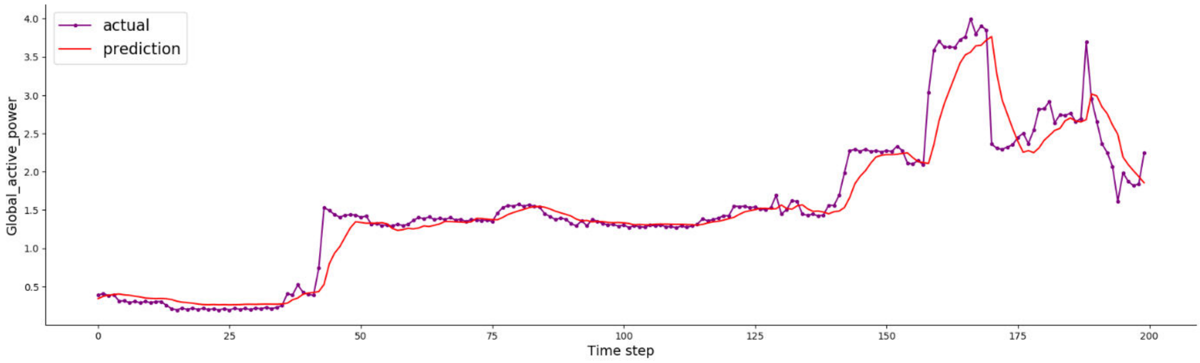

Figure 6: Performance of proposed LSTM model w.r.t. actual and prediction values of power consumption.

Figure 5 illustrates the architecture of a Long Short-Term Memory (LSTM) neural network used for time series analysis. It shows the layers within the LSTM network. Additionally, it includes the training performance over 50 epochs with the corresponding training accuracy, showing how well the model is fitting the training data over epochs and the training loss, indicating how the error or loss decreases during training

5. CONCLUSION AND FUTURE SCOPE

Residential electricity consumption has emerged as a critical aspect of modern living, driven by the growing global population and ongoing urbanization trends. The escalating demand for electricity in households necessitates a deeper understanding of how residential power is used. This understanding is not only vital for energy providers but also for policymakers and individuals, as it empowers us to optimize energy consumption, enhance efficiency, and make informed choices regarding our energy usage. Traditionally, the analysis of time series data related to residential power consumption relied on basic statistical techniques and manual examination of data. However, in our contemporary landscape, there is a compelling need to explore and analyze this data in a more sophisticated and comprehensive manner. This project is dedicated to diving into the rich reservoir of time series data pertaining to residential electricity consumption, with the aim of extracting valuable insights that can transform our energy landscape.The significance of time series analysis in the context of residential electrical power consumption is multifaceted. Firstly, it serves as a beacon for energy efficiency by identifying opportunities for adopting energy-saving practices and incorporating cutting-edge technologies based on consumption patterns. Secondly, it facilitates robust load management and forecasting, equipping utility companies to efficiently manage supply and demand dynamics, thus averting the specter of blackouts and brownouts. Furthermore, by predicting peak demand, we can optimize energy generation and distribution, thereby reducing reliance on costly peak-load power plants. Time series analysis also pioneers the development of more effective demand response strategies, motivating consumers to adjust their electricity consumption during peak hours. This not only aids in balancing the grid but also ensures stable energy distribution. As we transition towards renewable energy sources, the exploration of time series data becomes even more imperative. It allows us to harmonize our energy consumption with the intermittent nature of renewable energy generation, fostering sustainable practices. Additionally, accurate billing and the design of efficient tariff structures are achievable through an in-depth comprehension of consumption patterns. This benefits both consumers and energy providers, as it facilitates fair and cost-effective billing, reinforcing the equitable distribution of energy resources.

REFERENCES

[1] Yu, J.; Chang, W.-S.; Dong, Y. Building energy prediction models and related uncertainties: A review. Buildings 2022, 12, 1284.

[2] Wang, Y. Application of deep learning model in building energy consumption prediction. Comput. Intell. Neurosci. 2022, 9, 4835259.

[3] Jiang, F.; Ma, J.; Li, Z.; Ding, Y. Prediction of energy use intensity of urban buildings using the semi-supervised deep learning model. Energy 2022, 249, 123631.

[4] Olu-Ajayi, R.; Alaka, H.; Sulaimon, I.; Sunmola, F.; Ajayi, S. Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. J. Build. Eng. 2022, 45, 103406.

[5] Jiang, F.; Ma, J.; Li, Z.; Ding, Y. Prediction of energy use intensity of urban buildings using the semi-supervised deep learning model. Energy 2022, 249, 123631.

[6] Dinmohammadi, F.; Han, Y.; Shafiee, M. Predicting Energy Consumption in Residential Buildings Using Advanced Machine Learning Algorithms. Energies 2023, 16, 3748. https://doi.org/10.3390/en16093748

[7] IEA. World Energy Outlook 2019; IEA: Paris, France, 2019; Available online: http://www.iea.org/reports/world-energy-outlook-2019 (accessed on 13 November 2019).
[8] Nejat, P.; Jomehzadeh, F.; Taheri, M.M.; Gohari, M.; Majid, M.Z.A. A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries). Renew. Sustain. Energy Rev. 2015, 43, 843–862.
[9] Deb, C.; Zhang, F.; Yang, J.; Lee, S.E.; Shah, K.W. A review on time series forecasting techniques for building energy consumption. Renew. Sustain. Energy Rev. 2017, 74, 902–924.
[10] Kim, T.-Y.; Cho, S.-B. Predicting residential energy consumption using CNN-LSTM neural networks. Energy 2019, 182, 72–81.
[11] Lago, J.; De Ridder, F.; De Schutter, B. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Appl. Energy 2018, 221, 386–405.
[12] Fan, H.; MacGill, I.; Sproul, A. Statistical analysis of driving factors of residential energy demand in the greater Sydney region, Australia. Energy Build. 2015, 105, 9–25.