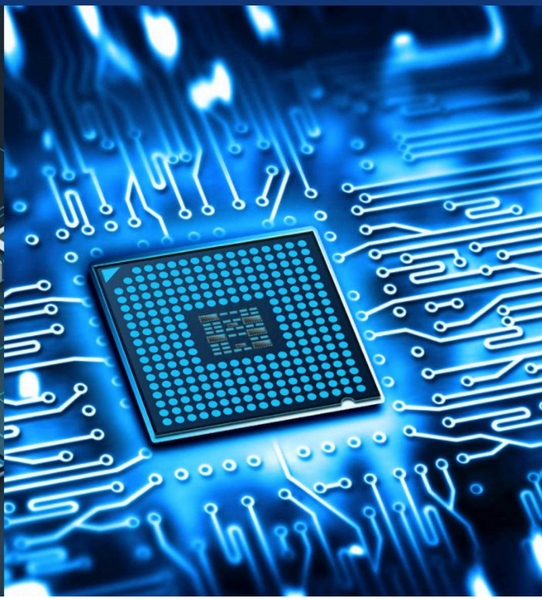
IJIRCCE

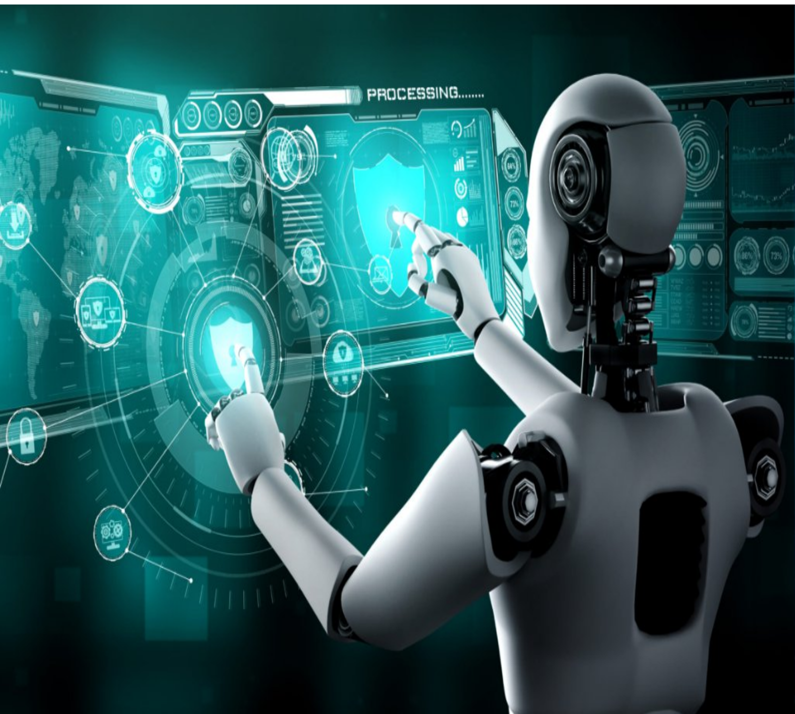
ISSN(O): 2320-9801

ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 14, Issue 4, April 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404021
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Energy Consumption Forecasting and Visualization for Smart Homes Using Data Analytics and Machine Learning

G.Nirmala[1], B. Deenaja[2], B. Trilok[3], B. Bandhavi[4], B. Praveen[5], B. Rama Jyothi[6]

Professor, Dept. of CSE, Sir CR Reddy College of Engineering, Eluru, India[1]

B. Tech Student, Dept. of CSE, Sir CR Reddy College of Engineering, Eluru, India[2,3,4,5,6]

ABSTRACT: Energy consumption forecasting is a critical component of smart home energy management systems, enabling efficient energy utilization and sustainable resource planning. This study presents a data-driven approach for forecasting electricity consumption using machine learning and deep learning techniques. A univariate time-series dataset comprising six years of hourly electricity consumption data is utilized for analysis.

Data preprocessing techniques, including resampling, feature engineering, and normalization, are applied to enhance model performance. Two predictive models—Random Forest Regressor and Long Short-Term Memory (LSTM) network are implemented and evaluated. The dataset is divided into training and testing sets using an 80:20 ratio, and model performance is assessed using metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).

Experimental results demonstrate that the LSTM model outperforms the Random Forest model by effectively capturing temporal dependencies and seasonal patterns in energy consumption data. The predicted results closely align with actual consumption trends, indicating high forecasting accuracy. The proposed approach can support smart grid systems, optimize energy distribution, and contribute to sustainable energy management in smart homes.

KEYWORDS: Energy Consumption Forecasting, Smart Homes, Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), Random Forest, Time Series Analysis, Data Analytics, Energy Management Systems, Electricity Consumption Prediction, Data Visualization, Smart Grid Systems.

I. INTRODUCTION

Energy consumption forecasting has become an important requirement in recent years due to the rapid growth in electricity demand and the increasing adoption of smart technologies. In modern smart homes and smart grid systems, efficient energy management depends on the ability to accurately predict future electricity usage. Without proper forecasting, energy providers may either generate excess power, leading to wastage, or produce insufficient energy, resulting in power shortages and instability in the system.

Traditional forecasting approaches mainly rely on statistical techniques such as linear regression and ARIMA. While these methods work reasonably well for simple and stable datasets, they are not effective in handling complex and dynamic patterns present in real-world electricity consumption data. Energy usage is influenced by multiple factors such as seasonal changes, daily routines, and varying consumption behaviours. These variations introduce nonlinear relationships and temporal dependencies that traditional models fail to capture accurately.

With the advancement of data analytics and machine learning, more intelligent approaches have been developed to address these challenges. Machine learning models can learn patterns directly from historical data and adapt to different consumption behaviours. In particular, deep learning techniques such as Long Short-Term Memory networks are capable of understanding sequential data and capturing long-term dependencies, making them suitable for time-series forecasting problems.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

This work focuses on predicting electricity consumption using a data-driven approach. The study makes use of historical electricity consumption data collected over several years and applies preprocessing techniques to prepare the data for analysis. Machine learning and deep learning models are then trained to identify patterns in the data and generate future consumption predictions.

Two models are considered in this work: Random Forest and Long Short-Term Memory. Random Forest is effective in handling nonlinear relationships and reducing overfitting, while LSTM is designed to model sequential data and capture temporal patterns. By analysing and comparing these models, the study aims to identify an effective approach for energy consumption forecasting.

The outcome of this work contributes to improving energy planning and management in smart environments. Accurate prediction of electricity consumption can support better decision-making for energy providers, help reduce operational costs, and promote efficient utilization of resources. This also plays a role in building sustainable energy systems and improving the reliability of smart grid infrastructure.

II. RELATED WORK

Energy consumption forecasting has been widely studied using statistical, machine learning, and deep learning approaches. Traditional models such as ARIMA rely on historical data and work well for linear patterns, but they are not effective in handling complex and nonlinear relationships [1].

Machine learning techniques like decision trees and Random Forest improve prediction by capturing nonlinear patterns in the data [2]. However, these models do not consider the sequential nature of time-series data, which limits their ability to model temporal dependencies.

To overcome this, deep learning models such as Long Short-Term Memory networks are used. LSTM can learn from sequential data and capture long-term dependencies and seasonal patterns, leading to better forecasting performance [3], [5], [6]. Despite their effectiveness, these models require large datasets and higher computational resources.

Recent studies focus on hybrid approaches that combine machine learning and deep learning models to improve accuracy [8]. However, such methods increase complexity and require careful tuning.

Based on these observations, this work combines Random Forest and LSTM to balance accuracy and efficiency, providing improved prediction performance while addressing the limitations of existing methods.

TABLE I: LITERATURE REVIEW

S No	Author(s) & Year	Method / Model Used	Theme / Approach	Key Findings	Limitations
1	J. Brownlee (2019) [1]	ARIMA, Statistical Models	Time-series forecasting using	Simple and effective for linear datasets	Cannot capture nonlinear patterns and
2	A. Geron (2019) [2]	Machine Learning Models (Decision Trees, RF)	historical data Data-driven prediction using ML techniques	Handles nonlinear relationships better than statistical models	complex seasonal variations Does not capture temporal dependencies effectively
3	M. Nachawati (2022) [3]	LSTM	Deep learning for sequential data forecasting	Captures long-term dependencies in time-series data	Requires large dataset and high computational cost
4	F. Chollet (2018) [5]	Deep Learning (Neural Networks)	Neural network-based prediction models	Learns complex patterns and improves prediction accuracy	Risk of overfitting and requires parameter tuning
5	M. Qureshi et al. (2024) [6]	Deep Learning Models	Advanced energy consumption forecasting	High accuracy in predicting energy demand	Computationally expensive and complex training

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

S No	Author(s) & Year	Method / Model Used	Theme / Approach	Key Findings	Limitations
6	C. V. S. Chandrika et al. (2024) [7]	LSTM-Based Models	Time-series energy prediction	Improved forecasting performance over traditional models	Needs large training data and longer training time
7	F. Zhou et al. (2024) [8]	Hybrid RF-LSTM	Combination of ML and DL models	Better accuracy by combining models	Increased complexity and tuning difficulty
8	S. Ul Rehman et al. (2025) [9]	Deep Learning Approaches	Smart building energy prediction	Effective for real-time energy forecasting	High computational requirements
9	Proposed Work	Random Forest + LSTM	Hybrid approach combining ML and DL	Captures both nonlinear and temporal patterns, improves accuracy	Reduced limitations compared to existing systems

TABLE I – LITERATURE REVIEW

III. PROPOSED SYSTEM

The proposed system aims to forecast electricity consumption using a combination of machine learning and deep learning techniques. It utilizes historical energy consumption data to identify patterns and generate accurate future predictions.

The dataset is collected from a reliable source [4] and pre-processed by converting timestamps, organizing data chronologically, and handling inconsistencies. Feature engineering is performed by extracting time-based features such as day, month, and weekday, along with lag features to capture past consumption patterns.

Two models are implemented: Random Forest and Long Short-Term Memory. Random Forest is used to model nonlinear relationships and provide stable predictions [2], while LSTM captures temporal dependencies in time-series data by learning from sequential patterns [5].

The dataset is split into training and testing sets for model development and evaluation. The performance of the models is measured using metrics such as Mean Squared Error, Root Mean Squared Error, and R-squared score.

By combining both models, the system improves prediction accuracy by capturing both nonlinear and temporal characteristics of electricity consumption, supporting efficient energy management.

A. DATASET INFORMATION:

Attribute	Description
Data Source	Fin grid Electricity Consumption Dataset [4]
Time Period	2016 – 2021
Total Records	Approximately 52,965
Frequency	Hourly
Data Type	Time-Series (Univariate)
Target Variable	Electricity Consumption (MWh)

TABLE II – DATASET INFORMATION

B. SYSTEM ARCHITECTURE: The working of the proposed system can be explained through the following step-by-step workflow, which corresponds to the system architecture diagram.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. **Data Collection**

The process begins with collecting electricity consumption data from a reliable source [4]. The dataset consists of time-stamped records that represent energy usage over a period of time. This historical data forms the foundation for training the forecasting models.

2. **Data Preprocessing**

Once the data is collected, it is processed to make it suitable for analysis. The timestamp column is converted into a standard datetime format, and the data is arranged in chronological order. Any inconsistencies or irrelevant entries are handled to ensure data quality.

3. **Feature Engineering**

In this step, additional features are created to improve model performance. Time-based features such as day, month, and weekday are extracted from the timestamp. Lag features are also generated using previous consumption values, which help the model understand past consumption patterns and dependencies.

4. **Data Transformation**

The dataset is further transformed by resampling the data, if required, to obtain meaningful intervals such as daily consumption. The data is then structured into input features and target variables for model training.

5. **Dataset Splitting**

The processed dataset is divided into training and testing sets. Typically, 80% of the data is used for training the models, while the remaining 20% is used for testing. This ensures that the model is evaluated on unseen data.

6. **Model Training– Random Forest**

The Random Forest model is trained using the training dataset. It builds multiple decision trees using different subsets of the data and combines their outputs to generate predictions. This helps in capturing nonlinear relationships in the data [2].

7. **Model Training– LSTM**

The LSTM model is trained using sequential data to capture temporal dependencies. The input data is structured into time steps, and the model learns patterns across different time intervals. This allows it to understand long-term trends in electricity consumption [5].

8. **Model Evaluation**

After training, both models are evaluated using the testing dataset. Performance metrics such as Mean Squared Error, Root Mean Squared Error, and R-squared score are used to measure how accurately the models predict electricity consumption.

9. **Prediction Generation**

The trained models are used to generate predictions for future electricity consumption. These predictions are based on learned patterns from historical data.

10. **Result Visualization**

The final step involves visualizing the results using graphs. Actual electricity consumption values are compared with predicted values to analyse model performance and understand prediction accuracy.

This step-by-step workflow clearly represents how the system processes data from input to final prediction, aligning with the system architecture diagram and demonstrating the complete functioning of the proposed model.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

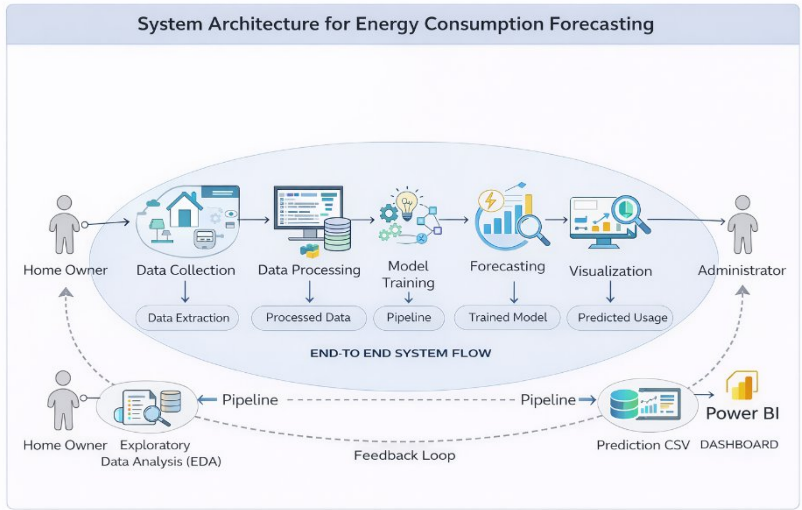


Fig: 1. System Architecture

C. LSTM MODEL : In this implementation, the LSTM layer is followed by a dropout layer. The purpose of the dropout layer is to reduce overfitting by randomly disabling a fraction of neurons during training. This improves the generalization capability of the model when applied to unseen data.

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Fig:2. RMSE formula

After the dropout layer, a dense (fully connected) layer is used to produce the final output. This layer takes the processed features from the LSTM layer and generates the predicted electricity consumption value.

The model is trained using a suitable loss function such as Mean Squared Error, which measures the difference between predicted and actual values. An optimizer like Adam is used to update the model weights during training. The training process continues for multiple epochs until the model learns the underlying patterns in the data.

Overall, the LSTM model structure enables effective learning of sequential patterns in electricity consumption data

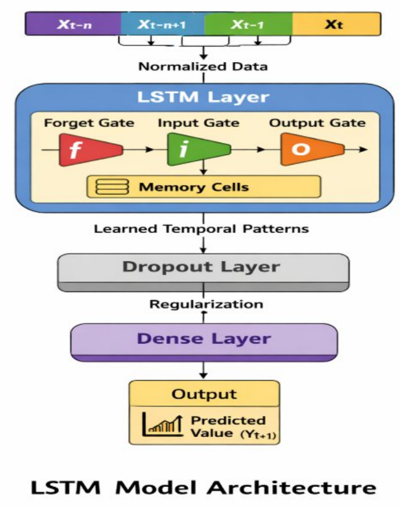


Fig: 3. LSTM MODEL ARCHITECTURE

IV. RESULTS AND EVALUATION

The analysis begins with visualizing the overall behaviour of the dataset. A time-series plot is used to represent electricity consumption over time, which helps in identifying trends, seasonality, and fluctuations in the data. This visualization provides an initial understanding of how energy usage varies across different periods.

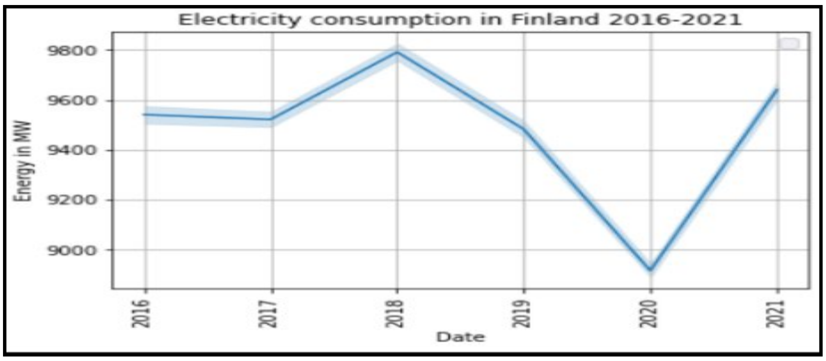
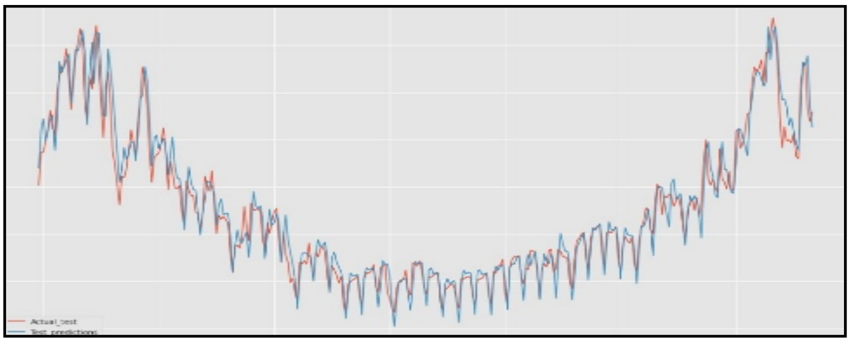


Fig:4. Time-series plot of electricity consumption



© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404021

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Fig: 5. Actual vs Predicted Electricity Consumption

V. DISCUSSION

The results indicate that both models are capable of learning patterns from historical electricity consumption data, but their performance differs based on how they handle temporal dependencies. Random Forest performs well in capturing nonlinear relationships and provides stable predictions, especially when the data does not have strong sequential dependencies. However, it does not fully utilize the time-based nature of the dataset.

In contrast, the LSTM model shows better performance in handling sequential data. It is able to capture long-term dependencies and seasonal variations more effectively, which leads to predictions that closely follow the actual consumption trends. The graphical analysis also supports this observation, where LSTM predictions align more closely with real values during fluctuations.

Overall, the comparison highlights that combining machine learning and deep learning approaches improves forecasting capability. The results are consistent with recent studies where LSTM-based and hybrid models outperform traditional methods in time-series forecasting tasks [6], [8].

VI. CONCLUSION

This work presented a data-driven approach for forecasting electricity consumption using Random Forest and LSTM models. The system was designed to analyse historical energy usage data and generate accurate predictions of future consumption.

The results demonstrate that the LSTM model provides better performance compared to Random Forest by effectively capturing temporal patterns in the data. The use of feature engineering and proper data preprocessing further improved the prediction accuracy.
The proposed approach can support efficient energy management in smart homes and smart grid systems by enabling better planning and utilization of resources. Overall, the study shows that integrating machine learning and deep learning techniques is a practical and effective solution for energy consumption forecasting.

REFERENCES

[1] J. Brownlee, “Time Series Forecasting with Machine Learning,” 2019.
Available: https://machinelearningmastery.com/time-series-forecasting-methods-in-python/
[2] A. Geron, “Hands-On Machine Learning with Scikit-Learn and TensorFlow,” O’Reilly Media, 2019.
[3] M. Nachawati, “Energy Consumption Prediction Using LSTM,” 2022.
Available: https://www.theseus.fi/handle/10024/748970
[4] Fingrid, “Electricity Consumption Dataset,” Finland Transmission System Operator.
Available: https://data.fingrid.fi/en/datasets/75
[5] F. Chollet, “Deep Learning with Python,” Manning Publications, 2018.
[6] M. Qureshi et al., “Deep Learning-Based Forecasting of Electricity Consumption,” Scientific Reports, 2024.
[7] C. V. S. Chandrika et al., “Advanced LSTM-Based Time Series Forecasting for Energy Consumption Management,” 2024.
[8] F. Zhou et al., “Prediction of Building Energy Consumption Using RF-LSTM,” Energy Reports, 2024.
[9] S. Ul Rehman et al., “Electricity Consumption Prediction in Smart Buildings Using Deep Learning Approaches,” 2025.
[10] F. Liu et al., “Short-Term Multi-Energy Consumption Forecasting,” Scientific Reports, 2024.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING