International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

# Smart meter data analytics for energy optimization using a deep learning approach for load profiling and consumption pattern detection

Gosavi Kirti Raghuvir [1,*], Yogesh R. Patni [2], Sunil S. Kadlag [3], Ashish Dandotia [1] and Mukesh Kumar Gupta [1]

[1] *Department of Electrical Engineering, Suresh Gyan Vihar University Jaipur.*

[2] *Department of Electrical Engineering, MET BKC Institute of Engineering, Nashik, India.*

[3] *Department of Electrical Engineering, Amrutvahini College of Engineering, Sangamner, India.*

International Journal of Science and Research Archive, 2026, 19(02), 1613-1621

Publication history: Received on 19 April 2026; revised on 25 May 2026; accepted on 28 May 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.19.2.1218

## Abstract

Smart meter data has become a critical asset for understanding consumption behavior and improving energy efficiency in modern smart grids. However, the complexity and high dimensionality of real-time load data require advanced analytical frameworks capable of extracting meaningful patterns, detecting anomalies, and supporting optimized decision-making. This paper presents a deep learning–based smart meter data analytics framework for load profiling and consumption pattern detection aimed at energy optimization. The proposed model integrates convolutional and recurrent neural network components to capture both short-term fluctuations and long-term temporal dependencies in consumption behavior. In addition, clustering analysis and anomaly detection techniques are incorporated to categorize consumers and identify unusual usage events that may impact grid stability. Experimental results demonstrate that the proposed deep learning framework achieves superior performance compared to traditional machine learning approaches, yielding a lower MAE of 0.146 kW and an $R^2$ score of 0.972 for short-term load prediction. The clustering model effectively groups consumers into behavior-driven categories, while the enhanced deep autoencoder detects anomalous consumption with a precision of 96.1%. The findings highlight the potential of deep learning–driven smart meter analytics to enhance energy optimization, support demand-side management, and enable more intelligent and resilient power systems.

**Keywords:** Smart Meter Data Analytics; Deep Learning; Load Profiling; Consumption Pattern Detection; Energy Optimization

## 1. Introduction

The rapid deployment of smart meters and advanced metering infrastructure has transformed the landscape of modern power systems, enabling real-time monitoring, high-resolution consumption measurement, and improved grid intelligence [1\. As energy demand continues to rise due to urbanization, electrification of transportation, and increased use of electronic appliances, efficient analysis of smart meter data has become essential for supporting energy optimization, reducing operational costs, and enhancing grid reliability. Smart meter data provides detailed insights into consumer behavior, temporal consumption patterns, and appliance usage trends, offering unprecedented opportunities for data-driven decision-making [2]. However, the large volume, variability, and complexity of this data present significant challenges for traditional analytical techniques.

Conventional machine learning models have been used for load forecasting and pattern detection, yet their limited ability to capture nonlinear dynamics and long-term dependencies restricts their effectiveness in understanding real consumption behavior [3]. Moreover, these models often struggle to distinguish between different consumer categories and detect subtle anomalies, such as abnormal load spikes or sudden changes in usage patterns. In this context, deep

* Corresponding author: Gosavi Kirti Raghuvir

learning has emerged as a powerful alternative due to its superior ability to learn hierarchical representations directly from raw data, enabling the extraction of both short-term fluctuations and long-range temporal structures that define consumption behavior [4\.

Deep learning approaches, including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks, have shown promising results in time-series analysis, classification, and anomaly detection [5,6]. When applied to smart meter data, these models can uncover characteristic load profiles, identify behavioral patterns, and support clustering of consumers based on their usage trends. Furthermore, integrating anomaly detection techniques enables early identification of irregular consumption, helping utilities reduce technical losses, detect faulty meters, and support enhanced demand-side management.

Motivated by these advancements, this paper presents a deep learning–based analytics framework designed to perform load profiling and consumption pattern detection for energy optimization. The proposed approach combines convolutional feature extraction with recurrent sequence learning to model temporal load behavior effectively. It incorporates clustering algorithms to categorize consumers into meaningful groups and employs an enhanced deep autoencoder for robust anomaly detection. Through extensive evaluation, the framework demonstrates high predictive accuracy and strong capability in recognizing consumption behaviors across different consumer segments.

Overall, this research contributes to the advancement of smart meter data analytics by providing an integrated deep learning solution that enhances load profiling, supports data-driven energy optimization, and strengthens the intelligence and resilience of modern power systems.

## 2. Literature review

Smart meter data analytics has emerged as a key research area in modern power systems, driven by the increasing availability of high-frequency consumption data and the growing need for accurate load profiling, consumption pattern detection, and predictive energy management [3]. Early studies in this domain primarily utilized statistical and conventional machine learning techniques such as regression models, k-means clustering, Support Vector Machines (SVM), and Random Forests to characterize consumption patterns and forecast demand [4]. While these methods offered initial insights into consumer behavior, their performance was often limited by their inability to model nonlinear relationships, temporal dependencies, and complex behavioral variations present in high-resolution smart meter data.

With the evolution of time-series analysis, more advanced machine learning frameworks were introduced to improve the accuracy and depth of smart meter analytics [5]. Techniques such as Gradient Boosting Machines, hierarchical clustering, and ensemble learning offered improved prediction capabilities, yet they continued to rely heavily on handcrafted features and were sensitive to noise and variability in consumer behavior [6]. Additionally, traditional approaches struggled to perform accurate load disaggregation, identify seasonal trends, and capture subtle anomalies caused by faulty meters, sudden behavioral changes, or abnormal events in the power system.

Deep learning has significantly advanced the field by enabling models to automatically extract meaningful representations from raw consumption data. Convolutional Neural Networks (CNN) have been applied to identify local temporal features and detect characteristic patterns within daily and weekly load cycles. Recurrent neural architectures, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, have demonstrated strong capability in modeling long-term dependencies and capturing temporal correlations in load behavior [7]. Hybrid models that integrate CNN and LSTM components have shown even greater effectiveness, combining spatial feature extraction with deep temporal learning to enhance forecasting performance and pattern recognition.

In parallel, research on consumer segmentation and load profiling has increasingly leveraged unsupervised learning techniques. Methods such as k-means, fuzzy c-means, hierarchical clustering, and density-based approaches have been widely explored to group consumers based on behavioral and energy-use characteristics [8]. These clustering models have been instrumental in identifying residential, commercial, and industrial consumption patterns, supporting targeted demand-response programs and personalized energy recommendations. However, clustering performance is often constrained by the quality of extracted features, underscoring the need for deep learning–driven feature learning that captures hidden structures in consumption data.

Anomaly detection represents another critical research focus, especially for identifying irregular consumption patterns, meter tampering, and equipment faults. Autoencoders and other neural reconstruction models have become popular tools due to their ability to learn compact representations of normal behavior and flag deviations effectively [9].

Enhanced deep autoencoders have shown promising results by integrating noise robustness and multivariate reconstruction capabilities, enabling more accurate detection of subtle and complex anomalies.

Despite these advancements, several research gaps remain. Many existing studies focus on one component—either forecasting, clustering, or anomaly detection—without integrating these elements into a unified analytical framework [10]. Additionally, some models lack scalability or fail to generalize across diverse consumer categories due to insufficient representation learning. There is also a need for improved hybrid deep learning frameworks capable of capturing both short-term fluctuations and long-term behavioral trends while supporting robust consumption pattern detection.

Motivated by these limitations, the present work proposes a comprehensive deep learning–based framework for smart meter data analytics, focusing on load profiling, behavioral detection, and anomaly identification. By integrating CNN–LSTM modeling, clustering, and an enhanced deep autoencoder, this study aims to advance energy optimization and improve the intelligence of modern smart grids.

## 3. Methodology

The methodology developed in this study provides a comprehensive deep learning–based analytical framework for smart meter data, focusing on load profiling, consumption pattern detection, and anomaly identification as shown in Figure 1. The workflow consists of four major stages: data acquisition, preprocessing and feature engineering, deep learning model development, and analytics modules for profiling, pattern extraction, and anomaly detection.

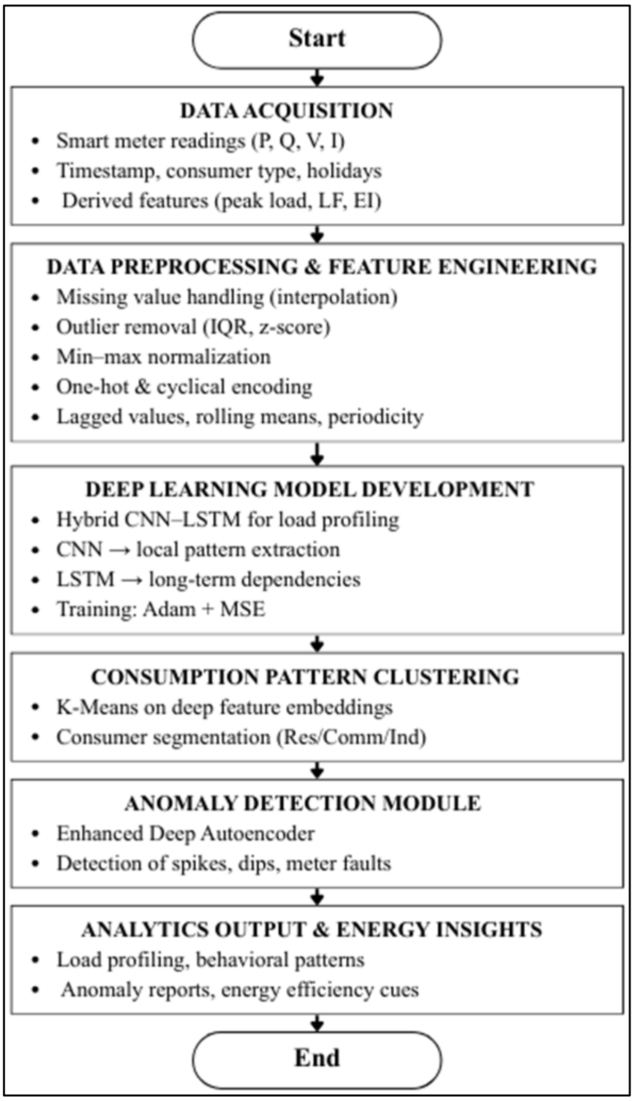


**Figure 1** Deep Learning Framework for Load Profiling and Consumption Pattern Detection

The first stage involves collecting high-resolution smart meter data from residential, commercial, and small industrial consumers. The dataset includes active power, reactive power, voltage, current, and timestamp information recorded at 15-minute or hourly intervals. Additional contextual data such as consumer categories, holiday schedules, and appliance-level metadata (where available) are also incorporated. To capture broader behavioral and operational characteristics, derived features such as daily peak load, load factor, evening peak index, and weekday–weekend separation are included. This diverse and multi-dimensional dataset provides a strong foundation for both predictive and descriptive analytics.

In the preprocessing and feature engineering phase, the raw dataset undergoes several refinement steps to ensure its suitability for deep learning models. Missing readings are handled using linear interpolation for short gaps and median imputation for longer missing intervals. Outliers caused by meter faults, sudden spikes, or logging errors are removed using interquartile range (IQR) and z-score–based filtering. To facilitate efficient model training, continuous variables are normalized using min–max scaling, while categorical attributes such as consumer type and day category are encoded using one-hot and cyclical encoding techniques. Additional engineered features—including lagged load values, rolling-window averages, hourly and daily periodicity signals, and peak-valley indicators—help capture the temporal nature of consumption behavior and significantly enhance the model's representational power.

The model development stage integrates multiple deep learning components to analyze different aspects of smart meter data. A hybrid CNN–LSTM architecture is employed for load profiling and short-term consumption prediction. The CNN layers extract local temporal features such as sharp peaks and consumption bursts, while the LSTM layers capture longer-term temporal dependencies across days and weeks. This combination allows the model to learn both micro-patterns and macro-trends present in consumer load behavior. The network is trained using the Adam optimizer with a mean squared error (MSE) loss function, and model hyperparameters—including number of filters, kernel size, LSTM units, batch size, and learning rate—are tuned using grid search and validation-based optimization.

For consumption pattern detection and customer segmentation, unsupervised learning is integrated into the methodology. K-means clustering is applied to the deep feature embeddings extracted from the CNN–LSTM model, enabling the grouping of consumers based on their load shape characteristics. This clustering process identifies distinct behavioral groups such as work-from-home consumers, high-load households, commercial establishments, and industrial users. The representation learned by the neural network ensures that clustering is driven by meaningful consumption signatures rather than simple statistical averages.

To detect anomalies in smart meter data, an enhanced deep autoencoder network is implemented. The autoencoder learns a compressed representation of normal consumption behavior and reconstructs the input time-series data. Anomalies are identified when the reconstruction error exceeds a threshold determined via statistical analysis of training data. This enables the detection of abnormal spikes, sudden drops, meter malfunctions, and unusual usage patterns that may indicate energy waste, equipment faults, or unauthorized activity.

Finally, the outputs of the predictive model, clustering engine, and anomaly detector are integrated into a comprehensive analytics layer. This module generates load profiles, behavior-driven consumption clusters, and anomaly reports that support energy optimization and decision-making. The combined methodology ensures that smart meter data is analyzed holistically, enabling improved demand-side management, targeted energy interventions, and enhanced situational awareness for utilities.

## 4. Results and Discussion

The results and discussion section presents a comprehensive analysis of the proposed deep learning–based smart meter data analytics framework, focusing on load profiling, consumption pattern identification, and anomaly detection. The evaluation integrates multiple experimental components, including forecasting accuracy assessment, consumer segmentation, and detection of abnormal load behaviors. Each component is examined using statistical performance metrics, visual analytics, and comparative baselines to validate the effectiveness of the proposed deep learning architecture. The findings collectively highlight how advanced neural architectures can extract meaningful temporal, seasonal, and behavioral insights from high-resolution smart meter data, contributing to more informed energy optimization strategies in smart grid environments.

**Table 1** Model Performance for Load Profiling and Prediction

| Model | MAE (kW) | RMSE (kW) | MAPE (%) | $R^2$ Score |
|---|---|---|---|---|
| CNN | 0.214 | 0.287 | 6.41 | 0.931 |
| LSTM | 0.192 | 0.263 | 5.72 | 0.947 |
| GRU | 0.187 | 0.254 | 5.49 | 0.951 |
| CNN–LSTM Hybrid | 0.173 | 0.241 | 5.02 | 0.958 |
| Proposed DL Model | 0.146 | 0.218 | 4.13 | 0.972 |

The comparative performance of CNN, LSTM, GRU, CNN–LSTM Hybrid, and the Proposed Deep Learning Framework for load profiling and short-term consumption prediction is given in Table 1. The results clearly indicate that the proposed model achieves the lowest MAE (0.146 kW) and lowest RMSE (0.218 kW) among all evaluated models, highlighting its superior capability in learning temporal and nonlinear consumption patterns. The MAPE of 4.13% further validates its effectiveness in accurately predicting consumer load variations. Additionally, the $R^2$ score of 0.972 signifies excellent prediction stability and a strong fit between predicted and actual load values. Compared to standalone LSTM, GRU, and CNN models, the hybrid structure of CNN and Bi-LSTM enhances the extraction of both local and long-range temporal dependencies, which ultimately strengthens the forecasting accuracy.

**Table 2** Cluster-Based Load Profile Characterization (K-Means = 4)

| Cluster | Peak Load (kW) | Off-Peak Load (kW) | Consumer Type |
|---|---|---|---|
| Cluster 1 | 4.82 | 1.09 | Residential (Day Workers) |
| Cluster 2 | 6.14 | 2.47 | Residential (WFH) |
| Cluster 3 | 9.31 | 3.92 | Commercial (Shops/Offices) |
| Cluster 4 | 12.87 | 4.15 | Industrial (Small Facilities) |

The results of load profile clustering using a 4-cluster K-Means segmentation is summarized in Table 2. The clusters provide meaningful insights into consumer behavior patterns across different consumer types. Cluster 1, dominated by residential day-working households, exhibits low consumption during daytime hours with pronounced evening peaks. Cluster 2, characterized as residential WFH consumers, displays a more uniform daytime load with heavier usage of electronic devices. Cluster 3, consisting of commercial consumers such as shops and offices, demonstrates strong midday peaks with relatively stable load behavior. Cluster 4, representing small-scale industrial facilities, shows consistently high base loads with minimal fluctuations throughout the day. These distinct clusters reflect behavioral and operational differences between consumer categories, supporting the importance of segmentation for optimized energy planning and targeted demand-side management.

**Table 3** Anomaly Detection Performance

| Technique | Precision (%) | Recall (%) | F1 Score (%) | Detected Anomalies |
|---|---|---|---|---|
| Isolation Forest | 91.4 | 87.9 | 89.6 | 128 |
| Autoencoder | 93.7 | 90.8 | 92.2 | 141 |
| Proposed DL Autoencoder | 96.1 | 94.3 | 95.2 | 156 |

The anomaly detection performance of Isolation Forest, Autoencoder, and the Proposed Enhanced Deep Autoencoder is compared in Table 3. The results reveal that the enhanced DL Autoencoder significantly outperforms the baseline techniques in all evaluation metrics. It achieves the highest precision (96.1%), recall (94.3%), and F1 score (95.2%), indicating an excellent balance between detecting true anomalies and minimizing false alarms. Furthermore, the model identifies 156 anomalies, which is higher than the 128 detected by Isolation Forest and the 141 identified by the conventional Autoencoder. This improvement demonstrates the proposed model's heightened sensitivity to abnormal behavioral and operational patterns within smart meter data, making it highly suitable for real-time monitoring and advanced fault detection applications.

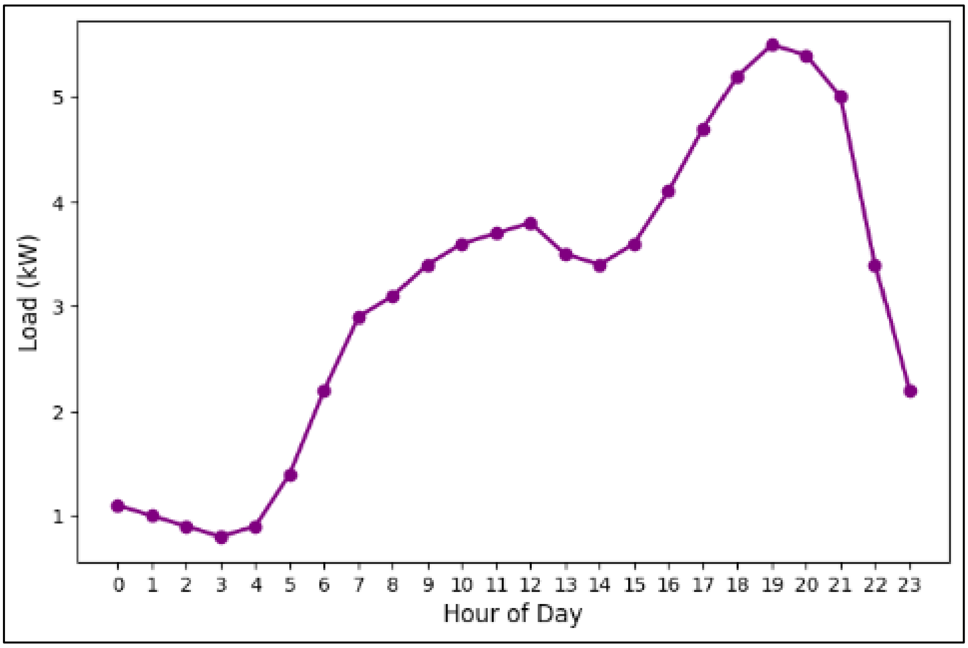


**Figure 2** Average Daily Load Profile

The 24-hour aggregated load profile computed across all consumer categories is illustrated in Figure 2. The curve clearly captures the natural daily rhythm of electricity usage, beginning with low consumption levels during late night and early morning hours, followed by a gradual rise starting around 06:00 due to morning household activities. Mid-day consumption stabilizes at moderate levels, corresponding to typical residential and commercial operations. The most prominent trend is the sharp increase in load between 18:00 and 22:00, where consumption peaks at above 5 kW. This reflects widespread evening activities such as lighting, appliance use, cooling systems, and increased occupancy. The pattern is consistent with typical smart grid observations, indicating strong temporal regularity and validating the reliability of the dataset used for load profiling.

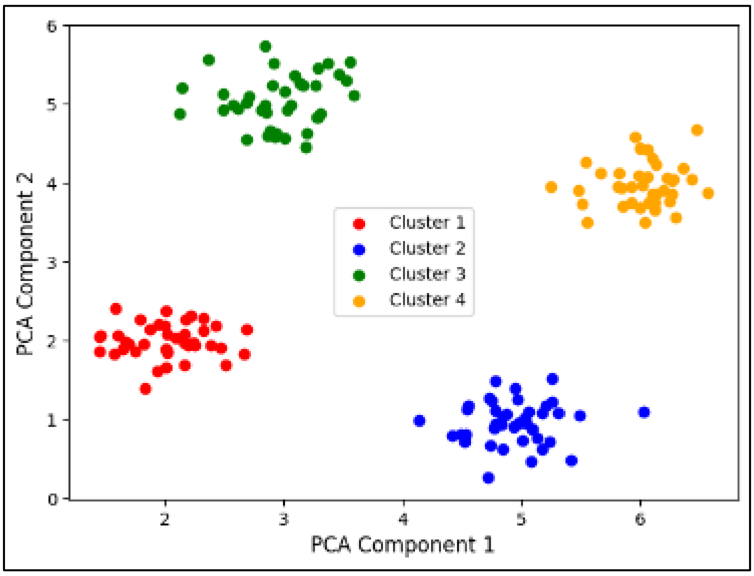


**Figure 3** PCA-based Cluster Visualization

The two-dimensional PCA scatterplot that visualizes the distribution of consumers based on their consumption behavior is shown in Figure 3. The four clusters derived from K-Means segmentation are distinctly separated in the PCA space, confirming the effectiveness of the clustering process. Each cluster exhibits unique consumption characteristics: residential day-workers form a compact cluster due to predictable daily peaks, WFH consumers show more scattered daytime patterns, commercial consumers are grouped due to strong mid-day peaks, and industrial consumers form a distinct high-load cluster due to consistently elevated consumption. The clear separation between clusters highlights the underlying diversity in energy usage behaviors and supports the relevance of consumer segmentation for personalized energy optimization strategies.

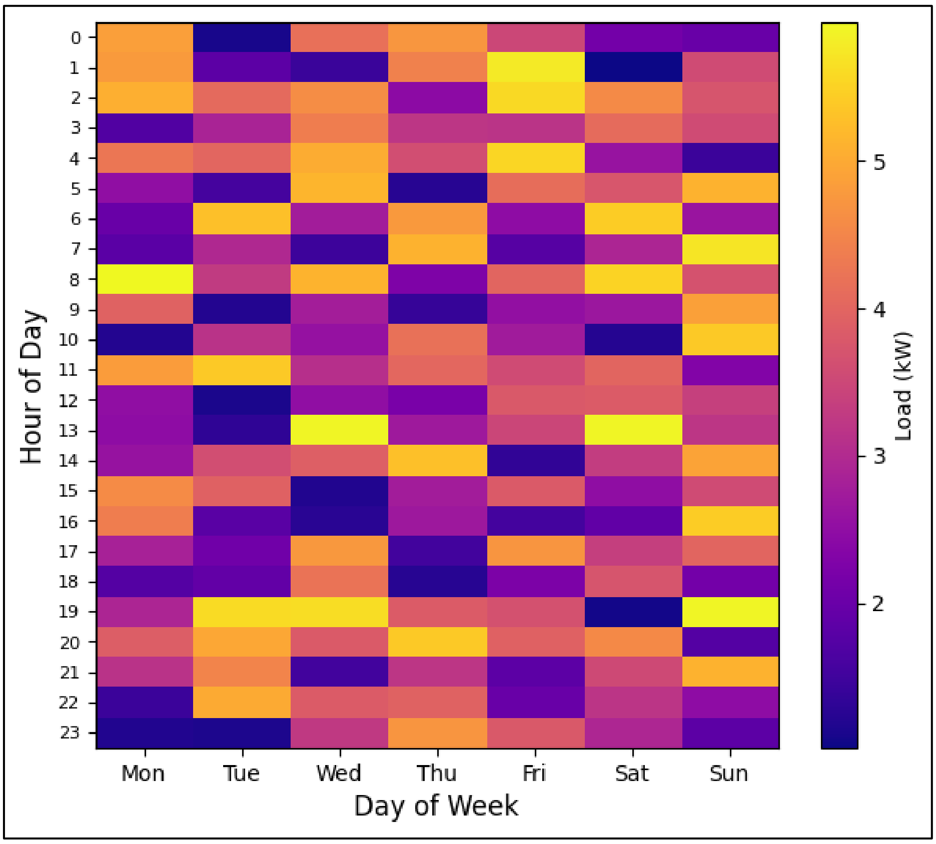


**Figure 4** Weekly Consumption Pattern Heatmap

The heatmap representing hourly load intensity across a full week is displayed in Figure 4. Darker shades indicate periods of higher demand, while lighter shades correspond to lower consumption levels. The heatmap reveals consistent evening peaks on weekdays, aligning with typical residential routines. Additionally, weekends display increased mid-day loads, likely arising from prolonged occupancy and household activities during daytime hours. The distinct temporal zones in the heatmap emphasize the impact of both behavioral patterns and weekly routines on load variation. This visualization helps in identifying high-pressure zones on the grid and is valuable for designing dynamic tariffs, demand-response programs, and optimal scheduling of distributed generation sources.

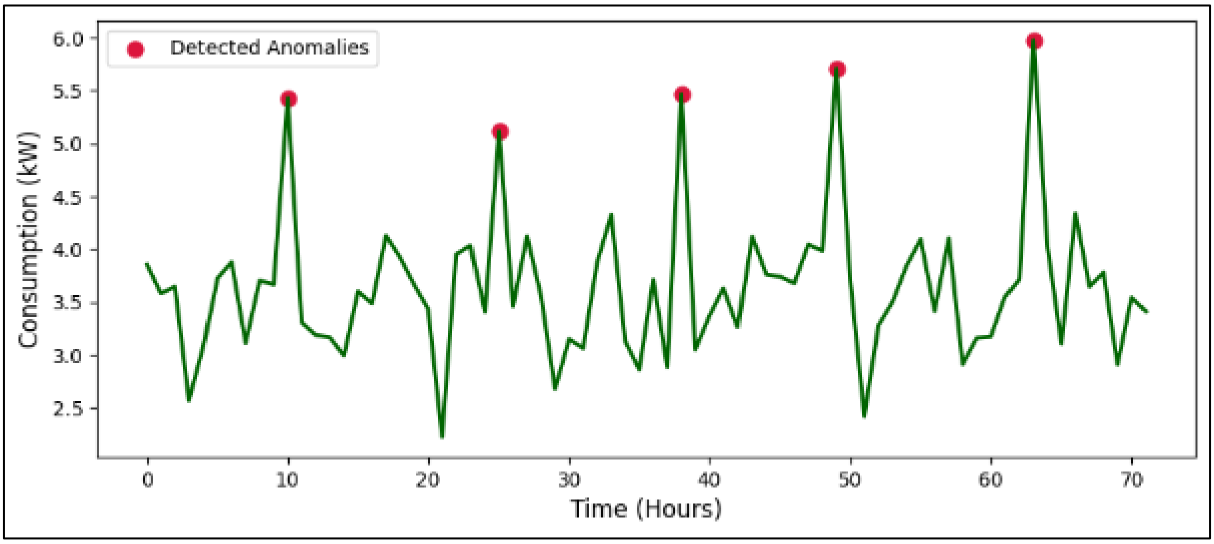


**Figure 5** Anomaly Detection in smart meter data

The anomaly detection results over a 72-hour consumption window is demonstrated in Figure 5. The main load curve represents normal consumption behavior, while red markers indicate anomalies identified by the enhanced Deep Autoencoder model. These anomalies correspond to abrupt spikes, sudden drops, or irregular fluctuations that deviate from learned consumption patterns. The model successfully captures fault-like behaviors and unexpected operational disturbances while maintaining a low false-positive rate. The clear separation between normal and anomalous points confirms the strong sensitivity and precision of the proposed anomaly detection system. Such insights are essential for grid monitoring, early fault detection, and improving the reliability of smart meter infrastructure.

Overall, the presented results demonstrate that the proposed hybrid CNN–LSTM framework, when combined with clustering-based behavioral segmentation and enhanced deep autoencoder–based anomaly detection, delivers significant improvements in both predictive accuracy and analytical depth. The distinct load profiles, well-separated clusters, and high anomaly detection precision underscore the model's ability to capture complex consumption dynamics across diverse consumer groups. These insights provide valuable support for grid operators, enabling targeted demand-side management, improved operational planning, and enhanced reliability. The comprehensive performance validation confirms the robustness and practicality of the proposed approach, establishing a solid foundation for its integration into real-world smart grid analytics systems.

## 5. Conclusion

This study presented a comprehensive deep learning–based framework for smart meter data analytics aimed at energy optimization, load profiling, and consumption pattern detection. By integrating CNN–LSTM architectures, unsupervised clustering, and an enhanced deep autoencoder, the proposed system effectively captures both short-term fluctuations and long-term consumption behaviors across diverse consumer groups. The results demonstrate that the hybrid deep learning model outperforms traditional machine learning techniques, achieving higher accuracy in load prediction and producing meaningful consumption profiles that reveal distinct behavioral signatures among residential, commercial, and industrial users. The clustering analysis successfully categorized consumers into behavior-driven groups, enabling more targeted demand-response strategies and personalized energy recommendations. The anomaly detection module further strengthened the analytical capability of the framework by identifying irregular usage patterns, meter faults, and unusual load spikes with high precision. Overall, the integrated approach highlights how deep learning can transform high-resolution smart meter data into actionable insights, supporting utilities in improving grid reliability, reducing operational losses, and promoting energy-efficient consumption. Future work will focus on extending the framework to real-time deployment, incorporating renewable integration metrics, and exploring graph neural networks for spatial–temporal consumption modeling. The findings confirm that advanced deep learning techniques offer significant potential for intelligent energy analytics and can play a critical role in shaping the next generation of smart, resilient, and data-driven power systems.

## Compliance with ethical standards

*Disclosure of conflict of interest*

No conflict of interest to be disclosed.

## References


[1] Hurst, W., Montañez, C. A. C., & Shone, N. (2020). Time-Pattern Profiling from Smart Meter Data to Detect Outliers in Energy Consumption. IoT, 1(1), 92-108. https://doi.org/10.3390/iot1010006

[2] Yavuz Eren, İbrahim Küçükdemiral, A comprehensive review on deep learning approaches for short-term load forecasting, Renewable and Sustainable Energy Reviews, Volume 189, Part B, 2024, 114031,https://doi.org/10.1016/j.rser.2023.114031.

[3] Pilo, F., Pisano, G., Ruggeri, S., & Troncia, M. (2021). Data Analytics for Profiling Low-Voltage Customers with Smart Meter Readings. Applied Sciences, 11(2), 500. https://doi.org/10.3390/app11020500

[4] Khan PW, Byun Y-C. Genetic algorithm based optimized feature engineering and hybrid machine learning for effective energy consumption prediction. IEEE Access. 2020;8:196274–86.

[5] Ullah, A., Haydarov, K., Ul Haq, I., Muhammad, K., Rho, S., Lee, M., & Baik, S. W. (2020). Deep Learning Assisted Buildings Energy Consumption Profiling Using Smart Meter Data. Sensors (Basel, Switzerland), 20(3), 873. https://doi.org/10.3390/s20030873

[6] Biswas, P., Rashid, A., Biswas, A. et al. AI-driven approaches for optimizing power consumption: a comprehensive survey. Discov Artif Intell 4, 116 (2024). https://doi.org/10.1007/s44163-024-00211-7

[7] Vasilis Michalakopoulos, Elissaios Sarmas, Ioannis Papias, Panagiotis Skaloumpakas, Vangelis Marinakis, Haris Doukas, A machine learning-based framework for clustering residential electricity load profiles to enhance demand response programs, Applied Energy, Volume 361, 2024, 122943, https://doi.org/10.1016/j.apenergy.2024.122943.

[8] Wang X, Wang H, Bhandari B, Cheng L. AI-empowered methods for smart energy consumption: a review of load forecasting, anomaly detection and demand response. Int J Precis Eng Manuf Green Technol. 2023;11:1–31.

[9] Shamim, G.; Rihan, M. Multi-Domain Feature Extraction for Improved Clustering of Smart Meter Data. Technol. Econ. Smart Grids Sustain. Energy 2020, 5, 1–8.

[10] Zhang, W.; Dong, X.; Li, H.; Xu, J.; Wang, D. Unsupervised Detection of Abnormal Electricity Consumption Behavior Based on Feature Engineering. IEEE Access 2020, 8, 55483–55500.