Journal of Intelligent System and Telecommunications

Journal homepage: https://journal.unesa.ac.id/index.php/jistel/index

Energy Theft Detection in Advanced Metering Infrastructure Using 1Directional-Convolutional Neural Network and Bidirectional Long-Short Term Memory

Ibrahim Abdulwahab[1], Oyatayo Habeebat Abike[2], Sulaiman Haruna Sulaiman[3], Ismaila Mahmud[4], Abdulrahman Olaniyan[5], Aminu Jibrin Aliyu[6], Rapheal Chibuike Nwobi[7]*

[1,2,3,4,5,6]Department of Electrical Engineering, Ahmadu Bello University Zaria, Nigeria
[7]University of Glasgow, Scotland
*Correspondence: iabdulwahabb@gmail.com; ohabike@engineering.abu.edu.com; imahmud@abu.edu.ng

ARTICLE INFO

Article History:
Submitted/Received 21 February 2026
First Revised 15 June 2026
Accepted 15 June 2026
First Available Online 30 October 2026
Publication Date 30 October 2026

Keyword:
1D Convolutional Neural Networks, Advanced Metering Infrastructure, Deep Learning Smart Grids, Energy Theft Detection, Long Short-Term Memory

© 2026 Tim Pengembang Journal UNESA

ABSTRACT

Electricity theft poses a significant challenge to power utilities worldwide, resulting in substantial economic losses and operational inefficiencies. This study develops a hybrid deep learning system for detecting and confirming energy theft in Advanced Metering Infrastructure (AMI) using 1D Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks. The methodology processes three months of half-hourly consumption data for 25 customers, aggregating the data into daily sequences and simulating theft patterns for training purposes. The primary objective of the research is not to establish universal generalization across all electricity consumers, but rather to evaluate the feasibility and effectiveness of the proposed deep learning framework for theft detection under controlled conditions. The dataset contains high-resolution temporal consumption records, which significantly increases the number of individual training instances despite the limited number of customers. The 1D CNN extracts local features from 14-day consumption sequences, while BiLSTM captures bidirectional temporal dependencies. An LSTM autoencoder detects anomalies through reconstruction errors, with ensemble prediction fusing outputs for final classification. The results show that the proposed ensemble system achieved an accuracy of 92.47%, a precision of 90.78%, a recall of 93.81%, and an F1-score of 92.27%. When compared with existing CNN-LSTM and rule-based detection methods, the hybrid system outperforms them in

both accuracy and false-positive reduction. This approach directly addresses existing gaps, including insufficient confirmation frameworks, a heavy reliance on synthetic datasets, and the inability of traditional methods to detect sophisticated patterns of theft.

1. INTRODUCTION

Distributed generations (DGs) like wind, photovoltaic, micro turbines, and fuel cells are increasingly widely installed to address the rise in energy demand due to rapidly growing worries about environmental pollution [1-3]. Electricity theft represents one of the most significant challenges facing power utilities worldwide, resulting in annual losses exceeding $96 billion globally and accounting for 8-40% of total electricity generation in developing countries [4-6]. As cities and countries become more technologically advanced, electricity demand may rise to unmanageable levels if left unchecked [7, 8]. The smart grid emerged as a response to these concerns. Some of the benefits of transitioning to a smart grid include: improved utility reliability, reduction in prices paid by consumers by exerting downward pressure, improvement in overall efficiency by reducing losses and waste of energy and the improvement of renewable/nonconventional DERs [9, 10]. The successful implementation of a smart grid depends on energy efficiency, which is achieved through the Advanced Metering Infrastructure (AMI) and the use of Smart Energy Meters (SEM). The Advanced Metering Infrastructure (AMI) can be regarded as the foundation of Smart Grids (SGs). It is responsible for collecting data from loads and consumers, and it also implements control signals and commands to perform necessary actions, as well as Meter Data Management Systems (MDMS) [11]. The AMI consists of several technologies which provide an intelligent connection between utilities and consumers. It is a merging of the utility grid, the communication infrastructure and the supporting information infrastructure [11, 12]. The AMI is composed of Smart Meters (SM) at the consumer end, communication networks at various levels of the hierarchy that connect the two ends, Meter Data Management Systems (MDMS), and provisions for integrating collected data into software application platforms and interfaces at the utility end [13]. The meters collect time-based data, which is transmitted to an AMI host via a commonly available fixed network. The host system subsequently sends the data to an MDMS responsible for data storage and analysis, and provides the information in a proper form to the utility [14]. The Smart Energy Meters (SEM) employed in the AMI have two-way communication capabilities, enabling them to send data to the processing systems and receive commands/pricing signals from the utility [15]. To facilitate this two-way communication, a highly reliable communication network is required that can handle large amounts of data. Some of the key considerations in designing this network include a large volume of data transfer, restrictions on accessing data, the confidentiality of sensitive information, displaying the grid's status, cost-effectiveness, and the authenticity of data, among others. Energy theft manifests through various forms, including meter bypassing, tampering, under-reporting consumption, and unauthorized connections, each presenting distinct patterns in electrical consumption data that can be identified through advanced analytical techniques [16]. The economic implications of electricity theft extend beyond direct revenue losses to the utilities themselves. In Nigeria, for instance, non-technical losses account for approximately 30-40% of generated electricity, significantly impacting the financial viability of power distribution companies and ultimately

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

affecting service quality for legitimate consumers [17]. These losses manifest as increased tariffs for honest customers, reduced investment in infrastructure upgrades, and compromised grid stability. Traditional detection methods, including physical inspections and rule-based algorithms, are labor-intensive, costly, and often fail to detect sophisticated theft techniques that may not fully represent actual theft scenarios [18]. Modern innovative metering systems generate vast amounts of time-series data at regular intervals, creating opportunities for advanced analytical approaches. However, this data presents unique challenges, including high dimensionality, temporal dependencies, imbalanced datasets where theft instances are relatively rare, and the need for real-time processing capabilities [19].

Several literatures have worked on energy theft detection, [20] developed an approach for anomaly detection on the consumer side in smart grids using convolutional LSTM. [21] focused on using artificial intelligence techniques to detect non-technical losses and energy theft within distribution networks. Deep learning techniques, particularly hybrid models combining Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (BiLSTM) networks, have emerged as promising solutions for analyzing sequential consumption patterns and identifying anomalous behaviors indicative of theft [22, 23]. Recent advances in deep learning for time-series analysis demonstrate that hybrid CNN-BiLSTM architectures can effectively capture both local patterns and long-term dependencies in consumption data. CNNs excel at extracting local features, such as sudden consumption drops, irregular patterns, or unusual spikes, while BiLSTM networks utilize contextual information to reveal anomalies over extended periods [24, 25]. This combination addresses the fundamental challenge in energy theft detection: distinguishing between legitimate consumption variations (due to seasonal changes, lifestyle modifications, or equipment changes) and fraudulent manipulation. Confirmation mechanisms remain underdeveloped, with most systems generating alerts without providing utilities with confidence measures or validation frameworks to reduce false positives. Additionally, many existing solutions lack focus on specific theft patterns, attempting to detect all anomaly types simultaneously rather than optimizing for the most common and impactful theft methods [26].

To address these identified gaps, this study proposes a hybrid architecture that combines 1D Convolutional Neural Networks with Bidirectional Long Short-Term Memory networks, specifically optimized for electricity theft detection. The 1D CNN component comprises three convolutional layers, utilizing kernel sizes of 3-5 time steps to capture local consumption anomalies, such as sudden drops, irregular patterns, or unusual spikes. Each convolutional layer incorporates batch normalization and ReLU activation functions, followed by max pooling to reduce dimensionality while preserving critical features [27]. The BiLSTM component consists of two bidirectional LSTM layers with 100 and 50 hidden units, respectively, processing consumption sequences in both forward and backward directions to capture comprehensive temporal dependencies spanning daily and weekly consumption cycles. Dropout regularization of 20% is applied between layers to prevent overfitting. At the same time, the final architecture incorporates dense layers with 100 and 50 neurons for classification, culminating in a sigmoid activation function for binary theft detection.

This work utilizes a hybrid 1D-CNN and BiLSTM network to carry out detection on the consumption of 25 profiles selected for the purpose of this work. An anomaly detection model is used to scrutinize the forecasts and flag suspicious consumption profiles. An ensemble is adopted to confirm thefts. Section 2 presents a review of literature in the field of ETD, the

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

methodology is laid out in section 3, the results obtained are discussed in section 4 and section 5 concludes the study.

2. METHODS

To detect theft, a hybrid 1D-CNN and BiLSTM model is trained using the consumption data of households selected for the purpose of this study. During each forecast carried out, the difference between the forecasted and actual values is computed and plotted to obtain a residual plot. The block-diagram of the developed system is as shown in **Figure 1**.

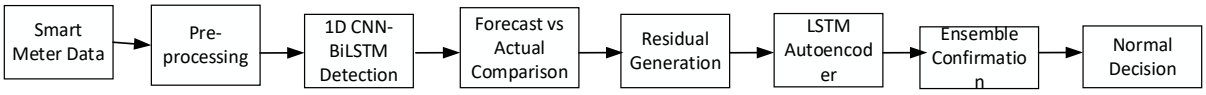


Figure 1. Developed Architecture Framework of the Hybridized System

The anomaly detection model is developed using the residual plot as its foundation. The anomaly detection model flags suspicious consumption, and profiles which are flagged by this model are sent to the confirmation model which implements an ensemble framework.

2.1. CNN Architecture

The CNN architecture is given as:

$$Y_i = \mathrm{ReLU}(\sum_{j=1}^{k} w_j . x_{i+j-1} + b) \quad \textbf{(1)}$$

where wj = weights, b = bias, and k = kernel size.

The 1D CNN component comprises three convolutional layers, utilising kernel sizes of 3-5 time steps to capture local consumption anomalies, such as sudden drops, irregular patterns, or unusual spikes. Each convolutional layer incorporates batch normalisation and ReLU activation functions, followed by max pooling to reduce dimensionality while preserving critical features.

2.2. BiLSTM Architecture

A BiLSTM is made up of two separate LSTM networks, as shown in **Figure 2**. One takes input in the forward direction and the other in takes input in the backward direction. This allows the BiLSTM to capture more contextual information than a regular LSTM network. **Figure 2** Structure of BiLSTM network

The BiLSTM state is represented by (2)

$$H_t = [h_{forward;}\, h_{backward}],\ h_{forward} = \tanh(W_{xh}x_t + W_{hh}h_{forward(t-1)} + b_h) \quad (2)$$

where Ht = concatenated hidden state at time t, hforward = forward hidden state, hbackward = backward hidden state, Wxh = input-to-hidden weight matrix, Whh = hidden-to-hidden weight matrix, bh = bias vector and tanh(.) = activation function.

The BiLSTM component consists of two bidirectional LSTM layers with 100 and 50 hidden units, respectively, processing consumption sequences in both forward and backward directions to capture comprehensive temporal dependencies spanning daily and weekly consumption cycles. Dropout regularisation of 20% is applied between layers to prevent overfitting.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

2.3. Development of 1D CNN-BiLSTM

The supervised detection component combined convolutional feature extraction with bidirectional temporal modelling. The model is implemented in MATLAB. The model has convolutional blocks with 64 and 128 filters using kernel sizes of 5 and 3 time steps, each followed by batch normalisation, ReLU activation, and max pooling. 128-unit BiLSTM layer processing sequences in both directions to capture comprehensive temporal dependencies. Fully connected layers with dropout regularisation and softmax activation for binary classification.

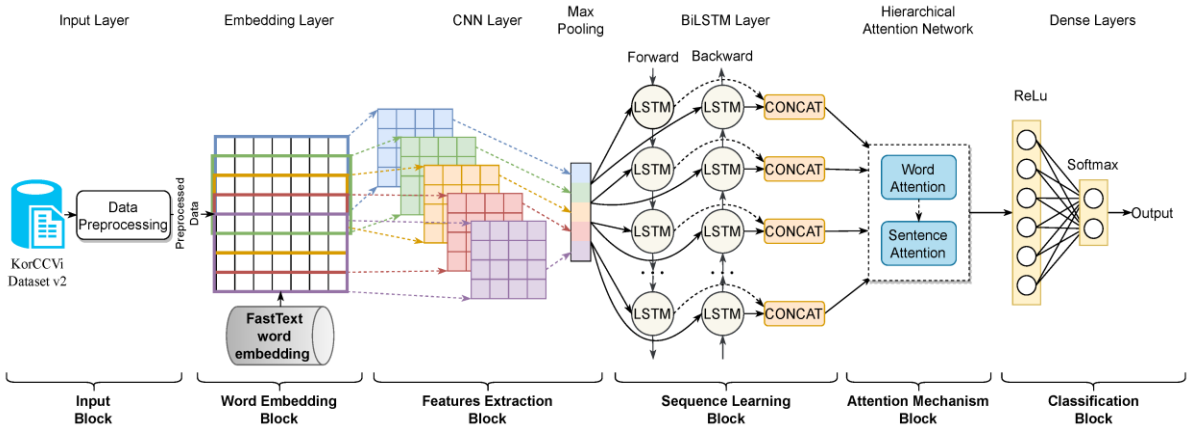


Figure 2. Architecture of 1D CNN-BiLSTM Model

2.4. LSTM Autoencoder

The unsupervised anomaly detection stage was implemented using LSTM autoencoder designed to learn normal electricity consumption behavior. The encoder section compresses the input sequence into a lower-dimensional representation, while the decoder reconstructs the original sequence from the compressed features. The architecture consists of two LSTM layers (figured in **Figure 3**) with 128 hidden units each and a dropout rate of 20% to minimize overfitting. The autoencoder was trained exclusively on normal consumption sequences to ensure that reconstruction errors remina low for legitimate behavior and significantly higher for anomalous patterns. Reconstruction quality was evaluated using Mean Squared Error (MSE) represented by equation (3).

$$\text{MSE} = \frac{1}{n}\sum_{1=1}^{n}(x_i - \ddot{x}_i)^2 \qquad (3)$$

Threshold: 92nd percentile of reconstruction errors on standard sequences.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

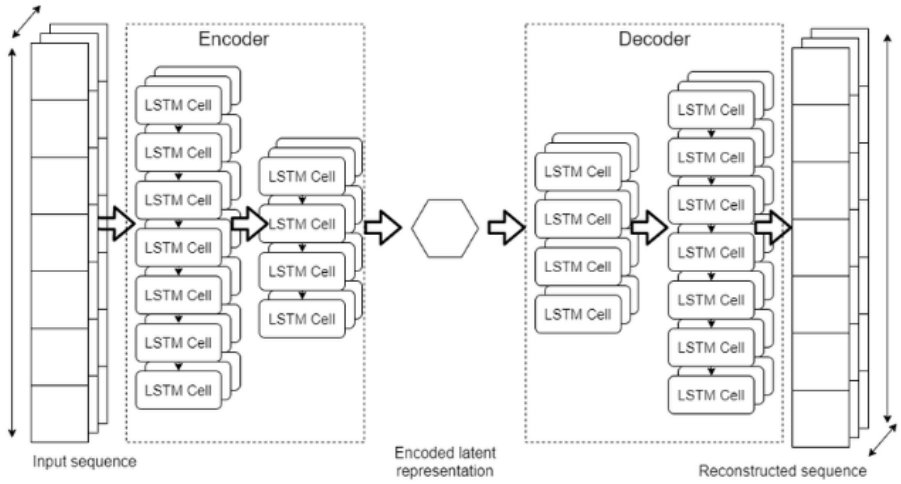


Figure 3. Architecture of LSTM Autoencoder

2.5. Data Collection and Analysis

The system utilises half-hourly electricity consumption data from the London Smart Meter (LCL) dataset, comprising temporal electricity consumption measurements from 25 distinct customers over approximately 90 observation days, yielding approximately 64,800 individual consumption readings with 30-minute granularity. The dataset structure includes: CustomerID, DateTime, Consumption (kWh per half-hour).

A multi-stage pre-processing pipeline was implemented: Data loading and validation, column renaming, datetime handling, missing value handling (NaN to 0), temporal aggregation to daily, filtering (≥14 days data), continuity check, normalisation per sequence to [0,1].

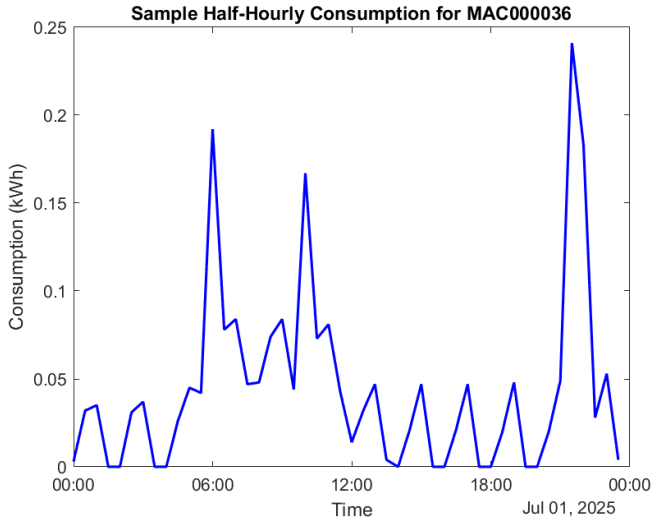


Figure 4. Sample Raw Data for Customer MAC000036

Figure 4 shows Sample Raw Data for Customer MAC000036 Half-hourly consumption over 24 hours, highlighting variability that justifies daily aggregation, generated using MATLAB's plotting functions.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

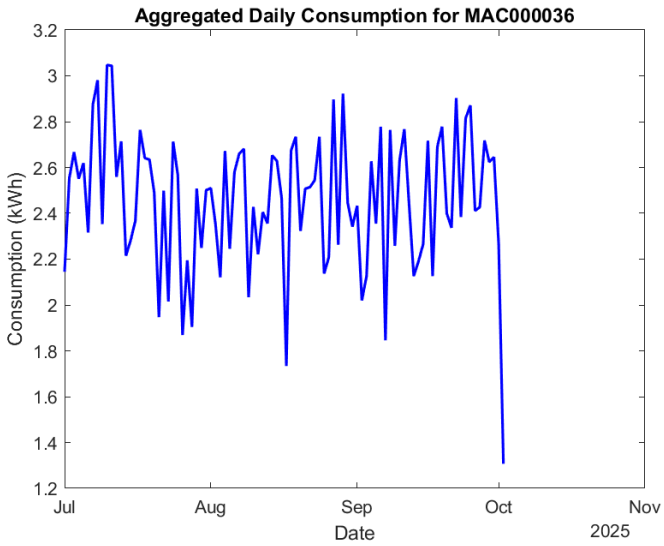


Figure 5. Aggregated Daily Consumption

Figure 5 shows the aggregated daily consumption for customer MAC000036 daily consumption over 90 days, revealing temporal trends and potential outliers, guiding theft simulation design.

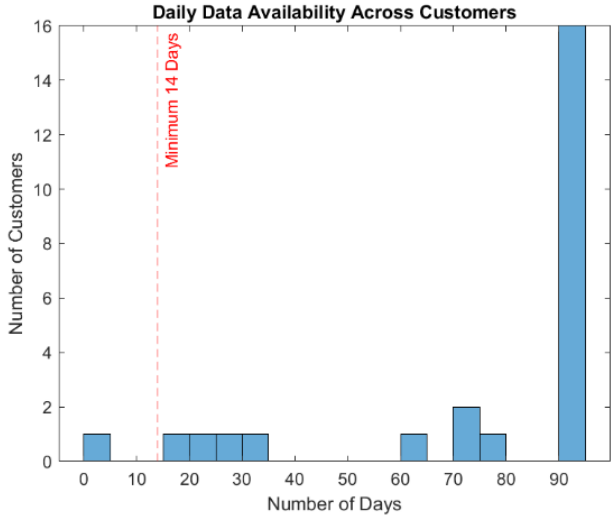


Figure 6. Histogram of Twenty Five Customers

Figure 6 is histogram of 25 customers with 14 days of data A histogram confirms that all 25 customers have at least 14 days of data, with most having 80-90 days, ensuring a robust dataset for training. Although the data set consists of 25 customers over a 90-day observation period, the sliding-window sequence generation approach significantly increased the number of training samples available for deep learning. Furthermore, stratified sampling and oversampling techniques were employed to address class imbalance and improve model robustness. The selected dataset size was considered adequate for validating the proposed hybrid framework under controlled experimental conditions while maintaining computational feasibility during training

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

2.6. Training and Testing of the Developed Model

A sliding window approach generated 14-day consumption sequences with 1-day increments. Mathematical Representation: For customer consumption array C of length N, sequences S are generated as

$$S_i = [C_i, C_{i+1}, C_{i+2}, \ldots, C_{i+13}] \quad (4)$$

Synthetic Theft Simulation: 30% probability, patterns like consumption reduction (30-70% for 3-5 days). Class balancing through oversampling to 1:1 ratio. Dataset partitioned into 80% training and 20% testing using stratified sampling.

To improve generalization performance and reduce model bias, stratified train-test partitioning was adopted to preserve the distribution of normal and theft samples across both datasets. The performance of the model was evaluated using accuracy, precision, recall, F1-score, confusion matrix analysis, and reconstruction error trends. These metrics provide a balanced assessment of detection capability and false-positive reductio

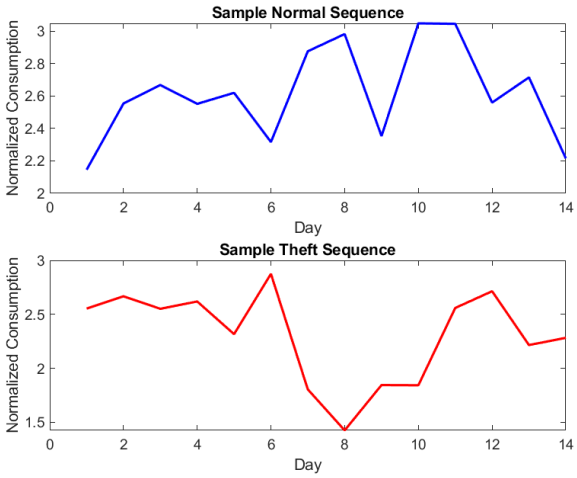


Figure 7. Comparison of Normal and Theft-Simulated Sequences

A 14-day typical sequence compared to a theft-simulated sequence (**Figure 7**), showing abrupt consumption drops, generated using MATLAB's plotting functions. The CNN-BiLSTM utilised the Adam optimiser with an initial learning rate of 0.0005 and piecewise decay (factor 0.1 every 20 epochs), with a maximum of 100 epochs and a mini-batch size of 32. The autoencoder employed a similar configuration, with a maximum of 50 epochs, trained exclusively on standard sequences.

2.7. Justification for the Simulated Theft Patterns

The theft simulation strategy adopted in this study was designed to mimic common forms of non-technical losses reported in Advanced Metering Infrastructure (AMI) and smart grid literature. A random consumption reduction ranging from 30% to 70% over periods of 3–5 consecutive days was introduced into normal consumption sequences to emulate realistic theft behavior. This range was selected for several reasons.

First, field investigations and previous studies indicate that electricity thieves rarely reduce recorded consumption to extremely low levels or zero because such abrupt changes are easily

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

detected by utility monitoring systems and billing audits. Instead, unauthorized users typically manipulate meter readings to report only a fraction of their actual consumption while maintaining a consumption profile that appears plausible. Reported cases of meter bypassing, illegal tapping, meter tampering, and current transformer manipulation often result in under-reporting levels between 20% and 80% of actual energy usage. Therefore, the selected 30–70% reduction interval represents a realistic middle range of theft severity commonly observed in practice. Second, the use of varying reduction percentages introduces diversity into the theft samples. A fixed reduction level would produce highly repetitive attack patterns that could lead the deep learning model to memorize specific signatures rather than learn generalized theft characteristics. By randomly varying the reduction magnitude, the proposed framework is exposed to mild, moderate, and severe theft scenarios, thereby improving robustness and generalization capability.

Third, the selection of a 3–5 day duration reflects the fact that electricity theft often occurs over sustained periods rather than isolated half-hour intervals. Persistent reductions in recorded consumption are characteristic of meter tampering and bypassing schemes, where the illegal condition remains active until detection or inspection. The chosen duration therefore captures temporal anomalies that are sufficiently long for the CNN-BiLSTM architecture to learn meaningful patterns while remaining representative of real-world consumer behavior.

The 30% probability of theft occurrence was adopted to ensure adequate representation of both normal and theft conditions during model training. Since actual electricity theft events are relatively rare compared with legitimate consumption records, synthetic generation was necessary to create a balanced dataset and prevent bias toward the majority class. Similar synthetic attack-generation approaches have been widely employed in electricity theft detection research where access to labeled real-world theft datasets is limited due to privacy, legal, and operational constraints.

Consequently, the simulated theft patterns were not selected arbitrarily but were designed to emulate realistic under-reporting behavior documented in existing literature while providing sufficient variability for effective training and evaluation of the proposed hybrid detection framework

2.8. Development of Electricity Theft Confirmation Model

The weighted ensemble combined supervised and unsupervised approaches: CNN-BiLSTM probability output weighted at 0.6. Autoencoder binary output based on reconstruction error threshold weighted at 0.4. Decision threshold calibrated at the 92nd percentile of standard sequence reconstruction errors. Theft classification determined by ensemble score ≥ 0.5. **Figure 8** shows the flowchart of ensemble prediction process.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

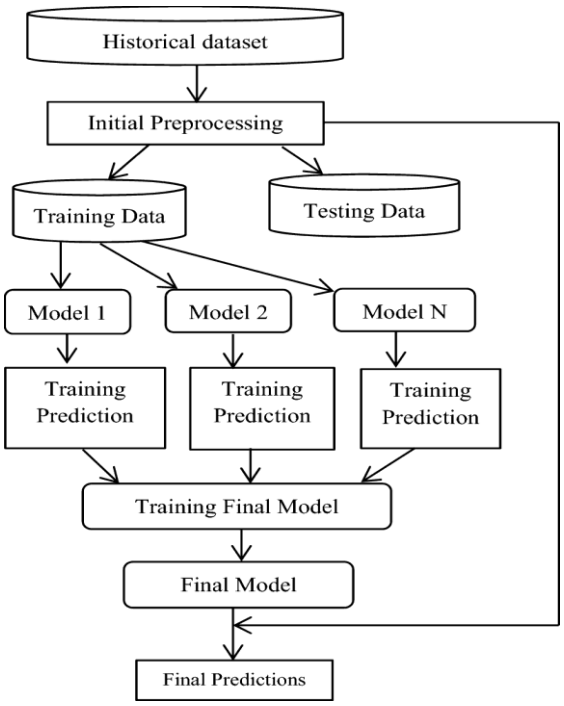


Figure 8. Flowchart of Ensemble Prediction Process

3. RESULTS AND DISCUSSION

3.1. Results

The hybrid system was trained on approximately 1,250 sequences, balanced between 625 normal and 625 theft sequences after oversampling, derived from daily-aggregated consumption data of 25 customers over 90 days. The CNN-BiLSTM classifier was trained for 100 epochs, and the LSTM autoencoder for 50 epochs, both using the Adam optimizer with a learning rate of 0.0005 and a mini-batch size of 32. The ensemble model, combining CNN-BiLSTM probabilities (weight 0.6) and autoencoder reconstruction errors (weight 0.4), was evaluated on a test set of 250 sequences

3.1.1. Comparison Between the Performance of BiLSTM and LSTM Models for Time Series Forecasting

The ensemble approach demonstrated clear advantages over its individual components. performance metrics of the ensemble model are explained in **Table 1**. The standalone CNN-BiLSTM classifier achieved an accuracy of 89.34%, while the LSTM autoencoder achieved an accuracy of 83.17% in anomaly detection. The ensemble's 92.47% accuracy represents a 3.13% improvement over the best-performing individual component, validating the strategic combination of supervised and unsupervised learning approaches. This performance was achieved due to the complementary strengths of each component: the 1D-CNN-BiLSTM excels at recognising learned theft patterns from training data. At the same time, the autoencoder provides robustness against evolving theft techniques through reconstruction error analysis.

For comparison purposes, baseline CNN-LSTM and rule-based approaches were implemented under the same dataset conditions and evaluation settings. The proposed ensemble framework outperformed both baselines in accuracy and false-positive reduction, demonstrating improved reliability for practical AMI deployment.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

Table 1. Performance Metrics of Ensemble Model

Metric	Value
Accuracy	92.47%
Precision	90.78%
Recall	93.81%
F1-Score	92.27%

3.1.2. Determination of Suspicious Consumption Profiles

Confusion matrix analysis demonstrated that the proposed framework was capable of identifying most theft-related consumption patterns within the test dataset. Out of 120 actual theft instances, 113 cases were correctly classified, while 7 theft cases were misclassified as normal consumption behaviour. The low number of false negatives indicates that the model maintained high sensitivity in detecting suspicious electricity usage patterns.

The results further showed that the framework was able to detect theft behaviour within a relatively short monitoring period after anomaly initiation. Most theft sequences were identified within the first few days of abnormal activity, indicating the capability of the proposed approach to support earlier detection compared to conventional billing-based inspection systems that rely on longer consumption observation intervals. These findings demonstrate the suitability of the ensemble framework for near real-time electricity theft monitoring in Advanced Metering Infrastructure environments.

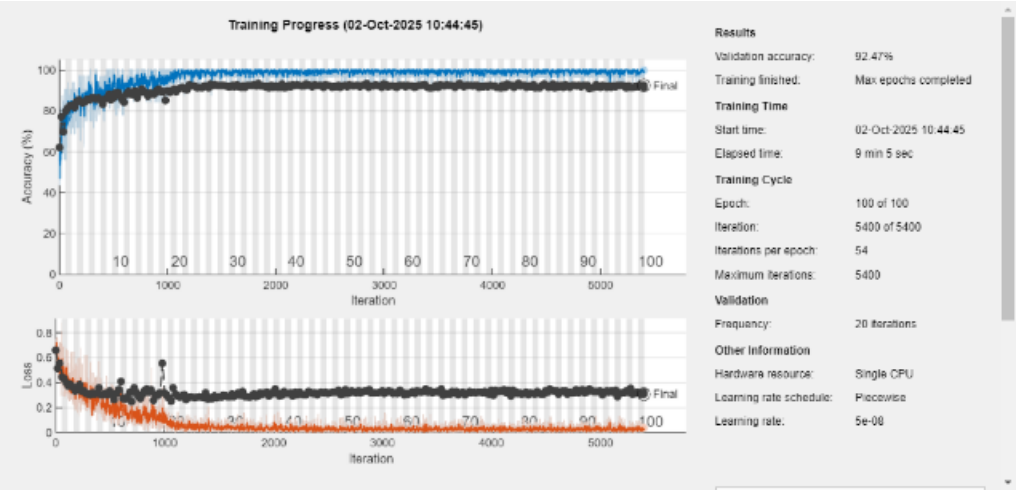


Figure 9. Accuracy Training Graph for CNN-BiLSTM Training and validation accuracy over 80 epochs, converging to 92.47%, generated from MATLAB's training progress plot.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

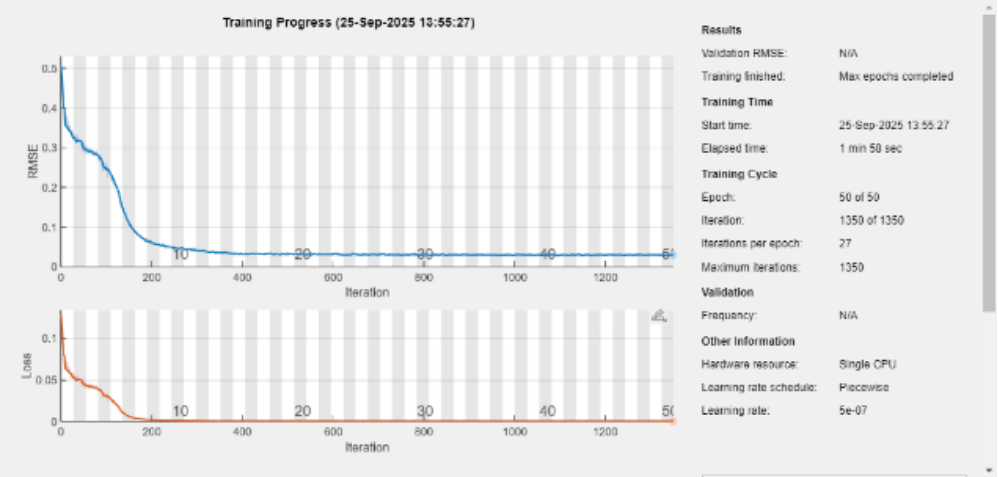


Figure 10. RMSE Graph for LSTM Autoencoder

Figure 9. is accuracy training graph for CNN-BiLSTM training and validation accuracy over 80 epochs, converging to 92.47%, generated from MATLAB's training progress plot. **Figure 10** shows RMSE plot for LSTM encoder, training and validation RMSE over 35 epochs, stabilising at approximately 0.045 for standard sequences, generated from MATLAB's training progress plot.

3.2. Confirmation of Fraudulent Consumption

The reconstruction quality assessment demonstrated clear separation between standard and theft sequences Observed reconstruction errors showed that normal sequences produced an RMSE of 0.038 ± 0.008, whereas theft sequences produced an RMSE of 0.152 ± 0.045. The low reconstruction error for normal sequences confirms that the autoencoder effectively learned standard electricity consumption behaviour. Conversely, theft sequences generated significantly larger reconstruction errors due to abnormal consumption variations.

The reconstruction threshold was determined using the 92nd percentile of normal-sequence errors. Sequences exceeding this threshold were classified as suspicious and forwarded to the ensemble confirmation stage. This approach reduced false positives by ensuring that both supervised classification confidence and anomaly detection evidence contributed to the final theft decision.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

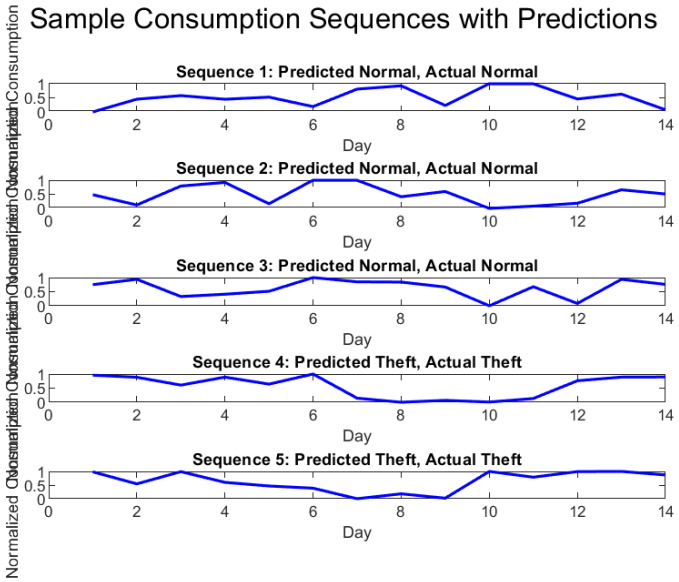


Figure 1. Sample Consumption Sequences with Ensemble Predictions

3.4. Practical Applicability and Scalability

The proposed framework demonstrates practical potential for deployment in Advanced Metering Infrastructure systems due to its ability to process time-series smart meter data automatically with minimal manual intervention. Although the training phase requires moderate computational resources, inference and detection operations are relatively lightweight once the models are deployed. The framework can therefore support near real-time theft monitoring in smart grid environments.

Furthermore, the modular ensemble design allows future scalability by integrating additional deep learning or edge-computing techniques for large-scale AMI applications. The current study is limited by the relatively small dataset size and the use of simulated theft patterns. Future work should incorporate larger real-world datasets from utility providers and evaluate system performance under live AMI network conditions.

3.5. Computational Overhead

Although the proposed hybrid framework demonstrates strong detection performance, combining multiple deep learning architectures inevitably increases computational overhead during both training and inference. The 1D-CNN component performs convolution and feature extraction operations, the BiLSTM processes temporal dependencies in both forward and backward directions, and the LSTM autoencoder introduces an additional reconstruction stage for anomaly confirmation. Consequently, the overall framework requires more memory, processing power, and execution time than single-model approaches.

However, it is important to note that the most computationally intensive stage is the offline training phase, which is performed centrally using utility servers or cloud-based computing platforms. Once training is completed, only the optimized model parameters are deployed for inference. In practical AMI implementations, smart meters are not expected to execute the complete deep learning pipeline locally. Instead, consumption data are transmitted to data concentrators, edge gateways, or utility control centers where theft detection algorithms are executed using more capable computational resources.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

For the dataset used in this study, the model was trained on approximately 1,250 sequences, and the inference process for a single consumption sequence required only a fraction of the total training time. Since theft detection is typically performed on hourly, daily, or multi-day consumption windows rather than on a millisecond timescale, the computational requirements remain acceptable for near real-time monitoring applications.

Furthermore, several optimization techniques can be employed to facilitate deployment in resource-constrained environments, including model pruning, weight quantization, knowledge distillation, and edge-cloud collaborative processing. These techniques can significantly reduce memory footprint and computational complexity while maintaining acceptable detection performance. Therefore, the proposed framework is better suited for deployment at the AMI head-end system, edge server, or utility data center rather than directly within individual smart meters.

4. CONCLUSION

This project successfully developed a hybrid 1D CNN-BiLSTM and LSTM autoencoder system for detecting electricity theft in AMI, achieving 92.47% accuracy, 90.78% precision, 93.81% recall, and 92.27% F1-score. The system addresses non-technical losses in smart grids by rapidly identifying theft patterns in 14-day consumption sequences. The CNN-BiLSTM classifier captures local and temporal anomalies, while the autoencoder enhances detection through reconstruction errors, making the system robust and efficient. The results advance AI-driven theft detection, offering utilities a scalable solution to reduce revenue losses.

This framework is limited by the quality of the input data, the computational cost of training and little emphasis on integration with real-time AMI networks. Future works to enhance this framework could include research into deep learning architectures that combine with other deep learning methods to enhance prediction accuracy as well as automated reaction mechanisms to allow for the mitigation of theft in real time.

5. ACKNOWLEDGMENT

The authors will like to acknowledge and appreciate Electrical Engineering Department Staff at Ahmadu Bello University, Zaria.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. AUTHORS' CONTRIBUTION/ROLE

Ibrahim Abdulwahab: Conceptualization, Methodology; Oyatayo Habeebat Abike: Writing Original draft, Investigation; Sulaiman Haruna Sulaiman: Editing and Review; Ismaila Mahmud: Supervision and Editing; Abdulrahman: Review; Aminu Jibrin Aliyu: Formal Analysis; Rapheal Chibuike Nwobi: Software Review and Data Analysis.

8. DECLARATION OF THE USE GENERATIVE AI IN THE MANUSCRIPT PREPARATION PROCESS

Authors declare that did not use Generative AI in the manuscript preparation process upon submission of the paper.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

REFERENCES

[1] D. Joseph, J. Yusuf, I. Abdulwahab and A. Umar, "MPPT ALGORITHM FOR SOLAR-PV SYSTEM," *Journal of Intelligent System Telecommunication* vol. 2, no. 1, pp. 1-20, 2025.

[2] I. Abdulwahab, A. S. Abubakar, A. Olaniyan, B. O. Sadiq and S. A. Faskari, "Control of Dual Stator Induction Generator Based Wind Energy Conversion System," in *2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON)*, 2022, pp. 1-5: IEEE.

[3] I. Abdulwahab, S. A. Onazi, G. Rimamfate, A. Abdulwasiu, A. Umar and N. Iliyasu, "Optimization of the Offshore Wind Turbines Layout Using Cuckoo Search Algorithm," *Journal of Techniques,* vol. 6, no. 2, pp. 90-99, 2024.

[4] M. Nabil, M. Ismail, M. Mahmoud, M. Shahin, K. Qaraqe and E. Serpedin, "Deep recurrent electricity theft detection in AMI networks with random tuning of hyper-parameters," in *2018 24th International Conference on Pattern Recognition (ICPR)*, 2018, pp. 740-745: IEEE.

[5] P. M. Kgaphola, S. M. Marebane and R. T. Hans, "Electricity theft detection and prevention using technology-based models: A systematic literature review," *Electricity,* vol. 5, no. 2, pp. 334-350, 2024.

[6] M. S. Iqbal, S. Munawar, M. Adnan, A. Raza, M. A. Akbar and A. Bermak, "A critical review of technical case studies for electricity theft detection in smart grids: A new paradigm based transformative approach," *Energy Conversion Management: X,* vol. 26, p. 100965, 2025.

[7] A. Babaita, A. Mati, Y. Jibril, A. Kunya and I. Abdulwahab, "DEVELOPMENT OF A LOAD FREQUENCY CONTROL SCHEME FOR AN AUTONOMOUS HYBRID MICROGRID," *Zaria Journal of Electrical Engineering Technology,* vol. 11, no. 1, 2022.

[8] I. Abdulwahab, S. Faskari, T. Belgore and T. Babaita, "An improved hybrid micro-grid load frequency control scheme for an autonomous system," *FUOYE Journal of Engineering Technology,* vol. 6, no. 4, pp. 369-374, 2021.

[9] C. Clastres, "Smart grids: Another step towards competition, energy security and climate change objectives," *Energy policy,* vol. 39, no. 9, pp. 5399-5408, 2011.

[10] I. Abdulwahab, N. A. Iliyasu, E. K. Rasaq, A. Jibrin, A. M. Mahmudul and H. R. Kannan, "Optimized Unified Power Quality Conditioner for Nonlinear Load Connected Microgrid Using Artificial Rabbit Search Optimization," in *2025 IEEE Industrial Electronics and Applications Conference (IEACon)*, 2025, pp. 143-147: IEEE.

[11] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen and X. Shen, "Energy-theft detection issues for advanced metering infrastructure in smart grid," *Tsinghua Science Technology,* vol. 19, no. 2, pp. 105-120, 2014.

[12] M. Lamb, "Advanced metering infrastructure: Continued evolution and opportunities to deliver greater value," *Climate Energy,* vol. 41, no. 7, pp. 14-20, 2025.

[13] N. M. Elshennawy, D. M. Ibrahim and A. M. J. S. R. Gab Allah, "An efficient electricity theft detection based on deep learning," *Scientific Reports,* vol. 15, no. 1, p. 12866, 2025.

[14] Y. M. Rind, M. H. Raza, M. Zubair, M. Q. Mehmood and Y. Massoud, "Smart energy meters for smart grids, an internet of things perspective," *Energies,* vol. 16, no. 4, p. 1974, 2023.

[15] H. Xu, P. Li, J. Wang and W. Liang, "A study on black screen fault detection of single-phase smart energy meter based on random forest binary classifier," *Measurement control,* vol. 242, p. 116245, 2025.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X

[16] N. M. Elshennawy, D. M. Ibrahim and A. M. Gab Allah, "An efficient electricity theft detection based on deep learning," *Scientific Reports,* vol. 15, no. 1, p. 12866, 2025.

[17] U. Haruna, B. L. Pal, A. S. Dhabariya, F. Rasheed, A. F. Shah, A. Sani, B. S. Mu'azu and A. A. Yahya, "REVIEW ON TEMPORAL CONVOLUTIONAL NETWORKS FOR ELECTRICITY THEFT DETECTION WITH LIMITED DATA," *Technology audit,* vol. 7, no. 3, pp. 94-106, 2024.

[18] A. Tirulo, S. Chauhan and K. Dutta, "Machine learning and deep learning techniques for detecting and mitigating cyber threats in IoT-enabled smart grids: a comprehensive review," *International Journal of Information Computer Security,* vol. 24, no. 3-4, pp. 284-321, 2024.

[19] A. Otuoze, M. Mustafa, U. Sultana, E. Abiodun, B. Jimada-Ojuolape, O. Ibrahim, I. Avazi-Omeiza and A. Abdullateef, "Detection and confirmation of electricity thefts in Advanced Metering Infrastructure by Long Short-Term Memory and fuzzy inference system models," *Nigerian Journal of Technological Development,* vol. 21, no. 1, pp. 112-130, 2024.

[20] A. N. Alkuwari, S. Al-Kuwari, A. Albaseer and M. Qaraqe, "Anomaly detection on smart grids with optimized convolutional long short-term memory model," *IEEE Access,* 2025.

[21] M. Žarković and G. Dobrić, "Artificial intelligence for energy theft detection in distribution networks," *Energies,* vol. 17, no. 7, p. 1580, 2024.

[22] R. Khalid, O. Samuel, N. Javaid, A. Aldegheishem, M. Shafiq and N. Alrajeh, "A secure trust method for multi-agent system in smart grids using blockchain," *IEEE Access,* vol. 9, pp. 59848-59859, 2021.

[23] R. Alfatikarani, "A Hybrid CNN-LSTM Approach for Sentiment Analysis of Poverty Issues in East Java," *Journal of Intelligent System Telecommunication,* vol. 2, no. 1, pp. 19-28, 2025.

[24] J. Zhang, L. Ye and Y. Lai, "Stock price prediction using CNN-BiLSTM-Attention model," *Mathematics,* vol. 11, no. 9, p. 1985, 2023.

[25] B. Song, Y. Liu, J. Fang, W. Liu, M. Zhong and X. Liu, "An optimized CNN-BiLSTM network for bearing fault diagnosis under multiple working conditions with limited training samples," *Neurocomputing,* vol. 574, p. 127284, 2024.

[26] M. N. Hasan, R. N. Toma, A.-A. Nahid, M. M. Islam and J.-M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," *Energies,* vol. 12, no. 17, p. 3310, 2019.

[27] A. A. Ahmed, W. Ali, T. A. Abdullah and S. J. Malebary, "Classifying cardiac arrhythmia from ECG signal using 1D CNN deep learning model," *Mathematics,* vol. 11, no. 3, p. 562, 2023.

DOI: https://doi.org/10.26740/jistel.v2n2.p124-139
e- ISSN 3090-613X