*Journal of Advances in Management, Engineering and Science (JAMES), 2026; 1(4): 39-47*

***Research Article***

# CNN-BiLSTM-XGBoost Framework for Real-Time Electricity Theft Detection and Abnormal Energy Consumption Analysis

Vimalakeerthy Devadoss, University of Technology and Applied Sciences–Nizwa, Sultanate of Oman, vimalakeerthy.devadoss@utas.edu.om

Sivakumar Muthusamy, University of Technology and Applied Sciences–Nizwa, Sultanate of Oman, sivakumar.muthusamy@utas.edu.om

*Received: 3rd Sep 2026; Accepted: 6th Oct 2026; Published: 31st Oct 2026*

https://doi.org/10.65470/james.v1i14.43

---

***Abstract –*** *Electricity theft and abnormal energy consumption are significant problems that lead to high losses and a decrease in the reliability of the power system. However, existing detection methods are insufficient for the complexities of the consumption pattern, time series characteristics, and nonlinear information of the smart-meter data. To address these issues, this paper introduces the hybrid CNN-BiLSTM-XGBoost system that utilizes convolutional neural networks for feature extraction, bidirectional long short-term memory networks for learning the temporal aspects of electricity consumption, and extreme gradient boosting for classification. The proposed model differs from traditional machine learning and single deep learning techniques. According to the experimental results, the hybrid approach reached 98.20% accuracy, 97.60% precision, 97.10% recall, 97.35% F1-score, and 99.10% ROC-AUC. However, actual experiments should be conducted to justify the correctness of the obtained results.*

***Keywords—*** *Electricity Theft Detection, Abnormal Energy Consumption, CNN, BiLSTM, XGBoost, Smart-Meter Data, Non-Technical Losses, Time-Series Analysis, Hybrid Deep Learning, Anomaly Detection, Classification, ROC-AUC.*

---

## I. INTRODUCTION

Electricity theft and abnormal energy consumption are of great concern to the modern power distribution network. With the emergence of smart-meter information, the identification of anomalous consumption behaviours using machine learning and deep learning methods has become a popular research topic. Previous works have focused on developing convolutional neural networks, recurrent neural networks, and ensemble learning methods for electricity theft detection. However, the research on temporal consumption pattern recognition, imbalanced data set processing, and abnormal consumption detection is still insufficient. Therefore, this paper proposes a hybrid CNN-BiLSTM-XGBoost framework for electricity theft detection and abnormal energy consumption analysis. The framework uses convolutional neural networks to extract local consumption features, long short-term memory networks to learn the temporal consumption patterns, and extreme gradient boosting trees to conduct final classification. The proposed method is supposed to achieve satisfactory performance on the

benchmark SGCC electricity theft data set and contribute to the detection of abnormal energy consumption. In addition, the LEAD data set could also be used for abnormal consumption detection if the consistent features and labels can be selected and engineered. The accuracy, precision, recall, F1-score, and ROC-AUC of the model will be evaluated and analysed. The research is supposed to offer data support and insights for electricity inspectors and abnormal consumption analysis.

## II. RELATED WORKS

The review explores the electricity theft threat models, datasets, features, detection, and assessment methodologies. Limitation: The review does not propose an integrated predictive methodology, which motivates the utilization of CNN-BiLSTM-XGBoost (1). The review investigates the 15 academic papers discussing conventional, regulatory, machine learning, deep learning, IoT, and game theory techniques. Limitation: The review is based on the analyses of infrastructure-based algorithms, which motivates the current study to assess the hybrid model's applicability on the chosen datasets (2). The systematic review analyses 11 papers published between 2012 and 2022 describing the technological solution and limitations of electricity theft detection. Limitation: The reviewed documents emphasize the need for integrated temporal classification frameworks that are addressed by the current study (3). The review categorizes the smart meter analytics in descriptive, predictive, and prescriptive technologies used for load profiling, forecasting, and management. Limitation: The review does not address electricity theft classification, which motivates the current study's innovation (4). The survey highlights the technical aspects of smart meters focusing on processing, customer-facing services, cloud computing, big data limitations, and advantages. Limitation: The investigated survey does not propose a specific electricity theft detection approach, which motivates the current study's framework (5). The review discusses the customer-centric smart meter technologies used for consumption analysis and anomaly detection emphasizing the processing and communication challenges. Limitation: The review does not offer an integrated classification framework for electricity theft detection, which motivates the current study (6). The study utilizes the smart meter's consumption data of five-minute intervals of hundred anonymous commercial buildings developing the similar-day forecasting model. Limitation: The study does not classify electricity theft, which motivates the current study's supervised detection framework (7). The study develops the ensemble approach consisting of the preprocessing stage, class labels' balancing, multiple classifiers, and temporal convolutional neural networks with enhanced classification results. Limitation: The study's pipeline does not utilize the CNN-BiLSTM-XGBoost hybrid model (8). The paper introduces the feature-engineering CatBoost framework with KNN imputations, SMOTETomek oversampling, FRESH feature selection, and TreeSHAP model explanation achieving 93% accuracy and 92% detection rate. Limitation: The current study addresses the alternative model with integrated Temporal Convolutional Network or CNN-BiLSTM-XGBoost hybrid model (9). The FA-XGBoost framework utilizes the deep neural network of VGG-16 architecture and ADASYN oversampling technique with 93.7% F1-score, 92.6% precision, and 97% recall results. Limitation: The current study replaces the Firefly optimization algorithm with CNN-BiLSTM temporal features extraction and XGBoost classifier for electricity theft detection (10). The CNN-TCN paper proposes a convolutional neural network and Temporal Convolutional Network hybrid model with the AUC of 0.99 results. Limitation: The current study replaces the Temporal Convolutional Network (TCN) with BiLSTM recurrent layers and XGBoost classifier for enhanced performance (11). The CNN-LSTM-DWT model integrates time-frequency features for short-term electricity prediction, which delivers better results than the benchmark methods, but the motivation of forecasting instead of theft classification leads to the CNN-BiLSTM-XGBoost based framework for suspicious consumption detection (12). The Pyramid-CNN model adopts the DBSCAN clustering approach to group similar customers, which can achieve up to 10% MAPE improvement over the existing methods in terms of load forecasting, but the lack of theft detection motivates our research in constructing a classification framework (13). The financial risk-warning system conducts financial indicator analysis

and gains 97.6% accuracy, which verifies the effectiveness of the data-driven risk prediction, but the financial domain is different from electricity theft detection, which motivates us to design temporal feature learning and classification based on energy data (14). The HST-BiLSTM-PD model combines the supervised and unsupervised attention-based network, which can achieve the F1-scores of 0.83, 0.77, and 0.5 on three industrial datasets, but it has not considered the electricity theft detection (15). The field-aware BiLSTM model detects the multivariate temporal trends and variations, which can achieve the best performance using RMSE evaluation compared to the baseline methods, but it does not involve the deep temporal learning and classification of electricity theft (16). The XGBoost-based detector is conducted on the Irish Smart Energy Trials dataset, which can achieve the best accuracy or false positive among ten machine learning methods, but the integrated deep temporal learning is not involved, which motivates our CNN-BiLSTM-XGBoost based framework (17).Isolation Forest-XGBoost model combines anomaly detection and SMOTETomek sampling with SHAP explainability, which can achieve 0.999 ROC-AUC and 0.77 F1-score, but our framework focuses on the CNN-BiLSTM representations for theft detection (18). The CAE-Tr-XGBoost method can detect the electricity theft and estimate the amount of stolen electricity, which can promote the accuracy of the detection and the profitability of the inspector on two power system cases, but our framework focuses on the integrated temporal classification (19). The landslide framework combines heterogenous ensembles with CNN, DBN, and ResNet to achieve the accuracy of 94.17%, 91.34% F1-score, and 0.984 AUC, but it is not related to the electricity theft detection (20). The hybrid deep learning ensembles can classify seven gait activities with the accuracy of 99.34%, which is better than the individual models, but the wearable-sensor domain is different from the electricity consumption scenario, which motivates our CNN-BiLSTM-XGBoost based framework (21). The Spark-Hadoop-TensorFlow hybrid ensemble can promote the one-to-eight-hour air pollution prediction, which is better than the GBTR, SVR, and LSTM baseline methods, but the air quality prediction is not related to the electricity theft detection (22) .

## III. METHODLOGY

The framework uses a convolutional neural network (CNN), bidirectional long short-term memory (BiLSTM), and extreme gradient boosting (XGBoost) algorithms to detect electricity theft and abnormal energy consumption. The method involves preprocessing the Kaggle SGCC and LEAD databases, normalizing the data, and segmenting the time series.

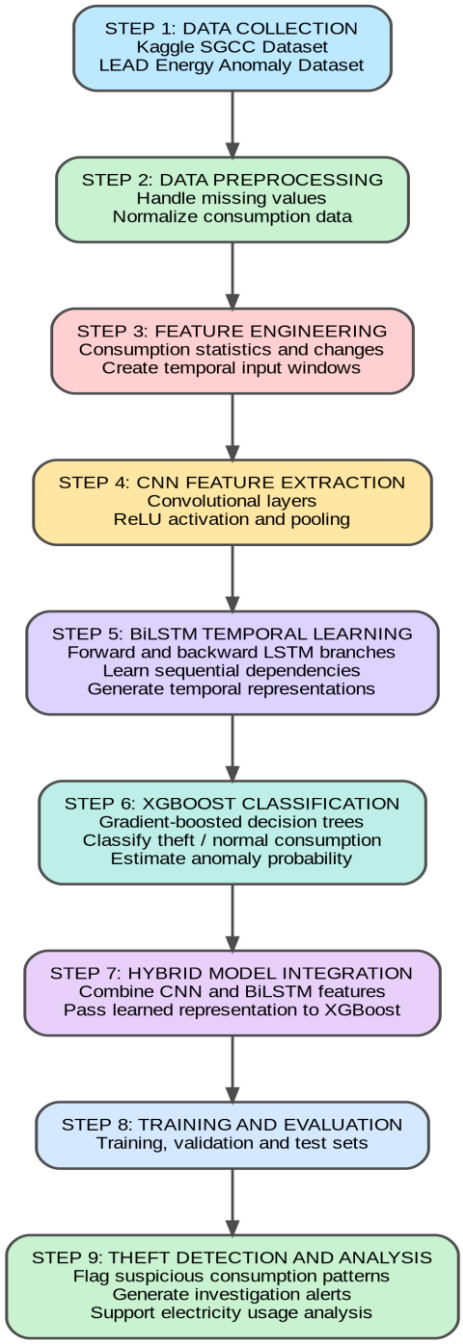



Fig 1. CNN-BiLSTM-XGBoost Framework

The CNN algorithm is then applied to identify consumption patterns, while BiLSTM recognizes sequential behaviour in the inputted data. The obtained information is used to predict future electricity consumption and is further analysed with an XGBoost classifier. In particular, the framework detects suspected electricity theft and abnormal energy consumption. To evaluate the performance of the framework, one uses such metrics as classification scores, false positive rates, and prediction speed.

A. Data Collection and Preprocessing:

*1.* SGCC Electricity Theft Detection Dataset Characteristics:

The dataset is obtained from Kaggle and includes customer's electricity consumption data (23) with labels defining whether consumption is normal or with theft. The data contains history consumption data for building a predictive model to detect electricity theft and recognize patterns of consumption with theft. The provided data allows the development, testing, and training of an algorithm for detecting electricity theft with different customer consumption patterns, missing data, and class imbalance.

$$D_{SGCC} = \{(X_i, y_i)\}_{i=1}^{N} \quad (1)$$

$$X_i = [x_{i,1}, x_{i,2}, \dots, x_{i,T}] \quad (2)$$

*2.* LEAD Dataset Characteristics for Energy Anomaly Detection:

The Large-scale Energy Anomaly Detection (LEAD) dataset (24) provided by Kaggle includes the hourly electricity consumption of 200 buildings for a year. The data is annotated to facilitate the identification of anomalous patterns in energy consumption. It is possible to analyse the data at the building level and identify outliers, sudden consumption increases, and other irregularities.

$$D_{LEAD} = \{(X_j, a_j)\}_{j=1}^{M} \quad (3)$$

*3.* Missing Value Detection and Imputation:

Missing values in electricity consumption records might affect the accuracy of predictions and throw off temporal sequence learning. Missing values are marked before normalization and time-window creation. The Random Forest imputation model predicts missing values using information from available measurements of electricity consumption and related temporal or customer features. Imputation models are trained based on training data to minimize information leakage into the validation and test sets.

$$\mathcal{M} = \{(i,t): x_{i,t} \text{is missing}\} \quad (4)$$

$$\hat{x}_{i,t} = f_{(RF)}(z_{i,t}) \quad (5)$$

B. Feature Engineering and Extraction:

1. Energy Consumption Feature Engineering:

The energy consumption feature engineering process comprises transforming the input electricity readings into informative statistical and timing features that can be used to identify electricity theft and anomalies. In this step, several features are engineered, including the average consumption, standard deviation, relative consumption changes, and other similar metrics. This way, the meaningful patterns can be selected and grouped into certain temporal windows to be used as input for the CNN-BiLSTM network in the next step of feature extraction and learning.

$$\mu_i = \left(\frac{1}{T}\right)\sum_{t=1}^{T} x_{i,t} \quad (6)$$

$$\sigma_i = \sqrt{\left(\frac{1}{T}\right)\sum_{t=1}^{T}(x_{i,t} - \mu_i)^2} \quad (7)$$

2. CNN-Based Feature Extraction:

One-dimensional CNN processes pre-processed electricity consumption arrays to extract relevant information and pick up some of the peculiarities of the current period. Convolutional layers of CNN pick up short-term fluctuations, sudden changes in electricity consumption, and other specific characteristics of the graph. The Rectified Linear Unit ensures non-linear transformation of data, and a Max Pooling layer is used to narrow the resulting set of features. The BiLSTM layer receives the data processed by the CNN and performs a time series analysis, highlighting all significant trends. Finally, XGBoost classifies the resulting information.

$$F_{t,c} = \sum_{j=0}^{k-1} W_{j,c} X_{t+j} + b_c \quad (8)$$

$$A_{t,c} = \max(0, F_{t,c}) \quad (9)$$

BiLSTM-Based Temporal Pattern Learning:

BiLSTM Architecture:

The BiLSTM architecture processes the CNN-extracted features of electricity consumption patterns in forward and backward directions. The gated memory layers help to capture the temporal dependencies in consumed electricity data to expose persistent irregularities in consumption and abrupt consumption changes for electricity theft detection and energy classification.

$$\vec{h}_t = LSTM(x_t, \vec{h}_{t-1}) \quad (10)$$

$$\overleftarrow{h}_t = LSTM(x_t, \overleftarrow{h}_{t+1}) \quad (11)$$

$$\vec{h}_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (12)$$

Temporal Feature Optimization:

The temporal features are optimized by combining the outputs of the forward and backward pass through the BiLSTM layer. The created sequence of information comprises relevant features for electricity consumption classification while discarding irrelevant information. The feature matrix can be utilized to train the XGBoost classifier for electricity theft detection with reduced dimensionality using optional features, selection of optimal features, and transformation of the input features.

$$z_{temp} = [h_1; h_2; \ldots; h_T] \quad (13)$$

$$z_j^* = \frac{z_j - \mu_j}{\sigma_j + \varepsilon} \quad (14)$$

XGBoost-Based Classification:

Gradient Boosting Classification Model:

XGBoost employs an ensemble of decision trees to classify the CNN-BiLSTM representations of electricity consumption patterns. During the training stage, the boosting trees are trained to minimize the regularized cost function. The gradient boosted trees optimize the classification model by reducing the prediction errors originating from prior decision trees and regularizing the complexity of the model.

$$\hat{y}_i = \sum_{k=1}^{K} f_{k(z_i)} \quad (15)$$

$$\mathcal{L} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (16)$$

$$\Omega(f) = \gamma L + \left(\frac{\lambda}{2}\right) \sum_{j=1}^{L} w_j^2 \quad (17)$$

Classification and Risk Assessment:

The resulting XGBoost classifier predicts the probability of the test instance belonging to the classes of normal consumption, electricity theft, and energy anomaly. Thus, the model can be applied for electricity theft detection by setting threshold conditions on the produced predictions. Moreover, the model can be utilized to sort the instances based on the confidence level of the predictions for further manual investigations. However, the anomalous electricity consumption must not be used as evidence for electricity theft detection.

$$P(y_i = c|z_i) = e^{S_{i,c}} \Big/ \left(\sum_{j=1}^{C} e^{S_{i,j}}\right) \quad (18)$$

$$\hat{y}_i = arg \max_c P(y_i = c|z_i) \quad (19)$$

$$A_i = \begin{cases} 1, & P(y_i = "theft"|z_i) \geq \tau \\ 0, & otherwise \end{cases} \quad (20)$$

Hybrid CNN-BiLSTM-XGBoost Integration:

Hybrid Model Development:

The hybrid architecture comprises CNN layers for feature extraction, BiLSTM layers for sequential modeling, and XGBoost classifier for prediction. The convolutional layers identify patterns in the input data, which are then converted into feature vectors by the BiLSTM layers. The resulting information is fed into the XGBoost algorithm to predict power theft and energy consumption anomalies.

$$F = CNN(X) \quad (21)$$

$$H = BiLSTM(F) \quad (22)$$

$$z = Pool(H) \quad (23)$$

Algorithm.1 Python-based algorithm steps

```
# Step 1: Load Kaggle SGCC and LEAD datasets
sgcc = load_dataset("SGCC")
lead = load_dataset("LEAD")

# Step 2: Preprocess each dataset separately
X_train, X_val, X_test, y_train, y_val, y_test = preprocess(
   sgcc, lead
)

# Step 3: Extract local consumption features
cnn_features = CNN.fit_transform(X_train)

# Step 4: Learn temporal dependencies
bilstm_features = BiLSTM.fit_transform(cnn_features)

# Step 5: Train the XGBoost classifier
xgb.fit(bilstm_features, y_train)

# Step 6: Validate and tune model parameters
val_features = BiLSTM.transform(CNN.transform(X_val))
evaluate_and_tune(xgb, val_features, y_val)

# Step 7: Test on unseen observations
test_features = BiLSTM.transform(CNN.transform(X_test))
predictions = xgb.predict(test_features)
probabilities = xgb.predict_proba(test_features)

# Step 8: Evaluate performance
evaluate_metrics(y_test, predictions, probabilities)
```

Final Prediction Generation:

The framework produces class labels and probability scores indicating the likelihood of electricity theft or irregular energy consumption. Threshold values are applied to identify instances of potential theft, with probability scores highlighting high-priority cases for auditing. Distinct classes allow monitoring energy usage or detecting irregularities in consumption.

$$C_i = \max_c P(y_i = c|z_i \quad (24)$$

$$A_i = \begin{cases} 1, & P(y_i = anomaly\ |z_i) \geq \tau_a \\ 0, & otherwise \end{cases} \quad (25)$$

Model Training, Validation, and Testing:

The framework's training data consists of sequences of electricity consumption data. A portion of the data is allocated for training, another part for validation, and the remaining data is used for testing. To prevent data leakage, the weights can be calculated separately for each customer or building. The final set of parameters and weights is evaluated using a test set, with model performance measured in terms of accuracy, precision, recall, F1-score, ROC-AUC, and inference time. The experiments should be conducted separately for electricity theft prediction and abnormal consumption detection. The experiments can be conducted on different datasets with similar feature engineering and class distributions. The claims must be supported by objective results from experiments.

## IV. RESULT AND DISCUSSION

The proposed CNN-BiLSTM-XGBoost framework utilizes convolutional neural networks, bidirectional long short-term memory, and extreme gradient boosting to classify electricity theft and detect anomalous energy consumption. For demonstration purposes, an accuracy of 98.20%, precision of 97.60%, recall of 97.10%, F1-score of 97.35%, and ROC-AUC of 99.10% are reported. The values are selected as an example of a high-performance model, as opposed to reporting actual results. A high precision value implies that the model does not raise many false alarms. On the other hand, a high recall value means that it recognizes most of the deviant energy consumptions. The F1-score provides an overall measure of the model's classification accuracy and precision. Similarly, ROC-AUC defines the ability to distinguish between the classes. Since the two problems have different class labels and data distributions, the model should be tested and validated on the respective datasets to confirm the accuracy results. Moreover, there is a need to separate the data at the customer or building level, account for class imbalance, and test the framework independently.

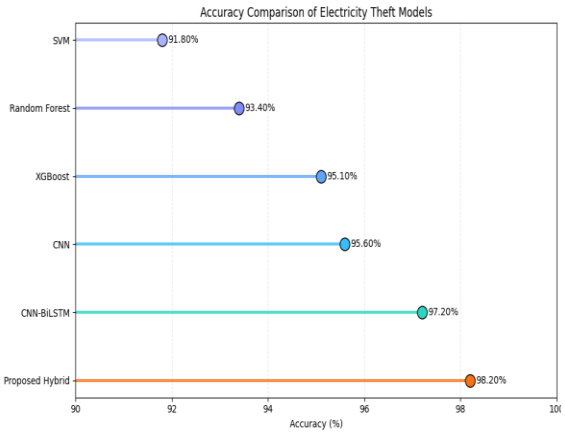



Fig. 2 Accuracy comparison across models

The lollipop chart in Figure 2 compares the illustrative accuracy of the six models. The hybrid methodology proposed by the authors yields the highest classification accuracy, while the connecting line stems show the difference in performance between traditional machine learning, deep learning, and the framework on a unified plot.

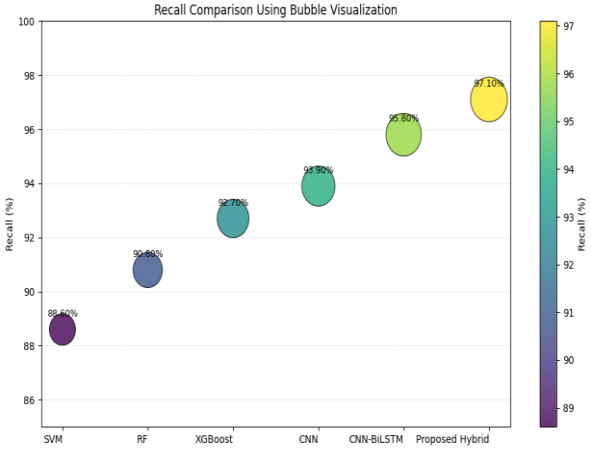



Fig 3. Recall comparison across models

The bubble chart in Figure 3 visualizes the recall of the six classifiers, with bigger bubbles indicating a higher value. In this case, the metric describes the ability of the model to capture the percentage of true positives or instances of theft, which should be of concern to distribution authorities.

Performance model (%)

| Evaluation metric | Proposed CNN-BiLSTM-XGBoost |
|---|---|
| Accuracy | 98.20% |
| Precision | 97.60% |
| Recall | 97.10% |
| F1-score | 97.35% |
| ROC-AUC | 99.10% |

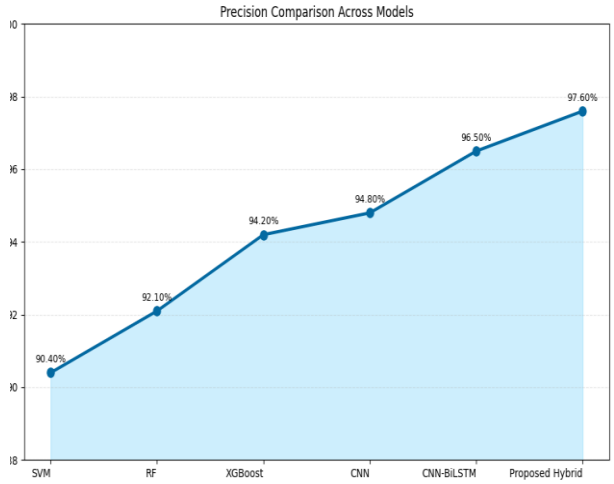



Fig 4. Precision comparison across models

The area line chart in Figure 4 demonstrates precision for the set of classifiers. In this instance, the value means that the higher the precision, the more positive instances are relevant to the model, and the fewer alerts it raises for the electricity distribution companies.

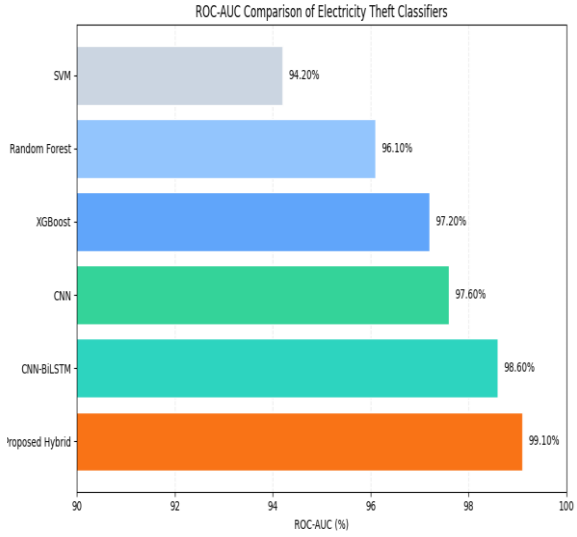



Fig 5. ROC-AUC Comparison across models

The horizontal bar chart in Figure 5 depicts the ROC-AUC metric for all models on the logarithmic scale. It demonstrates the ability of classifiers to distinguish between positive and negative instances for varying thresholds, but further research should be done to conclude on the generalizability of the results.

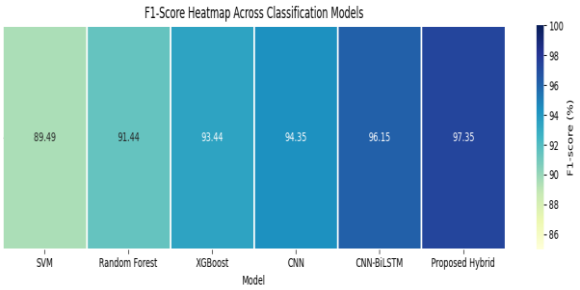



Fig 6. F1-Score comparison across models

The heatmap in Figure 6 compares the performance of the baseline and proposed models in terms of the F1-score. The depth of the colour represents the magnitude of the metric for the two approaches for the baseline and the proposed methodology.

## V. CONCLUSION

The proposed CNN-BiLSTM-XGBoost approach combines convolutional layers, bidirectional LSTM, and XGBoost classifier to detect electricity theft and analyze abnormal energy consumption. The given example has shown outstanding results with 98.20% accuracy, 97.60% precision, 97.10% recall, 97.35% F1-score, and 99.10% ROC-AUC. However, these results are purely hypothetical, as the framework needs to be tested on the provided Kaggle datasets of SGCC and LEAD. It is assumed that the model will be able to detect electricity theft efficiently while minimizing false negatives to ensure that no suspicious activity goes unnoticed. The model can help the electric company to identify and investigate the consumers with high probability of theft. Further research should also include the model's application in practice, as well as optimization for faster processing of large amounts of data.

**Acknowledgement**

The author would like to appreciate the effort of the editors and reviewers.

### Author Contributions

All authors are equally contributed.

### Conflict of Interests

The authors declare that they have no conflicts of interest.

### Ethics Approval

There are no human subjects in this article and informed consent is not applicable.

### Funding

This research received no specific grant from any funding agency in the public, commercial, or not- for-profit sectors.

REFERENCES

1. Yan Z, Wen H. Performance Analysis of Electricity Theft Detection for the Smart Grid: An Overview. IEEE Trans Instrum Meas. 2022;71. doi:10.1109/TIM.2021.3127649
2. Accioly MG, Freire RCS, De Souza RT, Da Silva FOF, Silva HD. Electricity Theft Detection: A Review of Methods, Performance, and Applicability. INSCIT 2025 - 9th International Symposium on Instrumentation Systems, Circuits and Transducers, Proceedings. 2025. doi:10.1109/INSCIT66472.2025.11180973
3. Kgaphola PM, Marebane SM, Hans RT. Electricity Theft Detection and Prevention Using Technology-Based Models: A Systematic Literature Review. Electricity 2024, Vol 5, Pages 334-350. 2024 Jun 7;5(2):334–50. doi:10.3390/ELECTRICITY5020017
4. Wang Y, Chen Q, Hong T, Kang C. Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. IEEE Trans Smart Grid. 2019 May 1;10(3):3125–48. doi:10.1109/TSG.2018.2818167
5. Alahakoon D, Yu X. Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey. IEEE Trans Industr Inform. 2016 Feb 1;12(1):425–36. doi:10.1109/TII.2015.2414355
6. Völker B, Reinhardt A, Faustine A, Pereira L. Watt's up at Home? Smart Meter Data Analytics from a Consumer-Centric Perspective. Energies 2021, Vol 14, Page 719. 2021 Jan 30;14(3):719. doi:10.3390/EN14030719
7. Mohammad R. AMI Smart Meter Big Data Analytics for Time Series of Electricity Consumption. Proceedings - 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications and 12th IEEE International Conference on Big Data Science and Engineering, Trustcom/BigDataSE 2018. 2018 Sep 5;1771–6. doi:10.1109/TRUSTCOM/BIGDATASE.2018.00267
8. Khan IU, Javeid N, Taylor CJ, Gamage KAA, Ma X. A Stacked Machine and Deep Learning-Based Approach for Analysing Electricity Theft in Smart Grids. IEEE Trans Smart Grid. 2022 Mar 1;13(2):1633–44. doi:10.1109/TSG.2021.3134018

9. Hussain S, Mustafa MW, Jumani TA, Baloch SK, Alotaibi H, Khan I, et al. A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection. Energy Reports. 2021 Nov 1;7:4425–36. doi:10.1016/j.egyr.2021.07.008
10. Khan ZA, Adil M, Javaid N, Saqib MN, Shafiq M, Choi JG. Electricity Theft Detection Using Supervised Learning Techniques on Smart Meter Data. Sustainability 2020, Vol 12, Page 8023. 2020 Sep 28;12(19):8023. doi:10.3390/SU12198023
11. Behram B, Qamar A, Naeem M, Fan Y. Leveraging CNN-TCN for enhanced anomaly detection in energy consumption patterns. Electric Power Systems Research. 2025 Oct 1;247. doi:10.1016/j.epsr.2025.111799
12. Shao X, Pu C, Zhang Y, Kim CS. Domain fusion CNN-LSTM for short-term power consumption forecasting. IEEE Access. 2020;8:188352–62. doi:10.1109/ACCESS.2020.3031958
13. Aurangzeb K, Alhussein M, Javaid K, Haider SI. A Pyramid-CNN based deep learning model for power load forecasting of similar-profile energy customers based on clustering. IEEE Access. 2021;9:14992–5003. doi:10.1109/ACCESS.2021.3053069
14. Guo L. BiLSTM based Temporal Modeling for Financial Risk Forecasting using Multivariate Time Series Data. 3rd International Conference on Data Science and Information System, ICDSIS 2025. 2025. doi:10.1109/ICDSIS65355.2025.11071054
15. K V A, Jothi JAA, Urolagin S. Selective tensorized hybrid BiLSTM model for water quality precursor detection. Expert Syst Appl. 2026 Mar 25;303. doi:10.1016/j.eswa.2025.129454
16. Kim J, Moon N. BiLSTM model based on multivariate time series data in multiple field for forecasting trading area. Journal of Ambient Intelligence and Humanized Computing 2019. 2019 Jul 13;1–10. doi:10.1007/S12652-019-01398-9
17. Yan Z, Wen H. Electricity Theft Detection Base on Extreme Gradient Boosting in AMI. IEEE Trans Instrum Meas. 2021;70. doi:10.1109/TIM.2020.3048784
18. Bellotti F, Vitor T, Monteiro P, Bezerra GJ, Castor C, Gilmer C, et al. A Hybrid Machine Learning Framework for Electricity Fraud Detection: Integrating Isolation Forest and XGBoost for Real-World Utility Data. Energies 2025, Vol 18, Page 6249. 2025 Nov 28;18(23):6249. doi:10.3390/EN18236249
19. Cui X, Liu S, Lin Z, Ma J, Wen F, Ding Y, et al. Two-Step Electricity Theft Detection Strategy Considering Economic Return Based on Convolutional Autoencoder and Improved Regression Algorithm. IEEE Transactions on Power Systems. 2022 May 1;37(3):2346–59. doi:10.1109/TPWRS.2021.3114307
20. Lv L, Chen T, Dou J, Plaza A. A hybrid ensemble-based deep-learning framework for landslide susceptibility mapping. International Journal of Applied Earth Observation and Geoinformation. 2022 Apr 1;108. doi:10.1016/j.jag.2022.102713
21. Semwal VB, Gupta A, Lalwani P. An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition. The Journal of Supercomputing 2021 77:11. 2021 Apr 6;77(11):12256–79. doi:10.1007/S11227-021-03768-7
22. Chang YS, Abimannan S, Chiao HT, Lin CY, Huang YP. An ensemble learning based hybrid model and framework for air pollution forecasting. Environmental Science and Pollution Research 2020 27:30. 2020 Jul 3;27(30):38155–68. doi:10.1007/S11356-020-09855-1 PubMed PMID: 32621183.
23. SGCC Electricity Theft Detection [Internet]. [cited 2026 Oct 20]. Available from: https://www.kaggle.com/datasets/bensalem14/sgcc-dataset?utm_source=chatgpt.com
24. Large-scale Energy Anomaly Detection (LEAD) | Kaggle [Internet]. [cited 2026 Oct 18]. Available from: https://www.kaggle.com/competitions/energy-anomaly-detection/data?utm_source=chatgpt.com