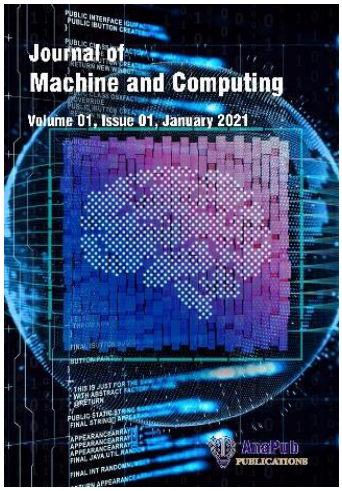
Journal Pre-proof

A Novel Deep Adaptive Feature Learning Framework for Efficient Electricity Theft Detection in Smart Grids



Deepa K R and Thillaiarasu N

DOI: 10.53759/7669/jmc202505177

Reference: JMC202505177

Journal: Journal of Machine and Computing.

Received 23 April 2025

Revised from 06 June 2025

Accepted 29 July 2025

Please cite this article as: Deepa K R and Thillaiarasu N, ‘‘A Novel Deep Adaptive Feature Learning Framework for Efficient Electricity Theft Detection in Smart Grids”, Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505177.

AnaPub
PUBLICATIONS

A Novel Deep Adaptive Feature Learning Framework for Efficient Electricity Theft Detection in Smart Grids

Deepa K R
School of Computing and Information Technology, REVA University, Bangalore -64
deepas.ravi@gmail.com

Thillaiarasu N
School of Computing and Information Technology,REVA University, Bangalore -64
Thillai888@gmail.com

Corresponding author:Deepa K R

Abstract

Electricity theft in **smart grids** poses a significant threat to energy security, leading to billions in financial losses and grid instability worldwide. Traditional detection methods, including hardware-based solutions and machine learning (ML) models, are often costly, reliant on labeled data, and lack scalability. Deep learning (DL) approaches, while more advanced, face challenges such as overfitting to static datasets and inefficiency in adapting to evolving consumption patterns and new cyberattacks, requiring frequent and computationally expensive retraining.In this context, we propose a novel deep learning framework, **Deep Adaptive Feature Learning for Theft Detection (DAFL-TD)**, tailored for smart grid environments. The architecture of DAFL-TD integrates a **Temporal Feature Extraction Network (TFEN)**, which captures temporal dependencies in electricity usage, with an **Adaptive Feature Learning Network (AFLN)** that leverages both labeled and unlabeled data for adaptive feature extraction and classification. The novelty of DAFL-TD lies in its ability to handle fluctuating, imbalanced data and dynamically update its feature representation without the need for extensive retraining, making it highly scalable for real-world smart grid applications. Extensive evaluations on the **State Grid Corporation of China (SGCC)** dataset demonstrate that DAFL-TD achieves a **13.84% improvement in AUC** compared to state-of-the-art models, along with superior precision as measured by MAP metrics. These results underline the efficacy of DAFL-TD as a robust, scalable, and efficient solution for real-time electricity theft detection, significantly enhancing the resilience and security of smart grids.

Keywords: Smart Grids, Electricity Theft Detection, Deep Learning, Temporal Feature Extraction, Adaptive Feature Learning

1. Introduction

The smart grid (SG) is a cutting-edge improvement over conventional electricity grid systems that aims to improve and regulate grid operations, guarantee dependable energy distribution, and evaluate the system's overall performance. The transmission and distribution networks, the advanced metering infrastructure (AMI) network, the electricity producing stations, and the system operator (SO) are the

parts that make up the SG architecture. The purpose of the AMI is to enable effective two-way communication between the smart meters (SMs) placed in residential properties and the System Manager (SM) [1], [2]. Unlike traditional monthly billing, SG collects comprehensive electricity use data from SMs, which is collected at intervals of a few minutes. It then sends this information to the SM using AMI. These signals may be used by the SM for effective control of electricity generation, load forecasting and monitoring, dynamic price computation for billing usage, and demand response management. Figure 1 presents the smart grid.

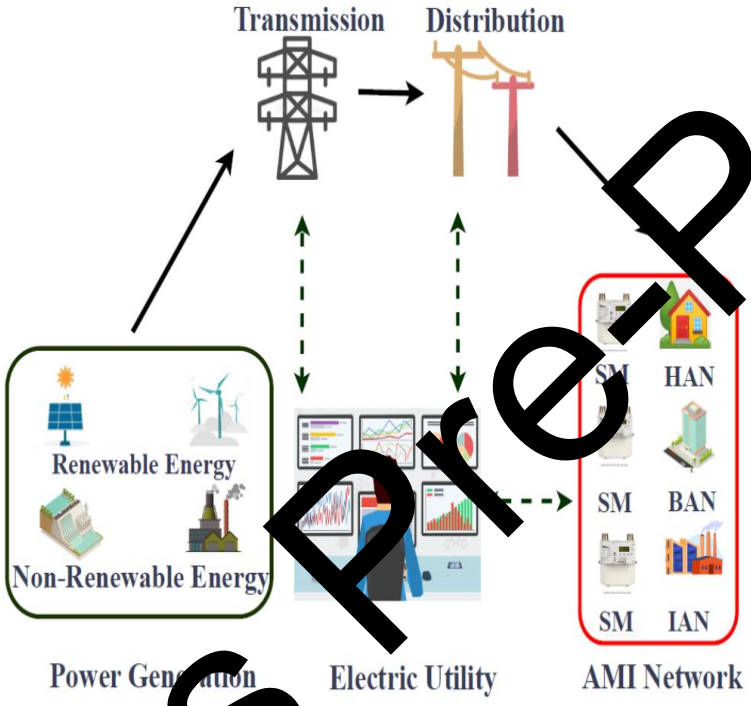


Figure 1 smart grid

Daily survival cannot be accomplished only by the usage of energy. Energy losses usually occur in energy distribution, transmission, and generating systems. Technical Losses (TLs) and Non-Technical Losses (NTLs) are the two types of electrical losses [1]. Net total losses (NTLs) are defined as the difference between total losses and electrical thefts, which account for the majority of total losses. The TL is essential to the flow of electricity since it is triggered by internal mechanisms found in the gearbox lining, transformers, and other electricity plant components. The calculation or assessment of non-technical losses arising from deliberate manipulation of reported electricity use figures is a major difficulty. This leads to incorrect billing, which may have serious negative financial and economic effects on many nations across the world. According to available statistics, non-technical problems cost the US, UK, and Canada, respectively, $6 billion, $173 million, and $100 million annually in financial losses [3]. Furthermore, a recent research conducted across 138 countries shows that non-technical losses and electricity theft cause utilities to lose $101.2 billion in revenue each year [4]. In addition to causing monetary losses, electricity theft raises the possibility of blackouts by causing instability and disturbances in the grid [5]. Improving the electrical grid's intelligence and resistance to these kinds of assaults is crucial. This has led to the research on electricity theft detection (ETD) as one of the major research area for enhancing the smart grid performance [3][4].

Traditionally, methods based on hardware have been used to detect electricity theft. These systems achieve high levels of precision through the use of specialist equipment for the investigation of customer behavior or the state of electrical networks [4]. It did not fulfill electricity firms' criteria because of its high implementation and maintenance costs, low universality, and limited scalability. Data-driven approaches are becoming the industry standard for electricity theft detection because they make use of a wide range of real-time metrics. In order to determine consumption patterns, historical data is often analyzed using machine learning techniques [6] [7]. Electricity theft may be detected by integrating algorithms into an intelligent management platform.In order to address practical issues, artificial intelligence (AI) has been incorporated into the electrical sector in a number of ways [8], [9]. Artificial Intelligence (AI) improves the cost-effectiveness of electricity by adjusting to variations in weather-related electricity generation. The integration of supplementary renewable energy sources into the smart grid offers notable benefits. This method improves the resilience of the grid and makes it easier to identify equipment malfunctions and forecast electricity output and demand. In recent years, machine learning (ML) techniques have been increasingly applied to mitigate the adverse effects of electricity theft and related cyberattacks in smart grids (SGs). Both supervised and unsupervised ML approaches, including deep learning (DL) models, have shown promise in detecting theft patterns. However, these approaches come with several limitations. First, DL models are typically trained on static datasets, which can lead to overfitting, making them adept at recognizing specific patterns but less effective at generalizing to broader and evolving theft behaviors. Second, adapting these models to changing consumption patterns and emerging cyberattacks is inefficient, as it requires frequent retraining on both old and new data. This retraining process is not only time-consuming but also computationally intensive, particularly when dealing with large datasets typical in smart grid environments [10].Electricity theft in smart grids remains a major challenge, causing significant financial losses and threatening grid stability. Traditional detection methods, often reliant on **labeled data**, are limited by the difficulty and cost of obtaining confirmed theft cases, which usually require physical inspections or audits. Moreover, these approaches struggle to capture the **complex temporal patterns** inherent in electricity consumption data, leading to inaccurate detection of subtle or evolving theft behaviors. As smart grids generate vast amounts of data, there is an urgent need for a **deep learning-based solution** that can adaptively leverage **unlabeled data** to automatically learn and identify abnormal consumption patterns. The problem lies in the need for a **deep, adaptive learning framework** capable of detecting electricity theft in real-time, improving detection accuracy, minimizing false positives, and scaling efficiently within smart grid environments. This requires moving beyond traditional methods to develop a more sophisticated, data-driven approach based on **deep learning architectures [11] [12]**.

1.1 Motivation and contribution

Electricity theft is a growing concern in smart grids, leading to substantial financial losses and grid instability. Traditional methods, often hardware-based or reliant on labeled data, struggle with scalability, high costs, and the complexity of evolving theft patterns. With utilities losing billions annually and smart grids generating vast amounts of real-time data, there is a pressing need for advanced, adaptive detection models. This motivates the development of a deep learning framework capable of leveraging both labeled and unlabeled data to accurately detect and mitigate electricity theft in real-time, improving efficiency and reducing false positives. This research introduces a novel deep learning framework, **Deep Adaptive Feature Learning for Theft Detection (DAFL-TD)**, aimed at improving electricity theft detection in smart grids. The key contributions are:

1. **Novel Framework**: We propose DAFL-TD, which integrates labeled and unlabeled data, enabling effective detection of electricity theft while adapting to changing consumption patterns.
2. **Hybrid Model**: The framework combines a **Temporal Feature Extraction Network (TFEN)** with an **Adaptive Feature Learning Network (AFLN)** to enhance feature representation and classification accuracy.
3. **Data Augmentation and Robustness**: Our model employs advanced data augmentation techniques to handle noise, imbalanced datasets, and fluctuating time-series data, ensuring robustness in real-world environments.
4. **Performance Improvement**: Experimental results on the SGCC dataset show that our model outperforms state-of-the-art methods, achieving higher accuracy and reduced false positives in detecting electricity theft

2 Related Work

In [11], the authors presented a methodology for predicting electricity theft utilizing data obtained from smart meters that monitor energy consumption. This technology enables energy supply companies to effectively address challenges related to inadequate electricity management, unexpected electricity consumption, and energy shortages. Convolutional Neural Networks (CNNs) were developed by scientists. The DL method maintains the critical characteristics of electricity consumption data by initially distinguishing between periodic energy, utilizing established methodologies. The findings indicate that the deep convolutional neural network (CNN) model surpasses earlier models, achieving the highest detection accuracy for energy theft. The results indicate that anomalous immobility behavior can be detected, and that an adaptive premises system is capable of consistently identifying it over an extended duration. This study presents a cost-effective, data-driven ETD approach that maintains ETD accuracy while substantially decreasing data labeling expenses. A deep active learning (DAL)[12] system designed for intellectuals is utilized to implement the process with precision. The DAL approach efficiently selects the optimal samples for the ETD model. The effectiveness of the proposed method is demonstrated through the experimental results derived from a real ETD dataset provided by the State Grid Corporation of China. Reference [13] details an investigation conducted by the authors into instances of electricity theft within the distributed generation (DG) sector. By conducting a thorough examination of distributed generating units that utilize renewable energy sources, certain consumers exploit smart meters for this violation to generate a misleading perception of heightened electricity consumption, resulting in overpayments to the utility provider. Techniques for identifying risky conduct are examined through the application of deep machine learning methodologies. The paper [14] presents a deep reinforcement learning (DRL) technique aimed at addressing the issue of electricity theft, utilizing samples derived from real-world datasets. A number of additional cases utilize the proposed methodology. A global detection model is constructed using a double deep Q network (DDQN) and a deep Q network (DQN), employing various deep neural network topologies. The global detector alters the consumption patterns of current customers and increases the complexity of security protocols in response to newly introduced threats. The results indicate that the proposed DRL method is capable of effectively identifying new consumption patterns. In the referenced work [15], the author employs convolutional neural networks (CNN) and long short-term memory (LSTM) architectures to extract abstract features from electricity usage data. The prototype for each class is generated by calculating the parameters of the abstract feature, which is subsequently utilized to predict the labels of unknown data.

[16] presents an examination of the effects of backdoor assaults in ETD for the first time, along with a proposed feature attention distillation defensive method. To enable adversaries to bypass ETD, it is essential to conduct a thorough analysis of the attack surface during the current model training process. Malicious backdoors can be integrated for specific triggers. The evaluation of six widely used ANN-based models is subsequently conducted. Research indicates that attackers can successfully bypass the back doored ETD models in over 90.53% of instances, resulting in substantial losses for energy suppliers.transfer learning-driven approach was introduced in [17], which aims to boost detection accuracy in cases with limited samples. This method transfers a model trained in a data-rich location (source domain) to another with fewer samples (target domain), addressing the issue of data scarcity in ETD. In [18], the authors proposed a novel method combining Omni-Scale CNN (OS-CNN) with AutoGB to tackle challenges in time-series data and class imbalance. They employed the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) for data interpolation and SMOTEENN for data resampling, ensuring effective coverage of diverse time series scales. A cost-effective data-driven approach is presented in [19], utilizing a deep active learning (DAL) scheme that reduces data labeling costs while maintaining detection accuracy. By integrating CNN learning with Monte Carlo dropout-based Bayesian active query, this approach efficiently selects valuable instances for model training.In [20], the correlation between water and electricity (W&E) usage is analyzed as a basis for a new ETD method. By using the mutual information coefficient (MIC) to model W&E usage correlations and applying a wavelet clustering algorithm, the authors propose a multisource ETD method that clusters power distribution users based on their MIC values.

Despite advancements in machine learning for electricity theft detection, existing models struggle with overfitting to static datasets and lack adaptability to evolving consumption patterns. Current approaches often require costly and time-consuming retraining to address new theft behaviors and cyberattacks. There is a clear need for a more flexible, scalable solution that can generalize across dynamic patterns and efficiently handle large-scale smart grid data without frequent retraining.

3 Proposed Methodology

The proposed study aims at developing a model that focuses on using information including labelled as well as unlabeled data for distinguishing between pattern that are normal as well as electricity theft while considering electricity load information samples. The study proposes a model termed as Deep Adaptive Feature Learning for Theft Detection (DAFL-TD). The characteristics of this DAFL-TD model include effectively using of data for learning the supervised representing in prior tasks and further transfer to task downstream through tuning of model metrics. For further enhancement considering classification that leads to the usage of information, the proposed model uses anTFEN, aAFLN model for training as well as a classification unit, which is described in the figure 2 given below.

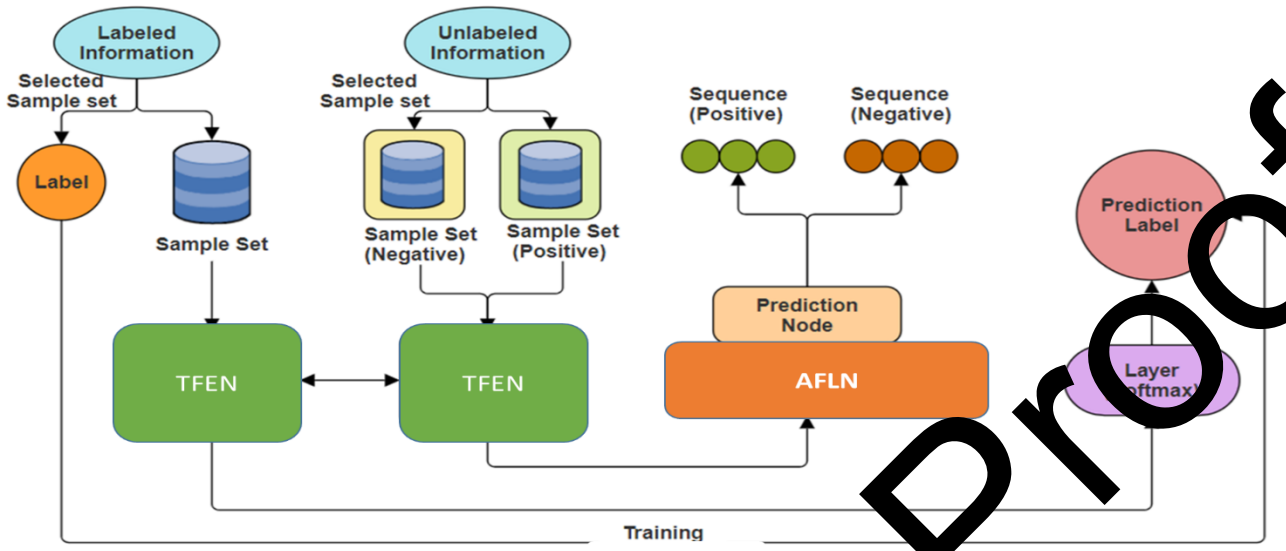


Figure 2 Adaptive Feature Learning for Theft Detection (DAFL-TD)

3.1 Pre-processing Module

The input information used by the model is matched using the inputs $(z_k, a_k) belongs\ to\ F_N$ that are marked and the inputs $(z_k)\ belongs\ to\ F_W$ that are unmarked. Here, $F_W = z_{k=1}^{V}$ is an energy usage sequence having V, z_k, a_k as length expressing the attribute labels of the samples. Z^{augmen_1} denotes the time sequence one and Z^{augmen_2} is used to represent the other time sequence both have distinct improvised intensities that are developed using data augmentation. The possible expressions that are retrieved using the TFEN$h_\emptyset$, the attributes of $h_\emptyset$ are updated by predicting the using the AFLN loss function. Considering the supervised classification method, the TFEN along with the weights that are prior trained is refined via the labelled information and is lastly classified using a conversion node. The classification loss function is formulated as given below

$$N_{classification} = -|F_N|^{-1} \sum_{K=1}^{|F_N|} a_k \cdot \log(r_k) \tag{1}$$

In this case, loss function for cross entropy is denoted as $N_{classifcation}$, the main aim of this is the optimization and reduction of divergence MN.

The augmentation of information acts as an essential unit in the task of AFLNlearning, that acknowledges the lack of information as well as enhances the diversity of information. The proposed method utilizes various augmentation methods that developed different perturbations to the source information as well as generation of datasets from various outlooks. This method helps in AFLN learning to increase consistency of different perspectives as well as learning of invariant attribute expressions. It also uses methods of augmentation including scaling, negating, time period shift and permutations.

The information is improvised using noise by addition of random noises to the sequential information leading to a Gaussian distribution. This aids the proposed model in generalization of noisy information better in real-time. The conversions on the scale are utilized for enhancement of information by scale adjustment of the information. The rearrangement of information is termed as permutation used to rearrange the structure of data that aids the model is learning various techniques of arranging information. Positive information sets are converted to negative samples that aids the proposed model to better identify the various classes. Time period shifting is used to improve the information by transfer of information within the time sequence that helps the model to gather the data fluctuating patterns at different instances of time.

3.2 Temporal Feature Extraction Network (TFEN)

Considering a group a electricity sequential information that is labeled denoted as $(z_k^N, a_k^N) belongs\ to\ F_N$, the proposed model TFEN$h_\emptyset$ is utilized for training for retrieval of local attributes. The TFEN a network of stacked attributes that has four layers. Every layer has a 1D layer of convolution, layer for batch normalization as well as an activation layer (ReLU) and lastly a pooling layer. The concluding layer is linked to a softmax for basic tasks of classification. This method is comparison to a completely linked layer has increased efficiency while considering parameters and the capabilities of retrieving attributes of higher levels. Assume we have electricity datasets $(Z_v^1, A^1), (Z_v^2, A^2), \ldots., (Z_v^p, A^p) belongs\ to\ Sequence v^p$, having $Z_v^1 = (z_1, z_2, \ldots, z_v) belongs\ to\ \mathbb{T}^{1\times v}$ denotes a set of electricity load having length and $A^p\ belongs\ to\ [1, O]$ represents the relating sequential time label having bias count of labels for various classifications. The data attributes are attained using the equation $b_k^e = b(Z_v^k)$. During the process of AFLN learning, the initial input Z_v^k undergoes the process of augmentation to produce two information samples having various perturbations, then input into $h_\emptyset$ formulated as:

$$h_{\emptyset_k} = Pooling(ReLU(BatchNor(Y \oplus z + d))) \quad (2)$$

Here, $BatchNor$ is used to denote the layer of Batch Normalization, the parameters Y and d represent the attributes of the convolutional model, the $Pooling$ represents $MaxPooling$ layer that utilizes highest strategy, the activation function $ReLU$ is situated between the $Pooling$ and$BatchNor$ layer.

3.3 Adaptive Feature Learning Network (AFLN)

The coding unit of the AFLN training uses Deep Learning Adaptive Sequential Feature Network (ASFN) and designs the attention scheme that is termed as Adaptive Sequential Feature Network (ASFN) in the proposed study. It is observed to have increased benefits for feature retrieval as compared to traditional prior networks, that also have anTFEN, attention scheme as well as classification model. The ASFN is made up of five built up unidirectional Gated Recurrent Units. This is a unique neural network that is recurrent having parameters of smaller sizes having a usage that is easier in comparison to Long Short-Term Memory. Consider v as a time period step, the input information for the TFEN of ASFN Units z belongs to $\mathbb{T}^{P\times N}$, where the vector size is denoted as P and the sequential time step has a length of N. The sequential inputs are of various lengths having particular batches for the training procedure, the

ASFN will not omit them but will describe a sequential input using batch z_k belongs to $\mathbb{T}^{D\times P\times \tilde{N}}$, here the measure of the sequence that is longest is expressed as $\tilde{N}$. Here, $\tilde{N}$is patched with 0 while ending.

In this case, we assume z_v as the input data, the concealed vector of the prior time step denoted j_{v-1} is also utilized as the input. The ASFN is evaluated using the concealed vector j_v for time period v. This is formulated as given below:

$$t_v = \delta\left((d_z^t + z_v Y_z^t) + \left(d_j^t + j_{v-1} Y_j^t\right)\right) \quad (3)$$

$$w_v = \delta\left((d_z^w + z_v Y_z^w) + \left(d_j^w + j_{v-1} Y_j^w\right)\right) \quad (4)$$

$$e_v = \tanh\left((d_z^e + z_v Y_z^e) + t_v\left(d_j^e + j_{v-1} Y_j^e\right)\right) \quad (5)$$

$$j_v = w_v \circ j_{v-1} + (1 - w_v) \circ e_v \quad (6)$$

Considering the above equations 3,4,5,6, the activation function δ is sigmoid given as $\delta(z) = (1 + e^{-z})^{-1}$, the gates denoting update, reset as well as candidate is given as w_v, t_v and e_v, respectively. The weight vectors that can be trained are given as Y_r^S and the bias vectors are given as d_r^S. Consider a sequence $z\ belongs\ to\ \mathbb{T}^{P\times N}$, the equations 3 to 6 are utilized by the TFEN to result in the conceal layer matrix $\tilde{j}\ belongs\ to\ \mathbb{T}^{128\times N}$. The attention scheme of ASFN uses compressed expression of sequential electric loads that are a result of the previous Adaptive Sequential Feature Network (ASFN) layer denoted as $\tilde{j}$, furthermore it aims on global parameters. The attention scheme utilizes parameters that are trainable represented as $\bar{J}$ for evaluation of the attention vector $e\ belongs\ to\ \mathbb{T}^{128}$ which is formulated in the equation (7) given

$$e = \bar{J}\left(E\left(\sum_{v=0}^{N-1} E j_{N-1}^V Y_e j_v\right)^{-1}\right) \quad (7)$$

The above given equation (7) uses a concealed layer for the TFEN network for computation of the attention vector as well as the final time period step of the TFEN concealed layer j_{N-1} that encapsulates data from the layers $N-1$ that follows. After obtaining the attention vector e, a new contextual attribute vector is developed by combining e and j_{N-1} that is used for tasks of classification. This method focuses on joining the last concealed layer of the encoded layer, hence improvising the diversity of information. The benefits of the Adaptive Sequential Feature Network included its acceptance of different length of inputs. The quantity of data that is grasped by j_{N-1} could differ depending on the

length of the sequences, exposing the model susceptible to different sequential lengths. To resolve this, the proposed model uses a set of parameters that are shared at the process of training, this gives context to the concealed state of the TFEN. The usage of the concealed state is performed directly or in a combined manner with the contextual state is determined by these parameters. On computation of contextual vector e, the evaluation of e' denoting auxiliary contextual vector is performed along with the evaluation of the attention scheme output:

$$e' = \mathbb{H}_{\mathcal{T}}(e, j_{N-1}) \quad (8)$$

$$Output_{\mathcal{T}} = [e; e']$$

Here, the Adaptive Sequential Feature Network (ASFN) is expressed as $\mathbb{H}_{attention}$ that utilizes contextual vector e and concealed state of prior j_{N-1} as input. Two completely linked layers are used for classification as well as activation layer of ReLU is used for the output layer. Lastly a softmax layer is utilized for probability distribution given as: $\hat{a} = Softmax(H_2(ReLU(H_1(Output_{attention}))))$.

The benefits of using the Adaptive Sequential Feature Network (ASFN), it depends only on the last time period step that decreases the complexity of computation. The effects that are caused by the difference in length of the sequence be reduced by evaluation of contextual vectors. The attention scheme aids in keeping the zero-sequence unaltered, hence small trainings quantitatively can be performed on smaller sequences having various lengths.

The AFLN training unit is an essential part of the proposed model Deep Adaptive Feature Learning for Theft Detection (DAFL-TD) The electricity load that is unlabelled denoted as $z_k^W \; belongs \; to \; F_W$ as a prior task used in supervised learning for AFLN learning. While considering the process of augmentation, the data augmentation has two various types of improvisations for inputs Z_v^k to result in $Z_v^{k1}$and Z_v^{k2}. Considering various types of augmentation, the proposed model can concatenate them to result is adequate negative as well as positive pairs of information samples. The TFEN considers these segments in relation to time sequence as inputs and therefore retrieves attribute expressions given as $b_v^{k+} = h\emptyset(Z_v^k)$and $b_v^{k-} = h\emptyset(Z_v^k)$. Further using the inputs b_v^{k+} and b_v^{k-} in the AFLN training unit.

We consider the initial length of the sample as N_u and the index of segmentation as K_u. $b_v^{k+}$and$b_v^{k-}$with the index K_u is divided as the historical_seqb_v^{k+}, the predicted_sequenceb_v^{k-} for index time interval $[K_u, K_u + T_u]$ that is implemented for the future of AFLN prediction. Once the sequential sets are generated from the samples that are both negative as well as positive, the historical_seq is first stored into the Adaptive Sequential Feature Network stack along with the attention scheme. The Adaptive Sequential Feature Network (ASFN) retrieves the possible attributes e_r^{k+} denoting the $future\_sequence$ which is then used to generate the $future_sequence\, S^{k+}$having length N_u viaa sequential string of non-linear conversion elements.

On using Noise Contrastive loss function, we estimate the mutual data of sequences that have been predicted such as S^{k+} and S^{k-}. The complete procedure uses the network via this function and the parameters of the network are also updated for further learning.

The figure 3 given below shows e_r^{k+} used for predicting the sequential attribute for time ranging from $K_u$to $(K_u + N_u)$ post eh indexing K_u. We assume that the prediction is expressed as $(S_r^{k+}, S_{r+1}^{k+}, \ldots, S_{r+N_u}^{k+})$, insider the similar range of index $[K_u, K_u + N_u]$,S_r^{k+} $belongs\ to\ b_v^-$.The network parameters estimated using the mutual data of S_r^{k+} and S_r^{k-}as

$$h_m\left(e_r^{k+}, S_{r+m}^{k+}\right) = exponent((Y_m(e_r^{k+}))^V S_{r+m}^{k+}) \quad (9)$$

Here, the contextual vector is represented as e_r^{k+} that is gathered from the attribute vector via the Adaptive Sequential Feature Network (ASFN). The linear attribute Y_m is used by mapping e_r^{k+} to the similar size as $S_{r+m}^{k+}, m\ belongs\ to\ [1, N_u]$. Also, the negative data sequential feature b_v^- retrieved from a completely linked layer is encoded to result in S_r^{k-} and S_r^{k+} $belongs\ to\ b_v^+$ for mutual data AFLN prediction.

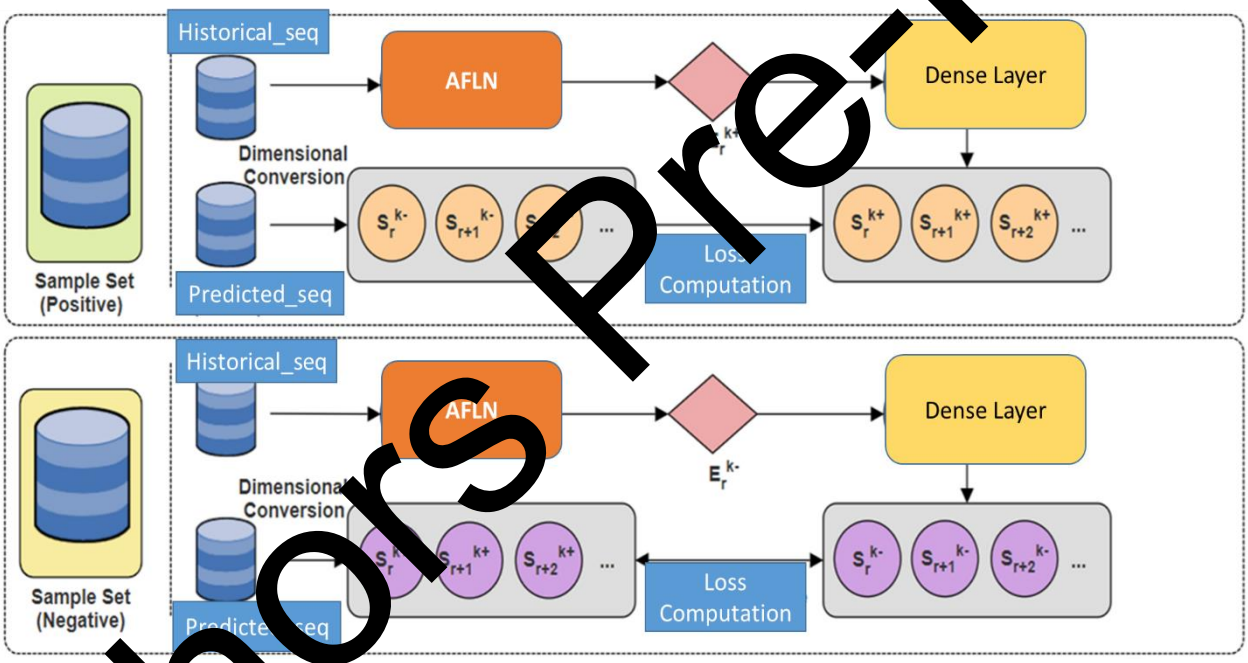


Figure 3Adaptive Feature Learning Network (AFLN)

Therefore, there are two loss functions that have to be essentially evaluated

$$N_{VE}^{+} = -(M)^{-1} \quad (10)$$

$$\sum_{m=1}^{M} log\left(\left(\mathbb{E}((Y_m(e_r^{k+}))^V S_{r+m}^{k+})\right)\left(\sum\nolimits_{p\ belongs\ to\ P_{r,r+N_u}} \mathbb{E}((Y_m(e_r^{k+}))^V S_{r+m}^{k+})\right)^{-1}\right)$$

$$N_{VE}^{-} = -(M)^{-1} \tag{11}$$

$$\sum_{m=1}^{M} log\left(\left(\mathbb{E}((Y_m(e_r^{k-}))^V S_{r+m}^{k-})\right)\left(\sum_{p\ belongs\ to\ P_{r,r+N_u}} \mathbb{E}((Y_m(e_r^{k-}))^V S_{r+m}^{k-})\right)^{-1}\right)$$

3.4 Adaptive Loss Training Module

Here, the AFLN loss for prediction is given as $N_{VE}^{+}$and N_{VE}^{-} that is utilized to increase the dot product of predicted as well as real expressions of the same sets of samples while reducing the dot product for predicted expressions as well as for other data sets within the batch. A context AFLN loss is introduced for increasing the similarity for samples that are positive and reduced the similarity for samples that are negative. e_r^{k+}ande_r^{k-} are both introduced by the Adaptive Sequential Feature Network, the enter a conversion block network sequentially having weights that are shared, this results in a similarity function given as: $similarity(w,x) = \left((w^v x)(\|w\|\|x\|)^{-1}\right)$, where w,x denotes the two vectors $(\|w\|\|x\|)$. The concluding formulation for the loss function is as given below

$$N_{ee} = -\sum_{k=1}^{P} log\left(\left(\mathbb{E}(similarity(h_r^{k+}, h_r^{k-})\right.\right.$$
$$\left.\left./\varphi)\right)\left(\sum_{o=1}^{2P} \beth_{[0\neq k]}\, \mathbb{E}(similarity(h_r^{k}, h_r^{o})/\varphi\right)^{-1}\right) \tag{12}$$

For the equation 12 given above N_{ee} is used to denote the loss function for single pairs, the function used for indication is represented by $\beth$, while $o\ is\ not\ equatl\ to\ k$. There is a coefficient used for temperature that is defined by φ, this is mainly used to increase the output of softmax. Hence, the negative log of softmax is used for loss function. The total loss is made up of the AFLN loss function in relation to time as well as contextual which is given as :

$$N_{loss} = \gamma_1 \cdot (N_{VE}^{+} + N_{VE}^{-}) + \gamma_2 \cdot N_{ee} \tag{13}$$

Where, γ_1 and γ_2 are scalar constant hyper parameters used to show the relating weights for various losses at every time period. The combination of the AFLN loss relating to time as well as the AFLN loss contextually, has increased significant attributes that are distinct for positive as well as negative samples for learning, hence the attribute TFEN Completely linked layer and the parameters of the neural network for AFLN learning of the ASFN is updated. This is described in detail in the algorithm 1 given below

Algorithm 1 **Detection of Electricity theft Electricity patterns using AFLN Learning**

Step 1 **Input:** Batch dimension P,structure V, TFEN attribute function h, constant φ,TFENfunction$i$ (autoregressive)

Step 2 **Output:** Optimal ideal or approximately ideal state of neural network

Step 3 For dataset sampled mini batch $\{z_m\}_{m=1}^{P}, \{a_m\}_{m=1}^{P}$ do

Step 4 For all $m\ belongs\ to\ [1, \dots, P]$do

Step 5 If $TrainingPhase$ is self-supervised learning do

Step 6 Set the information augmentation function $v \sim V, v' \sim V$

Step 7 Generation of Positive data samples $z_m^+ \leftarrow v(z_m)$

Step 8 Generation of Negative data samples $z_m^- \leftarrow v(z_m)$

Step 9 Attribute feature learning for positive data samples $e_m^+ \leftarrow h(z_m^+)$

Step 10 Attribute feature learning for negative data samples $e_m^- \leftarrow h(z_m^-)$

Step 11 Positive data sample feature learning (auto regressive)$h_m^+ \leftarrow i(e_m^+)$

Step 12 Negative data sample feature learning (auto regressive)$h_m^- \leftarrow i(e_m^-)$

Step 13 Initializing index score and length of sample sequence (r, N_u)

Step 14 Set equation (10)

Step 15 Set equation (11)

Step 16 Set equation (12)

Step 17 Compute $Loss$ using equation (13)

Step 18 Optimization parameters of neural network as well as TFEN

Step 19 Else

Step 20 Reduce the weighting information of TFEN

Step 21 $$Predictions_m, a_m \leftarrow h(z_m)$$

Step 22 Computation of $N_{Loss} = CrossentropyLoss(Predictions_m, a_m)$

Step 23 Reduce $N_{Loss}$and TFEN optimization

Step 24 End If

Step 25 End For

Step 26 End For

4 Performance Evaluation

In the performance analysis of electricity theft detection, multiple models are compared using the State Grid Corporation of China (SGCC) dataset. Evaluation metrics such as Area Under the Curve (AUC) and Mean Average Precision (MAP) were used to assess the performance of models like CNN, SVM, and advanced hybrid models like GCN-CNN, DAFL-TD, and LSTM-RUSBoost the existing system is compared with the proposed model and the results are evaluated in the form of graphs and table.

4.1 Dataset Details

The dataset in this research [21] the State Grid Corporation of China (SGCC) and includes electricity consumption data from 42,372 customers over a period of 1,035 days, spanning from January 1, 2014, to October 31, 2016. Among these customers, 38,757 are classified as normal consumers, while 3,615 are identified as electricity thieves. The dataset captures daily electricity usage patterns, allowing for the identification of abnormalities associated with electricity theft, where customers manipulate their consumption to reduce recorded usage. To improve the dataset's usability for analysis, preprocessing steps are performed to handle missing values and outliers, ensuring data quality. This preprocessing includes interpolation for missing values and outlier mitigation using the Three-sigma rule. The dataset is also normalized using Min-Max scaling to ensure consistency for machine learning model training. This comprehensive dataset is critical in training models aimed at detecting electricity theft by identifying irregular consumption patterns.

4.2 Evaluation metrics

The SGCC dataset served as the sole data source utilized in the trials conducted for this investigation. The model achieves an accuracy rate of 91.4% by classifying all users as normal, despite the presence of data imbalance. The primary reason for this is that the actual dataset comprises a significantly higher number of average user samples compared to instances of electricity thieves. Consequently, it would be overly simplistic to evaluate the model's quality solely based on its accuracy. In unbalanced classification tasks, model performance is frequently assessed using metrics such as mean average precision (MAP) and area under the curve (AUC). This enables the evaluation of the model's effectiveness in a manner that aligns more closely with established scientific principles. AUC serves as a critical evaluation metric for classification tasks. The AUC value represents the likelihood that a randomly chosen stolen sample will have a higher ranking than a randomly chosen normal sample. The formula for calculating AUC is given as follows as shown below:

$$\mathrm{AUC} = \frac{\sum_{k \in \mathrm{theftclass}} \mathrm{Rank}_k - \frac{M(1+M)}{2}}{M * N} \qquad (14)$$

M denotes the total number of larceny samples, N denotes the total number of normal samples, and the rank value for each sample is indicated by rank. MAP is commonly utilized to evaluate the effectiveness of information retrieval. The system is designed to conduct a comprehensive evaluation of the model's ability to detect rare events in imbalanced datasets. Before the assessment procedure utilizing MAP, the labels of the test set are organized according to the prediction score. The selection of the top N labels is based on performance evaluation. The definition of accuracy is initially presented at n, denoted by P@n.

$$P@n = \frac{Y_n}{n} \quad (15)$$

Ynrepresents the number of correctly detected cases of electricity theft that occurred before location n. Next, we define MAP@N as the average of all labels, taking into account just the first N labels. P@nscenarios. Its value is determined using the following formula:

$$MAP@P = \frac{\sum_{k=1}^{t} R@p_k}{t} \quad (16)$$

where r represents the quantity of individuals engaged in electricity theft within the leading N categories, and$p_k(k = 1 \dots \dots . t)$indicates the corresponding ranking of each instance of electricity theft. The @ symbol in $MAP@P$ The evaluation metric focuses on precision in identifying the top N most likely electricity theft cases, meaning that the metric assesses how accurately the model pinpoints the most probable instances of electricity theft from a ranked list. By concentrating on precision for the top N cases, the metric ensures that the model is particularly effective at identifying the highest-priority theft cases, minimizing false positives, and improving the efficiency of theft detection efforts. This targeted precision can help utility companies allocate resources more effectively to investigate and prevent theft., with $R@p_k$Presenting the accuracy at each position in the ranking.

4.3 State-of-art methods

This paper evaluates the performance of the proposed method with various state-of-art techniques, including CNN-LSTM, CNN-RF, LSTM-RUSBoost, self-attention, and GCN-CNN, in addition to several classical techniques including SVM, OPF, MLP, and CNN.

• SVM [22]: The noteworthy capability of SVM is attributed to the application of nonlinear separating hypersurfaces. The detection of electricity theft by this method has been thoroughly verified.

• OPF [23]: The entire graph is divided into optimal path trees in order to address the classification problem. Every user in the OPF (Optimum Path Forest) framework is considered a distinct node. The model uses the ideal route tree to classify these nodes.
MLP [24]: A multilayer perceptron (MLP) is a kind of feedforward neural network that consists of an output layer, several hidden layers, and a superficial input layer. The MLP integrates several linear layers and activation functions to provide classification results. The MLP is specifically designed to evaluate input data that is arranged into 1035 columns in a single row.
• CNN [25]: Convolutional neural networks (CNNs) are a subclass of artificial neural networks (ANNs) that perform convolutions in at least one of their hidden layers as opposed to utilizing standard matrix multiplication. WDCNN eliminates the CNN element while leaving all other settings same.
WDCNN [25]: This method accurately detects theft from both a depth and a breadth perspective by combining convolutional neural networks (CNN) with fully connected layers.
• CNN-LSTM [26]: This method combines an architecture for long short-term memory (LSTM) with a convolutional neural network (CNN). The construction has seven hidden layers. Each of the first four layers has twenty feature maps that are employed in convolution operations. The remaining layers use 10, 5, and 100 neurons, respectively, to perform the LSTM operations.

• LSTM-RUSBoost [28]: This method blends LSTM and RUSBoost. For feature refinement, the LSTM is utilized, and for data balancing, the RUSBoost method. The RUSBoost method performs better when parameter optimization is done using the bat algorithm.
• CNN-RF [27]: This model was developed by fusing the CNN and RF classifiers. Before submitting the 40 data points to the RF model for classification, the CNN examines them to produce new feature vectors.
• Self-attention [29]: This model includes a multi-head self-attention mechanism connected to dilated convolution. Significant performance benefits are obtained by creating a binary channel and employing a 1 x 1 convolutional kernel to locate missing data.
• GCN-CNN [30]: This method uses the K-Nearest Neighbors (KNN) methodology to statically generate the adjacency matrix by combining spectrum-based GCN with CNN.

4.4 Results

Table 1 performance evaluations of various model

Methods	**Training ratio 50% AUC**	**Training ratio 50% MAP@100**	**Training ratio 50% MAP@200**	**Training ratio 60% AUC**	**Training ratio 60% MAP@100**	**Training ratio 60% MAP@200**	**Training ratio 70% AUC**	**Training ratio 70% MAP@100**	**Training ratio 70% MAP@200**
SVM	0.718	0.686	0.597	0.731	0.719	0.607	0.727	0.724	0.607
OPF	0.737	0.701	0.681	0.753	0.723	0.711	0.747	0.713	0.711
MLP	0.743	0.919	0.888	0.757	0.909	0.873	0.754	0.923	0.877
CNN	0.773	0.82	0.842	0.771	0.839	0.843	0.781	0.875	0.924
WDCNN	0.776	0.94	0.896	0.792	0.955	0.929	0.786	0.968	0.932
CNN-LSTM	0.801	0.798	0.821	0.808	0.812	0.812	0.807	0.81	0.810
LSTM-RUSBoost	0.861	0.80	0.882	0.803	0.803	0.818	0.793	0.793	0.793
CNN-RF	0.868	0.869	0.872	0.878	0.878	0.870	0.822	0.879	0.876
Self-attention	0.868	0.881	0.824	0.888	0.888	0.871	0.826	0.892	0.872
GCN-CNN	0.891	0.909	0.776	0.962	0.931	0.787	0.787	0.981	0.954
DARGNN [ES] [31]	0.844	0.902	0.859	0.964	0.939	0.849	0.849	0.963	0.926
DAFL-TD	0.867	0.923	0.879	0.987	0.967	0.867	0.867	0.987	0.968

4.4.1 Training ratio @ 50%

The comparison of AUC improvements across the methods highlights key performance enhancements as shown in figure 4. The largest increase is seen between CNN-RF and DGRGNN [ES], showing a significant boost in classification accuracy due to the use of graph-based techniques. Other notable improvements occur between MLP and CNN, as well as Self-attention and DAFL-TD, both indicating substantial gains in performance with these advanced models. Moderate improvements are seen when transitioning from GCN-CNN to CNN-LSTM and from DGRGNN [ES] to LSTM-RUSBoost, reflecting the benefits of incorporating temporal processing and better handling of imbalanced data. Smaller increases, such as those between CNN and WDCNN or WDCNN and GCN-CNN, suggest only minimal enhancements. Overall, the comparison shows a consistent rise in performance, with the most advanced hybrid models offering the greatest gains in AUC.

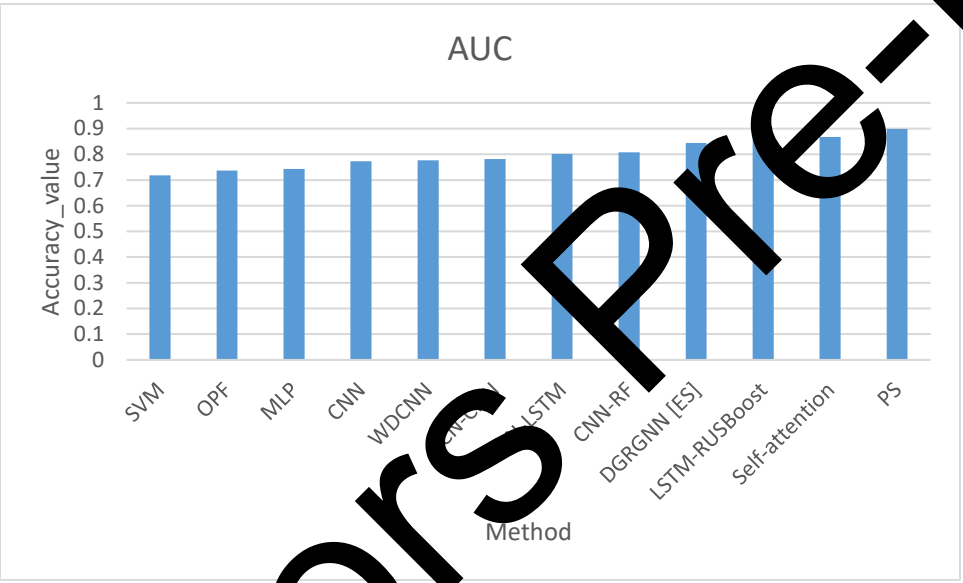


Figure 4 AUC for training ratio @50%

The comparison of MAP@100 and MAP@200 values across methods shows varying levels of improvement at training ratio @ 50 % as shown in figure 5. Notably, WDCNN achieves the highest MAP@100 (0.94), indicating strong performance, while MLP and CNN also exhibit high MAP scores at both cutoffs, reflecting their solid classification abilities. CNN-LSTM and LSTM-RUSBoost show stronger performance at MAP@200, with LSTM-RUSBoost improving significantly from MAP@100 to MAP@200. Conversely, methods like GCN-CNN perform well at MAP@100 but experience a decline at MAP@200, suggesting potential overfitting. DGRGNN [ES] and DAFL-TD show consistent and robust performance across both MAP scores, with DAFL-TD slightly outperforming DGRGNN at MAP@100. Self-attention shows solid performance at MAP@100 but decreases at MAP@200, while CNN-RF provides competitive performance, though data for MAP@200 is unavailable. Overall, WDCNN, MLP, and CNN perform consistently well, while methods like LSTM-RUSBoost and GCN-CNN show more variability depending on the evaluation metric.

Figure 5 MAP @100 and 200 comparison of existing with proposed for training ratio @50%

4.4.2 Training ratio 60% AUC

The AUC comparison at a 60% training ratio highlights significant performance improvements across the methods as shown in figure 6. DAFL-TD achieves the highest AUC at 0.987, showcasing its superior classification capabilities. GCN-CNN and DGRGNN [ES] follow closely with AUCs of 0.962 and 0.964, respectively, indicating the strong performance of graph-based models. CNN-RF and Self-attention also perform well, with AUCs of 0.878 and 0.888, demonstrating the effectiveness of hybrid models. CNN-LSTM shows a notable increase over traditional CNN architectures, achieving an AUC of 0.81. WDCNN improves over CNN, reaching 0.792, while LSTM-RUSBoost shows competitive results with an AUC of 0.803. SVM, OPF, and MLP show lower performance, with OPF slightly outperforming MLP and SVM. Overall, the analysis reveals a clear trend of higher AUC values for more complex and hybrid models, with DAFL-TD and graph-based methods leading the performance.

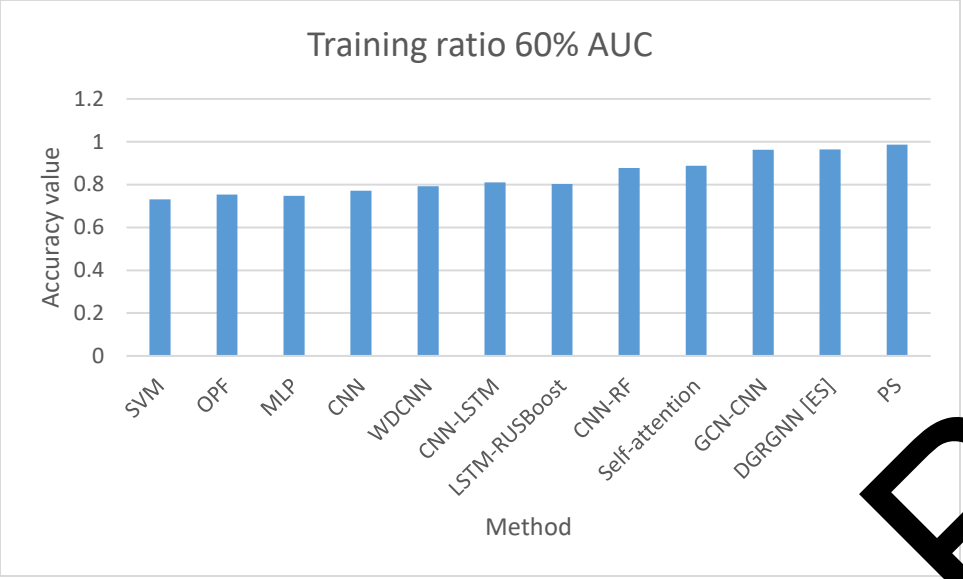


Figure 6 AUC for training ratio @50%

The MAP@100 and MAP@200 analysis at a 60% training ratio demonstrates notable improvements across models as shown in figure 7. DAFL-TD achieves the highest scores, with MAP@100 at 0.967 and MAP@200 at 0.978, indicating exceptional performance and consistency at both evaluation metrics. DGRGNN [ES] and GCN-CNN follow closely, with DGRGNN [ES] achieving 0.964 at MAP@100 and 0.939 at MAP@200, and GCN-CNN showing strong performance with 0.962 and 0.931, respectively. WDCNN and CNN continue to perform well, with WDCNN scoring 0.955 at MAP@100 and 0.929 at MAP@200, slightly outperforming CNN. Hybrid models like CNN-RF and Self-attention show competitive results, especially CNN-RF with 0.878 and 0.8[illegible]. Meanwhile, CNN-LSTM and LSTM-RUSBoost show more balanced but lower performance across both metrics, indicating moderate improvement. OPF and MLP display reasonable MAP scores, particularly OPF's steady improvement from MAP@100 to MAP@200. SVM shows the lowest values, suggesting limited effectiveness compared to more advanced models. Overall, DAFL-TD and graph-based models lead in terms of accuracy and consistency across both evaluation metrics.

Figure 7 MAP @100 and 200 comparison of existing with proposed for training ratio @70%

4.4.3 Training ratio @ 70%

The AUC analysis in figure 8 reveals a steady progression in model performance, with DAFL-TD achieving the highest AUC at 0.867, indicating its superior classification ability. DGRGNN [ES] follows with an AUC of 0.849, demonstrating the effectiveness of graph-based methods. GCN-CNN and Self-attention also perform well, with AUCs of 0.826 and 0.822, respectively, reflecting the strong potential of hybrid models. CNN-RF shows a significant improvement over simpler models, reaching 0.807. Traditional deep learning models like CNN, WDCNN, and CNN-LSTM exhibit moderate improvements, with AUCs ranging from 0.781 to 0.787, while LSTM-RUSBoost slightly outperforms them with an AUC of 0.793. MLP, OPF, and SVM represent lower AUCs, with SVM trailing at 0.727. Overall, the analysis shows consistent improvement in AUC as the models evolve in complexity, with DAFL-TD and DGRGNN [ES] leading the way.

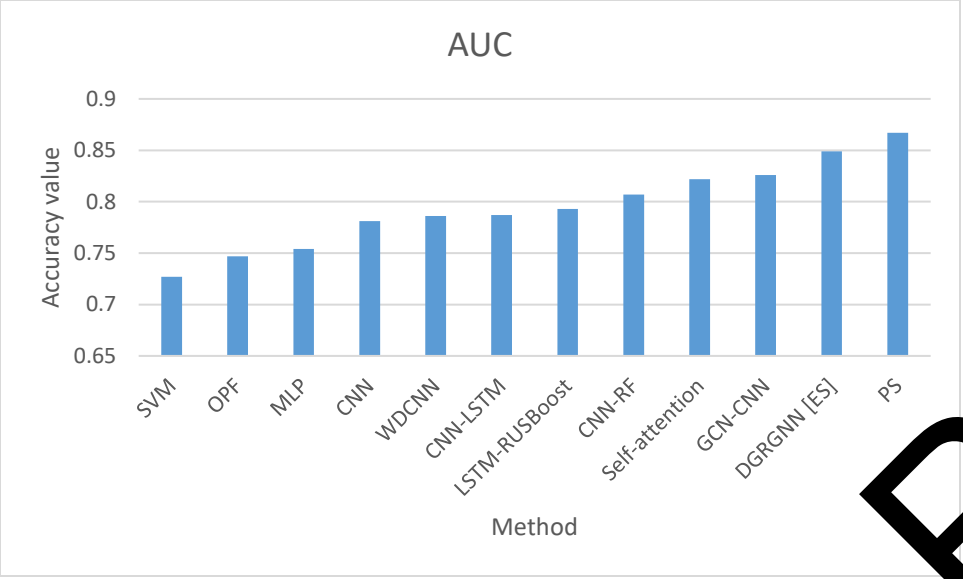


Figure 8 AUC for training ratio @70%

The MAP@100 and MAP@200 comparison at a 70% training ratio reveals significant performance improvements, particularly in advanced models as shown in Figure 9. DAFL-TD achieves the highest MAP scores, with 0.987 at MAP@100 and 0.956 at MAP@200, demonstrating exceptional accuracy and consistency. GCN-CNN follows closely with 0.981 at MAP@100 and 0.954 at MAP@200, reflecting the strong capability of graph-based methods. WDCNN and CNN also perform well, with WDCNN slightly outperforming CNN, especially at MAP@100 (0.968 vs. 0.955). While CNN-RF, Self-attention, and DGRGNN [ES] show solid performance, DGRGNN [ES] experiences a slight drop at MAP@200, indicating a small decline in consistency. Models like CNN-LSTM and LSTM-RUSBoost maintain lower MAP scores across both metrics, while MLP shows good performance, especially at MAP@100 (0.923). OPF and SVM, on the other hand, show the lowest results, with minimal improvements. Overall, DAFL-TD, GCN-CNN, and WDCNN dominate in performance, while more traditional models lag behind in comparison.

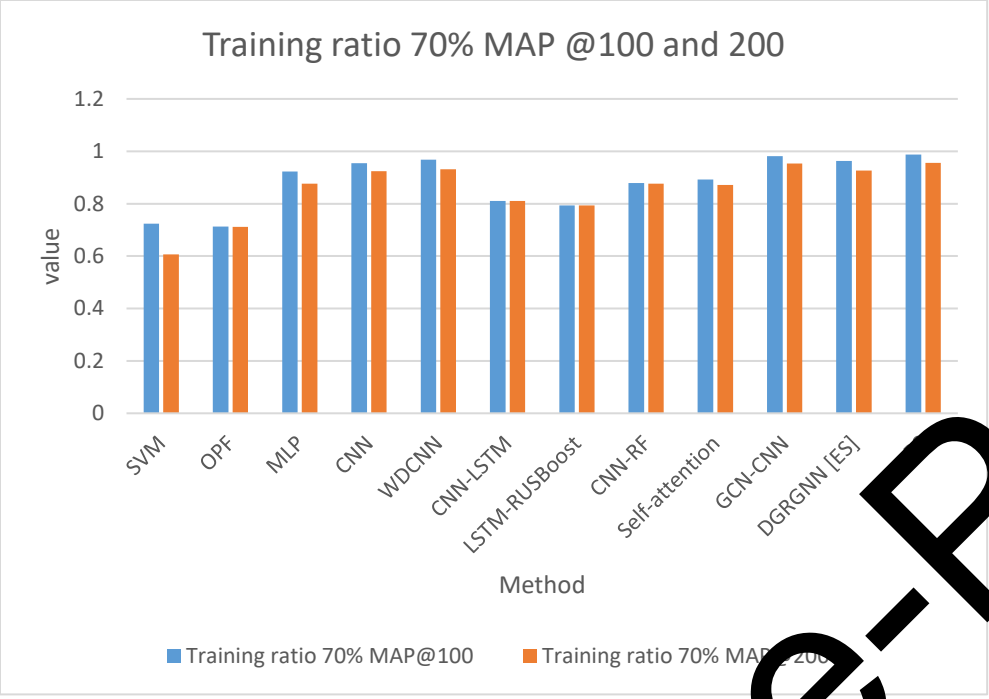


Figure 9 MAP @100 and 200 comparison of existing with proposed for training ratio @70%

4.5 Comparative Analysis

The performance of the proposed **Deep Adaptive Feature Learning for Theft Detection (DAFL-TD)** model was evaluated and compared with several state-of-the-art methods, including traditional machine learning models (SVM, OPF, MLP), deep learning-based models (CNN, CNN-LSTM, WDCNN), and hybrid models (Self-attention, GCN-CNN, CNN-RF, LSTM-RUSBoost). The evaluation metrics used were **Area Under the Curve (AUC)** and **Mean Average Precision (MAP)** at different training ratios (50%, 60%, and 70%) to ensure a comprehensive performance comparison. The proposed DAFL-TD model demonstrates significant improvements across all key metrics compared to the next best-performing models:

AUC: The **13.84% improvement** in AUC underscores the DAFL-TD model's ability to differentiate between theft and non-theft instances more effectively than other models. A higher AUC suggests a more reliable and accurate theft detection system that can better handle real-world complexities, including evolving theft behaviors.

MAP@100: While the improvement in **MAP@100** is marginal at **0.61%**, this still highlights DAFL-TD's precision in detecting the top 100 most suspicious cases. In practical terms, this means that the DAFL-TD model provides slightly more accurate theft detection rankings, which can be critical for resource allocation in theft investigations.

MAP@200: The **3.46% improvement** in **MAP@200** demonstrates that DAFL-TD not only excels at detecting the top 100 cases but also maintains high precision when identifying the top 200 cases. This shows that DAFL-TD's performance is consistent and scalable across larger datasets, which is crucial for electricity theft detection at scale.

Conclusion

In this study, we introduced a novel deep learning framework, Deep Adaptive Feature Learning for Theft Detection (DAFL-TD), aimed at improving electricity theft detection in smart grids. The proposed framework leverages both labeled and unlabeled data, utilizing a combination of Temporal Feature Extraction Network (TFEN) and Adaptive Feature Learning Network (AFLN) to enhance feature representation and classification accuracy. By incorporating advanced data augmentation techniques and robust feature extraction methods, the model effectively handles imbalanced datasets and fluctuating time-series data. Experimental results demonstrated that the DAFL-TD model outperforms state-of-the-art methods in terms of accuracy and reducing false positives, as validated on the SGCC dataset. This research provides a scalable, efficient, and adaptive solution for real-time electricity theft detection, offering utilities a valuable tool to mitigate financial losses and improve grid stability. Future work may explore integrating real-time data streams and further optimizing the model for broader deployment in smart grid environments.

References

[1] M. Ahmed *et al*., "Energy Theft Detection in Smart Grids: Taxonomy, Comparative Analysis, Challenges, and Future Research Directions," in *IEEE/CAA Journal of AutomaticaSinica*, vol. 9, no. 4, pp. 578-600, April 2022, doi: 10.1109/JAS.2022.105404.

[2] Z. Yan and H. Wen, "Performance Analysis of Electricity Theft Detection for the Smart Grid: An Overview," in *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-28, 2022, Art no. 2502928, doi: 10.1109/TIM.2021.3127649.

[3] X. Xia, Y. Xiao, W. Liang and J. Hui, "Detection Methods in Smart Meters for Electricity Thefts: A Survey," *Proceedings of the IEEE*, vol. 110, no. 2, pp. 273-319, Feb. 2022, doi: 10.1109/JPROC.2021.3139754.

[4] Deepa, K.R., Thillaiarasu, N. "Integrated Architecture for Smart Grid Energy Management: Deep Attention-Enhanced Sequence-to-Sequence Model with Energy-Aware Optimized Reinforcement Learning for Demand Response. SN COMPUT. SCI. 5, 1017 (2024). https://doi.org/10.1007/s42979-024-03305-2.

[5] A. Takiddin, M. Ismail, M. Nabil, M. M. E. A. Mahmoud, and E. Serpedin, ''Detecting electricity theft cyber-attacks in AMI networks using deep vector embeddings,'' IEEE Syst. J., vol. , no. 3, pp. 4189–4198, Sep. 2021.

[6] M. M. Badr, M. I. Ibrahem, M. Mahmoud, M. M. Fouda, F. Alsolami, and W. Alasmary, ''Detection of false-reading attacks in smart grid net metering system,'' IEEE Internet Things J., vol. 9, no. 2, pp. 1386–1401, Jan. 2022.

[7] L. J. Lepolesa, S. Achari, and L. Cheng, ''Electricity theft detection in smart grids based on deep neural network,'' IEEE Access, vol. 10, pp. 39638–39655, 2022.

[8] Elgarhy, M. M. Badr, M. M. E. A. Mahmoud, M. M. Fouda, M. Alsabaan and H. A. Kholidy, "Clustering and Ensemble Based Approach for Securing Electricity Theft Detectors Against

Evasion Attacks," in *IEEE Access*, vol. 11, pp. 112147-112164, 2023, doi: 10.1109/ACCESS.2023.3318111

[9] J. Anin, M. J. Khan, O. Abdelsalam, M. Nabil, F. Hu and A. Alsharif, "Efficient and Privacy-Preserving ConvLSTM-based Detection of Electricity Theft Cyber-Attacks in Smart Grids," in *IEEE Access*, doi: 10.1109/ACCESS.2024.3478068.

[10] Aldegheishem, M. Anwar, N. Javaid, N. Alrajeh, M. Shafiq and H. Ahmed, "Towards Sustainable Energy Efficiency With Intelligent Electricity Theft Detection in Smart Grids Emphasising Enhanced Neural Networks," in *IEEE Access*, vol. 9, pp. 25036-25061, 2021, doi: 10.1109/ACCESS.2021.3056566.

[11] E. U. Haq, C. Pei, R. Zhang, H. Jianjun, and F. Ahmad, ‘‘Electricity-theft detection for smart grid security using smart meter data: A deep-CNN based approach,’’ Energy Rep., vol. 9, pp. 634–643, Mar. 2023.

[12] L. Zhu, W. Wen, J. Li, C. Zhang, B. Zhou, and Z. Shuai, ‘‘Deep active learning-enabled cost-effective electricity theft detection in smart grids,’’ IEEE Trans. Ind. Informat., 2023.

[13] M. Ismail, M. F. Shaaban, M. Naidu, and E. Serpedin, ‘‘Deep learning detection of electricity theft cyber-attacks in renewable distributed generation,’’ IEEE Trans. Smart Grid, vol. 11, no. 4, pp. 3428–3437, Jul. 2020.

[14] D. K. R and M. M. Kodabagi, "Data Analytics Challenges and Needs in Smart Grid for Smart Energy Management," *2024 Second International Conference on Advances in Information Technology (ICAIT)*, Chikkamagaluru, Karnataka, India, 2024, pp. 1-6, doi: 10.1109/ICAIT61638.2024.10690834..

[15] Li, Shizhong, et al. "Feature Attention Distillation Defense for Backdoor Attack in Artificial Neural Network-Based Electricity Theft Detection." IEEE Internet of Things Journal (2024).

[16] W. Liao *et al*., "Transfer Learning-Driven Electricity Theft Detection in Small-Sample Cases," in *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1-13, 2024, Art no. 2532013, doi: 10.1109/TIM.2024.3470066.

[17] S. Zhu, Z. Xue and Y. Li, "Electricity Theft Detection in Smart Grids Based on Omni-Scale CNN and AutoXGB," in *IEEE Access*, vol. 12, pp. 15477-15492, 2024, doi: 10.1109/ACCESS.2024.3358683.

[18] L. Zhu, W. Wen, J. Li, C. Zhang, B. Zhou and Z. Shuai, "Deep Active Learning-Enabled Cost-Effective Electricity Theft Detection in Smart Grids," in *IEEE Transactions on Industrial Informatics*, vol. 20, no. 1, pp. 256-268, Jan. 2024, doi: 10.1109/TII.2023.3249212.

[19] W. Zhou *et al*., "Electricity Theft Detection of Residential Users With Correlation of Water and Electricity Usage," in *IEEE Transactions on Industrial Informatics*, vol. 20, no. 4, pp. 5339-5349, April 2024, doi: 10.1109/TII.2023.3332954.

[20] Ravishankar, H., et al. "Comparative Analysis and QoS Enhancement Through Novel Feedback Architecture." *2023 International Conference on Data Science and Network Security (ICDSNS)*. IEEE, 2023.

[21] M. Zanetti, E. Jamhour, M. Pellenz, and M. Penna, ‘‘A new SVM-based fraud detection model for AMI,’’ in Computer Safety, Reliability, and Security, A. Skavhaug, J. Guiochet, and F. Bitsch, Eds. Cham, Switzerland: Springer, 2016, pp. 226–237.

[22] D. K. R, R. H, L. R, A. M, S. G and C. S. M, "Accuracy Enhance of Smart Energy Theft Detection Using Machine Learning Classifiers," *2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*, Bangalore, India, 2024, pp. 1-6, doi: 10.1109/ICITEICS61368.2024.10624908.

[23]B. C. Costa, B. L. A. Alberto, A. M. Portela, and E. O. Eler, ‘‘Fraud detection in electric power distribution networks using an ann-based knowledge-discovery process,’’ Int. J. Artif. Intell. Appl., vol. 4, no. 6, pp. 17–23, Nov. 2013.

[24]Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, ‘‘Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids,’’ IEEE Trans. Ind. Informat., vol. 14, no. 4, pp. 1606–1615, Apr. 2018.

[25]M. N. Hasan, R. N. Toma, A.-A. Nahid, M. 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, Aug. 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/17/3310