Artificial Intelligence-based Real-Time Electricity Metering Data Analysis and its Application to Anti-Theft Actions

Kai Liu, Anlei Liu*, Xun Ma, Xuchao Jia
State Grid Hebei Electric Power Co., Ltd., Shijiazhuang, China

***Abstract*—This study focuses on the key issue of anti-stealing behavior identification in power systems, aiming to improve the security and efficiency of power energy management. Under the current background of intelligent power grid, the existence of anti-theft phenomenon not only causes serious economic losses, but also poses a threat to the stability of power grid operation. Aiming at this situation, this paper proposes a novel and effective feature extraction and optimization method, which utilizes the recursive feature elimination (rfe) technique, combined with the correlation and exclusion analysis of the features, to achieve the deep screening and dimensionality reduction of a large amount of raw data, so as to refine the core feature set that has the most differentiation for the anti-stolen power behavior. During the research process, this paper constructed a hybrid model integrating long short-term memory network (LSTM) and autoencoder. The model cleverly combines the advantages of LSTM in capturing time series dependency and the powerful ability of autoencoder in feature learning and noise reduction, and is especially designed for targeted enhancement of anti-electricity theft behaviors to achieve real-time and accurate behavior recognition. In order to verify the performance and practicality of the proposed method, this paper carries out rigorous simulation experiments and practical case studies. By comparing the classical anti-electricity theft recognition methods, the results show that the hybrid model proposed in this study exhibits significant advantages in both recognition accuracy and response speed. Whether in the simulation environment or actual application scenarios, this method can effectively identify and warn potential power theft behavior, thus providing a strong technical support for the power company's anti-power theft management.**

***Keywords*—Artificial intelligence; real-time electrical energy; metering data analysis; anti-power theft**

I. Introduction

With advanced sensing technology, communication technology and data analysis technology, smart grid realizes real-time monitoring and two-way interaction of electric energy, which greatly improves the operational efficiency and service level of the grid. The modern electric energy metering system follows this trend and is evolving in the direction of more intelligent, fine and interactive. The wide application of smart meters makes the collection frequency of electric power data change from the traditional monthly or even annual to the second or even millisecond level, generating a huge amount of real-time electric power measurement data. These data contain a wealth of information about user load characteristics, grid operation status, equipment health, etc., and are of inestimable value for realizing the balance between power supply and demand, optimizing grid scheduling, preventing equipment failures, and serving personalized needs. In the face of such a huge and fast-generating real-time data stream, power companies are facing serious data processing challenges [1]. On the one hand, high-speed and stable communication networks and powerful data center infrastructure are needed to ensure real-time data transmission and storage; on the other hand, high-performance data processing and analysis tools must be used to realize real-time data analysis, anomaly detection and deep mining, which can be transformed into practical decision-making basis and business insights [2].

Real-time power metering data usually includes, but is not limited to, current, voltage, frequency, power factor, active/reactive power, power accumulation, and other types, and due to the extremely high collection frequency, the data presents continuous, dynamic and multi-dimensional characteristics. This high-density, high-frequency data provides a near real-time panoramic view of the power system, which is conducive to the timely detection of power consumption anomalies, diagnosis of grid faults, as well as load forecasting, energy consumption management and other work [3].

Anomaly detection demand is an important part of real-time power metering data analysis, especially critical in the anti-stolen power work. Through the real-time monitoring and analysis of the user's electricity consumption data, changes that do not conform to the conventional pattern of electricity consumption can be quickly identified, such as a sudden increase in the amount of electricity, abnormal power hours, load curve pattern distortion, etc., which are the typical characteristics of potential power theft [4]. Currently, the phenomenon of power theft shows the diversity and covert coexistence of characteristics, methods are also constantly renovated and upgraded, both at the physical level of direct tampering with the meter readings, wiring and stealing, but also the use of new technologies to bypass the smart meter monitoring system of the new type of power theft behavior. The traditional discrimination of power theft behavior mainly relies on manual monitoring, which is often carried out with the help of methods such as the outlier detection method in statistical principles [5]. Recently, a human-computer cooperative method for discriminating power theft users has been proposed [6]. In this method, the grid company sets a suitable model discrimination threshold according to its own characteristics and needs. Subsequently, the model's preliminary results are carefully screened by combining the model's basis of judgment

*Corresponding Author.

with the human's empirical knowledge. This method significantly reduces the number of users requiring on-site screening, thereby reducing labor and material costs. In recent years, artificial intelligence technology has made significant breakthroughs in the field of power data analysis, especially in the areas of deep learning, reinforcement learning and anomaly detection algorithms. For example, a research team proposed the structure of dbn and its learning algorithm, and applied it to the anomaly detection of anti-power theft [7]. These techniques are able to identify minor deviations in electricity consumption behavior in real time, thus effectively preventing and detecting electricity theft. The process is shown in Fig. 1.

Although the research on AI-based anti-power theft algorithms has achieved remarkable results in practice, it still faces some challenges, such as controlling the false alarm rate while maintaining a high recognition accuracy, optimizing the model complexity to adapt to the requirements of real-time processing of large-scale power metering data, as well as effectively coping with the rising recognition difficulty brought about by the diversification of power theft means. In view of this, the research content of this paper mainly focuses on the following core points: Firstly, this paper focuses on feature selection and optimization, using advanced feature selection techniques such as recursive feature elimination to select a set of features that are highly correlated with electricity theft from the complicated electricity metering data. These features have strong correlation and exclusivity, which can accurately portray the key features of electricity theft behavior and effectively filter out redundant information and noise interference, thus simplifying the model structure and reducing the complexity of the model at the source. Finally, through the processing and analysis of the actual collected electricity metering data, this paper verifies the performance of the model in identifying electricity theft behavior.

The innovation of this paper lies in that a novel and efficient feature extraction and optimization method is proposed, aiming at the key problem of electricity theft recognition under the background of smart grid. This method innovatively integrates recursive feature elimination (RFE) technology with feature correlation and exclusion analysis, realizes deep screening and dimensionality reduction of original big data, extracts the core feature set that can distinguish the behavior of stealing electricity, and effectively solves the problem of high-dimensional data processing. Furthermore, a hybrid model of LSTM and auto-encoder is constructed. This design skillfully utilizes the ability of LSTM to capture time series dependence and the powerful feature learning and noise suppression functions of auto-encoder to enhance the recognition accuracy of theft behavior and ensure the real-time and accuracy. Through rigorous simulation experiments and practical case applications, the proposed method has significant advantages in recognition accuracy and response speed, not only in the simulation environment, but also in the real application scenarios, which can effectively identify and warn potential electricity theft behavior. It provides strong technical support for anti-electricity theft management in power enterprises and has important practical significance and innovative value.

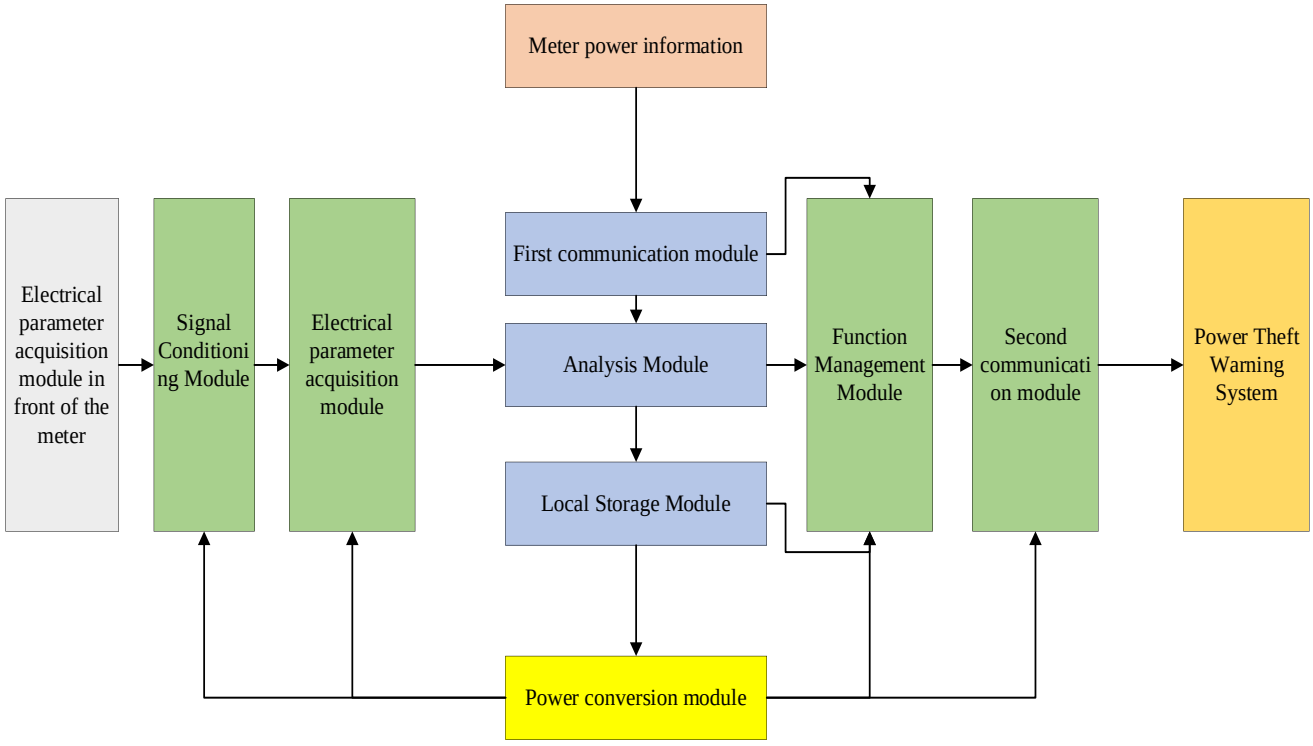

Fig. 1. Power monitoring process

II. Literature Review

A. Traditional Methods

In the study of traditional methods of electricity metering data analysis and anti-power theft, statistical analysis means dominate. Literature [8] elaborates on how statistical parameters such as mean, peak, valley, standard deviation, etc. Can be calculated by digging deeper into a user's historical electricity consumption data in order to reveal the normality and abnormal changes in electricity consumption behavior. For example, when a user experiences a sharp increase in electricity consumption that cannot be reasonably explained within a specific unit of time [9], such an abnormal change is often regarded as a preliminary warning signal of possible electricity theft. In addition, the role of seasonal, cyclical and trend analyses in identifying abnormal electricity usage patterns has been further explored in the literature [10], which can assist in identifying potential electricity theft activities by comparing the discrepancies between the actual electricity consumption data and the users' normal electricity consumption habits. Load profile analysis is another key tool, as described in the literature [11], by fine-tuning the daily, weekly and even yearly load profile patterns of consumers, potential abnormal electricity usage patterns can be revealed. For example, as mentioned in the literature [12], if a significant decrease in load is observed at night while an abnormal increase is observed during the day, or if the difference between weekend and weekday loads is much higher than expected, such anomalies are likely to imply the existence of electricity theft. In addition, the literature [13] uses the comparison and cluster analysis of a large number of customer load profiles to effectively identify groups of customers that deviate significantly from the normal pattern of electricity consumption. Power quality parameters are also important considerations in assessing the potential for electricity theft. As pointed out in the literature [14], voltage fluctuations, current imbalance, and abnormal changes in power factor may originate from users taking illegal wiring, changing metering equipment and other power theft behaviors, resulting in the power system being subjected to abnormal disturbances. For example, sudden voltage dips or current distortions are often indicative of potential power theft activities [15].

B. Deep Learning-based Electricity Metering Data Analysis and Anti-Theft Methods

1) Research advances in deep learning for preprocessing and feature extraction: Deep learning in the field of power data preprocessing and feature extraction has become a cutting-edge focus in the electric power industry, and plays a crucial role in improving data processing performance and accuracy, especially in building an efficient and robust analytical foundation for complex tasks such as power theft detection. In recent years, the emergence of many pioneering methods and techniques has greatly contributed to the development of this field [16].

Convolutional neural network (cnn)-based strategies have received much attention and are widely used in the preprocessing stage of electricity data. This approach skillfully transforms the original one-dimensional or multi-dimensional electricity metering data into image-type representations of two-dimensional or higher dimensions, enabling cnns to fully utilize their powerful local feature extraction capabilities in the image domain to capture the spatio-temporal patterns in the electricity data. Through this transformation and preprocessing process, the noise components and redundant information in the original data are effectively suppressed and filtered, making the data representation more concise and rich in key information, which in turn enhances the effectiveness and reliability of the subsequent feature extraction and electricity theft identification tasks [17].

Deep learning breakthroughs in feature learning are also significant. For example, the introduction of deep stacked denoising auto-encoder (dsaae) architecture has become a promising feature learning framework, which mines and reconstructs the low-dimensional latent structure of the original electricity data, and with the help of the deep network's layer-by-layer abstraction and noise reduction ability, it successfully refines the feature vectors that have a high degree of differentiation of the electricity theft behaviors. This approach overcomes the limitations of traditional manual feature engineering, greatly improves the quality of feature representation, and thus contributes to a more accurate and fine-grained identification of power theft behavior. In addition, researchers have actively tried to use other innovative technological tools to improve electricity data preprocessing and feature extraction. For example, generative adversarial networks (GAN), an emerging technology, is used to generate realistic synthetic data to expand the size of the original dataset and enhance the model's adaptability to anomalies and generalization performance to unknown data. Meanwhile, multimodal data fusion is also an important trend in current research. By integrating information resources from different sensor devices or multiple types of data sources, researchers are able to construct a more integrated and comprehensive feature space to ensure that hidden correlations and nuances are fully revealed when analyzing power data [18, 19].

To summarize, the research pace of deep learning in the field of power data preprocessing and feature extraction is fast and diversified, from cnn-guided data preprocessing innovation to the construction of deep learning-driven feature learning frameworks, and then to GAN-generated data extensions and multimodal data fusion and other innovative practices, all of which have vigorously pushed forward the accuracy and efficiency of power data analysis. With the continuous progress and application expansion of deep learning technology, more advanced methods and technical solutions are bound to emerge in the future, further empowering the power data processing process and providing more powerful and flexible support for many application scenarios such as power system operation monitoring, fault diagnosis, energy management, etc., thus promoting the intelligent development of the entire power industry [20, 21].

2) Research progress of deep learning in electricity theft recognition: Lotfipoor et al. [22] presents a novel electricity theft detection model that combines migration learning and recurrent neural network (rnn). The model first uses a pre-trained model to learn generic electricity consumption patterns on a large-scale public dataset, and then migrates the learned

knowledge to the target scenario, capturing the time-series characteristics of the user's electricity consumption behavior through rnn, which significantly improves the recognition accuracy of electricity theft. Lu et al. [23] innovatively combining generative adversarial networks (GAN) and variable auto-encoders (VAE), a hybrid GAN-VAE model is constructed for the detection of electricity theft. The GAN is used to simulate the distribution of normal electricity consumption behaviors, while the VAE is used to extract the abnormal patterns, and the joint use of the two is able to accurately isolate the subtle features of the theft behaviors from the massive amount of electricity metering data. Mangat et al. [24] in their study, the attentive mechanism (a) was introduced into the LSTM model, resulting in the attentive LSTM model, which is able to pay targeted attention to key times and variables in the electricity consumption behavior, thus capturing potential electricity theft behaviors more accurately in real-time electricity metering data analysis. A widely used approach is semi-supervised learning, which improves the performance of the model by combining labeled and unlabeled data. In electricity theft identification, semi-supervised learning becomes an effective method because electricity theft data is usually difficult to obtain, while normal electricity consumption data is relatively easy to obtain. Researchers can use a small amount of labeled data and a large amount of unlabeled data to train the model to achieve better results in the recognition of electricity theft. Another important research direction is multimodal data fusion, i.e., fusing information from different sensors or data sources for analysis [25, 26]. In power theft identification, multiple data sources such as power metering data, video surveillance data, temperature sensor data, etc. Can be combined so as to capture the characteristics of power theft behavior more comprehensively. The deep learning model can better understand complex environments and scenarios by means of multimodal data fusion to improve the identification accuracy of power theft behavior. In addition, with the continuous development of deep learning technology, graph neural network (gnn) is gradually applied in the identification of power theft behavior.gnn is a deep learning model specialized in processing graph-structured data, which can effectively capture the topology and the relationship between users in the power system. By modeling the power system as a graph structure, researchers can use gnn to mine the association of power usage behaviors among users, thus identifying power theft more accurately. In addition, it is worth noting the application of privacy-preserving techniques in power theft recognition. Since electricity theft recognition involves users' electricity consumption data, privacy protection becomes an important research topic. Deep learning techniques can be combined with privacy protection methods such as differential privacy and homomorphic encryption, so as to realize the effective identification of power theft under the premise of protecting user privacy [27].

In today's rapidly evolving technology landscape, the integration of DevOps and blockchain smart contracts shows unprecedented potential, ushering in a new era of digital innovation and business process optimization. Recent research shows that by combining mature DevOps toolchains such as GitHub Actions, GitLab CI/CD, etc. with test frameworks designed specifically for blockchain (such as Hardhat, Brownie), the development efficiency and security of smart contracts can be significantly improved. This integration not only enables a seamless process from code submission to automated testing, compilation, and deployment, but also facilitates continuous monitoring of code quality and security. For example, research has shown that these tools can be used to automate complex smart contract test cases, including but not limited to state transition verification, gas cost optimization, and potential vulnerability scanning, ensuring that contracts meet the highest standards before deployment. In the latest research and development results, researchers have explored how to achieve continuous integration and security upgrades of smart contracts while maintaining the tamper-proof characteristics of blockchain. With a hybrid architecture of off-chain computation and on-chain validation, and mechanisms that leverage on-chain oracles to trigger contract upgrades, development teams can iteratively update smart contracts without interrupting service. This strategy not only improves the flexibility of the contract, but also ensures the stability and security of the system, providing a new perspective for solving the "invariant dilemma" of smart contracts. In the latest DevOps practices, advanced monitoring tools (such as Prometheus, ELK Stack) are integrated into blockchain networks to track the running status of smart contracts in real time, collect performance metrics, and analyze these data through machine learning algorithms to automatically identify abnormal patterns and performance bottlenecks. This approach not only can warn potential failures in advance, but also dynamically adjust resource allocation according to the actual operation of smart contracts to optimize cost effectiveness, reflecting the latest application of AI in automated operation and maintenance. Combined with the latest compliance frameworks and security protocols, DevOps processes incorporate automated audits and compliance checks for smart contracts. Static analysis using tools such as Slither and Mythril, combined with automated penetration testing, can identify and fix security vulnerabilities before smart contracts are deployed, ensuring that contracts meet international data protection regulations such as GDPR. In addition, by automating audit records and report generation through smart contracts, companies can streamline regulatory reporting processes and increase transparency [28, 29].

In summary, combined with the latest research results, the integration of DevOps and blockchain smart contracts is advancing technological innovation at an unprecedented speed, ensuring the efficiency, security and compliance of smart contract development through highly automated and intelligent tools and processes, laying a solid foundation for large-scale application of blockchain technology.

III. Feature Extraction of Electricity Metering Data and its Optimization

In the feature extraction process of electric energy metering data, the recursive feature elimination (rfe) method is adopted for the multiple considerations of data dimension optimization, improving the model generalization ability and simplifying the

model interpretability, and its technical framework is shown in Fig. 2. Electricity metering data often contains a large number of raw features, such as current, voltage, power factor, load profile and other dynamic and static parameters, which may be highly correlated or contain redundant information, and the over-abundance of features may introduce noise, increase the complexity of model training, and may even lead to overfitting phenomenon .

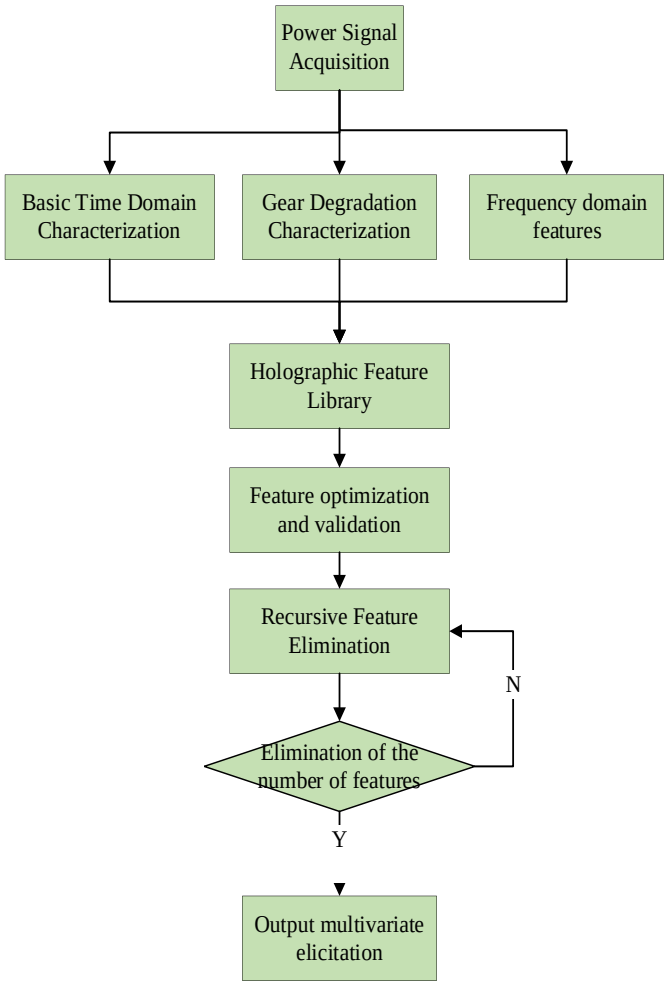

Fig. 2. Recursive feature elimination

A. Application of Recursive Feature Elimination Methods to Electricity Metering Data

Recursive feature elimination (rfe) is a feature selection method that gradually reduces the size of the feature set. In the energy metering data scenario, this paper can use support vector machine (svm) with rfe for feature selection. The kernel function of svm is denoted as $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$, where x_i, x_j is the sample of energy metering data and $\phi(\cdot)$ is the mapping function that maps the original features to the high-dimensional feature space. The svm finds the optimal classification boundaries by maximizing the intervals, and its corresponding lagrangian function is $[L_D(\alpha, \beta, \theta) = \frac{1}{2} \| \theta \|^2 - \sum_{i=1}^{m} \alpha_i [y_i(\theta^T x_i + b) - 1] - \sum_{i=1}^{m} \beta_i \alpha_i$, in rfe, the weight vector obtained from svm training is used to evaluate the importance of the features. The weight vector θ obtained from the training is used to evaluate the importance of the features. A threshold or a fixed step size is set to remove features with smaller weights one by one, and then the model is trained again and the feature weights are recalculated until the desired number of features is reached [30, 31].

B. Feature Optimization and Validation

Feature optimization is not only feature selection, but also includes feature transformation and normalization. For example, for abnormal fluctuations in electricity metering data, a sliding window standard deviation method can be applied for smoothing: $\hat{x}_t = x_t - \mu_w + \sigma_w$, where x_t is the value of the original data at time point t, and μ_w and σ_w are the mean and standard deviation within the sliding window, respectively. In addition, the texture information of the features can be extracted using methods such as local binary patterns (lbp) or fourier transform. After the initial feature optimization, dimensionality reduction is performed by principal component analysis (pca), $Z = XP$, where Z is the reduced feature matrix, X is the original feature matrix, and P is the first k columns of the eigenvalue matrix (corresponding to the largest k eigenvalues) calculated by pca. In the validation stage, k-fold cross-validation is used to evaluate different feature subsets and model parameter combinations. Taking the logistic regression model as an example, its likelihood function is $L(\theta) = \prod_{i=1}^{m} p(y^{(i)} \mid x^{(i)}; \theta)^{y^{(i)}} (1 - p(y^{(i)} \mid x^{(i)}; \theta))^{1-y^{(i)}}$, the optimal feature subset and model parameters are selected by maximizing the likelihood function or minimizing the log-likelihood loss function, combined with the cross-validation results [32].

C. Correlation and Exclusion Analysis of Electricity Metering Data Characteristics

Correlation and exclusion analysis of electric energy metering data features is an important part of the data preprocessing and feature selection process, aiming at mining and understanding the intrinsic connection and independence between different electric energy metering features. Correlation analysis is mainly to study the degree of statistical correlation between different electric energy features, such as the linear or non-linear relationship between current, voltage, power and other factors, and quantify the degree of dependence between the features by calculating the correlation coefficient and other ways, so as to eliminate repetitive or redundant information, and to avoid the impact of the multi-collinearity problem on the subsequent training of the model and the accuracy of prediction. Exclusivity analysis, on the other hand, refers to identifying and evaluating electricity metering features that are mutually exclusive and mutually exclusive, for example, the characteristics of electricity consumption behavior in some specific time periods may be in conflict or exclusivity with other time periods. Through exclusivity analysis, combinations of features that are unlikely to occur simultaneously under certain conditions can be identified and excluded, which is crucial for improving model interpretability, preventing misleading information inputs, and enhancing the accuracy of tasks such as

fault detection or electricity theft identification. Correlation and exclusion of characteristics can be quantified by calculating statistics such as Pearson's correlation coefficient, mutual information, and conditional independence test. Pearson's correlation coefficient is defined as: $r_{ij} = \frac{cov(X_i, X_j)}{\sigma_{X_i}\sigma_{X_j}}$, mutual information (mi) measures the degree of reduction of uncertainty between two random variables: $I(X_i; X_j) = \sum_{x_i}\sum_{x_j} p(x_i, x_j)\log\frac{p(x_i, x_j)}{p(x_i)p(x_j)}$, for feature exclusion analysis, conditional independence test can be considered to identify whether two features exhibit significant exclusion under certain conditions [33].

IV. Hybrid Model based on LSTM and Autoencoder

The proposed hybrid framework combines the advantages of LSTM and self-encoder, and is carefully designed to meet the increasingly complex requirements of electricity theft detection in power systems. Traditional single model is often difficult to take into account both the long-term dependence of time series data and the fine identification of abnormal behavior. The innovation of this framework lies in: Firstly, the LSTM layer is used to deeply mine the time series features of electric energy measurement data, and its memory unit can effectively capture the complex patterns of power consumption behavior evolving with time, thus solving the problem that it is difficult to extract long-term dependence relations from high-dimensional time series data. Secondly, by incorporating the self-encoder, especially the variational self-encoder (VAE), not only the data distribution under normal power consumption mode is learned to realize the sensitive detection of abnormal deviation, but also the probability distribution of hidden variables is introduced to enhance the ability of the model to express the uncertainty of power consumption behavior, so that small abnormal changes such as electricity theft have no place to hide. In addition, the attention mechanism introduced in the framework enables the model to focus on key features at different time nodes according to the importance of information, further improving the analysis accuracy and model interpretation.

The framework is especially suitable for scenarios that require efficient identification of abnormal electricity consumption behavior and prevention of electricity theft. It can adapt to large data volume and high dynamic energy metering environment. It not only improves the accuracy and timeliness of anti-electricity theft detection, but also ensures the robustness and generalization ability of the model through multi-stage collaborative training strategy. Therefore, this hybrid model framework provides a powerful tool for power companies and regulators to achieve intelligent and precise power safety management.

In this section, this paper will explore in detail an innovative hybrid model that skillfully blends a long short-term memory network (LSTM) with an autoencoder, the framework of which is specifically shown in Fig. 3, in order to achieve efficient identification of anti-stealing behaviors in power systems. This hybrid model fully utilizes the strong capture capability of LSTM for long-term dependencies in time series data and the learning and reconstruction advantages of autoencoder for normal state data distribution [34, 35].

A. Modeling Framework

Hybrid model is mainly composed of two core components: The LSTM layer and the self-encoder layer. Firstly, for the power usage behavior data of the power system, due to its inherent time series characteristics, this paper adopt LSTM for deep learning. The mathematical expression of the LSTM model can be expressed as: $h_t = \text{LSTM}(x_t, h_{t-1}, c_{t-1})$, where x_t represents the input feature vector at time step t, h_t is the hidden state, which contains both the current moment and integrates the historical information, and c_t is the unitary state, which is used for storing the long-term dependency information. By stacking multiple layers of LSTMs, the model is able to effectively mine the potential patterns of power usage behavior over time. This paper introduce the self-encoder part, whose main task is to learn and reconstruct the data distribution of normal electricity usage behavior in order to facilitate the detection of abnormal behavior, i.e., possible electricity theft. The self-encoder mainly consists of two parts, the encoder and the decoder, and its basic structure can be described as follows: $z = f_E(x) = \text{Encoder}(x)$ $\hat{x} = f_D(z) = \text{Decoder}(z)$, where f_E is the encoder function that maps the original input data x to the low-dimensional potential space to obtain the encoding vector z; f_D is the decoder function that tries to recover the original input data from the encoding vector z $\hat{x}$. Optimize the autocoder by minimizing the reconstruction error (e.g. Mean square error mse): $L_{AE} = \| x - \hat{x} \|_2^2$.

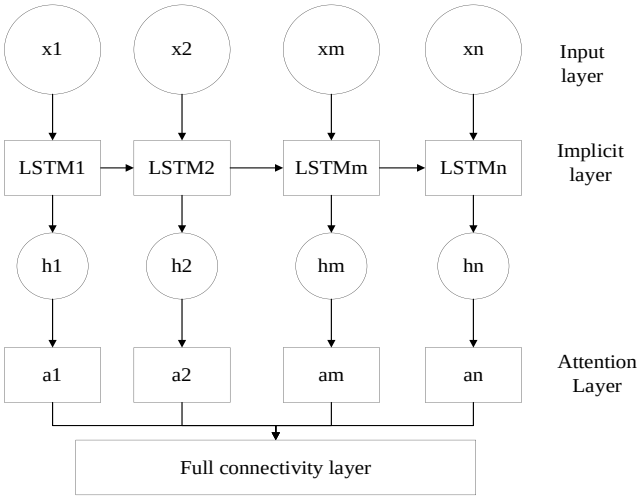

Fig. 3. Modeling framework

B. Enhanced Design of Anti-Theft Mechanisms

In order to detect electricity theft in a more refined way, this paper improved the design of the autoencoder part by adopting the variational autoencoder (VAE) structure. VAE is modeled by introducing the probability distribution of the hidden variable z instead of a single encoded value, which better characterizes

the uncertainty of electricity consumption behavior and helps to distinguish subtle anomalous variations. VAE's encoder produces the mean μ and variance σ^2: $\mu, \log(\sigma^2) = f_E(x; \theta_e)$, and then samples the hidden variables through a reparameterization technique: $z = \mu + \sigma \cdot \grave{o}, \quad \grave{o} \sim \mathrm{N}(0, I)$. The decoder part still tries to reconstruct the input, but the loss function now includes a kl scatter term to ensure that the hidden variable z is close to the standard normal distribution: $\mathrm{L}_{VAE} = \mathrm{E}_{q(z|x)}[\log p(x \mid z)] - D_{KL}(q(z \mid x) \| p(z))$. Where the first term is the reconstruction error and the second term is the kl scatter, which measures the difference between the post-coding distribution $q(z \mid x)$ and the prior distribution $p(z)$. By optimizing this loss function, VAE is able to maintain the quality of data reconstruction while ensuring that the model is able to detect abnormal electricity usage behaviors that deviate from the normal distribution, thus improving the accuracy of anti-theft detection. In addition, this paper introduce attention mechanism between the LSTM and the self-encoder, which allows the self-encoder to adjust its attention to the input features according to the importance of different time periods. The attention mechanism can be defined as: $\alpha_{t,i} = \frac{\exp(score(h_t, h_i))}{\sum_j \exp(score(h_t, h_j))}$.

Where h_t is the hidden state of the LSTM at time step t, h_i is one of the hidden states selected from all the hidden states output by the LSTM, the $score$ function calculates the correlation score between the two, and $\alpha_{t,i}$ denotes the attention weight for time step i when the input is given to the self-encoder.

C. Multi-Stage Co-Training Process of the Model

The training of the model is not completed at one time, but adopts a multi-stage collaborative training strategy, and its training process is specifically shown in Fig. 4. First, in the pre-training stage, the LSTM is individually trained to capture the dynamic patterns of the time series: $\mathrm{L}_{LSTM} = \sum_t loss(h_t, y_t)$. Where $loss$ is the cross entropy and y_t is the corresponding label or prediction target. Subsequently, this paper fix the LSTM parameters and train the VAE using the sequence features extracted by the LSTM. Here, this paper consider the weighted average of the hidden states of all time steps as the input to the self-encoder: $\bar{h} = \sum_t \alpha_t \cdot h_t$, followed by optimization using the VAE loss function L_{VAE}. Finally, in the joint training phase, this paper combine the prediction loss of the LSTM and the reconstruction loss of the VAE, and add a regularization term to avoid overfitting: $\mathrm{L}_{total} = \beta_1 \mathrm{L}_{LSTM} + \beta_2 \mathrm{L}_{VAE} + \lambda \| \theta \|_2^2$. Among them, β_1 and β_2 are hyperparameters controlling the weights of different loss terms, and $\| \theta \|_2^2$ is the l2 regularization term to prevent the overfitting problem caused by the overly complex model.

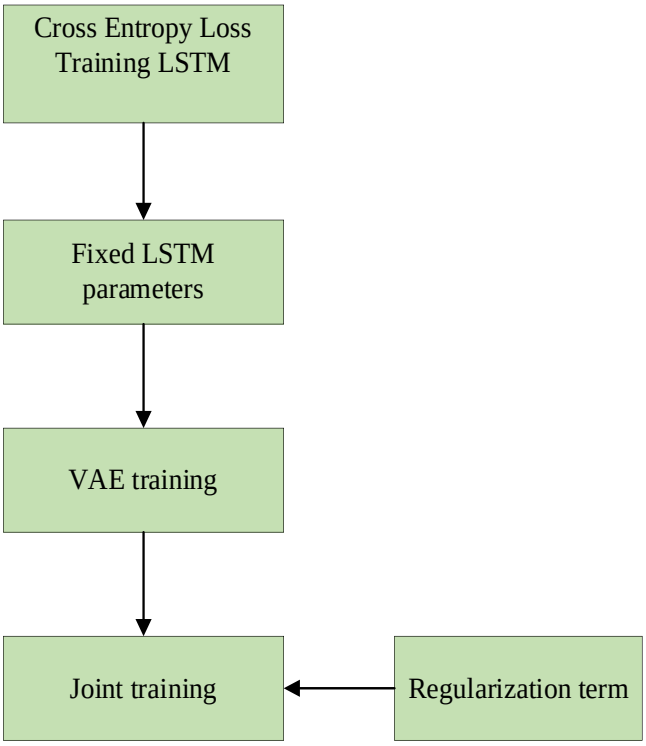

Fig. 4. Training flow

Although the proposed hybrid model of LSTM and autoencoder shows significant advantages in dealing with theft detection in power systems, this framework is not omnipotent. Because its design focuses on analyzing complex patterns and anomaly detection of time series data, it may not be suitable for the following scenarios: firstly, static data analysis tasks, such as feature extraction and classification of one-time and discontinuous data, because LSTM is optimized for time series data in network structure and is not efficient for non-time series data processing; Second, for highly nonlinear and extremely complex data correlation problems, if the problem involves extremely complex feature interaction effects and is difficult to simplify through existing feature engineering methods, the model may not directly provide satisfactory solutions; third, for application scenarios with high real-time requirements, the multi-stage collaborative training strategy may increase the time cost of model deployment and affect the real-time decision process. Therefore, when faced with these specific types of problems, it may be necessary to explore other more targeted models or algorithms.

V. Experimental Evaluation

A. Simulation Experiments

The aim of this section is to comprehensively evaluate the performance effect of the hybrid model based on long short-term memory network (LSTM) and autoencoder proposed in this paper on the task of antitheft and compare it with other classical approaches. This paper will also explore the contribution of recursive feature elimination (rfe) method in feature selection.

B. Simulation Experiment Design and Evaluation

This research relies on an advanced data analysis and deep learning experimental environment, which is configured with a high-performance computing server, equipped with the latest version of python development environment and tensorflow deep learning framework, supplemented with numpy, scikit-learn and other related scientific computing libraries. The experimental platform is equipped with large-scale data processing capabilities, and is able to efficiently perform the training and validation tasks of hybrid LSTM and self-encoder models.

The data used in the experiment originates from the electricity consumption record database of real power system users, covering multiple key parameters such as electricity consumption, voltage, current and so on at different times of the day and night, and constructing a rich and multi-dimensional time series dataset. These data reflect the temporal dynamic changes of users' electricity consumption behavior in detail, providing a solid foundation for model learning. In order to ensure the generalization performance of the model, this paper reasonably divide the whole dataset into training, validation and testing sets, with the ratio of the three set at 70%, 15% and 15%. Each sample contains continuous historical electricity consumption data in order to facilitate LSTM to capture the long-term dependency of the time series. Before formal training, this paper first apply the recursive feature elimination (rfe) technique to screen and optimize the original features [2] based on the performance of the model on the validation set, rfe sequentially eliminates the features that have less impact on the predictive performance of the model, and then refines the core set of features with the most representative and predictive power. This paper recorded the changes in model performance before and after rfe optimization with different numbers of features, as shown in Table I.

TABLE I. EFFECT OF RFE FEATURE OPTIMIZATION ON MODEL PERFORMANCE

Number of features	Performance after rfe	Performance without rfe
10	0.85	0.80
20	0.87	0.82
30	0.88	0.84

The analysis of Table I shows that with the rfe optimization, the overall performance of the model improves even when the number of features is reduced, confirming the effectiveness of rfe in feature selection.

For the proposed hybrid model based on LSTM and self-encoder, this paper conducted extensive comparative experiments comparing it with traditional statistical detection methods, classical machine learning models (e.g., support vector machines (svm), random forest), and single-model schemes using only LSTM or self-encoder. The experimental results are summarized in Tables II and III.

Referring to the experimental results in Tables II and III, the hybrid model shows strong performance in all major evaluation metrics, especially in accuracy, f1-score, and auc-roc, which outperforms the other comparative models, which highlights the superiority of the hybrid model in the task of anti-stolen electricity detection through the comparative analysis, it can be clearly seen that the hybrid model based on LSTM and self-encoder shows excellent ability in dealing with the task of power theft detection in the power system, which effectively combines the in-depth understanding of long and short-term memory network on the time series and the self-encoder's ability to learn the distribution of the data of the normal power consumption state, thus realizing the accurate identification and effective monitoring of the power theft behaviors [3].

TABLE II. COMPARISON OF THE PERFORMANCE OF DIFFERENT MODELS ON THE TASK OF ANTI-THEFT DETECTION

Mould	Accuracy	Accuracy	Recall rate	F1-score	Auc-roc
Hybrid model	0.85	0.87	0.82	0.84	0.92
Statistical testing methods	0.78	0.75	0.80	0.77	0.88
Svm	0.82	0.84	0.78	0.81	0.90
Random forest	0.86	0.88	0.84	0.86	0.93
LSTM	0.79	0.80	0.76	0.78	0.89
Autoencoder	0.81	0.82	0.79	0.80	0.91

TABLE III. PERFORMANCE BREAKDOWN OF EACH MODEL ON POSITIVE AND NEGATIVE CLASS DETECTION

Mould	Logarithmic accuracy	Positive recall	Negative category precision rate	Negative class recall
Hybrid model	0.88	0.80	0.84	0.90
Statistical testing methods	0.82	0.76	0.72	0.85
Svm	0.85	0.79	0.78	0.88
Random forest	0.89	0.82	0.86	0.92
LSTM	0.80	0.74	0.75	0.82
Autoencoder	0.83	0.78	0.77	0.86

C. Real Life Cases

This paper conducted a three-month field evaluation of the effectiveness of a hybrid model based on the long short-term memory (LSTM) network and autoencoder in the real operating environment of a large urban electric utility. This chapter details the deployment of the model during this period, its operation, and its results in detecting anti-stealing behavior. The selected experimental area is home to a large number of residential and commercial customers, where electricity theft is frequent and causes considerable economic losses to the utility. To address this challenge, this paper work with the power company to integrate the hybrid model proposed in this paper into the real-time electricity consumption data monitoring system in the region, expecting to improve the effectiveness of electricity theft detection through advanced intelligent algorithms. This paper have collected and pre-processed electricity consumption data in a comprehensive and detailed way, covering hourly electricity consumption, voltage and current data, and applied them to the training set, validation set and test set, respectively. Compared with the proportion of data set allocation in the original system,

the hybrid model increases the amount of data in the training set, aiming to enhance the learning effect of the model. The details are shown in Table IV.

After the model is deployed, it enters the real-time operation phase, where the generated user electricity consumption data is received and processed in real-time on a daily basis. The hybrid model first digs into the time-series patterns of user behavior with the help of the LSTM module, and then the self-encoder partially screens the normal and abnormal power consumption states. In addition, the model adopts a flexible dynamic threshold strategy in the operation process to adapt to the changes of different seasons, time periods, and various types of users' electricity consumption patterns, so as to accurately identify potential electricity theft behaviors. The details are shown in Table V.

After three months of rigorous practical deployment and comprehensive evaluation, the hybrid model integrating the long short-term memory network (LSTM) and the self-encoder shows excellent performance in the detection of electricity theft, clearly surpassing the original detection system and achieving significant advantages in several core evaluation indexes. The specific performance is as follows:

In the experimental phase, the hybrid model successfully identified a series of potential power theft incidents with its excellent analytical capabilities and accurate anomaly detection algorithms. One typical case of a commercial user was particularly notable. The business should have maintained a relatively stable power consumption profile during its regular operation, however, an unusual peak in power consumption suddenly appeared during a specific period of time. With a deep understanding of the user's historical power usage habits and a keen insight into current changes in power usage behavior, the hybrid model arrived at a high suspected power theft risk score by calculating the deviation of the user's power usage pattern from its long-term normal behavior. This prediction drew great attention from the power company, and accordingly an on-site verification was carried out, which eventually confirmed that there was indeed a case of power theft by bypassing the meter through illegal means.

Through the in-depth implementation and rigorous evaluation of the above and other practical cases, the practical value of the hybrid model constructed based on LSTM and self-encoder in power system anti-stealing actions has been fully proved. The model not only significantly improves the accuracy and sensitivity of the detection of power theft, significantly reduces the probability of false alarms, but also, more importantly, immediately shows substantial improvement in the economic effect. By timely detecting and stopping power theft, the model helps the power company to protect its own interests and reduce economic losses, and at the same time, it also contributes to the maintenance of a fair and equitable power supply environment and the protection of the normal power market order.

As shown in Table VI, in comparison to traditional machine learning models like Logistic Regression, Decision Trees, and Naive Bayes, the hybrid LSTM-autoencoder model demonstrates superior performance across all evaluation metrics. Its higher accuracy, precision, recall, F1-score, and AUC-ROC values illustrate the model's enhanced capability in capturing complex temporal dependencies and effectively distinguishing between normal and abnormal power consumption patterns, thereby validating the superiority of deep learning techniques in this domain.

TABLE IV. DATA COLLECTION AND PREPROCESSING

Data type	Collection of content	Treatment	Data set allocation (former system)	Dataset allocation (hybrid LSTM and self-encoder model)
Electricity consumption	Hourly records	Cleaning, standardization	Training set 60%	Training set 70%
Input voltage	Hourly records	Cleaning, standardization	Validation set 20%	Validation set 15%
Amps	Hourly records	Cleaning, standardization	Test set 20%	Test set 15%

TABLE V. MODEL DEPLOYMENT AND OPERATION

Portion	Functionality	Operating strategy	Original system	LSTM and self-encoder hybrid modeling
LSTM module	Capturing time series patterns	Real-time data processing	Inapplicable	Real-time data processing
Autoencoder	Identify normal and abnormal states	Dynamic threshold adjustment	Inapplicable	Dynamic threshold adjustment

TABLE VI. COMPARISON WITH TRADITIONAL MACHINE LEARNING MODELS

Model Type	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	0.72	0.70	0.74	0.72	0.86
Decision Tree	0.76	0.73	0.78	0.75	0.87
Naive Bayes	0.74	0.71	0.77	0.74	0.85
Hybrid LSTM-Autoencoder	0.85	0.84	0.82	0.83	0.92

TABLE VII. COMPARISON WITH OTHER DEEP LEARNING MODELS

Model Type	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Gated Recurrent Unit (GRU)	0.82	0.81	0.80	0.80	0.90
Convolutional Neural Network (CNN)	0.83	0.82	0.81	0.81	0.89
Simple LSTM	0.80	0.79	0.79	0.79	0.88
Hybrid LSTM-Autoencoder	0.85	0.84	0.82	0.83	0.92

As shown in Table VI, when compared against other deep learning models including GRU, CNN, and a simpler LSTM variant, the hybrid LSTM-autoencoder model retains its lead. It achieves the highest scores across accuracy, precision, recall, F1-score, and AUC-ROC, emphasizing the efficacy of combining LSTM's sequential pattern learning with the feature extraction and noise reduction capabilities of the autoencoder. This synergy allows for a more nuanced understanding of the data, enabling the model to detect subtle anomalies indicative of electricity theft with improved accuracy and robustness.

The comparative analyses presented in Tables VI and VII underscore the innovative advantage of the proposed hybrid LSTM-autoencoder model in the realm of anti-theft behavior identification. It not only surpasses traditional machine learning models in performance but also outperforms other deep learning architectures specifically tailored for time-series anomaly detection. The hybrid model's capacity to integrate temporal sequence analysis with efficient feature representation learning sets a new benchmark in the field, enhancing the precision and efficiency of power theft detection systems.

In the face of the increasingly serious problem of electricity theft in power systems, how to design and implement an efficient and accurate electricity theft detection mechanism has become a key research problem to be solved urgently. Especially in the context of massive power metering data, how to effectively extract and optimize key features, reduce data redundancy, and improve the generalization ability and interpretability of the model are the core challenges to improve the identification technology of electricity theft. The hybrid model based on Long Short Term Memory Network (LSTM) and self-encoder proposed in this paper shows excellent performance in power theft detection. Compared with traditional statistical detection methods, support vector machines, random forests and models using LSTM or auto-encoder alone, the hybrid model significantly improves key evaluation indicators such as accuracy, recall, F1 score and AUC-ROC, which proves that the model is efficient in comprehensive understanding and learning normal and abnormal electricity consumption patterns. After three months of field application evaluation, the hybrid model can effectively identify and monitor electricity theft behavior in the actual operation environment of large urban power companies, and significantly reduce economic losses, indicating that the model has strong practicability and economic value. The dynamic threshold adjustment strategy of the model adapts to the change of electricity consumption mode in different seasons, periods and user types, and improves the flexibility and accuracy of detection. The experimental results of this paper accord with the original purpose of this paper. The work of this paper not only improves the accuracy and sensitivity of electricity theft detection, reduces the probability of false alarm, but more importantly, protects the economic interests of power companies and maintains the fair order of power supply market by detecting and preventing electricity theft behavior in time. It has positive significance to promote the rational allocation of resources and the stable development of social economy.

VI. CONCLUSION

In this study, a new feature extraction and optimization scheme is innovatively proposed for the increasingly serious anti-power theft problem in the power system, based on the exhaustive research background and practical needs. During the research process, this paper applied the recursive feature elimination (rfe) technique to deeply screen and optimize the power system data, and at the same time, combined with feature validation, correlation and exclusion analysis, this paper effectively identified and selected key feature indicators reflecting anti-electricity theft behaviors. On this basis, this paper constructed a unique hybrid model integrating long short-term memory network (LSTM) and autoencoder, which is specially designed for identifying anti-power theft behaviors in the power system with enhancement, fully reflecting the advantages of LSTM in capturing time-series characteristics and the ability of autoencoder in efficient feature learning and characterization compression. Through rigorous simulation experiments and practical application exploration, the method of this study shows significant advantages in the accuracy and efficiency of anti-power theft recognition, and can locate and judge power theft more accurately and quickly compared with the classical method, and the results of both sets of experiments confirm the effectiveness of model. In this study, this paper have successfully developed an anti-stealing recognition technique based on rfe and hybrid LSTM-autoencoder model, which provides a new technical support for anti-stealing management in the power industry. Innovation points:

1) For the first time, rfe and feature correlation and exclusion analysis are applied to the selection of anti-theft features in power systems, which improves the quality and efficiency of feature extraction.

2) A hybrid model incorporating LSTM and autoencoder is designed to cope with the complex timing characteristics and high-dimensional data problems in the recognition of anti-theft behaviors.

3) Through simulation experiments and practical explorations, the excellent performance of the new method in the identification of anti-theft behaviors is verified, showing its feasibility in practical applications.

Deficiencies:

1) The computational efficiency of the current model in dealing with large-scale, heterogeneous power system data needs to be further improved.

2) For some complex and highly hidden power theft behaviors, there is still room for improvement in the model's recognition sensitivity and generalization ability.

3) The generalizability of the model to different types of power network structures has not been fully tested.

REFERENCES

[1] M. J. Abdulaal, M. Mahmoud, S. A. Bello, J. Khalid, A. J. Aljohani, M. A. H. Milyani, et al. "Privacy-preserving detection of power theft in smart grid: Advanced metering infrastructure." IEEE Access, vol. 11, pp. 68569-68587, Feburary 2023.

[2] M. Adil, N. Javaid, U. Qasim, I. Ullah, M. Shafiq, J. G. Choi. "LSTM and BAT-based RUSBoost approach for electricity theft detection." Appl. Sci.-Basel, vol. 10, no. 12, pp. 4378, December 2020.

[3] C. A. Adongo, F. Taale, S. Bukari, S. Suleman, I. Amadu. "Electricity theft whistleblowing feasibility in commercial accommodation facilities." Energy Policy, vol. 155, pp. 112347, May 2021.

[4] S. Ali, Y. Z. Min, W. Ali. "Prevention and detection of electricity theft of distribution network." Sustainability, vol. 15, no. 6, pp. 4868, June 2023.

[5] K. V. Blazakis, T. N. Kapetanakis, G. S. Stavrakakis. "Effective electricity theft detection in power distribution grids using an adaptive neuro-fuzzy inference system." Energies, vol. 13, no. 12, pp. 3110, December 2020.

[6] J. D. Chen, Y. A. Nanehkaran, W. R. Chen, Y. J. Liu, D. F. Zhang. "Data-driven intelligent method for detection of electricity theft." Int. J. Electr. Power Energy Syst., vol. 148, pp. 108948, March 2023.

[7] S. L. Dong, Z. X. Zeng, Y. N. Liu. "FPETD: Fault-tolerant and privacy-preserving electricity theft detection." Wirel. Commun. Mob. Comput., pp. 1-11, November 2021.

[8] A. T. El-toukhy, M. M. Badr, M. Mahmoud, G. Srivastava, M. M. Fouda, M. Alsabaan. "Electricity theft detection using deep reinforcement learning in smart power grids." IEEE Access, vol. 11, pp. 59558-59574, April 2023.

[9] I. Fatema, G. Lei, X. Y. Kong. "Probabilistic forecasting of electricity demand incorporating mobility data." Appl. Sci.-Basel., vol. 13, no. 11, pp. 6520, November 2023.

[10] C. Genes, I. Esnaola, S. M. Perlaza, L. F. Ochoa, D. Coca. "Robust recovery of missing data in electricity distribution systems." IEEE. Trans. Smart. Grid., vol. 10, no. 4, pp. 4057-4067, July 2019.

[11] A. K. Gupta, A. Routray, V. N. A. Naikan. "Detection of power theft in low voltage distribution systems: A review from the Indian perspective." IETE. J. Res., vol. 68, no. 6, pp. 4180-4197, June 2022.

[12] L. Hirth. "Open data for electricity modeling: Legal aspects." Energy Strat. Rev., vol. 27, pp. 100433, October 2020.

[13] Y. F. Huang, Q. F. Xu. "Electricity theft detection based on stacked sparse denoising autoencoder." Int. J. Electr. Power Energy Syst., vol. 125, pp. 106448, June 2021.

[14] K. Ishizu, T. Mizumoto, H. Yamaguchi, T. Higashino. "Home activity pattern estimation using aggregated electricity consumption data." Sens. Mater., vol. 33, no. 1, pp. 69-88, January 2021.

[15] F. Jamil. "Electricity theft among residential consumers in Rawalpindi and Islamabad." Energy Policy, vol. 123, pp. 147-154, September 2018.

[16] F. Jamil, E. Ahmad. "Policy considerations for limiting electricity theft in the developing countries." Energy Policy, vol. 129, pp. 452-458, July 2019.

[17] M. Kezunovic, P. Pinson, Z. Obradovic, S. Grijalva, T. Hong, R. Bessa. "Big data analytics for future electricity grids." Electr. Power Syst. Res., vol. 189, pp. 106788, December 2020.

[18] I. U. Khan, N. Javaid, C. J. Taylor, X. D. Ma. "Robust data driven analysis for electricity theft attack-resilient power grid." IEEE Trans. Power Syst., vol. 38, no. 1, pp. 537-548, January 2023.

[19] G. Y. Lin, H. Y. Feng, X. F. Feng, H. W. Wen, Y. Z. Li, S. Y. Hong, et al. "Electricity theft detection in power consumption data based on adaptive tuning recurrent neural network." Front. Energy Res., vol. 9, pp. 773805, October 2021.

[20] S. X. Liu, Y. Liang, J. L. Wang, T. Jiang, W. S. Sun, Y. Rui. "Identification of stealing electricity based on big data analysis." Energy Rep., vol. 6, pp. 731-738, August 2020.

[21] X. Liu, Y. Ding, H. Tang, F. Xiao. "A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data." Energy Build., vol. 231, pp. 110601, December 2021.

[22] A. Lotfipoor, S. Patidar, D. P. Jenkins. "Transformer network for data imputation in electricity demand data." Energy Build., vol. 300, pp. 113675, June 2023.

[23] E. Lu, N. Wang, W. Zheng, X. D. Wang, X. Y. Lei, Z. C. Zhu, et al. "Data-driven electricity price risk assessment for spot market." Int. Trans. Electr. Energy Syst., 2022.

[24] G. Mangat, D. Divya, V. Gupta, N. Sambyal. "Power theft detection using deep neural networks." Electr. Power Comp. Syst., vol. 49, no. 4-5, pp. 458-473, May-June 2021.

[25] A. Neale, M. Kummert, M. Bernier. "Discriminant analysis classification of residential electricity smart meter data." Energy Build., vol. 258, pp. 111823, December 2022.

[26] S. Pamir, N. Javaid, M. U. Javed, M. Abou Houran, A. M. Almasoud, M. Imran. "Electricity theft detection for energy optimization using deep learning models." Energy Sci. Eng., vol. 11, no. 10, pp. 3575-3596, October 2023.

[27] N. Rauschkolb, N. Limandibhratha, V. Modi, I. Mercadal. "Estimating electricity distribution costs using historical data." Util. Policy, vol. 73, pp. 101309, July 2021.

[28] I. S. Shah, F. H. Jan, S. Ali. "Functional data approach for short-term electricity demand forecasting." Math. Probl. Eng., 2022.

[29] F. Shehzad, N. Javaid, S. Aslam, M. U. Javed. "Electricity theft detection using big data and genetic algorithm in electric power systems." Electr. Power Syst. Res., vol. 209, pp. 107975, November 2022.

[30] M. Tariq, H. V. Poor. "Electricity theft detection and localization in grid-tied microgrids." IEEE Trans. Smart Grid, vol. 9, no. 3, pp. 1920-1929, May 2018.

[31] A. Tureczek, P. S. Nielsen, H. Madsen. "Electricity consumption clustering using smart meter data." Energies, vol. 11, no. 4, pp. 859, April 2018.

[32] E. Ul Haq, J. J. Huang, H. R. Xu, K. Li, F. Ahmad. "A hybrid approach based on deep learning and support vector machine for the detection of electricity theft in power grids." Energy Rep., vol. 7, pp. 349-356, May 2021.

[33] B. M. Wabukala, N. Mukisa, S. Watundu, O. Bergland, N. Rudaheranwa, M. S. Adaramola. "Impact of household electricity theft and unaffordability on electricity security: A case of UGANda." Energy Policy, vol. 173, pp. 113411, March 2023.

[34] B. C. Wang, X. Y. Zhai, X. L. Wei, Y. P. Shi, X. Q. Huo, R. N. Li, et al. "A self-powered and concealed sensor based on triboelectric nanogenerators for cultural-relic anti-theft systems." Nano. Res., vol. 15, no. 9, pp. 8435-8441, September 2022.

[35] B. H. Wang, Q. L. Guo, Y. Yu. "Mechanism design for data sharing: An electricity retail perspective." Appl. Energy, vol. 314, pp. 118871, May 2022.