**Electricity Theft Detection Strategy**

**– Powered by KAIFA Empower DataPro**

***Empower DataPro for digital transformation and smart energy management***

10

Version 1.0

**Historical Reference**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Version/Status** | **Author** | **Participant** | **Release Date** | **Remark** |
| Rev 1.0 | KAIFA |  | 2025/06/13 | final version |
|  |  |  |  |  |

# List of Contents

[Historical Reference 2](#_Toc200709635)

[List of Contents 3](#_Toc200709636)

[1. Introduction 5](#_Toc200709637)

[1.1 Background 5](#_Toc200709638)

[1.2 Objectives 6](#_Toc200709639)

[2. Overall Framework of the Action Plan 6](#_Toc200709640)

[2.1 Phased Implementation Roadmap 6](#_Toc200709641)

[2.2 Key Technical Supports 7](#_Toc200709642)

[3. Detailed Implementation Roadmap and Business Work Division 7](#_Toc200709643)

[3.1 Data Integration and Preparation Phase 8](#_Toc200709647)

[3.2 Big Data Analytics and Model Building Phase 9](#_Toc200709648)

[3.3 Inspection Execution and Field Verification Phase 11](#_Toc200709649)

[3.4 Evaluation and Continuous Optimization Phase 12](#_Toc200709650)

# 

# Introduction

This proposal presents a comprehensive, systematic, and highly feasible action plan to combat electricity theft within power utilities. Given the pressing issues of high line losses and rampant theft in power systems, the plan fully leverages the advanced data analytics and artificial intelligence capabilities provided by KAIFA Empower DataPro to develop a full-cycle solution encompassing data collection, intelligent analysis, targeted inspections, and continuous optimization.

By establishing and continuously updating an expert rule base for abnormal electricity usage analysis, along with the development and refinement of machine learning models, the plan aims to automate and accurately detect electricity theft leads. This will significantly enhance the overall efficiency and precision of inspection efforts. The proposal outlines a phased implementation roadmap, clearly defined business responsibilities, and the pivotal role of big data analytics. Successful execution of this plan is expected to greatly reduce power line losses, recover significant economic losses, reinforce fairness and transparency in the electricity market, and advance the modernization of national energy management, thereby ensuring a more reliable power supply for Uzbekistan’s socio-economic development.

## Background

Electricity is a core driver of modern socio-economic development, and its strategic importance is indisputable. However, electricity theft—a serious illegal activity—directly results in substantial financial losses for utilities, reduced tax revenues, and undermines market fairness, grid stability, and the rights of compliant electricity users.

Traditional methods of theft inspection have shown clear limitations:

* **Manual inspections and random checks** are inefficient and limited in scope, unable to address the vast amount of user data and the growing sophistication of theft techniques.
* **Delayed detection** means theft is usually discovered only after it has caused ongoing losses, severely limiting preventive measures.
* **Experience-dependence** restricts standardization and large-scale replication, as inspections rely heavily on individual expertise and intuition.

While Uzbekistan’s power system has accumulated substantial operational data through the deployment of Advanced Metering Infrastructure (AMI) in recent years, these data remain fragmented across systems (e.g., CAS, CRM, billing), forming "data silos" that hinder cross-system, cross-functional analysis. This fragmentation limits the full value of the data and highlights the urgent need for a technologically advanced, innovative solution to tackle electricity theft and enhance decision-making and operational management.

*KAIFA Empower DataPro*, a big data and AI platform tailored for the digital transformation of the power sector, offers robust capabilities for integrating heterogeneous data sources, large-scale storage, parallel computing, machine learning, and large language models (LLMs). This provides a strong technical foundation for intelligent theft detection.

## Objectives

The core objectives of this Anti-Theft Action Plan are as follows:

1. **Reduction of Line Loss Rates**: By accurately identifying and curbing electricity theft, line losses are expected to drop significantly—potentially saving millions of USD annually.
2. **Optimization of Inspection Efficiency and Accuracy**: Big data analytics and AI will enable automated and intelligent detection of suspicious activities, reducing the average response time from seven days to less than two, thus boosting efficiency by over 200%.
3. **Development of Intelligent Analysis Capabilities**: Establish and continuously improve an expert rule base and machine learning models tailored to Uzbekistan’s specific context for intelligent theft detection.
4. **Standardization of Business Processes**: Streamline and optimize the entire inspection workflow—from anomaly detection and task dispatching to on-site verification and feedback—to achieve closed-loop management.
5. **Unlocking Data Value**: Break down data silos, integrate multi-source data, and fully exploit its value in inspection, management, and strategic decision-making.
6. **Maintaining Market Fairness**: Effectively combat electricity theft to uphold fair electricity access and market integrity for all users.

# Overall Framework of the Action Plan

This plan is built upon the KAIFA Empower DataPro platform and is centered around data-driven, intelligent analytics for theft detection. The overall framework includes the following phases:

## Phased Implementation Roadmap

The plan will follow a four-stage approach to ensure systematic and stable deployment:

**Phase I: Data Integration & Preparation (Months 1–3)**

* **Objective**: Collect, clean, standardize, and store multi-source heterogeneous data to build a centralized data warehouse.
* **Key Task**: Eliminate data silos and lay a solid foundation for subsequent analytics.

**Phase II: Big Data Analysis & Model Development (Months 3–9)**

* **Objective**: Establish an expert rule base for abnormal usage, develop and deploy theft detection models, and automate anomaly identification.
* **Key Task**: Leverage DataPro’s AI computing to convert expert knowledge and historical data into intelligent insights.

**Phase III: Field Inspections & On-Site Validation (From Month 6 Onward)**

* **Objective**: Convert analytics-based leads into actionable inspection work orders for on-site verification.
* **Key Task**: Ensure seamless coordination between online analysis and offline inspection to boost precision and efficiency.

**Phase IV: Performance Evaluation & Continuous Optimization (Ongoing)**

* **Objective**: Establish feedback mechanisms to evaluate model accuracy and iteratively optimize rules and models to adapt to evolving theft techniques and usage patterns.
* **Key Task**: Achieve dynamic adaptability through closed-loop system management.

## Key Technical Supports

The implementation of this plan will rely heavily on the following core technical capabilities of the KAIFA Empower DataPro platform:

* **Massive Data Storage and Processing**: Supports petabyte-scale data storage and large-scale parallel computing (batch and real-time), capable of processing AMI, customer, and geospatial data.
* **Multi-Source Data Integration**: Offers versatile tools and methods (ETL, API, file import) to efficiently extract data from HES, MDM, CRM, billing, GIS, and other systems.
* **Data Warehouse Construction**: Builds a unified data model aligned with IEC CIM standards (ODS, DWD, ADS layers), and provides robust metric and tag management to ensure data quality and reusability.
* **AI Computing Capabilities**: Incorporates traditional machine learning algorithms (e.g., K-means, Isolation Forest, Random Forest, XGBoost, SVM) and LLM-based intelligence for various anomaly detection models.
* **Application Management Center**: Includes prebuilt apps for line loss analysis, usage anomaly detection, and equipment diagnostics, with support for custom application development.

# ****Detailed Implementation Roadmap and Business Work Division****

This section elaborates on the specific tasks, business responsibilities, and expected outcomes across each project phase.



## Data Integration and Preparation Phase

**Objective**:  
To build a high-quality, multi-dimensional, and real-time updated data foundation for electricity theft inspection.

**Work Tasks**:

1. **Data Source Identification and Integration**:  
   Precisely identify all data sources closely related to electricity theft inspection, including but not limited to:
   * **Meter Data**: Real-time and historical voltage, current, power, and energy data, mainly from the AMI system.
   * **Electricity Usage Behavior**: Usage time, geographic location, appliance type, payment records, customer complaints, etc., typically from CRM, billing, and marketing systems.
   * **Historical Theft Case Data**: Previous case details, including theft methods, occurrence time, and involved users, mainly from inspection systems and archives.
   * **External Environmental Data**: Weather (temperature, humidity), holidays, GIS-based geographic info, and regional economic data.  
     Leverage **KAIFA DataPro’s** multi-source data integration capabilities to ingest the above through ETL tools, APIs, or file uploads.
2. **Data Preprocessing and Quality Management**:
   * **Data Cleaning**: Handle missing values (e.g., interpolation), outliers (e.g., 3σ rule, isolation forest), and duplicates to ensure data integrity and consistency.
   * **Format Transformation and Standardization**: Normalize formats across data sources for model training (e.g., unify time formats, apply One-Hot encoding).
   * **Data Warehouse Modeling**: Build a theft detection data warehouse following the ODS-DWD-ADS layered architecture, with dimension modeling by user, time, and device.
   * **Data Pipeline Development and Scheduling**: Create ETL pipelines and fine-grained scheduling strategies (incremental updates, batch, real-time streaming) to ensure data flows efficiently and reliably.

**Business Work Division**:

* **IT/Data Team**: Responsible for data ingestion, ETL, warehouse modeling, pipeline development, and data scheduling.
* **Metering Department**: Provide meter data interfaces and assist in data validation.
* **Customer Service/Marketing Department**: Supply user profiles, payment records, complaints, and usage behavior.
* **Inspection Department**: Contribute historical case data and assist in defining business rules for cleaning and anomaly handling.

**Expected Outcomes**:

* A unified, multi-source, structured and unstructured big data platform for theft inspection.
* Standardized data formats, models, and governance frameworks for subsequent intelligent analytics.

## Big Data Analytics and Model Building Phase

**Objective**:  
To establish intelligent and high-accuracy electricity theft detection capabilities, transitioning from manual to data-driven identification.

**Work Tasks**:

1. **Expert Rule Base for Abnormal Usage Detection**:
   * Collaborate with domain experts (inspectors, metering specialists, engineers) to identify key characteristics of electricity theft.
   * Define rules using structured formats (e.g., JSON) via the DataPro “Expert Rule Base” module, with centralized management and versioning.

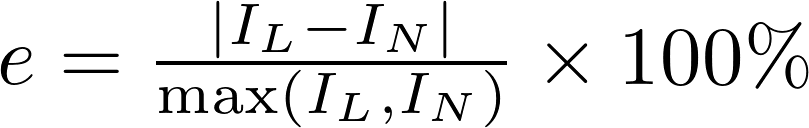
**Example Rules**:

* + **Current Bypass Model**: Triggered when current change rate and line loss exceed thresholds, with neutral/live current imbalance.

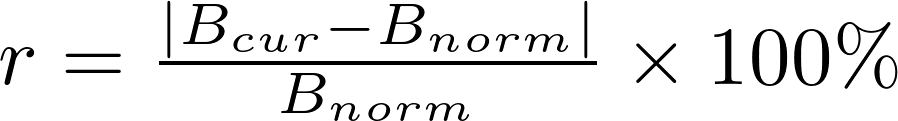
wpsoffice

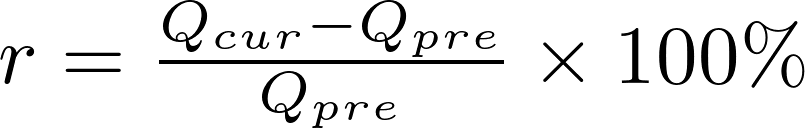
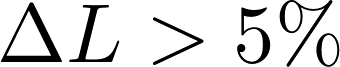
wpsoffice

* + **Voltage Loss Model**: Based on abnormal voltage deviation >20% and weak correlation between voltage and power.
  + **Neutral and Live Wire Current Imbalance Model:** Triggered when the current imbalance rate between the neutral and live wires continuously exceeds 15% for more than 30 minutes.



* + **Magnetic Interference Model**: Magnetic field rate of change exceeds 50%.



* + **Meter Tampering**: Abnormal meter\_open\_count or irregular open\_time\_distribution.
  + **Data Tampering**: Discrepancy between remote and local data surpasses set thresholds.
  + **Line Loss Correlation**: Assume the theoretical line loss rate in a given area is , and the actual line loss rate is , with the difference defined as .Define the user's electricity consumption variation rate as , where  is the current consumption and  is the consumption from the previous period. When  (where L is a configurable threshold based on actual conditions), and , the user is flagged as a suspected electricity theft case related to abnormal line loss.

1. **Model Marketplace & Usage Anomaly Modeling**:
   * Create a **Model Marketplace** in DataPro for managing all anomaly detection models, including rule-based, unsupervised, and supervised models.
   * **Model Types**:
     + **Rule-Based**: Automating expert rules.
     + **Unsupervised**: Detect unknown theft patterns using K-means, Isolation Forest, One-Class SVM, PCA+Mahalanobis, DBSCAN, etc.
     + **Supervised**: Train models like Random Forest, XGBoost, SVM, Logistic Regression, Neural Networks using labeled theft data.
   * **Unified Input/Output Interfaces**: Standardize model interfaces to facilitate integration and maintenance.
2. **Integration of Meter Device Fault Diagnosis**:
   * Theft often correlates with device anomalies. Integrate meter fault diagnostics into the detection system.
   * Develop a **Device Fault Rule Base** using domain knowledge (e.g., communication errors, incorrect phase sequence, missing voltage/current).
   * **Fault Models**:
     + Inputs: Power outages, phase loss, alarm frequency, etc.
     + Outputs: Probability, type, model name, version.
   * **Linkage Mechanism**: Automatically tag users with faulty meters as suspicious and correlate with usage anomaly models for more accurate theft detection.
3. **Model Iteration and Optimization Mechanism**:
   * Close-loop cycle: Experts create rules → Data team implements model → Model runs and produces leads → Inspectors verify on-site → Feedback used for model optimization.
   * Feedback accuracy feeds into performance scoring to boost future prediction precision, customized for the local context (e.g., Uzbekistan).

**Business Work Division**:

* **Data Science Team**: Responsible for digitizing rules, model development, training, deployment, and continuous performance tuning.
* **Domain Experts (Inspection/Metering)**: Provide rules, historical cases, on-site feedback, and support in model validation.
* **IT Team**: Manage AI infrastructure, model runtime environment, and tools.

**Expected Outcomes**:

* An intelligent model system covering various theft types and meter anomalies.
* Automated, accurate clue generation to drastically reduce manual analysis.
* A data-driven, model-supported decision-making framework.

## Inspection Execution and Field Verification Phase

**Objective**:  
Convert intelligent leads into efficient on-site inspections to ensure timely resolution of theft cases.

**Work Tasks**:

1. **Anomaly Report and Lead Display**:
   * System generates a **"User Abnormality List"** with:
     + User info (ID, address, contact).
     + Risk level (high/medium/low).
     + Associated model name/version.
     + Anomaly type (theft/device fault).
     + Key abnormal features (e.g., usage curve spike, neutral/live current imbalance).
   * Interactive interface allows inspectors to review detailed usage data (graphs, event logs) for remote pre-analysis.
2. **Inspection Work Order Generation & Dispatch**:
   * For high-risk users, the system enables one-click **work order generation**.
   * Work order includes: user profile, diagnosis results, recommended inspection focus, historical data.
   * Automatically dispatched to inspectors via mobile app or internal system.
3. **On-Site Verification and Handling**:
   * Field personnel inspect based on the work order, using tools and professional judgment.
   * Possible actions: manual reading, fee recovery, equipment replacement, evidence collection.
   * All results (e.g., confirmed theft, type, penalties) entered in real time via mobile terminal or inspection system.

**Business Work Division**:

* **Inspection Department**: Execute work orders, verify on-site, collect evidence, record outcomes.
* **IT/Data Team**: Maintain anomaly list, manage work order integration, provide system support.
* **Customer Service**: Handle user inquiries/complaints and provide user archives.

**Expected Outcomes**:

* Fast and precise field inspections based on intelligent leads.
* Significantly improved inspection efficiency and resource allocation.
* Digital, standardized management of field verification results.

## Evaluation and Continuous Optimization Phase

**Objective**:  
Establish a feedback loop to ensure continued improvement and adaptability of the inspection system.

**Work Tasks**:

1. **Closed-Loop Accuracy Evaluation**:
   * Compare field verification results (confirmed theft vs. false positives) with model predictions.
   * Evaluate model performance using metrics like accuracy, recall, F1-Score.
   * Assign positive scores to high-performing models, guiding future iterations.
2. **Periodic Reporting and Analysis**:
   * Generate **comprehensive reports** with:
     + Number of leads detected, thefts confirmed, economic loss recovered.
     + Model accuracy trends, theft type distribution, high-risk region analysis.
   * Use DataPro’s **BI Dashboard** for visualizing KPIs and supporting management decisions.
   * Hold regular review meetings to discuss insights, system gaps, and risks.
3. **Ongoing Optimization and Iteration**:
   * **Rule Adjustment**: Refine expert rules based on field feedback.
   * **Model Retraining**: Periodically retrain models with newly labeled data to capture evolving theft patterns.
   * **New Model R&D**: Actively develop models for emerging patterns or high-incidence zones.
   * **Business Process Optimization**: Improve workflows based on system effectiveness and feedback.

**Business Work Division**:

* **Management**: Review reports, provide strategic guidance, allocate resources.
* **Data Science Team**: Evaluate model performance, retrain models, generate analytical reports.
* **Inspection Department**: Deliver accurate feedback and participate in model improvement discussions.
* **IT Team**: Ensure system performance, stability, and data flow continuity.

**Expected Outcomes**:

* A full **“Detect–Verify–Feedback–Optimize”** closed-loop inspection management system.
* An evolving, self-learning inspection platform that adapts to changing theft methods.
* Continual improvements in inspection precision and recovery effectiveness.