Electricity Theft Detection Agent

فريق حماة الكهرباء

Digitopia Competition

Meet the Team: The Sentinels

Mahmoud Hamdy - Team Lead & AI Engineer

- **Contributions:** System architecture, research, and model training & validation.

Mohamed Hamdy - Grid Specialist

- **Contributions:** Electrical grid design, simulation, and topology configuration.

Youssef Romany - Data Engineer

- **Contributions:** Data generation and preprocessing pipeline.

Outline

- Problem Statement
- Challenges & Requirements
- Our Solution's Core Features
- Technical Solution
- Business Model
- Live Demo

The Problem in Numbers: A National Crisis

Legal Impact

513K+

Theft Cases Prosecuted

Economic Drain

1.2B

EGP in Fines 6 wks

Grid Pressure

12%+

YoY Peak Load

This isn't just data it's a national challenge demanding a smarter solution.

A National Priority

”لو أن نصف هذه السرقات لم تكن موجودة، لما كانت هناك مشاكل في إمدادات الكهرباء مرة أخرى.”

– دولة رئيس الوزراء مصطفى مدبولي، ٢٠٢٤

Challenge :1 The Infrastructure Reality

THE CHALLENGE

Egypt's electrical infrastructure is largely traditional, relying on transformers and underground cables.

THE SOLUTION MUST BE:

Specfiically designed to target this environment.

Challenge :2 Missing & Unreliable Network Data

THE CHALLENGE

Most grid parameters—like transformer voltages or underground cable lengths—are either hard to measure or often inaccurate/missing in practice.

THE SOLUTION MUST BE:

Able to operate effectively without relying on detailed grid measurements.

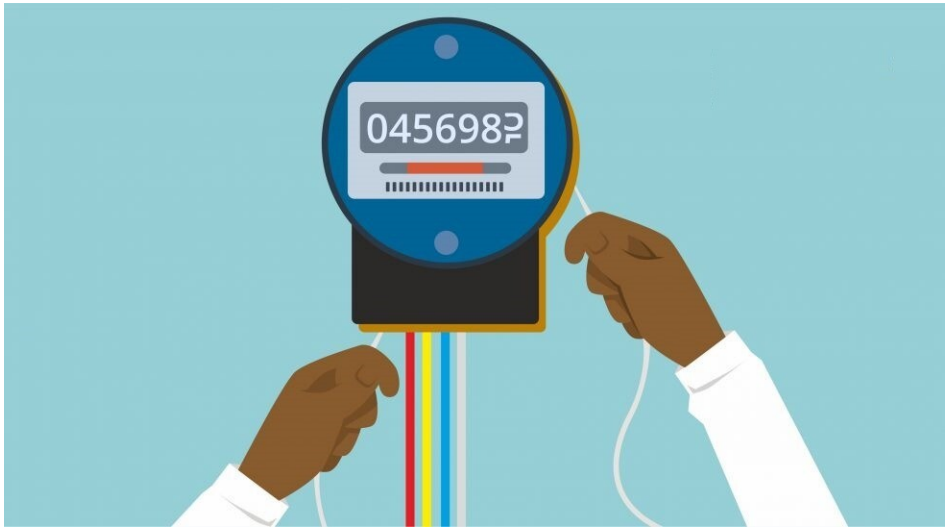
Challenge :3 Diverse Theft Methods

THE CHALLENGE

Electricity theft takes many forms: some manipulate meter electronics, while others bypass the meter completely using hidden direct connections.

THE SOLUTION MUST BE:

Capable of detecting both tampering and meter bypass.



Challenge :4 Legitimate vs. Illegitimate Behavior

THE CHALLENGE

A customer's consumption can spike for many legitimate reasons (e.g., new appliances, holidays). A naive system would flag these as theft, creating a flood of false alarms.

THE SOLUTION MUST BE:

Intelligent enough to distinguish between normal lfie events and actual theft.

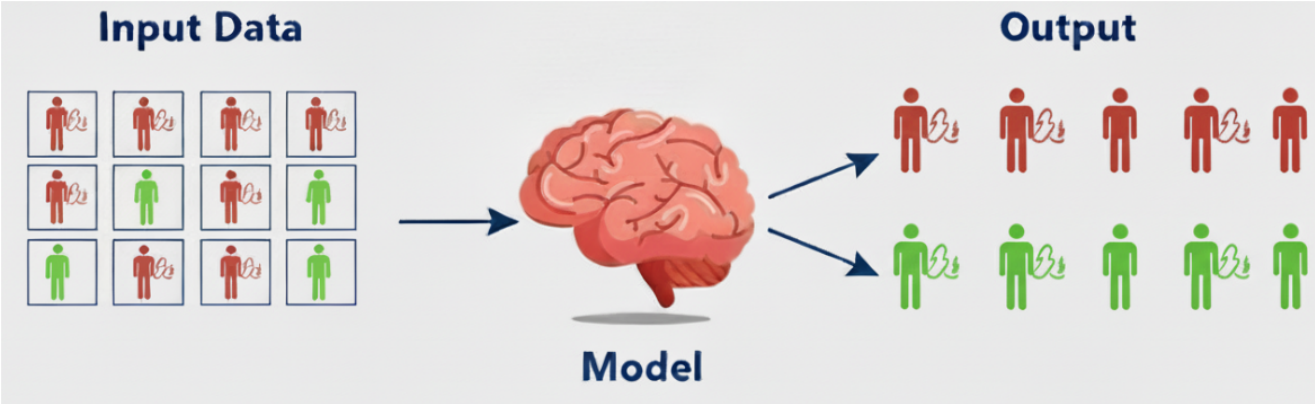
Challenge :5 The ”Ground Truth” Challenge

THE CHALLENGE

No ”ground truth” dataset exists that labels electricity readings as ’theft’ or ’legitimate’. This makes traditional supervised learning models impossible to train.

THE SOLUTION MUST BE:

Able to work in an unsupervised manner, detecting anomalies without needing pre-labeled data.



Challenge :6 The Cost & Feasibility Challenge

THE CHALLENGE

Any solution that requires expensive hardware upgrades to the grid is impractical for large-scale, nationwide deployment.

THE SOLUTION MUST BE:

A software solution that integrates with the existing infrastructure, requiring no physical modfiications.

Guess What ?

It is our solution

Our Solution's Core Features

Infrastructure-Ready

Grid-Agnostic

All-Method Theft Detection

Behavior-Aware

Unsupervised

Technical Outline: The Core Physical Principle

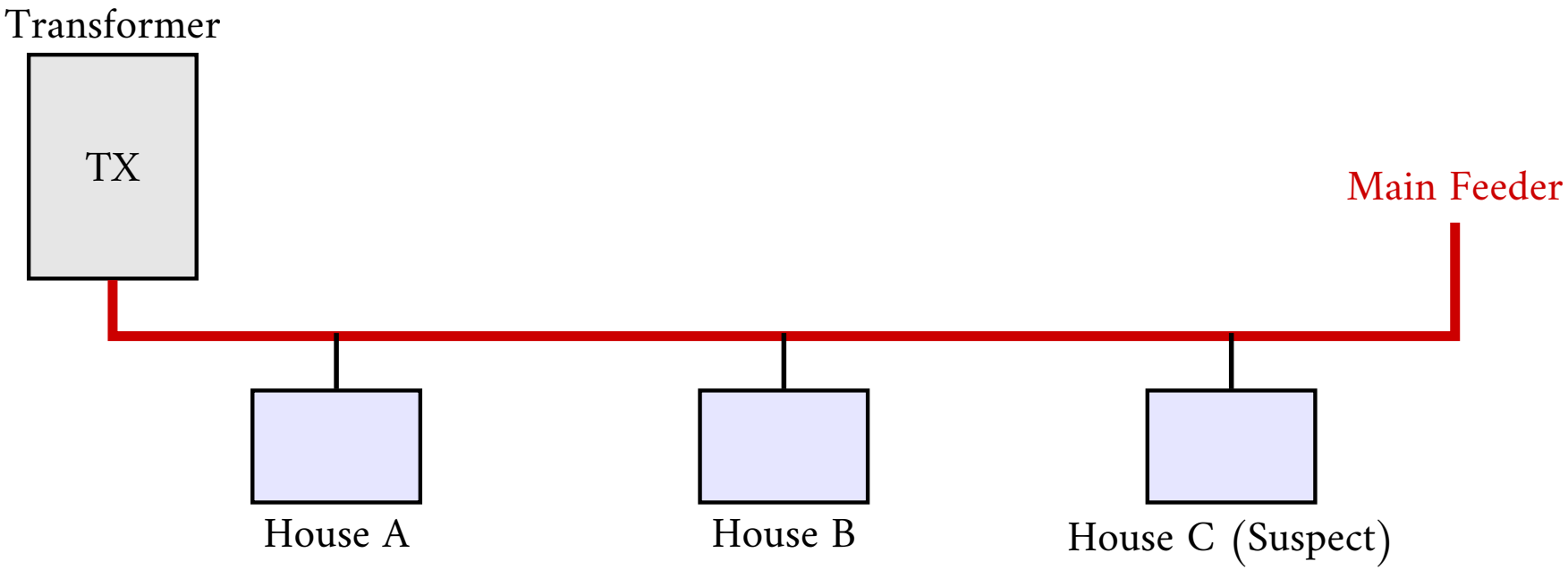


Figure: A simplfiied model of the low-voltage distribution grid.

The Core Physical Principle: Three Governing Equations

1. التأثير الذاتي

الجهد عندك = جهد المصدر - (استهلاكك × مقاومة السلك)

2. التأثير المشترك

الجهد عندك = الجهد عند نقطة التغذية - (استهلاكك × مقاومة سلك بيتك)

3. النموذج المتكامل

استهلاكك = (معامل أول × جهدك) + (معامل ثاني × مجموع استهلاك جيرانك) + ثابت

Model Application: Theft Detection Using Residuals

كيف نكتشف السرقة؟

نقارن الاستهلاك المتوقع من النموذج مع الاستهلاك الحقيقي المسجل

الفرق = الاستهلاك الحقيقي - الاستهلاك المتوقع

- إذا كان الفرق سالباً —-> احتمال سرقة كهرباء
- إذا كان الفرق موجباً —-> استهلاك طبيعي

لتطبيق هذا النموذج نحتاج إلى 3 معلومات أساسية :

1. قيمة الاستهلاك الفعلي (kWh) لكل بيت
2. قيمة الجهد الكهربائي (Voltage) لكل بيت
3. معرفة مجموعة البيوت المتصلة على نفس الفازة

Dataset Generation: Synthetic Power Grid Simulation

Phase :1 Virtual Network Construction

- **Tool:** pandapower Python library
- **Network:** 11kV/0.4kV transformer with 30 customers
- **Topology:** 3-phase LV distribution grid
- **Parameters:** Real Egyptian network specfiications

Dataset Generation: Synthetic Power Grid Simulation

Phase :2 Realistic Load Profile Generation

- **Duration:** 3 months of 15-minute interval data
- **Profiles:** Daily consumption patterns with peaks
- **Power Flow:** Voltage calculations using AC analysis
- **Output:** Consumption (kWh) + Voltage (V) per customer

Phase :3 Theft Cases Injection

- **Method:** Reduce reported consumption while keeping actual load
- **Patterns:** Various theft percentages ,20%) ,50% (80%
- **Result:** Ground truth labels for testing & evaluation

Model Implementation & Testing Flow

Step :1 Data Preprocessing & Feature Engineering

- **Data Loading:** final_dataset_with_theft.csv 3) months data)
- **Data Split:** Train (Jan-Feb clean data) + Test (March with theft)
- **Feature Engineering:** Created Total_Phase_Consumption_kw
- **Model Features:** Voltage_V + Total_Phase_Consumption_kw

Step :2 Unsupervised Learning & Prediction

- **Training:** Individual LinearRegression model per customer 30) models)
- **Learning:** (Voltage + Neighbors' Consumption) Expected Consumption
- **Prediction:** Model predicts consumption for each customer in test data
- **Key Point:** No theft labels used during training!

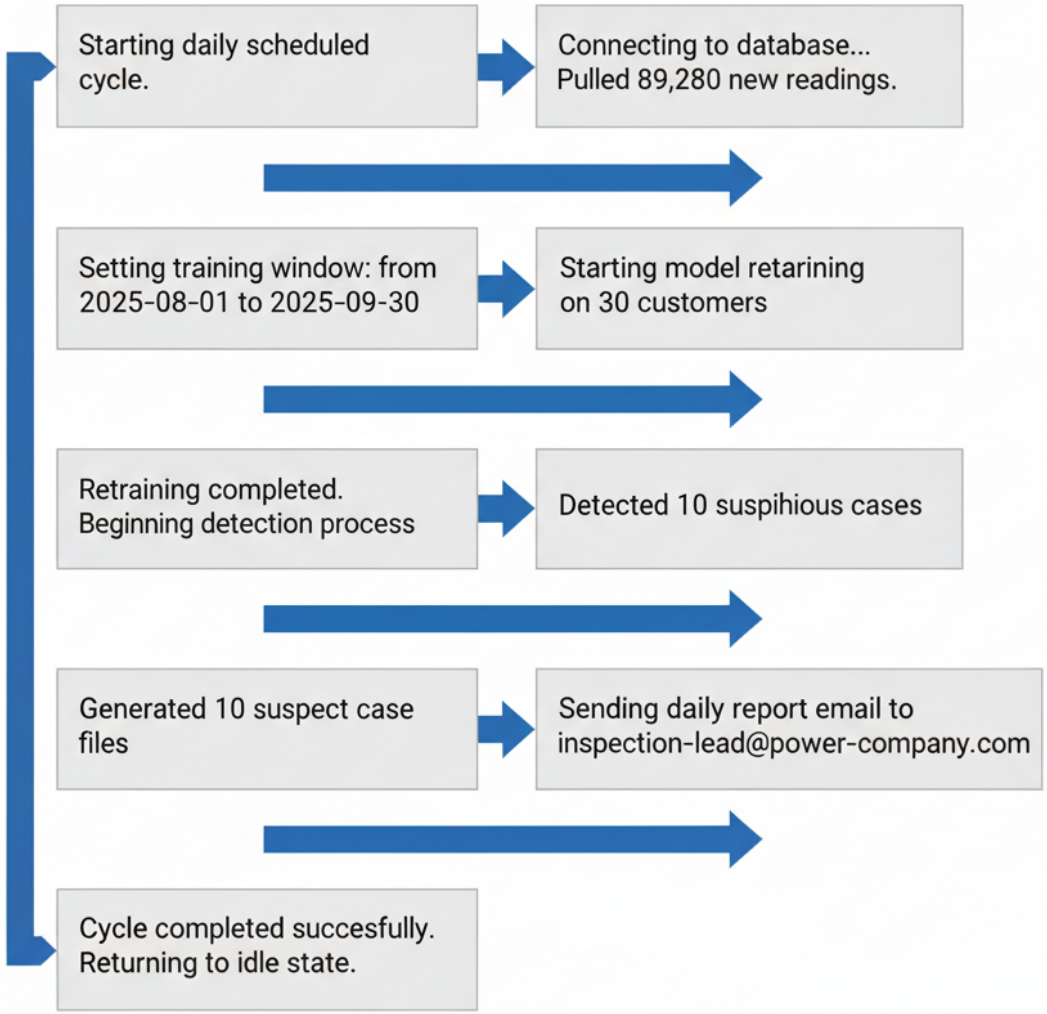
Residuals Analysis & Theft Detection Results

Residuals Analysis & Theft Detection

- **Residuals:** Actual Consumption - Predicted Consumption
- **Anomaly Score:** Sum of negative residuals per customer
- **Result:** Customer #5 ranked highest suspicious!
- **Validation:** Successfully detected injected thief
- **Success Rate:** Model correctly identfiied the sole thief without labels

Model correctly identfiied the sole thief : 100%

Real-World Deployment: Daily Agent Cycle



System Architecture: Agent Deployment

Business Model : Market & Financials

Customer Segments

- National & regional power utilities
- Private distribution companies
- Smart-grid service integrators

Channels & Customer Relations

- Direct B2B sales + pilot projects
- Annual SaaS contracts
- Dedicated customer-success team

Cost Structure

- R&D and algorithm development

Revenue Streams

- Annual SaaS licence
- X% share of recovered losses
- One-off setup fee
- Premium support packages

Business Model : Foundation & Operations

Key Partners

- Distribution utilities
- Smart-meter vendors
- Cloud & data-platform providers
- Field-service contractors

Key Resources

- Voltage–consumption ML engine
- Secure cloud infrastructure
- Utility integrations (API/OPC-UA)
- Data-science & support teams

Key Activities

- Daily data ingestion & model retraining
- Residual-based anomaly detection
- Alert generation & report delivery
- Ongoing model optimisation

Value Proposition

- 24/7 unsupervised theft detection
- Up to 5% yearly revenue recovery
- 80% cut in on-site inspection cost
- No new hardware—uses existing meters

Let's Watch the Demo

References

- Yu, N., Shah, S., Johnson, R., Sherick, R., & Hong, M. .(2019) *A Physically Inspired Data-driven Model for Electricity Theft Detection with Smart Meter Data*. IEEE Transactions on Industrial Informatics, ,(9)15 .5076-5088https://drive.google.com/file/d/17bBjldyOLbJ24QfwIiYE8cdIjfK29JSV/view?usp=drive_link

- Pandapower Development Team. .(2023) *pandapower - Convenient Power System Modelling and Analysis*. Version .2.13.1 https://www.pandapower.org/

- Scikit-learn Development Team. .(2023) *scikit-learn: Machine Learning in Python*. Version .1.3.0 Journal of Machine Learning Research, ,12 .2825-2830

Thank You!

Contact Information:
Email: matterm75@gmail.com

Protecting Power Grids with AI Innovation