🛡 Energy Shield AI

Electricity Theft & Anomaly Detection System

🔨 Project Overview

Energy Shield AI is an end-to-end, production-oriented system designed to **detect electricity anomalies and potential theft** using smart-meter data.
It combines **large-scale data engineering**, **unsupervised machine learning**, and **human-interpretable explainability** into a deployable solution.

The system is inspired by real-world utility challenges where **ground-truth theft labels are scarce**, making traditional supervised learning unreliable.

📋 Dataset Description

- **Source:** London Smart Meter Dataset (Kaggle)
- **Scale:**
  - ~3.47 million records
  - 71 engineered features
- **Granularity:**
  - Daily household consumption derived from **48 half-hour readings (`hh_0 – hh_47`)**

Key Feature Categories

1. **Consumption Statistics**

   - `daily_sum`, `daily_mean`, `daily_std`
   - `energy_min`, `energy_max`, `energy_median`
   - `peak_intensity`

2. **Behavioral Indicators**

   - `flatness_index`
   - `peer_ratio`

3. **Environmental Context**

   - `temperatureMin`, `temperatureMax`, `temp_gap`

4. **Socio-Economic Signal**

   - `Acorn_grouped`, `Acorn_encoded`

These features enable **behavioral profiling**, not just raw consumption analysis.

🧠 Modeling Strategy

Why Unsupervised Learning?

- Electricity theft is **rare and weakly labeled**
- Behavior changes over time
- Patterns differ by household type

➡ The problem is best framed as **anomaly detection**

🏅 Model Selection: Isolation Forest

Isolation Forest was chosen as the core model because:

- Designed specifically for anomaly detection
- Scales to **millions of rows**
- Works well with engineered tabular features
- Requires no labeled theft data
- Widely used in industry for fraud & energy analytics

Training Highlights

- Input: 15 carefully selected behavioral features
- Scaling: StandardScaler applied and persisted
- Output:
    - `-1` → Anomalous (suspicious behavior)
    - `+1` → Normal

⚙ Feature Engineering (Core Signals)

Feature	Meaning
Peer Ratio	Household usage vs similar socio-economic group
Flatness Index	Unnaturally constant consumption (tampering signal)
Peak Intensity	Strength of daily consumption peaks
Temp Gap	Sensitivity to environmental change

> Most detection power comes from **feature design**, not model complexity.

🔌 Backend Architecture (FastAPI)

The inference layer is built using **FastAPI**, enabling:

- Low-latency predictions
- Structured input validation (Pydantic)
- Production-grade deployment

API Flow

1. Receive household summary metrics

2. Reconstruct engineered features

3. Apply saved scaler

4. Run Isolation Forest inference

5. Return:

   - `status` (Normal / Suspicious)

   - `risk_score` (0–100)

   - `peer_ratio`

   - Explainable summary

🎨 Frontend Interface (Streamlit)

A clean, professional **Streamlit dashboard** provides:

UI Sections

1. **Input Simulation Panel**

   - Energy usage
   - Variability
   - Temperature
   - Socio-economic group

2. **Final Verdict**

   - Clear Normal / Suspicious status
   - Risk score gauge

3. **Key Risk Indicators**

   - Peer comparison
   - Weather sensitivity
   - Consumption consistency

4. **Explainable AI (XAI) Table**

   - Human-readable reasoning
   - Feature-level interpretation

The UI is designed for **utility analysts**, not ML engineers.

🔍 Explainability (XAI)

Rather than black-box scores, the system explains:

- *Why* a household was flagged
- *Which behaviors* triggered concern
- *How* it compares to similar households

This builds **trust**, which is essential in governance and billing decisions.

🏭 Real-World Applicability

Energy Shield AI can be used for:

- Electricity theft detection
- Meter tampering alerts
- Abnormal usage monitoring
- Revenue loss prevention
- Governance & compliance audits

It mirrors how **real power utilities** deploy AI systems.

🏁 Final Outcome

Energy Shield AI demonstrates:

- Scalable ML on real-world data
- Correct use of unsupervised learning
- Strong feature engineering
- Production-ready API
- Interpretable, analyst-friendly UI

> This is not just a model — it is a **complete, deployable AI system**.

🚀 Future Enhancements

- Household history timeline
- Batch CSV inspections
- SHAP-based explanations
- Cloud & Docker deployment
- Continuous learning pipeline

Author:
Energy Shield AI — Smart Meter Governance System