From Anomaly to Novelty: Active Detection and Adaptive Response in Smart Grids

Leann Alhashishi, KMA Solaiman

lalhash1@umbc.edu, ksolaima@umbc.edu

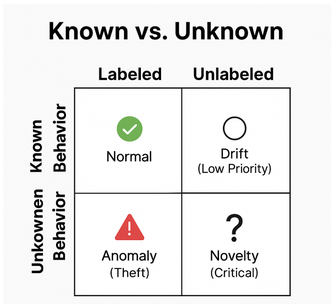
University of Maryland, Baltimore County

1. Why Detecting Novelty Matters

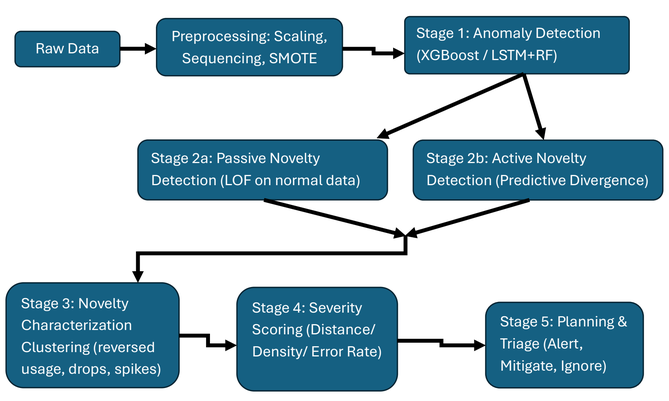
Cyber-physical systems must detect not just **known threats** (like electricity theft), but also **emerging unknown behaviors** — those that aren't labeled but may be disruptive.

Traditional anomaly detection stops at known risks. But what happens when the behavior is *unseen*, unlabeled, and structurally different?

Our Goal: Build a detection pipeline that separates anomaly from novelty — enabling **adaptive control and real-time triage** in electricity grids.



*Known vs. Unknown: Our model targets critical unknown, unlabeled behaviors — i.e., **novelty**.*



Overview of the Novelty-aware Smart Grid Pipeline

2. Anomaly Detection in Smart Grids

Dataset Overview

SGCC Smart Grid Dataset (2014–2016)

- **Users:** 1,037
- **Daily Samples:** 42,372
- **Label:** FLAG = 1 for electricity theft (8.6%)

Purpose: Detect known theft cases (FLAG=1) using supervised learning — to reduce false negatives.

Model 1: XGBoost

- Feature-based gradient boosting + SMOTE for class imbalance + Threshold tuning

Model 2: LSTM + Random Forest (RF)

- Input: 30-day consumption sequence
- LSTM encodes temporal behavior + RF predicts theft from LSTM embedding

3. Passive Novelty Detection

Purpose: Detect previously unseen consumption shifts not captured by labeled training data.

Model: Unsupervised detection using **Local Outlier Factor (LOF)**

- Trained on PCA-reduced features from FLAG = 0 (normal) users
- **Injected novel patterns** (e.g., spikes, sine) into normal and anomalous users to evaluate LOF sensitivity
- Detects structural outliers via local density

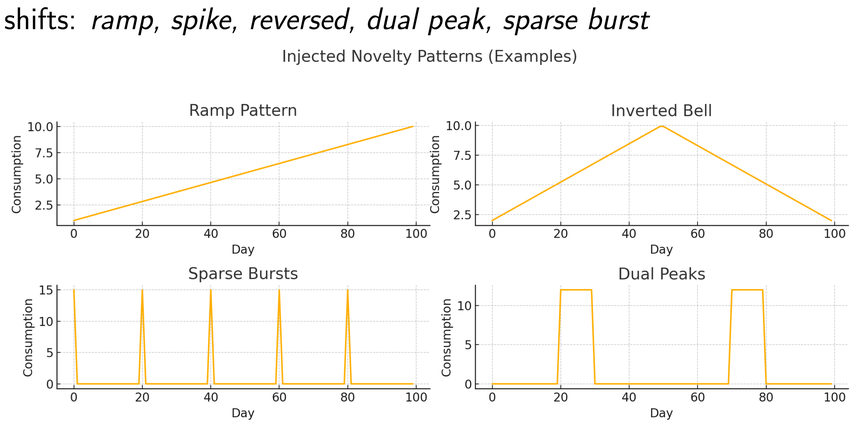
Model Type	Accuracy	Recall	Precision	F1 Score	AUROC
XGBoost (Anomaly)	80%	66%	39%	49%	81%
LSTM+RF (Anomaly)	91.4%	5%	44%	9%	74%
LOF (Novelty)	69.7%	88.9%	20%	32.7%	–

Takeaway:

- LSTM+RF achieves high accuracy but fails to detect most thefts (poor recall).
- LOF captures structural novelty with excellent recall — ideal for exploratory triage.

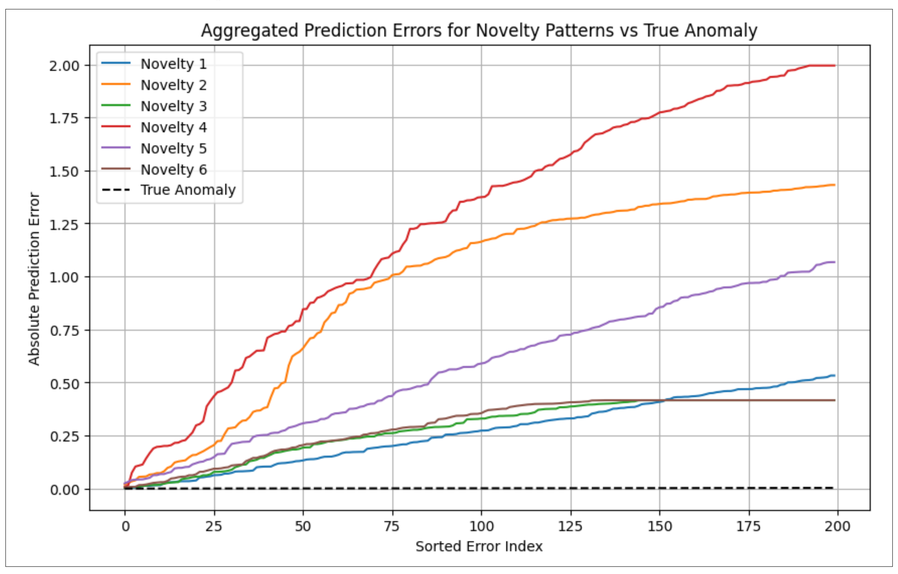
4. Active Novelty Detection via Predictive Divergence (**ELSTM**)

- ELSTM = LSTM + MLP model trained on FLAG=0 (normal) users
- **States**: 7-day consumption history with engineered features –
  - Holiday flag, day-of-week, 3-day rolling avg, change ratio
- Predict: [states] → next state (next-day consumption)
- Injected novelty patterns simulate plausible unseen consumption shifts: *ramp, spike, reversed, dual peak, sparse burst*



Injected patterns used to evaluate both LOF and ELSTM: representative of real-world novelty types.

- Prediction error (on consumption pattern of normal users, anomalies and novelties) used to quantify behavioral divergence
- Novelty patterns consistently caused higher error than FLAG=1 theft cases
- Evaluated MLP and Ridge models (non-temporal baselines); failed to detect novelty

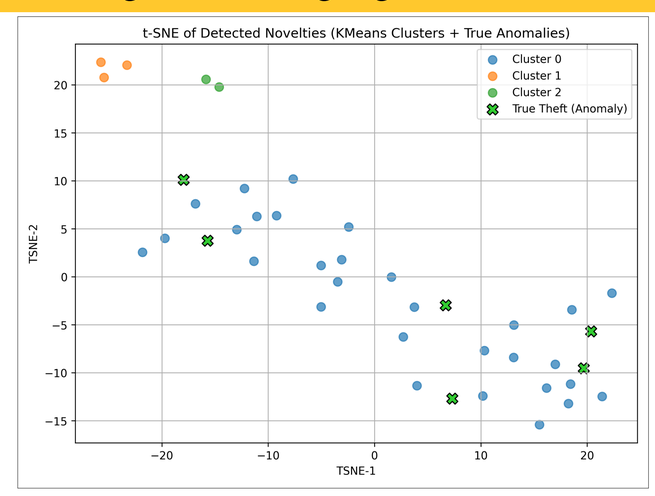


With ELSTM, **novelties show higher divergence than known anomalies for multiple users**.

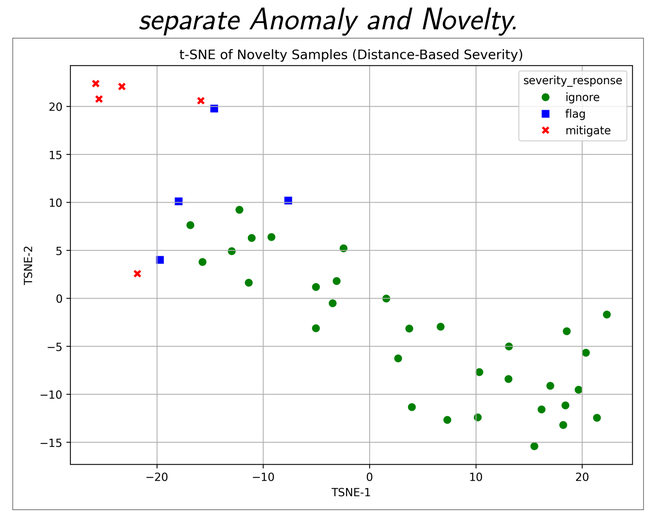
5. Novelty Characterization & Response

- **Clustering:** Applied KMeans to LOF-flagged novelties projected via PCA+t-SNE
- **Structure Discovery:** Three distinct clusters reflect varied novel behavior: burst, reversed, inverted bell
- **Label-Free Interpretation:** Clustering helps separate novelty types without any supervision
- **Planning Implication:** These groupings can guide mitigation or investigation actions
- **Severity Scoring:** Each novelty sample is assigned a severity tier based on structural divergence
- **Severity tiers:**
  Top 10% = High , 80–90% = Medium, Below 80% = Low
- Threshold-based triage system for aligning with control actions:
  High severity → Trigger **Mitigation**
  Medium severity → **Flag** for manual review
  Low severity → **Ignore** as benign drift
- Enables integration with planning/control systems (e.g., energy routing, alerting)
- Severity scoring shows clear structural separation — enabling actionable triage.

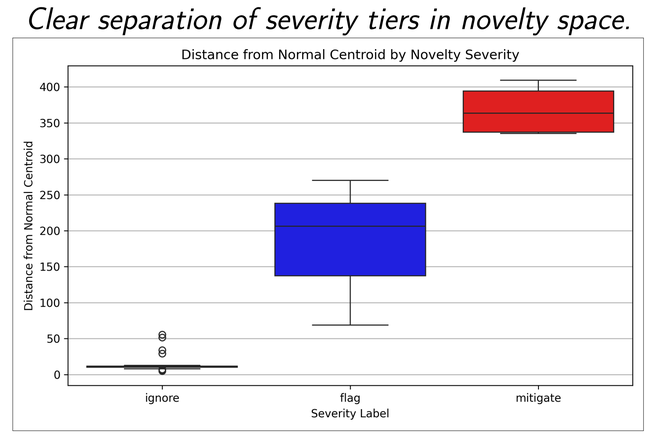
Control Triage & Planning Signal



LOF-flagged novelties grouped by KMeans. LOF struggles to separate Anomaly and Novelty.



Clear separation of severity tiers in novelty space.



Severity tiers are statistically distinct with minimal overlap.

6. Takeaways & Future Work

Key Insights:

Novelty ≠ Anomaly

Unlabeled, unseen shifts require different strategies than known threats.

- Predictive error from **LSTM + engineered features** flags novelties without labels
- **LOF + clustering** uncover structure and behavioral causes without supervision

Triage & Planning:

- Severity scoring enables mitigation, review, or dismissal of novel patterns
- Detection connects to planning — a step toward adaptive, control-aware CPS

Next Steps:

- Integrate with decision-aware systems (e.g., routing where actions influence future behavior)
- Expand to include feedback and action-driven adaptation
- Apply to real-world novelty: concept drift, attacks, and mass-scale behavior shifts

Acknowledgment

This project was led and developed by Dr. KMA Solaiman, including the full pipeline design, active detection, and poster creation. Leann contributed substantially to anomaly detection, novelty modeling, and implementation under Dr. Solaiman's direction. We thank Pooja Guttal for earlier contributions to anomaly detection.

H.A.R.M.O.N.I. Lab

Human-Aligned, Resilient, Multimodal, Open-ended, Novelty-Informed Intelligence