

Executive Brief: Intelligent Forecasting & Anomaly Detection Pipeline

Author: Piyush Kaushik Bhattacharyya

Purpose

This project was developed as part of the **Logic Leap AI** interview challenge to showcase the ability to **design, implement, and productionize a data-driven decision-support system**. The platform integrates **time series forecasting, anomaly detection, and automated reporting** for industrial operations, empowering stakeholders to anticipate demand, optimize resource usage, and detect operational risks early.

System Architecture

1. Data Ingestion and Validation

- **Inputs:** daily operations (units_produced, power_kwh, downtime) + site metadata.
- Validates schema; auto-detects latest dataset for reliability.

2. Feature Engineering

- **Calendar:** day-of-week, week, month, weekend flag.
- **Rolling stats:** 3–28 day means & stds for production/energy baselines.
- **Metadata join:** region, capacity, categorical encodings.

3. Forecasting Engine

- **Models:** Seasonal Naive (baseline) + Gradient Boosting/XGBoost (per-site, per-metric).
- **Horizon:** 14-day forecasts with expanding-window backtesting.
- **Metrics:** MAE & MAPE vs. baseline.

4. Anomaly Detection

- **Method:** STL decomposition residuals + robust Z-score ($|z| \geq 3$).
- Alerts include site, date, metric, observed vs. expected, residual, score.

5. Orchestration & API

- **Pipeline automation:** ingestion -> features -> forecasting -> anomalies.
- **FastAPI endpoints:**
 - **/run** : Execute pipeline
 - **/download/all** : Full ZIP of outputs
 - **/download/{file}** : Forecasts / metrics / anomalies

Strategic Impact

- **Operational foresight:** 14-day production & power forecasts for planning.
- **Risk mitigation:** Early anomaly alerts reduce downtime & inefficiency.
- **Risk mitigation:** Early anomaly alerts reduce downtime & inefficiency.
- **Scalability:** Modular design