What problem would you like to solve using anomaly detection?

Anomaly Detection and RUL Prediction for Aircraft Engines

- Using NASA’s C-MAPSS Turbofan Dataset
- Demo for Honeywell Aircraft Monitoring

Problem Statement

- Aircraft engines degrade over time -> unexpected failures = catastrophic risk
- Need early anomaly detection and RUL estimation to:
  - Increase safety
  - Reduce maintenance costs
  - Enable predictive maintenance(vs reactive)

The Dataset (C-MAPSS)

- Simulated jet engine degradation(NASA, 2008).
- 100+ engines, run to failure.
- Each cycle includes:
  - 3 operational settings
  - 21 sensors(pressures, temperatures, fan speeds).
- Provided true RUL labels for test set.
- Challenge: signals are noisy, redundant, nonlinear.

Industry Approaches

- Rule-based threshold -> simple, miss gradual faults.
- Classical ML (regression, forests) -> supervised, need failure labels.
- Deep Learning(LSTM, Autoencoders) -> learn hidden temporal degradation
- Honeywell interest: move from thresholds -> data-driven models

What I did

- Data Exploration
  - Identified informative sensors
  - Dropped flat/noisy signals
  - Built RUL labels
- Methods implemented
  - Isolation Forest(unsupervised)
  - LSTM Autoencoder(unsupervised)
  - Gradient Boosting and LSTM for RUL(supervised)
- Deployment
  - Interactive Streamlit dashboard for demo

Isolation Forest(Baseline)

- Trained only on healthy windows
- Scored all cycles for anomaly likelihood
- EWMA smoothing + percentile threshold
- Result
  - Detected anomalies with ~20-30 cycles lead time
  - Some false alarms early in life

LSTM Autoencoder

- Learned to reconstruct healthy sensor sequences
- High reconstruction error -> anomaly
- Advantages:
  - Captures temporal drift better than IsolationForest.
  - Lower false alarm rate
- Result
  - Clear anomaly signal as engine nears failure
  - Early warning before RUL <= 30 cycles

RUL Prediction

- Supervised Regression
  - Gradient Boosting on engineered features
  - LSTM regressor on raw sequences
- Metric: MAE (cycles)
- Results(FD001):
  - GBM MAE ≈ 25 cycles
  - LSTM MAE ≈ 20 cycles
- Accurate enough to plan maintenance windows

Dashboard Demo

- Built in Streamlit
- Features
  - Select engine ID
  - Plot anomaly scores vs cycles
  - Show alerts + threshold crossing
  - Overlay RUL curve
- Extensible for integration into Honeywell systems

Benefits

- Safety: early warnings before failure
- Cost savings: avoid premature replacement and reduce unplanned downtime
- Scalability: unsupervised methods don’t require failure labels
- Transparency: explain anomalies by top contributing sensors.
- Demo-ready: interactive tool shows practical applicability

Next Steps

- Extend to other FD002-FD004 datasets(multiple conditions, fault modes).
- Integrate with real Honeywell sensor feeds.
- Deploy as a real-time dashboard with alert notifications.
- Explore explainable AI for regular/engineer trust.

Closing

“From noisy sensor data to actionable insights:
Early anomaly detection and RUL prediction for safer skies.”