Turbofan Engine RUL Prediction & Anomaly Detection Report

Project Objective

Develop a machine learning pipeline to:

- Predict Remaining Useful Life (RUL) of turbofan engines.

- Detect anomalies based on prediction errors.

Data & Features

- Dataset: NASA CMAPSS Turbofan Engine Degradation Dataset (FD001)

- Rolling features: 5-cycle rolling mean and std of sensor readings

- Final feature count: 47 (after adding rolling features and dropping constants)

Final Model

- Model: Random Forest Regressor

- R^2: 0.7431

- MAE: 23.59

- RMSE: 34.26

- Best features: Rolling mean & std of key sensors (e.g., sensor_4, sensor_11, sensor_9)

Anomaly Detection

- Method: IQR-based thresholding of prediction error

- Threshold: 35.68

- Total anomalies detected: 1306

Visual Insights

- Prediction Error over Cycles: See `cycle_anomaly_plot.png`

- Prediction Error by Unit Number: See `unit_anomaly_plot.png`

Exported Files

- `anomalies_detected.csv`: List of detected anomalies with unit, cycle, RUL, prediction, error

---

Next Steps

- Optional hyperparameter tuning

- Alternative anomaly thresholds (e.g. z-score)

- Deployment as a script or microservice