

Hybrid Predictive Maintenance System

Overview

This project implements a multi-stage degradation modeling and risk-aware predictive maintenance system using the NASA CMAPSS Turbofan Engine Degradation Dataset. Instead of traditional binary failure prediction, we model degradation as a 5-stage process and combine classification, regression, and risk scoring for more actionable maintenance alerts.

Features

- Multi-stage degradation modeling (5 distinct health stages)
- Unsupervised stage discovery using KMeans and Agglomerative Clustering
- Real-time health state classification with Random Forest, Logistic Regression, and SVC
- Time-to-next-stage regression predicts remaining time in current health state
- Dual risk scoring system combines probability and time estimates
- Optimized alert thresholds for maintenance decision support

Dataset

The NASA CMAPSS (Commercial Modular Aero-Propulsion System Simulation) dataset consists of run-to-failure data for turbofan engines operating under various conditions

The dataset includes:

- 4 subsets (FD001-FD004) with different operating conditions and fault modes
- 21 sensor measurements plus 3 operational setting features
- Train (full run-to-failure) and test (truncated) datasets with truth values

Dataset available at: [NASA Prognostics Data Repository](#)

Requirements

`numpy>=1.19.0` , `pandas>=1.0.0`

`scikit-learn>=0.24.0` , `matplotlib>=3.3.0`

`seaborn>=0.11.0` , `jupyter>=1.0.0`

Contributors

Tejaansh Sara - se23ucse174 : 35% - implementing all phases of code

Sejal - se23ucse209 : 27.5% - help in writing report and classifying data

Arushi K- se23ucse031 : 27.5% - code for regression and visualisation

Srivanth - se23ucse166 : 10% - eda and data analysis