Machine Failure Prediction

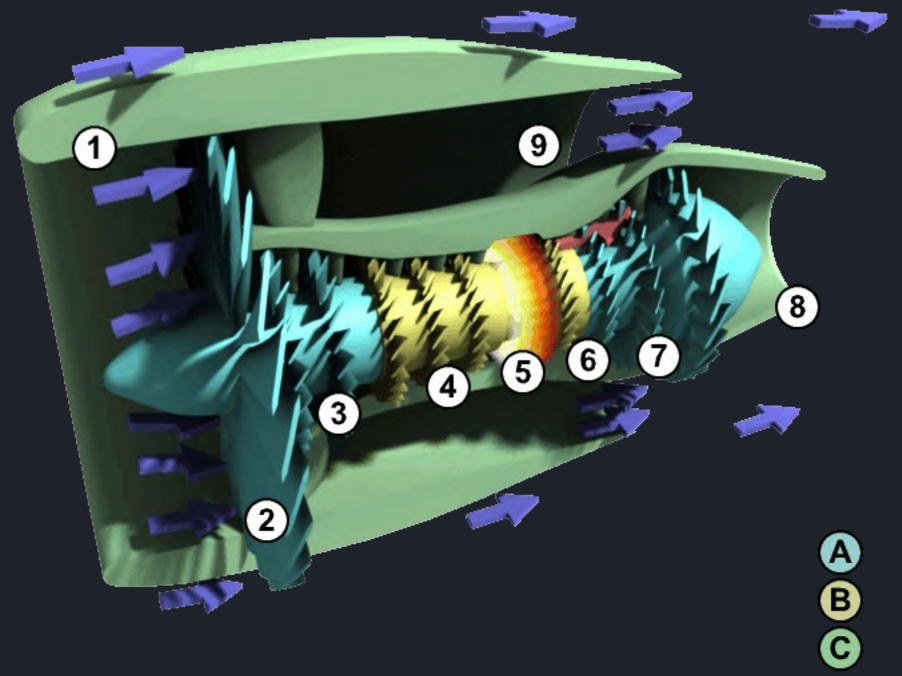
PROGNOSTICS HEALTH MANAGEMENT 8 (PHM08) CHALLENGE

Contents

Introduction

PROGNOSTICS AND HEALTH MANAGEMENT

"Prognostics and Health Management (PHM) is a multidisciplinary field aiming at maintaining physical systems in their optimal functioning conditions."



https://en.wikipedia.org/wiki/Turbofan

Background

THE CHALLENGE

Train a machine learning model to identify machine failures.

Dataset: PHM08 by NASA Prognostics Center of Excellence (PCoE).

EXPERIMENTAL SCENARIO

Multiple Multivariate time series data.

Flight engines of the same types with initial wear.

Engines are simulated with 3 operational settings that have substantial effect on engine performance.

Sensors recording various parameters are also contaminated with noise.

Symbol	*Description*	*Units*
Parameters available to participants as sensor data		
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	--
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass Ratio	--
farB	Burner fuel-air ratio	--
htBleed	Bleed Enthalpy	--
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s
Parameters for calculating the Health Index		
T48 (EGT)	Total temperature at HPT outlet	°R
SmFan	Fan stall margin	--
SmLPC	LPC stall margin	--
SmHPC	HPC stall margin	--

Source: [1]

Related Work

18+ publications with PHM08 dataset. [2]

Target problem solved:

1. Classification
2. Prognostics

Uncertainty management:

Probability theory, set-membership, belief functions, others

Performance measures:

PHM08 score, MSE, MAE, precision, others

Methods:

RNN, KF, MLP, genetic algo, Ensemble, logistic regression, Baysian updating, similarity matching, etc.

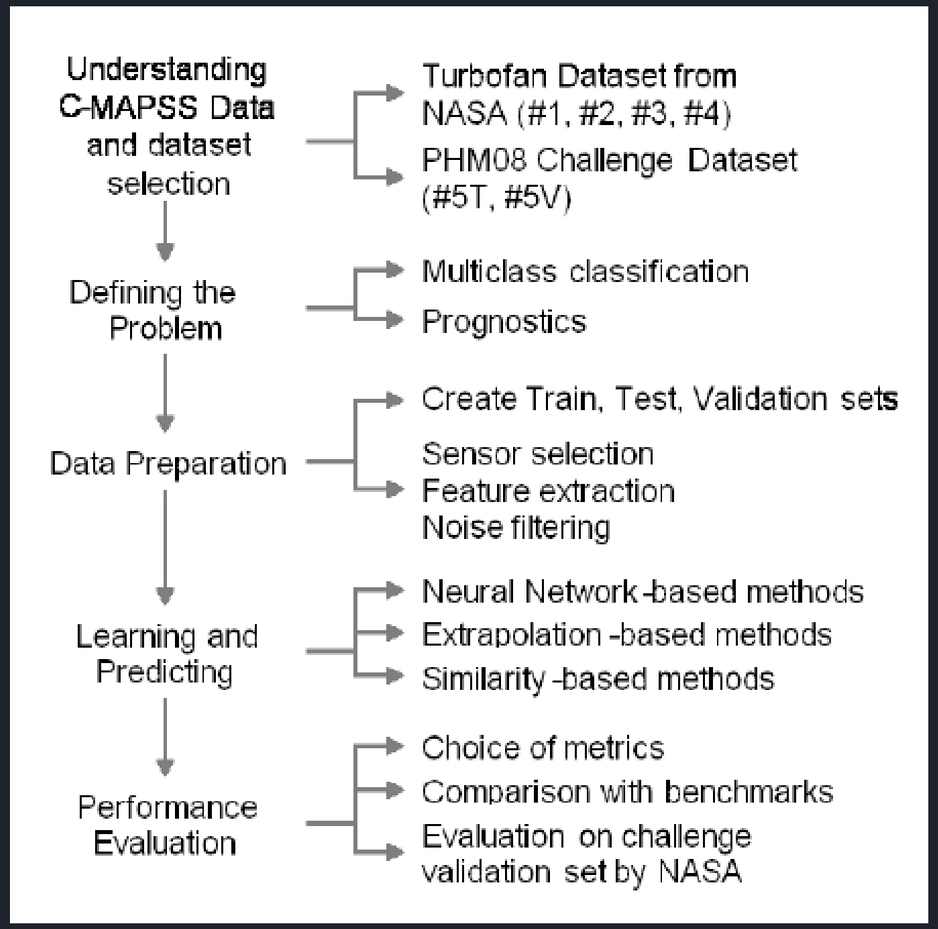


Figure: Guidelines to Using C-MAPSS Datasets.
Source: [2]

Steps in Jupyter notebook

- Understanding Data
- Defining problem
- Data Preparation
  - Data Splits
  - Sensor Selection
  - Normalization
- Learning and Predicting
- Performance Evaluation

Dataset

- No missing or undefined data.
- Train.txt has trajectories until failure.
- unit, time, op. setting 1, op. setting 2, op. setting 3, sensor 1,....26

Data preparation

RUL added

RUL = max cycles - current cycle

Splitting

- 70:15:15 ratio.
- No shuffling.
- Keep full trajectory.

Train trajectories

218

Size

~9MB

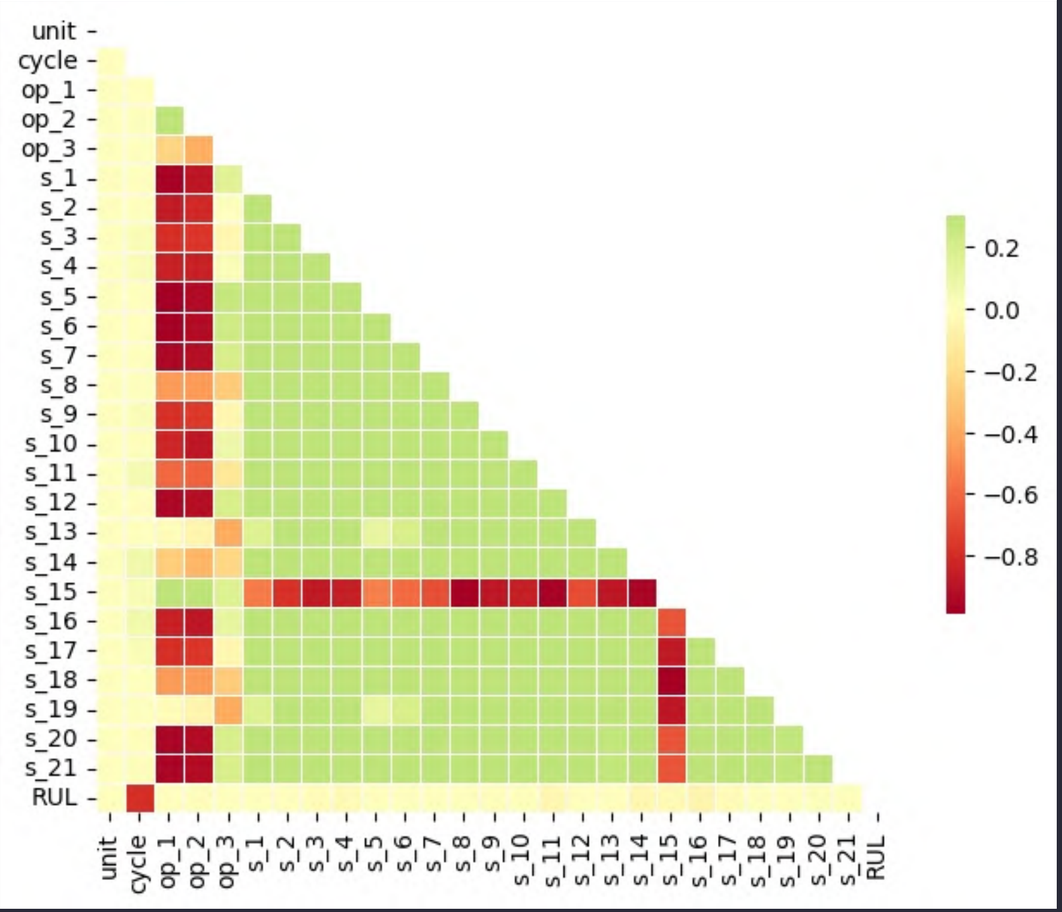
Rows

45,918

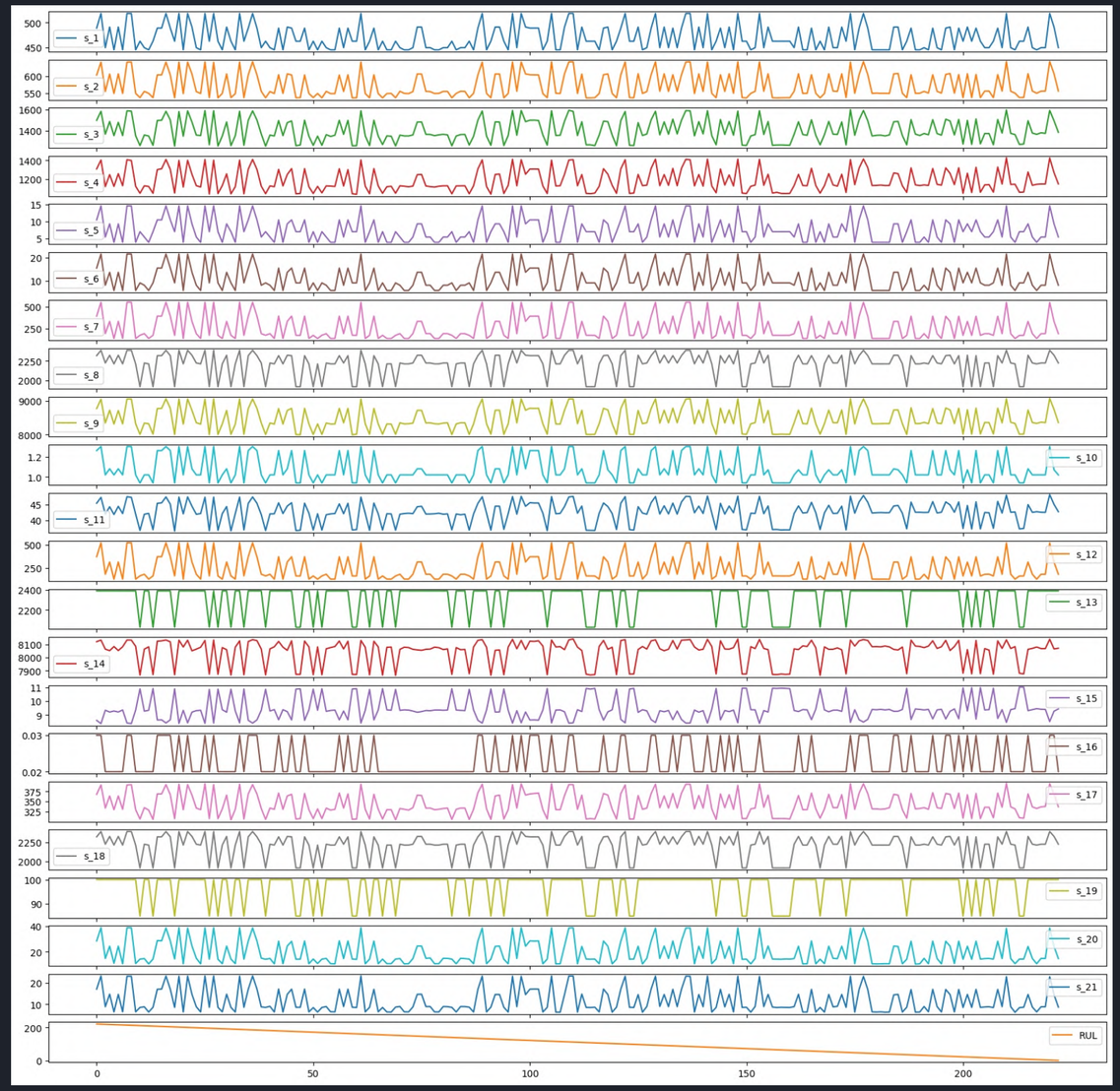
Columns

26

Correlation matrix

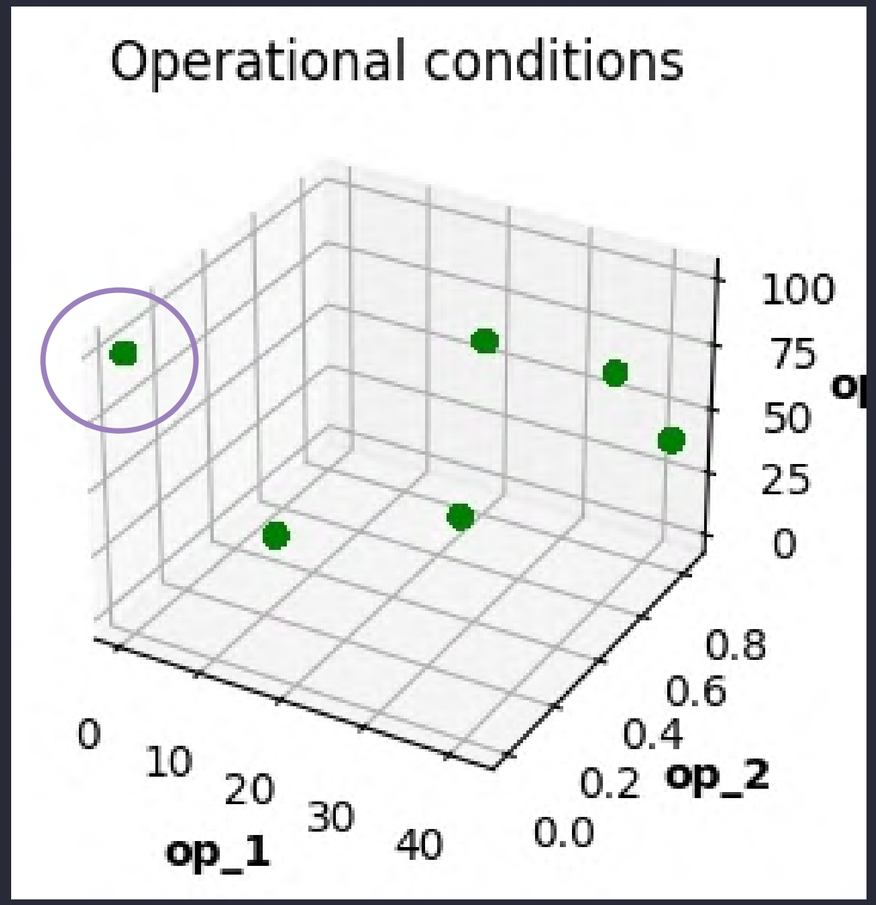


Sensor Data: Unit = 1, all cycles

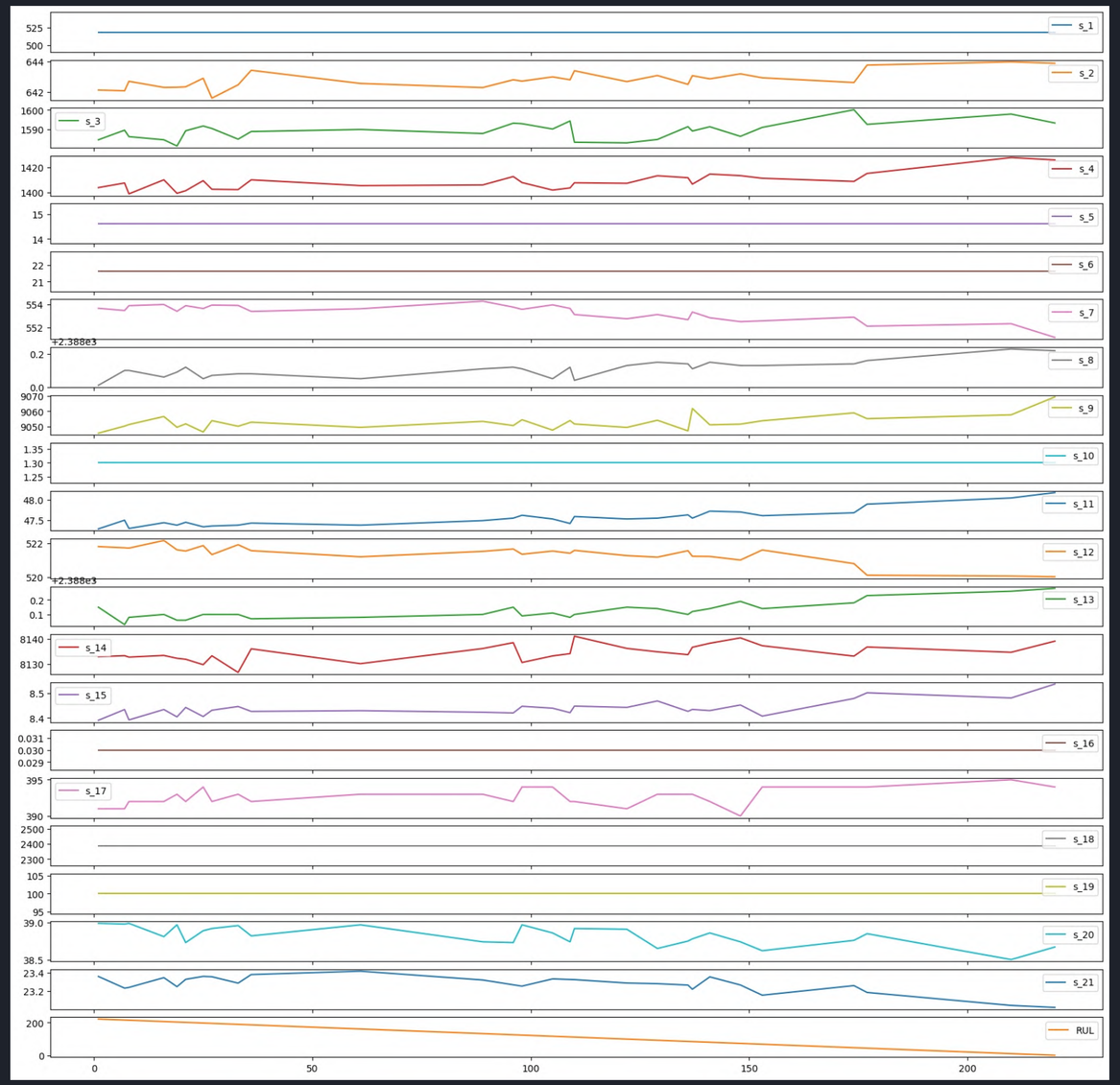


Operational Conditions: Kmeans clustering

- add modes as a feature



Sensor Data : Unit = 1, regime = 1 cycles



- sensor group types:

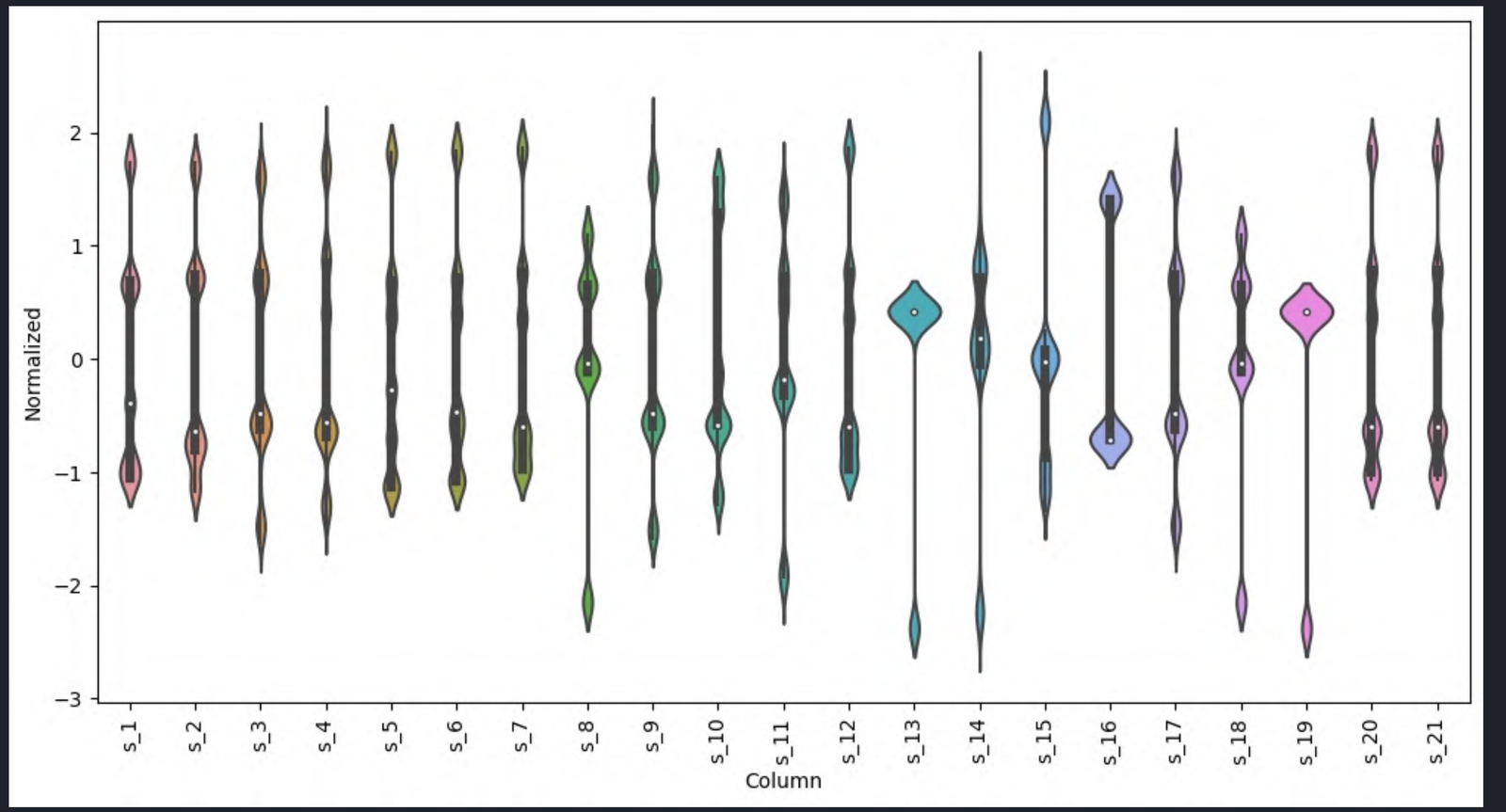
  Increasing, decreasing, constant

Normalization

mean = train.mean()

std = train.std()

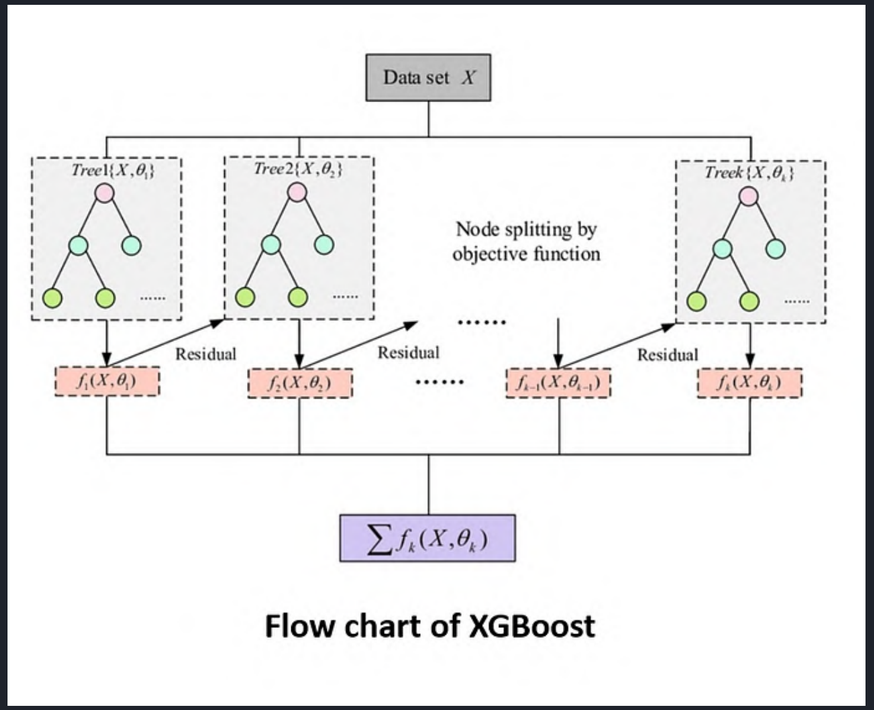
new val = (old val -mean)/std



Methods

EXTREME GRADIENT BOOSTING REGRESSOR

Problem definition is of Prognostics. Classification methods work on window approach and find a label for the next window.



Flow chart of XGBoost

Approach 1:

Use sensor as inputs, HIC as latent variable, RUL as output.

Bayesian Probability.

Approach 2:

Sequence models like RNN/LSTM or CNN(1D) with window to find label to next window. labels: healthy, failure, normal degradation, accelerated degradation.

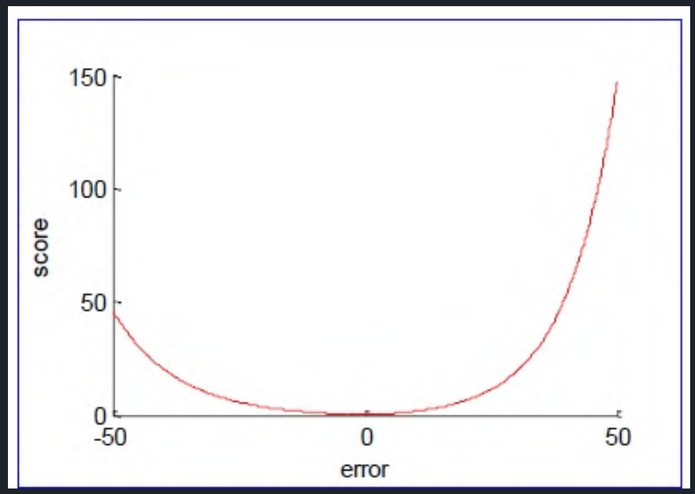
Approach 3:

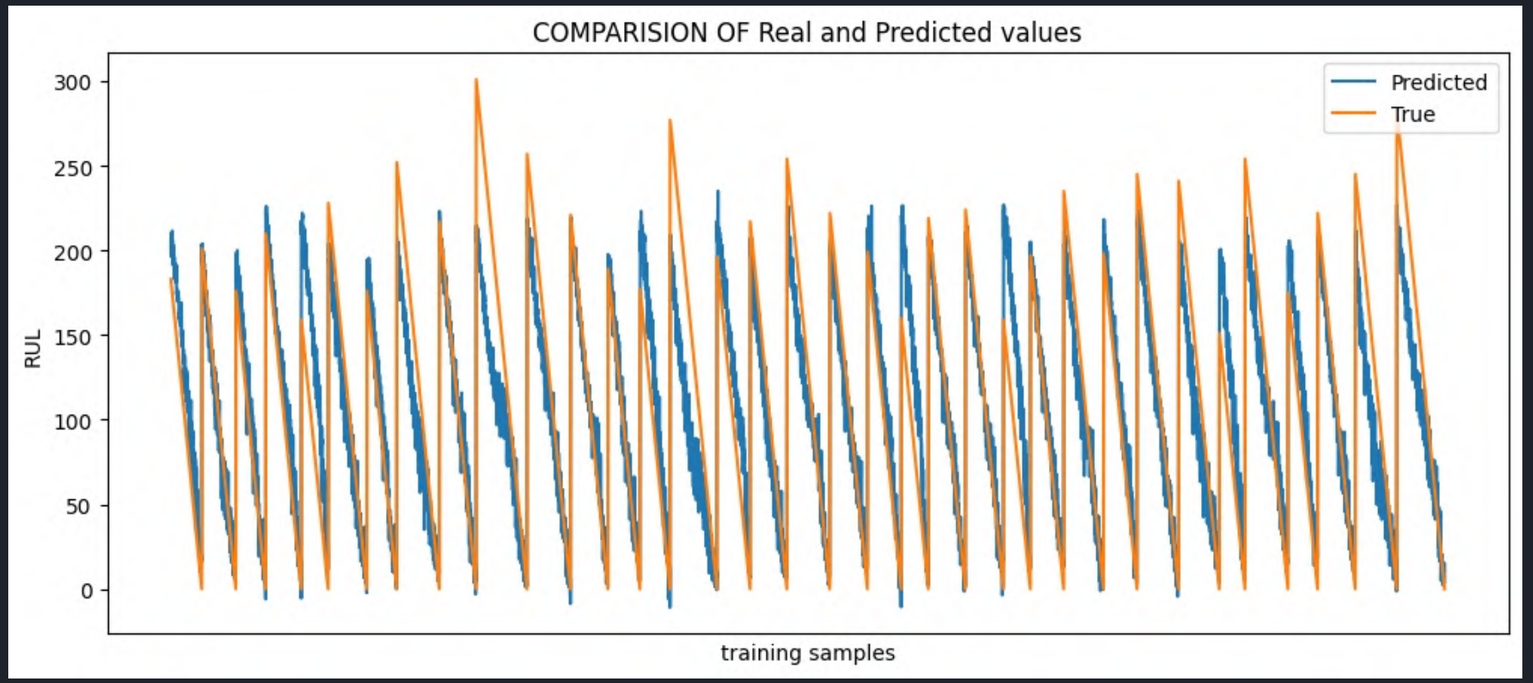
Ensemble models with regression trees.

Evaluation

Metrics :

- PHM08 score: 1054983.68
- mean absolute error (MAE): 22.51
- root mean squared error (RMSE): 29.91
- R2 score 0.8





Future Work

- Cluster engines/cycles in 6 regimes and then perform RUL separately.
- Implement PCA and select most imporant sensors, since Corr. matrix was inconclusive.
- Remove noise. moving avg. norm?
- Implement MLP/DNN. MLP + Kalman filter?

Thank You

JAYANT YADAV

References:

[1] Saxena A, Goebel K, Simon D, Eklund N. Damage propagation modeling for aircraft engine run-to-failure simulation. In2008 international conference on prognostics and health management 2008 Oct 6 (pp. 1-9). IEEE.

[2] Ramasso E, Saxena A. Review and analysis of algorithmic approaches developed for prognostics on CMAPSS dataset. InAnnual Conference of the Prognostics and Health Management Society 2014. 2014 Sep 29.