Data

Data set overview

Four data sets

	1 Operating Condition	6 Operating Conditions
HPC Degradation Failure	**FD001**	**FD002**
HPC Degradation and Fan Degradation	**FD003**	**FD004**

We chose to focus on HPC Degradation only, so we used data sets FD001 and FD002.

Training data format

Sensor readings per engine/cycle

Engine	Time in Cycles	Sensor 1 Reading	…	Sensor 26 Reading
1	**1**	34.99	…	1.02
1	**2**	41.67	…	1.02
…	**…**	…	…	…
1	**143**	63.12	…	1.26

In Train, there is one record per cycle until failure.

Test set is same format, except data for each unit truncates at a mystery point in time before failure.

Target definition

In Train data, we calculated for engine *i* in time period *t*:

Cycles until Failure$_{it}$ = (Max cycle$_{i}$) - (Current cycle)$_{it}$

Why XGboost?

Can be used for regression-type prediction

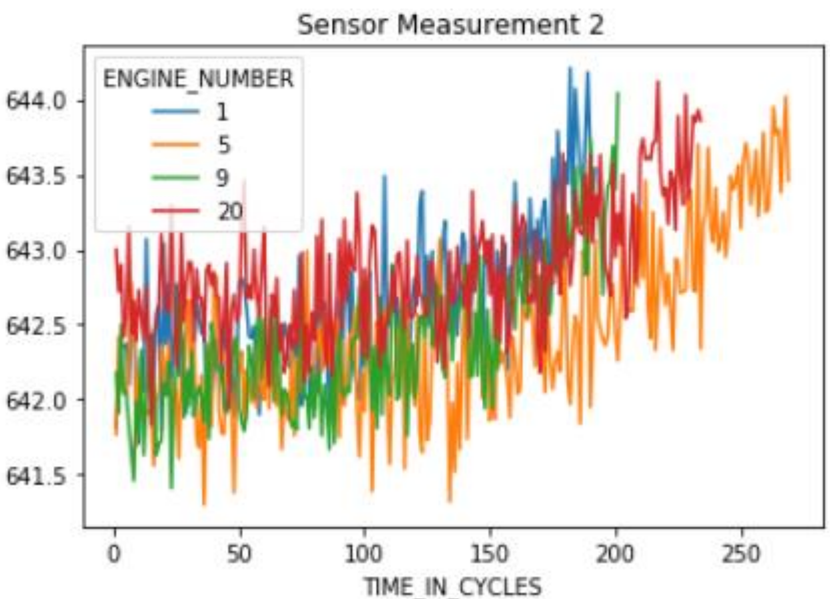
Not much needed for data prep – doesn't need standardizing, and our data was already all numeric

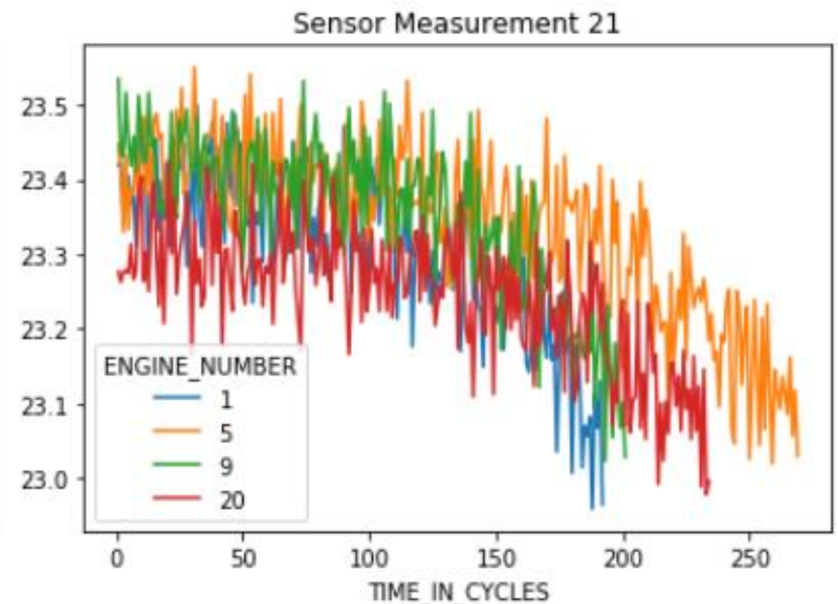
Robust to outliers

Speedy

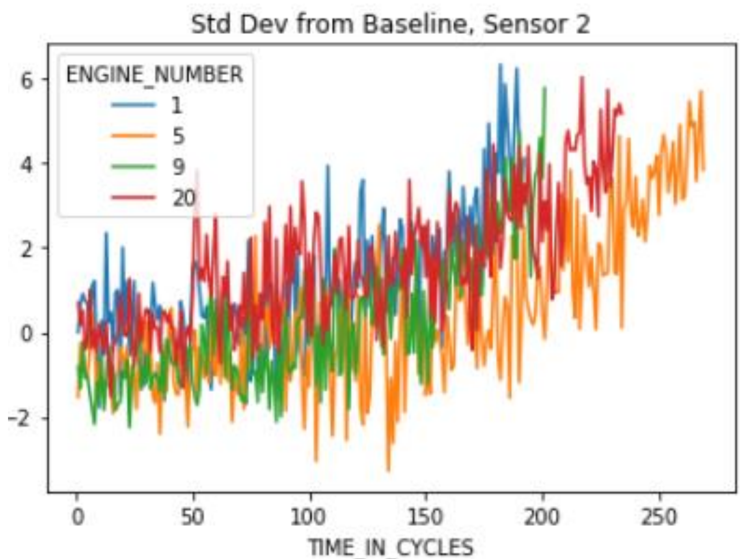
Data exploration

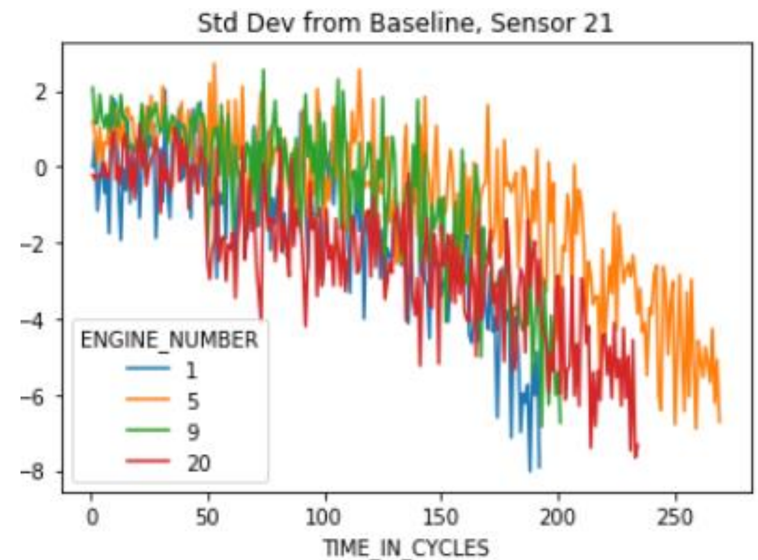
- Sensor readings gradually increase or decrease over time as they approach failure
- The readings are very noisy

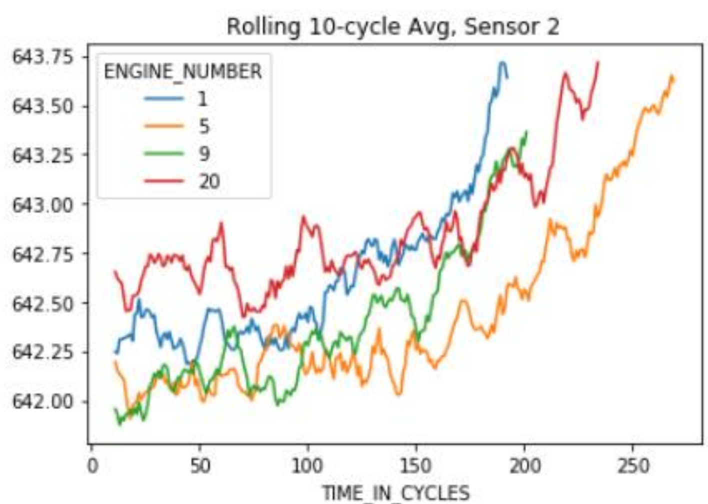


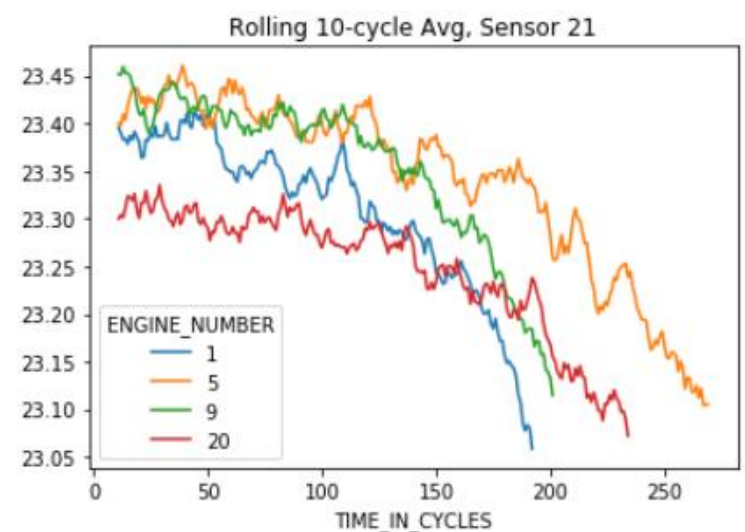


Feature engineering

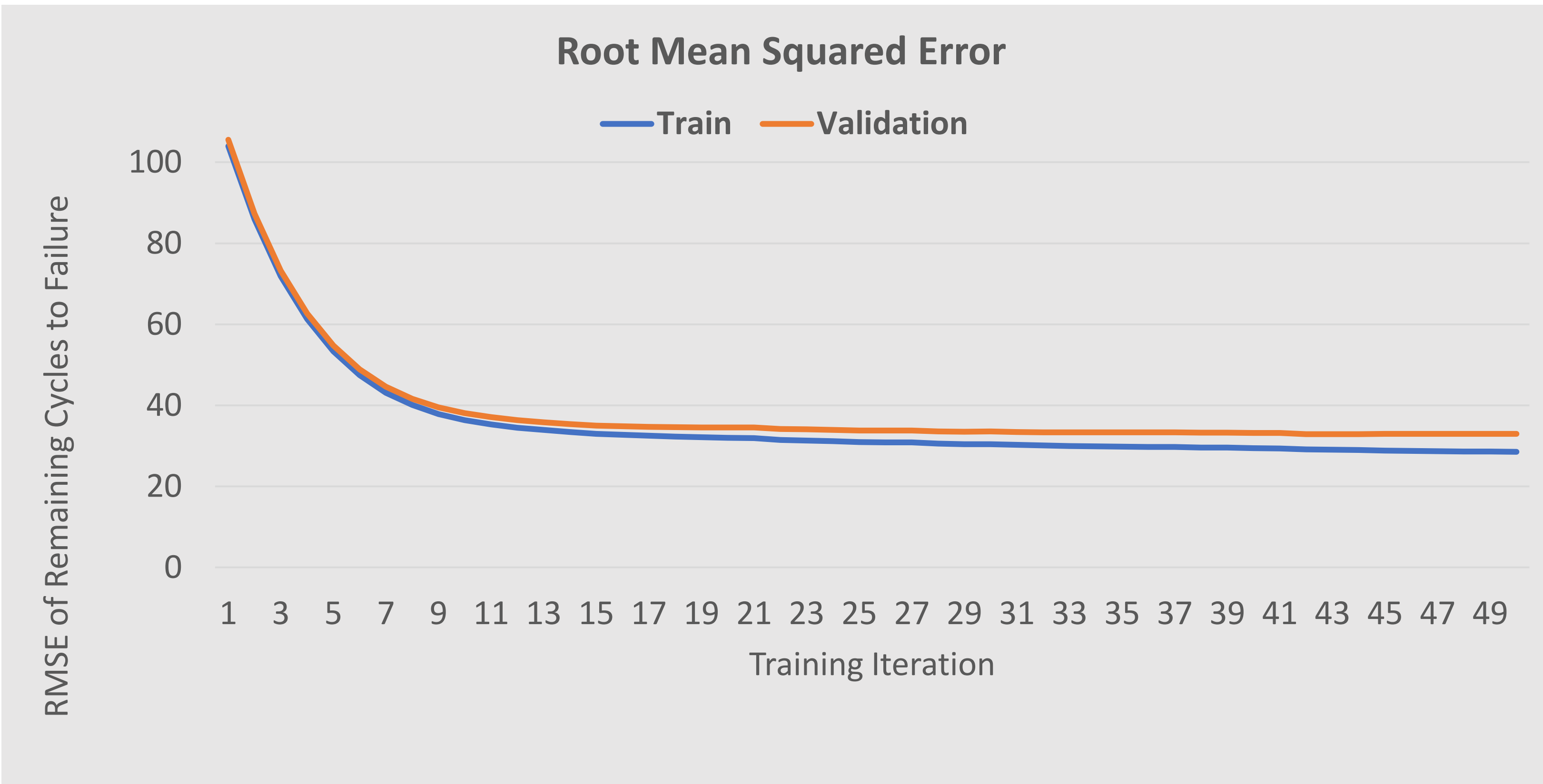






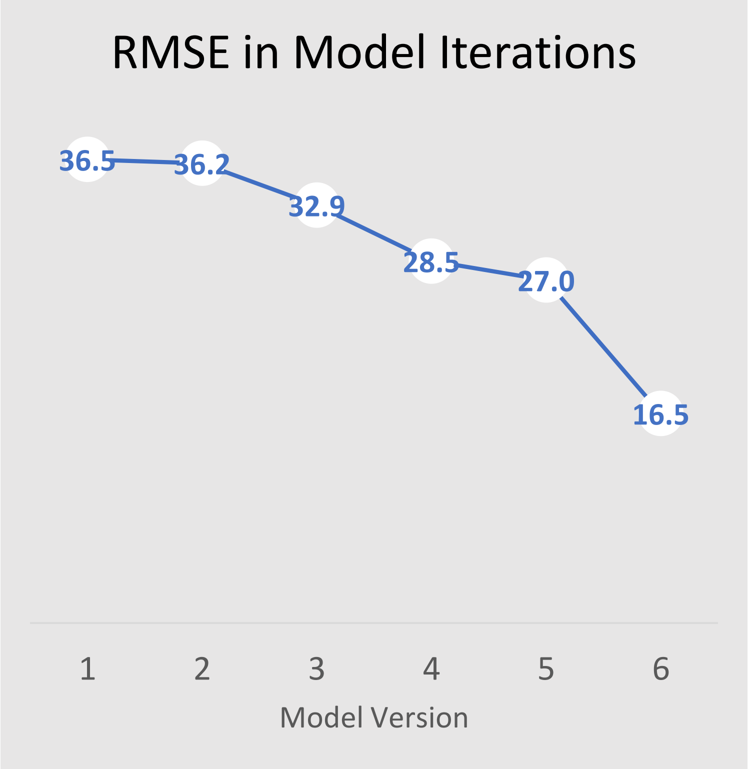


XGboost 1st training evaluation



Iterative improvement through feature engineering + hyperparameter tuning

Model Version	Model Description	Validation RMSE for Dataset FD001
1	Limited feature engineering, No hyperparameter tuning	**36.5**
2	Limited feature engineering, With hyperparameter tuning	**36.2**
3	Feature engineering round 1 (std deviation from baseline)	**32.9**
4	Feature engineering round 2 (rolling avg sensor readings)	**28.5**
5	Feature engineering round 3 (rolling avg of std deviation from baseline)	**27.0**
6	With hyperparameter tuning	**16.5**



“If we had time…”

Additional feature engineering

- There are papers with some cool feature engineering to identify signals in each engine’s sensor data over time

Clustering of similarly behaving engine units

- Can be used to further segment population
- Or as an input in predictive model

Try alternate populations

- We trained 1 model for the 1 operating condition data set (FD001) and a separate model for the 6 operating condition data set (FD002).
- Can we successfully combine those populations in one model?

Try alternate target definitions

- We completely ignored Fan Degradation failure type.
- Can we successfully model both failure types with the same model? Or are we more accurate training separate models?

Try alternate modeling algorithms

- Linear Learner, etc.

Try alternate evaluation metric

- It may make sense to train to minimize Mean Absolute Percent Error (MAPE) instead of RMSE, because as you get closer to failure, a small number of cycles off is more of a problem than when you’re farther from failure.

Further evaluation of our residuals (errors)

- To diagnose where we perform well/poorly