Prediciting Engine Failure | ML Model Results

ISSUE / PROBLEM

NASA is looking for data professionals to look into their jet engine simulated data and predict the number of remaining operational cycles before the jet engine reaches failure.

This is important because we want to understand when the failure may occur around in order to **save lives**, and economically **save on costs**.

There are four datasets that NASA wants experimented and this summary is **reporting on FD001**.

IMPACT

→ The model overall helps predict the time cycles that the engine may approximately fail.

→ Margin of error should also be taken into consideration as the model doesn't predict perfectly of when the engine will fail, but has a good idea of the timeframe the engine will fail.

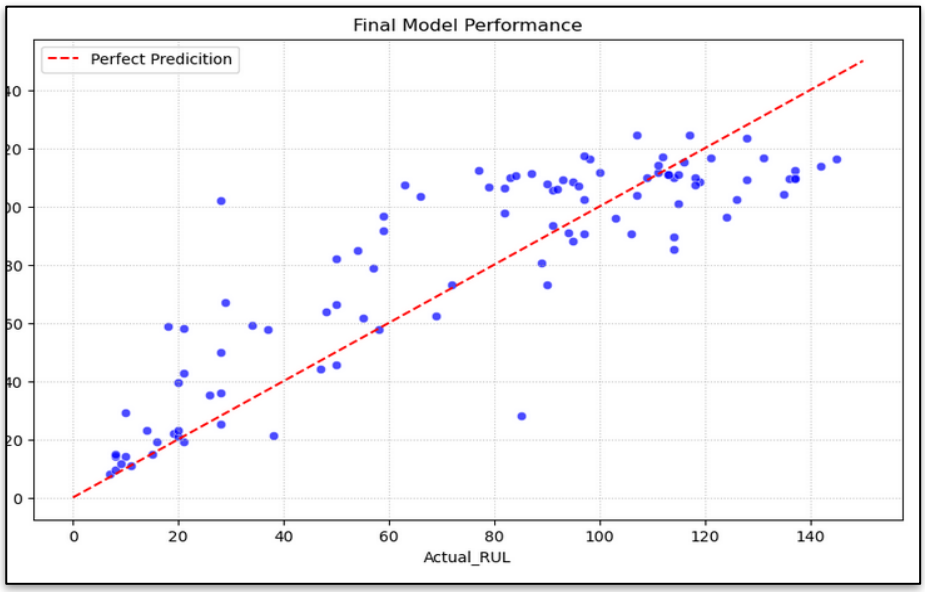
→ The results will be able to be used to determine when to run maintenance checks on the engines or replace them entirely.

RESPONSE

- To predict the **"Remaining Useful Life" (RUL)** I decided to use the XGBoost model. In order to obtain the best predictive results, I also used a GridSearch in order to tune the model to its best capabilities.

- Using the training set provided I was able to establish a RUL column of data and analyze where the data looked to be giving me the most useful information to focus my attention. I was able to focus my attention on three sensors and their information to guide where RUL may be coming to an end. Once this was established, I was able to use those insights to model what's in the test data and come to my results.

KEY INSIGHTS



- **Engineered features included 3 sensors:** s_7, s_8, s_11.

- **The XGBoost model was able to produce predicitive results with a root mean squared error of 20.72.** Meaning the model's predictions are approximately **20.72 units away** from the actual observed values.

- The graph above shows a visual of the predictive measures shown on the y-axis, with the "Actual_RUL" is shown on the x-axis. The red line indicates perfect prediction. The scatterplot reveals points that are closer to the line means the better the results.