Data Analysis and Model Development Report

- **Malay Vyas**

1. Data Collection

There were 3 types of dataset available for predictive maintenance; out of that ai4i2020 dataset was used as it has 10,000 product information. So it was good for training and testing purposes. Silicon wafer data was also collected but electronic maintenance was somehow getting diverted from actual 3 types of dataset what we already collected. Hence Silicon wafer data is kept as reference purpose.
Below is the feature description of the dataset used for analysis and model development.

a. Feature Description:

1) **Product ID**: consisting of a letter L, M, or H for low (50% of all products), medium (30%) and high (20%) as product quality variants and a variant-specific serial number.

2) **Type**: just the product type L, M or H from column 2.

3) **Air Temperature [K]**: generated using a random walk process later normalized to a standard deviation of 2 K around 300 K.

4) **Process Temperature [K]:** generated using a random walk process normalized to a standard deviation of 1 K, added to the air temperature plus 10 K.

5) **Rotational Speed [rpm]:** calculated from a power of 2860 W, overlaid with a normally distributed noise.

6) **Torque [Nm]:** torque values are normally distributed around 40 Nm with a SD = 10 Nm and no negative values.

7) **Tool Wear [min]:** (breakdown and gradual failure of a cutting tool due to regular operation) The quality variants H/M/L add 5/3/2 minutes of tool wear to the used tool in the process.

8) **Machine failure:** Machine Failure label that indicates, whether the machine has failed in this particular datapoint for any of the following failure modes are true.

Generated Data Report:

Generated_Data_Report.pdf

Link: https://github.com/MalayVyas/Maintenance-Prediction/blob/main/Generated_Data_Report.pdf

Notebook File Link:

https://github.com/MalayVyas/Maintenance-Prediction/blob/main/Maintanence_Prediction%20(3).ipynb