**PREDICTIVE MAINTENANCE**

**Abstract :**

Goal of this project is to predict values of the two output variables 'GT Compressor decay state coefficient' and 'GT turbine decay state coefficient' based on the values of 16 features representing meassurements in the vessel and specifically the engine compartment. Basic data exploration steps were conducted, the dataset was cleaned and normalized. After that the dataset was split into train and test dataset. The machine learning models applied were trained on the training set and then evaluated on the test dataset. Results of the analysis can be used to evaluate the state the engine, specifically the gas turbine compressor and turbines and to implement condition based or predictive maintenance on the vessel system. Algorithms applied were linear regression and descision tree regressor. Both algorithms performed well on the given test dataset. Decision tree based regression predicted the values of the outcome variables better that linear regression. Accuracy of the algorithms was measured by the mean squared error.

**Problem Statement :**

To develop a reliable and accurate model that can predict the failure of engine components before they occur. This will help to minimize downtime, reduce maintenance costs, and prevent catastrophic failures that could result in safety hazards. The model should be able to analyze sensor data, detect anomalies, and provide actionable insights to maintenance teams to proactively schedule maintenance or repairs. The challenge lies in developing a model that is accurate enough to provide reliable predictions, yet not overly complex, so it can be easily implemented and maintained.

**Algorithms Used :**

1. Linear Regression
2. Decision Tree Regressor

**Conclusion :**

We see in the plots above that the decision tree regressor uses features differntly for predicting our two ouput variables. For GT compressor decay it is the variable 'gt\_comp\_outlet\_airtemp' that is most important for the algorithm. 'gt\_comp\_outlet\_airpressure' is of highest importance for predicting GT turbine decay. Both algorithms, linear regression and decision tree regression provide good results in predicting the values of our two output variables based on the 14 features available in the dataset used. The difference in RSME, our measure of accuracy for the algorithms, is very small for GT turbine decay (only about 0.0002) and a little larger for GT compressor decay (0.0038). The Naval Propulsion Plants Data Set yields interesting opportunities for machine learning applications. The application of other regression algorithms like GLM,GAM, quantile regression or SVM as well as a combination of several models could be the basis of a system able to assist companies in maintaining production machines. Predictive maintenance is a hot topic nowadays for any company with machines that cost a lot of money when not producing goods. Implementing a system that is able to predict a possible machine failure in time can cut down repair costs and be an important part of a efficient production environment.