Engine Condition Trend Monitoring using Machine Learning Algorithms

**Shreyas S**

Machine Learning Intern, Faststream Technologies,

Bangalore, Karnataka, India

hello@shreyas.im

**Sudhanva N**

Machine Learning Inter, Faststream Technologies,

Bangalore, Karnataka, India

[nsudhanva@gmail.com](mailto:nsudhanva@gmail.com)

**Vinay Bansal**

Chief Executive Officer, Faststream Technologies,

Bangalore, Karnataka, India

[vinay@faststreamtech.com](mailto:vinay@faststreamtech.com)

**Vinod Agrawal**

Chief Technology Officer, Faststream Technologies,

Bangalore, Karnataka, India

vinod.agrawal@faststreamtech.com

**Abstract: Aircraft Engine Condition Trend Monitoring (ECTM) helps assist operators in managing the health of the turbine engines. Trend Monitoring is the process of monitoring the performance for engines over a period of time. By tracking a number of parameters from the turbine engine, it is possible to predict a failure before it occurs. This paper focuses on predicting the probability of failure of an aircraft turbine engine based on five parameters – Fan Blade’s noise, Exhaust Gas Temperature, Fuel Flow, Low Pressure Fan Speed and High Pressure Rotor Speed. Machine Learning Regression algorithms such as Multiple Linear Regression, Decision Tree Regression and Random Forest Regression are used for prediction purposes. Finally, comparison of all the three algorithms are made by calculating the R2 value.**

***Keywords: Engine Condition Trend Monitoring, Machine Learning, Regression, Aircraft Engine***

1. **INTRODUCTION**

Aircraft engines are composed of multiple moving and stationary components constituting a highly complex system. [1] Displays in the cockpit indicate performance of the engine using vital information [2] such as exhaust gas temperature, fuel flow, low pressure fan speed, high pressure rotor etc. Along with these, there are a lot of other parameters such as Fan Blade’s noise, vibrations, oil supply to bearings, air temperature and pressure etc., that can serve as an early indication of failure. There are different sensors for each of these components that record values over a period of time. These recorded values help in the prognosis of failure of an aircraft engine. It is extremely crucial for an aircraft manufacturer to know when any component in an engine might fail. With proper training, the analysis can be manually done by humans, but the risk for error is great. Hence, Engine Condition Trend Monitoring (ECTM) tools are used for predicting the failure of a components before they occur.

There are a lot of different ways in which ECTM can be implemented. In this paper, we have considered five parameters – Fan Blade’s noise, Exhaust Gas Temperature, Fuel Flow, Low Pressure Fan Speed and High Pressure Rotor Speed. Using historical data of these five parameters and how they affected the probability of failure of the engine, we perform predictive analysis using three Machine Learning Regression algorithms names Multiple Linear Regression, Decision Tree Regression and Random Forest regression. Performances of each of these algorithms have been compared by computing their R2 values.

The paper is organized as follows: Section II contains literature survey, Section III deals with the methodology, Section IV presents the implementation and pseudo code, Section V deals with Performance Analysis where every algorithm’s performance is measured and compared. Section VI finally concludes the work.

1. **LITERATURE SURVEY**

There is a lot of research work going on for Engine Condition Trend Monitoring.

QI Shufen et al. [3] have implemented an open layered architecture with intelligent reasoning to enhance the accuracy of fault diagnosis and reduce false reports. P Janssens et al. [4] have used neural networks and expert system techniques to enhance Engine Condition Monitoring Process to ensure better consistent results. Andrew K S Jardine et al. [5] have attempted to summarise and review the recent research and developments in diagnostics and prognostics of mechanical systems implementing Condition Based Monitoring with focus on techniques, technologies and algorithms for data processing and maintenance decision-making. K Mathioudakis et al. [6] have presented methods of analyzing aerothermodynamic performance measurement data for assessing the condition of components of a gas turbine. Louis A Urban [7] has presented an approach to turbine gas path analysis and monitoring that permits the isolation of single of simultaneous multiple engine faults, with a quantitative assessment of their relative severity.

1. **METHADOLOGY**

There are a lot of steps that go into Engine Condition Trend Monitoring. The architectural diagram for the same is given below.

Acquiring Field Data for Training

Train Machine Learning Algorithms with the field data

Save the Model

Acquire Field Data for Testing

Fit the saved Model to the Test Data

Predict the Probability of Failure

Visualize the Predictions

Compute R2 Score

Figure 1. Architectural Diagram

* The first step is to acquire historical data from multiple aircraft engines. We call this data as the Training Set. This data is used to train the Machine Learning Model by applying different algorithms.
* The trained model is then saved.
* Next step is to acquire some more new data for validating out trained machine learning model. This data is referred to as Test Set.
* The machine learning model is then loaded to the Test Set.
* Probability of failure of the engine is then predicted using the data from the Test Set.
* To validate the results, graph is plotted against Age of the engine and Probability of Failure with one plot representing nominal values and another plot representing the predictions from the Test Set.
* Finally R2 Score is computed for every algorithm and the results are compared.

1. **IMPLEMENTATION**

A few assumptions have been made due to the absence of available data.

Data has been generated based on nominal value and random degradation. A minimum and maximum value of each component has been preset. An increase in the value of any one of the components beyond its maximum value affects the probability of failure of that component. The failure of each component constitutes to the total probability of failure of the whole Turbine Engine. The turbine engine fails completely when the probability reaches one.

Data for one thousand engines with random degradation have been generated as Training Set which is used to train the Machine Learning Models.

Data for another engine in 10 different scenarios with random degradation has been generated as Test Set to validate the trained Machine Learning Model.

Details of all the steps involved in implementing ECTM is given below.

1. Generating Data

Data flow diagram for generating data is given below:

Importing Libraries

Set minimum and maximum values for each parameter

i = 1000 ?

i = 1

Generate values for each parameter with random degradation

i = i + 1

Stop

Generate probability of failure for each engine based on assumptions

Figure 2. Data flow for Generating Data

Pseudo code for the same is as follows:

Table 1. Pseudo code for Generating Data

1. Importing Libraries
2. Set the minimum and maximum values

for each parameter

1. Loop 1000 times

Generate values for each parameter with random degradation

Generate probability of failure for

each engine based on assumptions

1. Training Machine Learning Models

Data flow diagram for Training Machine Learning Models is given below:

Importing Libraries

Importing Dataset

Encoding Categorical Data

Avoid Dummy Variable Trap

Splitting Dataset

Applying Machine Learning Algorithms

Saving the Model

Figure 3. Data flow for Training Machine Learning Models

Pseudo Code for the same is given below:

Table 2. Pseudo code for Training Machine Learning Models

1. Importing Libraries
2. Importing Dataset
3. Encoding Categorical Data
4. Avoid dummy variable trap
5. Splitting dataset
6. Applying Machine Learning Algorithm
7. Saving the model

We have used three Machine Learning algorithms to train our models. They are as follows:

1. Multiple Linear Regression:

Multiple Linear Regression is a linear approach to modelling a relationship between a scalar response (dependent variable) and explanatory variables (independent variables). [8]

The model for Multiple Linear Regression, given n observations is

i = 1, 2, … , n

: Predicted value (dependent variable)

through : p distinct independent variables

: value of y when all the independent variables ( through ) are equal to zero

1. Decision Tree Regression:

Decision Tree is a learning technique where information extracted is organized from a training set in a hierarchical structured composed of nodes and ramifications. It is easy to understand the results for decision trees as the output of the decision tree can be organized in the form of tree or rules. [9]

The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches, each representing values for the attribute tested. Leaf node represents a decision on the numerical target. The topmost decision node in a tree which corresponds to the best predictor called root node.

1. Random Forest Regression:

The Random Forest Regression is a type of Machine Learning Algorithm that makes predictions by combining decisions from a sequence of base models. More formally we can write this class of models as:

*g(x) = f0(x) + f1(x) + f2(x) + …*

where the final model *g* is the sum of simple base models *fi*. Here, each base classifier is a simple Decision Tree.

1. Validating Machine Learning Models

Data flow diagram for Validating Machine Learning Models is given below:

Importing Libraries

Importing Dataset

Encoding Categorical Data

Avoid Dummy Variable Trap

Fitting Machine Learning Model to the Dataset

Computing Results

Figure 4. Data flow for Validating Machine Learning Models

Pseudo Code for the same is given below:

Table 3. Pseudo code for Training Machine Learning Models

1. Importing Libraries
2. Importing Dataset
3. Encoding Categorical Data
4. Avoid dummy variable trap
5. Fitting Machine Learning Model to the Dataset
6. Computing Results

References

1. <https://ntrs.nasa.gov/search.jsp?R=20100024368>
2. <https://arc.aiaa.org/doi/abs/10.2514/6.1999-2528>

6.http://journals.sagepub.com/doi/abs/10.1243/0957650011538442

7. <https://arc.aiaa.org/doi/abs/10.2514/3.60240>

8. <http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2008/2008-19-multiple-linear-regression.pdf>

9. https://www.sciencedirect.com/science/article/pii/S0957417411009237