

Remaining Useful Life Prediction of Aircraft Turbofan Engine Based on Random Forest Feature Selection and Multi-Layer Perceptron

Abstract

[2] Wang et al. This research paper proposes a method for predicting the remaining service life (RUL) of turbojet engines with the aim of improving engine safety and reducing maintenance costs the paper focused on the problem that engine monitoring data is complex, non-linear, and contains a large number of characteristics making the **RUL** prediction process difficult therefore the researchers used the Random Forest algorithm to select the important characteristics from the sensor data and then used the SES algorithm to smooth the data and reduce the effect of noise. An MLP model was then built to predict remaining life. Bayesian parameter updating was used to improve the model's parameters and reduce the RMSE value. The model was tested using the C MAPSS dataset of turbofan engines, and the results showed that the model improved prediction accuracy, with the RMSE value in the FD001 set decreasing by 6.1% compared to several other methods

Introduction

[2] Wang et al. explain in the introduction that aircraft engines are among the most important components of aviation equipment because they directly affect aircraft safety and flight performance therefore PHEM techniques are used to monitor engine condition and predict remaining service life this helps support predictive maintenance and improve equipment reliability and safety the paper also mentions that methods for predicting remaining service life are divided into three main approaches physical **model-based methods**, **data-based methods**, and **hybrid methods** due to the difficulty of building an accurate physical model of the engine data-based methods and machine learning have become important topics in this field therefore this paper proposes a model based on Random Forest for selecting influential characteristics then improved MLP using Bayesian updating to predict the remaining service life of aircraft engines

Related Work

[2] Wang et al. The paper reviewed previous studies in the field of predicting the remaining lifespan of aircraft engines it noted that some studies used deep learning models such as MLP, DLSTM, and CNN RNN to analyze sensor data and predict engine condition the paper also explained that some researchers used methods for extracting characteristics such as k-means and Health Index construction methods to reduce the dimensions of the data and improve prediction quality It further mentioned similarity methods such as **Similarity Matching** and **Dynamic Time Warping** Previous methods were used to compare current engine data with historical data but the problem with many of them is that they can be affected by the large number of characteristics the difficulty in selecting the most important sensors, or the noise present in the data therefore this paper aims to improve upon previous work by selecting the important characteristics using a **Random Forest** then smoothing the data and finally using an improved MLP model to predict remaining life

Methodology

[2] Wang et al .The methodology of this paper relies on the use of the C MAPSS turbofan engine dataset provided by NASA this dataset simulates the degradation of aircraft engines under various operating conditions and contains four subsets: FD001, FD002, FD003, and FD004 these subsets differ in the number of operating conditions and failure patterns in this paper this data was used to test the model's ability to predict the remaining engine lifespan initially the researchers used the Random Forest algorithm to select the most important characteristics from the sensor data the importance of 21 engine data characteristics was evaluated and then characteristics with a weak impact on lifespan were removed leaving only 14 characteristics as inputs to the model the algorithm was then used **SES** was used to smooth the data and reduce noise MinMax normalization was then used to convert the data values to a uniform range between 0 and 1 so that value variations wouldn't affect model training the MLP model was then trained and the researchers used Bayesian parameter updating to select the best parameters for the model to minimize its RMSE values the model was evaluated using metrics such as RMSE, MAE, Score, and R2

Results

[2] Wang et al . The results showed that the proposed model performed well in predicting the remaining service life of turbofan engines. The model was tested on four groups: FD001, FD002, FD003, and FD004 the paper explained that prediction in FD001 and FD003 was easier and better because the data was under a single operating condition, while prediction in FD002 and FD004 was more difficult because it included multiple operating conditions, which increased the complexity of the prediction the paper also compared the proposed model with other methods such as **Ridge Regression, Lasso Regression, LGBM**, and some previous methods, stating that the results predicted by the proposed model were closer to the actual values of the remaining service life. Among the most important numerical results, the RMSE value in group FD001 decreased by 6.1% compared to other modern methods, and the R2 value increased by 2.4% this indicates that using Random Forest for feature selection with Bayesian optimized MLP helped improve the prediction accuracy

Paper	All Models	Datasets	Dataset Link	Limitation	Best model	results
Wang et al. [2]	Proposed RF + SES + Bayesian optimized MLP Comparison models Ridge Regression Lasso Regression LGBM KNR	NASA C MAPSS FD001 FD002 FD003 FD004			RF feature selection + SES + Bayesian optimized MLP	FD001 RMSE decreased by 6.1 percent and R2 increased by 2.4 percent compared with other state of the art methods

Ref

[2] Wang, H., Li, D., Li, D., Liu, C., Yang, X., and Zhu, G. "Remaining Useful Life Prediction of Aircraft Turbofan Engine Based on Random Forest Feature Selection and Multi-Layer Perceptron." *Applied Sciences*, vol. 13, no. 12, Article 7186, 2023. DOI: 10.3390/app13127186.