

A Deep Learning Model for Remaining Useful Life Prediction of Aircraft Turbofan Engine on C-MAPSS Dataset

Abstract

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. The impact of the industrial revolution on the safety, efficiency, and reliability of industrial machinery is significant in commercial sectors this system allows us to predict the remaining lifespan of industrial equipment in a timely manner and take timely maintenance actions thus reducing downtime due to sudden breakdowns and increasing business revenue over the past decade methods for predicting the lifespan of industrial equipment have garnered considerable attention from researchers various deep learning-based techniques for lifespan prediction exist utilizing Long Term Memory Networks (**LSTMs**) to improve prediction quality this paper proposes a model that combines efficient preprocessing with an LSTM network and one of the key factors in lifespan prediction The remaining assumption is to determine the starting point of turbine engine failure the paper focuses on using and analysing turbine engine sensor data to select the most influential data in the prediction process and reduce irrelevant data The proposed model aims to improve the accuracy and quality of the prediction of the remaining lifespan by reducing errors in the results and comparing them with previous methods used in the same field

Introduction

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. Automation in the modern industrial revolution has impacted the field of industrial machinery safety making maintenance crucial across various sectors the paper explains that predictive maintenance before a breakdown occurs allows for RUL (Residual Lifespan) procedures to be implemented which help estimate the remaining service life of equipment timely maintenance reduces unexpected downtime thus lowering maintenance costs and preventing the impact on profits caused by sudden work and production stoppages the paper focuses on turbine engines used in aircraft because they require continuous monitoring due to their importance to safety the study indicates that there are several methods for predicting remaining service life which can be divided into three categories main methods: These are methods based on **physical models**, **data-based methods**, and **hybrid methods** due to the availability of sensor data and the development of artificial intelligence technologies learning in predicting lifespan has become common therefore, this paper proposes a model based on **LSTM** and **CNN** with preprocessing steps such as sensor selection applying **LSTM** to the data network and determining the starting point for engine failure in order to improve the quality and accuracy of predicting the remaining lifespan of the engine

Related Work

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. This research paper reviewed previous work and studies in the field of predicting the lifespan of turbine engines initially there were traditional methods based on mathematical models such as the Kalman Filter, Extended Kalman Filter, and Particle Filter this is to analyze sensor data and link it to the engine condition some studies used statistical models such as Weibull and Eyring to predict the remaining engine lifespan after that machine learning-based methods such as ARIMA and SVR emerged With the development of deep learning researchers used models such as Deep Belief Network, CNN, LSTM, Bi-LSTM, and CNN-LSTM these models proved their ability to extract hidden patterns from sensor time data and improve prediction accuracy the paper also explained that some previous studies used a linear failure model with LSTM and achieved good results Therefore this paper came to improve previous studies by integrating preprocessing steps with the LSTM network to improve the quality and accuracy of predicting the engine lifespan

Methodology

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. The methodology of this paper relies on the use of the NASA C-MAPSS dataset a simulation of aircraft turbo engines this dataset comprises four subsets: FD001, FD002, FD003, and FD004 these subsets include readings from various sensors with different operating cycles and lifespans the primary objective is to predict the remaining engine lifespan before failure initially the researchers used correlation analysis to select the important sensors that have a clear impact on the prediction and are strongly correlated with the remaining engine lifespan sensors with no clear impact on the prediction were excluded the selected sensors were then passed through a Moving Median Filter to remove outliers from the data The data were then printed using Z-score normalization because the sensor values vary in range which can affect the model training accuracy An improved piecewise linear degradation model was also used to determine the starting point of motor failure as motors do not fail on their first operating cycle finally the processed data and updated remaining life labels were used to train the LSTM network and the model was evaluated using RMSE and the Score Function

Results

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. The results showed that the proposed model performed well in predicting the remaining service life of turbo engines. The model was tested on all four C-MAPSS subsets, using different LSTM models with varying settings to achieve the best results the initial service life values for subsets FD001, FD002, FD003, and FD004 were 78, 103, 79, and 87, respectively the model also achieved the best values which were 7.78 for FD001, 17.64 for FD002, 8.30 for FD003 and 17.63 for FD004 the results indicated that the model performed better in FD001 and FD003 because they were less complex while FD002 and FD004 were more difficult due to more demanding operating conditions and irregular sensor data behaviour overall the paper demonstrated that integrating preprocessing steps with a network LSTM and the use of a better method to determine the starting point of faults helped improve the accuracy of the prediction compared to previous methods

Conclusion

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. The paper concluded that using the LSTM model with preprocessing steps yielded good results in predicting the remaining service life of turbo engines steps such as selecting critical sensors and refining data significantly improved modeling performance and the improved segmented linear failure model played a crucial role in identifying the starting point of engine degradation this significantly impacted the accuracy of predicting the remaining lifespan the paper also indicated that using a single model for all data might not be a good option because each subset has different characteristics therefore using a separate model for each subset yields better results the results also showed that the proposed model achieved higher accuracy compared to some previous studies using C-MAPSS data In the future the paper suggests applying this approach to real-world problems and other instruments using more advanced deep learning algorithms to handle complex and realistic sensor data

Asif et al. [1]	Deep LSTM with four hyperparameter configurations	NASA C-MAPSS	NASA C-MAPSS Dataset		Deep LSTM with preprocessing	FD001 RMSE=7.78, FD002 RMSE=17.64, FD003 RMSE=8.30, FD004 RMSE=17.63
-----------------	---	--------------	----------------------	--	------------------------------	---

Ref

[1] Asif, O., Haider, S. A., Naqvi, S. R., Zaki, J. F. W., Kwak, K.-S., and Islam, S. M. R. "A Deep Learning Model for Remaining Useful Life Prediction of Aircraft Turbofan Engine on C-MAPSS Dataset." IEEE Access, vol. 10, 2022, pp. 95425–95440. DOI: 10.1109/ACCESS.2022.3203406.

Remaining Useful Life Estimation of Turbofan Engines with Deep Learning Using Change-Point Detection Based Labeling and Feature Engineering

Abstract

[3]Ensarioğlu K İnkaya T and Emel E . This paper proposes a model for predicting the remaining lifespan of turbofan engines using deep learning predicting remaining lifespan is a challenging problem in engine health monitoring because sensor data requires thorough processing before model training researchers employed a method based on differential feature construction change point detection and then used a hybrid 1D CNN LSTM deep learning model to predict remaining lifespan the model was evaluated using the FD001 C MAPSS dataset the results showed that the proposed model outperformed some machine learning and deep learning models achieving an RMSE of 16 and a score of 437.

Introduction

[3]Ensarioğlu K İnkaya T and Emel E . The paper states that turbofan engines are critical components in aircraft and that predicting their remaining service life helps support predictive maintenance and reduce unexpected failures it also notes that most recent studies in this field rely on data-driven models because sensor data contains crucial information about engine degradation the paper highlights the problem that many studies consistently use piecewise linear labeling assigning the same initial service life value to all engines during their normal operating period this may not accurately represent the state of degradation therefore the paper suggests using change Point Detection to better identify the starting point of degradation for each engine and then using a 1D CNN LSTM model to improve the accuracy of predicting remaining service life.

Related Work

[3]Ensarioğlu K İnkaya T and Emel E. Previous studies in the field of predicting the remaining lifespan of turbofan engines have shown that many studies have used data driven models such as machine learning and deep learning to analyze sensor data and predict remaining lifespan the paper also mentions that some studies used LSTM because it is suitable for handling temporal data while others used CNN to extract features from sensor data some studies relied on methods for classifying or dividing deterioration stages or used methods to extract features before prediction however the problem is that many of the previous methods rely on fixed labels for remaining life or do not adequately utilize the development of new features therefore this paper aims to improve upon previous work by integrating feature engineering with change-point detection-based labeling and a 1D CNN-LSTM model.

Methodology

[3]Ensarioğlu K İnkaya T and Emel E. The C MAPSS FD001 turbofan engine dataset was used to evaluate the proposed model. The methodology focuses on three main steps first new features were constructed from the sensor data using difference-based feature construction to extract clearer information about engine condition changes then change point detection based on piecewise linear labeling was used to identify the starting point of deterioration instead of relying on a fixed value for initial age subsequently a 1D CNN-LSTM hybrid deep learning model was trained the CNN helps extract important features from the data while the LSTM helps understand temporal changes in the sensor data the proposed model was then compared with machine learning and deep learning models both with and without the new features to determine the impact of the processing steps on prediction accuracy.

Results

[3]Ensarioğlu K İnkaya T and Emel E .The results showed that the 1D CNN LSTM model with features built using difference feature analysis achieved the best performance in predicting the remaining service life of turbofan engines on the FD001 set of C MAPSS data. The paper achieved an RMSE score of 16.1 and a Score of 437.2 the results also showed that using feature engineering with change point detection based labeling helped improve prediction accuracy compared to models that did not use these steps the paper stated that the proposed model outperformed some previous machine learning and deep learning models indicating that identifying the point of deterioration and building appropriate features helps improve predictive maintenance results.

Conclusion

[3]Ensarioğlu K İnkaya T and Emel E .The study concluded that integrating change point detection with feature engineering and a 1D CNN LSTM model improves the accuracy of predicting the remaining lifespan of turbofan engines the change point detection method allowed for more realistic remaining lifespan designations compared to using a fixed value for all engines the new features also enhanced the model's ability to understand engine degradation the paper indicated that the proposed model performed well on FD001 data from C MAPSS but its main limitation was its focus on only a single subset therefore future testing of the method on more complex subsets and multiple operating conditions is recommended to improve the model's generalizability.

Paper	All Models	Datasets	Dataset Link	Limitation	Best model	results
Ensarioğlu et al [3]	1D CNN LSTM and comparison with machine learning and deep learning models with and without difference feature	C MAPSS FD001			1D CNN LSTM with difference feature and change point based labeling	RMSE 16.1 and Score 437.2

Ref

[3]Ensarioğlu K İnkaya T and Emel E “Remaining Useful Life Estimation of Turbofan Engines with Deep Learning Using Change Point Detection Based Labeling and Feature Engineering” Applied Sciences vol 13 no 21 Article 11893 2023 DOI 10.3390/app132111893

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.