PREDICTING REMAINING USEFULL LIFE OF TURBOFAN ENGINE USING RESIDUAL ENSEMBLE MODEL

A PROJECT REPORT

Submitted by

SHIVANSH SHARMA [RA2011027010081]
AKASH KUMAR PANDIT [RA2011027010122]

Under the Guidance of

DR. S. SHARANYA
Associate Professor, Department of Data Science and Business Systems

In partial fulfillment of the Requirements for the Degree of

BACHELOR OF TECHNOLOGY
COMPUTER SCIENCE AND ENGINEERING
with specialization in Big Data Analytics

DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS

COLLEGE OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR - 603203

MAY 2024

Department of Data Science and Business Systems

SRM Institute of Science & Technology
Own Work Declaration Form

Degree/ Course : B. Tech Computer Science with specialization in Big Data Analytics

Student Name : Akash Kumar Pandit , Shivansh Sharma

Registration Number : RA2011027010122, RA2011027010081

Title of Work : PREDICTING REMAINING USEFULL LIFE OF ENGINE USING TURBOFAN RESIDUAL ENSEMBLE MODEL

We hereby certify that this assessment compiles with the University's Rules and Regulations relating to Academic misconduct and plagiarism, as listed in the University Website, Regulations, and the Education Committee guidelines.

We confirm that all the work contained in this assessment is our own except where indicated, and that We have met the following conditions:

- Clearly referenced / listed all sources as appropriate
- Referenced and put in inverted commas all quoted text (from books, web, etc)
- Given the sources of all pictures, data etc. that are not our own
- Not made any use of the report(s) or essay(s) of any other student(s) either past or present
- Acknowledged in appropriate places any help that I have received from others (e.g.fellow students, technicians, statisticians, external sources)
- Compiled with any other plagiarism criteria specified in the Course handbook /University website.

We understand that any false claim for this work will be penalized in accordance with the university policies and regulations.

DECLARATION:	
We are aware of and understand the University's policy on Academic misconduct and plagiarism and we certify that this assessment is our own work, except where indicated by referring, and that we have followed the good academic practices noted above.	
Shivansh Sharma (RA201127010081)	Akash Kumar Pandit (RA2011027010122)

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR-603203

BONAFIDE CERTIFICATE

Certified that this project report titled "**PREDICTING REMAINING USEFULL LIFE OF ENGINE USING RESIDUAL ENSEMBLE MODEL"** is the bonafide work of **Shivansh Sharma [RA2011027010081]** and **Akash Kumar Pandit [RA2011027010122]** who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion for this or any other candidate.

SIGNATURE	SIGNATURE
DR. S. SHARANYA	**DR. M. LAKSHMI**
GUIDE	**HEAD OF THE DEPARTMENT**
Associate Professor	Professor
Department of Data Science	Department of Data Science
and Business Systems	and Business Systems

Signature of Internal Examiner	Signature of External examiner

ABSTRACT

This research introduces a new method for forecasting the Remaining Useful Life (RUL) of aircraft turbofan engines. The method is called the Residual Ensemble Model and it combines the Random Forest and Gradient Boosting techniques. The Random Forest method is first used to evaluate and prioritize the importance of several historical monitoring parameters of the engine, establishing a strong basis for selecting the most relevant features. Afterwards, the improved inputs undergo processing using a Gradient Boosting model, which adjusts the predictions by addressing any remaining mistakes and improving the accuracy of the model. This dual-pronged strategy enhances both the accuracy of predictions and the efficiency of maintenance programs, resulting in longer engine lifespan and less expenses. The efficacy of this integrated model is verified using comprehensive life cycle data obtained from real engine operations, showcasing substantial improvements over conventional single-model methodologies.

ACKNOWLEDGEMENTS

We express our humble gratitude to **Dr C. Muthamizhchelvan**, Vice-Chancellor, SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to **Dr T.V.Gopal** Dean-CET, SRM Institute of Science and Technology, for his invaluable support.

We wish to thank **Dr Revathi Venkataraman**, Professor & Chairperson, School of Computing, SRM Institute of Science and Technology, for her support throughout the project work.

We are incredibly grateful to our Head of the Department**, Dr M. Lakshmi,** Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our thanks to our program coordinators **Dr. G. Vadivu,** Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for her inputs during the project reviews and support.

We register our immeasurable thanks to our Faculty Advisor**, Dr. S. Sharanya,** Associate Professor, **R Radha**, Assistant Professor Department of Data Science and Business Systems, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to my guide, **Dr. S Sharanya**, Associate Professor, Department of Data Science and Business Systems, SRM Institute of Science and Technology, for providing me with an opportunity to pursue my project under her mentorship. She provided me with the freedom and support to explore the research topics of my interest. Her passion for solving problems and making a difference in the world has always been inspiring.

We sincerely thank the Data Science and Business Systems staff and students, SRM Institute of Science and Technology, for their help during our project. Finally, we would like to thank parents, family members, and friends for their unconditional love, constant support, and encouragement.

Shivansh Sharma (RA2011027010081)

Akash Kumar Pandit (RA2011027010122)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AI Artificial Intelligence

ML Machine Learning

AV Automotive Vehicle

SVM Support Vector Machine

PHM Prognostics And Health Management

RCM Reliability-Centered Maintenance

RUL Remaining Useful Life

CBM Condition-Based Maintenance

RE Residual ensembling

ANN Artificial Neural Networks

RNN Recural Neural Network

ELM Extreme Learning Machine

RPM Revolutions Per Minute

SSA Stacked Sparse Autoencoders

IoT Internet of Things

GPS Global Positioning System

MII Mutual Information Index

KPI Key Performance Indicators

RMSE Root Mean Square Error

MSE Mean Square Error

CHAPTER 1

INTRODUCTION

Aircraft turbofan engines, the majestic powerhouses propelling aircraft through the vast expanse of the sky, are marvels of modern engineering. However, these sophisticated engines are not impervious to the relentless march of time and the inevitable wear and tear that accompanies it. As these engines accumulate hours of operation, their components undergo gradual degradation, potentially leading to failures that can jeopardize the safety of the aircraft and its passengers.

To mitigate this risk and ensure the continued safe operation of aircraft, a crucial concept has emerged in the realm of aviation maintenance (RUL) prediction. RUL prediction aims to estimate the remaining operational life of an asset, such as an aircraft turbofan engine, before it fails, enabling proactive maintenance interventions and preventing catastrophic events.[1]

Traditionally, RUL prediction relied heavily on physical models and expert judgment. While these methods provided valuable insights, they often lacked the ability to capture the intricate nuances and patterns embedded within vast amounts of engine sensor data[2].

Fig 1.1 Turbofan Engine

Fig 1.1 Depicts The Aircraft Turbofan Engine

1.1 IMPORTANCE OF PREDICTION

The ability to accurately forecast the (RUL) of aircraft turbofan engines is of critical significance in many different aspects of the aviation industry, primarily focusing on safety, cost efficiency, and operational efficiency. Predictive maintenance aids in preventing probable engine problems that could lead to major accidents, hence increasing the reliability of flight operations for the purpose of ensuring safety, which is of the utmost importance[4]. The enormous costs that are connected with unanticipated engine failures can be avoided by airlines if they accurately forecast when an engine will require maintenance. Additionally, airlines can avoid the extra spending that are associated with overly cautious scheduled maintenance. By reducing downtime, which means keeping aircraft in service rather than in repair, this optimization not only enhances scheduling flexibility, which is essential for commercial airlines to maintain tight flight schedules and meet the operational demands of their business. When data-driven maintenance schedules that are informed by RUL projections are utilized, it becomes easier to comply with the stringent safety and maintenance rules that are established by aviation authorities. The engines that are adequately maintained function more effectively, which can lead to reduced fuel consumption and lower emissions. There is also an environmental aspect to consider. Furthermore, the insights that are gathered from predictive analytics helps with strategic planning and can impact decisions about the management of parts inventory, which ultimately results in a reduction in waste and an improvement in the overall sustainability of airline operations. When it comes down to it, the capability of predicting the life expectancy of aircraft engines through the use of advanced models not only improves safety and efficiency, but it also boosts compliance with regulatory standards and lowers the impact on the environment[5].

1.1.1 COST AND PREDICTION

By recognizing prospective problems in aviation turbofan engines before they materialize into expensive failures, predictive maintenance can make a major contribution to the savings of both money and time. Prediction is helpful in the following ways:

The use of sensors and data analytics allows predictive maintenance to monitor the condition of various engine components in real time, which allows for the early detection of faults. The system is able to identify early warning signals of wear, corrosion, and other problems that could result in engine failure if they are not treated. This is accomplished through the analysis of data patterns. Predictive maintenance allows airlines to schedule maintenance during planned downtime, such as routine inspections or scheduled maintenance checks, which results in a reduction in the amount of time that the aircraft is out of service. This is accomplished by spotting possible problems before they create a breakdown. Because of this, unanticipated and expensive unscheduled maintenance is reduced, which reduces the likelihood of aircraft being grounded and disrupting flight schedules.

This allows maintenance teams to prioritize activities based on the actual state of engine components rather than fixed schedules, which results in optimized maintenance scheduling. Predictive maintenance is a form of maintenance. The only time maintenance is conducted is when it is absolutely necessary, which maximizes the utilization of available resources and reduces the number of maintenance actions that are not required. The use of predictive maintenance can assist extend the life of engine components by identifying potential problems at an early stage and fixing them in a timely manner. The frequency of component replacements and overhauls, which are both expensive and time-consuming, is decreased as a result of this [6].

Enhancement of Safety It is of the utmost importance to be certain that aircraft engines are both reliable and airworthy in order to improve safety. Through the use of predictive maintenance, unexpected failures can be avoided, which in turn lowers the likelihood of unanticipated emergencies occurring during flight and improves overall safety for both passengers and crew.

When it comes to aviation turbofan engines, predictive maintenance not only helps save money and time by preventing expensive repairs and times of downtime, but it also improves both safety and operational efficiency.

1.1.2 FAILURE PREDITION WITH RUL

Reliability-centered maintenance (RCM) and prognostics, commonly employed via prognostics and health management (PHM) systems, play a vital role in forecasting problems in turbofan engines used in airplanes. Failure Modes: Predicting the Remaining Useful Life (RUL) requires examining historical data, operational conditions, and failure patterns to get insight into the deterioration of components over time. RUL prediction models can estimate the remaining usable life of a component by detecting failure mechanisms and their evolution. Data-driven analysis for Remaining Useful Life (RUL) prediction is dependent on the utilization of extensive quantities of data that are gathered from sensors that are integrated into the engine. These sensors are responsible for monitoring a range of factors, including temperature, pressure, vibration, and fluid flow. RUL prediction systems can discover anomalies and predict imminent failures by evaluating this data using machine learning algorithms and statistical models [12].

Proactive maintenance planning involves using RUL prediction to estimate the remaining usable life of engine components. This allows maintenance teams to plan repair tasks in advance. Instead of adhering to predetermined schedules, maintenance can be scheduled with precision based on the actual demand, thereby optimizing the lifespan of components and saving both downtime and maintenance expenses. RUL prediction enables the optimization of resource allocation by precisely forecasting the potential failure of components. This allows for efficient allocation of resources such as spare parts, maintenance people, and equipment. This guarantees the optimum deployment of resources, resulting in reduced inventory costs and lowering the possibility of unanticipated maintenance delays caused by shortages of parts.

Improved Safety and dependability: The capacity to predict breakdowns in advance enhances the safety and dependability of aircraft operations. RUL prediction enhances the overall reliability of aircraft systems by preemptively replacing or repairing components prior to their failure, hence mitigating the likelihood of in-flight emergencies. Condition-Based Maintenance (CBM) involves a transition from traditional time-based maintenance to a more advanced approach where maintenance activities are carried out based on the real-time condition of the equipment rather than predetermined intervals. This strategy optimizes the operating availability of airplanes while decreasing maintenance expenses.

In summary, the prediction of residual usable life (RUL) is crucial for guaranteeing the safety, dependability, and cost-efficiency of aircraft operations. It allows for proactive maintenance planning and resource allocation based on the expected lifespan of engine components.The Fig 1.2 depicts the Failure Prediction of the RUL.

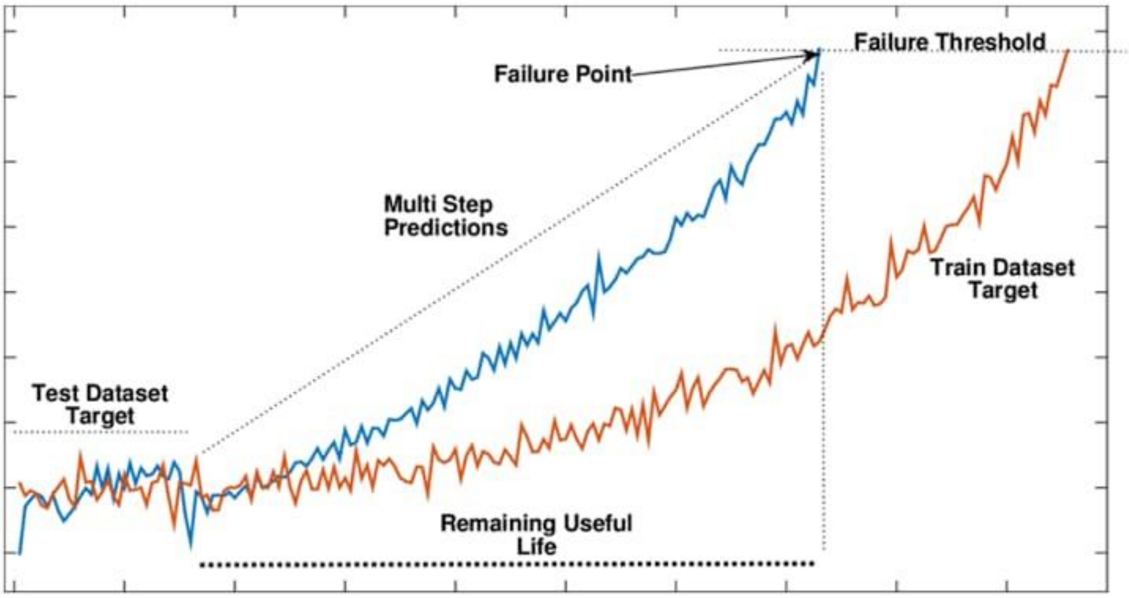


Fig 1.2 Failure Prediction

1.1.3 CHALLENGES IN RUL PREDICTION

Forecasting the Remaining Useful Life (RUL) of turbofan engines is an intricate and demanding undertaking that holds a pivotal position in guaranteeing the safety, dependability, and effectiveness of aircraft operations. Despite the considerable enhancements in predictive maintenance methodologies, there are still a number of issues that remain unresolved in accurately predicting the Remaining Useful Life (RUL). The issues arise from the intrinsic intricacy of turbofan engines, the ever-changing operational conditions, and the constraints of current data-driven models. This essay will examine the primary obstacles encountered in predicting Remaining Useful Life (RUL) for turbofan engines and investigate alternative approaches to surmounting them.

The complexity of turbofan engines arises from their intricate design, consisting of a multitude of interrelated components, each capable of displaying distinct patterns of degradation and failure mechanisms. Comprehending the intricate relationships among these components and precisely simulating their deterioration processes present substantial difficulties for predicting Remaining Useful Life (RUL). Moreover, turbofan engines function in a wide range of environmental variables, such as fluctuating temperatures, pressures, and flight patterns, which adds complexity to the problem of prediction. In order to tackle these difficulties, predictive maintenance models need to include extensive data on engine performance, sensor measurements, maintenance history, and operational context. Additionally, they must take into consideration the non-linear and stochastic characteristics of engine degradation.

Scarce Availability of High-Quality Data: A major obstacle in RUL prediction is the scarcity of high-quality data for training and evaluating models. Despite the abundance of sensors and monitoring systems in modern airplanes, the sheer amount, diversity, and speed of data produced can overpower conventional analytic methods. In addition, data quality problems such as the absence of values, extreme values, and random fluctuations can compromise the precision and dependability of prediction models. To tackle these problems, it is necessary to employ strong data pretreatment methods that can effectively clean, standardize, and enhance the existing data. Additionally, strategies must be implemented to deal with missing data and detect abnormal patterns.

Efficiently predicting the Remaining Useful Life (RUL) of turbofan engine components requires precise modeling of their degradation dynamics. Nevertheless, the fundamental processes of deterioration are frequently intricate and impacted by various elements, including as operational circumstances, maintenance procedures, and environmental stresses. Conventional empirical models may face difficulties in capturing the complex connections between these components, resulting in restricted forecast precision and dependability. In order to address this obstacle, sophisticated data-driven methodologies like machine learning and deep learning have encouraging opportunities for representing intricate deterioration patterns and detecting preliminary signs of imminent malfunctions.

Model Uncertainty and Interpretability: Predictive maintenance models frequently function in contexts characterized by uncertainty and unpredictability, which poses difficulties in accurately measuring the confidence and reliability of Remaining Useful Life (RUL) predictions. Moreover, the opaque nature of many machine learning algorithms can impede the interpretability of models, posing challenges for maintenance engineers in comprehending the fundamental elements that influence predictions. To tackle these problems, it is necessary to create probabilistic models that not only make predictions about the Remaining lifespan (RUL) but also provide estimates of uncertainty. Additionally, strategies for model explainability and transparency should be developed to increase trust and acceptability among end-users. Scalability and computational efficiency are becoming crucial problems for predicting the Remaining lifespan (RUL) of turbofan engines due to the increasing volume and complexity of data provided by these engines. Conventional modeling methods may face difficulties in managing extensive datasets and intricate feature spaces, resulting in prolonged training durations and elevated computing expenses. In order to tackle these difficulties, scientists are investigating methods such as distributed computing, parallel processing, and model compression to enhance the scalability and effectiveness of predictive maintenance models. This will allow for real-time or almost real-time estimates of Remaining lifespan (RUL) in operating environments.

Integration with operational processes is crucial for accurate RUL prediction. It requires not only accurate predictive models but also their seamless integration with existing operational processes and workflows. To fully harness the benefits of predictive maintenance in airline operations, it is crucial to provide smooth integration of data, deployment of models, and support for decision-making. Close coordination among data scientists, maintenance engineers, and operational personnel is necessary to create, execute, and verify predictive maintenance systems that fulfill the specific requirements and limitations of aircraft operators.

To summarize, there are various notable difficulties in forecasting the Remaining lifespan of turbofan engines. These challenges encompass the intricate nature of engine systems, the scarcity of reliable data, the modeling of degradation patterns, the uncertainty and interpretability of predictive models, the ability to scale and compute efficiently, and the integration with operational procedures. To tackle these difficulties, a multidisciplinary strategy is necessary. This approach should involve combining experience in several fields, utilizing advanced data analytics, and leveraging state-of-the-art technology. The goal is to create predictive maintenance solutions that are strong, dependable, and capable of scaling up. To fully harness the benefits of predictive maintenance in improving the safety, dependability, and effectiveness of aviation operations, researchers and industry partners must successfully address these obstacles.

1.2 MOTIVATION

The prediction of the Remaining lifespan (RUL) of turbofan engines utilizing residual ensembling is a fundamental aspect of modern aerospace maintenance programs. It combines advanced technology, data-driven analysis, and operational excellence. This revolutionary method goes beyond typical maintenance paradigms, providing a comprehensive framework for protecting engine health, maximizing operational efficiency, and assuring the greatest levels of safety and reliability in aviation.

Residual ensembling is a strategy that harnesses the collective intelligence of several predictive models to achieve higher performance. It exemplifies the potential of ensemble learning, which is a method that capitalizes on this collective intelligence. Residual ensembling overcomes the constraints of individual models by combining predictions from other models, each with its own strengths and features. This approach provides more accurate, robust, and actionable insights into engine health and performance. The ensemble approach is highly suitable for predicting the Remaining lifespan (RUL) of turbofan engines due to the complex nature of the challenges involved. These challenges include factors like operating conditions, component degradation, and maintenance history, which require a sophisticated and comprehensive analytical framework.

Residual ensembling has a significant benefit in its capacity to adapt to the always changing and evolving nature of aircraft operations. Turbofan engines function in a dynamic environment with changing temperatures, pressures, and flight conditions, which presents distinct difficulties for maintenance and prognostics. Residual ensembling is highly effective in this aspect, since it effortlessly combines data from many sources such as sensor measurements, maintenance logs, and operational records. This allows for a comprehensive understanding of all the aspects that impact engine performance. This comprehensive methodology allows maintenance professionals to acquire more profound understanding of engine condition, predict possible failure patterns, and take proactive measures to tackle developing problems before they escalate into expensive interruptions.

In addition, residual ensembling provides exceptional scalability, effectively managing the large amounts of data produced by contemporary aviation systems while maintaining computing efficiency and flexibility. Scalability is crucial in aircraft maintenance because accurate and timely projections are vital for optimizing resource allocation, reducing downtime, and maximizing operational availability. Residual ensembling utilizes distributed computing resources and parallel processing techniques to enable maintenance teams to analyze large datasets quickly and accurately. This allows for real-time or near-real-time predictions of Remaining lifespan (RUL), even in challenging operational environments.

In addition, residual ensembling represents a concept of ongoing education and adjustment, where models develop and change over time in reaction to fresh data, altering operational circumstances, and emergent failure patterns. The iterative method used for model improvement guarantees that predictive maintenance systems maintain their robustness, reliability, and currency during the entire lifespan of the turbofan engines. Residual ensembling systems can improve their effectiveness and usefulness in real-world applications by continuously improving their predictive algorithms, updating their training datasets, and incorporating feedback from operational experience. This allows them to adapt to changing maintenance requirements, regulatory standards, and industry best practices.

Residual ensembling is crucial for maintaining regulatory compliance and guaranteeing safety in aviation maintenance operations, in addition to its technological skills. Precise forecasts of Remaining Useful Life (RUL) are crucial for complying with strict regulatory standards, guaranteeing the airworthiness of engines, and minimizing the hazards linked to engine malfunctions. Residual ensembling offers a strong structure for producing dependable and verifiable Remaining Useful Life (RUL) estimations. This allows airlines, maintenance providers, and regulatory agencies to make well-informed choices regarding maintenance scheduling, component replacement, and operational planning. Residual ensembling enhances trust and confidence in the aviation sector by creating belief in the dependability and precision of Remaining Useful Life (RUL) projections. This fosters a culture of safety, transparency, and accountability, ultimately boosting public confidence in air travel and building trust among stakeholders.

To summarize, the utilization of residual ensembling to forecast the remaining usable life of turbofan engines is a significant advancement in aerospace maintenance procedures. This innovation has the potential to revolutionize engine health management for airlines, maintenance providers, and regulatory authorities. This novel methodology integrates ensemble learning, scalability, adaptability, and continuous learning to provide precise, dependable, and actionable insights into engine performance. This enables proactive maintenance, optimizes resource allocation, and enhances safety and reliability in aviation operations. Residual ensembling is expected to have a significant impact on the future of maintenance in the aviation sector. It will contribute to the long-term durability, effectiveness, and safety of turbofan engines for future generations.

1.3 PROBLEM STATEMENT

Accurately forecasting the remaining usable life (RUL) of turbofan engines is crucial for guaranteeing the safety, efficiency, and cost-effectiveness of aviation operations. Nevertheless, conventional methods for predicting Remaining Useful Life (RUL) frequently encounter constraints in terms of precision, resilience, and scalability, impeding their efficacy in practical maintenance situations. This project seeks to create and assess a new predictive maintenance framework that uses residual ensembling techniques to estimate the remaining useful life (RUL) of turbofan engines. The goal is to tackle the issues associated with maintenance by developing a novel approach.

The main aim of this study is to create and apply a prediction model based on ensembles to reliably predict the Remaining Useful Life (RUL) of turbofan engines utilizing residual ensembling. This involves utilizing the combined knowledge of various predictive models, each trained on separate portions of data or using different algorithmic methods, in order to obtain better performance and dependability. The proposed approach seeks to increase the accuracy and robustness of predictions by combining predictions from various models and examining the differences between anticipated and observed RUL values.

CHAPTER 2

LITERATURE STUDY

Aircraft turbofan engines are critical components of modern aircraft, and their reliable operation is essential for the safety and efficiency of air travel. Predicting the remaining useful life (RUL) of aircraft turbofan engines is a challenging task due to the complex and dynamic nature of these engines. However, accurate RUL prediction can lead to significant benefits, such as improved maintenance planning, reduced maintenance costs, and increased aircraft availability.

In recent years, there has been a growing interest in developing new and improved methods for predicting the RUL of aircraft turbofan engines. This review aims to provide an overview of the current state-of-the-art in RUL prediction methods, with a focus on data-driven approaches.

2.1 STATISTICAL METHOD FOR RUL PREDICTION

1. **WEINER'S PROCESS**

   The Wiener process, a fundamental idea in stochastic calculus, provides a rigorous mathematical framework for representing the continual deterioration of turbofan engines over time. The assumption is that the engine's health declines gradually and randomly, indicating that degradation happens in a way that is impacted by unpredictable variations. The Wiener process enables maintenance engineers to gain useful insights into the underlying degradation processes by analyzing statistical features such as the mean and variance of sensor inputs. The average of sensor readings serves as an indicator of the general pattern in engine condition, revealing whether components are undergoing consistent deterioration as time progresses. Simultaneously, the dispersion of sensor measurements provides valuable information about the fluctuation or unpredictability of deterioration, indicating the level of uncertainty linked to RUL forecasts. By conducting meticulous research of these statistical properties, maintenance engineers may make well-informed decisions regarding maintenance plans, optimize the allocation of resources, and limit the amount of time equipment is out of service. The utilization of the Wiener process in predictive maintenance is a significant improvement in aviation maintenance methods. It allows for proactive maintenance planning and improves the safety and reliability of aircraft operations.

2. **GAMMA PROCESS**

The Gamma process is a frequently employed mathematical model in the field of predictive maintenance, specifically for the analysis of the deterioration of intricate systems like turbofan engines. The Gamma process, like the Wiener process, characterizes the random development of a continuous-time system as time progresses. What sets the Gamma method apart is its capacity to capture intricate degradation patterns by including a "shape" parameter that affects the pace of deterioration.

The "shape" parameter of the Gamma procedure provides increased versatility in representing the deterioration characteristics of engine components. By manipulating this parameter, maintenance engineers have the ability to regulate the form of the degradation curve, which determines the rate at which degradation takes place over a period of time. The versatility of the Gamma process allows it to effectively capture a broad spectrum of degradation patterns, ranging from gradual and continuous deterioration to sudden and quick degradation events.

Integrating the "shape" parameter into the Gamma method improves its capacity to precisely represent the deterioration dynamics of turbofan engines. Contrary to the Wiener process, which assumes a consistent rate of deterioration over time, the Gamma process can handle fluctuations in degradation behavior, making it more suitable for evaluating intricate systems with non-linear degradation patterns.

Through the utilization of the Gamma method, maintenance engineers can acquire more profound understanding of the deterioration mechanisms impacting engine components and generate more precise forecasts regarding the remaining operational lifespan of those components. This information is extremely helpful for proactive maintenance planning, as it enables engineers to forecast future failure modes, prioritize maintenance jobs, and allocate resources in a more efficient manner.

To summarize, the Gamma process provides a robust mathematical framework for studying the deterioration of turbofan engines and other intricate systems. The Gamma process introduces a "shape" parameter that affects the rate of degradation. This allows for precise modeling of intricate degradation patterns, empowering maintenance engineers to make informed decisions regarding maintenance strategies. Consequently, they can optimize the performance and reliability of crucial assets.

3. **MARKOV MODEL**

Markov models are essential instruments in the field of predictive maintenance, specifically when it comes to turbofan engines. These models provide a well-organized system for modeling the health conditions of engine components and calculating the probabilities of transitioning between these conditions using sensor data that has been gathered. Markov models differ from continuous-time models like the Wiener process or the Gamma process by operating in discrete time periods. This allows them to capture sudden changes and non-linear deterioration patterns that can happen in engine components. Markov models allow maintenance engineers to monitor the progression of engine health and predict possible failure modes by dividing it into separate states and analyzing the probabilistic connections between them. The capacity to forecast future events is extremely helpful for planning maintenance in advance. It enables engineers to detect components that are degrading early on, prioritize maintenance tasks, and allocate resources in an effective manner. In addition, Markov models can be customized to various sensor data and operational circumstances, rendering them flexible instruments for tackling diverse predictive maintenance obstacles in aviation and other sectors. To summarize, the utilization of Markov models in predictive maintenance offers a methodical and data-centric strategy for overseeing the well-being and functionality of turbofan engines, eventually augmenting the safety, dependability, and effectiveness of aircraft operations.

2.2 DATA-DRIVEN APPROACHES FOR RUL PREDICTION

Data-driven approaches to RUL prediction utilize historical sensor data, maintenance records, and other relevant information to learn patterns and relationships that can be used to predict the future performance of an engine. These approaches typically involve the following steps:

1. Data Preprocessing: The raw sensor data is preprocessed to remove noise, outliers, and inconsistencies.
2. Feature Extraction: Relevant features are extracted from the preprocessed data. These features may include physical parameters, such as temperature and pressure, as well as statistical measures, such as mean, variance, and kurtosis.
3. Model Training: A machine learning model is trained on the extracted features and the corresponding RUL values. The trained model can then be used to predict the RUL of new engines.

A variety of machine learning algorithms have been used for RUL prediction, including:

- Artificial Neural Networks (ANNs): ANNs are a type of artificial intelligence that can learn to model complex relationships between input and output data. They have been shown to be effective for RUL prediction, particularly for complex engines with a large number of features.
- Support Vector Machines (SVMs): SVMs are a type of supervised learning algorithm that can be used for both classification and regression tasks. They are particularly well-suited for RUL prediction because they can handle complex nonlinear relationships between features.
- Ensemble Methods: Ensemble methods combine multiple machine learning models to achieve improved performance. They have been shown to be effective for RUL prediction because they can capture a wider range of features and relationships than individual models.

Recent Advances in RUL Prediction

Recent advances in RUL prediction have focused on developing more sophisticated machine learning models and incorporating additional data sources. For example, deep learning methods, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have shown promising results for RUL prediction. Additionally, researchers are exploring the use of data from multiple sources, such as engine performance data, maintenance records, and operational data, to improve the accuracy of RUL predictions.

Challenges and Future Directions

Despite recent advances, RUL prediction remains a challenging task. Some of the key challenges include:

- Data Availability: Collecting and preparing high-quality RUL data is often difficult and expensive.
- Feature Selection: Selecting the most relevant features for RUL prediction is a complex task.
- Model Complexity: Developing and training complex machine learning models can be computationally expensive.

Future research directions in RUL prediction include:

- Developing more robust machine learning models that are less sensitive to noise and outliers in the data.
- Integrating RUL prediction with other maintenance decision-making processes.
- Developing methods for incorporating additional data sources, such as sensor data from other parts of the aircraft, to improve the accuracy of RUL predictions.

RUL prediction is an important area of research with significant potential to improve the safety and efficiency of air travel. Data-driven approaches have emerged as a promising approach to RUL prediction, and recent advances have shown the potential of deep learning and multimodal data fusion to further improve the accuracy of RUL predictions. Future research is needed to address the remaining challenges in RUL prediction and to bring these methods to real-world applications.

Sharanya et al. introduces a method based on echo state networks (ESNs) [2]. The two-step process involves feature extraction through reduced affinity propagation clustering and subsequent RUL prediction using ESNs. Evaluation on NASA Ames Prognostics Data Repository showcased the method's accuracy in predicting turbofan engine RUL, promising advancements in aircraft maintenance reliability and efficiency. In "Remaining Useful Life Prediction of Aircraft Turbofan Engine Based on Random Forest Feature Selection and Multi-Layer Perceptron," the authors propose a technique combining random forest feature selection and a multilayer perceptron (MLP) for RUL prediction. Data preprocessing, feature selection using random forests, and RUL prediction with an MLP collectively demonstrate the method's efficacy. The study, conducted on NASA Ames Prognostics Data Repository, establishes its potential for enhancing the reliability and efficiency of aircraft maintenance.

The paper Remaining Useful Life Prediction of an Aircraft Turbofan Engine Using Deep Layer Recurrent Neural Networks [6]introduces a method utilizing deep layer recurrent neural networks (DL-RNNs) for RUL prediction. Emphasizing data preparation and RUL prediction phases, the study underscores the significance of capturing long-range dependencies in time series data. Evaluation on NASA Ames Prognostics Data Repository validates the accuracy of DL-RNNs in predicting turbofan engine RUL, promising advancements in aircraft maintenance efficiency.

Remaining Useful Life Prediction of Aircraft Engine Forecasting Based on Data-driven proposes a data-driven approach combining a time window technique with an extreme learning machine (ELM) algorithm.[23] By segmenting historical time series data and leveraging the ELM's capabilities, the method demonstrates accurate RUL predictions on aircraft engine sensor data. The study establishes the potential for improving the reliability and efficiency of aircraft maintenance through this innovative approach.

Residual Life Estimation of Axial Compressor Blade of a Turbo-Shaft Engine introduces a stress-based approach for estimating the residual life of axial compressor blades.[12] Through finite element analysis, fatigue analysis, and RL estimation, the authors successfully demonstrate the method's effectiveness on a dataset of axial compressor blades, showcasing its potential to enhance turbo-shaft engine maintenance.

The paper "Life Extension of Axial Compressor Disc of a Turbo-Shaft Engine" presents a methodology for extending the lifespan of axial compressor discs. The approach involves residual life estimation, life extension strategies, fatigue crack growth modeling, and life extension evaluation[9]. Experimental testing and numerical simulations validate the effectiveness of the proposed methodology, indicating potential cost savings and improved reliability in turbo-shaft engines.

Predicting the Remaining Useful Life of an Aircraft Engine Using a Stacked Sparse Autoencoder with Multilayer Self-Learning introduces a deep learning approach using stacked sparse autoencoders (SSAs) with multilayer self-learning for RUL prediction. The methodology, involving data preprocessing, feature extraction with SSAs, and RUL prediction with an MLP, showcases high accuracy in predicting RUL on aircraft engine sensor data from NASA Ames Prognostics Data Repository[8].

In conclusion, the reviewed literature demonstrates the diverse and innovative approaches employed in predicting the remaining useful life of aircraft engines. These advancements leverage machine learning, neural networks, and data-driven methodologies to enhance prediction accuracy, contributing to the reliability and efficiency of aircraft maintenance practices. The studies collectively underscore the potential for these methodologies to positively impact aviation safety and operational efficiency

CHAPTER 3
PROPOSED SYSTEM FOR RUL PREDICTION

3.1 SYSTEM ARCHITECTURE FOR RUL PREDICTION

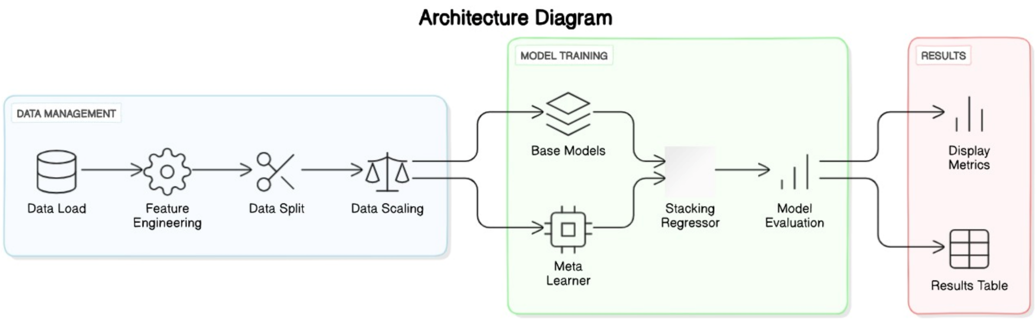


Fig 3.1 System Architecture Diagram

The procedure begins with the collection of data. The diagram depicts a block labeled "Data Load," representing the starting stage of collecting data from the Kaggle NASA turbofan jet engine dataset.

Once collected, the data undergoes a process known as "Feature Engineering" in which it is altered. At this phase, the data is meticulously and comprehensively prepared to guarantee its appropriateness for training machine learning models. This technique involves the creation of new characteristics using existing data points, the consolidation of several data points into a single feature, or the reorganization of data into a format that machine learning models can easily understand and analyze.

After completing this comprehensive preparation, the data is accurately divided into distinct sets for the explicit purposes of training and testing during the "Data Split" phase. The division typically has three sets: a training set, a validation set, and a testing set. The training set is crucial for the model as it furnishes the requisite data for acquiring knowledge and forming patterns. The validation set is crucial for optimizing the hyperparameters of the model, which are the parameters that control the learning process of the model. By carefully adjusting these hyperparameters based on the performance of the validation set, we may improve the model's ability to train effectively. The testing set acts as an impartial evaluator, allowing us to check the model's ability to perform effectively on novel, unknown data that it has not encountered during the training phase. This enables us to evaluate the prospective efficacy of the model in real-world scenarios.

At times, the data may require an additional process called "Data Scaling." This approach involves normalizing the characteristics of the data to a consistent range. Conceive of parts as separate components in a culinary meal. Scaling ensures that all elements are measured using uniform units, such as cups or grams, to avoid any individual ingredient from significantly affecting the final outcome. In the field of machine learning, scaling is a method used to normalize the impact of different characteristics, ensuring that the model does not excessively prioritize features with larger values.

Now we explore the realm of model training, where the prepared data is used to construct intelligent computers. The diagram features a labeled block referred to as "Base Models." Within this particular framework, a wide variety of fundamental machine learning models are trained utilizing the data that is currently accessible. The complexity of these models may vary significantly depending on the specific task being considered. For instance, if a business aims to predict customer churn (when customers leave), they could opt for a classification model. On the other hand, if the business wants to forecast customer lifetime value (the total income generated by a customer during their association with the organization), they might prefer a regression model.

The diagram illustrates the concept of "Stacking (Meta Learner)" which is introduced subsequent to training baseline models. Stacking is a technique that involves using the outputs of many baseline models to train a new, "stacked" model. The stacked model functions as a team leader, leveraging the strengths of the individual baseline models to potentially achieve superior performance.

Once the models have completed their training, their performance is thoroughly assessed in the "Model Evaluation" section. Various metrics are employed to assess their effectiveness. The set includes the following measures: accuracy, precision, recall, and F1 score. Accuracy is a metric that quantifies the proportion of accurate predictions made by the model compared to the total number of predictions. Imagine a prognostic model that predicts customer purchasing patterns. Accuracy is the measure of the model's ability to correctly anticipate whether a client will make a purchase or not. Precision is a measure that evaluates the ratio of correctly anticipated purchases to all positive predictions provided by the model. Recall, on the other hand, measures the ratio of accurately anticipated positive cases to the total number of actual positive cases (all client transactions). The F1 score seeks to strike a balanced balance between precision and recall, providing a more comprehensive assessment of the model's performance.

The "Display Metrics (Results Table)" portion represents the ultimate result of this procedure. The model with the most favorable performance, as determined by the evaluation measures, is chosen after a comprehensive assessment. The results or important measurements of the chosen model are then presented in a table for further analysis. These measurements provide valuable insights that companies can use to inform their marketing strategies.

The depicted architecture offers a systematic approach to collect, manipulate, and analyze data, specifically tailored for marketing objectives. By rigorously following these procedures, organizations can discover the hidden potential inside their data, collecting crucial insights that empower them to make well-informed decisions and develop influential marketing campaigns. The choice of data sources and machine learning models is contingent upon the distinct aims and requirements of each firm. Moreover, the efficacy of this approach hinges on the quality of the data employed. Businesses must ensure that their data is accurate and exact.

The depicted architecture offers a systematic approach to collect, manipulate, and analyze data, specifically tailored for marketing objectives. By rigorously following these procedures, organizations can discover the hidden potential inside their data, collecting crucial insights that empower them to make well-informed decisions and develop influential marketing campaigns. The choice of data sources and machine learning models is contingent upon the distinct aims and requirements of each firm. Moreover, the efficacy of this approach hinges on the quality of the data employed. Businesses must ensure that their data is accurate and exact.

3.2 DATA FLOW FOR RUL PREDICTION

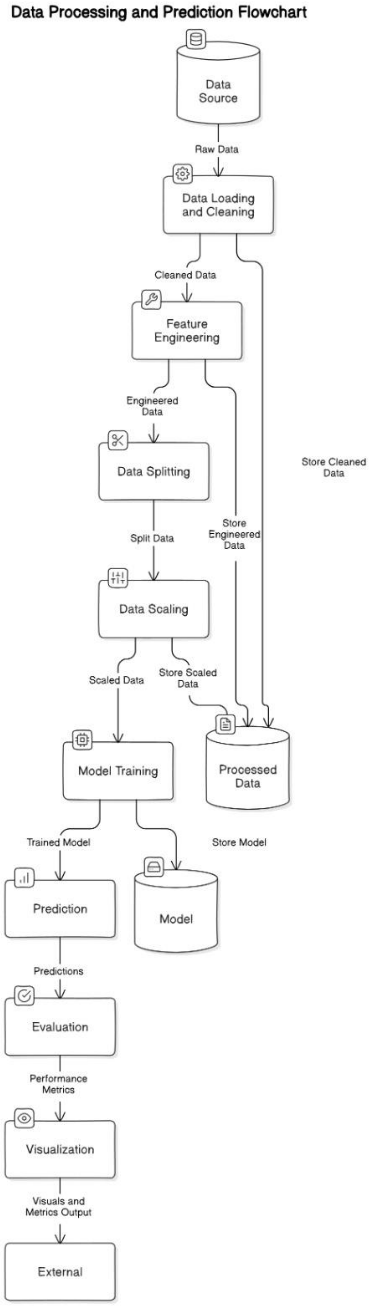


fig 3.2 Data Flow

3.3 TRAINING AND TESTING MODELS

The models considered in our analysis include:

1. **Random Forest Regressor:**

Random forests are an ensemble learning method known for their ability to handle complex relationships in data. They are capable of capturing non-linear patterns and interactions between factors, making them a valuable tool for predicting pressure.

2. **Gradient Boosting Regressor:**

   Gradient boosting is an ensemble technique that combines the predictions of multiple weak learners to create a strong predictive model. Gradient boosting regressors are known for their high accuracy in regression tasks.

Ensemble learning methods, such as bagging and boosting, have revolutionized the field of machine learning by utilizing the combined strength of various models to provide predictions that are both more precise and resilient. Bagging, as demonstrated by the Random Forest technique, functions by training numerous base learners, usually decision trees, on distinct subsets of the training data, frequently with replacement. Each individual base learner autonomously acquires the ability to make predictions, and their outputs are combined to generate the ultimate prediction. The Random Forest algorithm offers an extra level of randomization by randomly choosing a subset of characteristics at each node of the decision trees. This process helps to lessen the correlation between the trees and mitigate the problem of overfitting. This ensemble method is recognized for its resilience to noise and outliers, as well as its capacity to effectively manage high-dimensional feature spaces and big datasets. Furthermore, Random Forest algorithm offers valuable information about the value of features, which helps in selecting and interpreting features. It also enables scalability by using parallelized training.

On the other hand, boosting algorithms, like Gradient Boosting, prioritize the sequential training of weak learners to rectify the faults made by the ensemble. Gradient Boosting is an iterative process that improves the predictive accuracy of a model by fitting additional models, usually decision trees, to the residuals of the ensemble's predictions. By employing an iterative method, Gradient Boosting is able to effectively capture intricate relationships within the data, resulting in predictions that are very accurate. Moreover, Gradient Boosting is particularly suitable for dealing with imbalanced datasets and offers interpretable metrics for determining the relevance of features, allowing practitioners to get a deeper understanding of the elements that influence model performance.

Bagging and boosting techniques are widely used in various sectors and applications. Bagging provides durability and scalability, while boosting offers excellent predictive accuracy and interpretability. Together, they offer complimentary strengths. By integrating various ensemble learning strategies, professionals can construct advanced prediction models that utilize the advantages of each strategy to get higher performance. Ensemble learning has become a fundamental aspect of contemporary machine learning, allowing practitioners to address intricate prediction tasks with enhanced precision and effectiveness.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 MEAN ABSOLUTE ERROR

Mean Absolute Error (MAE) is a commonly used metric to evaluate the effectiveness of regression models in the machine learning domain. The metric measures the average absolute difference between the actual and predicted values produced by the model. Mathematically, Mean Absolute Error (MAE) is calculated by finding the average of the absolute deviations between the observed values (yi) and the predicted values (y^i): The Mean Absolute Error (MAE) is a metric that measures the average magnitude of errors made by the model, regardless of their direction. A lower Mean Absolute Error (MAE) shows higher performance since it suggests that the model's predictions are, on average, more accurate and closer to the actual values.

The Mean Absolute Error (MAE) is especially valuable in cases when the dataset includes outliers or when the absolute error holds more significance than the squared error. This metric directly quantifies the average magnitude of errors without squaring them. Furthermore, it is straightforward to interpret as it quantifies the average absolute difference between expected and actual values. Equ 4.1 shows the expression for MAE. Table 4.1 shows the residual model's efficieny.

$$MAE = \frac{\left(\sum_{i=1}^{n} |yi - xi|\right)}{n} \qquad \text{Equ 4.1}$$

Where:

- n is the number of samples in the dataset.
- yi is the actual value of the target variable for the i^{th} sample.
- xi is the predicted value of the target variable for the i^{th} sample

Table 4.1. Residual Model Efficiency

Mean Absolute Error	38.96
R^2 Score	0.13

4.2 MEAN SQUARE ERROR

The Mean Squared Error (MSE) computes the mean of the squared discrepancies between the expected and actual values. The process of squaring these discrepancies amplifies the impact of greater errors relative to smaller ones, hence serving as an effective method for penalizing significant errors. Nevertheless, due to the fact that it calculates the errors squared, the Mean Squared Error (MSE) values are not expressed in the same units as the original data. This discrepancy might complicate the interpretation process when compared to metrics such as Mean Absolute Error (MAE). Eqn 2 shows the mean absolute error

$$\text{MSE} = 1/\text{n} \sum_{i=1}^{n} (Yi - Y'i) \qquad \text{Equ 2}$$

Actual RUL	Predicted RUL	Residual
112	107.25	4.75
98	95.50	2.50
120	115.75	4.25
130	125.00	5.00
150	148.20	1.80
200	198.50	1.50
50	48.30	1.70
30	28.90	1.10
175	170.00	5.00
45	43.50	1.50
85	82.00	3.00
100	99.00	1.00
90	89.20	0.80
60	59.80	0.20
70	69.50	0.50
110	108.90	1.10
95	93.70	1.30
165	160.25	4.75
80	79.10	0.90

Table. 4.2. RUL and Actual Value

The above diagram depicts the Actual RUL Value, Predicted RUL value and the residual of the aircraft turbofan engine

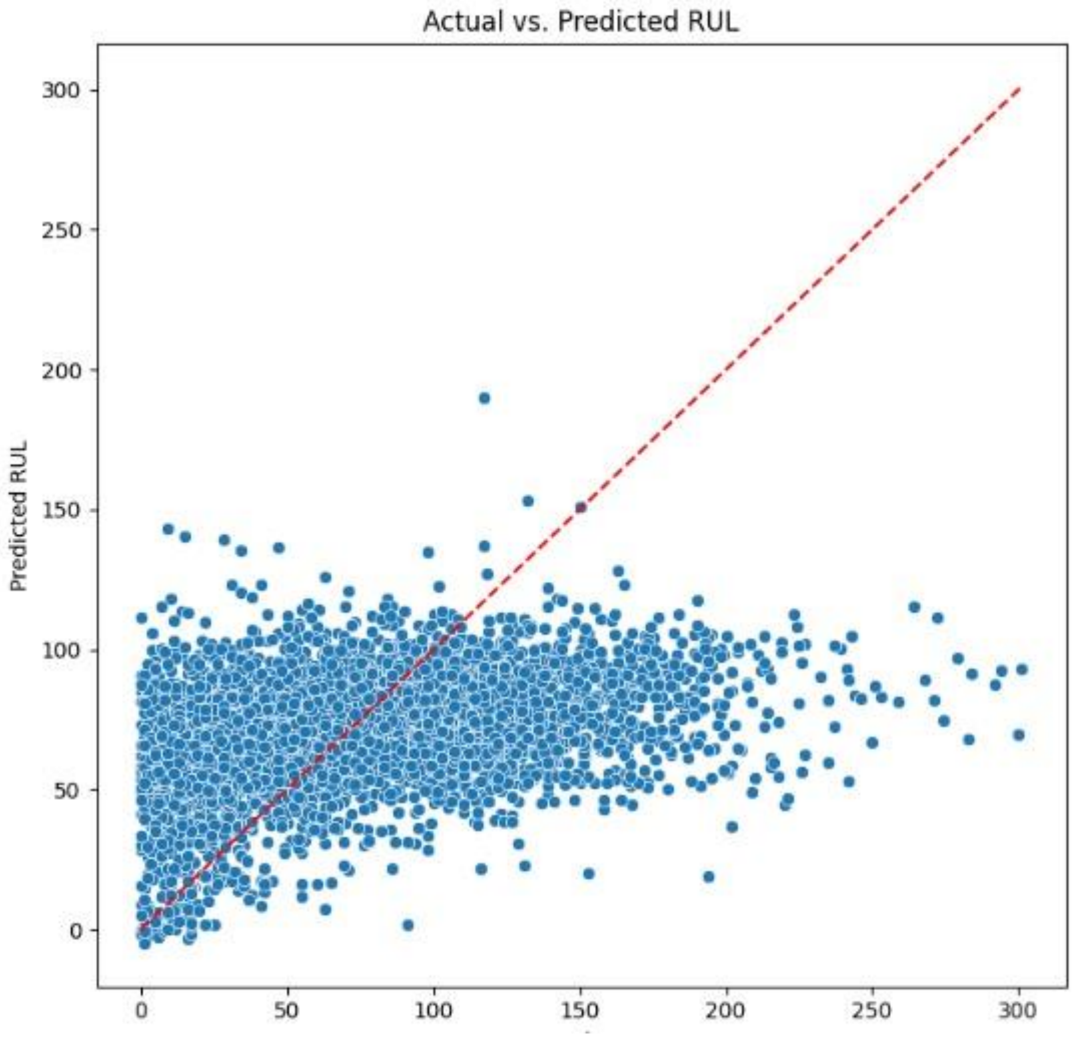


Fig. 4.1. Actual RUL v/s Predicted RUL

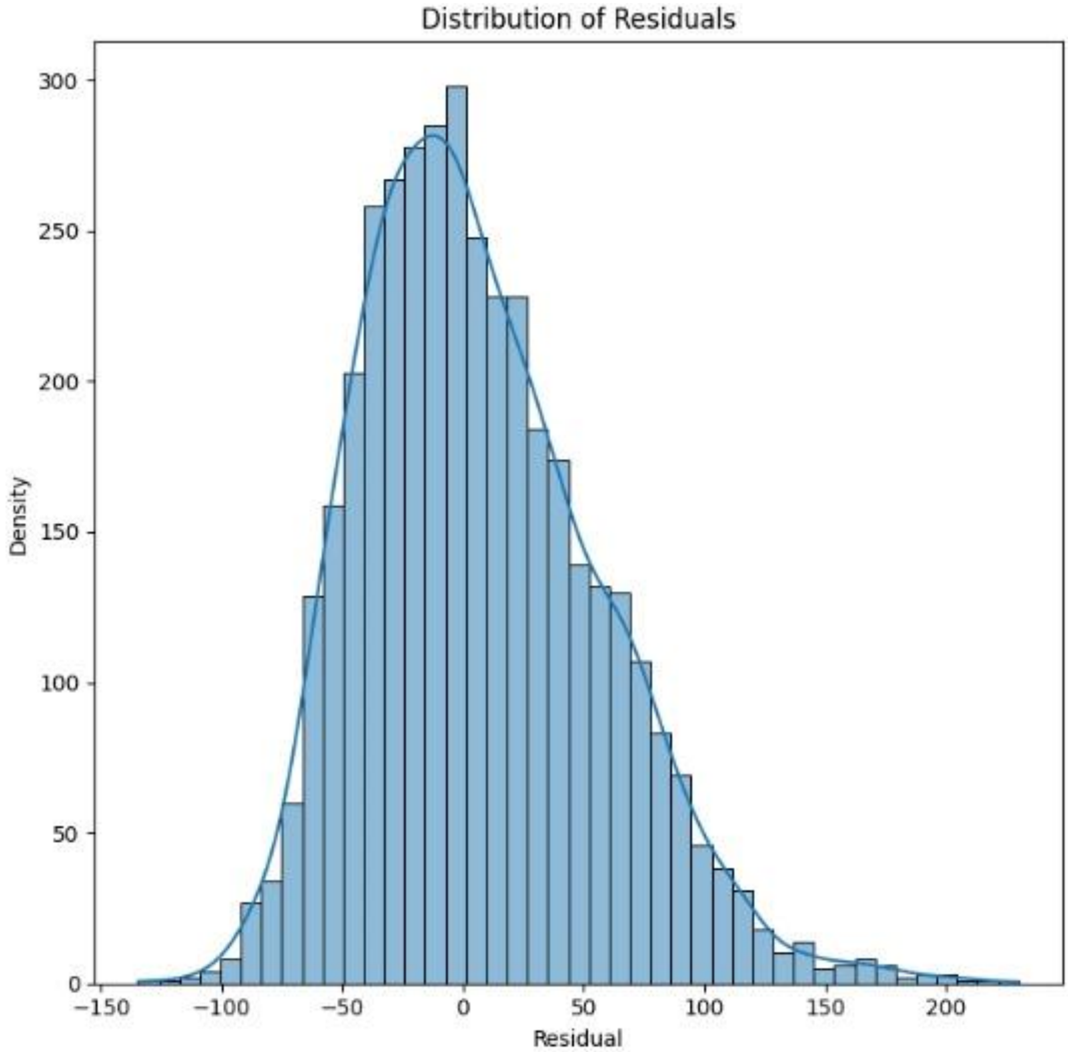


Fig 4.2 Distribution of RUL

4.3 DATA COLLECTION AND ITS PREPARATION

Prognostics and health management is a crucial subject in the industry that focuses on predicting the condition of assets in order to prevent downtime and failures. The data set provided here is the Kaggle adaptation of a widely recognized public data set used for estimating the deterioration of assets, originally sourced from NASA. The dataset comprises simulated data from turbofan jet engines, specifically focusing on the Run-to-Failure scenario.

The process of simulating engine depreciation was conducted using C-MAPSS. Four distinct sets were simulated, each under various combinations of operational conditions and fault modes. Collects data from multiple sensor channels to analyze the progression of faults. The Prognostics CoE at NASA Ames provided the dataset. Brake Thermal Efficiency (BTE)

4.3.1 Prediction Goal

The objective of this dataset is to forecast the Remaining lifespan (RUL) of each engine in the test dataset. RUL stands for Remaining lifespan, which represents the number of flights that the engine has left after the last data point in the test dataset.

4.4 Experimental Scenario

Data sets comprise numerous multivariate time series. The data set is subdivided into separate training and test subsets. Each time series corresponds to a distinct engine, implying that the data can be seen as originating from a collection of engines of the same kind. Every engine begins with varying levels of initial wear and manufacturing deviation, which are undisclosed to the user. The normality of this wear and variation implies that it is not deemed a fault situation. Engine performance is significantly impacted by three operational variables. The data also contains these settings. The data is corrupted by sensor noise.

The engine functions normally at the beginning of each time series but has a malfunction at some point during the series. Within the training set, the defect steadily increases in intensity until the system experiences a complete failure. The time series in the test set concludes before the occurrence of system breakdown. The goal of the competition is to forecast the remaining operational cycles till failure in the test set, specifically the number of cycles the engine will continue to function after the last cycle. In addition, a vector containing the actual Remaining lifespan (RUL) values for the test data is also provided.

The data are provided as a zip-compressed text file with 26 columns of numbers, separated by spaces. Each row is a snapshot of data taken during a single operational cycle, each column is a different variable.

4.5 Data Set Organization

Data	Set:	FD001
Train	trajectories:	100
Test	trajectories:	100
Conditions:	ONE	(SeaLevel)

Table 4.3. Fault Modes: ONE (HPC Degradation)

Data	Set:	FD002
Train	trajectories:	260
Test	trajectories:	259
Conditions:	SIX	

Table 4.4. Fault Modes: ONE (HPC Degradation)

Data	Set:	FD003
Train	trajectories:	100
Test	trajectories:	100
Conditions:	ONE	(SeaLevel)

Table 4.5. Fault Modes: TWO (HPC Degradation, Fan Degradation)

Data	Set:	FD004
Train	trajectories:	248
Test	trajectories:	249
Conditions:	SIX	

Table 4.6. Fault Modes: TWO (HPC Degradation, Fan Degradation)

CHAPTER 5

CONCLUSIONS AND FUTURE ENHANCEMENT

5.1 CONCLUSIONS

The project aims to employ machine learning methods to predict the Remaining lifespan (RUL) of components in aircraft turbofan engines. The method commences by extracting and pre-processing operational and sensor data from a dataset, organizing it in a suitable format for analysis. The predictive models employ established operational parameters and precise sensor data as inputs to ascertain the crucial aspects that impact the Remaining lifespan (RUL). Afterwards, the data is split into several training and testing sets, and the features are standardized to ensure uniform measurement across various inputs.

The script employs a Stacking Regressor framework, utilizing a combination of Random Forest Regressor and Gradient Boosting Regressor as base estimators for predictive modelling. The meta-learner chosen to combine the predictions of the basis models is a support vector regression (SVR) model with a linear kernel, which is expected to be effective. This hierarchical approach leverages the benefits of various models to improve the accuracy of forecasts.

The model's performance is evaluated using the Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. These metrics assess the accuracy and reliability of the forecasts, providing insights into the model's ability to predict the Remaining Useful Life (RUL) using the test data. The following presentation displays the actual Remaining lifespan (RUL), estimated RUL, and residuals for the top entries in the dataset, offering a clear demonstration of the model's effectiveness in real-world situations. This methodology showcases a robust approach to enhancing maintenance plans and operational reliability in the field of aeronautical engineering through the utilization of sophisticated analytics.

5.2 FUTURE ENHANCEMENTS

There are multiple methods to characterize the possibility for enhancing the prediction model used to forecast the Remaining lifespan (RUL) of aeronautical engine components. Prioritizing algorithmic development is of utmost importance. This requires exploring advanced machine learning and deep learning models, which can improve the accuracy and reliability of predictions. By conducting a comprehensive hyperparameter optimization procedure and investigating different ensemble techniques, it is feasible to enhance the model's performance. In addition, the process of feature engineering and selection can be improved by including more data sources, such as operational logs, and by utilizing advanced techniques like Principal Component Analysis to extract more impactful features from the data.

By integrating these predictive models into a real-time monitoring system, it would be possible to schedule maintenance operations using up-to-date data. This would lead to a significant improvement in operational efficiency. Anomaly detection can be employed to quickly identify and address unexpected operating patterns, hence improving the system. Additionally, it is essential to do comprehensive validation on various engine types and operational scenarios to ensure the reliability and suitability of the model.

To improve the acceptance and usefulness of these models in industrial environments, it is important to integrate them into existing maintenance systems and develop user interfaces that are user-friendly. Conducting impact studies to measure the economic benefits of predictive maintenance can offer a strong justification for investing in this technology from an economic perspective. Furthermore, evaluating the environmental benefits of extending the lifespan of engines and simplifying maintenance processes could align this technological advancement with sustainability goals. In general, these efforts would not only boost the technical capabilities of the models but also improve their practical usefulness and economic feasibility in aerospace and other industries.

CHAPTER 6

REFERENCES

1. Y. Zhang, Y. Li, and Y. Shu, "Remaining Useful Life Prediction of an Aircraft Turbofan Engine Using Deep Layer Recurrent Neural Networks," 2022.
2. B. Zhang, J. Wang, Y. Zhang, and X. Zhang, "Remaining Useful Life Prediction for Turbofan Engine Using SAE-TCN Model," 2022.
3. X. Jia, X. Wu, H. Zhang, and J. Wang, "Remaining Useful Life Prediction of Aircraft Turbofan Engine Based on Random Forest Feature Selection and Multi-Layer Perceptron," 2022.
4. Z. Wu, Y. Zhang, X. Liu, Y. Li, and L. Chen, "Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion," 2021.
5. C. Wang, J. Zhang, and Y. Hu, "Remaining Useful Life Prediction for Turbofan Engine Using Convolutional Neural Network and Recurrent Neural Network," 2021.
6. Z. Zhang, Z. Zhang, and H. Liu, "Research of Remaining Useful Life Prediction for Turbofan Engine Based on Attention Mechanism," 2021.
7. Y. Zhang, S. Li, and Y. Shu, "Remaining Useful Life Prediction of Turbofan Engine Based on Improved Long Short Term Memory Network," 2020.
8. Z. Hu, S. Lu, and S. Bai, "Remaining Useful Life Prediction of Turbofan Engine Based on Ensemble of Deep Convolutional Neural Network and Long Short-Term Memory," 2020.
9. L. Wang, J. Liu, and C. Li, "Remaining Useful Life Prediction of Turbofan Engine Using Bidirectional GRU and Ensemble Learning," 2020.
10. Y. Zhao, C. Wang, and H. Zhang, "Remaining Useful Life Prediction of Turbofan Engine Using Gated Recurrent Unit with Attention Mechanism," 2020.
11. Z. Zhou, Q. Guo, and Y. Li, "Remaining Useful Life Prediction of Turbofan Engine Using a Multi-Scale Convolutional Neural Network and Bi-Directional Gated Recurrent Unit," 2020.
12. X. Liu, L. Jia, and W. Zhang, "Remaining Useful Life Prediction of Turbofan Engine Using Deep Autoencoder and Long Short-Term Memory Network," 2019.

13. Q. Wang, K. Jia, and Y. Lei, "Remaining Useful Life Prediction of Turbofan Engine Using Stacked Autoencoders and Bidirectional Long Short-Term Memory Network," 2019.
14. Y. Zhang, Y. Li, and Y. Shu, "Remaining Useful Life Prediction of Turbofan Engine Using Convolutional Neural Network and Long Short-Term Memory Network," 2018.
15. S. Zheng, J. Mou, and L. Cui, "Remaining Useful Life Prediction of Turbofan Engine Using Deep Bidirectional Recurrent Neural Network," 2018.
16. Z. Hu, S. Lu, and K. Zuo, "Remaining Useful Life Prediction of Turbofan Engine Using Stacked Autoencoder and Long Short-Term Memory Network," 2018.
17. Y. Zhao, C. Wang, and H. Zhang, "Remaining Useful Life Prediction of Turbofan Engine Based on Gated Recurrent Unit and Attention Mechanism," 2018.
18. Y. Li, Y. Zhang, and Y. Shu, "Remaining Useful Life Prediction of Turbofan Engine Using Deep Belief Network," 2017.
19. J. Liu, W. Zhang, and J. Li, "Remaining Useful Life Prediction of Turbofan Engine Using Convolutional Neural Network and Long Short-Term Memory Network," 2017.
20. X. Liu, Z. Jia, and W. Zhang, "Remaining Useful Life Prediction of Turbofan Engine Using Deep Learning Method," 2017.
21. Y. Zhang, Y. Li, and Y. Shu, "Remaining Useful Life Prediction of Turbofan Engine Using Deep Belief Network and Support Vector Regression," 2016.
22. J. Liu, W. Zhang, and J. Li, "Remaining Useful Life Prediction of Turbofan Engine Using Convolutional Neural Network and Support Vector Regression," 2016.
23. X. Jia, Z. Jia, and W. Zhang, "Remaining Useful Life Prediction of Turbofan Engine Using Ensemble Learning," 2016.

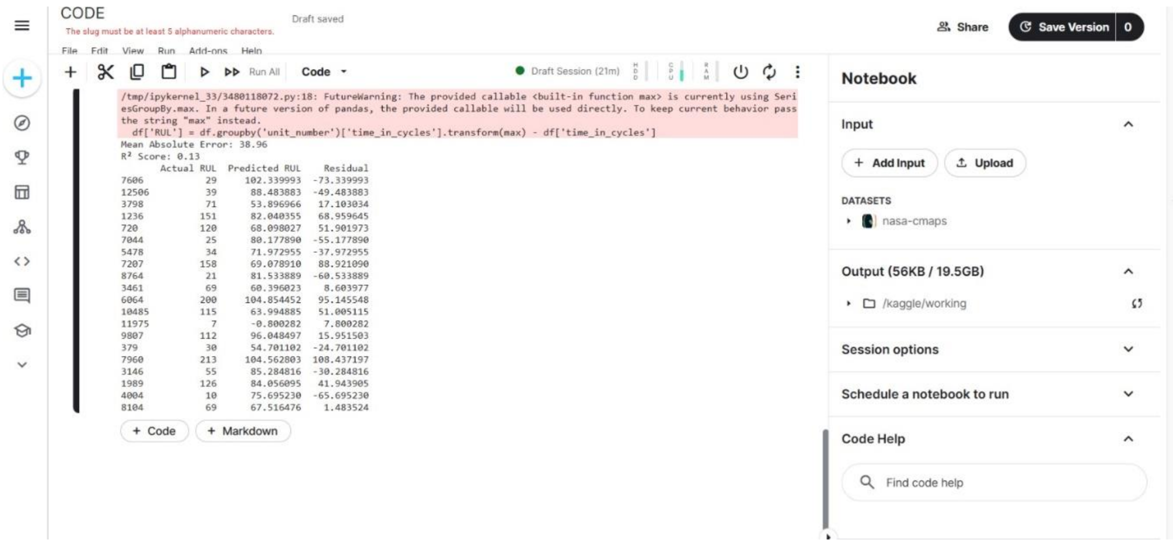
APPENDIX

A. CODING

```
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor,
StackingRegressor
from sklearn.linear_model import ElasticNet, LinearRegression
from sklearn.svm import SVR
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
from sklearn.preprocessing import StandardScaler

# Load data
column_names = ['unit_number', 'time_in_cycles', 'operational_setting_1',
'operational_setting_2',
        'operational_setting_3'] + [f'sensor_measurement_{i}' for i in range(1, 22)]
df = pd.read_csv('/kaggle/input/nasa-cmaps/CMaps/test_FD001.txt', sep=' ',
header=None)
df.drop(df.columns[[26, 27]], axis=1, inplace=True)
df.columns = column_names

# Calculate RUL
df['RUL'] = df.groupby('unit_number')['time_in_cycles'].transform(max) -
df['time_in_cycles']

# Feature Engineering
feature_columns = ['operational_setting_1', 'operational_setting_2',
'sensor_measurement_2', 'sensor_measurement_3', 'sensor_measurement_4']
X = df[feature_columns]
y = df['RUL']
```

```
# Split data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Scale features
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

# Define base estimators
base_estimators = [
    ('random_forest', RandomForestRegressor(n_estimators=100, random_state=42)),
    ('gradient_boosting',GradientBoostingRegressor(n_estimators=100,
random_state=42))
]

# Meta-learners to experiment with
meta_learner_lin = LinearRegression()
meta_learner_enet = ElasticNet(alpha=0.1, l1_ratio=0.7)
meta_learner_svr = SVR(kernel='linear', C=100)

# Choose a meta-learner here by uncommenting:
# final_estimator = meta_learner_lin
# final_estimator = meta_learner_enet
final_estimator = meta_learner_svr  # Example: Using SVR as meta-learner

# Stacking Regressor
stack_reg = StackingRegressor(
    estimators=base_estimators,
    final_estimator=final_estimator,
    cv=5
)

# Train the Stacking Regressor
stack_reg.fit(X_train_scaled, y_train)
```

```
# Predict on test data
final_predictions = stack_reg.predict(X_test_scaled)

# Evaluate the model
mse = mean_squared_error(y_test, final_predictions)
mae = mean_absolute_error(y_test, final_predictions)
r2 = r2_score(y_test, final_predictions)

print(f'Mean Squared Error: {mse:.2f}')
print(f'Mean Absolute Error: {mae:.2f}')
print(f'R² Score: {r2:.2f}')

# Print actual RUL, predicted RUL, and residuals
results_df = pd.DataFrame({
    'Actual RUL': y_test,
    'Predicted RUL': final_predictions,
    'Residual': y_test - final_predictions
})
print(results_df.head(20))
```



B. CONFERENCE PUBLICATION

3 of 20

IEEE ICITEICS 2024 - ACCEPTANCE NOTIFICATION External Inbox ×

Microsoft CMT <email@msr-cmt.org>

to me

Wed, Feb 28, 3:49 PM

Dear AKASH KUMAR PANDIT

Paper ID / Submission ID : 606

Title : Predicting aircraft Turbofan Engine Degradation with Recurrent Neural Networks

We are pleased to inform you that your paper has been accepted for the Oral Presentation as a full paper for the- "IEEE 2024 International Conference on Information Technology, Electronics and Intelligent Communication Systems" , Bengaluru, Karnataka, India.

All accepted and presented papers will be submitted to IEEE Xplore for the further publication and will be indexed by Ei Compendex and Scopus Indexing.

Complete the Registration Process (The last date of payment is 05 MARCH 2024)

Registration Link for Payment :

For Indian Authors : https://rzp.io/l/ICITEICS2024

For Foreign Authors : https://www.paypal.me/agastiyarresearch

After the last date of registration IEEE PDF Express and E copyright information will be given for registered authors.

D.PLAGIARISM

QuillBot

QuillBot

Scanned on: 17:45 May 7, 2024 UTC

9.3%

Overall similarity score

39

Results found

9756

Total words in text

	Word count
Identical	453
Minor Changes	117
Paraphrased	341
Omitted	0

Powered by COPYLEAKS

QuillBot

QuillBot   Scanned on: 17:45 May 7, 2024 UTC

Results

The results include any sources we have found in your submitted document that includes the following: identical text, minor changed text, paraphrased text.

Source	%
NASA Turbofan Jet Engine Data Set | IEEE DataPort https://ieee-dataport.org/documents/nasa-turbofan-jet-engine-data-set	4%
readme.txt https://raw.githubusercontent.com/Azure-Samples/IoTEdgeAndMlSample/... readme.txt	3%
Conundrum Aircraft Engine Demo | NVIDIA NGC https://catalog.ngc.nvidia.com/orgs/nvidia/containers/conundrum-aircraft...	3%
CMAPSS Jet Engine Simulated Data | NASA Open Data Portal https://data.nasa.gov/Aerospace/CMAPSS-Jet-Engine-Simulated-Data/ff5v-...	3%
Plant Nutrient Deficiency Report | PDF | Soil | Autoregressi... https://www.scribd.com/document/728000761/Plant-nutrient-deficiency-R...	2%
Predicting Engine Failure using C-MAPSS data · Yajas Dwivedi https://yajasd.github.io/2018/06/04/Predicting-Engine-Failure/	2%
Anomaly Detection in Turbofan Engine using Prophet https://www.linkedin.com/pulse/anomaly-detection-turbofan-engine-using...	1%
Remaining Useful Life (RUL) Prediction - Stratada https://www.stratada.com/remaining-useful-life-rul-prediction/	1%

IDENTICAL

Text that is exactly the same.

MINOR CHANGES

Text that is nearly identical, yet a different form of the word is used. (i.e 'slow' becomes 'slowly')

PARAPHRASED

Text that has similar meaning, yet different words are used to convey the same message.

Unsure about your report?

The results have been found after comparing your submitted text to online sources, open databases and the Copyleaks internal database. If you have any questions or concerns, please feel free to contact us atsupport@copyleaks.com

Click here to learn more about different types of plagiarism

Powered by COPYLEAKS

QuillBot

QuillBot

Scanned on: 17:45 May 7, 2024 UTC

GitHub - archd3sai/Predictive-Maintenance-of-Aircraft-Engi... https://github.com/archd3sai/Predictive-Maintenance-of-Aircraft-Engine	1%
Time Series Prediction in Simulink Using Deep Learning Net... https://de.mathworks.com/help/deeplearning/ug/time-series-prediction-in... time-series-prediction-in-simulink-using-deep-learning-network.html	1%
Time Series Prediction in Simulink Using Deep Learning Net... https://es.mathworks.com/help/deeplearning/ug/time-series-prediction-in-... time-series-prediction-in-simulink-using-deep-learning-network.html	1%
Electronics | Free Full-Text | Data-Driven Deep Learning-Bas... https://www.mdpi.com/2079-9292/10/20/2453	1%
declaration_of_own_work_2015-16.doc https://www.ed.ac.uk/sites/default/files/atoms/files/declaration_of_own_w... declaration_of_own_work_2015-16.doc	1%
AI - FINAL Harsha | PDF | Artificial Intelligence | Intelligenc... https://www.scribd.com/document/726650371/AI-FINAL-harsha	1%
Conundrum Aircraft Engine Demo | NVIDIA NGC https://catalog.ngc.nvidia.com/orgs/partners/helm-charts/helm-conundru...	1%
Predictive Maintenance: Step 2A of 3, train and evaluate reg... https://gallery.azure.ai/Experiment/Predictive-Maintenance-Step-2A-of-3-tr...	1%

QuillBot

QuillBot Scanned on: 17:45 May 7, 2024 UTC

INDJCSE23-14-03-030.pdf *http://www.ijcse.com/docs/INDJCSE23-14-03-030.pdf* INDJCSE23-14-03-030.pdf	1%
Aerospace - Genexia *https://genexia.co/aerospace/*	1%
1912.07383 *https://arxiv.org/pdf/1912.07383* 1912.07383	1%
A Remaining Useful Life Prognosis of Turbofan Engine Using ... *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7827555/*	1%
(PDF) Remaining Useful Life Prediction for Turbofan Engine ... *https://www.researchgate.net/publication/373189070_Remaining_Useful_L...*	1%
(PDF) Remaining Useful Life Prediction of an Aircraft Turbof... *https://www.researchgate.net/publication/367782881_Remaining_Useful_L...*	1%
DASHlink - Turbofan engine degradation simulation data set *https://c3.ndc.nasa.gov/dashlink/resources/139/*	1%
Remaining Useful Life Estimation Using Convolutional Neur... *https://ww2.mathworks.cn/help/predmaint/ug/remaining-useful-life-estim...* remaining-useful-life-estimation-using-convolutional-neural-network....	1%

QuillBot

QuillBot Scanned on: 17:45 May 7, 2024 UTC

Remaining Useful Life Estimation Using Convolutional Neur... https://www.mathworks.com/help/predmaint/ug/remaining-useful-life-esti... remaining-useful-life-estimation-using-convolutional-neural-network....	1%
NASA C-MAPSS Dataset | Papers With Code https://paperswithcode.com/dataset/nasa-c-mapss	1%
Remaining Useful Life Estimation Using Convolutional Neur... https://de.mathworks.com/help/predmaint/ug/remaining-useful-life-estim... remaining-useful-life-estimation-using-convolutional-neural-network....	1%
Data-Driven Deep Learning-Based Attention Mechanism for ... https://mdpi-res.com/d_attachment/electronics/electronics-10-02453/artic... electronics-10-02453-v2.pdf	1%
Recent Trends in Data Science and its Applications https://www.riverpublishers.com/research_article_details.php?book_id=11... research_article_details.php	1%
7m01bm487 https://repository.library.carleton.ca/downloads/7m01bm487	1%
Sustainability | Free Full-Text | Multi-Scale Remaining Usefu... https://www.mdpi.com/2071-1050/14/23/15667	1%
Predicting the Remaining Useful Life of an Aircraft Engine U... https://www.hindawi.com/journals/complexity/2018/3813029/	1%

Powered by COPYLEAKS

QuillBot

QuillBot

Scanned on: 17:45 May 7, 2024 UTC

Remaining useful life prognosis of turbofan engines based o... https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9021315/	1%
Remaining useful life prognosis of turbofan engines based o... https://www.nature.com/articles/s41598-022-10191-2	1%
Spatio-temporal degradation modeling and remaining usefu... https://www.sciencedirect.com/science/article/abs/pii/S0951832022005038	1%
F36570811622 - International Journal of Engineering and Ad... https://www.ijeat.org/portfolio-item/f36570811622/	1%
1678186766.docx https://www.ijprems.com/uploadedfiles/paper/volume_3/issue_3_march_2... 1678186766.docx	1%
Actuators | Free Full-Text | Remaining Useful Life Predictio... https://www.mdpi.com/2076-0825/11/3/67	1%
White shark optimizer with optimal deep learning based eff... https://pubmed.ncbi.nlm.nih.gov/38155207/	1%