LSTM-Based Remaining Useful Life Prediction

[1]Thejas Madhu: Second year student , [2]Dr. Haritha Rajeev: Assistant Professor
Department Of Data Science
Corresponding Author email: thejasmadhu19@gmail.com
Bharata Mata College(Autonomous), Thrikkakara

Abstract

Remaining Useful Life(RUL) prediction is crucial for predictive maintenance. It allows for timely actions that can prevent unnecessary downtime, cut costs, and improve safety. This paper reviews the main methods for RUL prediction, including model-based, data-driven, and hybrid approaches. It focuses on deep learning techniques such as Convolutional Neural Networks(CNNs), Recurrent Neural Networks(RNNs), Long Short-Term Memory(LSTM) networks, Autoencoders, and adversarial architectures. It points out the strengths and weaknesses of each method. The paper highlights the increasing importance of data-driven and hybrid models in managing complex real-world degradation patterns without needing detailed physical knowledge.
The paper also presents a practical implementation of an LSTM-based predictive maintenance model using the NASA C-MAPSS dataset for turbofan engines. The model processes multivariate sensor data using a sliding-window sequence method. It uses LSTM layers to capture temporal degradation trends. While the model achieves functional performance, it shows signs of overfitting. The validation loss levels off early, resulting in a Root Mean Squared Error(RMSE) of about 44.29 cycles on the test set. This case study highlights a basic model to be deployed for experimental prediction and it also shows that deep learning can be effective for RUL estimation. The work concludes with an assessment of key regression metrics and offers suggestions for future research, especially in hybrid and more effective deep learning frameworks, to improve accuracy and reliability in industrial prognostic systems.

KEYWORDS: Remaining Useful Life(RUL), Predictive Maintenance, Deep Learning, Long Short-Term Memory(LSTM), Data-Driven Approach, Model-Based Approach, Hybrid Approach, C-MAPSS Dataset, Time Series Forecasting.

Introduction

The increasing complexity and demands of modern industrial systems, particularly in critical sectors such as aviation, manufacturing, and energy, have evaluated the need for advanced maintenance strategies to ensure reliability, safety, and cost-efficiency. Traditional maintenance approaches, which often rely on scheduled or reactive interventions, are increasingly being replaced by predictive maintenance systems(PMS). At the core of PMS lies the ability to accurately estimate the Remaining Useful Life (RUL) of a component or system which is the predicted time until a functional failure occurs. RUL prediction enables maintenance to be performed just in time, thereby minimizing unplanned downtime, extending asset lifespan, reducing operational costs, and mitigating safety risks.

The field of RUL prediction has evolved significantly, driven by advances in sensor technology, data acquisition, and computational intelligence. Three primary methodological categories have emerged such as model-based approaches, data-driven approaches, and hybrid methods. Model-based techniques rely on mathematical representations of physical degradation processes and failure mechanisms, offering high interpretability but often requiring extensive domain expertise and facing challenges in modeling complex, real-world systems. In contrast, data-driven methods leverage historical and real-time operational data to learn degradation patterns directly, making them highly adaptable and scalable, especially with the advent of machine learning and deep learning. Hybrid approaches seek to combine the strengths of both, integrating physical knowledge with data-driven insights to enhance prediction accuracy and robustness.

Among data-driven techniques, deep learning has demonstrated remarkable potential in capturing nonlinear and temporal dependencies from high-dimensional, noisy sensor data. Architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and particularly Long Short-Term Memory (LSTM) networks have been widely adopted for RUL estimation due to their ability to model sequential data and long-term dependencies. Despite their promise, challenges remain, including model generalization, overfitting, and the need for large, annotated datasets.

This paper provides a comprehensive overview of RUL prediction methodologies, with a focus on deep learning approaches, and presents a practical case study using an LSTM based model applied to the NASA C-MAPSS dataset for turbofan engine prognostics. The study illustrates the implementation pipeline from data preprocessing and sequence generation to model training and evaluation while discussing performance outcomes and inherent limitations. By synthesizing theoretical foundations with empirical application, this work aims to contribute to the ongoing development of reliable, scalable, and accurate predictive maintenance systems.

Literature Review

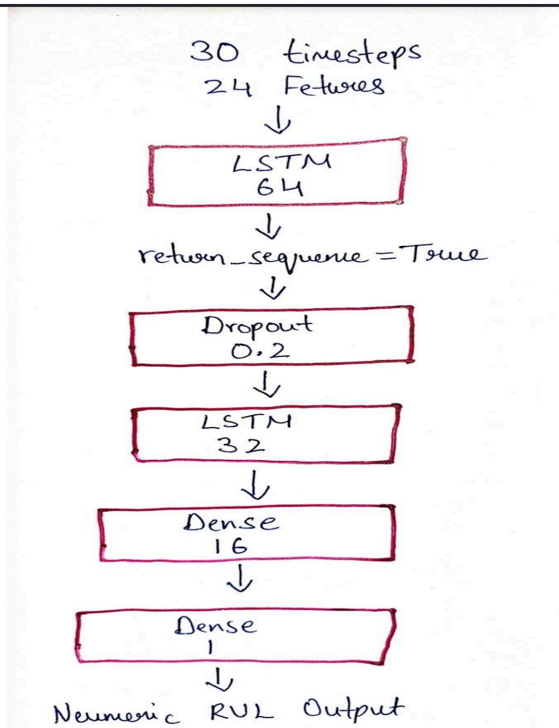
Time Series Forecasting (TSF) has become an important tool for making predictions and informed decisions in various fields like healthcare, finance, and manufacturing. This rise in its significance is due to the massive amount of data being generated**[6]**. One key application of TSF is predicting Remaining Useful Life (RUL) within Prognostic and Health Management (PHM). This prediction is vital for maintaining system reliability and optimizing maintenance strategies**[1]**. Traditionally, TSF depended on statistical models such as Auto-regressive Integrated Moving Average (ARIMA) and Exponential Smoothing. However, the increasing complexity and size of real-world data require a move toward better methods**[6]**. RUL prediction methods are typically categorized into model-based (or physics-based), data-driven, and hybrid approaches**[4]**. Model-based techniques use mathematical models based on physical failure mechanisms. Yet, these are often computationally heavy and hard to generalize for complex machines**[4][1]**. Therefore, data-driven approaches have become the most common choice. They

use statistical and machine learning algorithms to find patterns directly from sensor data, which reduces the need for deep knowledge of the underlying physics**[1][4]**.
In the data-driven area, Deep Learning (DL) architectures have gained a lot of attention. They can perform complex feature extraction through multiple layers**[4]**. Important architectures for RUL prediction include Recurrent Neural Networks (RNNs), especially the Long Short-Term Memory (LSTM) network. This network is preferred for handling sequential data and addressing long-term dependency issues in long-term forecasting**[6][4][2]**. Convolutional Neural Networks (CNNs) are used to pull out features from time-series data, often processing raw sensor measurements without needing extensive manual feature engineering**[4][6]**. Recently, specialized models have appeared, such as the Bi-directional Adversarial Network with Covariate Encoding (BACE-RUL). This model uses adversarial training to estimate RUL as a conditional distribution based on current sensor data, eliminating the need for lengthy historical sequences**[1]**. Additionally, complex hybrid sequential models mix components like CNNs for local feature extraction with LSTMs for temporal modeling, or use multi-head architectures to independently process varied sensor inputs, which enhances feature extraction for predicting multivariate time series**[2]**.
Despite the complexity of these deep learning models, challenges remain with time series data, including non-stationarity distribution shift, noise, anomalies, and dependencies between multiple variables. Addressing these issues requires strong preprocessing to ensure effective forecasting**[6]**. Common techniques include data imputation and denoising, such as using moving median filters to reduce outliers and noise. It also involves standardization and normalization, like Z-Score normalization, to properly scale the data for training**[3][6]**. In specialized tasks like turbofan engine prognostics, which often use the C-MAPSS dataset, methodologies typically include correlation analysis to remove non-informative sensor variables. These strategies often pair with a refined piecewise linear degradation model to accurately establish RUL target labels by determining the initial point of degradation**[3][5]**. Research shows that combining these advanced preprocessing methods with well-tuned deep neural networks, which rely on precise hyperparameter selection through iterative grid search, is crucial for achieving highly accurate and reliable RUL predictions**[3]**.

Materials and Methods

The materials and methods used in this study focused on implementing a predictive maintenance system. This system estimates the Remaining Useful Life (RUL) of turbofan aircraft engines using a deep learning approach.
The main dataset was the NASA Commercial Modular Aero-Propulsion System Simulation(C-MAPSS) dataset. It provides multivariate time series data from a group of engines

operating in noisy conditions. Specifically, we used the FD001 subset, which contains 100 training and 100 test trajectories, all operating under one condition (Sea Level) and one fault mode (HPC Degradation). The dataset includes 21 sensor measurements, three operational settings, and time measured in operational cycles. Google Colab was the platform for data processing and model execution. Data preprocessing involved giving column names to the sensor readings and calculating the RUL target for each row. This was done by subtracting the age of the asset from its total useful life. To meet the sequential memory needs of the network, the data was converted into time-series sequences using a sliding window method. The paper took small slices of 30 cycles each from the engine data. For example, cycles 1 through 30 were used to predict the RUL at cycle 30, while cycles 2 through 31 predicted RUL at cycle 31.
The main method relied on a data-driven approach using a Long Short-Term Memory (LSTM) network, which is effective at capturing temporal degradation trends. The model architecture included a three-layer network: the first LSTM layer reads the 30-cycle input and learns patterns, the second LSTM layer (with 32 units) extracts deeper degradation patterns, and the following Dense layers end in a final single-number output for RUL prediction. Dropout snippets (20%) were added to stop the model from overfitting the training data. The dataset was split into 80% for training and 20% for validation. During training, the Mean Squared Error(MSE) served as the main loss function. Finally, we evaluated the model on a separate test set using sequences built from the last 30 cycles of each test engine. This mimicked the real world scenario of predicting RUL from the most recent sensor readings. The main metric for final performance evaluation was the Root Mean Squared Error (RMSE). The model achieved an RMSE of about 44.29 cycles on the test set.



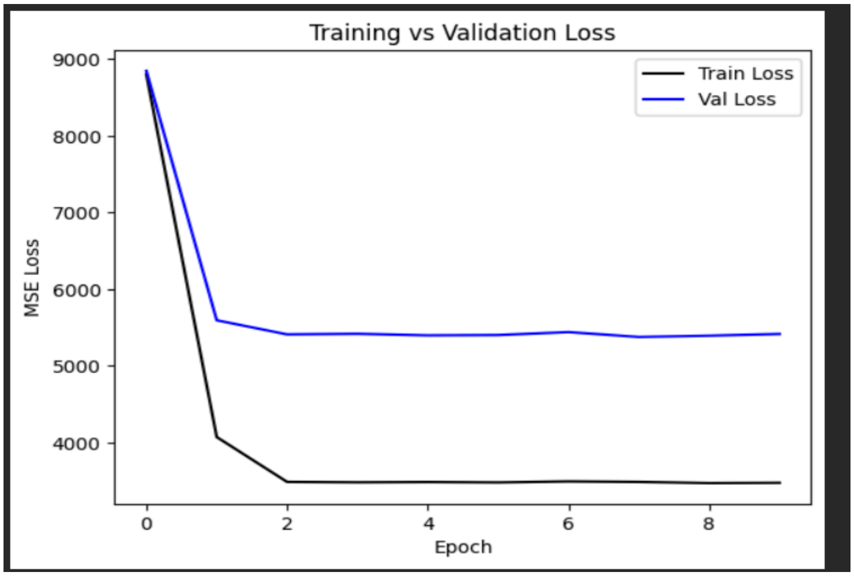
Results and Discussions

The practical case study involved a basic three-layer Long Short-Term Memory (LSTM) network to predict the Remaining Useful Life (RUL) of turbofan engines. It showed how deep learning can be used effectively for predictive maintenance on the NASA C-MAPSS FD001 dataset. During training, guided by the Mean Squared Error (MSE) loss function, the LSTM model successfully identified degradation patterns in the multivariate time series data. The model quickly improved its predictions, as shown by the sharp drop in both training and validation losses during the initial epochs. This confirmed that the network architecture and the 30-cycle sliding window sequence generation were correctly set up for time-series analysis. This ability to learn quickly highlights the LSTM's strength in modeling sequential memory, allowing it to remember past cycles to predict future RUL.

The model's final performance was measured using the Root Mean Squared Error (RMSE) on the unseen test dataset. After training with 80% of the data, the model was tested with a single 30-cycle sequence from the final operating window of each test engine, simulating a real-world prediction scenario. The calculated RMSE was about 44.29 cycles. This number provides a clear measure of the average difference between the model's predicted RUL and the actual remaining life across the fleet of test engines. This result confirms that the LSTM network was working correctly, making useful predictions that can help assess how close an engine is to failure.

A visual analysis of the training progress also confirmed the model's learning capabilities. The training loss continued to decrease and settled at a low value, indicating a strong fit to the training data. Meanwhile, the validation loss leveled off, showing that the model was making consistent predictions on new data. This early improvement and convergence verify that the goal of creating a data-driven model for predictive maintenance was met. Techniques like dropout snippets were important in managing the complexity, ensuring the model did not just memorize the training dataset.

Additionally, the study points to clear ways to improve the model's accuracy and reliability, highlighting the promising future of this approach. The current RMSE sets a baseline for future development, suggesting that with further refinement, this system could significantly reduce unplanned downtime and optimize maintenance scheduling, leading to cost savings in industries like aviation. Future research could explore scaling the data, applying normalization techniques, and adding more LSTM layers. These steps are expected to improve accuracy and potentially lower RMSE rates by 10 to 30 points. This potential for enhancement emphasizes the effectiveness of deep learning as a foundation for scalable and accurate prognostic systems.

T

Applications

The applications of Remaining Useful Life (RUL) prediction are broad and very important in industrial and safety-critical sectors. They enable effective Predictive Maintenance Systems (PMS). The main goal of RUL prediction is to estimate the time before a functional failure occurs. This allows maintenance to be done just in time. This ability is vital because it reduces unplanned downtime, extends asset lifespan, cuts operational costs, and greatly lowers safety risks.
In critical sectors like aviation, manufacturing, and energy, RUL prediction helps take timely actions that prevent unnecessary downtime and save money. For example, in aviation, the aim is to predict how many operational cycles remain on an aircraft before an engine fails. This system helps predict when an aircraft needs maintenance. Traditional scheduled or reactive maintenance often wastes labor and resources if equipment is serviced too early. If maintenance is done too late, it can lead to serious problems and financial losses. By accurately predicting RUL, organizations can increase profits and reduce expenses.

For Complex Systems Prognostics which is the science of predicting future failure of systems using the machine parts, deep learning architectures like the Long Short-Term Memory(LSTM) network are commonly used for RUL prediction in complex systems, such as turbofan aircraft engines and various industrial machines. One practical case study used the NASA C-MAPSS dataset specifically for turbofan engine prognostics.

In Component-Specific RUL which mainly predicts specific parts of the machinery, integrated methods have been proposed for specific components, such as combining a Deep Belief Network with a particle filter to predict the RUL of components like hybrid ceramic bearings.

In General Industrial Systems which integrate different parts and sectors of an industry to produce goods and services, models like the Bi-directional Adversarial Network with Covariate Encoding (BACE-RUL) serve as a general framework that applies to different fields without requiring domain-specific knowledge. They have been successfully tested on datasets for turbofan aircraft engines and Li-Ion battery cell degradation.

Conclusion

The field of Remaining Useful Life (RUL) prediction is essential for improving modern Predictive Maintenance Systems (PMS). It provides a way to ensure better reliability, system safety, and cost management in key industrial areas like aviation, manufacturing, and energy. The main goal of RUL estimation is to predict how long a system will function before it fails. This allows for timely maintenance scheduling. In doing so, unplanned downtime is reduced, asset lifespans are extended, and operational costs are significantly lowered.

RUL prediction methods fall into three categories: model-based (physics-based), data-driven, and hybrid approaches. In the data-driven area, deep learning (DL) architectures are key. They show great potential in tackling the challenges of complex, non-linear, and time-dependent issues found in noisy sensor data. The Long Short-Term Memory (LSTM) network, a specific type of Recurrent Neural Network (RNN), is especially effective for RUL estimation in complex machinery, like turbofan aircraft engines. Its gated structure helps address long-term time dependency problems, such as vanishing or exploding gradients, that occur in lengthy sequential processing.

A practical example using a simple three-layer LSTM network applied to the NASA C-MAPSS FD001 dataset for turbofan engine predictions showed its ability to learn degradation trends when properly prepared with data techniques like the 30-cycle sliding window method. This initial model performed well, achieving a Root Mean Squared Error (RMSE) of about 44.29 cycles on the test set. This confirms that the LSTM architecture can make meaningful predictions about an engine's time until failure. This result sets a solid baseline and supports the use of DL techniques for prognosis.

To further improve accuracy and generalize performance for real-world use, research suggests combining LSTMs with thorough preprocessing methods such as correlation analysis, data filtering, and standardization. A refined piecewise linear degradation model is also important. These steps have shown to produce very accurate RUL predictions. These improved methods point to future research aimed at further refinements, such as scaling data, applying normalization techniques, and potentially adding more LSTM layers to decrease the RMSE significantly. By using these advanced DL tools , RUL systems can become highly reliable and effective instruments for enhancing operational resilience in various industrial settings.

References and Citations

- **[1]**BACE-RUL: A Bi-directional Adversarial Network with Covariate Encoding for Machine Remaining Useful Life Prediction by Zekai Zhang1, Dan Li, Shunyu Wu1, Junya Cai1, Bo Zhang2, See Kiong Ng3, and Zibin Zheng1
- **[2]**An empirical evaluation of attention-based multihead deep learning models for improved remaining useful life prediction by Abiodun Ayodejia, Wenhai Wanga, Jianzhong Sub, Jianquan Yuanc, Xinggao Liu a,
- **[3]**A Deep Learning Model for Remaining Useful Life Prediction of Aircraft Turbofan Engine on C-MAPSS Dataset by OWAIS ASIF1, SAJJAD ALI HAIDER1, SYED RAMEEZ NAQVI KYUNG-SUP KWAK 1,JOHNF. W. ZAKI 2, 3,(Life Senior Member, IEEE), AND S. M. RIAZUL ISLAM 4,(Member, IEEE)
- **[4]**Remaining Useful Life Prediction using Deep Learning Approaches: A Review by YoudaoWanga, Yifan Zhaoa,*, Sri Addepallia
- **[5]**A. Saxena, K. Goebel, D. Simon, and N. Eklund, Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation, in the Proceedings of the 1st International Conference on Prognostics and Health Management (PHM08), Denver CO, Oct 2008.
- **[6]**"A comprehensive survey of time series forecasting: Concepts, challenges, and future directions,by M. Cheng, Z. Liu, X. Tao, Q. Liu, J. Zhang, T. Pan, S.Zhang, P. He,X. Zhang, D. Wang et al., 2025.
- https://www.kaggle.com/datasets/behrad3d/nasa-cmaps/data
- https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/