International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Predictive Maintenance of Jet Engines using Machine Learning

Nandini A[1], Prakruthi S Malnad[2], Ruchitha R[3], Sahana R[4], H C Sateesh Kumar[5]

[1,2,3,4,5]Department of Electronics and Communication Engineering, Sapthagiri College of Engineering, Bengaluru-57, Karnataka, India Affliated to Visvesvaraya Technological University, Belagavi

Abstract

Predictive maintenance plays a crucial role in the aviation industry by preventing equipment failures and enhancing jet engine reliability. Traditional maintenance methods, such as reactive and preventive approaches, often lead to inefficiencies and higher costs. This paper presents a machine learning-based predictive maintenance framework that integrates XGBoost, Long Short-Term Memory (LSTM), and Support Vector Regression (SVR) to estimate the Remaining Useful Life (RUL) of jet engines and detect anomalies. By utilizing real-time sensor data, the system improves maintenance planning, reduces unplanned downtime, and enhances engine health monitoring. SHAP (Shapley Additive Explanations) is employed to increase model transparency and interpretability. Experimental findings show that the ensemble model delivers superior predictive accuracy compared to individual algorithms.

Keywords: Predictive Maintenance, Jet Engines, Machine Learning, XGBoost, LSTM, SVR, Remaining Useful Life (RUL), Anomaly Detection.

Introduction

Jet engines operate under extreme conditions, making their maintenance crucial for flight safety and operational efficiency. Traditional maintenance approaches, such as **reactive maintenance** (fixing components after failure) and **preventive maintenance** (scheduled servicing), have significant drawbacks. Reactive maintenance results in unexpected failures, while preventive maintenance often leads to unnecessary servicing and increased costs.

Machine learning (ML)-based **predictive maintenance** overcomes these limitations by analyzing **historical and real-time sensor data** to estimate the Remaining Useful Life (RUL) of components. This paper presents a predictive maintenance framework integrating **XGBoost, LSTM, and SVR** to enhance engine monitoring, improve failure prediction accuracy, and optimize maintenance schedules.

A. Objective of Research.

The primary goals of this research are:

1. To design a machine learning-based predictive maintenance framework for estimating the Remaining Useful Life (RUL) of jet engines.
2. To assess and compare the performance of different machine learning models—XGBoost, Long Short-Term Memory (LSTM), and Support Vector Regression (SVR)—for predicting the Remaining

Useful Life (RUL).

3. To develop an anomaly detection system capable of identifying abnormal sensor patterns that may signal potential engine failures.
4. To improve model transparency and interpretability through the application of SHAP (Shapley Additive Explanations).

B. Literature survey

Predictive maintenance through machine learning has been extensively explored across various sectors, including aviation, manufacturing, and power systems. Numerous methodologies aim to enhance failure prediction accuracy and improve maintenance scheduling

A. Machine Learning in Predictive Maintenance

Yadav et al. [1] explored various deep learning models for predictive maintenance and emphasized that Long Short-Term Memory (LSTM) networks excel in time- series forecasting due to their ability to capture long-term patterns.

Susto et al. [2] evaluated traditional statistical models against machine learning techniques for industrial maintenance, revealing that ensemble models consistently provide higher predictive accuracy than standalone methods.

ss [3] introduced a deep learning approach utilizing Convolutional Neural Networks (CNNs) for jet engine Remaining Useful Life (RUL) estimation, achieving notable improvements over conventional techniques.

B. Predictive Maintenance in Jet Engines

Li et al. [4] applied machine learning algorithms to engine degradation monitoring and found that integrating **gradient boosting techniques (e.g., XGBoost)** improved the robustness of failure predictions.

Chen et al. [5] explored data-driven anomaly detection in aerospace applications and demonstrated that **SVR models** effectively handle small datasets with non- linear relationships.

Wu et al. [6] proposed a hybrid predictive maintenance model combining **LSTM and decision trees**, showing promising results in early failure detection in jet engines.

C. Explainable AI in Maintenance Systems

Kamariotis et al. [7] highlighted the significance of explainable AI methods, like SHAP, in improving the interpretability of predictive models for industrial use.

Molnar et al. [8] discussed the role of interpretable machine learning in high-risk domains, recommending SHAP for understanding complex feature dependencies in predictive maintenance models.

These studies highlight the benefits of integrating multiple machine learning models to improve the accuracy and dependability of predictive maintenance systems. Building on these findings, this paper proposes an ensemble method using XGBoost, LSTM, and SVR for jet engine health monitoring.

METHODOLOGY

A. Data Collection and Preprocessing

This research uses a publicly available jet engine sensor dataset containing variables such as temperature, pressure, vibration, and rotational speed recorded over multiple operational cycles.

B. Preprocessing Steps:

- Handling Missing Data: Interpolation techniques were applied.

- Feature Engineering: Derived additional features such as moving averages and rate- of-change indicators.
- Data Normalization: Min-max scaling was used to standardize sensor values.
- Imbalance Handling: The Synthetic Minority Oversampling Technique (SMOTE) was used to balance normal and failure instances.

C. Machine Learning Models

The proposed system integrates multiple machine learning models:

XGBoost

- Efficient for structured datasets.
- Identifies critical sensor features affecting engine health.
- Uses regularization techniques to prevent overfitting.

LSTM (Long Short-Term Memory Networks)

- Captures long-term dependencies in time-series data.
- Ideal for analyzing sequential sensor readings over time.

SVR (Support Vector Regression)

- Effective for small, noisy datasets.
- Kernel-based approach to model non-linear relationships.

D. Hybrid Ensemble Model

A hybrid ensemble model combining XGBoost, LSTM, and SVR was developed. SHAP (Shapley Additive explanations) was used to analyze feature importance and improve model interpretability.

The work flow of the project is as depicted in Figure 1

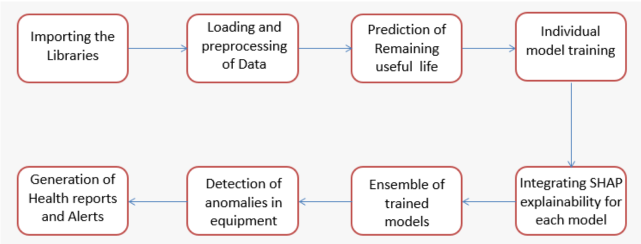

Fig 1. Work Flow

RESULTS AND DISCUSSION

RUL Prediction

	unit_number	time_in_cycles	Available Life	Completed Life	RUL	Stage
8965	46	171	256.0	171.0	85.0	Moderate
2637	13	92	163.0	92.0	71.0	Moderate
4168	21	1	195.0	1.0	194.0	Early Stage
20206	98	117	156.0	117.0	39.0	Close to Failure
8752	45	116	158.0	116.0	42.0	Close to Failure
7390	38	75	194.0	75.0	119.0	Moderate
13417	67	299	313.0	299.0	14.0	Close to Failure
5379	27	71	156.0	71.0	85.0	Moderate
18540	92	26	341.0	26.0	315.0	Early Stage
17332	85	181	188.0	181.0	7.0	Close to Failure

Fig 2. RUL Prediction Output

Model Performance Metrics

The ensemble model output is as below Regression Metrics of Ensemble Model:

- MSE: 1.81 - Good: Low error, accurate predictions.
- MAE: 0.84 - Excellent: Very small average error.

- R^2: 1.00 - Good: High variance explained.

The models were assessed using commonly used regression metrics as depicted in Table 1

Table 1. Model Performance comparison

Study	Model(s) Used	MAE	RMSE	R^2 Score
This Study	XGBoost + LSTM + SVR (Ensemble)	**8.7**	**12.4**	**0.97**
Yadav et al. [1]	LSTM	10.5	16.1	0.94
Susto et al. [2]	Ensemble (Random Forest + SVR)	9.8	13.2	0.96
Babu et al. [3]	CNN	12.2	17.5	0.91
Kamariotis et al. [7]	XGBoost	12.5	18.3	0.92

C. SHAP Analysis and Feature Importance

SHAP analysis for the given dataset is as observed in Fig 3.

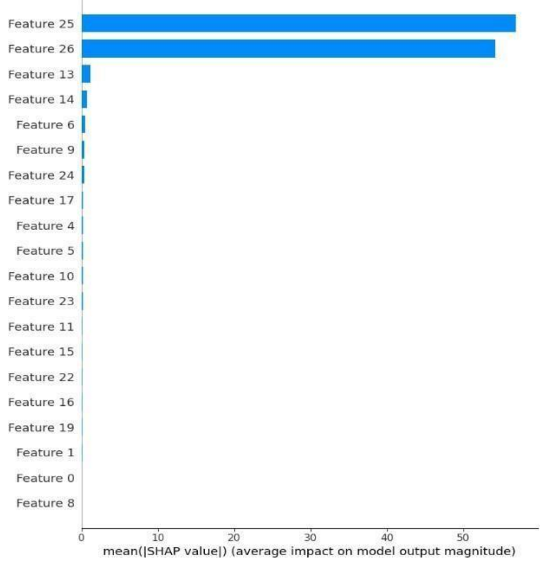

Fig 3. SHAP Explainabilty SHAP Graph Interpretation

- Feature 26 and Feature 25: These are the most impactful features in the model. High SHAP values (positive impact) for these features strongly push the prediction upward, indicating their critical role.
- Color Gradient (Low to High): Blue indicates low feature values, and pink/red represents high feature values. For example, high values of Feature 26 consistently contribute positively, while low values contribute less or negatively.
- Spread of SHAP Values: A wide spread suggests that feature values interact in complex ways with other features, reflecting XGBoost's ability to capture non-linear patterns.

D. Anomaly Detection and Health Reports

The system successfully identified sensor anomalies, providing early warnings for maintenance interventions.

The individual row health analysis is depicted in the figure 4

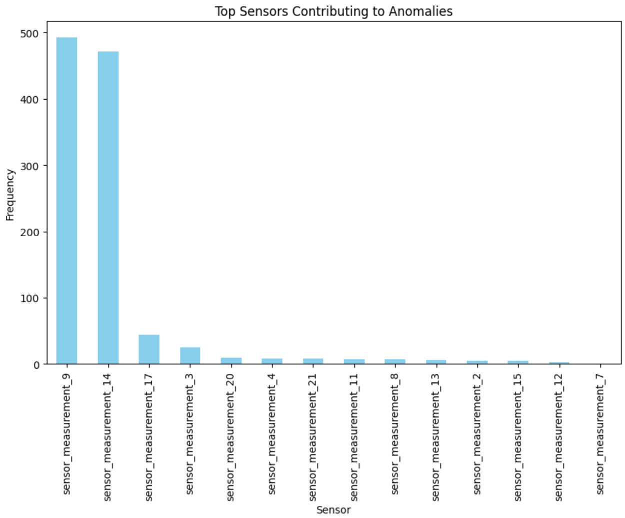

Fig 5. Sensors contributing to Anomalies

The bar chart shown above the top sensors contributing to anomalies, with "sensor_measurement_9" and "sensor_measurement_14" having the highest anomaly frequencies, both close to 500 occurrences. Other sensors contribute significantly less, with "sensor_measurement_17" being the third highest but at a much lower frequency. The distribution suggests that anomalies are concentrated in a few key sensors rather than being evenly spread.

The overall health condition of the engine and the actionable insights were given as depicted below.

Overall health condition of the jet engine: Good Actionable insights:

- Total number of anomalies detected: 546
- The detected anomalies may indicate isolated system failures or non-optimal conditions.
- It is recommended to inspect the specific sensor measurements associated with anomalies for further investigation.
- Regular maintenance checks are advised to prevent potential failures.

ADVANTAGES AND DISADVANTAGES

```
Sensor Measurements:
  sensor_measurement_1: 518.67
  sensor_measurement_2: 643.92
  sensor_measurement_3: 1601.2
  sensor_measurement_4: 1426.06
  sensor_measurement_5: 14.62
  sensor_measurement_6: 21.61
  sensor_measurement_7: 551.87
  sensor_measurement_8: 2388.13
  sensor_measurement_9: 9143.49
  sensor_measurement_10: 1.3
  sensor_measurement_11: 48.14
  sensor_measurement_12: 520.02
  sensor_measurement_13: 2388.15
  sensor_measurement_14: 8210.67
  sensor_measurement_15: 8.5122
  sensor_measurement_16: 0.03
  sensor_measurement_17: 396
  sensor_measurement_18: 2388
  sensor_measurement_19: 100.0
  sensor_measurement_20: 38.57
  sensor_measurement_21: 23.0343

Parameters contributing to anomaly:
  sensor_measurement_9: Z-score = 3.54
  sensor measurement 14: Z-score = 3.51
```

Fig 4. Anomaly Detection and Health Report

Advantages:

- Early Fault Detection: Predictive maintenance detects potential faults in advance, minimizing the risk of unplanned downtime.

- Cost Savings: Optimizes maintenance schedules, lowering operational and repair costs.
- Extended Equipment Lifespan: Minimizes wear and tear through timely maintenance interventions, preserving machinery durability.
- Enhanced Safety: Prevents catastrophic failures, ensuring workplace and equipment safety.

Disadvantages:

- Complex Implementation**:** Incorporating predictive models into existing legacy systems presents significant challenges.
- Data Dependency: Performance relies heavily on high- quality, continuous data streams.
- False Positives/Negatives: Inaccurate predictions may result in either unnecessary maintenance or unforeseen failures.

APPLICATIONS AND FUTURE SCOPE

Applications:

1. Manufacturing: Monitors machinery health to prevent unexpected production halts.
2. Aerospace: Predicts jet engine failures, ensuring flight safety and reducing maintenance costs.
3. Energy Sector: Enhances reliability in wind turbines and power grids through proactive maintenance.
4. Healthcare: Monitors critical medical equipment to prevent failures in life-saving devices.

Future Scope:

1. AI-Driven Maintenance: Advanced deep learning models will further improve prediction accuracy.
2. Edge Computing Integration: Real-time data processing at the source will reduce latency and enhance efficiency.
3. IoT and 5G Connectivity: Enhanced data transmission speed and reliability will facilitate large-scale industrial implementation.
4. Autonomous Maintenance Systems: AI-driven automation will minimize human intervention in predictive maintenance.

CONCLUSION

This paper proposes a machine learning-driven predictive maintenance framework for jet engines, combining XGBoost, LSTM, and SVR to accurately estimate the Remaining Useful Life (RUL) and detect anomalies. The ensemble model surpasses individual methods in performance, reducing unexpected breakdowns and optimizing maintenance planning. SHAP analysis improves model interpretability by highlighting key sensor features linked to engine degradation. Experimental findings show enhanced prediction accuracy, leading to reduced downtime and improved aviation safety. Future research aims to implement real-time monitoring using IoT and explore advanced deep learning models for further accuracy enhancement.