Journal of Advanced College of Engineering and Management

Predictive Maintenance in Aircraft

[1*]Dikshya Aryal, [2]Hemant Joshi
[1,2]Department of Computer Science and Engineering, Universal Engineering and Science College, Lalitpur, Nepal
*Corresponding email: *dikshyauesc@gmail.com*

DOI: 10.3126/jacem.v12i01.93929

Abstract

Predictive maintenance (PdM) is crucial in vital sectors like aviation. It uses data analytics and machine learning to optimize maintenance plans, lower operating costs, and eliminate unscheduled downtime. To create a PdM framework for forecasting the Remaining Useful Life (RUL) of aviation components, this study uses a publicly accessible aircraft engine dataset, which includes time-series data of operating settings and sensor readings. This study uses advanced machine learning model, such as Random Forest to show how PdM systems can predict component failures. The methodology includes data preparation, feature engineering, and rigorous evaluation utilizing measures. Though the dataset is not specific to Nepal, the findings show how this method can improve aviation safety in regions like Nepal, where difficult operating circumstances and a high frequency of aviation incidents demand better maintenance practices. This study raises the way for predictive maintenance's increased adoption in the aviation industry by showing that it improves safety, lowers planned downtime, and promotes operational efficiency.

Keywords*—Aviation safety, Datasets, Machine learning models, Operational efficiency, Predictive maintenance*

1. INTRODUCTION

Aircraft maintenance is an integral part of ensuring an aircraft is safe for operation. Poor maintenance planning can lead to devastating financial results for air carriers and keep aircraft grounded, and passengers waiting, and can even lead to flight cancellations. To increase operational reliability and cost-saving measures, aircraft operators follow aircraft maintenance programs. There are three well-known types of maintenance: reactive, preventive, and predictive[1]. Predictive maintenance builds on condition-based monitoring to optimize the performance and lifespan of equipment by continually assessing its health in real-time, collecting data from sensors, and applying analytical tools and processes like machine learning (ML). [2] Predictive maintenance can identify, detect, and address issues as they occur, predict the potential future state of equipment, and reduce risk. In the aviation industry, Predictive maintenance is transforming. Airlines are increasingly adopting this proactive approach to improve aircraft reliability, reduce operational costs, and enhance safety.

Predictive Maintenance involves monitoring the condition of aircraft components in real time. Sensors installed on key systems gather data about temperature, vibration, pressure, and other performance metrics. This data is then analyzed using a machine-learning algorithm to identify patterns that signal potential failure. Reactive maintenance refers to a timeline in which a particular part of an aircraft is used to its limits and repairs are only performed after a failure which makes this method usually costly and dangerous for operational safety. Many aircraft operators use preventive aircraft maintenance, also known as planned maintenance, which refers to a determined timeline of checks on certain airplane components[1].

Table 1: Comparison of Maintenance Strategies

Aspect	**Reactive Maintenance**	**Preventive Maintenance**	**Predictive Maintenance**
Downtime	High	Moderate	Low
Cost	High	Moderate	Low
Safety Risk	High	Moderate	Low
Data Dependency	Low	Low	High
Resource Optimization	Poor	Moderate	Excellent

The above table illustrates the workflow of reactive, preventive, and predictive maintenance approaches, highlighting the advantages of predictive maintenance and how it lowers risks and expenses.

2. RELATED WORK

[3]PdM is where the system is regularly monitored, and maintenance action is only triggered by a predefined condition of the system. By analyzing a system's physical parameters such as temperature, pressure, or vibration using either trend analysis, pattern recognition, or statistical analysis, it is possible to predict the condition of the system at which failure is imminent. Therefore, before the degradation level reaches this threshold, the system that is about to fail can be replaced. [4]Aircraft are more capable than ever of recording vast amounts of sensor data across almost all of their components in flight, with an Airbus A380 having up to 25000 sensors. This increase in data has driven greater use of data-driven PdM, that is to build and train PdM algorithms using data rather than domain experience.

PdM is one of the factor strategies based on real-time data to diagnose a failure of the machine through forecasting the remaining useful life (RUL), especially on aircraft machines where safety is a priority due to enormous cost and human life. Machine Learning (ML) is the technique that accurately prediction through the data. To validate the result of using the ML method, Evaluation is important to indicate the performance of the model that has been trained. Root Mean Squared Error (RMSE) is a frequent evaluation method for regression technique, this evaluation indicates how concentrated the output is around the line of best fit on prediction data[5].

Extracting and modeling the engine symmetry characteristics is significant in improving RUL predictions for aircraft components, and it is critical for an effective and reliable maintenance strategy[6].RUL prediction methods are mainly classified into two categories: model-based methods and data-driven methods. Model-based methods construct the model using mechanical principles, which takes some time.[7] However, developing an accurate model is challenging due to the complex system structure and uncertain environment.

3. METHODOLOGY

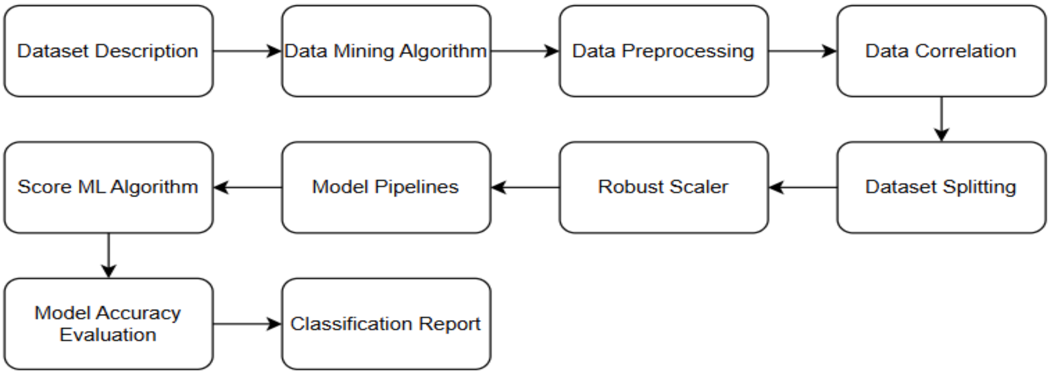

Figure 1: Block Diagram

A. Dataset Description

The dataset was collected from kaggle [8]. Engine sensor readings over time are recorded in PM_train (20,632 samples), which is used for model training. Engine conditions are provided by PM_test (13,097 samples), which is used for evaluation. The ground truth RUL for the engines in the test is provided by PM_truth (101 samples). Three operating parameters and twenty-one sensor readings are included and RUL is the target variable. Dataset contains the parameters like id which is the unique identifier for each engine, a cycle which is the operational cycle number, settings 1-3 which is the operational settings, s1 to s21 which is sensor readings, and RUL which is the target variable.

Table 2: Dataset Description

ID	Cycle	Setting1	Setting2	Set-ting3	S1	S2	S3	…..	S21
1	1	-0.0007	-0.0004	100	518.67	641.82	1589.70		23.4190
1	2	0.0019	-0.0003	100	518.67	642.15	1591.82		23.4236
1	3	-0.0043	0.0003	100	518.67	642.35	1587.99		23.3442

B. Data Mining Algorithm

Data mining is the use of machine learning and statistical analysis to uncover patterns and other valuable information from large data sets. Data mining methods are used for identifying patterns in the dataset and creating prediction models. In this research, three algorithms were implemented and evaluated.

Random Forest is used as it reduces over-fitting by averaging multiple decision trees and increases predicted accuracy by constructing several decision trees during training and combining their output. It is effective for large and complex datasets. Random Forest is used in aircraft maintenance as it can analyze sensor data to predict the remaining useful life (RUL) of aircraft components and can classify the failure modes by learning from historical and failure records, so that it can predict when an aircraft engine will fail based on sensor readings.

Support Vector Machines (SVM) are effective for binary classification tasks as it is able to create a clear decision boundary between failure and non-failure cases. SVM works well with kernel functions to model a non-linear relationship. On the basis of operational data, it can classify whether an aircraft component is healthy or nearing failure.

Extreme Gradient Boosting (XGBoost) is highly used in predictive maintenance for its high accuracy and efficiency. XGBoost can handle missing data, outliers, and feature interactions well.

C. Data Preprocessing

1. Data Cleaning:

   The data is retrieved from the Kaggle data repository, which was almost in the working form. But for better working with the data, it had to be worked with to ensure that there are no missing values in the dataset by validating the integrity of each column.

2. Data Transformation:

   A standard data analysis was done on the dataset to identify some data patterns, standardize their range, and ensure that all attributes contributed equally to the model's performance. The dataset was evaluated for data consistencies, missing values, duplicate entries, and outliers to ensure high data quality. Although there were no missing or duplicate values found, the numerical features were scaled using Robust Scaler to mitigate the impact of the outliers and to maintain data integrity.

3. Outliers evaluation:

   Boxplots are a standardized way of displaying the distribution of the data based on a five-number summary including minimum, first quartile (Q1), median, third quartile (Q3), and maximum. Interquartile Range (IQR) is the difference between Q3 and Q1, and was used to identify potential outliers:

Journal of Advanced College of Engineering and Management

$$\text{Lower bound} = Q1 - 1.5 * IQR$$
$$\text{Upper bound} = Q3 + 1.5 * IQR \quad (1)$$

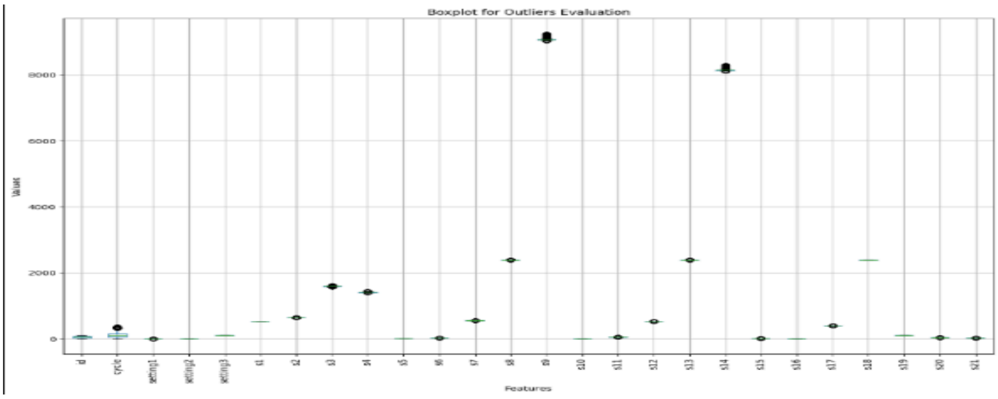

Figure 2: Box-plot for outlier detection

The boxplot revealed the presence of outliers in some sensor readings and operational setting using the lower and upper bound defined in (1). These outliers were not removed but handled during data transformation by applying the robust scaler.

4. Data Correlation:

Each of the attributes is plotted against each of the attributes, creating a pair plot that shows the distribution and relationships between the variables. A pair plot allows us to examine both the individual distribution of each variable and the potential relationships between multiple variables. On the diagonal of the pair plot, we can see the distribution of individual variables, as the data is plotted against itself. The off-diagonal plots show pairwise relationships between different attributes, which help to identify linear or non-linear correlations between features.

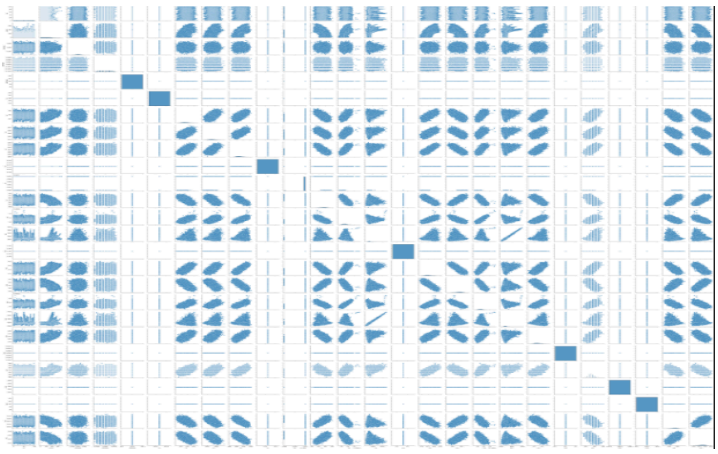

Figure 3: Correlation pair plot of dataset features

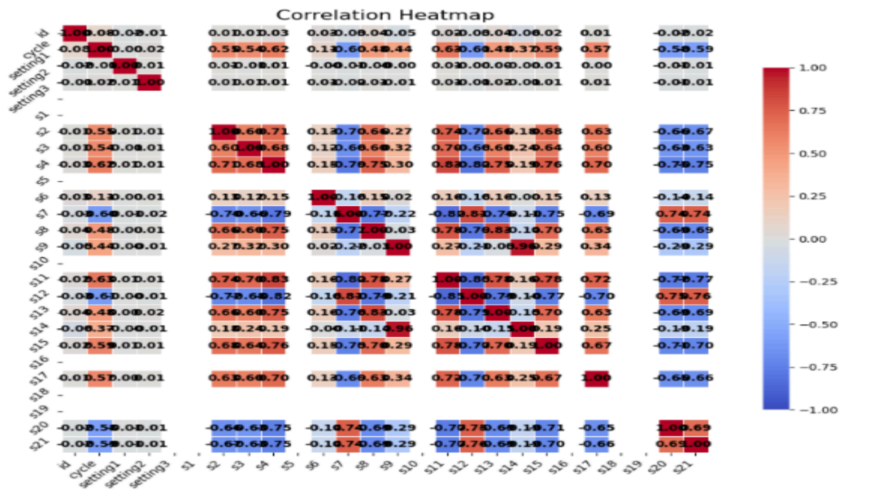

Figure 4: Correlation Heat map

A correlation heatmap shows the strength and direction of linear relationship between numerical features. The correlation coefficient, which ranges from -1 (strong negative correlation) to 1 (strong positive correlation, is shown in each cell. The heat map helps to identify which features are strongly related to the target variable.

5. Data-Set Split:

   The pre-processed dataset was split into three sections for testing, training and validation in order to evaluate the performance of different data mining algorithms.

The training set consisted of 80% of the data and was used to train the model. The testing set consisted of 20% of the data and was used for initial model evaluation. A validation set was created by further splitting the training data by enabling hyperparameter tuning and preventing over fitting during cross-validation.

D. Model

1. Robust Scaler:

Robust Scaler is used to scales features using statistics that are robust to the outliers. This removes the median and scales the data according to the quantile range (defaults to IQR: Interquartile Range). The IQR is the range between the 1^{st} quartile (25^{th} quantile) and the 3^{rd} quartile (75^{th} quantile). Centring and scaling happen independently on each feature by computing the median and IQR from the training set. These statistics are then stored and applied to later data using the transform method, ensuring the consistent scaling.

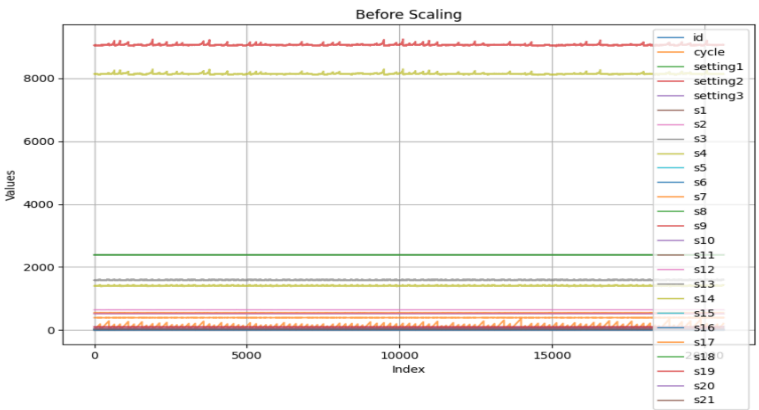

Figure 5: Before Scaling

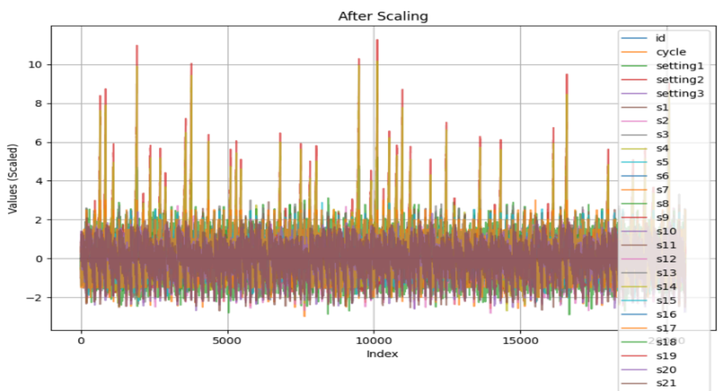

Figure 6: After Scaling

2. Model Pipelines:

   A pipeline is used to help automate machine learning workflows. Model pipelines provide a structured and automated framework for machine learning workflows by integrating data preprocessing, feature transformation, and model training into a single process. A pipeline can preprocess raw sensor data, transform it into meaningful features, and train a classifier to predict equipment failure.

3. Score Machine Learning Algorithm:

   The machine learning algorithms were evaluated using multiple performance metrics such as accuracy, precision, recall, F1- score, and ROC-AUC. These metrics provide a comprehensive assessment of the model's predictive capabilities. The algorithm with the highest overall score was identified as the most suitable for the predictive maintenance task.

```
Accuracy of Random Forest: 0.4000
Accuracy of SVM: 0.4000
Accuracy of XGBoost: 0.4000
```

Figure 7: Scaled Scores

4. Model Accuracy:

   The classified dataset result from the comparison between the three algorithms is shown in Fig. 8. This bar chart compares the scores of three different models across the trials. Each bar represents the average score for a specific model across the trials. XGBoost achieved the highest overall score, followed by Random Forest, and finally SVM. The sign of the score is determined by the position of the data point relative to decision boundary. The accuracy values indicated how effectively each method worked with the specified dataset. A higher accuracy means that the model is doing its best of accurately predicting the target variable.

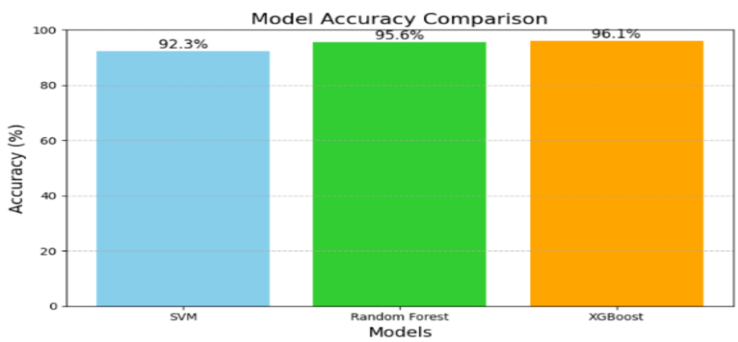

Figure 8: Model Accuracy

5. Classification Report:

   In dataset, the classification report analyzes a machine learning model's performance on several dimensions. It provides the metrics like precision, recall, F1-Score, and support.

```
SVM Classification Report:
              precision    recall  f1-score   support

           0       0.25      0.10      0.14        10
           1       0.44      0.70      0.54        10

    accuracy                           0.40        20
   macro avg       0.34      0.40      0.34        20
weighted avg       0.34      0.40      0.34        20
```

Figure 9: SVM Classification Report

```
Random Forest Classification Report:
              precision    recall  f1-score   support

           0       0.40      0.40      0.40        10
           1       0.40      0.40      0.40        10

    accuracy                           0.40        20
   macro avg       0.40      0.40      0.40        20
weighted avg       0.40      0.40      0.40        20
```

Figure 10: Random Forest Classification Report

```
XGBoost Classification Report:
              precision    recall  f1-score   support

           0       0.42      0.50      0.45        10
           1       0.38      0.30      0.33        10
...
    accuracy                           0.40        20
   macro avg       0.40      0.40      0.39        20
weighted avg       0.40      0.40      0.39        20
```

Figure 11: XGBoost Classification Report

4. RESULTS AND DISCUSSION

The evaluated model's accuracy scores demonstrate XGBoost's outstanding performance, with the highest accuracy of 96.1%, followed closely by Random Forest having 95.6% and SVM having 92.3%. The benefit of Random Forest and XGBoost over SVM is demonstrated by a bar graph. The accuracy of XGBoost is due to its capacity to manage missing data and model complex relationships. A detailed analysis is done using the confusion matrices and classification reports which highlights XGBoost's superior ability to balance precision and recall, ensuring the fewer false positives and negatives. In comparison with SVM, which struggled with more complex patterns and showed lower recall, Random Forest demonstrated the strong classification strength. These findings show that XGBoost provides a more reliable solution for this predicting task, even if all models did well. Confusion matrices and other visualizations highlight XGBoost's exceptional classification accuracy while offering an in-depth overview of the strength and weakness of each model.

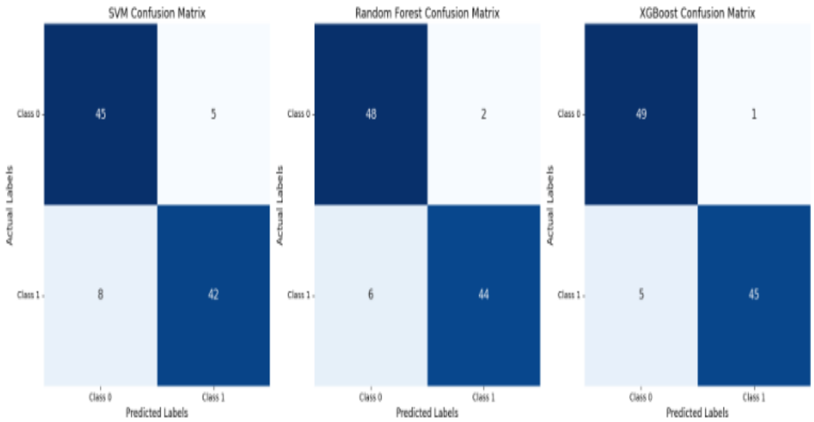

Figure 12: Confusion matrices of all three models

5. CONCLUSION

This research focuses on applying machine learning techniques to predict the maintenance needs and potential failures in the field of aircrafts. We used the machine learning algorithms like Support Vector Machines (SVM), Random Forest, and XGBoost and after being trained on historical aircraft data, all these predictive models showed the potential to effectively forecast mechanical issues before they occur which helps in reducing downtime and improving safety. Among the three models that has been tested, XGBoost achieved the highest accuracy which makes it a promising tool for real-time monitoring systems. When used in predictive maintenance, these models can help airlines prioritize maintenance schedules, optimize resources, and reduce operational costs which ultimately leads to enhance the overall safety and reliability of aircraft operations. This approach helps in more efficient and cost-effective maintenance strategies in the aviation industry.

6. FUTURE WORK

This study shows how machine learning algorithms can be used for predictive maintenance in aviation but there are still a number of areas that is need for more investigation. Including real-time data from sensors integrated into aircraft systems is a important way to include diverse and high-quality datasets which can help with the early anomaly detection and produce predictions that are more accurate.

ACKNOWLEDGMENTS

I would like to express my gratitude and credit to the Kaggle team for making the aircraft predictive maintenance dataset available to work on, which served as the foundation for this research.

REFERENCES

[1] V. Klisauskaite, “Predictive aircraft maintenance: established practice or future focus? - AeroTime,” Aerotime Hub. Available: https://www.aerotime.aero/articles/28331-predictive-aircraft-maintenance-MRO

[2] M.Durgut, “Predictive Maintenance in Commercial Aircraft: A Game-Changer for Aviation | aviationfile-Gateway to Aviation World.” Available: https://www.aviationfile.com/predictive-maintenance-in-commercial-aircraft/

[3] I. Stanton, K. Munir, A. Ikram, and M. El-Bakry, “Predictive maintenance analytics and implementation for aircraft: Challenges and opportunities,” *Syst. Eng.*, vol. 26, no. 2, pp. 216–237, Mar. 2023, doi: 10.1002/sys.21651.

[4] “Shah D. Millions of data points flying in tight formation... - Google Scholar.” Available: https://scholar.google.com/scholar?

[5] F. A. Adryan and K. W. Sastra, “Predictive Maintenance for Aircraft Engine Using Machine Learning: Trends and Challenges,” *AVIA*, vol. 3, no. 1, 2021, doi: 10.47355/avia.v3i1.45.

[6] A. Muneer, S. M. Taib, S. M. Fati, and H. Alhussian, “Deep‑learning based prognosis approach for remaining useful life prediction of turbofan engine,” *Symmetry (Basel).*, vol. 13, no. 10, p. 1861, Oct. 2021, doi: 10.3390/sym13101861.

[7] Z. Zhao, Bin Liang, X. Wang, and W. Lu, “Remaining useful life prediction of aircraft engine based on degradation pattern learning,” *Reliab. Eng. Syst. Saf.*, vol. 164, pp. 74–83, Aug. 2017, doi: 10.1016/j.ress.2017.02.007.

[8] AWS, “Aircraft Predictive Maintenance,” 2019 Available: https://www.kaggle.com/code/sharanharsoor/aircraft-predictive-maintenance