Comparative Study on the Prediction of Remaining Useful Life of an Aircraft Engine

[1]Arpitha V, [2]A. Soumi Narayani, [3]Abhilasha M S, [4]Varshini S A, [5]Sparsha Yadav M, [6]Dr.Revathi V, [7]Akshatha Venkatesh Prasad

[1,2,3,4,5]Student, Department of Computer Science, Dayananda Sagar University, Bangalore, India
[6]Associate Professor, Department of Computer Science, Dayananda Sagar University, Bangalore, India
[7]Student, System Engineering and Engineering Management, Fachhochschule Südwestfalen, Soest, Germany

---***---

Abstract - *Aviation Safety is a major component of the Aviation Industry. The practice of preventive maintenance and Prognostic Health Management is a new area for research in the Aviation industry. The improvement of different safety systems and scheduled maintenance of airplane engines is beneficial in reducing operational and maintenance costs. The abundant information available from flight systems through different sensors recording data at different situations is helpful to obtain data patterns. The study of these data patterns obtained from the complex sensor data, post data-normalization and the removal of noise, helps to evaluate the 'Remaining-Useful Life' (RUL) of the airplane engine. 'Supervised Machine Learning' is a division of machine learning where the model is trained with labelled training data. To train the model we will be using the CMAPSS dataset for RULprediction. This study aims to introduce a comparison between different supervised machine-learning models based on the RUL of the aircraft engines. The predictions made would be helpful to schedule maintenance of airplane engines only when required, preferably, once the threshold RUL value is reached.*

***Key Words*: Remaining Useful Life of an aircraft engine, PHM08 Prognostic Data Challenge, Supervised Machine Learning, CMAPSS.**

1. INTRODUCTION

Airplane systems are designed taking into accordance different failure scenarios that can occur during operation. Aviation safety is mainly dependent on the efficient and reliable operation of the engine as it is the most critical component. It provides thrust which is mainly required to keep the airplane air-borne. Hence, the prognostication of the RUL of the engine, ensures airplane's safety. 'Prognostics Health Management' (PHM) is a novel technology aimed at predicting the occurrence of failures in components and consequently minimizing unexpected down-times of airplane engines, thereby, reducing the maintenance and the operational cost.

The continuous operation of an aircraft engine makes it very vital for the different sensors to record data and report it to the main system. This ensures the engine is running both optimally and safely. The performance degradation process of the engine are due to different responses of various sensors due to noise/insensitivity, they show an uncertain tendency to degradation trend. The most sensitive sensors for this process are Temperature measurement sensors, Pressure measurement sensors, RPM measurement sensors, and Air Mass flow measurement sensors which are chosen as inputs to the RUL prediction.

Currently, approaches for the RUL prediction of systems are categorized as:

1. 'Physics based' models

Usually, Physics based model is used in the scenario with failed physical model of the components for predicting the RUL which is mostly dependent on the physics of failure of the components and they use failed historical samples or limited historical samples for the accurate prediction of RUL. However, for a complex system it is not economical to implement a physics-based approach.

2. 'Data driven' approach

This is a robust approach requiring less prior knowledge, but quite a bit of Experimentation and Computation. The data-driven method uses a physics-based model. For RUL estimate, it makes use of the monitored operational data on system health. When the failure physics of a system is sophisticated or unavailable, the data-driven method relies on the system's degradation procedure and easily available data. In comparison to a physics-based model, the data-driven method for a system gives precise 'RUL' predictions which can be applied conveniently and affordably. Furthermore, data-driven method can be separated bifurcated into 'statistical techniques' utilizing regression methods and 'AI techniques' involving neural networks and SVM.

3. Hybrid

The hybrid approach, combines physics-based and data-driven models to produce predictive predictions that are said to be more accurate and dependable.

2. LITERATURE SURVEY

Under RUL prediction, the journal paper [1] suggests a hybrid model of 'Deep convolution neural networks' (DCNN) with the 'Light gradient boosting machine' (LightGBM) algorithm to evaluate the RUL of high-dimensional and complicated data. Unlike the conventional 'Prognostics and Health Management' (PHM) methods, raw sensor data processing is not required here. Firstly, after normalization, DCNN uses the time window of raw sensor data asits input. The DCNN extracts info from the input data. Secondly, due to the constraintsof DCNN's connected layer, it is substituted by a powerful 'classifier-LightGBM' to increase accurateness of prediction. As the model facilitates extraction of 'failure features' from the sensordata as the engine degrades, the suggested method's prediction findings became increasingly accurate as the engine degrades.

The Fault Diagnosis and RUL prediction's research progress is discussed in the study [2]. The applications and structures of ML-based 'Fault Diagnosis' and 'RUL prediction' are studied, with the applications based on DL being summarized in greater detail.

The study by Shixin Ji and Xuehao Han [3] focuses on a PCA–BLSTM hybrid model meant for forecasting the RUL of an aviation engine. To commence, PCA is used to extract the most relevant information and minimize the data dimension from high-dimensional and complicated data that contains noise and some meaningless information that may impair the accuracy of RUL prediction. The RUL prediction was then achieved to mine the internal relationship of state monitoring data by combining it with BLSTM (multiple layers). This input layer also is used by BLSTM for both LSTM layers (forward and backward), and the LSTM outputs are combined to generate the final output. Thus the model's performance improvement is achieved as the BLSTM layer mixes previous and future data to study the internal relationship of the entire sequence.

Zhongzhe Chen, Shuchen Cao, Zijian Mao [4] published a paper that presented dual approaches for RUL estimation of an engine in various scenarios. With lots of run- to-failure data of referenced samples, the chief strategy uses a revised similarity-based method to estimate the RUL of the aero engine. The second scheme uses an SVM-based approach to estimate the RUL of the operational sample with less weakened performance data than the 'reference' systems, based on weakened / deteriorated data of samples without 'tending-to-failure' data. To look at the degradation trend of reference samples and estimate their lives, SVM is employed. The acquired weakening indicators of the reference samples are utilised to train the model using SVM and extract the performance-weakened pattern of these samples, as well as fit their degradation indicator relation curve w.r.t time. This functionis computed using the maximum likelihood estimation approach based on the curve. The reference systems are considered failure whenever they hit a certain threshold value, hence the lifespan of these reference samples is approximated in this way.

A unique architecture 'CNN-based regressor' was investigated in the research [5] to estimate complex system's RUL using multivariate time series data. The convolution and pooling layersare primarily used in this suggested deep architecture that captures significant patterns of sensor signals and at diverse time scales.

Zio et al. [6] study presents a similarity-based approach in which the trajectory patterns of the reference samples is compared to the evolution data using similarity analysis to predict the RUL, and the weighted total of their time to failure accounts for their similarity to the evolving pattern.

It is inferred that when there are many failed historical samples, it is more appropriate to use the modified similarity-based technique, whereas SVM methodology is appropriate when there are few failed historical reference samples. When the combination model of PCA-BLSTM is employed instead of LSTM / BLSTM, the model displays better performance and higher accuracy, providing a smart decision footing for aeroplane engine maintenance and management. Also, the accuracy of the RUL prediction is higher towards the stage of engine failure.

3. PROPOSED METHOD

The 'Remaining Useful Life' of an engine is the time length of the engine before it reaches failure. Prediction of the RUL of an engine is a comparative study of machine learning models to give the best accurate results.

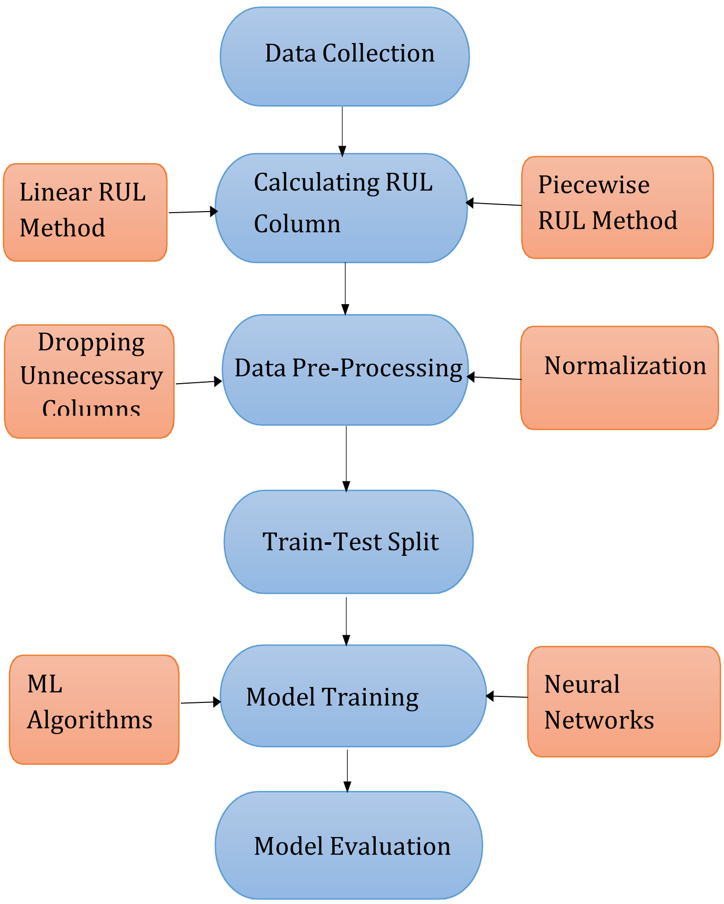


Fig -1: Flowchart of RUL Prediction of an aircraft engine

The above figure 1 shows the steps involved in predicting remaining useful life of an aircraft engine.

3.1 RUL DATASET

For this comparative study, the RUL dataset used is from the PHM08 challenge dataset givenby NASA. The data of turbofan engine from the Prognostic Center of Excellence of NASA AmesResearch Center (PCoE) was amassed and simulated using a tool developed by NASA themselves. This tool is C-MAPSS or the 'Commercial Modular Aero-Propulsion System Simulation' which was coded in MATLAB and Simulink Environments where the user can insert values as per their prerequisites into the customisable input parameters.

An engine of an aircraft is a critical component as its failure can cause loss of lives and forfeiture of property. The air that enters the engine through its inlet gets split. Some portionof it passes to the compressor and then into the burner where it is combined with fuel and combustion takes place. This hot air now goes through the high and low pressure turbine before passing out and providing the thrust required by the airplane to fly. The remaining airgoes around the engine and gets heated thereby generating thrust. Using the thrust from the engine and the free air, the aeroplane can fly. Hence, when such a component is used for an extended period, numerous factors affect its health and condition, therefore we require different sensors to monitor it. The dataset consists data of 21 sensory values, 1 engine ID, 1cycle, and 3 operational settings of many cycles. Based on the sensitivity to the performance degradation process, the sensors have been chosen. The sensors are based on:

- Temperature Measurement
- Pressure measurement
- RPM (Revolution per minute) measurement
- Air Mass flow measurement

3.2 CALCULATING RUL COLUMN

With this dataset, we can create a new RUL column which will be used as a label during the training period. To create the RUL column there are 2 methods:

1. Linear RUL – Calculated by taking the difference between the maximum cycle of the engine and each cycle of the engine. Formula is as follows -

$$\boxed{\textbf{RUL = max cycle - current cycle}} \quad (1)$$

2. Piecewise Linear RUL – Calculated by taking a limit of 120. Given the max life as 120, we calculate the cut off which is the difference between the maximum cycle of the engine and the max life. If the engine's cycle is less than the cut off, we append RUL as 120 otherwise it keeps decrementing by one. Formula is as follows –

```
max life =120
cut off = max cycle – max life
for i in range (max cycle):
    if i < cut off:
        RUL=120
    else:
        RUL = RUL - 1
```
(2)

3.3 DATA PRE-PROCESSING

- Dropping Unnecessary columns-

  Dropping of unnecessary columns is done based on the standard deviation. Having standard deviation near or equal to zero, have no changes in the values of that particular feature column. Therefore, by dropping those columns, the accuracy of prediction will not be affected.

- Normalizing the data-

  Data normalization is done to bring the dataset values to a standard range of 0 to 1 or -1 to 1, for simplifying the process of building the model. In this project, we have used minmax scaler for normalization. Minmax scalar normalization brings all the dataset values between the ranges of 0 to 1.

$$x_{minmax\_scaled} = \frac{x - min(x)}{max(x) - min(x)}$$
$$x \rightarrow Data\ set$$
$$max(x) \rightarrow maximum\ of\ dataset$$
$$min(x) \rightarrow mimimum\ of\ dataset \quad (3)$$

- Feature Selection-

  Recursive Feature Elimination is used to identify the key attributes (plus its combinations) that contribute the most to the target attribute prediction.

3.4 SPLITTING OF DATA

The RUL dataset is split into train and validation based on the engine ids as the model needs to train according to the degradation of RUL to retain the sequence of the number of cycles in each engine id for better training and testing accuracy.

The data is distributed in the proportion of 70:30 which has 70% of training and 30% testing data. In our dataset we have 218 engine ids which in turn have records for the different number of cycles in each engine id. Hence, we have 154 engine ids for train data and the remaining 64 engine ids records for test data.

3.5 BUILDING OF MODEL

After the creation of the RUL column, the dataset is used for the model training. The modelsused for the prediction are

1. Simple linear regression
2. Support vector regression
3. Random forest
4. Decision tree
5. Gradient boosting algorithm.
6. K-nearest neighbours
7. ANN
8. CNN

1. Simple Linear Regression –

This is a 'supervised machine learning' model where based on fitting a line to the observed data we find the association between the variables. (Relationship between independent and dependant variables).

$$y = \beta_0 + \beta_1 x$$
$$x(Cycle) \rightarrow independent\ variable$$
$$y(RUL) \rightarrow dependent\ variable$$
$$\beta_0 \rightarrow Intercept$$
$$\beta_1 \rightarrow Slope \quad (4)$$

In this project, for linear regression model the cycle column is the independent variable and RUL column is the dependent variable. It shows a negative linear relationship as thex-axis (cycle) increases, the y-axis (RUL) decreases.

2. Support Vector Regression –

In Support Vector Regression, a regression line is used to predict the target variable for the continuous data. The regression line in this project is the actual RUL with the help of which the target variable i.e., predicted RUL is predicted. Support Vector Regression is similar to Linear Regression. The equation, y= β1x+ β0 is referred to as the hyperplane in SVR. It also shows a negative linear relationship, just like in the Linear Regression model. The R2 score for Support Vector Regression in both Linear and Piece-wise RUL, is slightly higher than the R2 score for Linear Regression, which signifies that it can be preferred over linear regression.

3. Random Forest –

Average for the regression output is taken by building decision trees on different samples. We can notice that Random Forest has the highest R2 score in both linear and piece-wise RUL among all the models we use in this project. This is because it gives high accuracy of results through cross validation. Since Random forest is a collection of various decision trees, each of these trees pick a different sample of features. They hence have their individual predictions averaged to produce a single result. This makes Random forest better, because we have multiple trees in the forest rather than using a single decision tree.

4. Decision Tree –

The decision tree employs the tree representation to create a model to predict target variable value. The leaf node represents a class label and attributes are denoted on tree's internal node. A non-linear relationship is established amongst the dependent and independent variable. The R2 score for decision tree is really low, making it the least desirable model in this project. This is because decision trees can tend to be very unstable because any changes in the input can make the results different and it might also lead to over fitting at times.

5. Gradient Boosting Algorithm –

Here weak prediction models which are normally decision trees are grouped to form a prediction model. It is used to fit the model which predicts continuous values like house rates. Moreover it is very effective for tabular datasets. In this model, after selecting a weak learner we combine simple models one at a time, as more of these models are combined, the final complete model becomes a strong predictor. This model is preferred as it provides better accuracy and high flexibility and also because it requires minimal data pre-processing and handles missing data too.

6. K- Nearest Neighbours –

Values of new data points are predicted using feature similarity in this algorithm. Meaning, based on how closely the new point resembles to those in the training set the new point is assigned a value.

7. ANN –

Neural Networks learns the complex non-linear relationship between the features and target due to the existence of activation function in each layer. As can be seen from the name, this algorithm is loosely based on the working of neurons in human brain. Popularly called as ANN, this algorithm has the ability to rework to changes in input. This input is usually subjective to various criteria. This model is preferred over linear regression as, in linear regression, only linear relationships amid the features and targets are studied whereas in ANN, complex non-linear relationships amid features and targets are understood. It is so because of the existence of activation function in each layer.

8. CNN-

Convolution Neural Networks are a sort of artificial neural network that uses convolution instead of regular matrix multiplication in at least one of its layers. They have three main types of layers, which are Convolutional layer, Pooling layer, Fully-connected (FC) layer. The first layer of a convolutional network is the convolutional layer. The convolutional layers can be trailed by supplementary convolutional layers or pooling layers but the fully-connected layer is the finishing layer. CNN was inspired by the biological process of the connection arrangement between neurons that mirrors the arrangement of the animal visual cortex. In contrast to other Machine learning algorithms, CNNs use minimum pre-processing. This means that the network learns to optimise the filters by automatic learning, rather than hand-engineering them as in traditional methods.

3.6 MODEL EVALUATION

The estimated RUL is compared with the true RUL and based on the MSE, RMSE, MAE, and R2_Score, the accuracy of the model will be decided. Lower the value of Score/RMSE, the better the accuracy.

Comparing the accuracies given by each model, we can find out the most suited algorithm for the estimation of RUL.

3.6.1 R2_SCORE

We use R2 score to calculate the performance of the models used. The extent of the variation in the 'output dependent' attribute (RUL) which is predictable from the 'input independent' variables(cycle and sensor values) and is used to assess well-monitored results are produced by the model, depending on the deviation of results described by the models.

$$y_i = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$R2\ \text{Score} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

$\mathrm{y_i} \rightarrow$ actual data

$\hat{\mathrm{y}}_\mathrm{i} \rightarrow$ predicted data

$\overline{\mathrm{y}} \rightarrow$ mean value of sctual data (5)

3.6.2 RMSE

Root mean square error' is used to evaluate the quality of prediction of the RUL. It shows how far predictions deviate from measured absolute values using Euclidean distance. The residual (variation between prediction and absolute) for each data point is calculated, followed by calculation of the norm of residual for every data point, followed by the average of residuals and finally the square root of that average.

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}}$$

$\mathrm{y_i} \rightarrow$ actual data

$\hat{y}_\mathrm{i} \rightarrow$ predicted data (6)

3.6.3 MSE

MSE is the average of the square of the difference between the absolute values and the predicted values of the RUL.

$$MSE = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}$$

$y_i \rightarrow$ actual data

$\hat{y}_i \rightarrow$ predicted data (7)

3.6.4 MAE

We use it to measure accuracy for continuous variables. The equation for calculating MAE in words, is the average of the absolute values of the differences / variations between anticipated and the actual values of RUL.

$$MAE = \frac{|(y_i - y_p)|}{n}$$

$y_i \rightarrow$ actual data

$y_p \rightarrow$ predicted data (8)

SCORE	LINEAR RUL			PIECEWISE LINEAR RUL		
	SLR	SVR	Random Forest	SLR	SVR	Random Forest
R2_score	0.5278176443492266	0.5304208676024913	0.5909090332061397	0.6689369465039579	0.6999361699190959	0.7690137339715295
MSE	0.01663254600599008	0.01654084789400936	0.014410162185022972	0.036458932714344534	0.03304508574847353	0.02543779090459688
RMSE	0.12896722841865713	0.12861122771363845	0.12004233496988874	0.19094222349795903	0.18178307332772634	0.15949229105068646
MAE	0.10170745033447096	0.10109357492745841	0.08887709458189841	0.1555811505127454	0.14671670438356366	0.11553451359084407

Table 1: Model Evaluation measures of RUL Prediction using SLR, SVR and Random Forest

SCORE	LINEAR RUL		PIECEWISE LINEAR RUL	
	Decision Tree	KNN	Decision Tree	KNN
R2_score	0.19078301015534127	0.19090164113798658	0.5680890967511716	0.5670447150593613
MSE	0.028504535697591038	0.028500356940691356	0.047564989188167 01	0.047680003658757 17
RMSE	0.16883286320379406	0.16882048732512103	0.21809399163701648	0.21835751340120443
MAE	0.12283266791159564	0.12249988775684974	0.14064365349514085	0.14100254057520595

Table 2: Model Evaluation measures of RUL prediction using Decision tree, KNN

SCORE	LINEAR RUL			PIECEWISE LINEAR RUL		
	Gradient Boosting Algorithm	ANN	CNN	Gradient Boosting Algorithm	ANN	CNN
R2_score	0.5278176443492266	0.5587841608413399	0.5958098908821654	0.6689369465039579	0.7242895773236069	0.7303222672949369
MSE	0.01663254600599008	0.015541755543287454	0.014237530277477501	0.036458932714344534	0.03036312159527122	0.029698760424706917
RMSE	0.12896722841865713	0.12466657749087144	0.11932112251180635	0.19094222349795903	0.17425016957027967	0.17233328298592504
MAE	0.10170745033447096	0.0941223240083215	0.08933812518761788	0.1555811505127454	0.10754214509579904	0.12996610967352987

Table 3: Model Evaluation measures of RUL Prediction using Gradient Boosting, ANN and CNN

4. RESULTS

From Table 3, it is noted that as the number of cycles increases, the remaining useful life decreases. The cobalt colour line indicates the absolute values and the crimson curve indicates the predicted values of RUL values. It is inferred that for a given number of cycles, the curve remains almost the same, but after reaching a certain threshold value the RUL values start decreasing drastically which indicates that the engine failure is about to happen and hence it helps in predicting the engine failure before it happens.

We notice from the comparison study of using different algorithms that random forest using the piecewise RUL method is giving the best prediction accuracy of all the algorithms and K nearest neighbours and decision tree algorithm is giving the lowest accuracy compared to the other algorithms as shown in Table 4. Table 5 gives us the prediction curve based on the test data where the engine ids don't run to failure. The degradation will be more evident if each engine ran a higher number of cycles

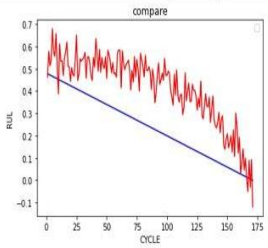
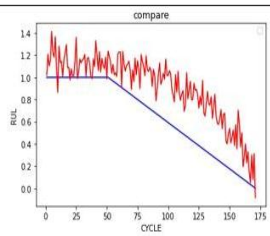
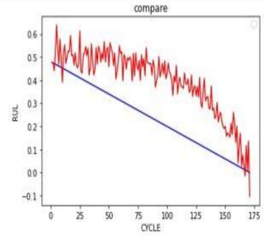
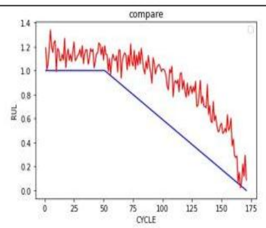
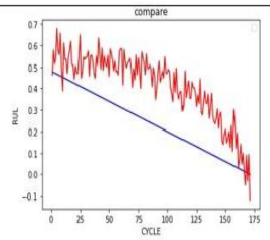
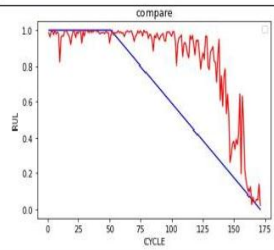
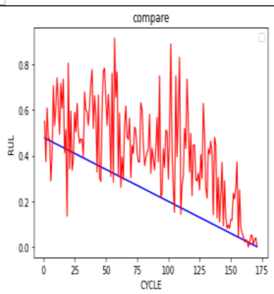
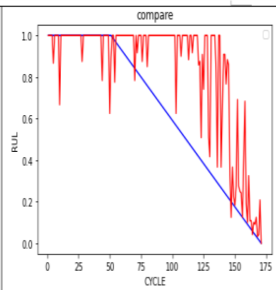
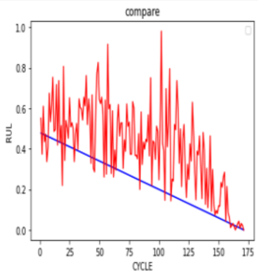
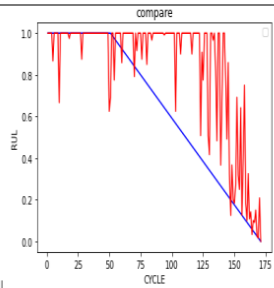
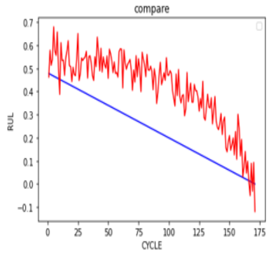
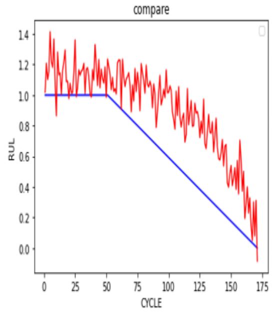
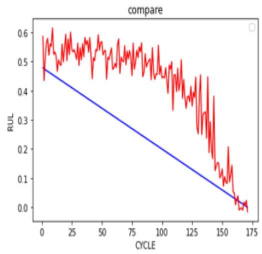
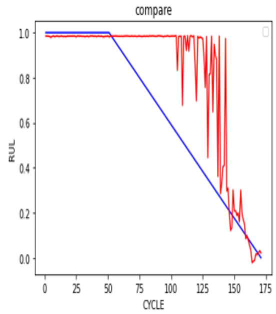
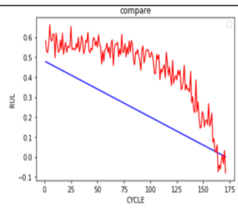
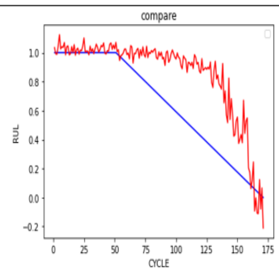
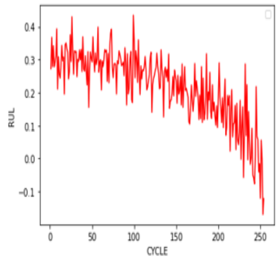
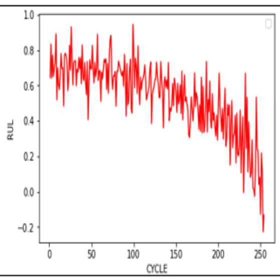
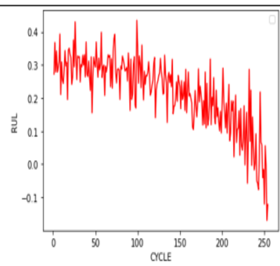
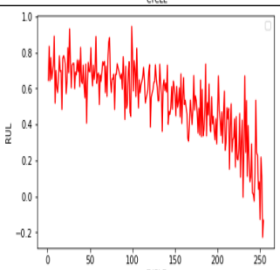
ALGORITHM	LINEAR RUL GRAPH	PIECEWISE LINEAR RUL GRAPH
SIMPLE LINEAR REGRESSION		
SUPPORT VECTOR REGRESSION		
RANDOM FOREST		
DECISION TREE		
KNN		
GRADIENT BOOSTING ALGORITHM		
ARTIFICIAL NEURAL NETWORK		
CONVOLUTIONAL NEURAL NETWORK		

Table 4: Actual RUL vs Predicted RUL for a single-engine id(e_id=154) for training data

ALGORITHM	LINEAR RUL GRAPH	PIECEWISE LINEAR RUL GRAPH
SIMPLE LINEAR REGRESSION		
SUPPORT VECTOR REGRESSION		

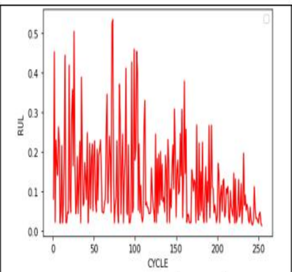
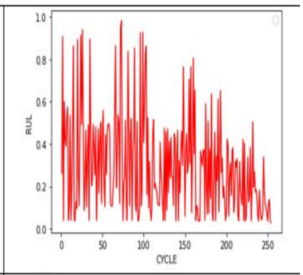
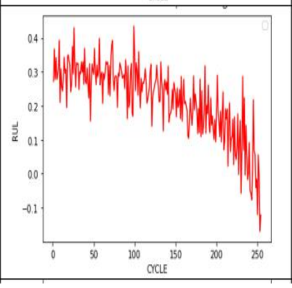
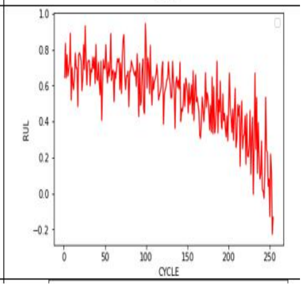
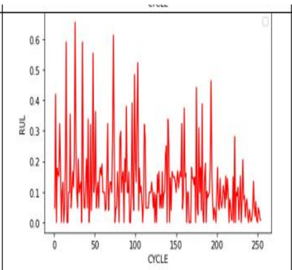
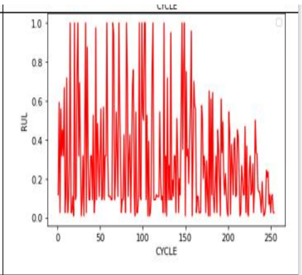
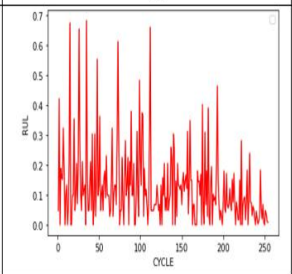
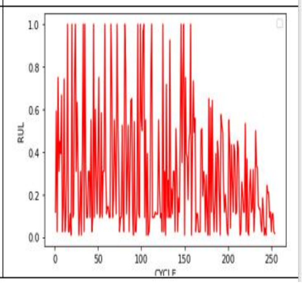
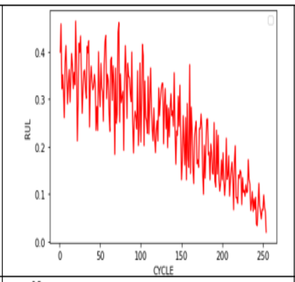
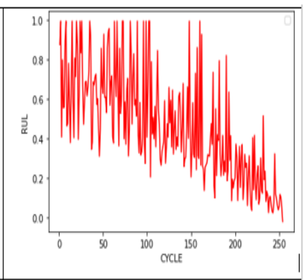
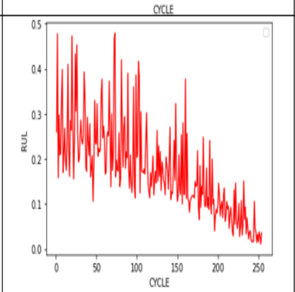
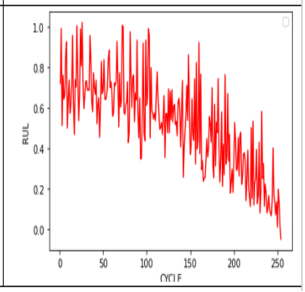
RANDOM FOREST		
GRADIENT BOOSTING ALGORITH M		
DECISION TREES		
K NEAREST NEIGHBOU RS		
ARTIFICIAL NEURAL NETWORK		
CONVOLU TIONAL NEURAL NETWORK		

Table 5: Predicted RUL for a single-engine id(e_id=5) for test data

5. CONCLUSIONS

This paper presents the comparative study of RUL estimation. For the comparative study, we have used various machine learning and deep learning algorithms such as simple linear regression, support vector regression, random forest, gradient boosting algorithm, K- nearest neighbours, decision tree, artificial neural network, and convolution neural network. RUL estimation helps in predicting the degradation of the engine, if the engine starts degrading then the engine needs to be changed before it runs to failure. In this project, we have used the training data and the testing data where the training data is further split into train and validation as the testing data does not run to failure and therefore cannot be compared with training data for accuracy. We train the models using the train data and evaluate based on the validation data from this we see that random forest based on piecewise linear method gives the best accuracy as seen in table 4 and evaluating the data based on the testing data we can see that there is a degradation trend of the engine as shown in Table 5. Therefore, estimation of RUL helps in the process of pre-aircraft check and reduces operation and maintenance costs.

REFERENCES

[1] Liu, L., Wang, L. & Yu, Z. "Remaining Useful Life Estimation of Aircraft Engines Based on Deep Convolution Neural Network and LightGBM Combination Model", Int J Comput Intell Syst 14, 165 (2021). https://doi.org/10.1007/s44196-021-00020-1

[2] Daan Ji, Chuang Wang, Jiahui Li & Hongli Dong (2021) A review: data driven-based fault diagnosis and RUL prediction of petroleum machinery and equipment, Systems Science & Control Engineering, 9:1, 724-747

[3] Shixin Ji, Xuehao Han, Yichun Hou, Yong Song, Qingfu Du (2020)." Remaining Useful Life Prediction of Airplane Engine Based on PCA–BLSTM", Sensors 20, no. 16: 4537. https://doi.org/10.3390/s20164537.

[4] Zhongzhe Chen, Shuchen Cao, Zijian Mao (2017). "Remaining Useful Life Estimation of Aircraft Engines Using a Modified Similarity and Supporting Vector Machine (SVM) Approach" Energies 11, no. 1: 28. https://doi.org/10.3390/en11010028.

[5] Babu, Giduthuri Sateesh, Peilin Zhao, and Xiao-Li Li. "Deep convolutional neural network- based regression approach for estimation of remaining useful life." International conference on database systems for advanced applications. Springer, Cham, 2016.

[6] Zio, E.; Maio, F.D. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of an unclear system. Reliab. Eng. Syst. Saf. 2001, 95, 49–57