# Regression and Classification Model Based Predictive Maintenance of Aircrafts Using Neural Network

## Humaira Maqbool[1], and Dr. Monika Mehra[2]

[1] M.Tech, Department of Electronics and Communication Engineering, RIMT University, Punjab, India
[2] Head of Department, Department of Electronics and Communication Engineering, RIMT University, Punjab, India

Correspondence should be addressed to Humaira Maqbool; 1996humairakhan@gmail.com

**ABSTRACT-** One of the key objectives of today's businesses and mills is to predict machine problems. Failures must be avoided, because downtimes represent expensive expenses and a loss of productivity. This is why the number of remaining cycles (RULs) until the failure occurs is vital in machine maintenance. The estimations of the RUL should be based on earlier observations, whenever possible under the same conditions. In the research of RUL estimates, the creation of systems that monitor current equipment conditions is becoming crucial. I employed Long Short Term Memory (LSTM) in my project to determine an aircraft's remaining usable lives. The aircraft's functioning condition is also forecast. The former is done by a regression method, using a classification methodology predicted by working circumstances. In order to estimate operating conditions and remaining usable life of the aircraft, data utilized for LSTM models training are derived from 21 aircraft sensor readings located at different locations with three distinct settings.

**KEYWORDS-** Artificial Intelligence, Long Short Term Memory, Neural networks, Regression, Classification, Remaining useful life.

## I. INTRODUCTION

Airlines experience cost pressures from several factors in the civil aviation business. The so-called 'low price' airlines for example allow cheap price public transport, leading to a significant decrease of the operating expenses of other airlines on the market. Efforts to reduce fuel consumption or new conceptions of finance for aircraft, for example, leasing, led to a position whereby only modest further improvements are predicted. The objective of reducing maintenance costs, by contrast, has lately been highlighted.

Maintenance expenditures are projected to represent up to 20% of the operational costs of airlines at an estimated US$40 billion a year, depending on the particular kind of aircraft and their age, globally. Civil aviation has an equally optimising problem between quality, time and cost of aircraft maintenance. The airworthiness and availability of aircraft are ensured by maintenance, repair and overhaul (MRO) operations (quality). The required ground times (time) should be minimised to further increase the operation of the aeroplanes while being as economical as possible (costs).
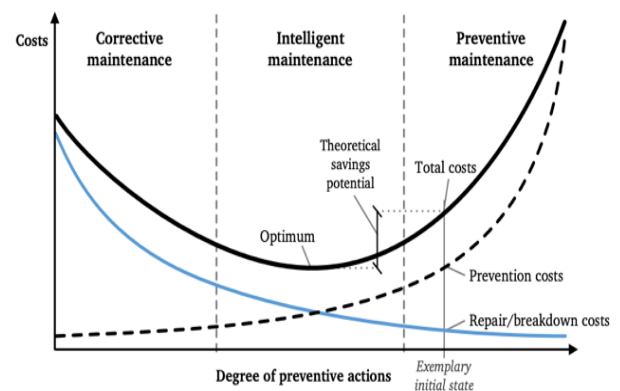


Figure 1. Impact of Maintenance Strategies on costs

Fig. 1 depicts the classification of the levels of prevention action ("act before failure") and their influence on costs in relation to the maintenance solutions. The degree of preventative maintenance measures is represented on the x-axis as part of all operations. If no preventative maintenance (left) is carried out, it is thought that maintenance is still necessary, but that the system and the component failure is remedial. Due to the nature of these accidents, most maintenance is performed on an as-needed basis, resulting in higher efforts (e.g. ad-hoc troubleshooting) and operational disruptions (e.g. flight delays), resulting in so-called breakdown costs. As a result, for larger degrees of preventative activities, the qualitative curve for repair and breakdown costs (thin line) shows a declining behaviour. Because most issues are repaired before any defect or failure, it is anticipated that nearly no corrective measures are

required if all maintenance operations are preventative (right). The appropriate preventive costs (dashed line) are hence directly proportionate to the frequency of action and hence likely to be infinitely uneconomic to prevent any corrective activity. When both qualitative cost curves are included (thick line), the quality relationship between the two classic techniques becomes optimal. The notion of forecasting consists of constant monitoring and evaluation of the present status of the system to provide information on whether maintenance actions are to be made (diagnostic) and when exactly such action is to be done (prognosis). It aims at overcoming the drawbacks of traditional maintenance concepts: first, it is believed to enhance the planning of current ad-hoc faults that have been corrected and so to lessen the impact of a breakdown. In contrast, it may minimise the number of useless preventative measures without providing any value or potentially causing further failures. However, there are still heated discussions about the particular influence of predictive maintenance principles on real-life maintenance. Due to recent improvements in technology and information technology, computerised monitoring and communication networks, data from a variety of equipment components for the automated failure detection and diagnosis (FDD) may now be collected from a large array of operations and process conditions[6].

Aircraft engine with minimum number sensors were created in the previous decades since there were no requirements of many additional sensor values that are important for the engine status of an aircraft. Now that we have all 21 new sensors within an aircraft engine, we need predictive maintenance to save us a great deal of time and effort to avoid needless maintenance. These Aircraft Engine sensors supply a great deal of historical data, which truly displays the engine's true status. Therefore, the manual time for engine work is reduced so that the maintenance service can be located and can handle the closed data and do predictive maintenance as necessary. The thorough study strategy to deploying Recurrent Neural Networks (RNN) and employing the Long term Memory Technique (LSTM) predicts the time-stamped data set, meaning that the data present looks at 60% of times before, to estimate the current useful life(RUL) of the aircraft engine and so 96%. The requirement to create the neuronal long-term memory (LSTM) network is necessary since it looks back and trains itself with all the data. This is obviously the greatest approach for neural networks to forecast future value in series of times. The first approach is to evaluate how long the aircraft has been able to remain usable (RUL) based on the number of cycles. Long-term memory is obviously meant to prevent long-term problems with dependency. The LSTM networks may be used to learn from and transform real-time data to a time series. The data were gathered from NASA's Prognostic Data Repository with 3 aircraft types Id1, Id2, Id3 and its sensor values as well as the number of death cycles. Based on the aforementioned data set of the NASA data repository, we present a method using LSTM to create a model that can provide an estimated RUL and the chance of an aircraft engine failure after n

series of cycles. Based on this very accurate methodology, greater service for the aircraft industry may be implied and used [2].

## II. OBJECTIVES

Following are the objectives of my work:
- To calculate the Remaining Useful Life of an aircraft. (Regression)
- To predict the working condition of an aircraft after the predefined working cycles. (Classification)

## III. LITREATURE REVIEW

In the last year the forecast has been based on machine learning, which self-learns data that involve multiple learning algorithms. In the data set the input variable (sensor readings) is separated into a train-test set that is expected to use RUL. The learning algorithm in the machine is taught using an exercise set and is tested in comparison with the test set. We utilise Deep Learning Approach to handle this scenario to overcome the drawback of low accuracy and less reliability of the machine learning technique.[1]

### A. Hardware and Software Required

1) *Hardware:* Minimum 2.8 GHz Processor Computer System or latest

2) *Software:* Windows 2000(Service Pack 4),XP, Vista or Windows 7, Anaconda and Spyder

3) *Libraries:* Keras, Tensor flow, Scikit Learn, Matplotlib, Numpy and Pandas

## IV. METHODOLOGY

In order to increase the performance of our application, we have set up the LSTM-neural network. This LSTM methodology involves time stamps to examine previous data and to prepare the data set in such a way that 45 values are given in each row for the original data sets.

### A. Input data

I downloaded the NASA Dataset, which comprises 21 sensor readings and three settings, setting2, and setting3.

### B. Data Pre-processing

One of the key issues of providing data for neural networks is pre-processing data, where data are manipulated and distinct non-used values, redundant values or null values are removed that are unwanted when the model's predict value is formed. The steps covered in the pre-processing of data are given below

- Null values are removed.
- The data are measured using the Scikit Learn package in order to make it easier to analyse the data for the LSTM network. The scaling process is carried out using the Min-

Max scaler package which scales data from 0 to 1 and so makes this data set totally comparable [1].

### C. Training and Testing

Data set is split into training data and testing data. 75% of the whole data is kept for training the LSTM model and rest 25% is kept for testing the Model. Scikit Learn libraries is used for this purpose.

### D. Output Accuracy

For Classification problem, accuracy is measured by plotting the confusion matrix and calculation values like Precision, Recall and F1 score.

A confusion matrix is a table typically used to describe the performance of a classification model on a series of test data that are known to be true values. With the exception of AUC, most four parameters may be computed using the left. The observations successfully anticipated and so represented in green are true positive and true negatives. We want to reduce false positive and false negative such that they are displayed in red. These are quite perplexing terms.



Figure 2: Confusion Matrix

- *True Positives (TP)* — These are the positive values correctly predicted which means that the true value of the class is yes and the predicted class value is yes as well. For example, if the actual class value shows that the passer survives and predicts the same class that informs you.
- *True Negatives (TN)* —These are the correctly forecast negative values, meaning that the true class value is no and the forecast class value is no. For example, if the class actually.

These values arise, if the real class contradicts the expected class, with false positive and false negatives.

- *False Positives (FP)* **-** When the class is not real and the class that is expected is yes. For instance, if the class really states this passenger hasn't survived, but the class forecasted informs you that it will survive.

- *False Negative (FN)* - If the real class is yes but the class is not expected. For example, if a true class figure 2 shows a survival and a forecasting class of this passenger, the passenger will die.
- *Accuracy* **–** Accuracy is the most basic measure of performance and is only a relationship to the total observations. You may believe that our model is superior if we have high accuracy.

$$\text{Accuracy} = TP+TN/TP+FP+FN+TN$$

- *Precision* -The ratio of positive observations accurately foreseen for total positive observations is precision.

$$\text{Precision} = TP/TP+FP$$

- *Recall-* (Sensitivity) – Recall is the proportion of positive observations that were anticipated correctly to all actual observations.

$$\text{Recall} = TP/TP+FN$$

- *F1 score* **-** F1 Score is the Precision and Recall weighted average. This score therefore takes into consideration both false positives and false negatives. It's not so easy to comprehend intuitively as accuracy, but F1 is frequently more valuable than precision particularly if the class distribution is unequal. Accuracy works best if false and false negatives are equally expensive. If the costs of false positives and false negatives differ much, both Precision and Recall should be considered.

$$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

For Regression problem,

The RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

It is one of the most frequent regression measures in both machine and statistics. Why is it so famous? It is quite easy to discern, one of the key reasons. This facilitates the use of methods such as gradient descent in combination with derived techniques.

Another essential feature of the RMSE is that because the mistakes are squared, the weight of the mistakes is significantly higher. So a 10-fold inaccuracy is 100-fold worse than a 1.

The MAE is simply defined as:

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j|$$

The inaccuracy is linear when utilising the MAE. Thus, a 10-fold inaccuracy is 10-fold worse than a 1. The error is specified in the same unit as the target variable in both

circumstances [4].

# V. RESULTS AND OBSERVATIONS

## A. Predicting Remaining Useful Life (Regression)

Our LSTM neural network methodology offers a lower MAE and $R^2$ error value which demonstrates that the RUL for this collection of data is close to the actual value predicted by the LSTM neural networks model. The actual and forecast points are the most precise. The diagram 5 presents the training data to clearlydemonstrate the model's precision.

Mean Absolute Error = 12. Coefficient of Determination $(R^2)$= 0.8134

The following plots (fig.3&4) show the trend of loss Function, Mean Absolute Error, $R^2$ and actual data compared to predicted data [5]:
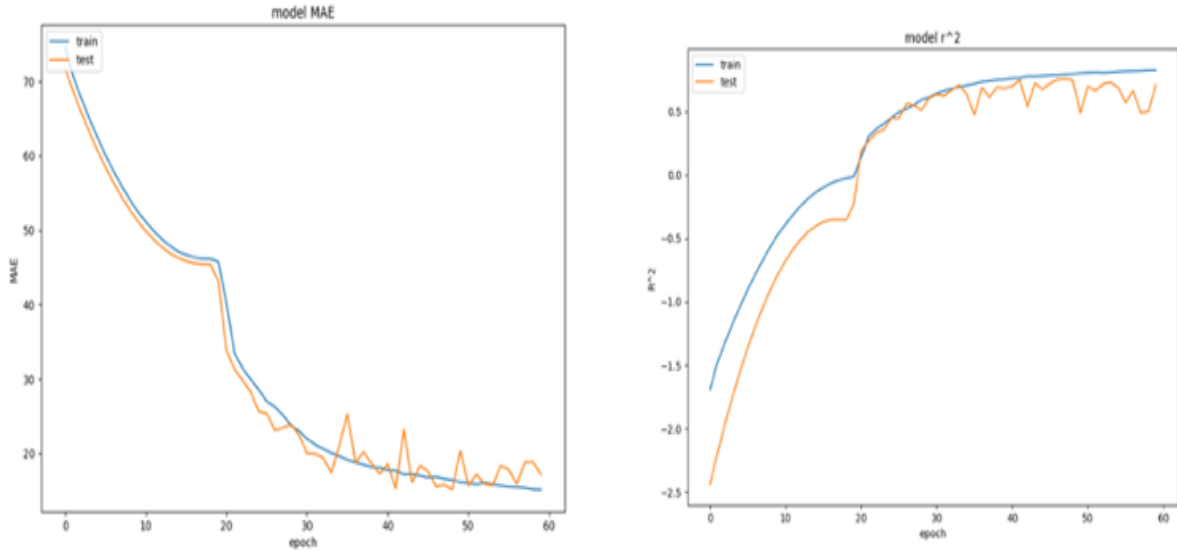


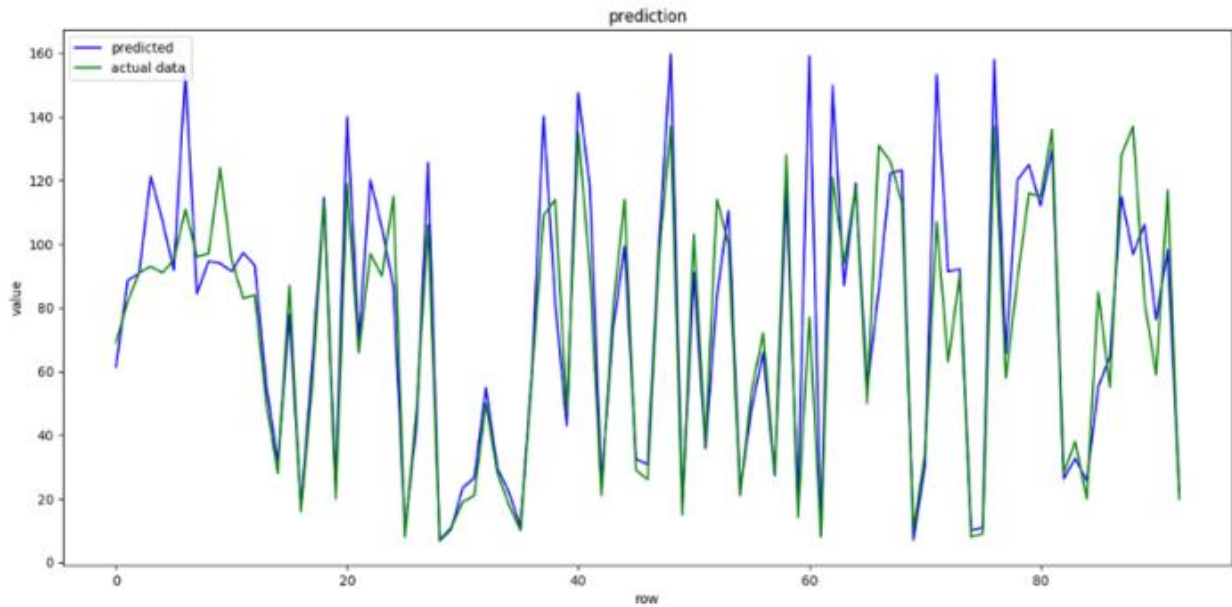Figure 3: Mean Absolute Error and Coefficient of Determination



Figure 4: Loss Function

### B. Predicting Working Condition of an aircraft (Classification)

```
Confusion matrix
- x-axis is true labels.
- y-axis is predicted labels
[[33 15]
 [ 4 41]]
Precision:  0.7321428571428571
 Recall:  0.9111111111111111
 F1-score: 0.8118811881188118
```

```
Confusion matrix
- x-axis is true labels.
- y-axis is predicted labels
[[45  3]
 [ 4 41]]
Precision:  0.9318181818181818
 Recall:  0.9111111111111111
 F1-score: 0.9213483146067416
```

Figure 5: Classification Results for Epoch =1 and Epoch =10

## VI. CONCLUSION

Our proposal LSTM neural network technique performs and has reduced MAE and R^2 values, showing that the RUL for this dataset set is extremely similar to the actual value predicted by the LSTM neural network model. The experimental findings indicate clearly that the suggested approach is able to properly forecast how long the aeroplane engines would last (RUL). In the instance of classification, our model achieved 93% correct results when it was 10 times and may be improved with additional epochs, which means that an aircraft's functioning state was more precisely predicted than the earlier techniques.[3]

## REFERENCES

[1] Charles M. Able, Alan H. Baydush, Callistus Nguyen, Jacob Gersh, Alois Ndlovu, Igor Rebo, Jeremy Booth, Mario Perez, Benjamin Sintay, and Michael T. Munley. 2016. A model for preemptive maintenance of medical linear accelerators-predictive maintenance. Radiation Oncology 11, 1 (2016), 36. https://doi.org/10.1186/s13014-016-0602-1

[2] https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/

[3] https://thedatascientist.com/performance-measures-rmse-mae/

[4] Khalid F Al-Raheem and Waleed Abdul-Karem. 2011. Rolling bearing fault diagnostics using artificial neural networks based on Laplace wavelet analysis. International Journal of Engineering, Science and Technology 2,6 (2011).https://doi.org/10.4314/ijest.v2i6.63730.

[5] H. O.A. Ahmed, M. L.D. Wong, and A. K. Nandi. 2018. Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features. Mechanical Systems and Signal Processing 99 (2018), 459–477. https://doi.org/10.1016/j.ymssp.2017.06.027

[6] Tsatsral Amarbayasgalan, Bilguun Jargalsaikhan, and Keun Ho Ryu. 2018. Unsupervised novelty detection using deep autoencoders with density based clustering. Applied Sciences (Switzerland) 8, 9 (2018), 1468. https://doi.org/10.3390/app8091468