Prediction of Remaining Useful Lifetime (RUL) of Turbofan Engine using Machine Learning

Vimala Mathew, Tom Toby, Vikram Singh, B Maheswara Rao, M Goutham Kumar
Information Technology Group, National Institute of Electronics and Information Technology,
NIT Campus PO, Calicut, Kerala, India
vimala@calicut.nielit.in

***Abstract*— Maintenance of equipment is a critical activity for any business involving machines. Predictive maintenance is the method of scheduling maintenance based on the prediction about the failure time of any equipment. The prediction can be done by analyzing the data measurements from the equipment. Machine learning is a technology by which the outcomes can be predicted based on a model prepared by training it on past input data and its output behavior. The model developed can be used to predict machine failure before it actually happens.**

There are different approaches available for developing a machine learning model. In this paper, a comparative study of existing set of machine learning algorithms to predict the Remaining Useful Lifetime of aircraft's turbo fan engine is done. The machine learning models were constructed based on the datasets from turbo fan engine data from the Prognostics Data Repository of NASA. Using a training set, a model was constructed and was verified with a test data set. The results obtained were compared with the actual results to calculate the accuracy and the algorithm that results in maximum accuracy is identified.

We have selected ten machine learning algorithms for comparing the prediction accuracy. The different algorithms were compared to obtain the prediction model having the closest prediction of remaining useful lifecycle in terms of number of life cycles.

Keywords*— *Predictive Maintenance, Machine Learning, Big Data Analytics, Random Forest.

I. INTRODUCTION

Prediction of equipment maintenance requirements helps the businesses to plan maintenance of equipment before the failure occurs. Predictive maintenance can save lot of time and energy as it avoids the need for doing unnecessary maintenance activities, as in the case of periodical maintenance [2], [6].

Nowadays sensors are fitted in equipment in factories and these sensors are capturing huge amount of data and it is being stored into large storage devices including cloud servers. The behavior of the sensor data contain useful information to decide what may lead to an equipment failure. So maintenance decision is now converted into a data analytics problem termed as predictive analytics [3], [5].

For an airplane to move through the air, some kind of propulsion system should generate a thrust. Most modern airliners use turbofan engines for this. Hence predicting the remaining useful lifetime of a turbofan engine is critical for scheduling aircraft maintenance. An attempt is made in this paper to analyze the historical data and evaluate different machine learning approaches best suited for accurately predicting the metric. The dataset used consists of run-to-failure sensor measurements from degrading turbofan engines.

Predictive maintenance has received much attention recently. There are two ways of doing predictive maintenance. The first is the classification approach in which it predicts the possibility of failure in next n-steps. Or it can be regression approach where it predicts the time left before the next failure, called Remaining Useful Life (RUL) [7].

This paper deals with the regression approach, mainly the machine learning algorithms used for prediction and a performance comparison of these algorithms are carried out.

II. METHODOLOGY

This section presents the methodology that has been followed. Section II-A describes the dataset used. Section II-B describes the method used for prediction as a flow chart. Section II-C includes the experimental set up. Section II-D lists the algorithms implemented for predicting the results.

A. Dataset

The dataset is taken from NASA data repository. The dataset selected includes the run-to-failure sensor measurements from degrading turbofan engines. Although the turbo fan engines are of same type, each engine starts with different degree of initial conditions and there are variations in the manufacturing process of the engines, which are not known to the user.

For the turbo fan engines under consideration, the performance of each engine can be changed by adjusting three operational settings. Each engine has 21 sensors collecting different measurements related to the engine state at runtime. The main characteristic of the dataset is that it is a time series data, the schema of which is included in Table I.

TABLE I. DATASET SCHEMA

Index	*Data Fields*	*Types*	*Descriptions*
1	Id	Integer	Aircraft Engine Identifier
2	Cycle	Integer	Time, in cycles
3	Setting1	Double	Operational Setting 1
4	Setting2	Double	Operational Setting 2
5	Setting3	Double	Operational Setting 3
6	S1	Double	Sensor Measurement 1
7	S2	Double	Sensor Measurement 2
8	…	..	..
9	S21	Double	Sensor Measurement 21

At the beginning of the time series, the engine's operation is normal. After many cycles, a fault is developed in the engine and gradually the engine fails.

Three data sets were provided as text files for training, testing and measurement of accuracy as part of our approach. The datasets available were

- Training data: It is the aircraft engine's run-to-failure data.
- Testing data: It is the aircraft engine's operating data without failure events recorded.
- Ground truth data: It contains the information of true remaining cycles for each engine in the testing data.

In the training set, the amount of fault grows in magnitude until the system fails. In the test set, the time series ends some time before the failure of the system.

Four such sets of data were used for conducting the study.

B. *Prediction using Machine Learning Algorithms*

Machine Learning (ML) enables computers to do self-learning from the data available, without any explicit programs. These machines can react to new data, based on its past learning. Prediction is one of the important applications of Machine Learning. There are two categories of ML. They are categorized as Supervised and Unsupervised [13]

In supervised learning, the dataset includes their output classes which are used to train the machine whereas in unsupervised learning output class information is not available; instead the data is partitioned into unknown classes. The methodology adopted in this paper is to do prediction based on supervised machine learning. The turbo fan engine dataset includes a training set and a test set. Both training and test data includes the actual RUL value which we plan to predict. The machine is trained using the training set and is tested for accuracy using the test data as explained in the flowchart in Fig. 1.

C. *Experimental Setup*

R programming environment was used to construct the models and test the results. The data sets were read from files and the algorithms were applied on it using R language. The advantage of R environment is that it is easy to port data to R, process and visualize the results. The results were stored in variables. These variables were plotted using the visualization methods available in R.

D. *Machine Learning Algorithms Evaluated*

The following machine learning algorithms were used for predicting the RUL.

1) Linear Regression: It is the most basic type of algorithm used for predictive analysis. Regression estimates

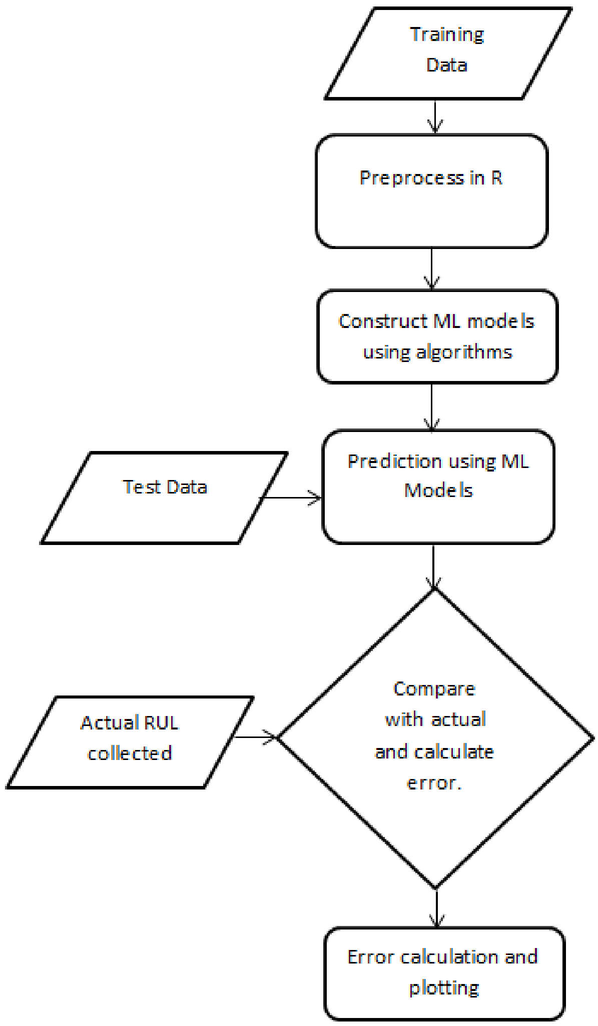


Fig. 1. Flowchart for the Machine Learning model

are used to explain the relationship between a dependent variable and one or more independent variables. In the problem, the RUL is considered to be dependent variable[12].

2) Decision Tree: In decision tree method, a tree is constructed as the predictive model. The branches of this tree illustrate the outcome of the decisions taken. The observations about an item can be converted to conclusions with the help of this decision tree.

3) Support Vector Machine (SVM): In SVM, the numeric input variables in the data which are in the different columns form an n-dimensional space. A hyper plane is a line that splits the input variable space. In SVM, a hyper plane is selected to best separate the points in the input variable space by their class [1]. The SVM algorithm is implemented in practice using a kernel.

4) Random Forest: Random forest is an ensemble learning method, It operates by constructing a set of decision trees at training time and outputting the mean prediction of the individual trees [8], [9].

5) K-Nearest Neighbors: The K nearest neighbor classifier classifies objects based on the nearest values with respect to features. It stores all available cases and classifies a new case based on the similarity to the stored classes [1].

6) The K Means algorithm: K Means algorithm is used to find groups in the data. Assuming there are K groups, the algorithm assigns each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity.

7) Gradient Boosting Method(GBM): Gradient Boosting Method is a forward learning ensemble method. It is based on the idea that good predictive results can be obtained through increasingly refined approximations. GBM sequentially builds regression trees on all the features of the dataset in a fully distributed way - each tree is built in parallel.

8) AdaBoost: AdaBoost, short for "Adaptive Boosting". Boosting, is a machine learning technique, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees, with each new model attempting to correct for the deficiencies in the previous model [4].

9) Deep Learning: Deep learning is based on artificial neural networks that are composed of many layers. In deep learning, features are learned at different levels of the hierarchy. Learning features like this, at multiple levels of abstraction, allow a system to learn complex functions mapping the input to the output directly from data.

10) Anova: Analysis of variance, is a statistical method in which the variation in a set of observations is divided into distinct components. It can be used to predict the value of a variable on the basis of one or more categorical predictor variables.

III. Results

Training and evaluation of the Machine Learning Model was carried out using four data sets, each containing data collected from 100-250 engines and each engine having 21 different sensor values. The schema of the data set is included in Table I. All ten algorithms were tested using all four datasets. The results include the graphs showing the relation between predicted RUL and actual RUL for each dataset as described below.

The first set of graphs includes the Machine ID against RUL plots for the 10 machine learning algorithms under consideration, covering both predicted and actual RUL values. The variation between predicted RULs and actual RULs of dataset1 are depicted in these graphs (Fig. 2). If actual and predicted points coincide, it indicates maximum accuracy.

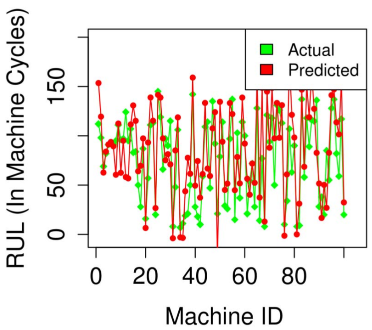
In the second set, the variation of actual RULs from the predicted of dataset1 are plotted in the form of a regression line. The plots of performance of 10 such algorithms are included in it (Fig. 3). For the most accurate algorithm, maximum points will be on the regression line.

In the third set, the Root Mean Square Error (RMSE) values between actual RUL and predicted RUL were plotted (Fig. 4). The RMSE of individual algorithms were compared for consistency between different data sets, by preparing a Root Mean Square Error Table and the same was also visualized as bar plots. The Table II describe the Root Mean Square Error Table and Fig. 4 represents the 4 barplots of RMSE values of each algorithm for the four datasets.

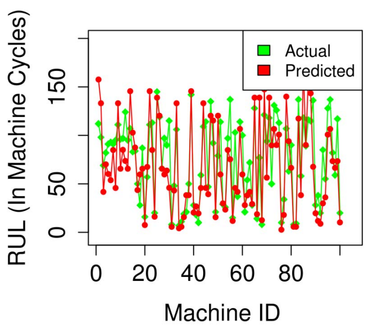
The objective was to obtain the algorithm having lowest RMSE Value. Even though cent percent accuracy cannot be achieved, error in prediction can be minimized by selecting the best algorithm. On observation it was found that random forest algorithm generates the least error.

TABLE II. ROOT MEAN SQUARE ERROR VALUES

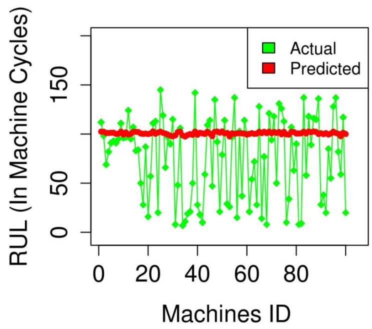
Algorithm	*Data Set 1*	*Data Set 2*	*Data Set 3*	*Data Set 4*	*Mean*
Linear Regression	29.91	31.49	45.64	39.81	36.71
Decision Tree	28.48	34.52	27.74	45.91	34.17
SVM	48.17	31.12	61.53	34.65	43.86
Random Forest	24.95	29.64	30.55	33.79	29.73
KNN	30.79	34.79	34.44	44.70	36.18
K Means	78.30	90.19	72.92	95.45	84.21
Gradient Boost	27.45	33.35	31.78	39.30	32.97
Ada boost	28.82	33.84	30.91	39.16	33.18
Deep Learning	29.62	42.41	46.82	38.11	39.24
Anova	33.50	41.14	35.46	51.44	40.38



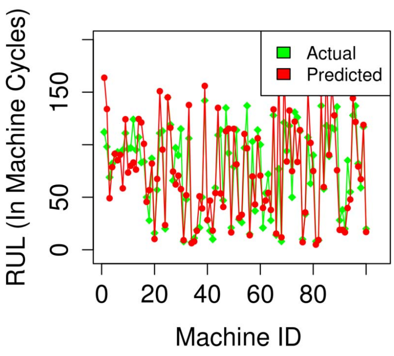
a. Linear Regression



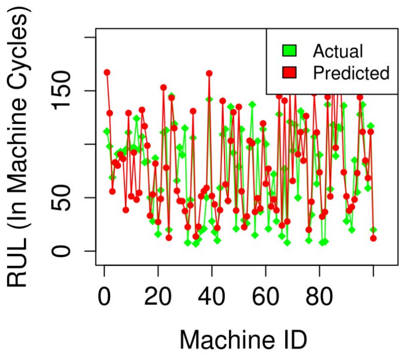
b. Decision Tree



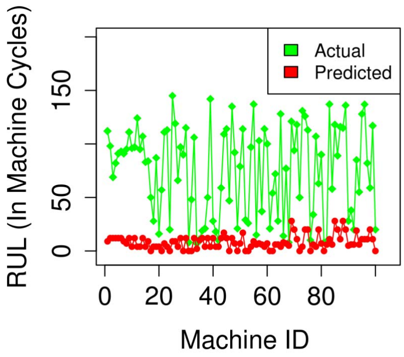
c. SVM



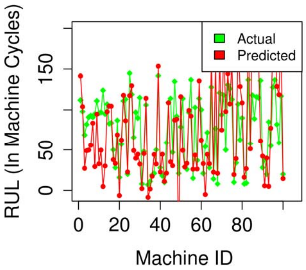
d. Random Forest



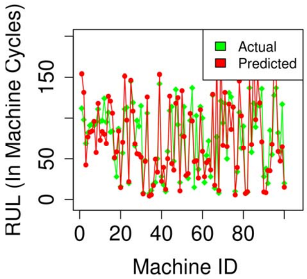
e. KNN



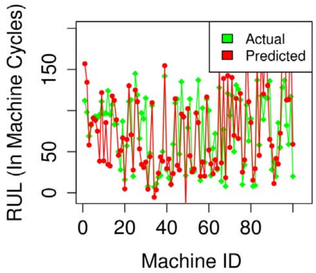
f. K Means



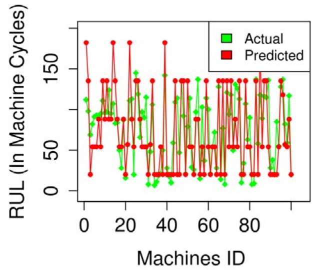
g. Gradient Boost



h. Ada Boost

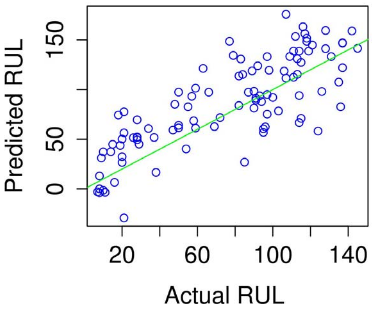


i. Deep Learning

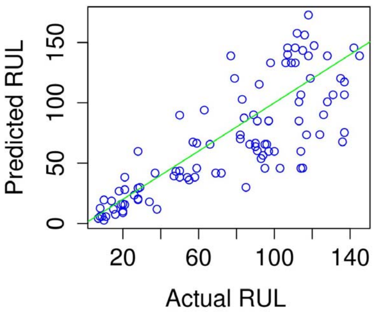


j. Annova

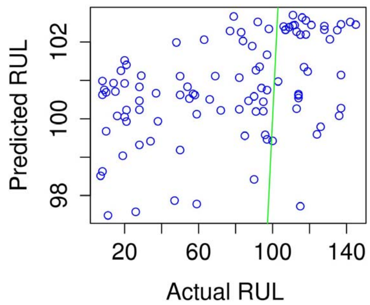
Fig. 2. Visualisation of Machine ID Vs. Remaining Useful Lifetime covering both predicted and actual values of dataset1



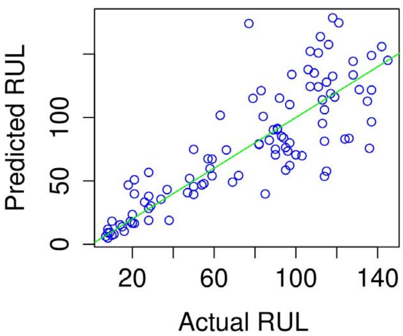
a. Linear Regression



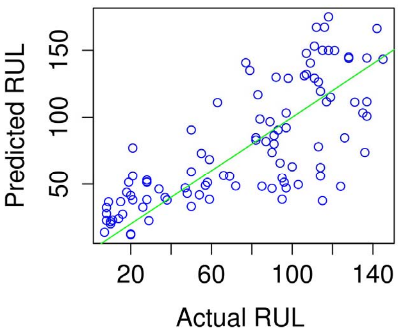
b. Decision Tree



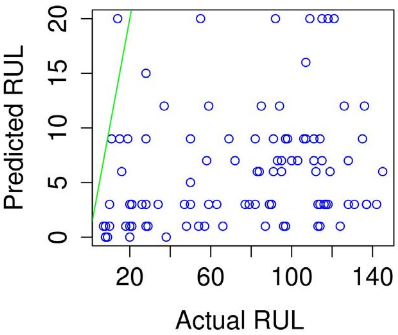
c. SVM



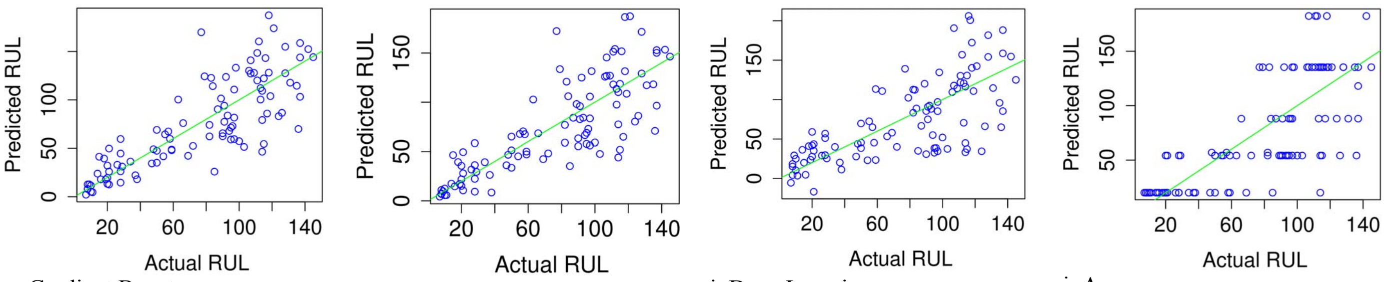
d. Random Forest



e. KNN

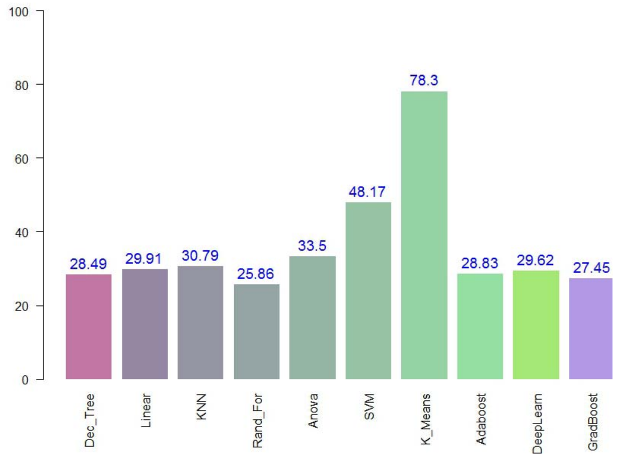


f. K Means

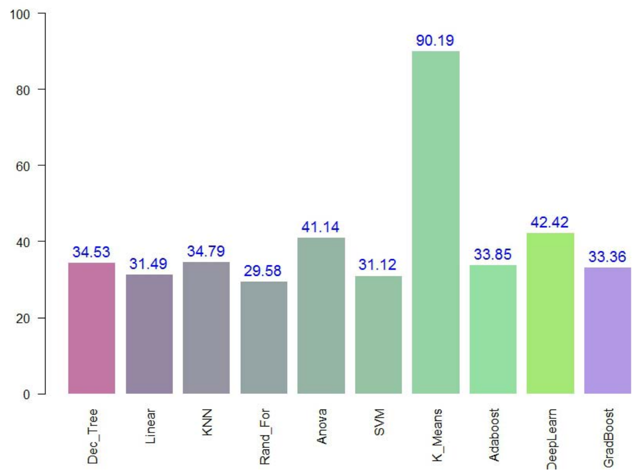


g. Gradient Boost h. AdaBoost i. Deep Learning j. Annova

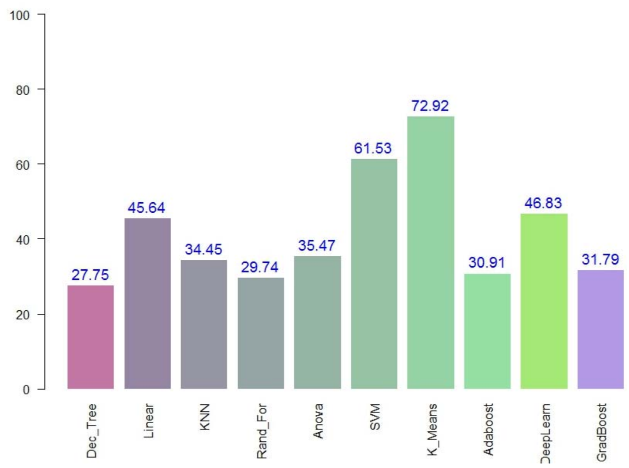
Fig. 3. Visualisation of the predicted Remaining Useful Lifetime Vs. actual values of dataset1



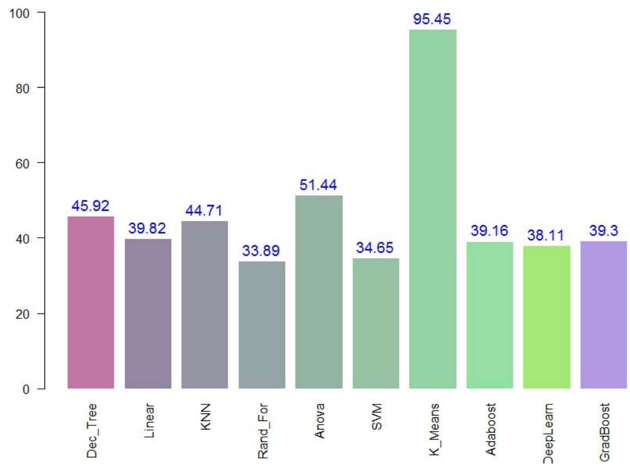
a. RMSE Plot for data set1



b. RMSE Plot for data set2



c. RMSE Plot for data set3



d. RMSE Plot for data set4

Fig. 4. Visualisation of Root Mean Square Error of RULs against Machine Ids for different algorithms on four different data sets

IV. DISCUSSION

The different algorithms were evaluated in the same way on the same data for consistent comparison of results. While predicting RUL, the main objective is to reduce the error between the actual RUL and the predicted RUL. For each dataset, the test results were compared with the actual values of the RUL available in the data set. The root mean squares of the errors were plotted and it is observed that the best results were obtained by random forest algorithm. Random forest algorithm captures the variance of several input variables at the same time and enables high number of observations to take part in the prediction.

It was observed that the performance of all ten algorithms were consistent in the four different datasets, generating proportional accuracy for the different algorithms tested.

V. Conclusion and Future Scope

The main objective of predictive maintenance is to predict the equipment failure. The Remaining Useful Lifetime prediction has been carried out so as to plan the maintenance requirements of the turbo fan engine. By doing predictive maintenance, failures can be predicted and maintenance can be scheduled in advance. This reduces the cost and effort for doing maintenance. It increases safety of employees and reduces lost production time. In our work, we have studied the performance of ten machine learning algorithms. In future, the algorithms can be tested for more real time data and always be one step ahead in predicting the maintenance requirements.

References

[1] W. F. Godoy, I. d. S. Nunes , A. Goedtel, R. H. Cunha Palacios and T. D. Lopes, "Application of intelligent tools to detect and classify broken rotor bars in three-phase induction motors fed by an inverter," IET Electric Power Applications, vol. 10, no. 5, pp. 430-439, 2016.

[2] M. Canizo, E. Onievay, A. Condez, S. Charramendietax and S. Trujillo, "Real-time predictive maintenance for wind turbines using big data frameworks," in IEEE International Conference on Prognostics and Health Management (ICPHM), 2017.

[3] C. Yang, Q. Chen, Y. Yang and N. Jiang, "Developing predictive models for time to failure estimation," in Proceedings of the 2016 IEEE 20th International Conference on Computer Supported Cooperation Work in Design, 2016.

[4] G. Wang, T. Xu, H. Wang and Y. Zou, "AdaBoost and least square based failure prediction of railway turnouts," in 9th International Symposium on Computational Intelligence and Design, 2016.

[5] J. J. A. Costello, G. M. West and S. D. J. McArthur, "Machine learning model for event-based prognostics in gas circulator condition monitoring," IEEE Transactions on Reliability, vol. DOI 10.1109/TR.2017.2727489, 2017.

[6] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, A. Beghi and G. A. Susto, "Machine learning for predictive maintenance: A multiple classifier approach," IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812-820, 2015.

[7] Z. Li and Q. He, "Prediction of railcar remaining useful life by multiple data source fusion," IEEE Transactions on Intelligent Transportaion Systems, vol. 16, no. 4, pp. 2226-2235, 2015.

[8] C. Jennings, D. Wu and J. Terpenny, "Forecasting obsolescence risk and product life cycle with machine learning," IEEE TRansactions on Components Packaging and Manufacturing Technology, vol. 6, no. 9, pp. 1428-1439, 2016.

[9] W. Lin, Z. Wu, L. Lin, A. Wen and J. Li, "An ensemble random forest algorithm for insurance big data analysis," IEEE Access, vol. 5, pp. 16568-16575, 2017.

[10] T. Pitakrat, A. v. H. Hoorn and L. Grunske, "A comparison of machine learning algorithms for proactive hard disk drive failure detection," Proceedings of the 4th international ACM Sigsoft symposium on Architecting critical systems, pp. 1-10, 2013.

[11] K. Koniuszewski and P. D. Domanskiy, "Reproduction of equipment wear characteristics with kernel regression," in 21st International Conference on Methods and Models in Automation and Robotics (MMAR), Miedzyzdroje, Poland, 2016.

[12] W. Ahmad, S. A. Khan and J.-M. Kim, "A hybrid prognostics technique for rolling element bearings using adaptive predictive models," IEEE Transactions on Industrial Electronics, Vols. PP, DOI: 10.1109/TIE.2017.2733487, no. 99, pp. 1-1, 2017.

[13] J. Samantaray, M. Armaco and M. l Armaco, "Big data capabilities applied to semiconductor manufacturing advanced process control," IEEE Transactions on Semiconductor manufacturing, vol. 29, no. 4, pp. 289-291, 2016.

[14] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation", Proceedings of the Ist International Conference on Prognostics and Health Management (PHM08), Denver, pp. 1-9, 2008