NASA Jet Engine Predictive Maintenance

Jay Trivedi

Abstract **- This project paper describes how machine learning algorithms can predict the future maintenance date based on analyzing engine performance data collected from sensors attached to every engine. We have pre-possessed the data and performed exploratory data analysis to find patterns and develop an efficient failure detection model by leveraging decision trees, random forest, SVM, and logistic regression. We have used F1, Recall, Precision, and ROC for performance evaluation to compare and evaluate models.**

Predictive Maintenance | Remaining Useful Life | Damage Modeling | Performance Evaluation

Introduction

Maintenance of equipment is critical for any business involving machines. Predictive maintenance is the method of scheduling maintenance based on the prediction of the failure time of any equipment. The machine learning models are constructed based on The datasets from turbofan engine data from the Prognostics Data Repository of NASA. Using a training set, a model was constructed and verified with a test data set. The results were compared with the actual results to calculate the accuracy, and the algorithm that maximized accuracy was identified. We have selected two machine learning algorithms for comparing the prediction accuracy.

Data Source

This dataset is taken from the Prognostic and Health Management (PHM) Competition 2008 conducted by NASA, a well-known publicly available dataset for equipment degradation modeling from NASA. This dataset includes 100 Turbofan engine's Run to Failure instances. The dataset contains 23 sensor data, cycles, three types of operational settings, and the serial numbers of engines.

Approach

The approach consists of a few data pre-processing and manipulation steps where certain irrelevant columns were removed based on the filter method's correlation metric and features having a low correlation with the target feature. Feature Engineering is performed by calculating the Remaining Useful Life (RUL) for each engine on each passing cycle. Based on our observations and exploratory data analysis(EDA), we concluded to approach this problem as a classification problem; otherwise, this problem can be approached as a regression task. Observations were plotted for several sensors, and we found a sudden drop in sensor values indicative of deterioration in the Turbojet engine. We devised a rule such that if the remaining useful life is greater than 30 cycles, label it as a normal or else alarming state. While performing classification, we again eliminated certain irrelevant features for the model building based on trends concerning RUL. Later, we scaled the selected features and made the data ready for model building. First, we created test-train data sets to build K-nearest neighbor, Logistic regression, Support vector machine, and decision tree classifiers. We performed Hyper-Parameter Tuning via GridSearch and used 5 Fold Cross Validation techniques to fine-tune each model's performance further. After evaluating the performance scores of each F1, Precision, Recall, and Accuracy classifier, we plotted them on the AUC-ROC curve. Finally, we evaluated the models on F1, Recall, Precision, and AUC-ROC metrics to determine the best-performing model.

Evaluation

Models are evaluated based on Accuracy, AUC-ROC, Precision, Recall, and F1 Score. Recall and AUC-ROC are prioritized over other metrics to minimize False Negatives. The results obtained are available in Figure 1.

	Accuracy	AUC ROC	Precision	Recall	F1 Score
Logistic Regression	0.94	0.88	0.84	0.78	0.81
KNN	0.94	0.84	0.86	0.71	0.78
SVC	0.94	0.87	0.86	0.76	0.81
Decision Tree	0.93	0.85	0.78	0.74	0.76

Fig. 1. Results

Best AUC-ROC and Recall are received for Logistic Regression.

Discussion

In the analysis, we assumed that the shift in sensor readings is related to an engine's remaining useful life, which we later verified by visualizing trends. We also assumed the operational settings do not affect RUL, and we confirmed this using the correlation heat map.

Conclusion

From a classification perspective, we achieved a decent result on the problem. Such modeling is limited since we initially assumed the remaining useful life depends on sensor readings which we later verified using correlation and trends. From a future work perspective, we can also approach this problem as a Regression task over RUL as the target feature. The dataset consists of 23 features, out of which several were found to be irrelevant. Another approach to this problem can be made by reducing the features by leveraging dimensionality reduction techniques like PCA and SVD, followed by the classification models.