



# FourthBrain

## Industry Project Proposal

Name of the Organization	Booz Allen – Aerospace Market – DSMS
Project location (city or timezone)	DC Metro / Eastern

### A. Project Description

#### Problem definition

*[50-100 word description of the problem which the candidates need to solve]*

The U.S. DoD's Joint Artificial Intelligence Center ([JAIC](#)) has designated Predictive Maintenance as one of its two founding National Mission Initiatives (NMIs). By leveraging rich data sources and machine learning techniques, predictive maintenance has huge potential to drastically reduce costs, increase mission readiness and even save lives of service members.

In this project, you will develop a state-of-the-art machine learning model against an open dataset to predict the remaining useful life of turbofan engines. Following MLOps best practices, experiments will be tracked in a central model repository and the final model will be containerized, deployable, and able to serve inferences via a REST API. The weightage to ML model building vs. deployment can vary based on student team progress.

#### Key Research Questions/ Technological constraints that the Project will Answer

1. What are the current state-of-the-art methods for predicting engine life that are published in the literature? (Benchmarking State of the Art)
2. Are you able to devise a better method by combining techniques from multiple papers or implementing your own technique? (ML Model building)
3. Once a model is selected, can this model be readily exposed to end users and services across an enterprise data platform? Can the model be easily updated if improvements are made? (ML Model deployment)



### Final deliverables at the end of the project

*[Please list the desired technical deliverables from the project team in as much detail as possible]*

The following should be delivered in one or more GitHub repositories:

1. Scripts to ingest and clean data, train a model, evaluate the predictions, and facilitate hyperparameter tuning
2. A script or instructions to deploy the best performing model as an inference service that returns predictions for a selected set of test data.

Additionally, contents of the model registry should be exported, compressed and delivered as a single file.

### Key activities/ technologies the project team may be expected to undertake/ work with

*[E.g. What kind of technology stack they will work with, the datasets they may need to work on, what kind of analysis they may be expected to undertake, etc.]*

Dataset:

- Primary: Turbofan Engine Degradation Dataset 1 – less data, more examples
- Secondary: [Turbofan Engine Degradation Dataset 2](#) – more data, less examples

Models and Methods: Downloadable with the datasets above.

Pre-Processing: pandas, NumPy, SciPy

Model Training: scikit-learn, Keras/Tensorflow or PyTorch/fast.ai

ML App building and deployment on AWS or GCP

### Expected learning outcomes

*[What do you expect the candidates to learn from the project. Please mention the technical skills they will imbibe over the project.]*

1. Applying variety of machine learning techniques to a real-world, domain specific, and high-impact problem
2. Developing repeatable and reproducible machine learning solutions using containerization technologies such as Docker
3. Employing best practices for tracking training and validation datasets, model architectures, parameters, and results



Desired Team Size (if any):	3
-----------------------------	---

**Desired selection process** (if any)

- Secret Clearance