HACKtheMACHINE
BOSTON 2017

Welcome to the Challenge!

We are very excited to have you join us for HACKtheMACHINE: “Data Science and the Seven Seas.”

The U.S. Navy has big datasets on machinery performance across our ships. When our data is evaluated, it may create new opportunities to understand and monitor the condition of specific machinery. Your analysis can result in lower fleet maintenance costs, and improved reliability and productivity. This weekend you will compete to design and solve a real-world predictive maintenance and anomaly detection problem in maritime machinery data.

The Challenge:

For this challenge, your primary data set will be from the US NAVY’s Integrated Condition Assessment System (ICAS).

ICAS is a remote equipment monitoring maintenance application for on-line automated machinery condition monitoring and assessment supporting Conditioned Based Maintenance (CBM). Integrated condition assessment and diagnostic monitoring systems minimizes costs, optimizes availability, and improves productivity. People monitoring the ICAS data rely heavily on the system reports telling them that something is wrong.

How can we use data science and machine learning to improve our predictive maintenance and avoid system downtime?

As an example, imagine there was a shaft break on a Navy ship while on deployment. Now place a sensor on that shaft that collects key performance metrics such as vibrations that are collected in ICAS. Using machine learning on the vibration data, anomaly detection or condition based maintenance criteria being met could be used to inform our sailors that a repair should be made before the shaft breaks or a part should be ordered for a future repair.

This challenge will test your machine learning skills to generate algorithms that will help the Navy chart a course to merge *Data Science and the Seven Seas*.

ICAS Resources:

What is ICAS? This link gives some elementary examples of how data collected from ICAS could be used to detect anomalies and save money: http://www.sae.org/events/dod/presentations/greatideas-diulio.pdf

HACKTHEMACHINE
BOSTON 2017

The Data Set

The primary data set for this challenge contains actual information on four different ships of three different classes with data from various systems of interest. Each system of interest contains multiple sensors. We've compiled **10 CSV files for your analysis**. There are approximately 3.2 million rows of data.

For this exercise, ship classes and names have been obfuscated. Each ship class is distinguished with a letter designation, currently A, B and C. Each individual ship has a numeral designator per class. For Example, our first ship will be Class A – Ship 1. Information on equipment and its corresponding sensors readings remains complete.

Please see Table 1 for types of equipment used on specific ship classes:

Table 1

Ship Class	Ship Hull #	Corresponding Systems of Interest
Class A	1	GTM, MRG
Class B	2	GTM, MRG
Class C	1, 2	MPDE, MRG, SSDG

Note on timestamps: While the ship equipment is operating nominally, according to its CDS configuration, the equipment timestamps are on a regular interval, often 10 minutes. This state is called "Trend" readings. When a change occurs and is out of the normal ranges for the configuration, an "Event" is triggered. Two minutes of data is collected prior to the event and three minutes after the event. This will also reset the regular interval timestamp to the last measurements taken.

Please note that Navy data has been designated as: DISTRIBUTION A. Approved for public release: distribution unlimited.

HACKTHEMACHINE
BOSTON 2017

Compute Resources

Teams will be provided with access to an Amazon Elastic Cloud Compute (AWS EC2) p2.xlarge instance that will provide you with a GPU and 4 vCPUs to support algorithm development as well as act as storage for the primary data set. If you wish to obtain the data by other means, please approach the onsite support at the event for alternative means.

If you are not already familiar with AWS EC2, here is more information on how to get started:
http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/concepts.html

To access the compute environment you will need to do the following:

1. Get the F5 Big IP Edge VPN Client, Navigate to https://vpn.boozallencsn.com get the client for your OS and install.
2. Open the Big IP Client and connect to the "CSN Network" it will ask for your username and password and you should be using your provided Booz Allen Account Credentials.
   - If you pre-registered for the event, you received an email with your credentials a few days ago.
   - If you did not pre-register, talk to the Hackathon staff who can get your account made onsite.
3. Now that your VPN is connected, you should be able to use your provided SSH key and provided IP address to SSH into your provided AWS EC2 instance.