

**GENERAL DYNAMICS**

Mission Systems

# COMPARATIVE EVALUATION OF ANOMALY DETECTION SOLUTIONS ON SATELLITE TELEMETRY

February 27th, 2024

Andrew Robbertz





# Satellite Telemetry Overview



1. Satellites generate thousands of telemetry variables each day (component statuses, voltages, temperatures, etc.)

## PROBLEM:

- Too much telemetry data for operators to monitor every channel
  - Some teams prioritize a subset of telemetry to review
  - Some teams use a 'review-what-you-can' model within fixed time
- Static thresholding software can help, but thresholds are unable to detect pattern changes and require adjustment over mission life
- Manual review by highly-trained Operators is expensive

## IMPACT:

- Anomalies may be overlooked on unmonitored channels or within static thresholds
- May impact mission availability, and result in increased costs for repairs
- Flight Operators reviewing data manually is much more time-consuming and expensive than reviewing the data with software

2. Ground station processes downlink data and parse out telemetry data



3. Satellite Operations teams review telemetry data for anomalies to identify risks

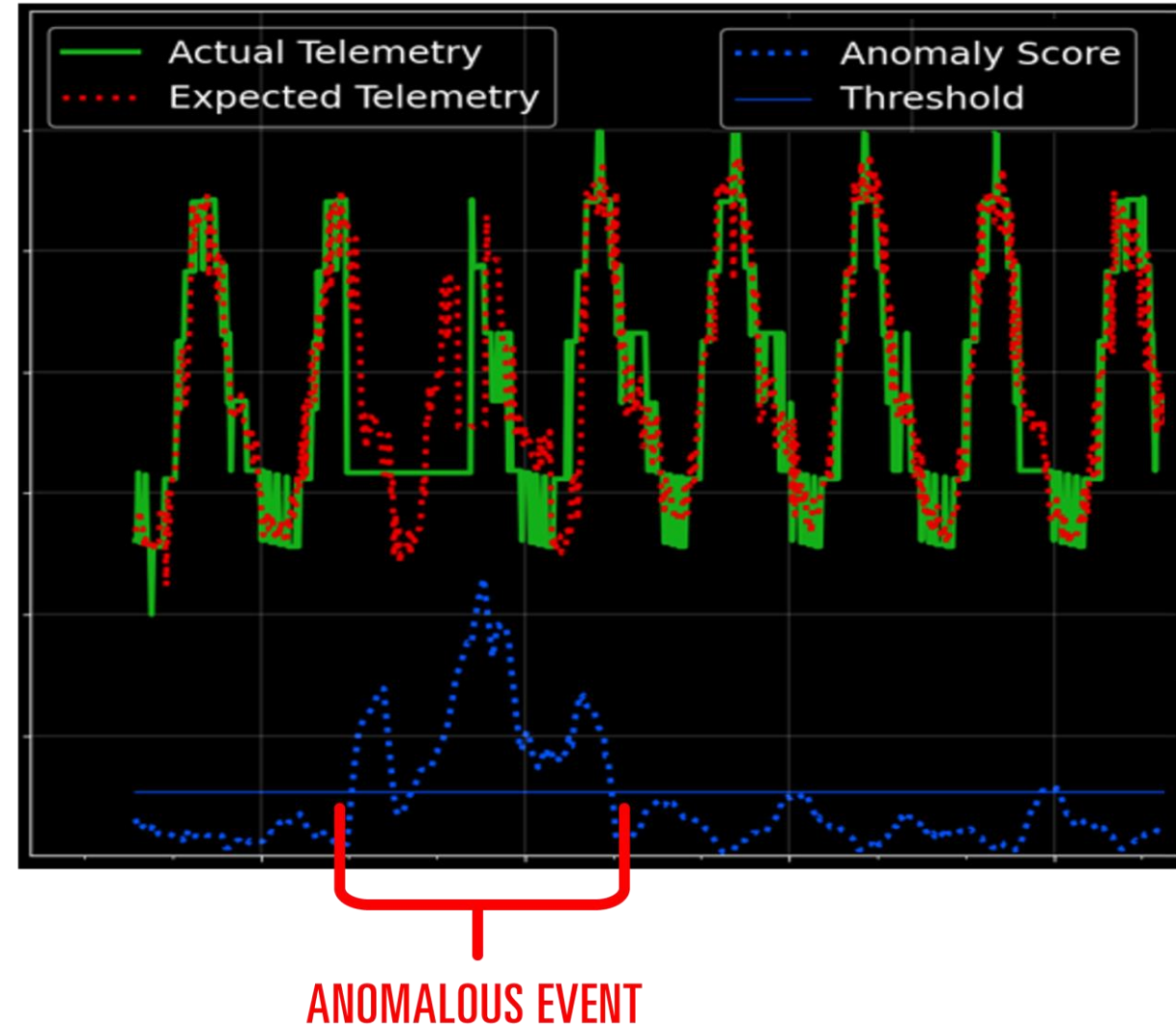
## OUR APPROACH

- Machine Learning (ML) can help solve this. But which ML tool is best?
- We provide a framework to evaluate ML performance.

# Machine Learning-Based Anomaly Detection



- Spacecraft are ideal candidates for ML-based telemetry analysis and anomaly detection (AD)
  - Highly instrumented
  - High periodicity
  - Limited interaction with external factors
- Previous research demonstrates that ML is effective at detecting anomalous events in satellite telemetry
  - Risk reduction through increased situational awareness, faster analysis of anomalies
  - Reduced downtime of satellite components
  - Cost savings through reduced labor hours to review telemetry
- Still using human-in-the-loop review
  - ML filters through large quantities of time-series telemetry
  - Imperative to have **Low False Positive Rate** to not overwhelm manual reviewers



Previous studies are relatively isolated, which raises unanswered question –  
**How do we tell which ML solutions are better than others for anomaly detection?**



# Problem Space



**PREVIOUS RESEARCH**



DATASET (1)  SOLUTION (A) =  RESULTS 1

DATASET (2)  SOLUTION (B) =  RESULTS 2

Previous research does not allow direct comparisons between ML solutions.

**OUR APPROACH**

 SOLUTION (A) =  RESULTS 1

 SOLUTION (B) =  RESULTS 2

Our approach allows direct performance comparisons.

**Machine Learning for Telemetry Analysis (MALTA) Project Objective:**

Create detailed, direct comparisons of anomaly detection solutions for future satellite flight operations missions (Not to identify single best solution for all scenarios)



# MALTA Evaluation Framework Methodology

1. Integrate standardized AD datasets for ML applications
  - Telemanom ([GitHub](#)) (Hundman et al., 2018)
  - LASP ([GitHub](#)) (Polson, 2019)
2. Integrate, optimize, and evaluate 6 unique ML anomaly detection solutions
  - Including COTS & Open-Source
  - Using multiple ML architectures and algorithms
3. Analyze qualitative and quantitative results for solution comparison

## EXPERIMENT CONFIGURATION

### DATASET CONFIGURATION

- Training Sample
- Test Sample

### ALGORITHM CONFIGURATION

- Hyperparameters
- Single-Variate vs. Multi-Variate

### EVALUATION FRAMEWORK

Dataset Adapters  
& Integrations

AD Solution  
Adapters &  
Integrations

### RESULTS ANALYSIS

- Hyperparameter Tuning
- Confusion Matrices
- Anomaly Score Distributions
- ROC Plots

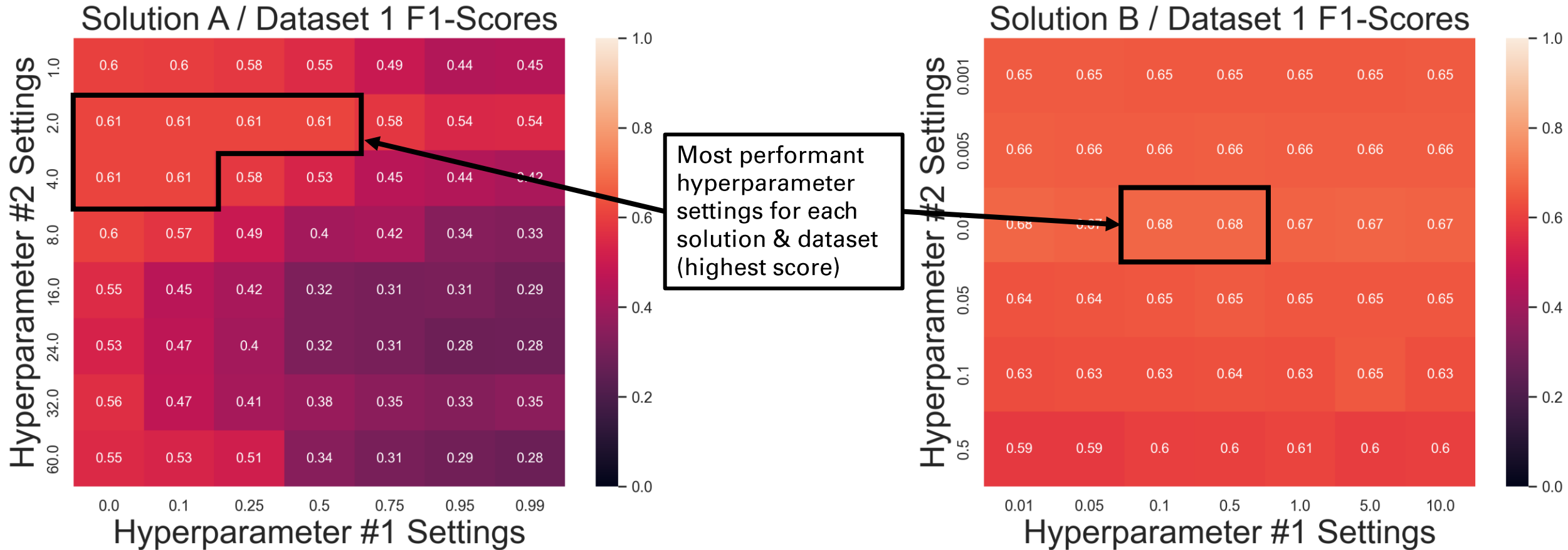
## RESULTS AGGREGATION

DATASET	SOLUTION	VARIATE	HYPERPARAMETERS	STATISTICS...
Dataset 1	Solution A	Single-Variate	Alpha=0.001	...
...	...	...	...	...
Dataset 2	Solution B	Multi-Variate	Epsilon=1	...

- Configurable to test any number of datasets and solutions during execution
- Extensible to new Solutions, and Datasets through common interfaces
- Interoperable to tune different features and parameters of different solutions

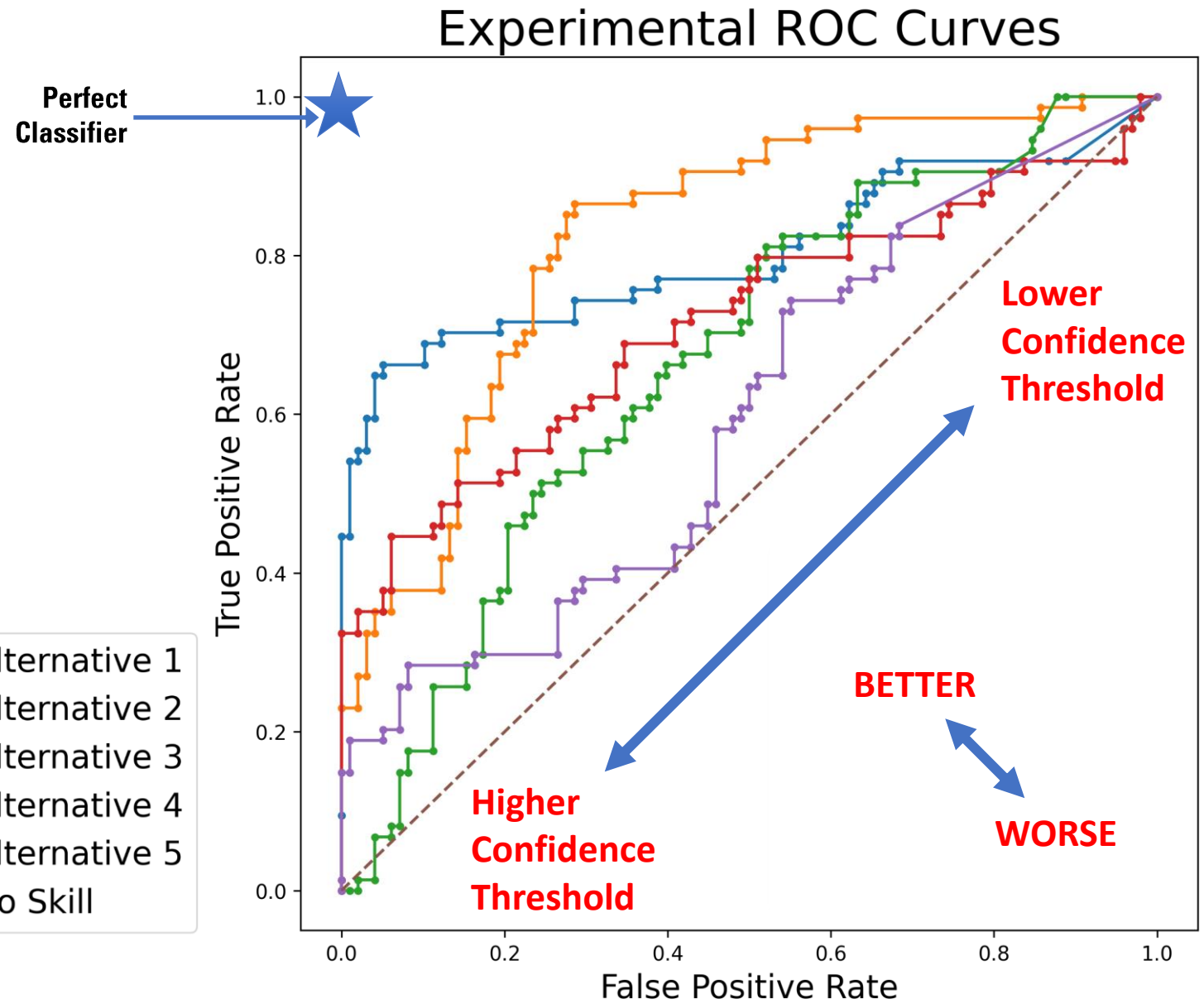
# Hyperparameter Fine-Tuning

- Enables finding the most performant "settings" for each solution
- Previous research fine-tuned solutions with only one dataset causing apples-to-oranges comparison
- MALTA built-in fine-tuning provides best apples-to-apples comparison between solutions



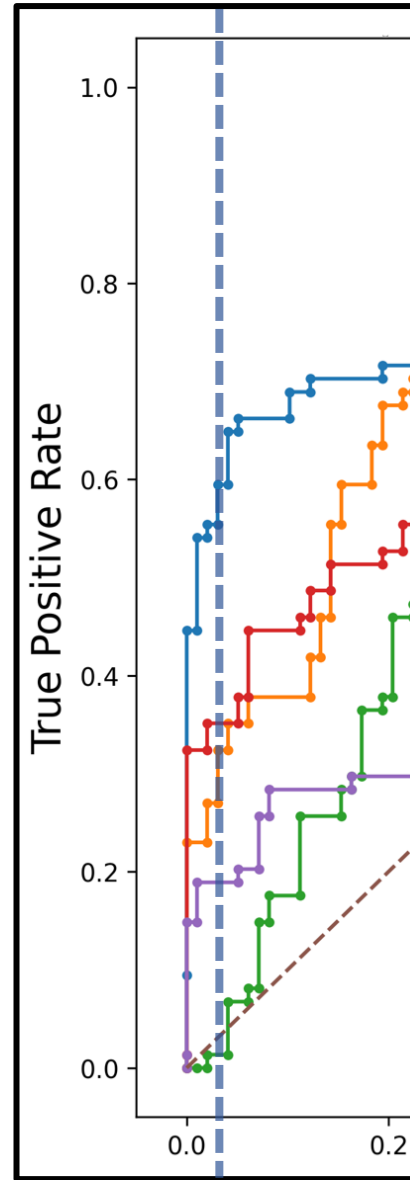
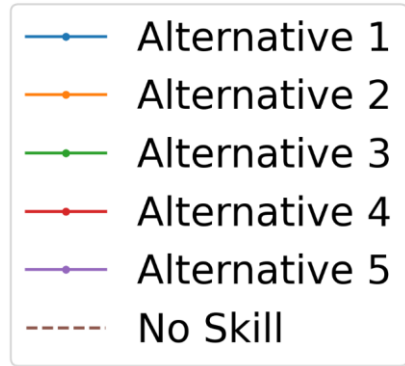
# MALTA Results

- Receiver Operating Characteristic (ROC) Plots show performance at varying anomaly score (confidence) thresholds
- Allows operations teams to identify solutions that best fit their needs
- Identify solutions with greatest performance at low false-positive-rates

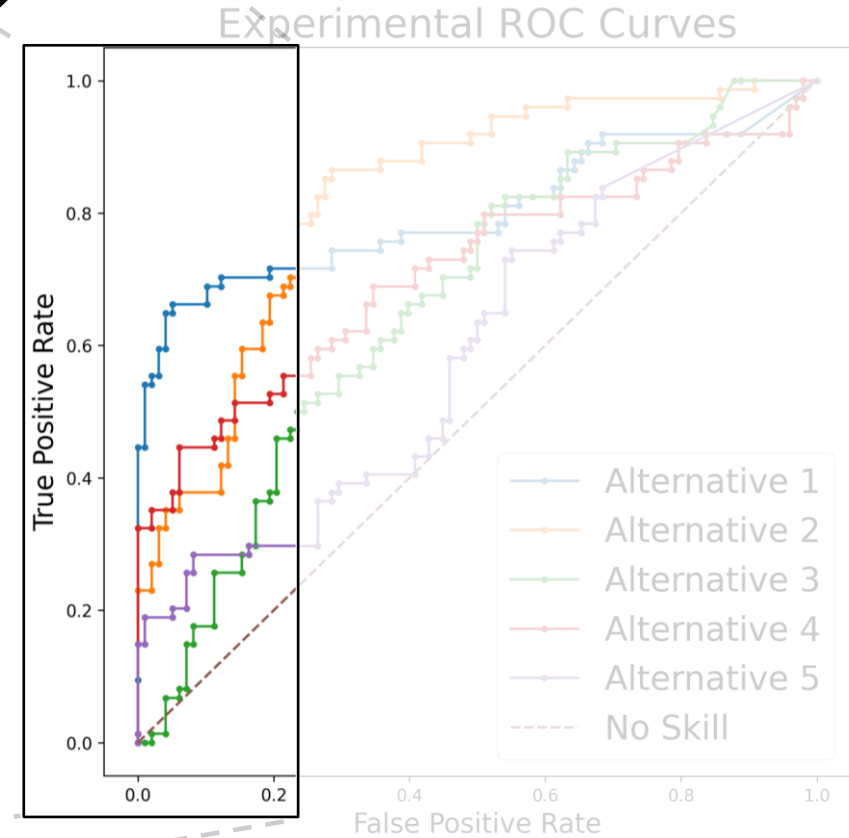


# MALTA Results

- Example Selection Criteria
  - 2000 telemetry variables to review
  - Cannot manually review >50 alerts per day
  - Requires a False-Positive-Rate (FPR) < 3%
- We can visually and programmatically examine results to find best solutions with low false-positive rates



Applying (Blue) False-Positive-Rate Threshold allows us to identify Solution A as best performing

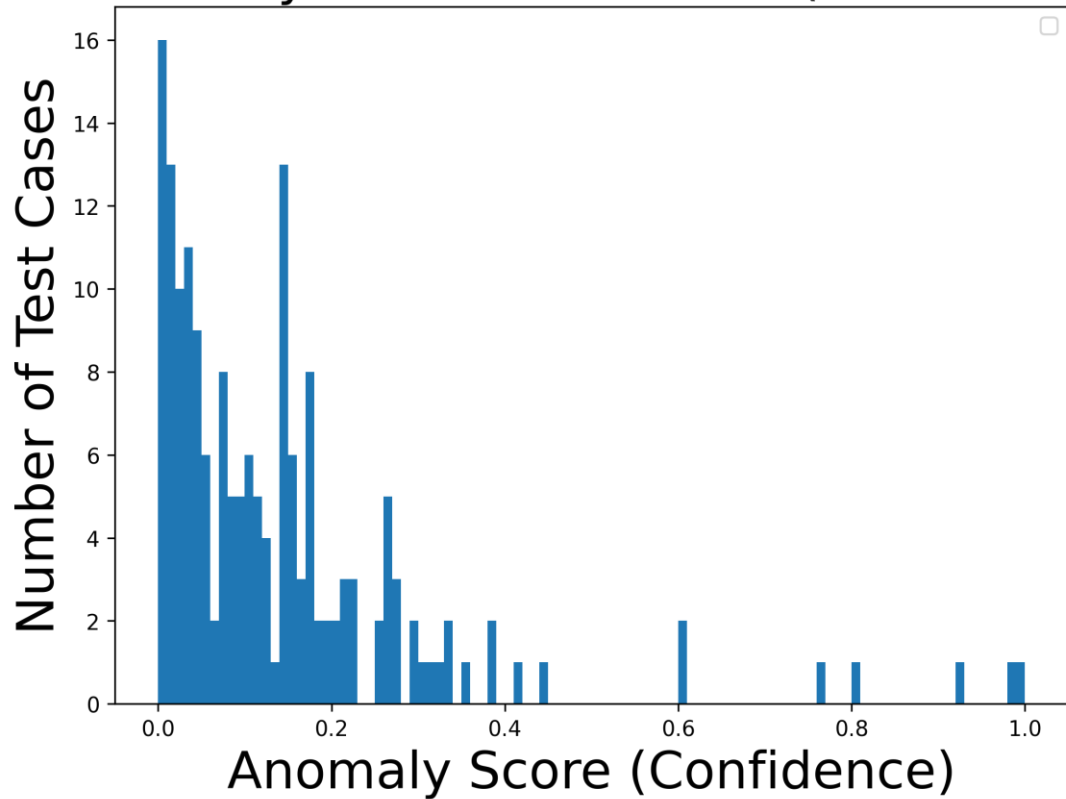




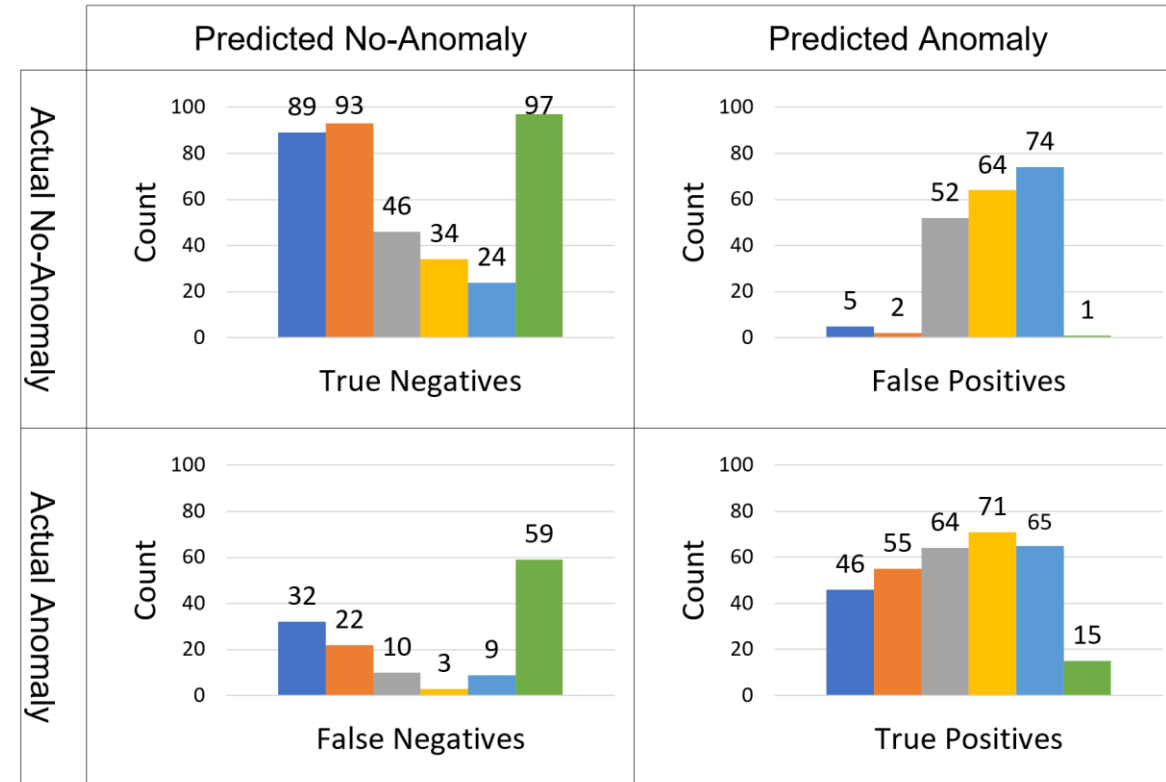
# MALTA Results

- Understand ML algorithm's "confidence" in its own classifications
- Smoother often better

Anomaly Score Distribution (Solution B)



- Combined Confusion Matrix allows us to directly compare class-by-class results at a given anomaly score threshold



# Future Opportunities

- Evaluate additional tools, additional datasets, more use cases
  - Focus on detection of long-term trend changes, recurring anomalies
- Investigate ML Ensemble Techniques to improve performance
- Extend MALTA Evaluation Framework to execute parallel evaluations of multiple solutions



# Thank you! / Questions?



**Thank You!**

