An Overview of Machine Learning Techniques for on-board anomaly detection in satellite telemetry

James Murphy

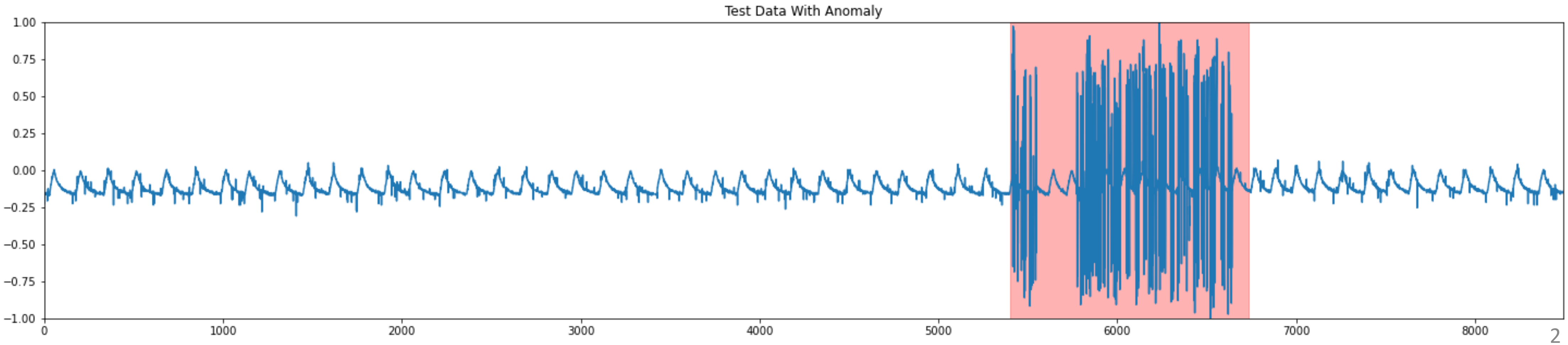
Space Software & AI Engineer, PhD Student

jmurphy@realtra.space

www.realtra.space

Telemetry Anomalies

- What is an anomaly?
  - Unexpected variance in telemetry data
- What does an anomaly look like?
  - Point/Contextual
- What causes anomalies?
  - The Space Environment



Experiment

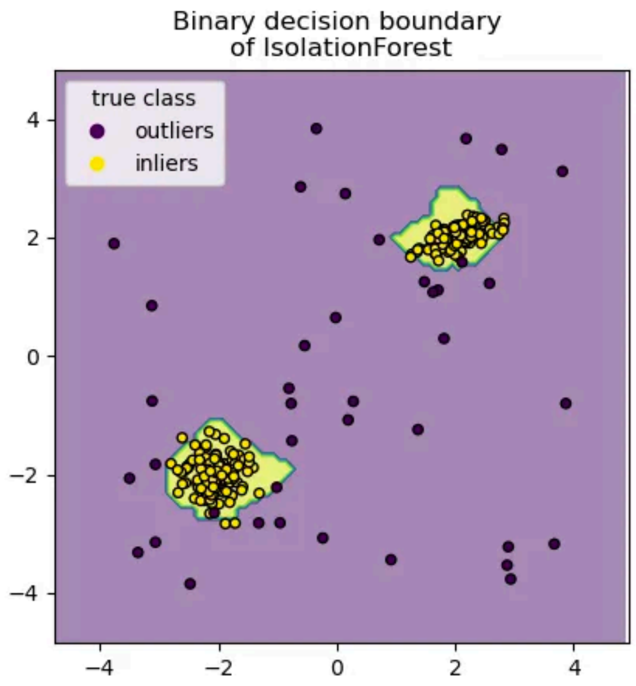
- Q1: Can Statistical, Classical and Deep Learning models reliably detect anomalies in a satellite dataset?
- Q2: Do Classical Learning and Deep Learning Methods outperform industry-standard statistical methods?
- Q3: Is there a model that outperforms the current industry standard?
- Q4: Is this model capable of being deployed on low-power devices currently deployed in space missions?

Current Anomaly Detection Methods

- Current Industry Standards
  - FDIR: Failure Detection, Isolation and Recovery
  - Thresholding on data streams
  - Statistical methods such as ARIMA
  - Ground Based Systems
- Drawbacks compared to Machine Learning alternatives
  - Generally cannot detect contextual anomalies
  - Incapable of multivariate analysis on multiple telemetry channels
  - Generally limited to ground operations and relies on satellite operators

Classical & Deep Learning Models

- Classical Learning
  - Clustering - K-Means Clustering
  - Decision Trees - Isolation Forest
- Deep Learning
  - Point Prediction - Long to Short Term Memory (LSTM) models
  - Autoencoder - Dense, Convolutional, Variational, LSTM
  - Transformer
  - Generative Adversarial Network
  - Hybrid - LSTM Encoder, Convolutional Decoder



Credit: Scikit-learn

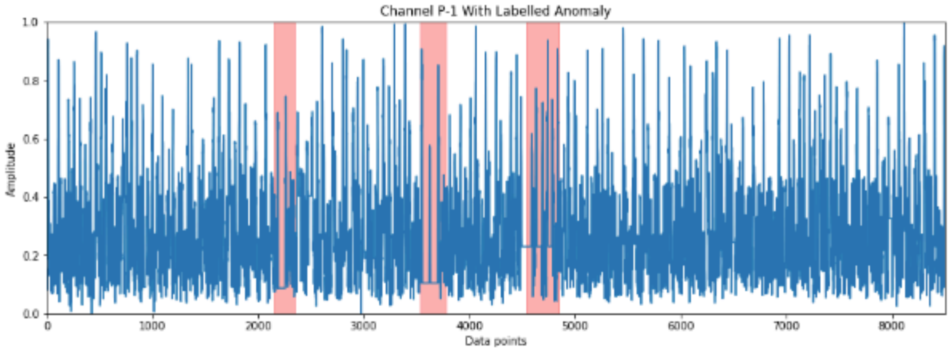
Dataset Used

- Published by Hundman et al. in 2018
- Comprised of 82 channels of satellite telemetry
- Ranges in size from 1500 to 9000 data points.
- Sourced from NASA's SMAP and MSL missions
- 1 to 3 labelled anomalies in each test dataset
- Multiple types of anomalies

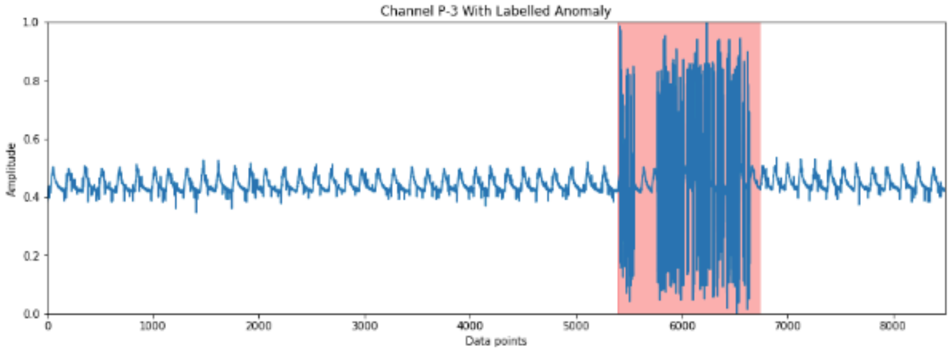
- This experiment took 5 channels from this dataset to train on with differing anomaly types and sources

Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018, July). Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 387-395).

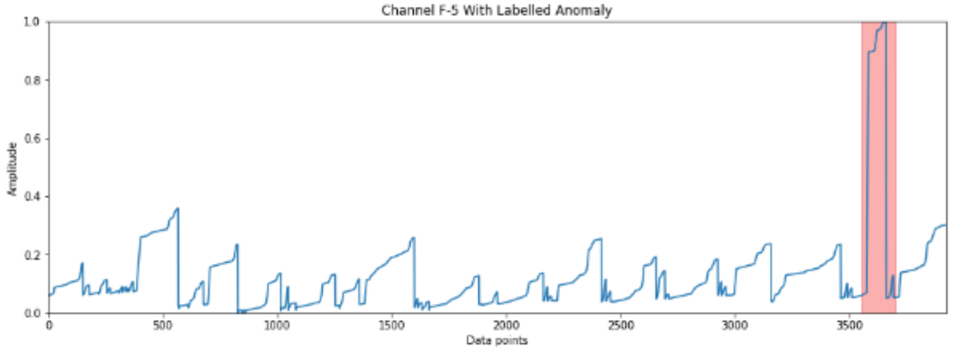
Dataset Used Contd.



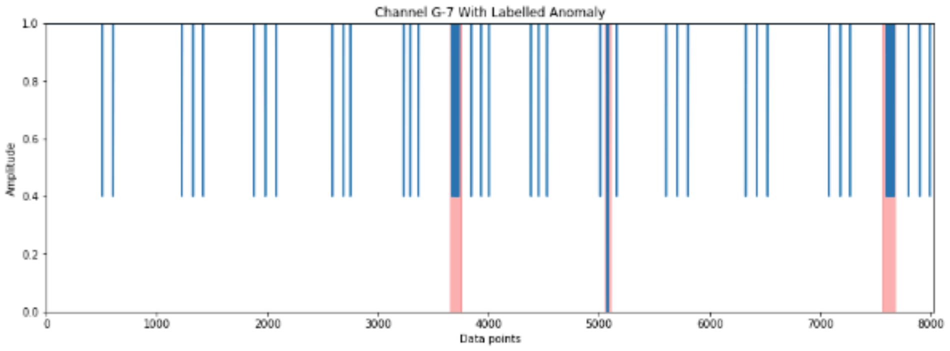
(a) P-1



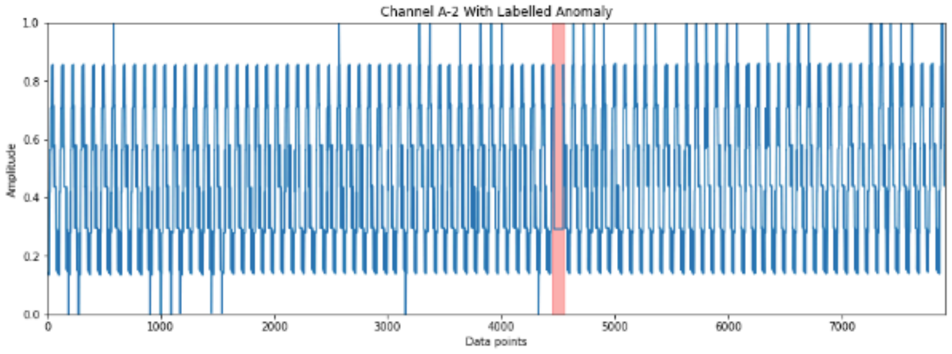
(b) P-3



(c) F-5



(d) G-7



(e) A-2

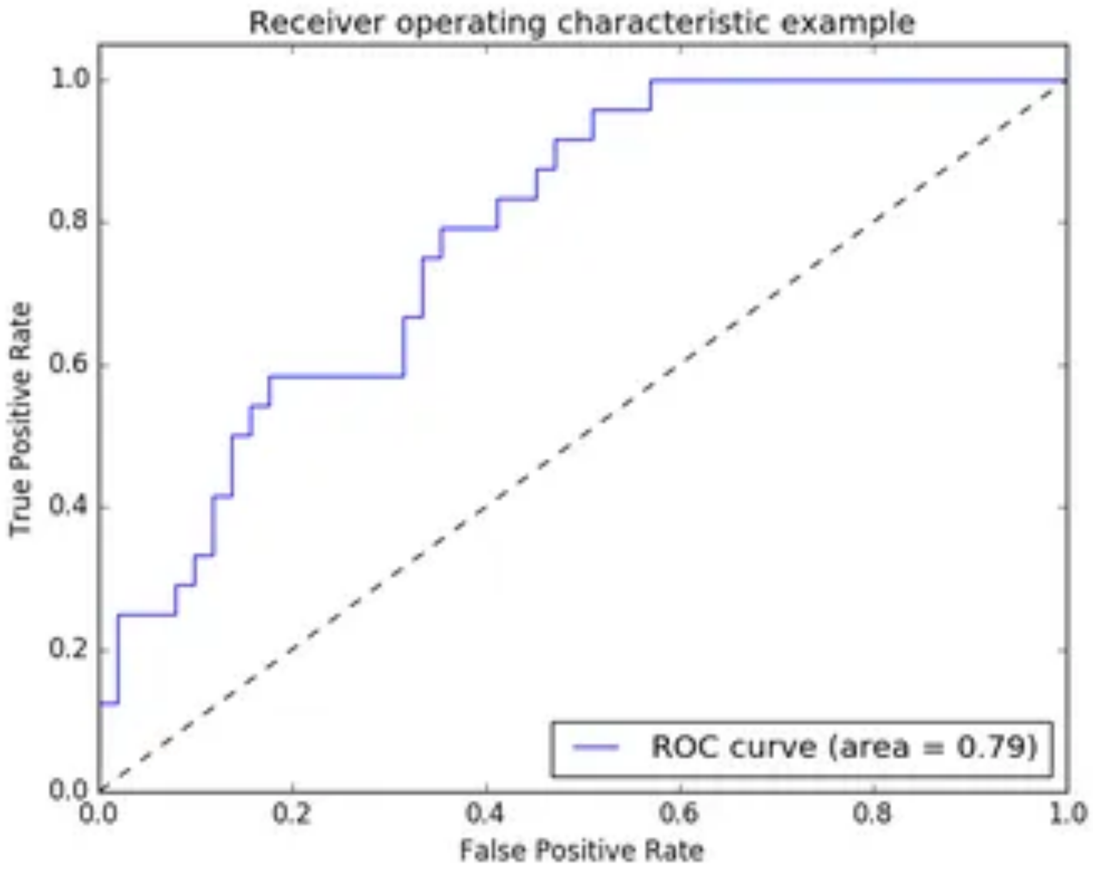
Training Step

- Preprocessing
  - Moving window approach
- Model specific preprocessing
  - Look back for LSTM models
- Model design
  - Classical Learning Parameters: K-means, Isolation Forest
  - Hidden Layers/Neurons
- Tuning
  - Grid Search
  - Kerastuner
- Testing

Metrics

ROC Curve

- AUC
  - The Area Under a Reciever Operator Characteristic Curve
- Rank
  - Ranked all models on each channel



Results

Model List	A-2		F-5		G-7		P-1		P-3		Average
	AUC ROC	*Rank*	*AUC ROC*	Rank	*AUC ROC*	Rank	*AUC ROC*	*Rank*	*AUC ROC*	*Rank*	**Rank**
ARIMA	0.492	10	0.690	10	0.677	9	0.513	3	0.723	10	8.4
KNN	0.808	4	0.860	5	0.846	3	0.510	4	0.987	4	4.0
Isolation Forest	0.717	7	0.776	8	0.808	8	0.499	6	0.940	8	7.4
Dense AE	0.761	6	0.857	6	0.840	4	0.654	2	0.981	5	4.6
Convolutional AE	0.917	2	0.863	3	0.813	6	0.318	10	0.990	3	4.8
Variational AE	0.794	5	0.770	9	0.830	5	0.432	8	0.973	6	6.6
LSTM AE	0.934	1	0.903	2	0.924	2	0.503	5	0.996	1	2.2
LSTM Predictor	0.372	11	0.682	11	0.337	10	0.421	9	0.599	11	10.4
Transformer	0.810	3	0.863	3	0.813	6	0.318	10	0.917	9	6.2
GAN	0.660	9	0.836	7	0.300	11	0.454	7	0.964	7	8.2
Hybrid	0.661	8	0.907	1	0.925	1	0.665	1	0.995	2	2.6

Conclusion

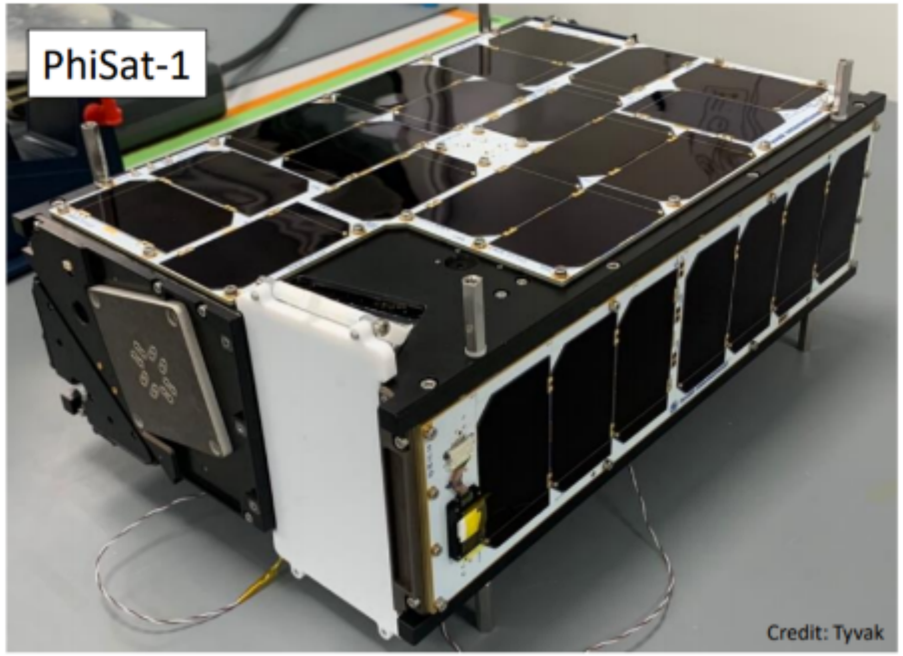
- LSTMs seem to show the most promise after multiple training rounds with parameter tuning
- The hybrid LSTM CNN model out classes the basic LSTM autoencoder, but fails entirely on A-2
- Other models perform well such as the Transformer model and CNN autoencoder
- Overall, the LSTM autoencoder and Hybrid autoencoder architectures have performed the best and will be the best basic models available for limited hardware
- Almost all deep learning and classical learning methods have outperformed the ARIMA methods

Future Work

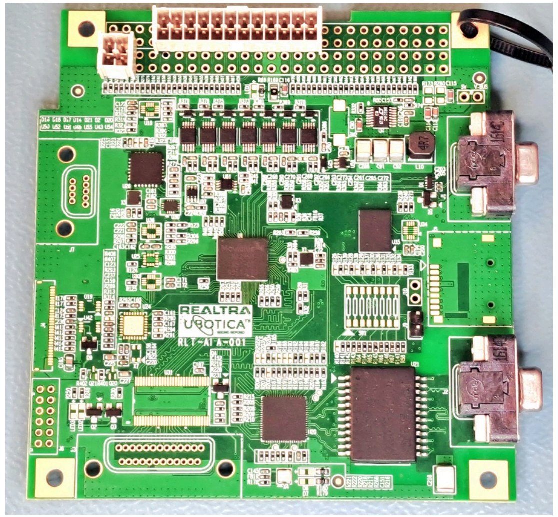
- Deployment on space hardware
  - Moving to an intel Myriad X as it has flight heritage
  - CogniSAT-XE2
- Multivariate model development
  - Generation of Datasets
  - Solnix Satellite Anomaly Dataset
  - Investigate multiple preprocessing options

Flight

- Phi-Sat1
  - Earth Observation
  - Intel Myriad 2
  - Commercially Available Chip
  - Successfully deployed CloudScout



Credit: Intel



Credit: Réaltra Space Systems Engineering

Thank You!

James Murphy

jmurphy@realtra.space

This work was supported by Réaltra Space Systems Engineering. This work was also supported by the Irish Research Council under the Employment-based postgraduate Scheme (IRC-EBPPG/2020/11).