

Wyzwania detekcji anomalii w szeregach czasowych na przykładzie misji satelitarnych Europejskiej Agencji Kosmicznej



Krzysztof Kotowski



Bogdan Ruszczak

21.11.2024, Kraków, SIGML 2024

info@kplabs.pl
www.kplabs.pl

KP Labs

Komputery, oprogramowanie i AI dla wymagających misji kosmicznych



Założone w 2016, status Centrum Badawczo-Rozwojowego od 2019



80+ pracowników



7 misji, 30+ projektów dla ESA/NASA



2300m² laboratoriów

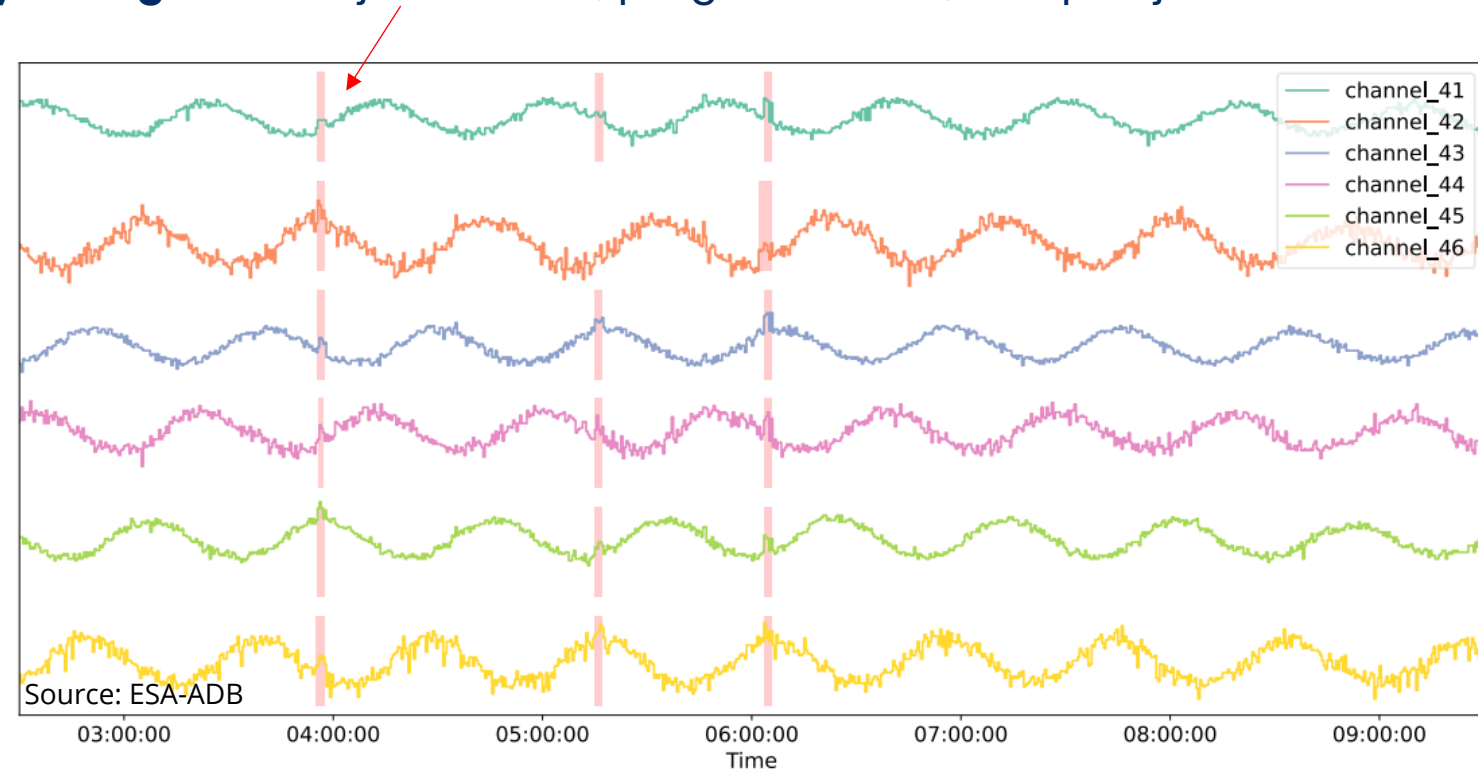
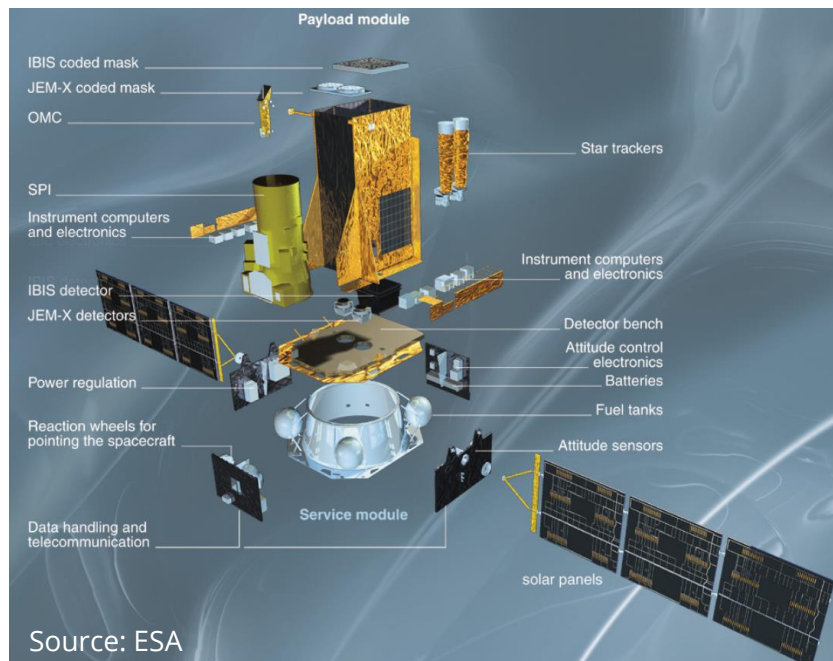


Gliwice

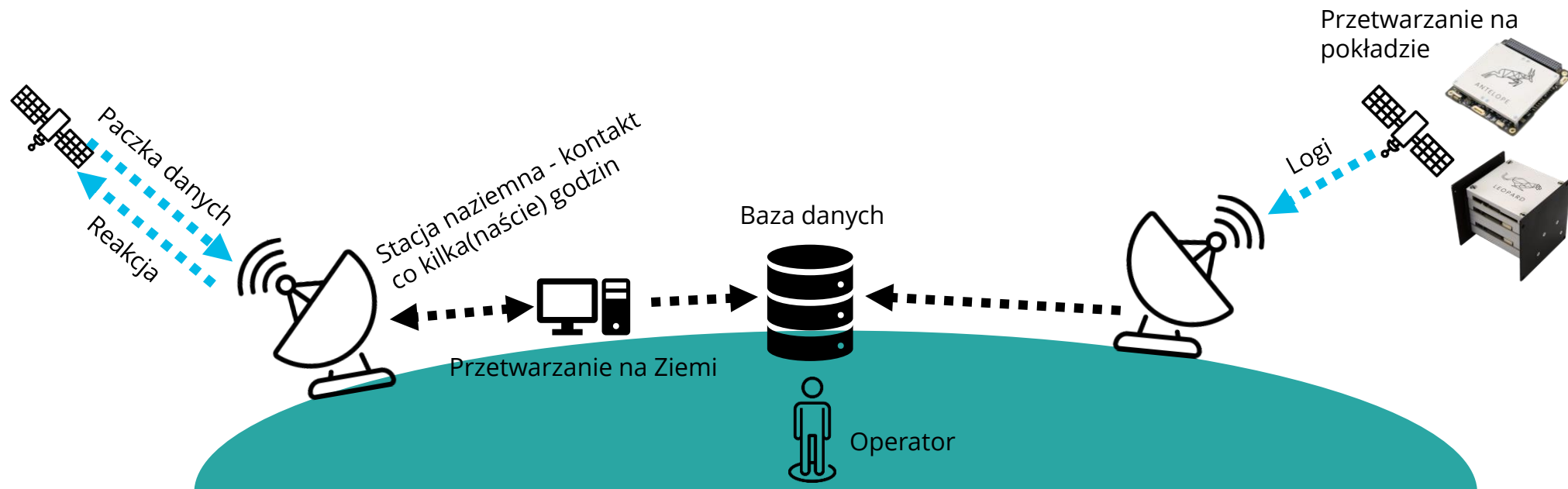


Dane telemetryczne z satelitów

- **Wielokanałowe szeregi czasowe** zawierające informacje na temat stanu satelity, np. pomiary z czujników (położenie, napięcie, temperatura, itp.), statusy, liczniki, komendy
- **Charakterystyka i wyzwania:** od kilkunastu do kilku tysięcy kanałów, złożone zależności pomiędzy kanałami, zmienne i nierównomierne próbkowanie, braki w danych, wpływ środowiska kosmicznego
- **Typowe zastosowania uczenia maszynowego:** detekcja anomalii, prognozowanie, kompresja



Przetwarzanie na Ziemi	Przetwarzanie na pokładzie
Czas reakcji nawet do kilkunastu godzin	Możliwość natychmiastowej reakcji, większa autonomia
Konieczność przesyłania kompletu danych przy każdym kontakcie ze stacją naziemną	Oszczędność pasma komunikacyjnego i pamięci
Dostęp do klastrów obliczeniowych na Ziemi (złożone modele, najnowsze metody, wyższa skuteczność)	Ograniczone możliwości sprzętowe na pokładzie (budżet energetyczny, chłodzenie, dostępność sprzętu i bibliotek)
ESA Anomaly Detection Benchmark (ESA-ADB)	OPS-SAT Anomaly Detection Benchmark (OPSSAT-AD)



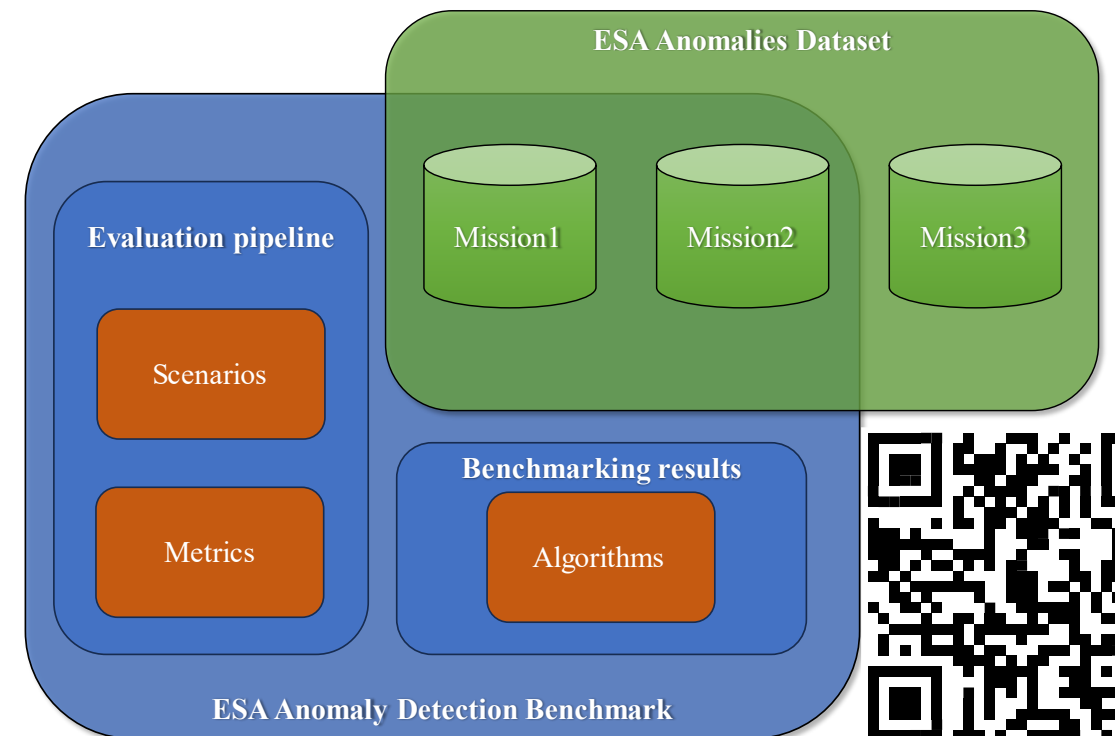
ESA Anomaly Detection Benchmark

Projekt we współpracy z Airbus, finansowany z ESA (kontrakt 4000137682/22/D/SR) jako część *Artificial Intelligence for Automation (A2I) Roadmap*

Wynik: duży zbiór danych z 3 satelitów ESA przygotowany z ich operatorami i wiarygodny proces walidacji algorytmów do detekcji anomalii

Rozwiązane wyzwania:

1. Istniejące zbiory danych i metryki nie odzwierciedlają rzeczywistych potrzeb operatorów
 - Wu, R., & Keogh, E., Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress. IEEE TKDE, 2023
 - małe i trywialne zbiory, niewłaściwe metryki, niespójne anotacje
2. Brak zdefiniowanej nomenklatury i taksonomii anomalii w danych telemetrycznych
3. Anotacje danych rzeczywistych wymagają dużych nakładów pracy i wiedzy eksperckiej

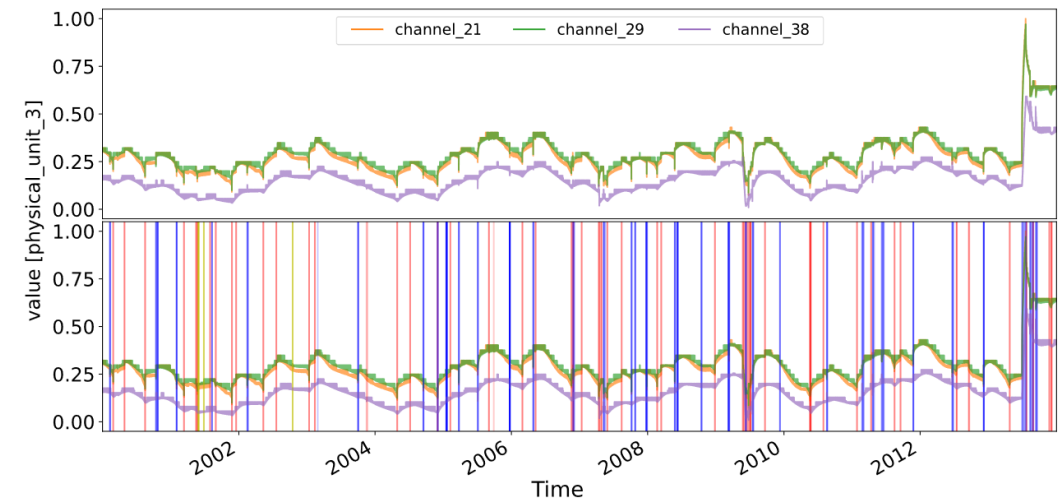


ESA Anomaly Detection Benchmark

Nowe wyzwania:

- Większość istniejących algorytmów nie osiąga zadowalających wyników lub nie jest przystosowana do tego rodzaju danych
 - Uruchomienie wyzwania na platformie Kaggle planowane w Q1 2025 żeby zachęcić do proponowania nowych rozwiązań

Algorithm	R1	R2	R3	R4	R5	R6	R7	R8	R9	Included in ESA-ADB
COPOD ⁶⁰	1	0.5	1	1	1	0	0	0	1	NO
HBOS ⁵³	1	1	0	1	0.5	0	0	0	1	YES
iForest ⁵⁰	1	1	1	1	0.5	0	0	0	1	YES
Windowed iForest ³⁰	1	1	1	1	0.5	0	0	0	0.5	SUBSETS
k-Means ⁷³	1	1	1	1	0.5	0	0	0	0.5	NO
KNN ³⁴	1	1	1	1	0.5	0	0.5	0	0.5	SUBSETS
LOF ⁷²	1	1	1	1	0.5	0	0	0	0.5	NO
Matrix Profile ^{36,37}	1	1	0.5	0	0.5	0	0.5	0	1	NO
PCC ³²	1	0.5	1	1	0.5	0	0	0	1	YES
Torsk ⁷⁴	0.5	1	1	1	1	0	0	0	0.5	NO
DAE ⁷⁶	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE ^{28*}	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE-ESA*	1	1	1	0.5	1	1	0	0	1	YES
GlobalSTD*	1	1	0	0.5	1	0	0	0	1	YES
Hybrid KNN ⁷⁷	1	1	1	0	0.5	0	0.5	0	0.5	NO
LSTM-AD ⁷⁸	0.5	1	1	0	0	0	0	0	0.5	NO
OmniAnomaly ²⁶	0.5	1	1	0	0.5	0	0	0	0.5	NO
RobustPCA ⁷⁵	0.5	0.5	1	0.5	0.5	0	0	0	1	NO
Telemanom ²	1	1	1	0	1	0	0	0	0.5	NO
Telemanom-ESA*	1	1	1	0.5	1	1	0	0	1	YES

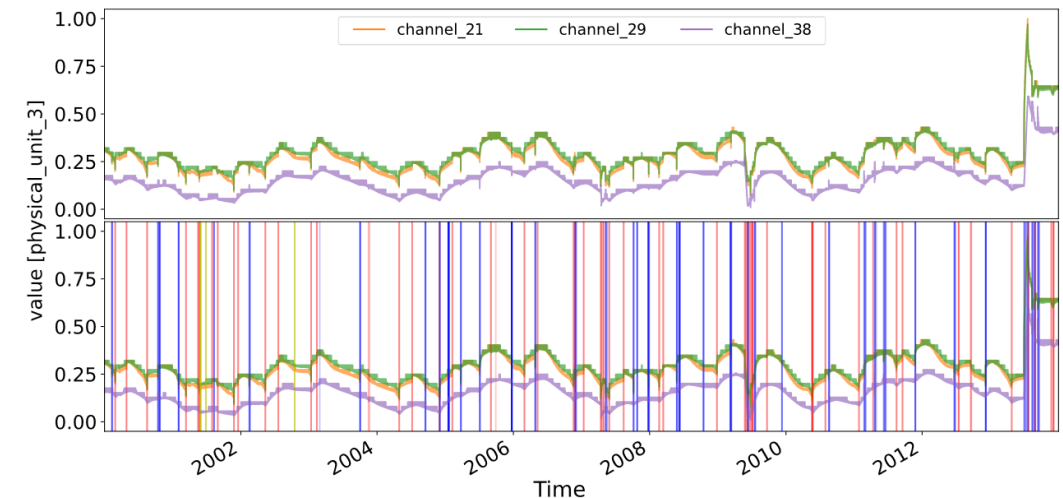


ESA Anomaly Detection Benchmark

Nowe wyzwania:

- Większość istniejących algorytmów nie osiąga zadowalających wyników lub nie jest przystosowana do tego rodzaju danych
 - Uruchomienie wyzwania na platformie Kaggle planowane w Q1 2025 żeby zachęcić do proponowania nowych rozwiązań
- Brak skutecznych algorytmów uczenia ciągłego (continual learning) i na małej liczbie próbek (few-shot learning) dla tego zadania
 - „Few-shot Anomaly Detection in Satellite Telemetry”, (ESA contract 4000141301)
 - „On-Board Continual Learning in SatCom Systems”, (ESA contract 4000142406, **Eutelsat OneWeb, OHB**)

Algorithm	R1	R2	R3	R4	R5	R6	R7	R8	R9	Included in ESA-ADB
COPOD ⁶⁰	1	0.5	1	1	1	0	0	0	1	NO
HBOS ⁵³	1	1	0	1	0.5	0	0	0	1	YES
iForest ⁵⁰	1	1	1	1	0.5	0	0	0	1	YES
Windowed iForest ³⁰	1	1	1	1	0.5	0	0	0	0.5	SUBSETS
k-Means ⁷³	1	1	1	1	0.5	0	0	0	0.5	NO
KNN ³⁴	1	1	1	1	0.5	0	0.5	0	0.5	SUBSETS
LOF ⁷²	1	1	1	1	0.5	0	0	0	0.5	NO
Matrix Profile ^{36,37}	1	1	0.5	0	0.5	0	0.5	0	1	NO
PCC ³²	1	0.5	1	1	0.5	0	0	0	1	YES
Torsk ⁷⁴	0.5	1	1	1	1	0	0	0	0.5	NO
DAE ⁷⁶	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE ^{28*}	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE-ESA*	1	1	1	0.5	1	1	0	0	1	YES
GlobalSTD*	1	1	0	0.5	1	0	0	0	1	YES
Hybrid KNN ⁷⁷	1	1	1	0	0.5	0	0.5	0	0.5	NO
LSTM-AD ⁷⁸	0.5	1	1	0	0	0	0	0	0.5	NO
OmniAnomaly ²⁶	0.5	1	1	0	0.5	0	0	0	0.5	NO
RobustPCA ⁷⁵	0.5	0.5	1	0.5	0.5	0	0	0	1	NO
Telemanom ²	1	1	1	0	1	0	0	0	0.5	NO
Telemanom-ESA*	1	1	1	0.5	1	1	0	0	1	YES

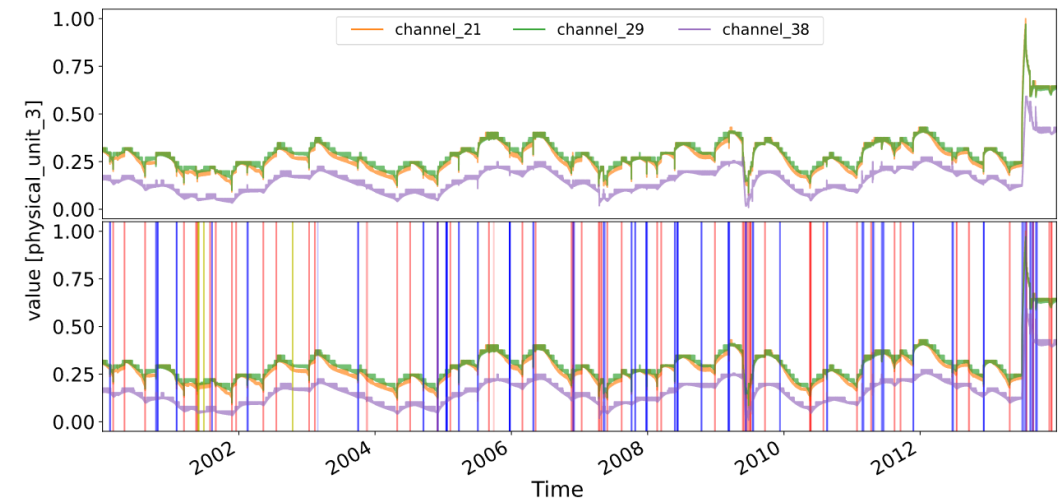


ESA Anomaly Detection Benchmark

Nowe wyzwania:

- Większość istniejących algorytmów nie osiąga zadowalających wyników lub nie jest przystosowana do tego rodzaju danych
 - Uruchomienie wyzwania na platformie Kaggle planowane w Q1 2025 żeby zachęcić do proponowania nowych rozwiązań
- Brak skutecznych algorytmów uczenia ciągłego (continual learning) i na małej liczbie próbek (few-shot learning) dla tego zadania
 - „Few-shot Anomaly Detection in Satellite Telemetry”, (ESA contract 4000141301)
 - „On-Board Continual Learning in SatCom Systems”, (ESA contract 4000142406, **Eutelsat OneWeb, OHB**)
- Wyjaśnialność modeli i zaufanie operatorów
 - „Pineberry: Explainable, Robust, and Secure AI for Demystifying Space Mission Operations”, (ESA contract 4000144194, **Politechnika Warszawska, zespół prof. Przemysława Biecka**)

Algorithm	R1	R2	R3	R4	R5	R6	R7	R8	R9	Included in ESA-ADB
COPOD ⁶⁰	1	0.5	1	1	1	0	0	0	1	NO
HBOS ⁵³	1	1	0	1	0.5	0	0	0	1	YES
iForest ⁵⁰	1	1	1	1	0.5	0	0	0	1	YES
Windowed iForest ⁵⁰	1	1	1	1	0.5	0	0	0	0.5	SUBSETS
k-Means ⁷³	1	1	1	1	0.5	0	0	0	0.5	NO
KNN ³⁴	1	1	1	1	0.5	0	0.5	0	0.5	SUBSETS
LOF ⁷²	1	1	1	1	0.5	0	0	0	0.5	NO
Matrix Profile ^{36,37}	1	1	0.5	0	0.5	0	0.5	0	1	NO
PCC ³²	1	0.5	1	1	0.5	0	0	0	1	YES
Torsk ⁷⁴	0.5	1	1	1	1	0	0	0	0.5	NO
DAE ⁷⁶	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE ^{28*}	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE-ESA*	1	1	1	0.5	1	1	0	0	1	YES
GlobalSTD*	1	1	0	0.5	1	0	0	0	1	YES
Hybrid KNN ⁷⁷	1	1	1	0	0.5	0	0.5	0	0.5	NO
LSTM-AD ⁷⁸	0.5	1	1	0	0	0	0	0	0.5	NO
OmniAnomaly ²⁶	0.5	1	1	0	0.5	0	0	0	0.5	NO
RobustPCA ⁷⁵	0.5	0.5	1	0.5	0.5	0	0	0	1	NO
Telemanom ²	1	1	1	0	1	0	0	0	0.5	NO
Telemanom-ESA*	1	1	1	0.5	1	1	0	0	1	YES

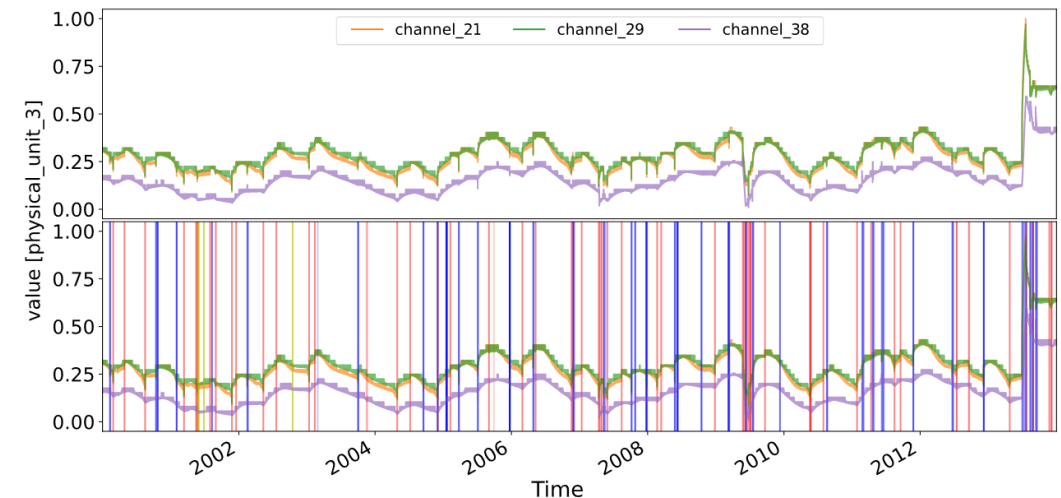


ESA Anomaly Detection Benchmark

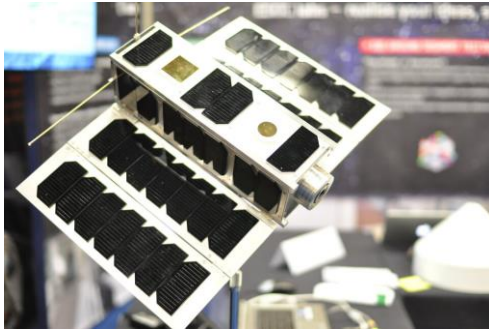
Nowe wyzwania:

- Większość istniejących algorytmów nie osiąga zadowalających wyników lub nie jest przystosowana do tego rodzaju danych
 - Uruchomienie wyzwania na platformie Kaggle planowane w Q1 2025 żeby zachęcić do proponowania nowych rozwiązań
- Brak skutecznych algorytmów uczenia ciągłego (continual learning) i na małej liczbie próbek (few-shot learning) dla tego zadania
 - „Few-shot Anomaly Detection in Satellite Telemetry”, (ESA contract 4000141301)
 - „On-Board Continual Learning in SatCom Systems”, (ESA contract 4000142406, **Eutelsat OneWeb, OHB**)
- Wyjaśnialność modeli i zaufanie operatorów
 - „Pineberry: Explainable, Robust, and Secure AI for Demystifying Space Mission Operations”, (ESA contract 4000144194, **Politechnika Warszawska, zespół prof. Przemysława Biecka**)
- Weryfikacja, walidacja i wdrożenie algorytmów uczenia maszynowego w praktyce operacyjnej
 - „Verification And Validation Methods For Machine Learning Algorithms”, (ESA przetarg T702-803SW)

Algorithm	R1	R2	R3	R4	R5	R6	R7	R8	R9	Included in ESA-ADB
COPOD ⁶⁰	1	0.5	1	1	1	0	0	0	1	NO
HBOS ⁵³	1	1	0	1	0.5	0	0	0	1	YES
iForest ⁵⁰	1	1	1	1	0.5	0	0	0	1	YES
Windowed iForest ³⁰	1	1	1	1	0.5	0	0	0	0.5	SUBSETS
k-Means ⁷³	1	1	1	1	0.5	0	0	0	0.5	NO
KNN ³⁴	1	1	1	1	0.5	0	0.5	0	0.5	SUBSETS
LOF ⁷²	1	1	1	1	0.5	0	0	0	0.5	NO
Matrix Profile ^{36,37}	1	1	0.5	0	0.5	0	0.5	0	1	NO
PCC ³²	1	0.5	1	1	0.5	0	0	0	1	YES
Torsk ⁷⁴	0.5	1	1	1	1	0	0	0	0.5	NO
DAE ⁷⁶	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE ^{28*}	0.5	1	1	0	1	0	0	0	0.5	NO
DC-VAE-ESA*	1	1	1	0.5	1	1	0	0	1	YES
GlobalSTD*	1	1	0	0.5	1	0	0	0	1	YES
Hybrid KNN ⁷⁷	1	1	1	0	0.5	0	0.5	0	0.5	NO
LSTM-AD ⁷⁸	0.5	1	1	0	0	0	0	0	0.5	NO
OmniAnomaly ²⁶	0.5	1	1	0	0.5	0	0	0	0.5	NO
RobustPCA ⁷⁵	0.5	0.5	1	0.5	0.5	0	0	0	1	NO
Telemanom ²	1	1	1	0	1	0	0	0	0.5	NO
Telemanom-ESA*	1	1	1	0.5	1	1	0	0	1	YES



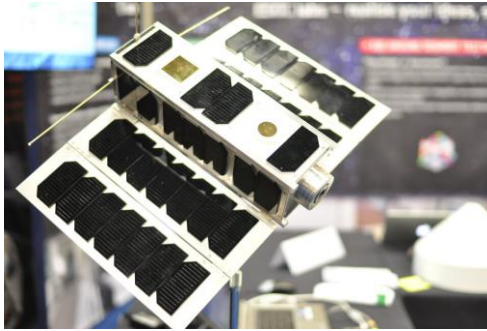
ESA OPS-SAT aircraft



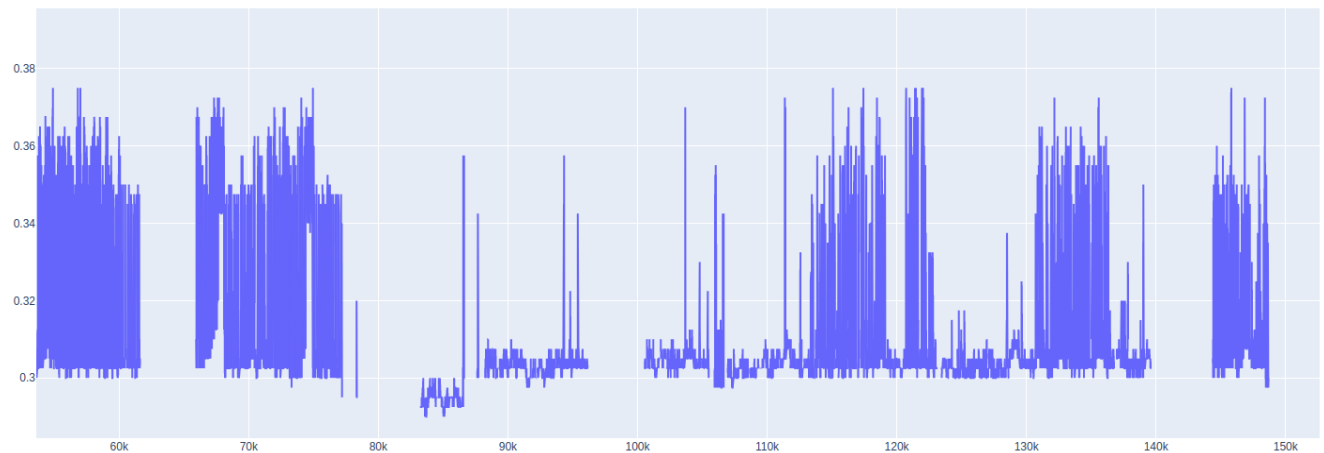
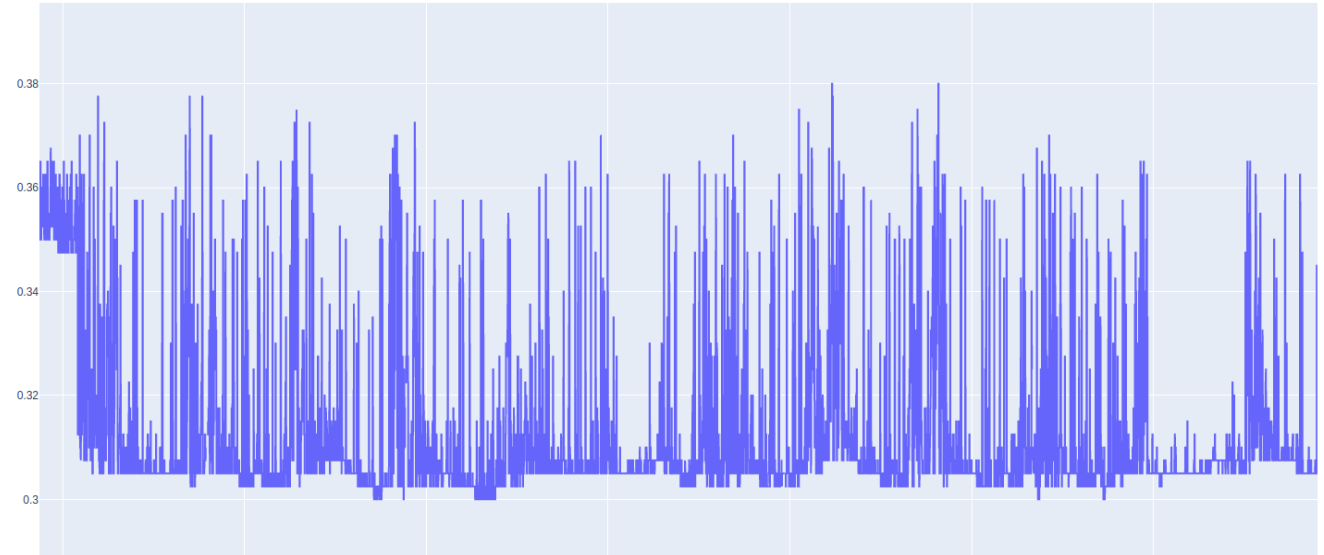
- *European Space Agency Flying Lab*
- Few thousands of numeric telemetry series
- Sampling rate 1s - 30s
- Commands sent from ground station



OPS-SAT telemetry

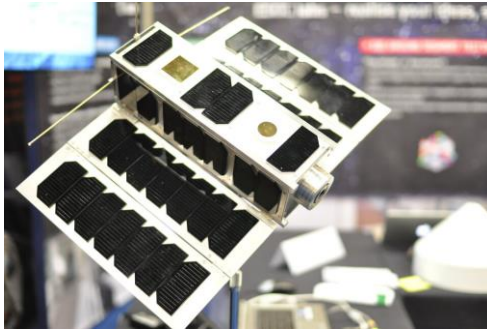


- Discontinuities in data
- Poses problem while processing with Machine Learning models



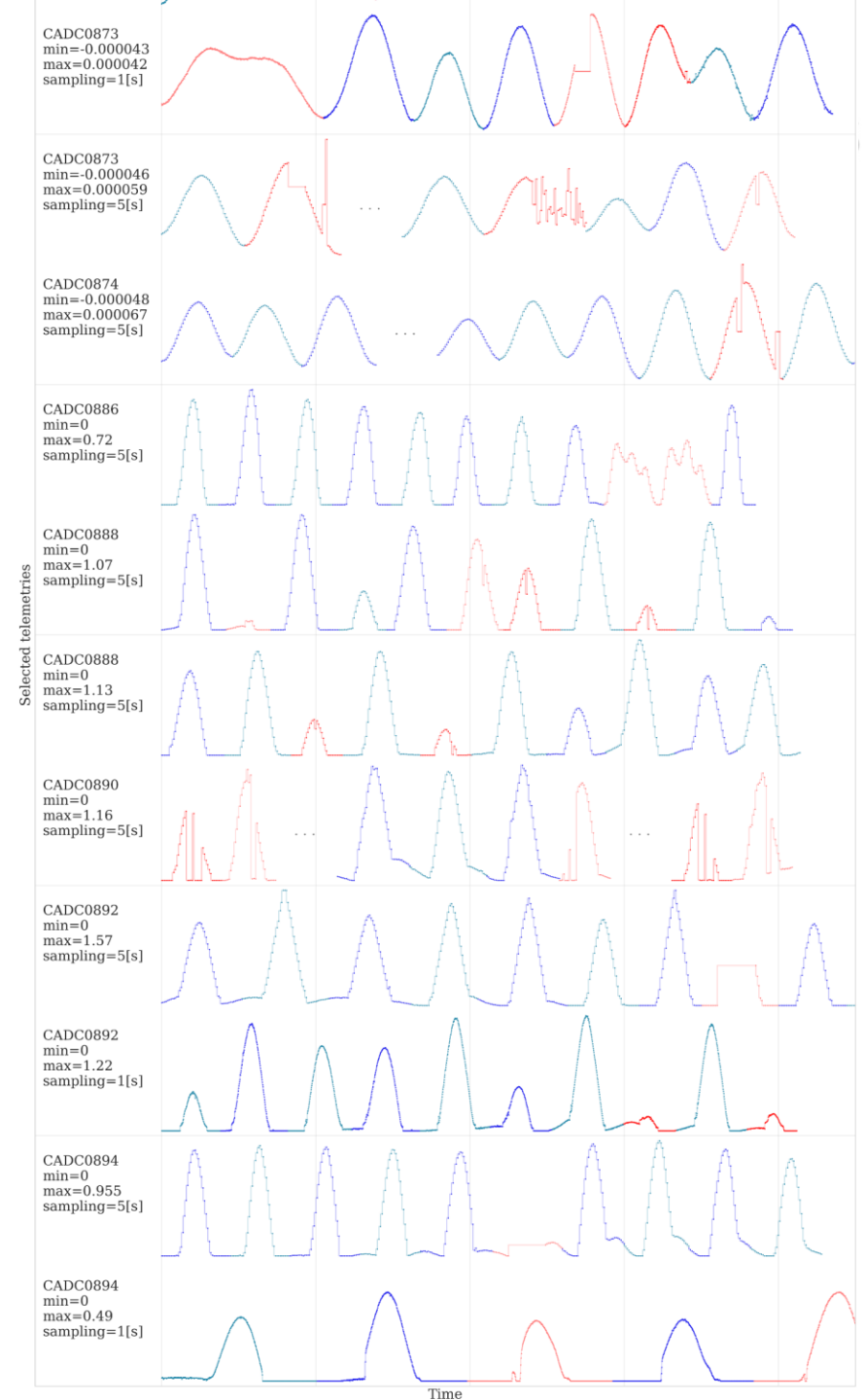
Values of SEP8260p: continuous fragment above, gaps visible below

OPSSAT-AD Dataset

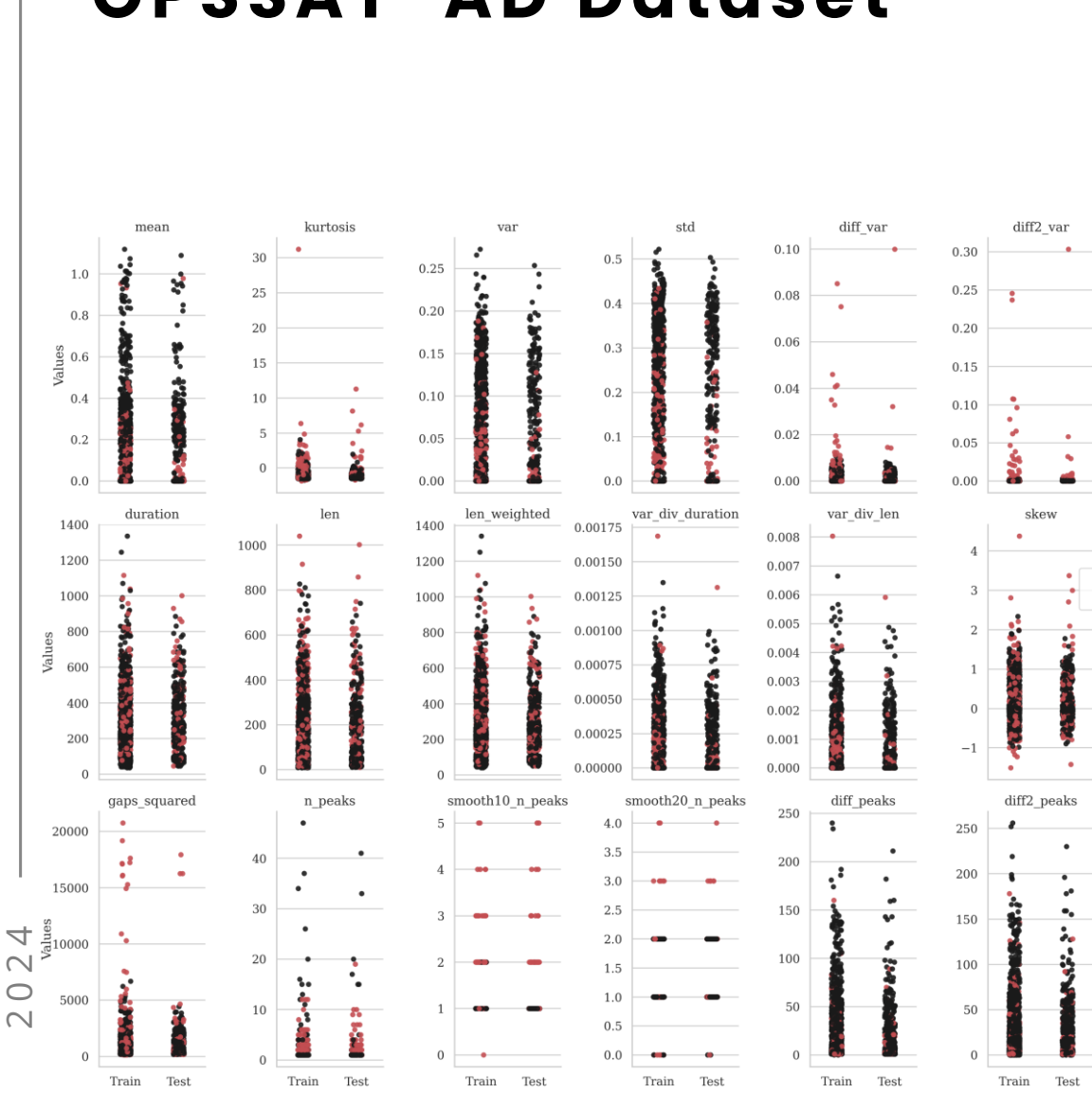
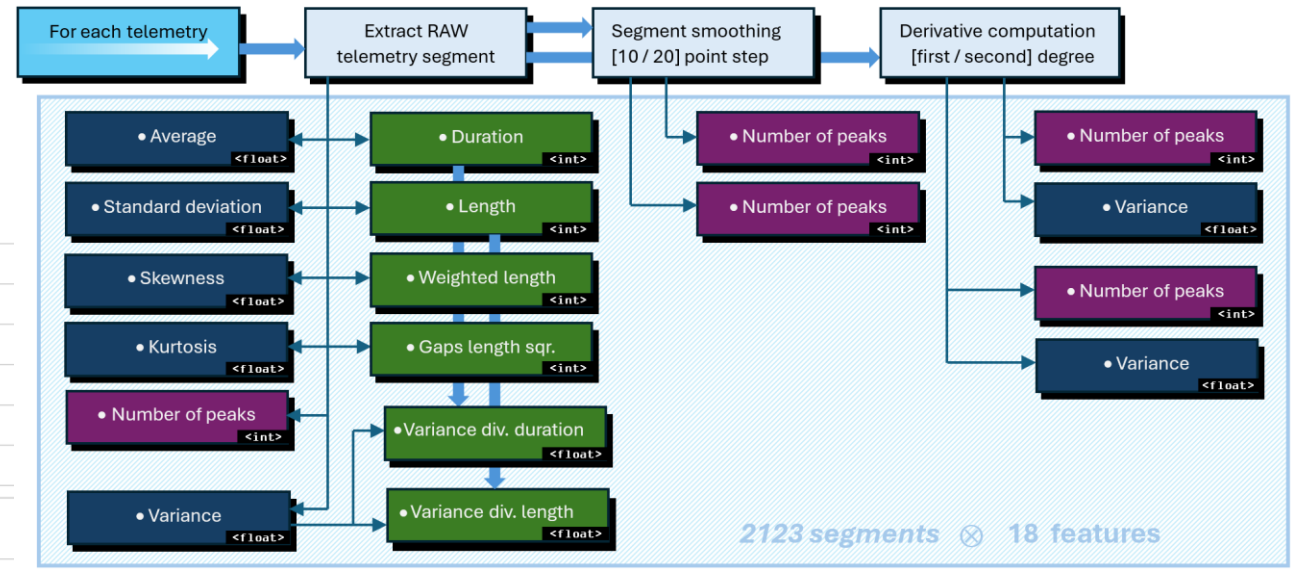
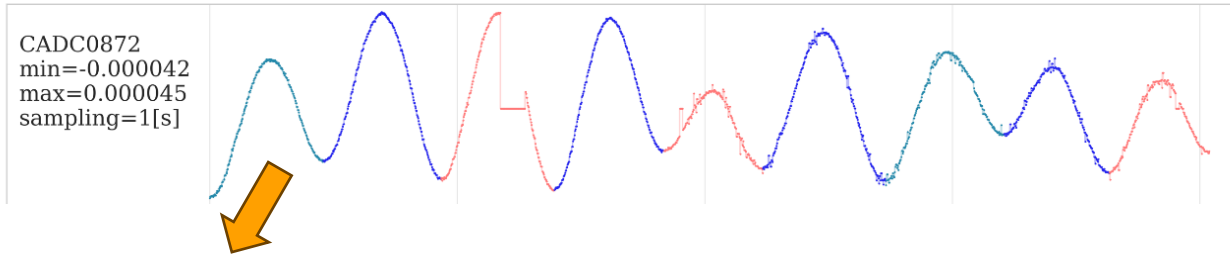


OPSSAT-AD:

- 2123 telemetry segments
- Nominal: 1689
- Anomalous: 434
- Training set: 1494
- Training set: 529 (25%)



OPSSAT-AD Dataset



2024

• synthetic features

• distribution of samples

OPSSAT-AD results

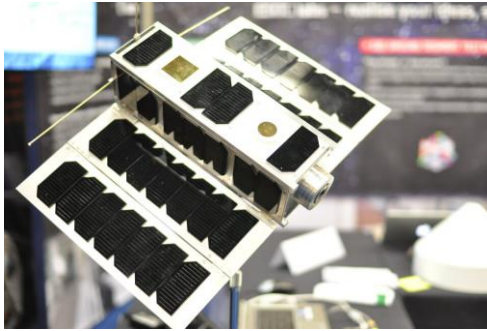
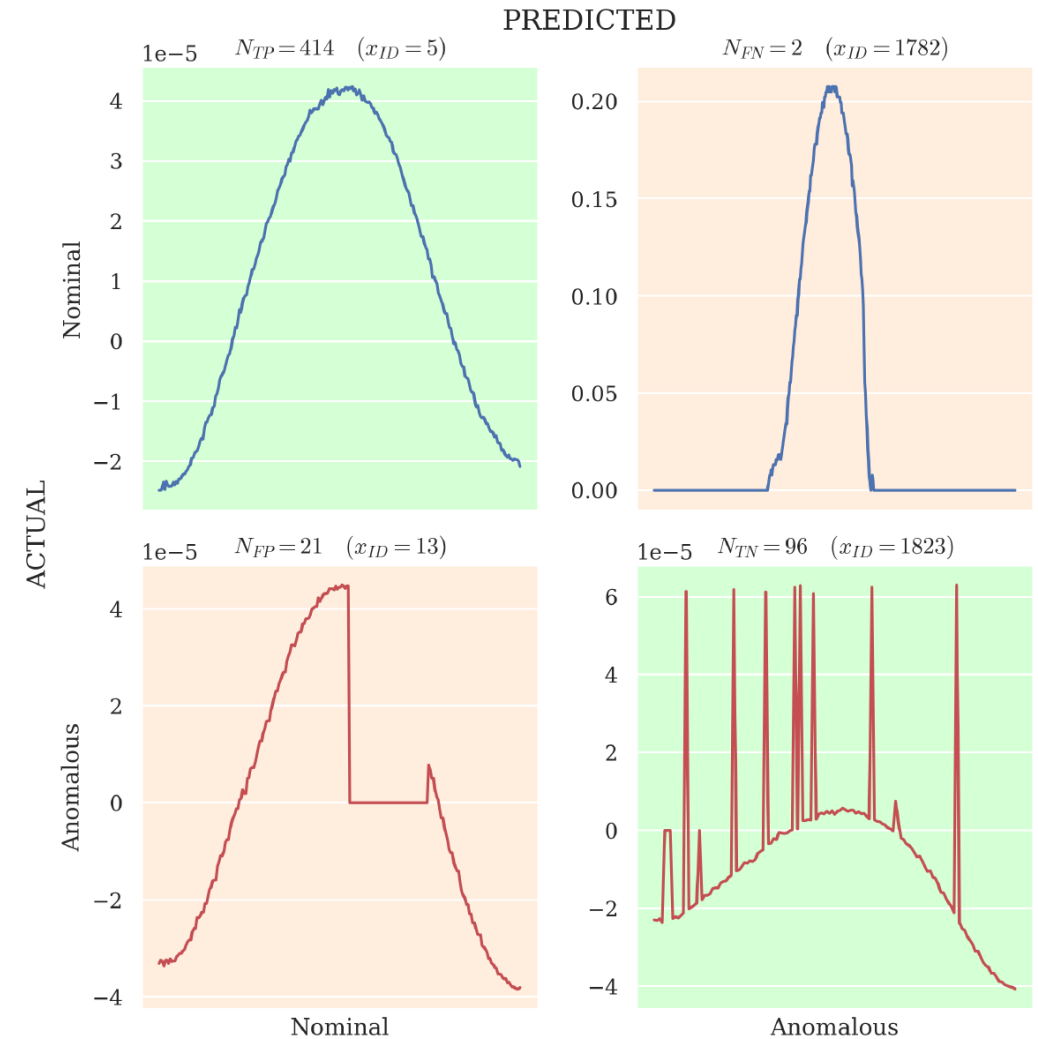


Table 4. The results (obtained over the test set) elaborated using all investigated classification models with optimized hyperparameters and trained over the augmented training set ($\omega_1 \wedge \omega_2 \wedge \omega_3$), sorted by r_ϕ . The best results are **boldfaced**.

Model	Accuracy	Precision	Recall	F_1	r_ϕ
Random forest	0.9843	1.0000	0.7317	0.8239	0.8261
Multi-layer perceptron	0.9792	0.8716	0.7437	0.7795	0.7815
Adaptive boosting	0.9812	0.9422	0.6858	0.7817	0.7893
SVM with linear kernel	0.9729	0.9097	0.6154	0.6683	0.6802
k -nearest neighbors	0.9491	0.8662	0.6637	0.7186	0.7008

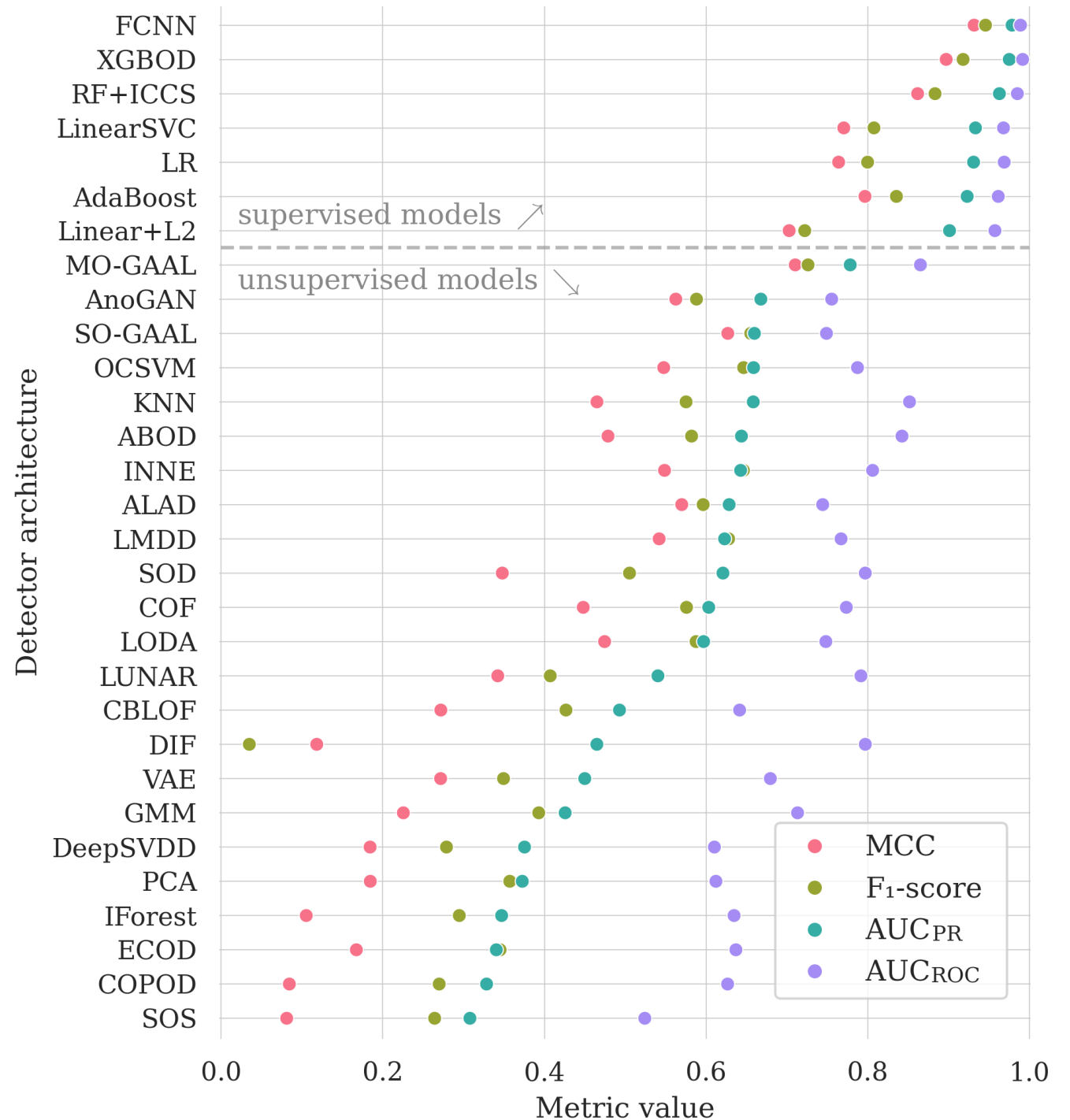


OPSSAT-AD results



- Full dataset release ★
 - github.com/kplabs-pl/AI-datasets
- + benchmarking results

B. Ruszczak, J. Nalepa, K. Kotowski, and D. Evans,
 “The OPS-SAT benchmark for detecting anomalies in satellite telemetry,”
 arXiv:2407.04730, 2024.



Dziękujemy za uwagę

Krzysztof Kotowski i Bogdan Ruszczak

kkotowski@kplabs.pl
b.ruszczak@po.edu.pl

- **Zbiory danych** →
 - **Kod** →
- **Artykuły naukowe** →

**dla detekcji anomalii
w szeregach czasowych
i innych zadań też 😊**



★ github.com/kplabs-pl/AI-datasets ★