AI-Powered Anomaly Detection for Satellite Telemetry

POSTER: AI-Powered Anomaly Detection for Satellite Telemetry

Federico Buccellato
federico.buccellato@polito.it
Politecnico di Torino
Turin, Italy

Davide Nicolini
davide.nicolini@polito.it
Politecnico di Torino
Turin, Italy

Eleonora Vacca
eleonora.vacca@polito.it
Politecnico di Torino
Turin, Italy

Corrado De Sio
corrado.desio@polito.it
Politecnico di Torino
Turin, Italy

Luca Sterpone
luca.sterpone@polito.it
Politecnico di Torino
Turin, Italy

ABSTRACT

Reliable anomaly detection in satellite telemetry is critical for mission success, yet traditional threshold-based methods struggle with complex and evolving patterns. This work presents machine learning (ML) techniques to analyze high-dimensional telemetry data. Evaluations of real-world satellite telemetry datasets demonstrate the potential of ML to enhance spacecraft health monitoring and reduce manual intervention.

CCS CONCEPTS

• **Hardware** → **Robustness**; **Safety critical systems**;

KEYWORDS

Satellite telemetry, Reliability, Machine Learning, Anomalies Detection

ACM Reference Format:

Federico Buccellato, Davide Nicolini, Eleonora Vacca, Corrado De Sio, and Luca Sterpone. 2025. POSTER: AI-Powered Anomaly Detection for Satellite Telemetry. In *Proceedings of May 28–30, 2025 (CF '25).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3719276.3727953

1 INTRODUCTION

Monitoring satellite telemetry time-series data for anomalies is a critical daily task performed by thousands of spacecraft operations engineers (SOEs) in mission control centers worldwide. It ensures the safe and continuous operation of scientific, communication, Earth observation, and navigation satellites. SOEs typically rely on automatic anomaly detection systems that flag deviations when measurements exceed predefined nominal thresholds or match known anomalous patterns [1] [4] . However, these rule-based systems struggle with more complex or previously unseen anomalies, often requiring manual identification, labor-intensive, error-prone, and costly process. Indeed, satellite telemetry data presents distinct challenges due to its high-dimensional and multivariate time-series nature, coupled with substantial data volume. Satellites remain operational for extended durations in space, continuously acquiring high-frequency telemetry from hundreds of subsystems. This results in inherently non-stationary data characteristics, influenced by varying sampling rates, intermittent data gaps, long-term trends associated with component degradation, and mission-dependent operational phases. In addition to queste complessità, the harsh space environment [2] introduces further variability, as radiation effects, extreme temperature fluctuations, and altri stressori ambientali possono introdurre rumore, errori e anomalie sia nelle letture dei sensori che nei canali di comunicazione. Therefore, traditional methods are often insufficient for detecting subtle, evolving, or previously unknown anomalies in such a complex dataset. In this scenario, space agencies including the European Space Agency (ESA) have been actively developing and testing advanced automatic anomaly detection systems. Specifically, ESA introduced a new dataset [3] containing annotated real-life telemetry from three different ESA missions aiming at promoting the development of new approaches in autonomous anomaly detection. Machine learning (ML) techniques offer a promising alternative by learning intricate patterns, adapting to non-stationary behaviors, and providing real-time anomaly detection with minimal human intervention [5]. This paper explores ML-driven approaches for anomaly detection in satellite telemetry, addressing the challenges of high-dimensional time-series data and discussing their potential for improving satellite health monitoring, fault diagnosis, and mission resilience.

CF '25, Cagliari, Italy,

ACM ISBN 979-8-4007-1528-0/2025/05
https://doi.org/10.1145/3719276.3727953

2 PROPOSED APPROACH

The proposed approach enables automatic anomaly detection in satellite telemetry data by adopting ML algorithms. The algorithms are developed on the ESA Anomalies Dataset (ESA-AD) [3], a collection of telemetry data from two ESA missions (M1 and M2). The dataset spans 168 months of satellite life for M1 and approximately 43 months for M2, encompassing over 700 million data points per mission. The dataset is organized in channels for each mission, each referring to satellite subsystems (i.e., power, thermal, attitude control, etc.). The dataset includes annotations for anomalies, rare nominal events, and communication gaps, along with the specific telemetry channels where anomalies were detected. ESA-AD accurately represents real-world telemetry characteristics, such as irregular timestamps, varying sampling rates, and overlapping anomaly segments. The initial step in model development involved refining the dataset through techniques such as normalization, noise reduction, and feature selection. Then, a design space exploration was conducted to evaluate different approaches. Seven models have been implemented and trained from scratch. The exploration ranged from traditional ML methods, such as Logistic Regression (LR), Support Vector Machines (SVM), Random Forest (RF), and Gaussian Naïve Bayes (GNB), Fully Connected Neural Network (FC–NN) to more advanced architectures including Transformers. The motivation behind this broad evaluation stemmed from the complexity of the telemetry data, which exhibits intricate temporal dynamics and interdependencies between multiple parameters. These characteristics pose challenges for models that struggle to capture long-term dependencies in multivariate time series. Consequently, both simpler models (e.g., GNB, which assumes feature independence) and more complex ones (e.g., Transformers, which leverage self-attention mechanisms to model long-range dependencies) were considered to assess their effectiveness in capturing the underlying data patterns. Among the implemented models, the best-performing ones are presented below, along with a detailed description of their characteristics.

Deep FC–NN: Consisting of 15 densely connected layers, this architecture effectively captures complex nonlinear relationships and mitigates overfitting through regularization techniques.

Transformer: This model is provided with initial dense layers for feature extraction, followed by a Transformer block employing multi-head attention mechanisms to extract contextual information and intricate dependencies within the data.

Shallow FC–NN: A compact model composed of 5 densely connected layers, designed to provide a balanced trade-off between computational simplicity and predictive accuracy. Subsequently, we conducted a thorough exploration of hyperparameter values. As a result, we selected a batch size of 32, a learning rate of 10^{-5}, 10 training epochs, and the Adam optimizer. This combination proved effective in maximizing the models' generalization capability.

3 EXPERIMENTAL RESULTS

The models have been trained and run on an NVIDIA GeForce RTX 4070. The comparison is proposed considering F0.5 score, precision, and recall, which are summarized in Table 1. As indicated by ESA [3], false positives in anomaly detection are more costly than false negatives due to their potential to trigger a complete system block. Therefore, the F0.5 score was selected as the evaluation metric.

Considering that, the **Deep FC–NN** achieved superior results, displaying high scores across all metrics due to its capacity to model complex nonlinearities and generalize effectively. The **Transformer-based Hybrid model** also performed well, leveraging attention mechanisms to capture complex relationships and dependencies within data. The **Shallow FC–NN** provided moderate performance, adequate for less computationally intensive scenarios but limited by its lower complexity. Conversely, traditional models underperformed significantly due to their linear limitations and inability to capture intricate relationships inherent to satellite telemetry data.

The ROC curves shown in Figure 1 provide a general summary of the best models' classification performance.

4 CONCLUSION

In this paper, we explored ML techniques to enable fast anomaly detection in satellite operations starting from telemetry data.

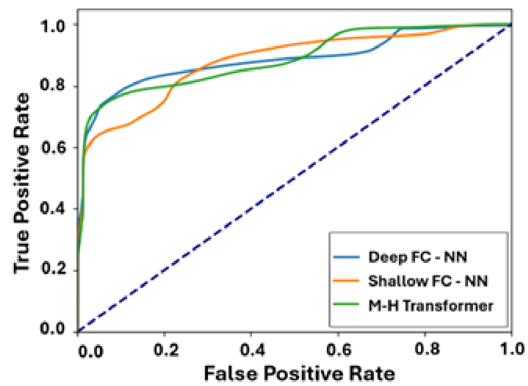

Figure 1: ROC curves displaying the performance of the top 3 best-performing models.

Table 1: Comparative performance results of evaluated models.

Model	F0.5-Score	Recall	Precision	Inference Time (s)
Deep FC–NN	0.7642	0.6366	0.8054	0.1072
Shallow FC–NN	0.6730	0.4555	0.7643	0.0742
M-H Transformer	0.7253	0.5619	0.7839	0.0708
SVM	0.4128	0.1693	0.6450	0.0076
RF	0.5360	0.2402	0.7756	0.0081
LR	0.4701	0.1929	0.7352	0.0002
GNB	0.3913	0.1456	0.6770	0.0003

Experimental results indicate that Deep FC–NN demonstrates effectiveness due to its robust modeling capabilities of complex patterns.

REFERENCES

[1] D. Liu et al. 2017. Satellite Telemetry Data Anomaly Detection with Hybrid Similarity Measures. In *2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*. Shanghai, China, 591–596. https://doi.org/10.1109/SDPC.2017.116

[2] E. Vacca et al. 2024. Radiation Environment and Effect Analysis of the Zodiac Pioneer Mission. In *75th International Astronautical Congress*. https://doi.org/10.52202/078369-0054

[3] K. Kotowsk et al. 2024. European Space Agency Benchmark for Anomaly Detection in Satellite Telemetry. https://doi.org/10.48550/arXiv.2406.17826

[4] Varsha Parthasarathy et al. 2025. A satellite fault detection system based on telemetry data using statistical process control and time-domain feature extraction. *Advances in Space Research* (2025).

[5] G. Li and J. J. Jung. 2023. Deep learning for anomaly detection in multivariate time series: Approaches, applications and challenges. *Information Fusion* 91 (2023), 93–102.