QCoKaIn
Hybrid Quantum Machine Learning with Causal Inference for Anomaly Detection

ESA Standard License, Credit: ESA / NASA

We develop and evaluate a hybrid Quantum Machine Learning (QML) algorithm for anomaly detection in telemetry data in combination with causal inference techniques. The hybrid algorithm will be trained to detect anomalies in data from the ISS Columbus module.

- Applications
- Quantum Machine Learning
- Anomaly Detection in Telemetry Data
- Causal Inference

Why use Quantum Machine Learning with Causal Inference?

Quantum machine learning (QML) is an emerging research field promising to solve computational problems intractable even to modern deep learning methods. Despite its potential to improve algorithm performance with enhanced computational power, applying QML alone still carries the same "black box" concerns as the classical machine learning. This includes the lack of interpretability and reliability.

With the use of causal inference (Fig. 1), QML algorithms can offer insights into why a certain outcome or a prediction had occurred. Moreover, by understanding the causal relationships between different input variables, we can determine the propagation of anomalies and gain knowledge of the underlying mechanisms of a given physical system, undeterred by its complexity (Fig. 2).

In this project, we develop a hybrid quantum anomaly detection architecture that uses techniques of causal inference and machine learning, applied to a coupled nonlinear dynamical system. These techniques allow for more robust and explainable data analysis. In addition to improved efficacy, this leads to a more reliable anomaly detection in high-volume telemetry data.

ISS Columbus Anomaly Detection Use Case

Our use case is centered on the Environmental Control and Life Support System (ECLSS) of the ISS Columbus module, which is responsible for maintaining a habitable environment for the crew. Currently, anomaly detection is a manual process where a flight control team monitors telemetry (TM) and creates an anomaly report upon detecting an issue.
The primary challenges for this use case stem from the inherent nature of the Columbus module. As a system built for human spaceflight, it is engineered for exceptional robustness, stability and redundancy. While this ensures crew safety, it creates significant data science hurdles:
The first is a **severe class imbalance** - The system is so reliable that anomalous events make up a tiny fraction of the overall dataset. This scarcity of failure examples makes it challenging to train an ML model to recognize what constitutes an anomaly.
A second challenge lies in the **dual nature of the data stability**. While the overall system can operate in different modes, leading to non-stationary behavior, many TM parameters within any given mode are very stable and static. Anomaly detection models learn from dynamic variations. Training on predominantly static signals makes it difficult to discern the subtle patterns that differentiate normal fluctuations from genuine faults.

Machine Learning Algorithm Selection

The analysis of the data generated by ECLSS submodule requires a data-driven approach due to its quantity, complexity, and the previously-mentioned challenges. In addition, we ultimately aim to develop a method that would exploit the full potential of quantum computers. However, we first had to establish a classical baseline for the given task. We therefore explored a suite of machine learning algorithms tailored for multivariate time series anomaly detection and prioritized the architectures suitable for quantum hybridization. Such architectures are currently based on Variational Autoencoders (VAEs)[1], Generative Adversarial Networks (GANs)[2], Transformers[3]. We also considered a method based on Long Short Term Memory networks (LSTMs)[4].

Following the ESA's comparative analysis[5], among the methods we appraised were:

- DC-VAE[6]
- DCT-GAN[7]
- TranAD[8]
- Telemanom[9]

Finally, as a classical baseline, we selected an advanced model due to its superior benchmark performance and explainability features.
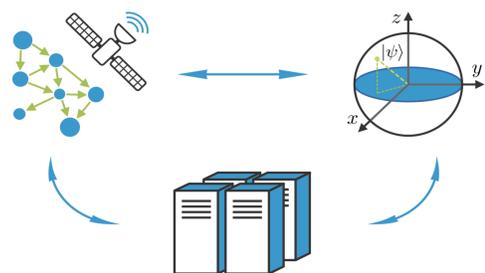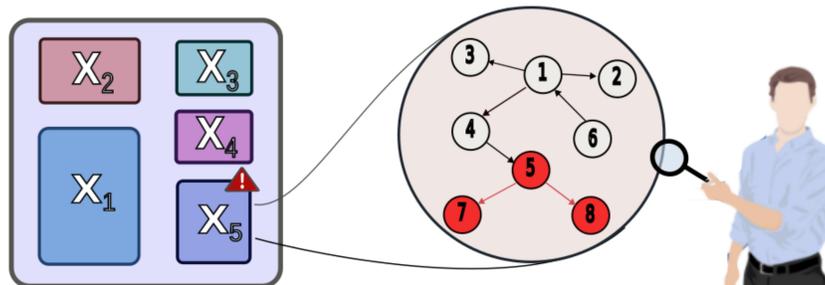

Fig. 1: The essential constituents of the project: Causal inference within the context of telemetry data (left), high-performance computing (middle), and quantum computing (right)

Fig.2: Monitoring anomalies: Analysis of different system components (left), anomaly detection (right)
(Credits: Airbus and DLR)

Challenges of Quantum Machine Learning on Telemetry Data

We will train our QML method on the telemetry data from the Columbus' ECLSS module. This highly complex system generates several thousand telemetry parameters, recorded with an average sampling rate of 1 Hz, which results in around 10 GB of data per year. Apart from the sheer magnitude, telemetry data comes with the following challenges for QML:

High dimensionality - within the scope of this project, we constrain ourselves to the analysis of just over 100 telemetry channels from the Columbus' Environmental Control and Life Support Subsystem (ECLSS). Nevertheless, running hybrid quantum simulations remains inefficient. We aim to tackle this issue by analyzing causal links between channels.
Diverse data structure and high **noise** levels can make training the hybrid algorithm difficult because they can reduce accuracy and cause overfitting. We address these challenges by thorough data preprocessing.
Non-stationarity and binary switching of the data can inhibit QML models by introducing distribution shifts and instability. Their resulting poor generalizability can be alleviated by using techniques of causal inference, for example finding the factors which drive the non-stationarity.

Main Objectives

In addition to addressing the data-related challenges, our main objective is to tackle the obstacles posed by the three main modelling stages underlying foundational QML algorithms. Those stages include definition of a **data encoding**, design of **trainable circuit layers** and determination of an expressive **measurement reading**. By addressing these challenges and carefully crafting the QML algorithm at each stage, we aim to create a QML solution for time series anomaly detection that will open the door for applications beyond simulation and enable an efficient use of quantum resources.

1) Quantum encoding (rich encoding scheme) – classical data needs to be encoded into quantum states. We intend to find an encoding scheme that balances the expressivity and trainability of the respective quantum model.

2) Trainable circuit layers (model architecture) – selecting suitable parameterized variational gates is cruicial for the QML model efficiency and applicability that exceeds simulations.

3) Measurement reading (efficient readout) – output of quantum circuits is measured multiple times to obtain classical predictions. The main objective is to optimize the number of shots, in order to accelerate training.

Potential Impact

- More reliable QML applications due to explainability gained by analyzing causal links between different telemetry channels and determining origins of detected anomalies
- Gaining new insights into the unique potential of QML algorithms for anomaly detection
- Exploring benefits and challenges of combining classical with quantum ML
- Monitoring and pin-pointing errors in complex dynamical systems
- Better data-driven understanding of those systems
- Reducing the rate of false negative detections

References
[1] Kingma et al., "Auto-Encoding Variational Bayes", 2013
[2] Goodfellow et al., "Generative Adversarial Networks", 2014
[3] Vaswani et al., "Attention Is All You Need", 2017
[4] Hochreiter et al., "Long Short-Term Memory", 1997
[5] Kotowski et al., "European Space Agency Benchmark for Anomaly Detection in Satellite Telemetry", 2024
[6] González et al., "One Model to Find Them All Deep Learning for Multivariate Time-Series Anomaly Detection in Mobile Network Data", 2023
[7] Li et al., "DCT-GAN: Dilated Convolutional Transformer-based GANfor Time Series Anomaly Detection", 2021
[8] Tuli et al., "Tranad: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data", 2022
[9] Hundman et al., "Detecting Spacecraft Anomalies using LSTMs and Nonparametric Dynamic Thresholding", 2018

More information about the project on our website