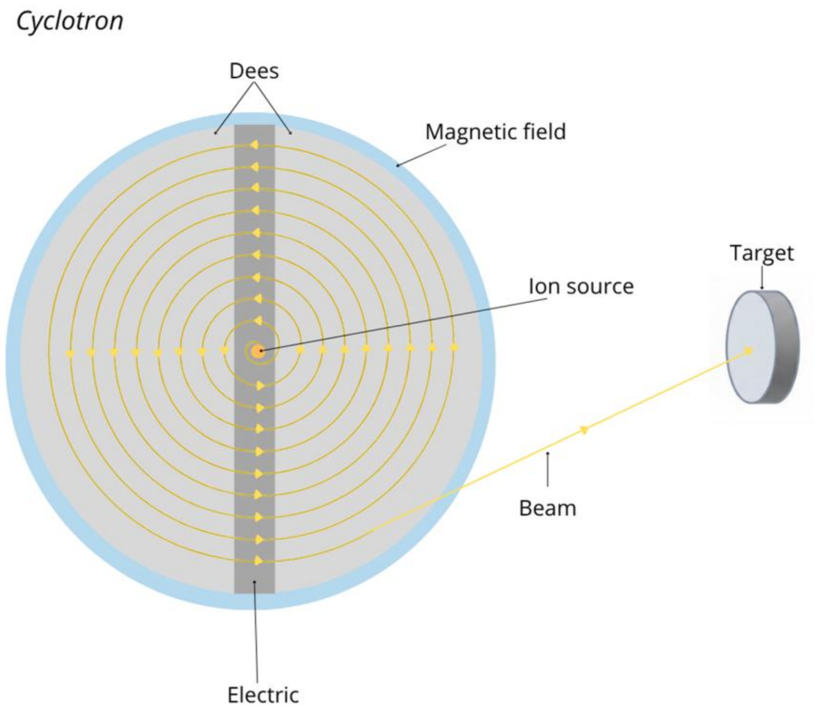
Hybrid Autoencoder-Isolation Forest Approach for Time Series Anomaly Detection in C70XP Cyclotron Operation Data at ARRONAX

Fatima Basbous (ARRONAX), Freddy Poirier (ARRONAX), Diana Mateus (LS2N), Ferid Haddad (ARRONAX)

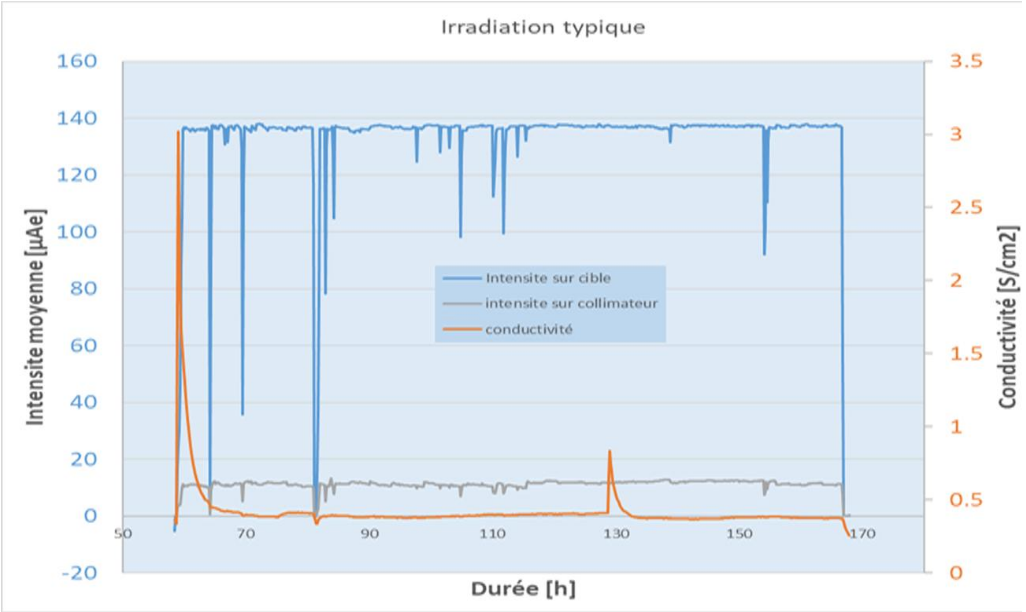
IN2P3/IRFU Machine Learning workshop 2025
Caen, France



Schematic of the proton beam from the source to the target.

EPICS

Experimental and Industrial Physical Control System



Example of several temporal variables acquired through EPICS. The proton beam intensity is generally stable, although it occasionally experiences stops, variations, and drops.

Anomaly detection

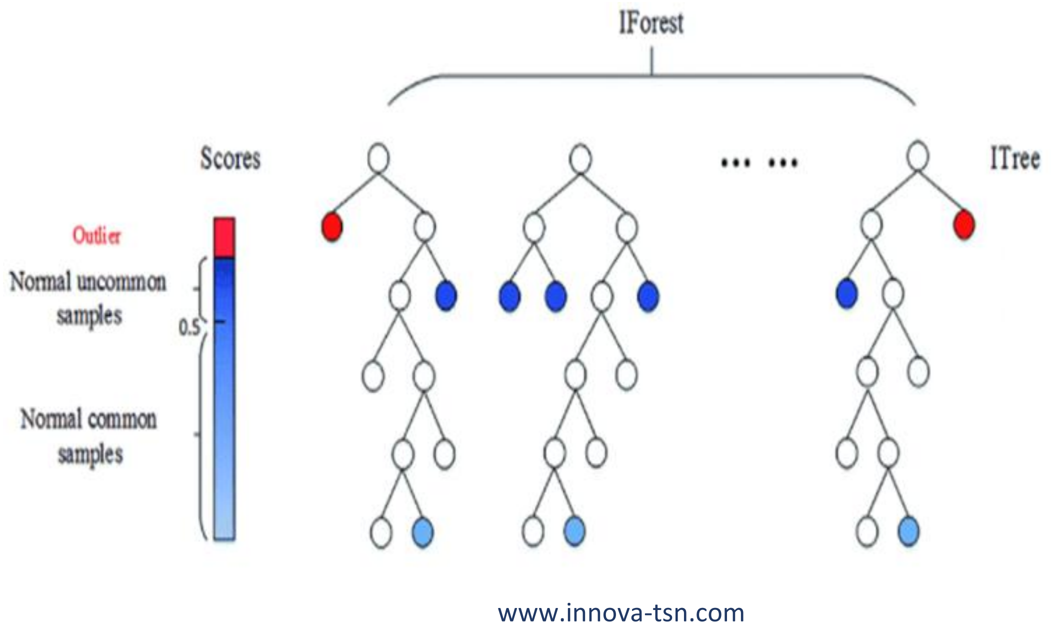
Isolation Forest (IF): A Reference Approach

Principle

IF isolates anomalies more easily than normal points by using random trees and assigning higher anomaly scores to points with shorter path lengths (*Liu et al., 2008*).

Strengths

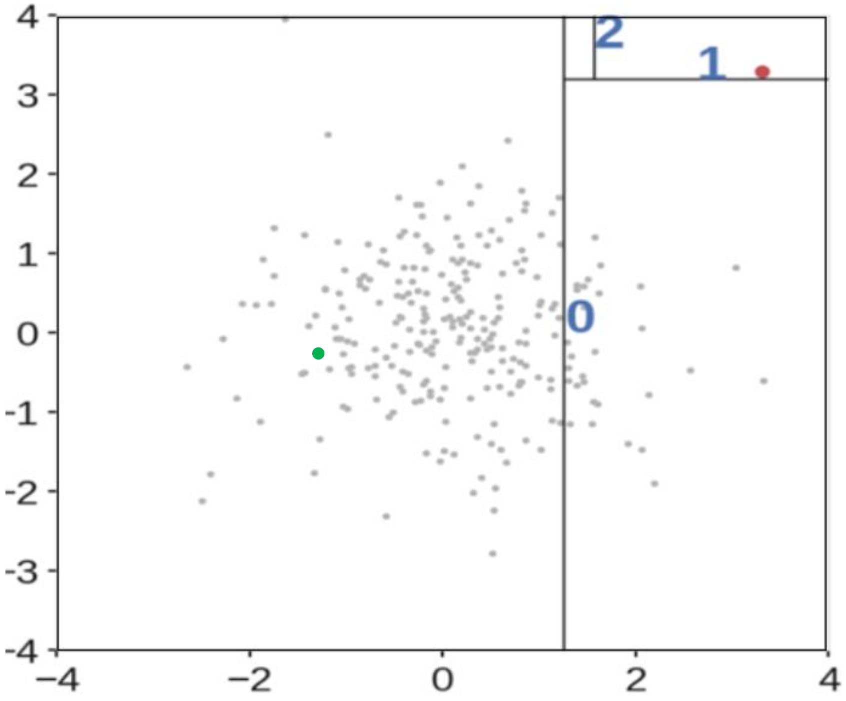
Efficiency, scalability, and widespread use in real-world applications *(Chua et al. (2024), Kumar et al. (2024), Zerkouk et al. (2023), Ahmed et al. (2019))*.



Mechanism of IF for Anomaly Detection.

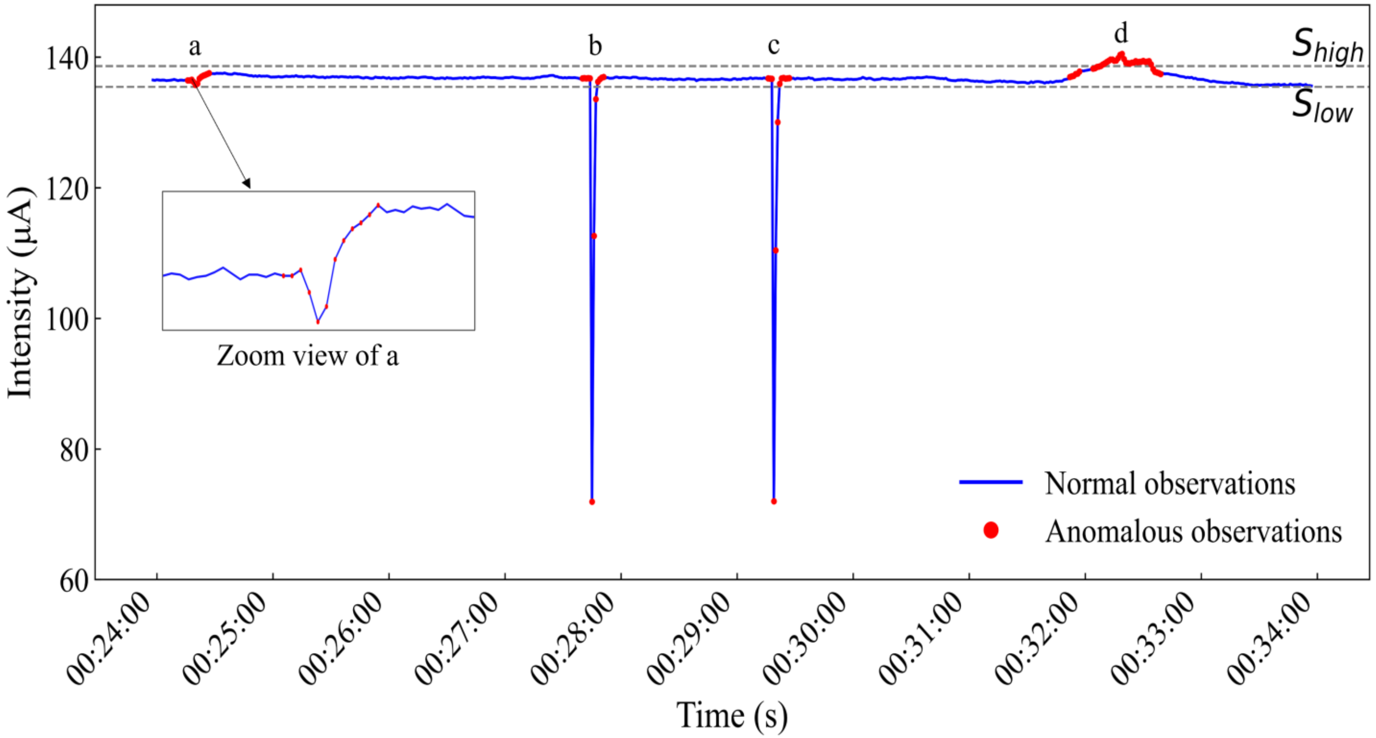
Research Problem

- **Limitation:** Difficulties with subtle anomalies due to axis-parallel partitions.
- **Global anomalies:** Large deviations beyond thresholds →visible at large scale.
- **Local anomalies:** Significant internal fluctuations even when values remain within thresholds →visible only when zoomed in.



analyticsvidhya.com

Example of anomaly point x_0 isolation from a set of points of a Gaussian distribution by axis-parallel partitioning using IF.



Example of beam intensity signal showing different types of anomalies (a-d): from subtle local variations to breakdowns.

Overcoming the Limitations of Isolation Forest

Three main research directions were identified in the literature.

Algorithm Improvement	Pre-processing	Post-processing
• Kim et al. (2024) • Xu et al. (2023) • Xiang et al. (2023) • Song et al. (2022) • Chater et al. (2022) • Mensi & Bicego (2021) • Hariri et al. (2021) • Lesouple et al. (2021) • Ding & Xing (2020) • Stripling et al. (2018)	• Chen et al. (2020) • Khan et al. (2019) • Puggini & McLoone (2018)	• Carletti et al. (2023) • Alsini et al. (2021) • Aminanto et al. (2020)

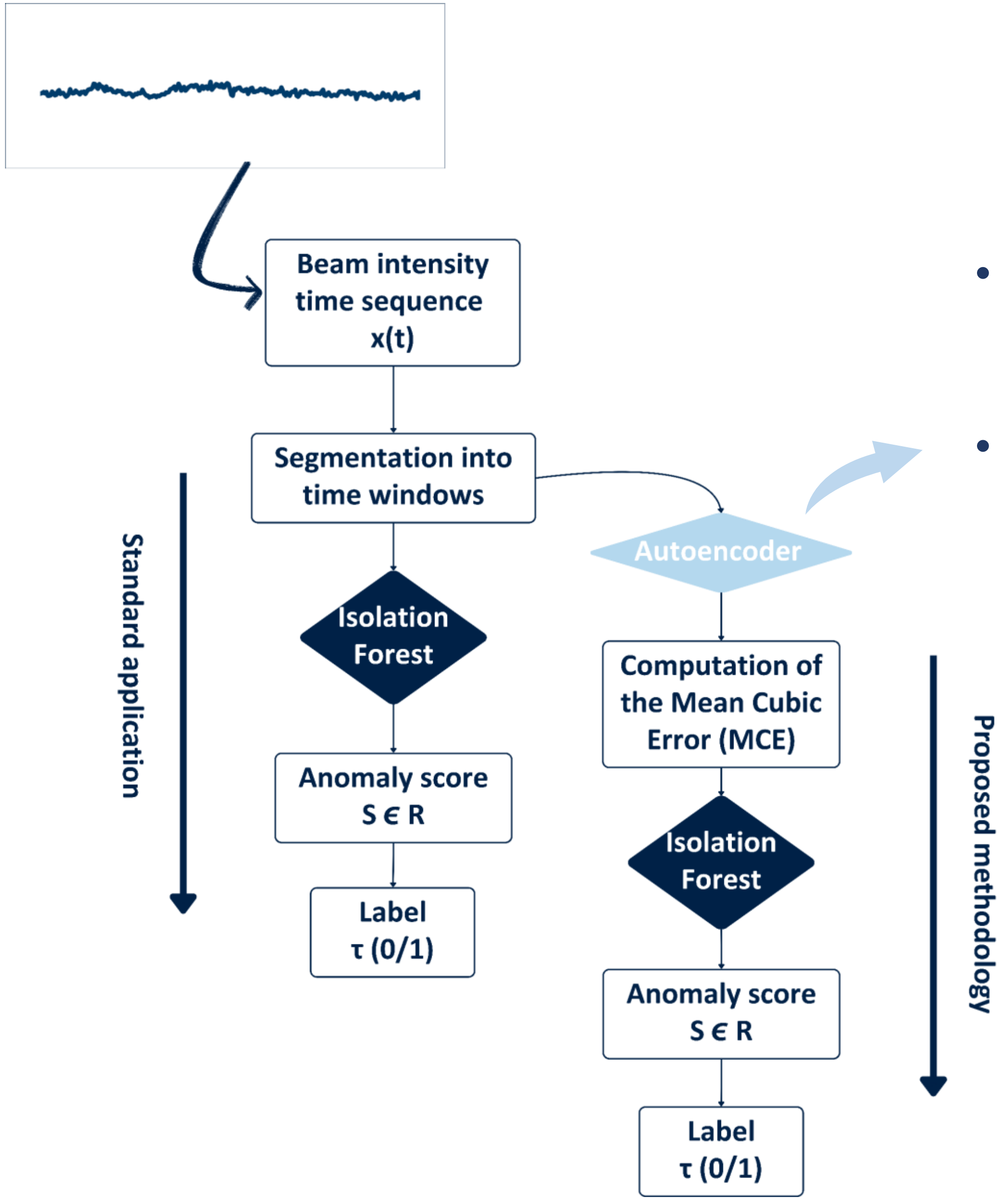
Identified Gap

- Few studies explore the transformation of the feature space.
- Data representation strongly influences the performance of Isolation Forest.

Proposition

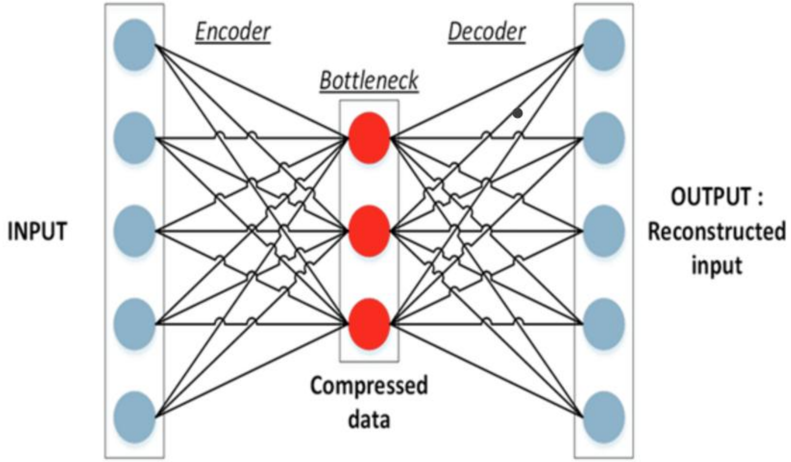
- Hybrid Autoencoder-Isolation Forest model

Proposed Methodology



Comparison between the standard Isolation Forest application and the proposed AE–IF methodology.

- Unsupervised neural networks designed to reconstruct their input.
- Learn to compress data into a lower-dimensional latent space and then decode it.



www.analyticssindiamag.com

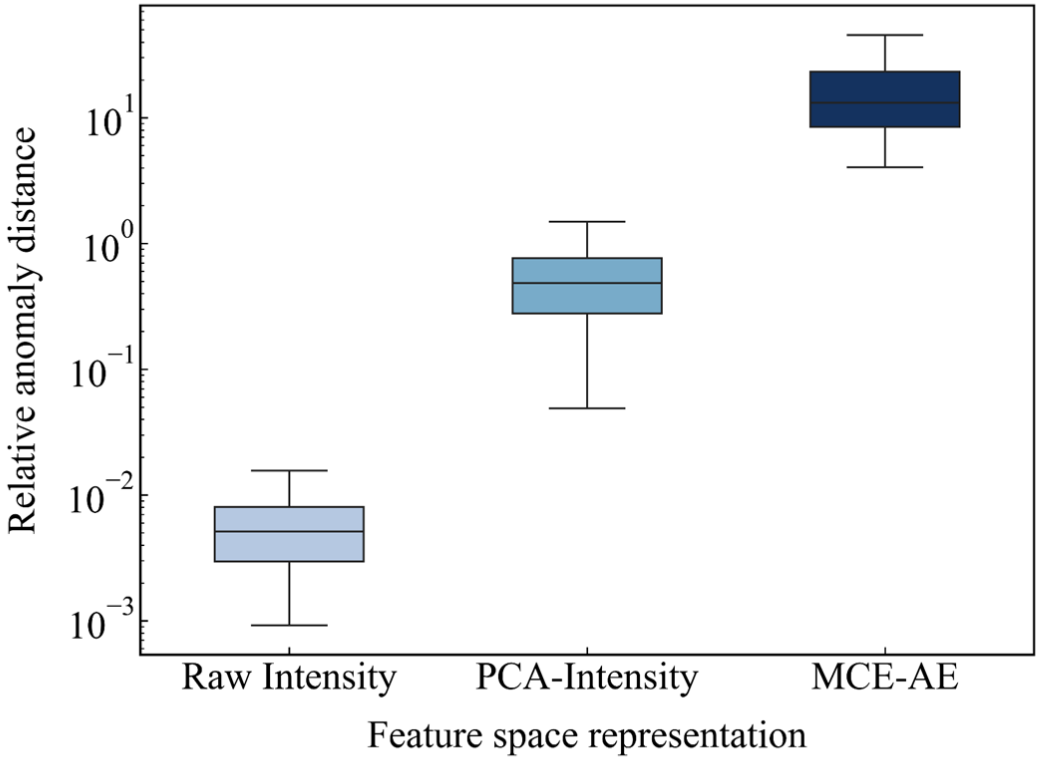
The basic structure of an autoencoder includes an encoder, a bottleneck and a decoder.

Reconstruction Error Space (MCE)

- Low values: Normal sequences
- High values: Global and local anomalies

- We compared three approaches: IF on the raw space, IF with Principal Component Analysis (PCA), and AE-IF on reconstruction errors, using 25 proton beam intensity datasets (7-10 days each), split 60% for development and 40% for final evaluation.

- Subtle anomaly distances from normal means were evaluated in three spaces: nearly indistinguishable in raw intensity (5.13×10^{-3}), better separated with PCA (4.87×10^{-1}), and clearly distinct with MCE-AE (1.32×10^{1}).



Anomaly separability across representation spaces.

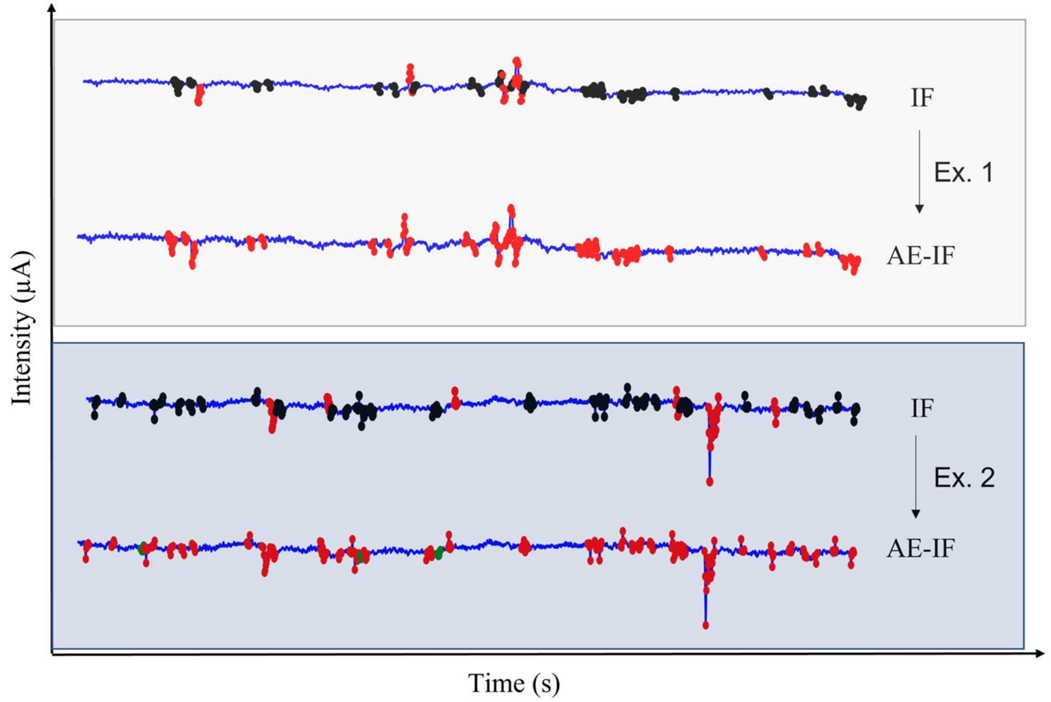
Average performance on test data (10 samples)

Method	Recall	Precision	F1-score	AUC–PR
IF	0.47	0.45	0.41	0.49
PCA–IF	0.67	0.63	0.61	0.74
AE–IF	**0.81**	**0.94**	**0.87**	**0.87**

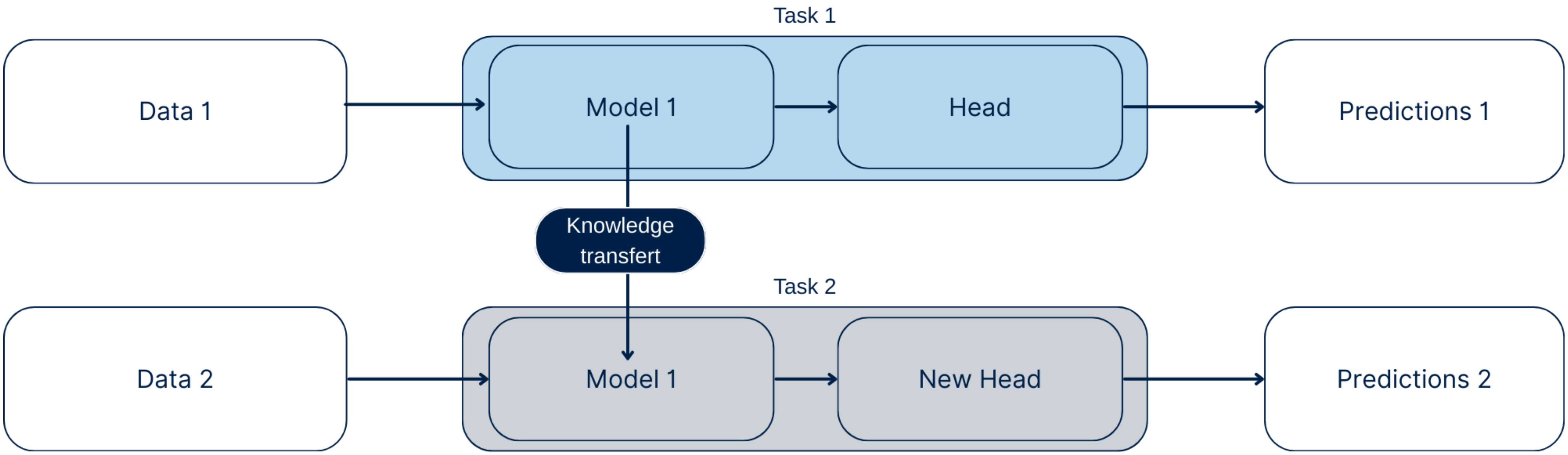
1. AE-IF outperforms the other approaches, combining high recall (0.81) and strong precision (0.94), resulting in the best F1-score (0.87) and AUC-PR (0.87).

2. A qualitative analysis on the temporal intensity data compares the standard IF application with the hybrid AE-IF model.

- The results show that the AE-IF model detects subtle anomalies missed by IF while maintaining correct classification of normal signals.

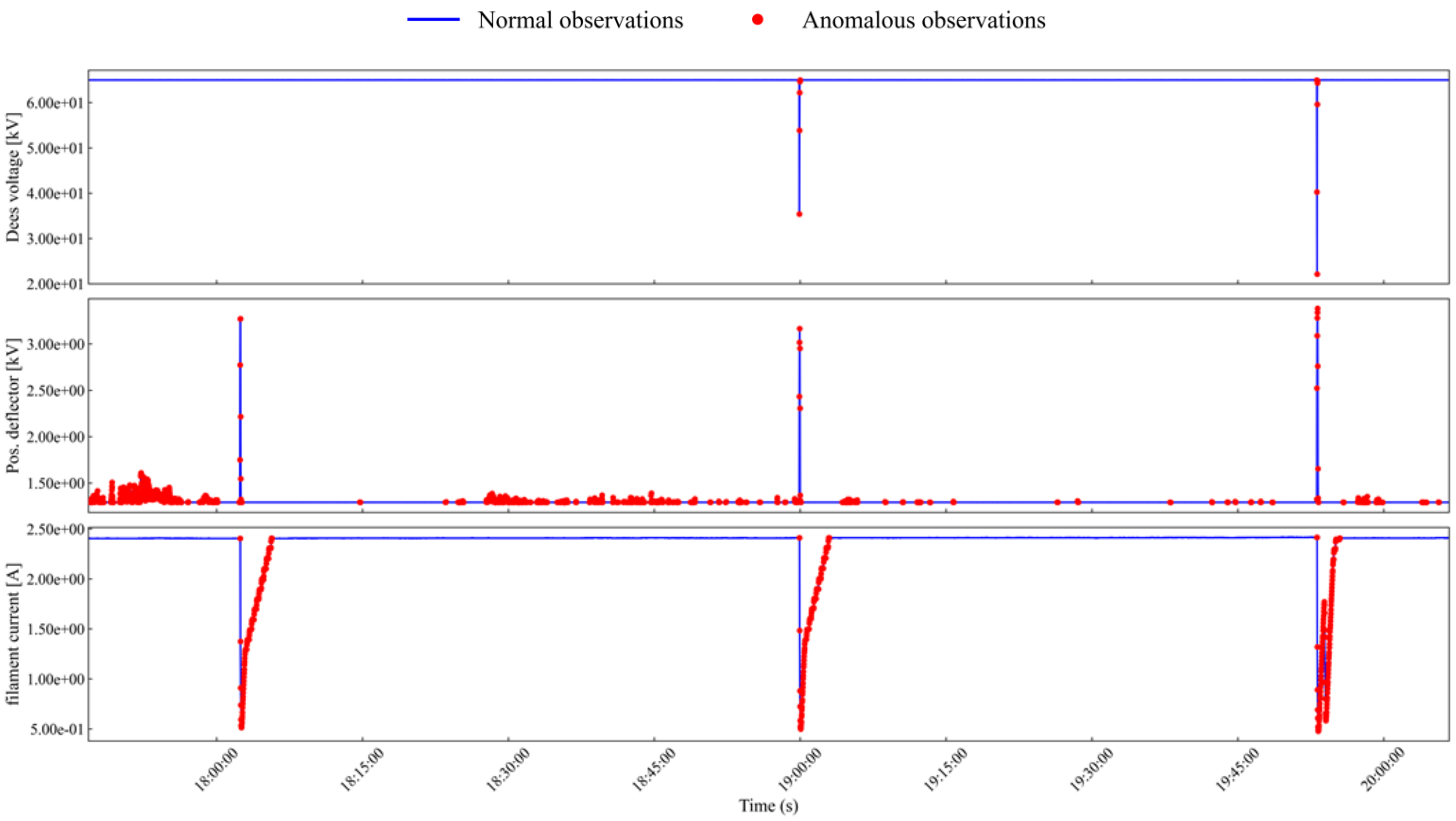


Detection comparison between IF and AE-IF.

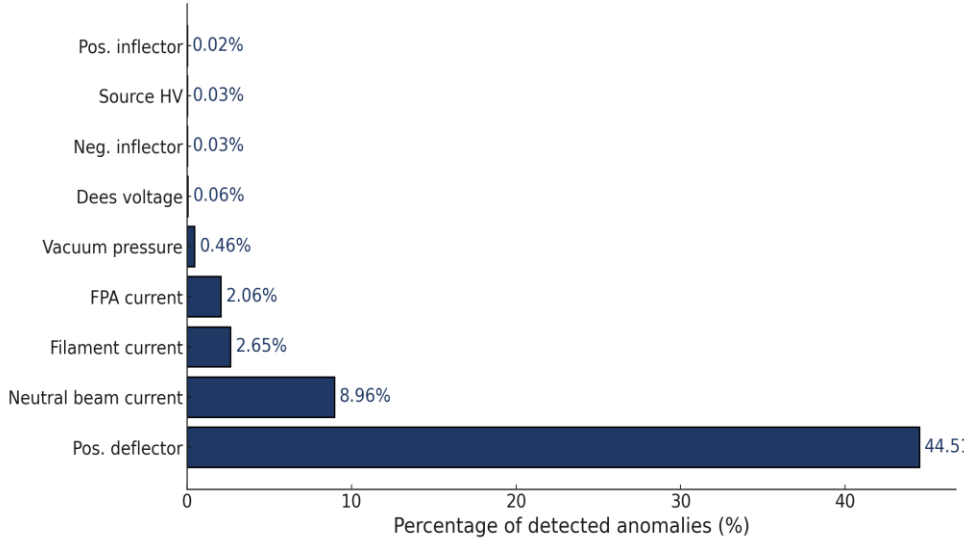


Architecture for transfer learning with head replacement

Preliminary Results of Transfer Learning



Time series visualization of selected PVs with detected anomalies



Anomaly detection percentage per PV

- Reusing a unified model through transfer learning reduces the optimisation effort, and early results show improved anomaly detection performance that generalises well across different types of PV signals.

AE-IF Contribution

- Significantly improves anomaly detection.
- Effective for both global and subtle (local) anomalies.

Transfer Learning

- Allows model reuse across different process variables.
- Reduces training and optimization effort.
- Improves generalization across heterogeneous signals.

Perspectives & Applications

- Stabilization of irradiation processes.
- Reduction of irradiation time for radionuclide production.
- Opens new possibilities for flash therapy.

Thank You!