EUROPEAN AI FOR
FUNDAMENTAL PHYSICS
CONFERENCE
EuCAIFCon 2025

Contribution ID: **151**   Type: **Poster** + **Flashtalk**

Background Enrichment to improve Anomaly Detection

Model-independent anomaly detection for Beyond the Standard Model (BSM) searches in high-energy physics faces significant challenges due to the lack of tractable methods to build rich background priors as well as inherent uncertainties in simulated background processes. Traditional unsupervised ML approaches to anomaly detection, commonly train models on background samples produced by a single physics generator (e.g., Pythia) using a fixed generator tune. This comes with the risk of overfitting to generator-specific features thereby increasing sensitivity to non-anomalous processes that deviate from the limited background representations. To address this, we present a novel method that enhances background modelling by aggregating samples from multiple generators (Pythia, Herwig, Sherpa) and generator tunes, capturing a broader spectrum of possible background variations. We train an unsupervised variational autoencoder (VAE) augmented with contrastive learning objectives, which enforce separation between latent space clusters corresponding to each generator and tune. This enriched background representation ensures that generator-specific features are encoded into the distinct clusters, while anomalous signals —which do not align with any generator's characteristics —are projected outside these regions. The resulting reduction in false positives thereby improves anomaly detection performance. We compare performance across different anomaly metrics, different VAE-based architectures like (but not restricted to) Normalizing Flow augmented VAEs to arrive at a broader picture of model-agnostic, VAE-based anomaly detection for new physics searches. We also attempt an empirical study of the use of contrastive methods to build small foundation models for broader Physics tasks. We present our work through an open-source python package-style repository called BEAD, which is designed to be a modular VAE-based anomaly detection toolkit.

AI keywords

VAE; Contrastive Methods; Anomaly Detection; Normalizing Flows

Primary authors: DOGLIONI, Caterina (University of Manchester); KAR, Deepak (University of Witwatersrand); JAWAHAR, Pratik; SINHA, Sukanya (University of Witwatersrand)

Presenter: JAWAHAR, Pratik

Track Classification: Patterns & Anomalies