EUROPEAN AI FOR
FUNDAMENTAL PHYSICS
CONFERENCE
EuCAIFCon 2025

Contribution ID: **58**   Type: **Parallel talk**

Transformer-based compression of big data

The storage, transmission and processing of data is a major challenge across many fields of physics and industry. Traditional generic data compression techniques are lossless, but are limited in performance and require additional computation.

BALER [1,2] is an open-source autoencoder-based framework for the development of tailored lossy data compression models suitable for data from multiple disciplines. BALER models can also be used in FPGAs to compress live data from detectors or other sources, potentially allowing for massive increases in network throughput.

This presentation will introduce BALER and discuss recent developments and results. These include the development and analysis of new transformer-based autoencoder models, the application of BALER to particle physics analyses and the resulting affect on discovery significance, and the evaluation of the energy consumption and sustainability of differing autoencoder compression models.

BALER is developed by a cross-disciplinary team of physicists, engineers, computer scientists and industry professionals, and has received substantial contributions from a large number of master's and doctoral students. BALER has received support from industry both in providing datasets to develop BALER, and to transfer industry best practices.

[1] https://arxiv.org/pdf/2305.02283.pdf
[2] https://github.com/baler-collaboration/baler

AI keywords

compression; sustainability; transformers; autoencoders; big data

Primary author: SMITH, James (University of Manchester)

Presenter: SMITH, James (University of Manchester)

Track Classification: Datasets & Ethics