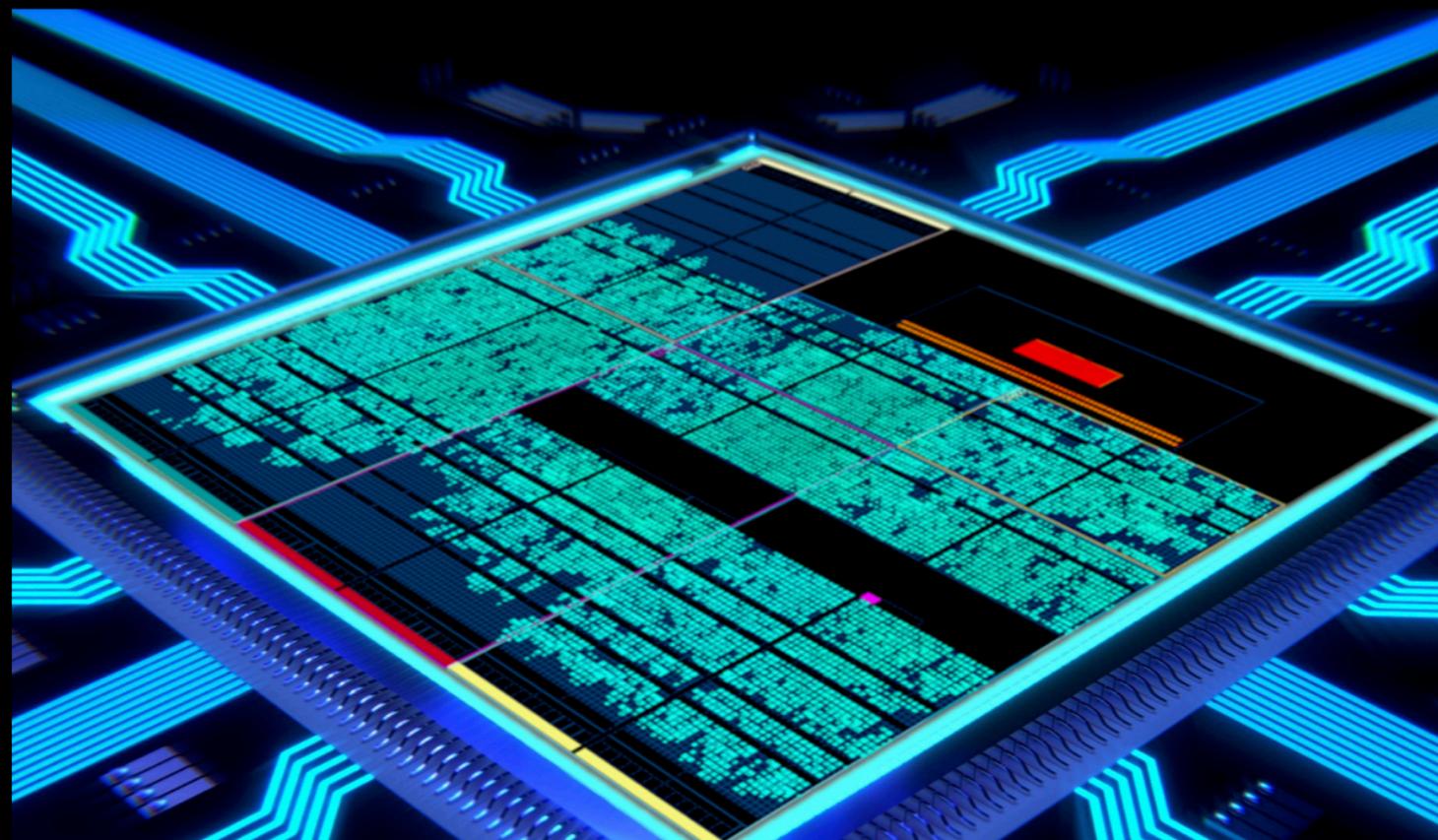


Exploring Data Challenges and Leveraging Codabench

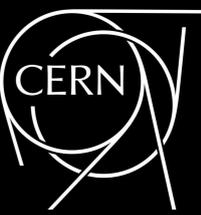
A Practical Journey With Unsupervised New Physics Detection at 40 MHz



Katya Govorkova
MIT

29 November 2023
AISSAI

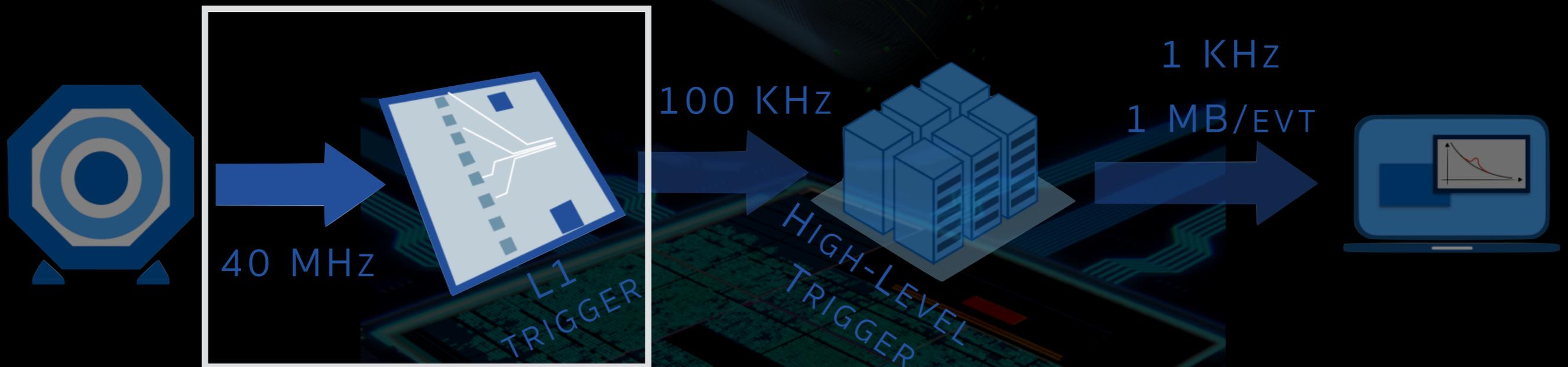
UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



CHALLENGE

Idea is to look for something **very rare and unusual** directly in the **Level-1 Trigger** without any signal hypothesis in mind

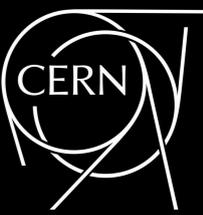
The **challenge** is to find **a-priori unknown** and **rare New Physics** hidden in a data sample dominated by ordinary Standard Model processes



The **deliverable** is a developed **algorithm** that can be deployed and run in L1 with strict **latency** requirement of **< 1 microsecond**

The **task** is therefore to design an architecture that maximises the **sensitivity for New Physics** but at the **lowest possible resource and latency** budget

UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



STEP-BY-STEP

- ▶ **Gather** in a team/by yourself
- ▶ **Get a cool name** for your team, for example team " **DeepAnomaly** " 🧑🧑
- ▶ Get yourself familiar with the **details on the challenge webpage**
- ▶ **Investigate** available **datasets** and **example codes**

SM Dataset: Training

BSM Dataset: A to 4 leptons

BSM Dataset: h^+ to $\tau\nu$

Blackbox Dataset

BSM Dataset: LQ to $b\tau$

BSM Dataset: h^0 to $\tau\tau$

- ▶ **Design** your AD model
- ▶ **Evaluate performance** and submit **results**
- ▶ **Best models** will be **published** in a White Paper (and perhaps deployed in L1 trigger of CMS!!)

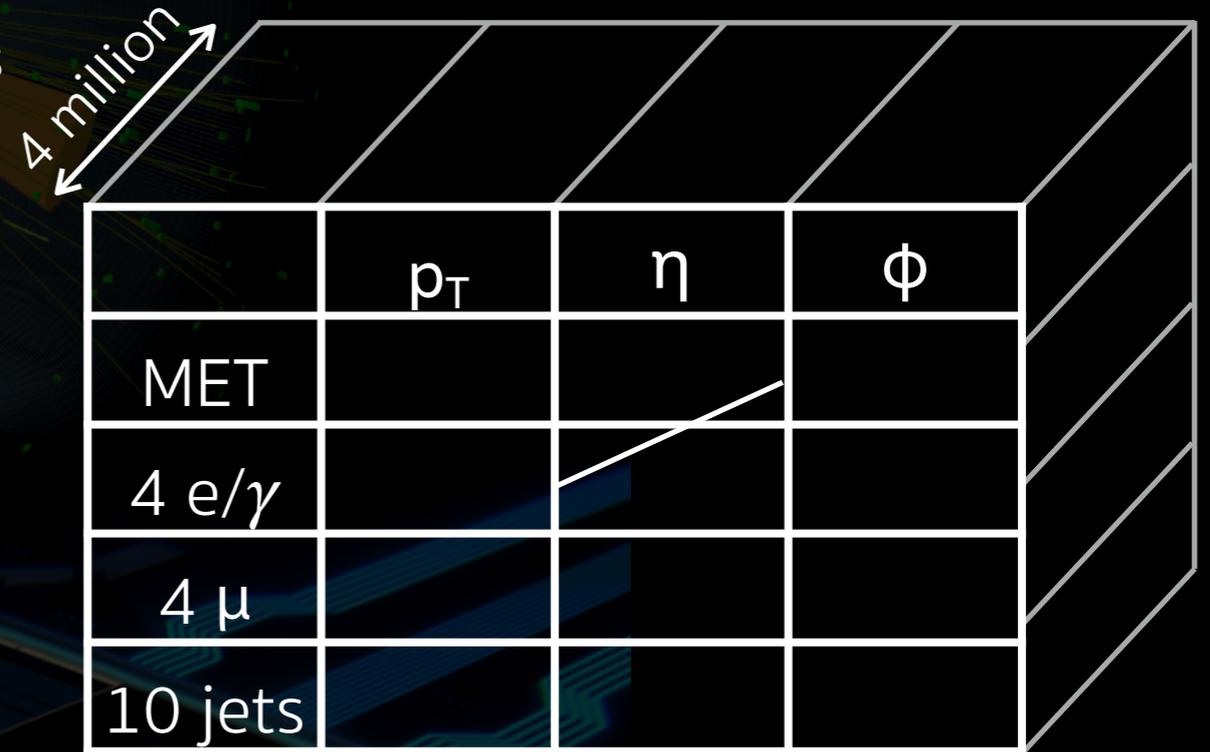
DATA SAMPLES

The data is represented as an array of MET, up to 4 e/γ , 4 μ and 10 jets each described by p_T , η and ϕ to mimic L1 data format

Train using provided **4 million** background-like events simulated with Delphes 

Events are **pre-filtered** to have **at least one lepton**

- ▶ Inclusive W production, with $W \rightarrow l\nu$ (59.2%)
- ▶ Inclusive Z production, with $Z \rightarrow ll$ (6.7%)
- ▶ tt production (0.3%)
- ▶ QCD multijet production (33.8%)

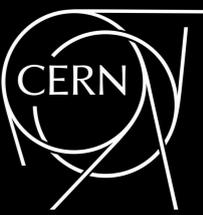


Evaluate performance on several different **New Physics simulated samples**

- ▶ Neutral scalar boson A, 50 GeV \rightarrow 4 l 
- ▶ Leptoquark, 80 GeV \rightarrow b τ 
- ▶ Scalar boson, 60 GeV \rightarrow $\tau\tau$ 
- ▶ Charged scalar boson, 60 GeV \rightarrow $\tau\nu$ 
- ▶ Black Box 

WE CHOOSE A LOW-
MASS, RARE SIGNALS THAT WOULD
PASS PRE-FELTIRING

OVERVIEW AND TIMELINE OF THE CHALLENGE



2020 — we had a dataset that represents data in the L1 trigger of CMS and we started working on an algorithm for the L1 trigger

February 2021 — developed an unsupervised Anomaly Detection (AD) algorithm
The idea was to release the datasets and set up a public challenge based on it

July 2021 — released [the datasets paper](#) on arXiv

July 2021 — published the datasets in Zenodo

July 2021 — set up a challenge using GitHub io <https://mpp-hep.github.io/ADC2021/>

July 2021 — announced the challenge at the [ML4Jets2021](#)

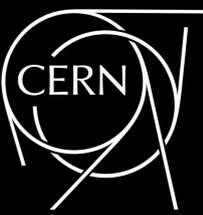
August 2021 — released [the AD algorithm paper](#) on arXiv

February 2022 — published the AD algorithm paper in nature Machine Intelligence

March 2022 — published the datasets paper in nature Scientific Data

2022/2023 — the data has been used to publish other research, but not to participate in the challenge

CHALLENGE SETUP

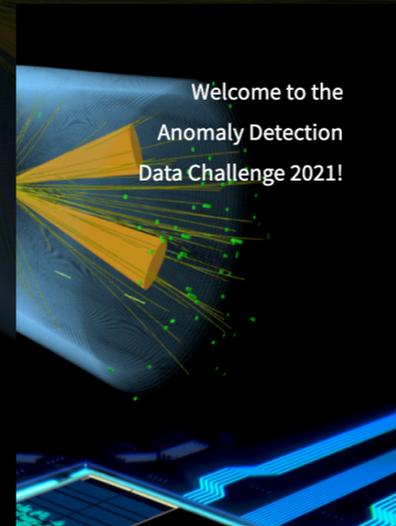


We provided

○ A website with all the details <https://mpp-hep.github.io/ADC2021/>

○ Datasets

1. Signal $h^+ \rightarrow \tau \nu$
2. Signal $h^0 \rightarrow \tau \tau$
3. Signal $LQ \rightarrow b \tau$
4. Signal $A \rightarrow 4 \text{ leptons}$
5. Training dataset
6. BlackBox dataset



Unsupervised New Physics detection at 40 MHz

In this challenge, you will develop algorithms for detecting New Physics by reformulating the problem as an out-of-distribution detection task. Armed with four-vectors of the highest-momentum jets, electrons, and muons produced in a LHC collision event, together with the missing transverse energy (missing E_T), the goal is to find a-priori unknown and rare New Physics hidden in a data sample dominated by ordinary Standard Model processes, using anomaly detection approaches.

The algorithms are intended to be deployed in the first stage of the real-time event filter processing system of LHC experiments (Level 1 or L1 trigger), where the available bandwidth, latency and resources are strictly limited. Such limitations constrain the design of the algorithm. To emulate the constraints in terms of bandwidth only the leading 10 jets, 4 muons, 4 electrons and the missing E_T will be provided to be used as input to the algorithm. Furthermore, only a maximum of X, Y, and Z bits are available for the representation of the η , ϕ , and the transverse momentum p_T of each physics object, respectively.

○ A paper with details about the data

○ Example algorithms

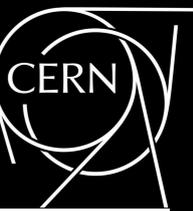
○ Example analysis pipeline

○ A script to estimate FLOPS

○ A repo to upload the contributions to <https://github.com/mpp-hep/ADC2021-results>

○ “Further reading” mostly about hls4ml and AD

UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



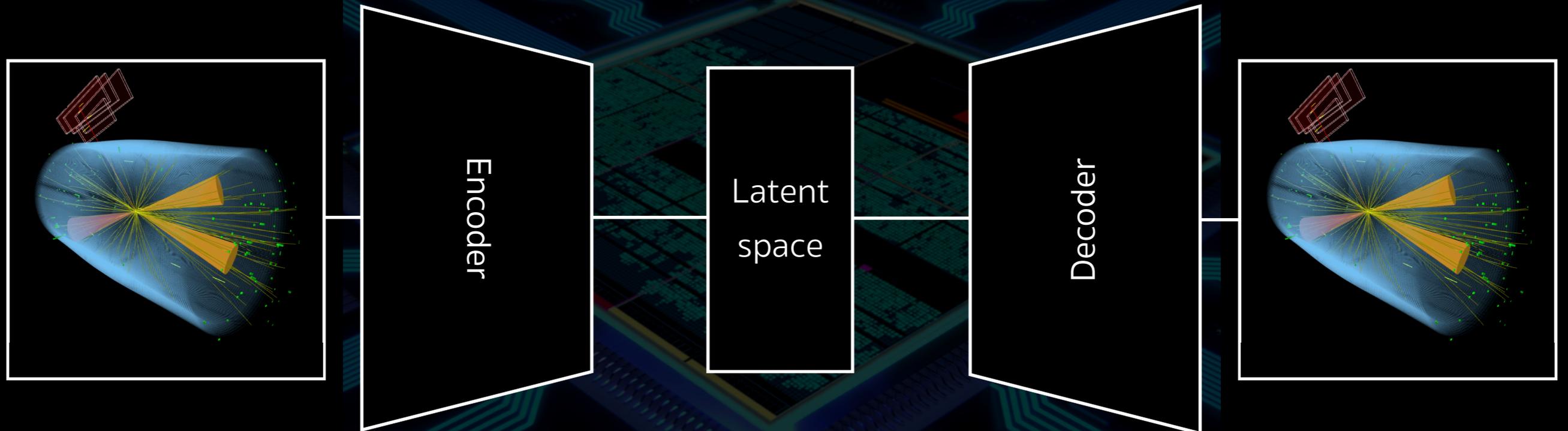
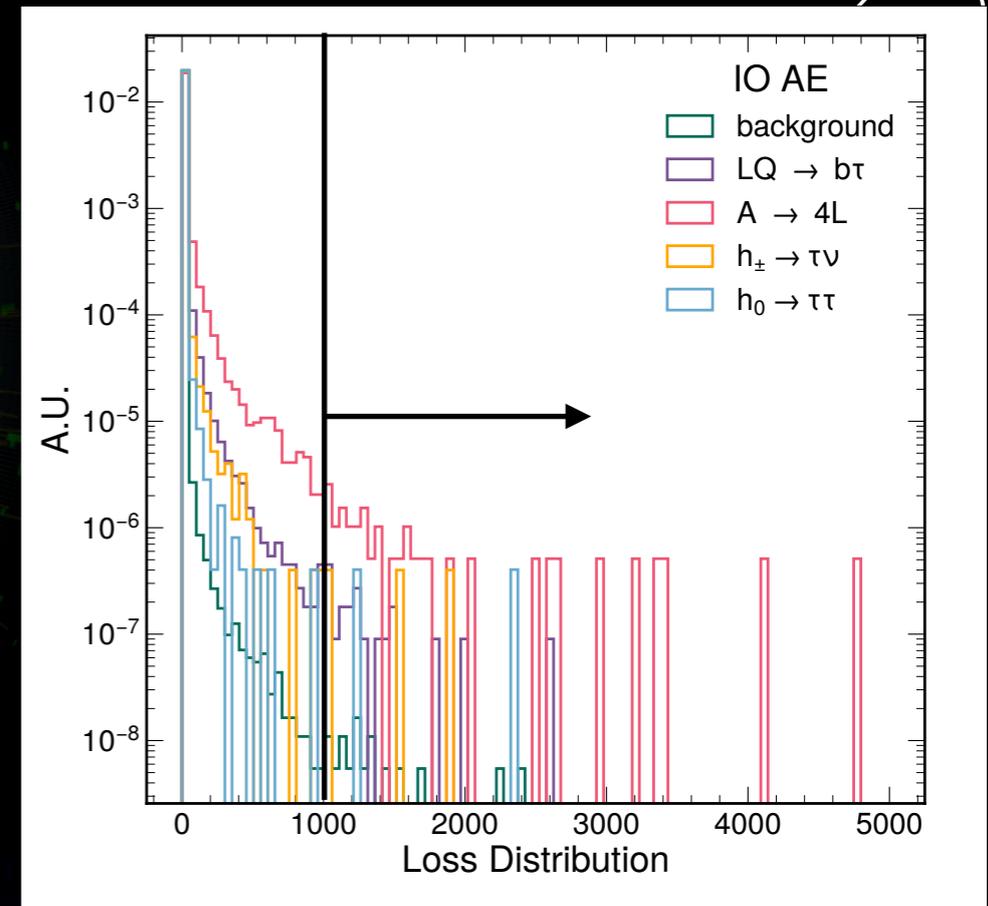
EXAMPLE TEAM "DEEPANOMALY" 🧑🧑🧒

DESIGNS ALGORITHM

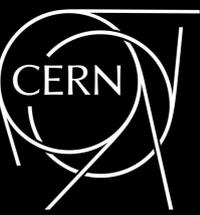
Decide on the **algorithm** that you want to explore

The Example Team "DeepAnomaly" has chosen **Autoencoders**

- ▶ Encode input in **smaller** dimensional space
- ▶ Train on typical LHC **background**
- ▶ **Anomalous** data will have **higher loss**
- ▶ **Calculating** the loss requires to **store the input** until the output is computed



UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



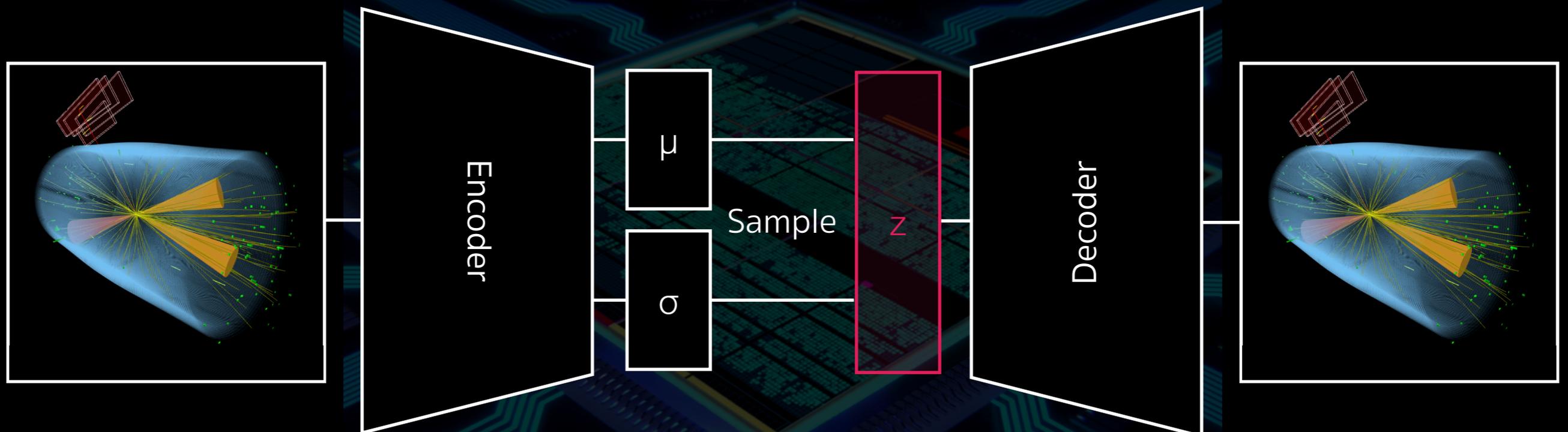
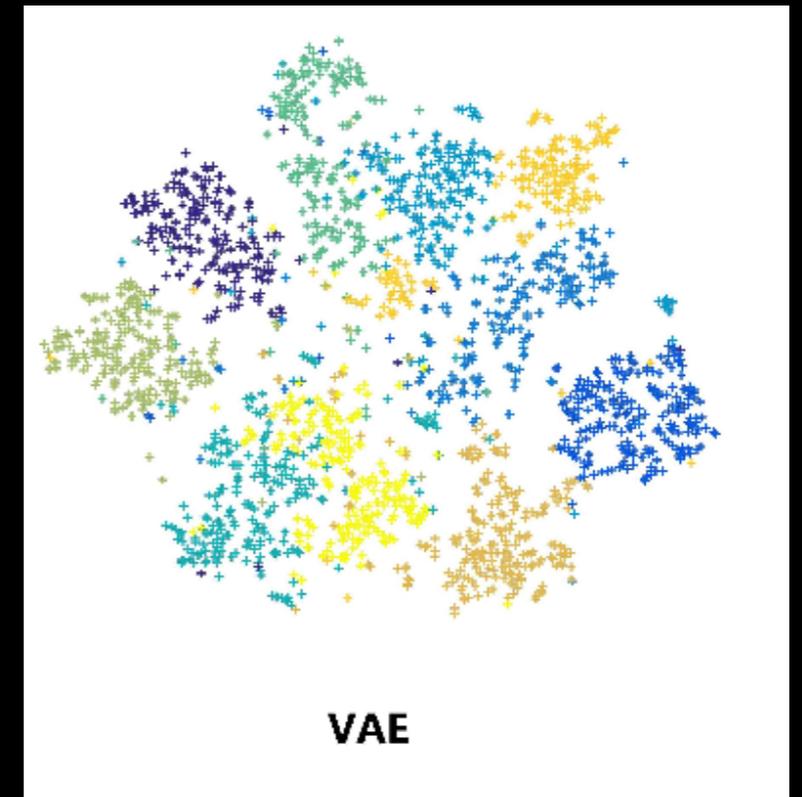
EXAMPLE TEAM "DEEPANOMALY" 🧑🧑🧒

DESIGNS ALGORITHM

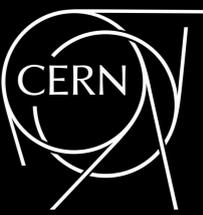
The "DeepAnomaly" team has also considers Variational Autoencoders

- ▶ The latent space is sampled from Encoder output
- ▶ Can be used to generate new samples
- ▶ Inference can be done only on the latent space
- ▶ No need to store input and deployment of Encoder is enough

(e.g. saves resources and latency in comparison to AE)



UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



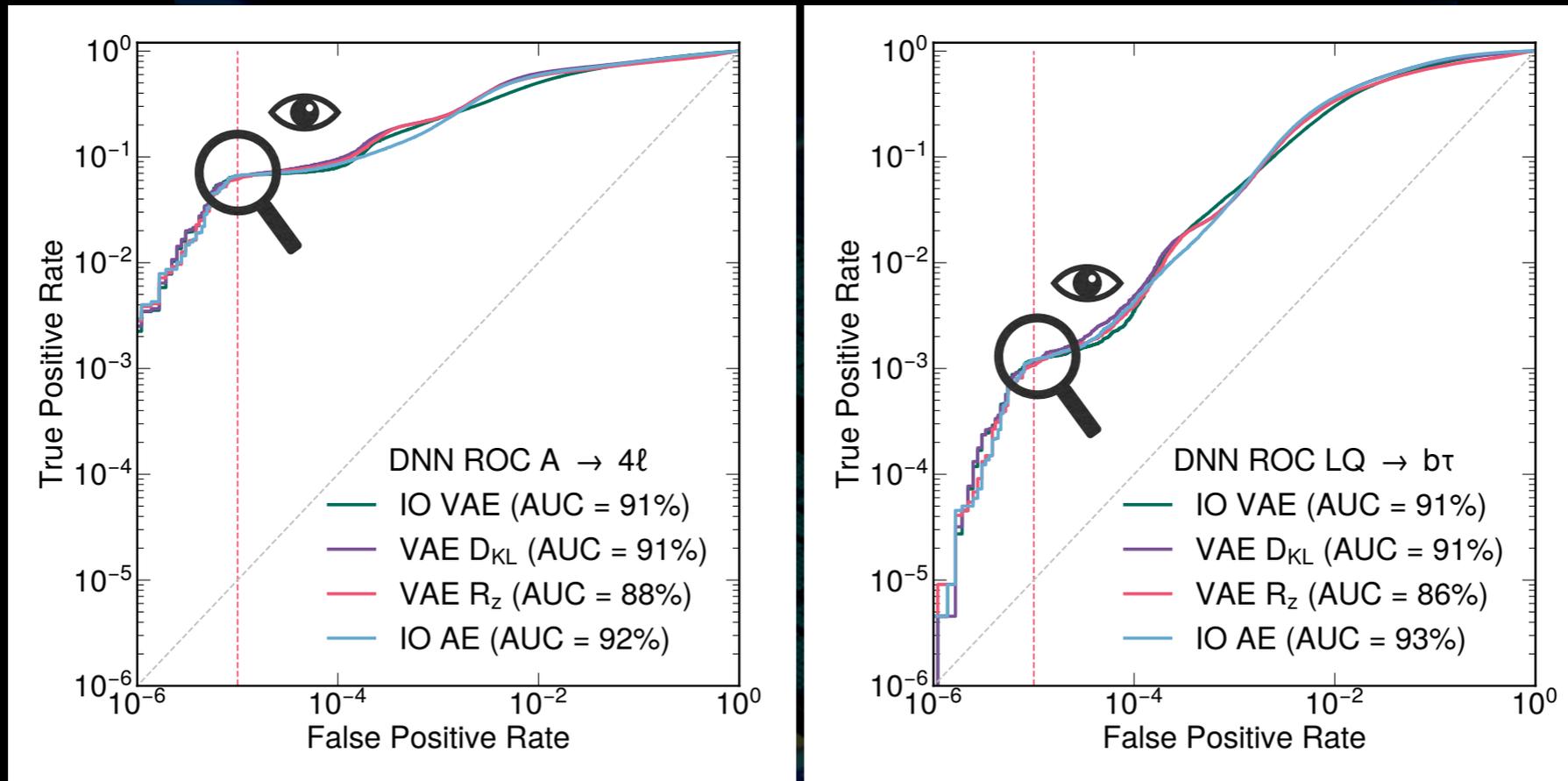
EXAMPLE TEAM "DEEPANOMALY" 🧑🧑🧒

EVALUATES PERFORMANCE ON SIGNAL SAMPLES

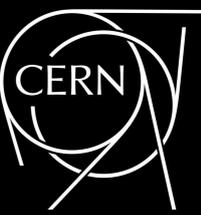
Goal is to maximise TPR at FPR 10^{-5} (roughly corresponding to the available output data rate budget for a trigger algorithm) for each of the provided anomaly

The Team "DeepAnomaly" checks AE vs VAE

- ▶ The **Inference** can be done only on the latent space, either with D_{KL} or R_z
Shoutout to DarkMachines for the R_z idea
- ▶ No need to store input and deployment of **Encoder** is enough



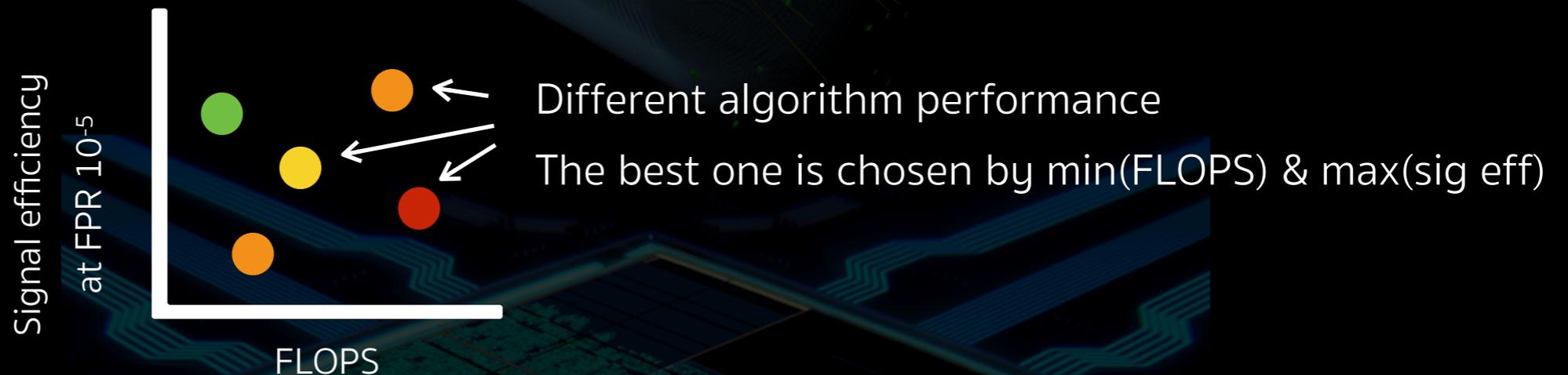
UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



EXAMPLE TEAM "DEEPANOMALY" 🧑🧑🧒

SUBMITS RESULTS

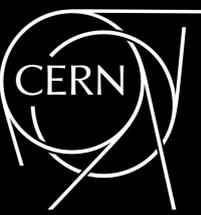
- ▶ For the **Blackbox** dataset get the event number of the **1000 most anomalous** events based on the algorithm metrics
- ▶ An estimate of the **algorithm efficiency** can be obtained by calculating the **floating-point operations per second (FLOPS)** [🔗](#)



- ▶ **The submission** should be in a form of a HDF5 file, [DeepAnomaly.h5](#), containing a **numpy array** with the **identification numbers** of each selected event, plus a dictionary with the algorithm deployment **performance** [🔗](#)

Upload contribution!

UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



EXAMPLE TEAM "DEEPANOMALY" 🧑🧑🧒

SUBMITS RESULTS

- ▶ For the **Blackbox** dataset get the event number of the **1000 most anomalous** events based on the algorithm metrics

- ▶ An estimate of the **algorithm efficiency** in terms of **point operations per second**

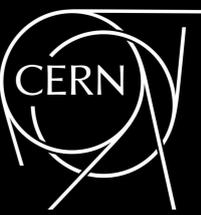


WE ONLY ACCEPTED THE FINAL SUBMISSIONS, WHILE MANY WERE ASKING US FOR AN INTERMEDIATE RESULT TO ITERATIVELY IMPROVE THE ALGORITHM

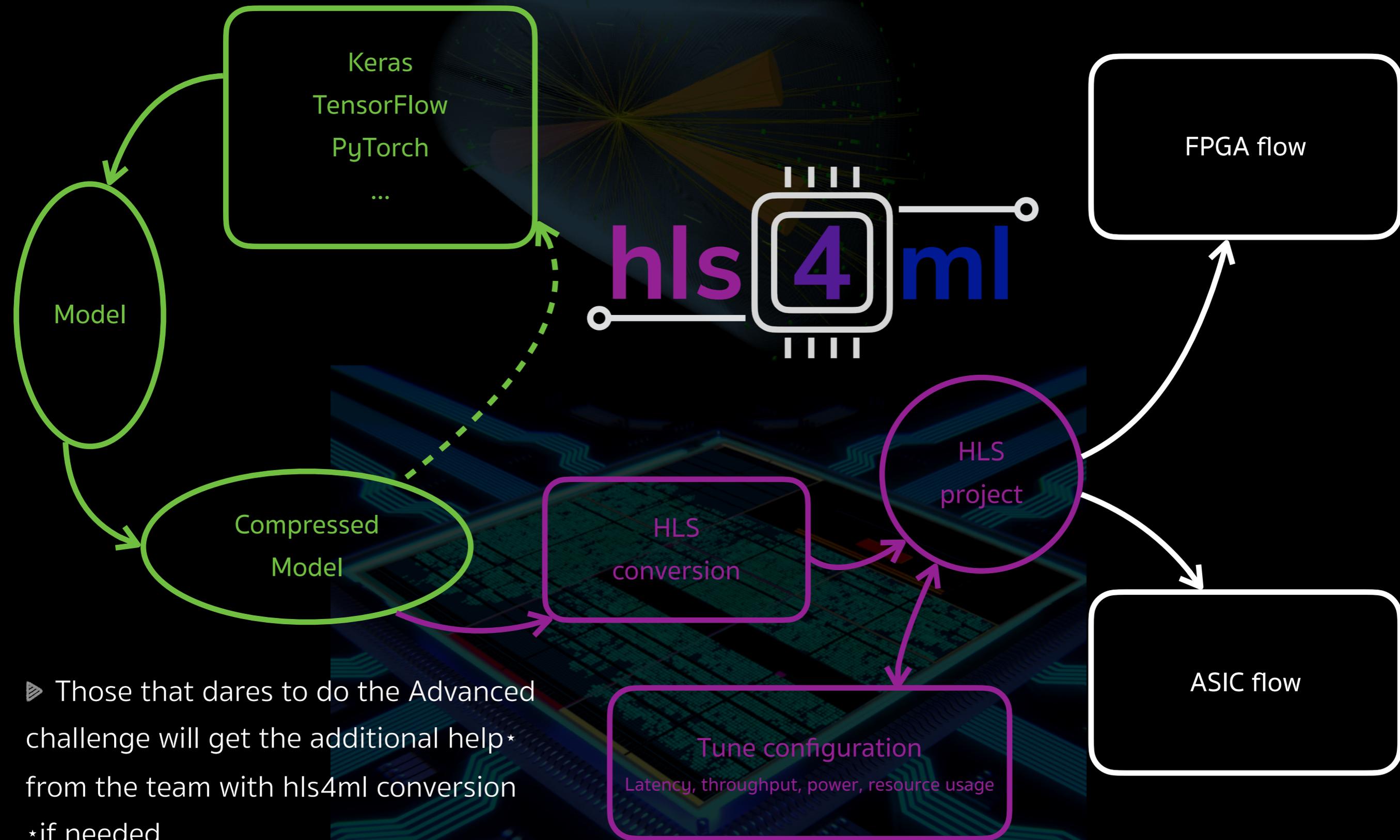
- ▶ The **submission** should be in a form of **numpy array** with the **identification numbers** and a **dictionary** with the algorithm deployment **performance** [🔗](#)

Upload contribution!

UNSUPERVISED NEW PHYSICS DETECTION AT 40 MHz



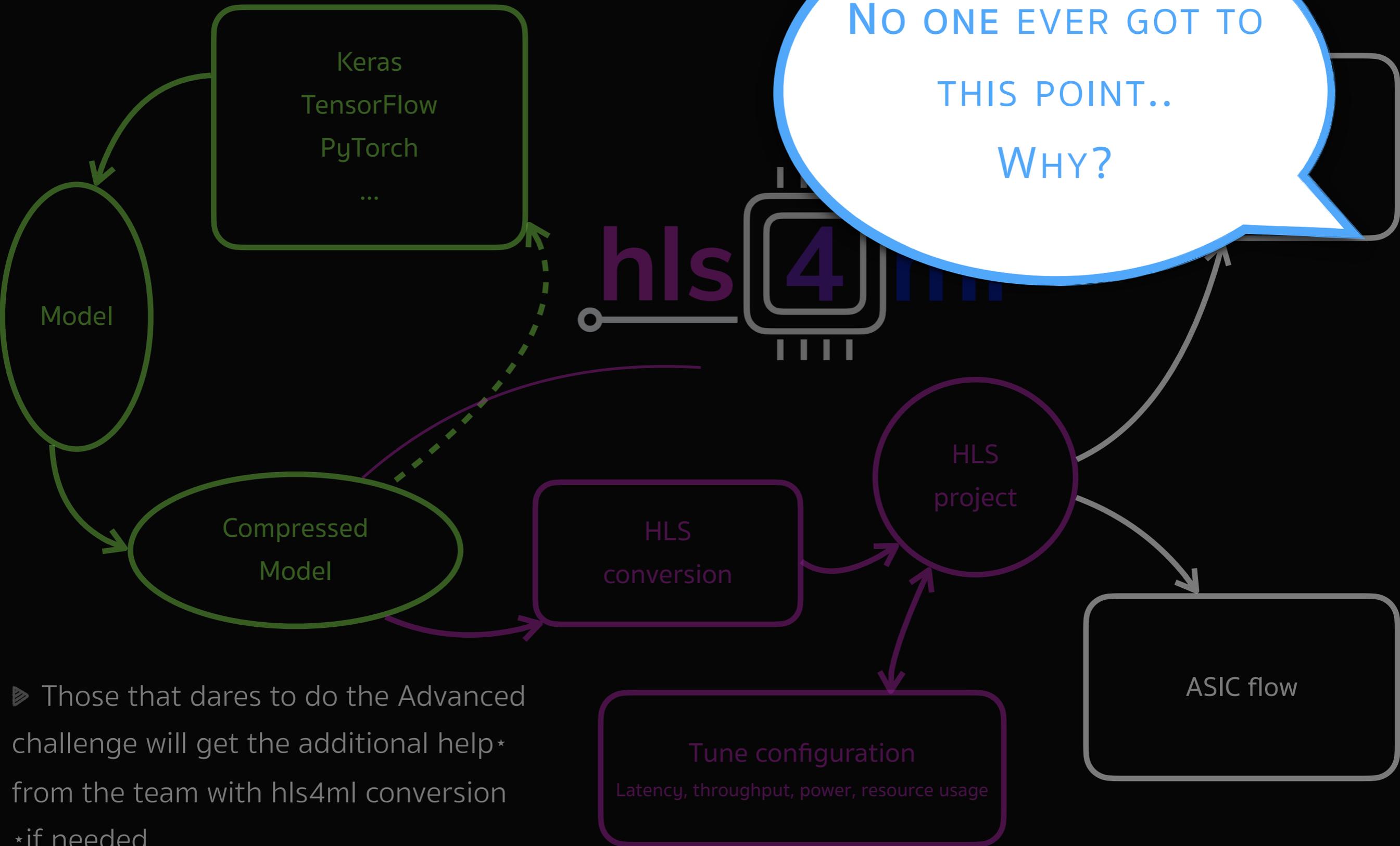
★ ADVANCED CHALLENGE



► Those that dares to do the Advanced challenge will get the additional help* from the team with hls4ml conversion
*if needed

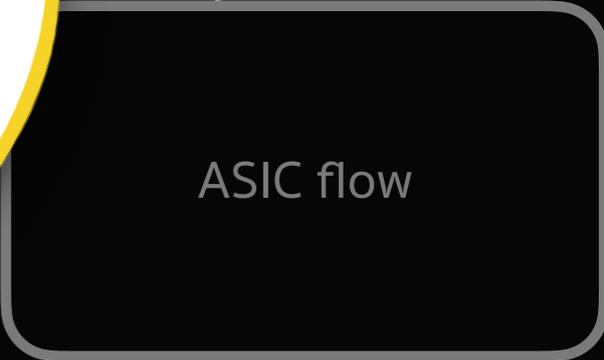
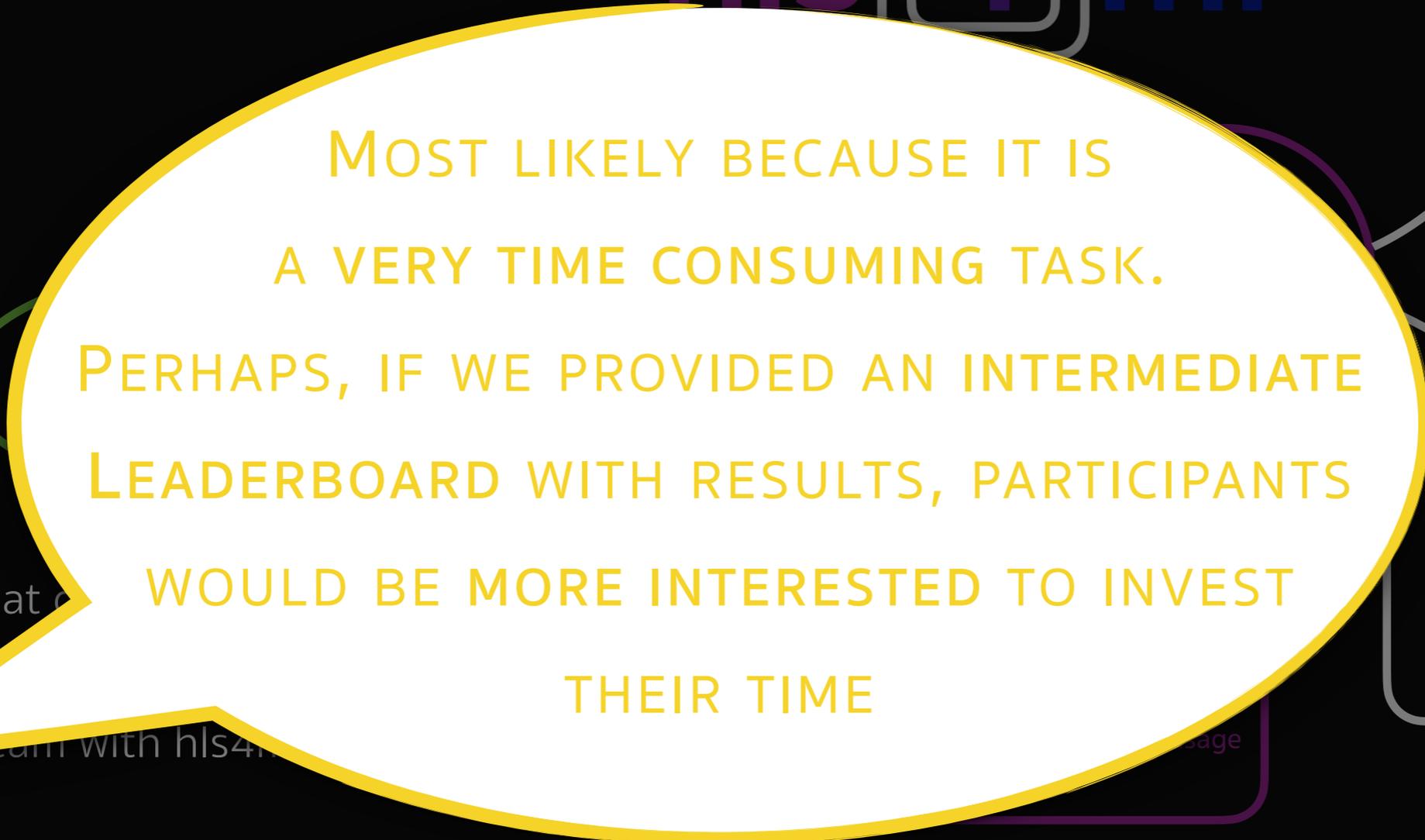
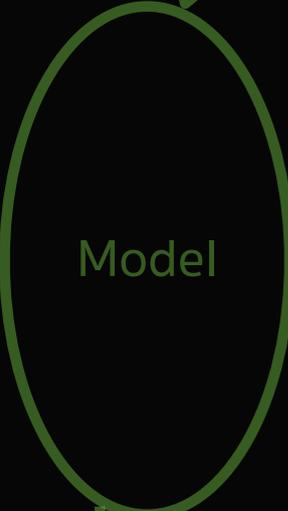
★ ADVANCED CHALLENGE

NO ONE EVER GOT TO THIS POINT.. WHY?



► Those that dares to do the Advanced challenge will get the additional help* from the team with hls4ml conversion
*if needed

★ ADVANCED CHALLENGE



► Those that... challenge... from... team with hls4... *if needed

THERE WAS ONLY ONE OFFICIAL
SUBMISSIONS IN THE GITHUB REPO
BUT SEVERAL PAPERS PUBLISHED USING
THE DATASET WE PROVIDED

ANOMALIES, REPRESENTATIONS, AND SELF-SUPERVISION

BARRY M. DILLON, LUIGI FAVARO, FRIEDRICH FEIDEN, TANMOY MODAK, TILMAN PLEHN

ONLINE-COMPATIBLE UNSUPERVISED NON-RESONANT ANOMALY DETECTION

VINICIUS MIKUNI, BENJAMIN NACHMAN, DAVID SHIH

NANOSECOND ANOMALY DETECTION WITH DECISION TREES FOR HIGH ENERGY PHYSICS AND REAL-TIME APPLICATION TO EXOTIC HIGGS DECAYS

STEPHEN ROCHE, QUINCY BAYER, BENJAMIN CARLSON, WILLIAM OULIGIAN, PAVEL SERHIAYENKA, JOERG STELZER, TAE MIN HONG

A platform to create your own challenges, see [the talk by Wahid](#)



Announcement

Welcome to Codabench!
Join the [Google group](#) to connect with the community!

Popular Benchmarks

	Machine Learning for Physical Simulation Challenge This competition aims at promoting the use of ML based surrogate models to solve physical ... <i>Organized by: systemx</i>	10 October 2023  60 Participants
	Track 1: Pedestrian Attribute Recognition - WACV'24 As a part of the WACV'2024 Pedestrian Attribute Recognition and Person Retrieval Challenge... <i>Organized by: juliojj</i>	9 September 2023 60 Participants
	AutoML Cup Phase 1 AutoML Cup Phase 1 <i>Organized by: automlcup</i>	6 June 2023 50 Participants
	(ended) Auto-Survey Challenge'23 Auto-Survey Challenge'23 <i>Organized by: fnachaleam</i>	7 July 2023  37 Participants

Featured Benchmarks

	RescueNet_VQA- Visual question Answering for Damage Assessment Visual question Answering for Damage Assessment <i>Organized by: binalab</i>	12 October 2023 6 Participants
	Cross-Domain MetaDL Any-Way Any-Shot Learning Competition with Novel Datasets from Practical Domains <i>Organized by: pavao</i>	15 November 2022 10 Participants
	STYLE-TRANS-FAIR STYLE-TRANS-FAIR competition M1[AI] 2022-2023 <i>Organized by: fnachaleam</i>	7 February 2023 7 Participants
	Style Trans-Fair Competition Style Trans-Fair competition M1 [AI] 2022-2023 <i>Organized by: ayoubhammal</i>	28 January 2023 11 Participants

Get Started



Participate

Find benchmarks that pique your interest! A benchmark allows you to test new algorithms against reference datasets OR (inverted benchmark) submit challenging data to reference algorithms.



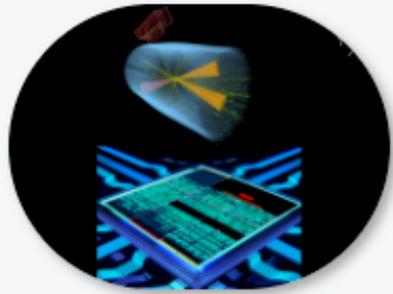
Organize

Organize a benchmark on Codabench. Start with our [tutorial](#).



Contribute

Interested in joining the development team? Join us on [Github](#) or [contact us](#) directly.



CLASSIFY STANDARD MODEL EVENTS IN HEP DATASET

5 PARTICIPANTS

1 SUBMISSIONS

ORGANIZED BY: Katyag
CURRENT PHASE ENDS: Never
CURRENT SERVER TIME: 19 November 2023 At 15:53 CET
Docker image: [codalab/codalab-legacy:gpu](https://codalab.com/codalab-legacy/gpu)

I've made a test challenge to learn how to use Codabench

Things that are very useful:

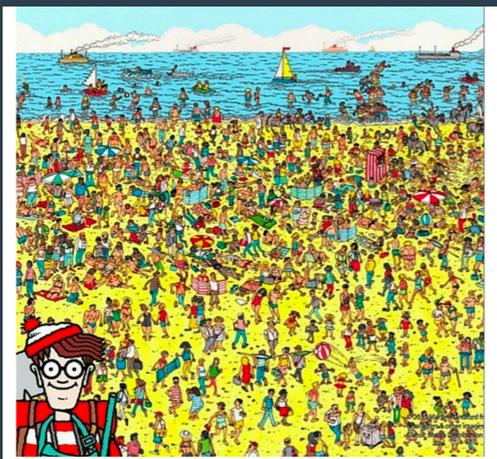
- ✓ Datasets+code examples in one place
- ✓ Clear deadline that is easily updated if need to
- ✓ Easy to reach out and to advertise the challenge to the communit
- ✓ Infinite submissions and automatic check of it
- ✓ Leaderboard
- ✓ Past challenges can be used as educational materials

PERHAPS IF OUR CHALLENGE WAS
HOSTED ON CODABENCH, WE WOULD GET MORE
SUBMISSIONS (AND WE COULD AUTOMATE
FLOPS ESTIMATION)

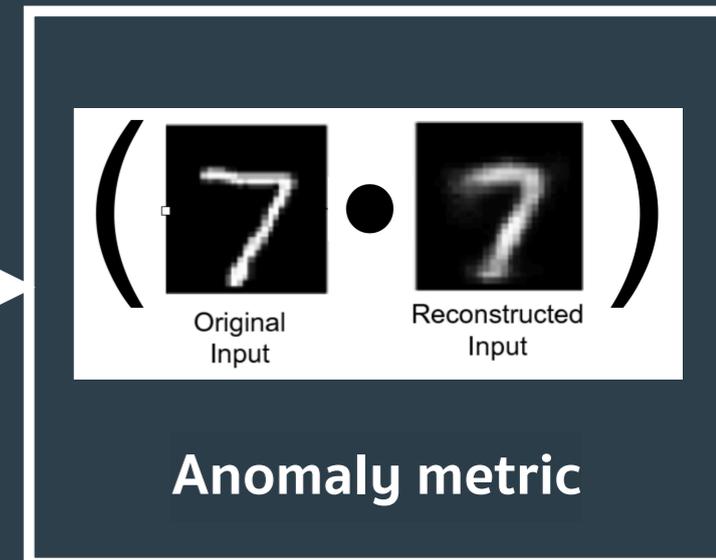
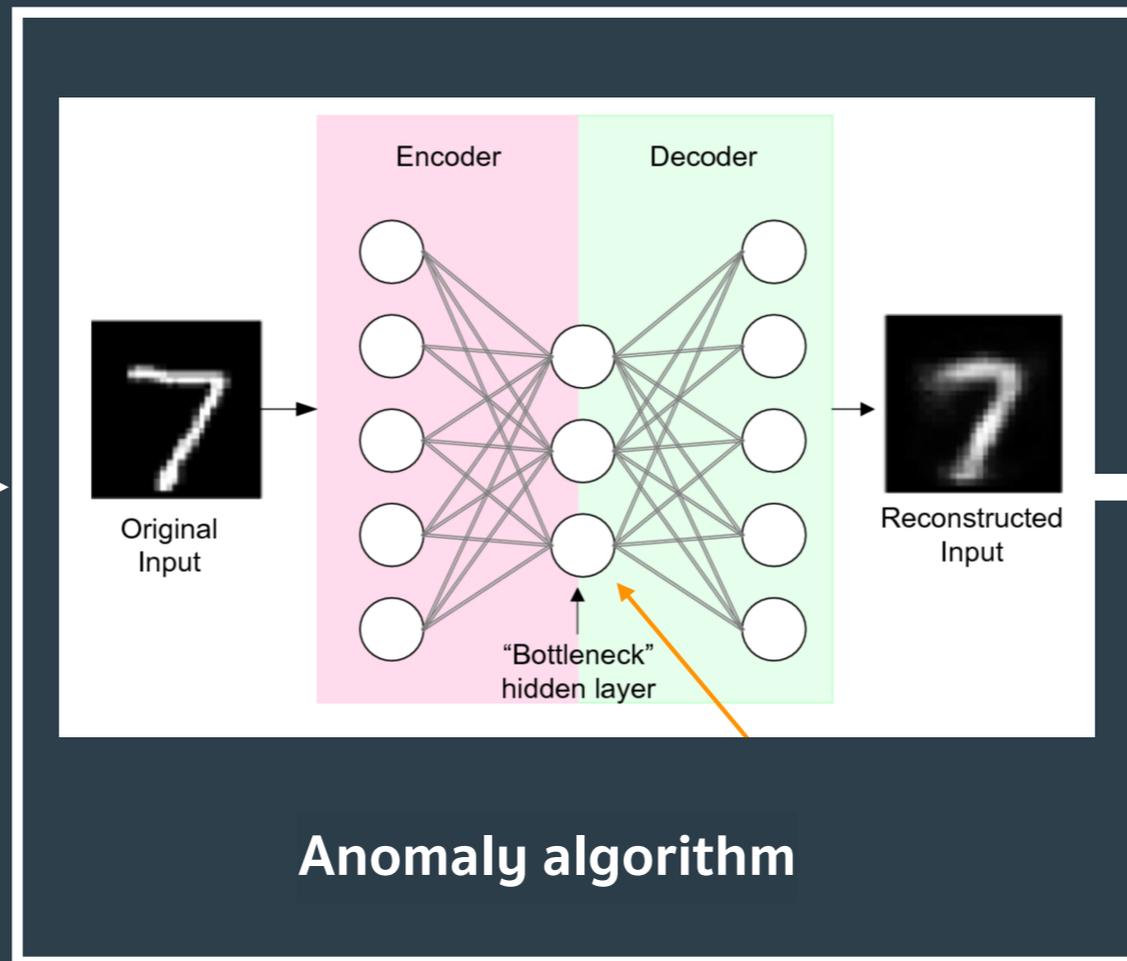
FUTURE ANOMALY DATA CHALLENGES



“Can we make an anomaly detection algorithm capable of anomaly detection in many domains? “



Many different datasets





At A3D3 we are looking to lead a challenge across NSF HDR initiative institutes on anomaly detection and would like to host this on [Codabench.org](https://codabench.org)

Possible Suggested datasets

- ❖ Imageomics: Images of animals for classification and new species
- ❖ Iguide: Rainfall data to predict anomalous local weather
- ❖ iHARP: Ice data from anomalous melts?
- ❖ ID4: new materials identification?
- ❖ A3D3: Astrophysical anomalies (LIGO/...)

AND A PRIZE!! 10K\$  FOR
DEVELOPMENT

The plan is to write two white papers

1. A white paper on making ML Workflows fair not just the dataset, but we want the whole workflow reproducible
2. A second whitepaper comparing and contrasting anomaly detection on the different datasets