MLOps for LHC Hardware Triggers

Physics result → reproducible search → hardware-ready winner

CERN openlab lightning talk

Maksymilian Artur Moskwytyn | Ioannis Xiotidis

83% fewer parameters with comparable anomaly performance

Controlled ablation: identical HF data, seed and training budget; minbias excluded from training

PARAMETERS

185,864 → 31,240

83.2% fewer parameters
in the candidate encoder

UNSEEN-MINBIAS AUC

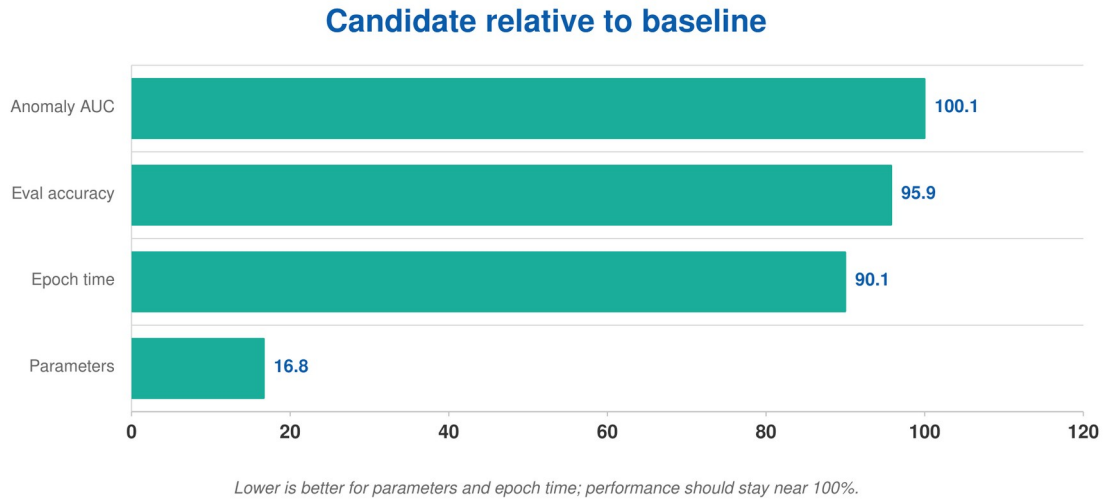
0.9953 → 0.9959

+0.00056; bootstrap interval
includes zero

EFF @99% BG REJ

91.0% → 92.9%

+1.85 pp; bootstrap interval
includes zero



Training control: eight background classes only; minbias is never seen.

Cost: 83.2% fewer parameters and 9.9% faster epochs. Classification accuracy changes by −1.88 pp.

Conclusion: comparable anomaly performance with a much smaller model. This was a controlled ablation—not yet an Optuna winner.

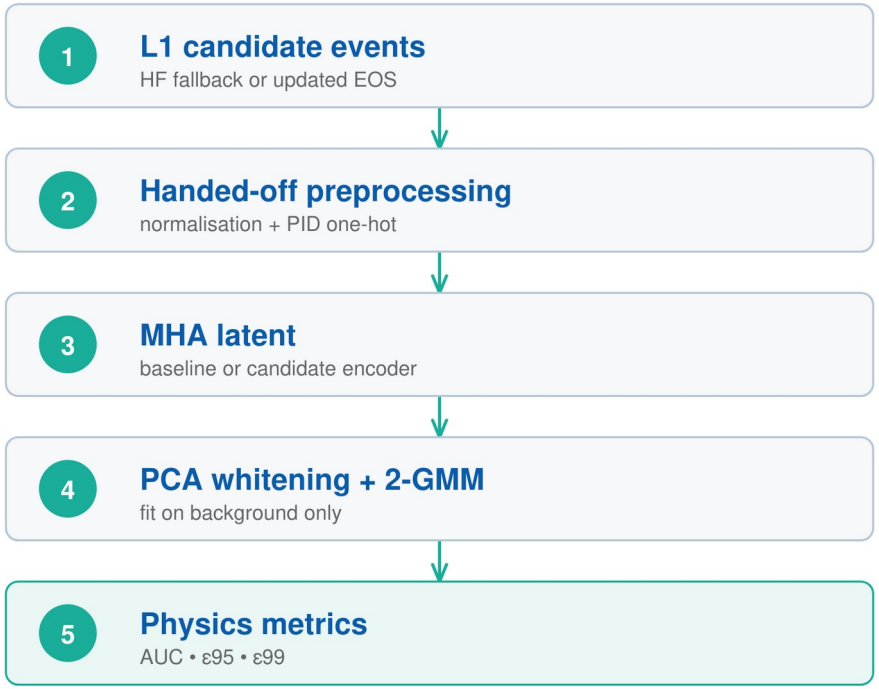
The physics comparison is reproducible end to end

Shared L1 preprocessing, MHA latents and auditable physics outputs

One path now covers the complete comparison:
raw L1 → preprocessing → MHA latent → PCA + 2-GMM → physics metrics.

The Hugging Face dataset enabled rapid development and the controlled baseline–candidate comparison.

EOS access is now available. Historical and corrected class mappings are separated for auditability, and final physics claims will be confirmed on corrected EOS data.



We packaged the experiment as a reusable optimization pipeline

Configuration, Optuna, hardware evidence and physics metrics share one workflow

1 **YAML config** — experiment contract

2 **HGQ2 QAT** — quantization + EBOPs

3 **Optuna studies** — resume + Pareto

4 **Hardware eval** — C-sim + Vitis

5 **Physics validation** — custom objectives

6 **MLflow winner** — 1-trial rerun

SEARCH SCALE

300 trials

resumable studies with configurable winner policies

TEST SUITE

323/323

passing after the physics integration

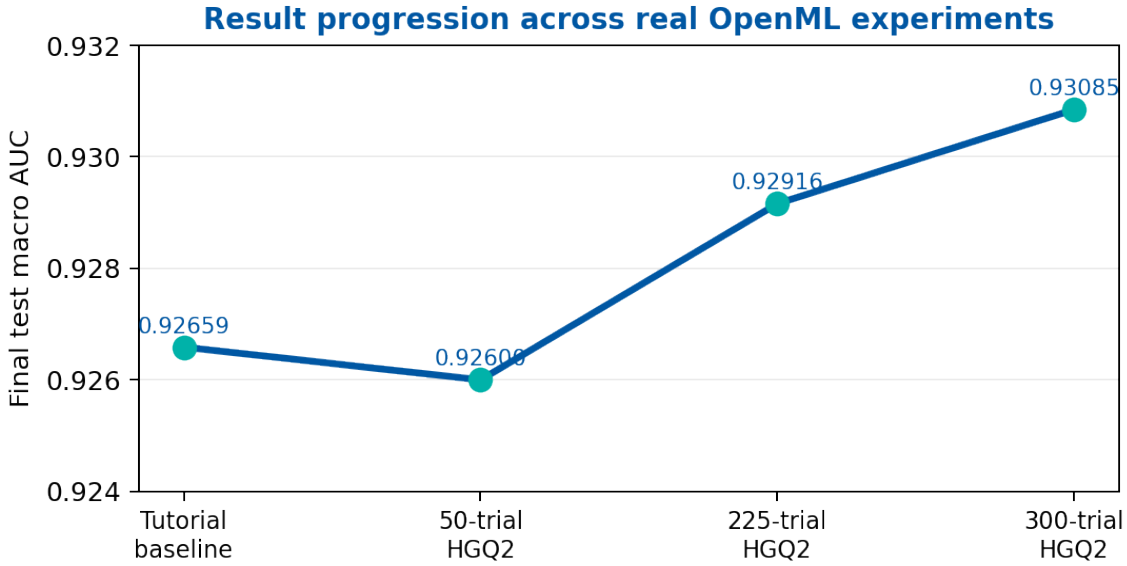
NEW USE CASE

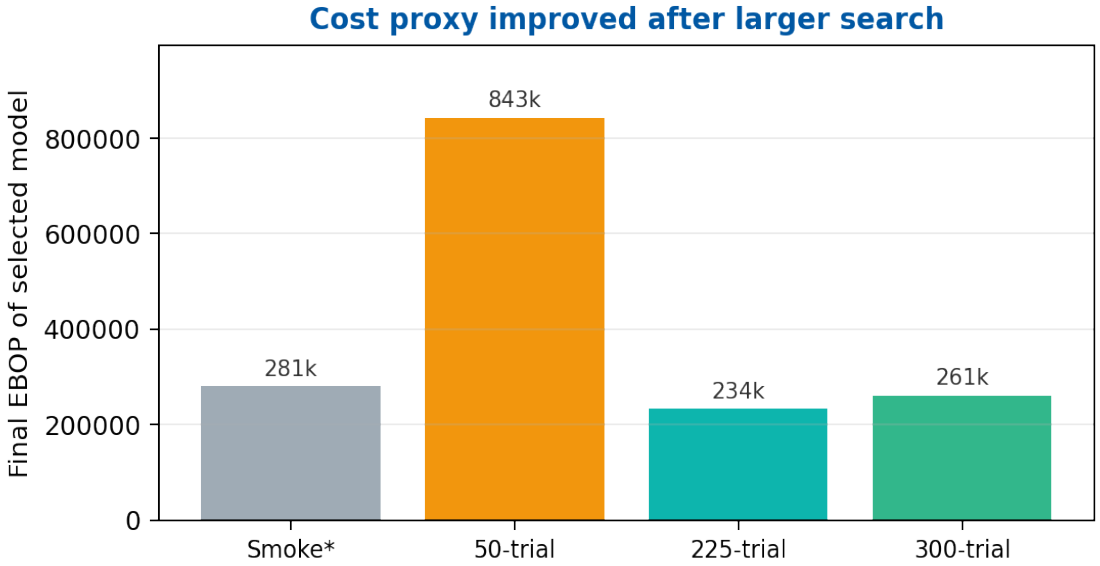
MHA → GMM

unseen-minbias anomaly evaluation

Hardware-aware search scales to 300 trials

OpenML benchmark: larger HGQ2 studies improved the selected trade-off





Macro AUC rose from 0.92659 to 0.93085 across the benchmark.

The 300-trial winner uses ≈261k EBOPs, versus ≈843k for the 50-trial winner.

Pareto export makes the quality–cost trade-off explicit and reproducible.

Next: make physics metrics Optuna objectives

The workflow is ready; systematic physics optimisation closes the loop

1	Config-driven search	Resumable Optuna/HGQ2 studies, multi-objective fronts and configurable winner policies.
2	Hardware evaluation	C-simulation and Vitis latency/resource reports use the same selected winner.
3	Reproducibility	MLflow publication and winner_config.yml support one-trial reruns.
4	Physics workflow	HF baseline/candidate training and unseen-minbias evaluation are reusable.
5	Next validation	Optimize AUC / ε95 / ε99 across trials and seeds, then confirm on corrected EOS data.

MODEL
−83%
parameters vs baseline

PHYSICS
0.9959
unseen-minbias AUC

QUALITY
323
tests passing

Thank you

MLOps for LHC Hardware Triggers • Optuna + HGQ2 + MLflow reproducibility

moskwytyn.maksymilian@gmail.com

Last slide
Last slide