Machine Learning based Global Particle Identification Algorithms at the LHCb Experiment

23rd International Conference on Computing in High Energy and Nuclear Physics

Denis Derkach[1], **Mikhail Hushchyn**[1,2], **Nikita Kazeev**[1,2,3]
on behalf of the LHCb collaboration

[1]National Research University Higher School of Economics
[2]Yandex School of Data Analysis
[3]Università degli Studidi Roma "La Sapienza"

Introduction

Particle identification (PID) plays a crucial role in LHCb analyses. The LHCb PID system is composed of two ring-imaging Cherenkov detectors (RICH), a series of muon chambers and a calorimeter system (ECAL and HCAL). Combining information from these subdetectors allows one to distinguish between various species of long-lived charged particles. Advanced machine learning techniques are employed to obtain the best PID performance and control systematic uncertainties in a data-driven way. This poster covers the major steps of the implementation, and highlights the PID performance achieved in Run 2.

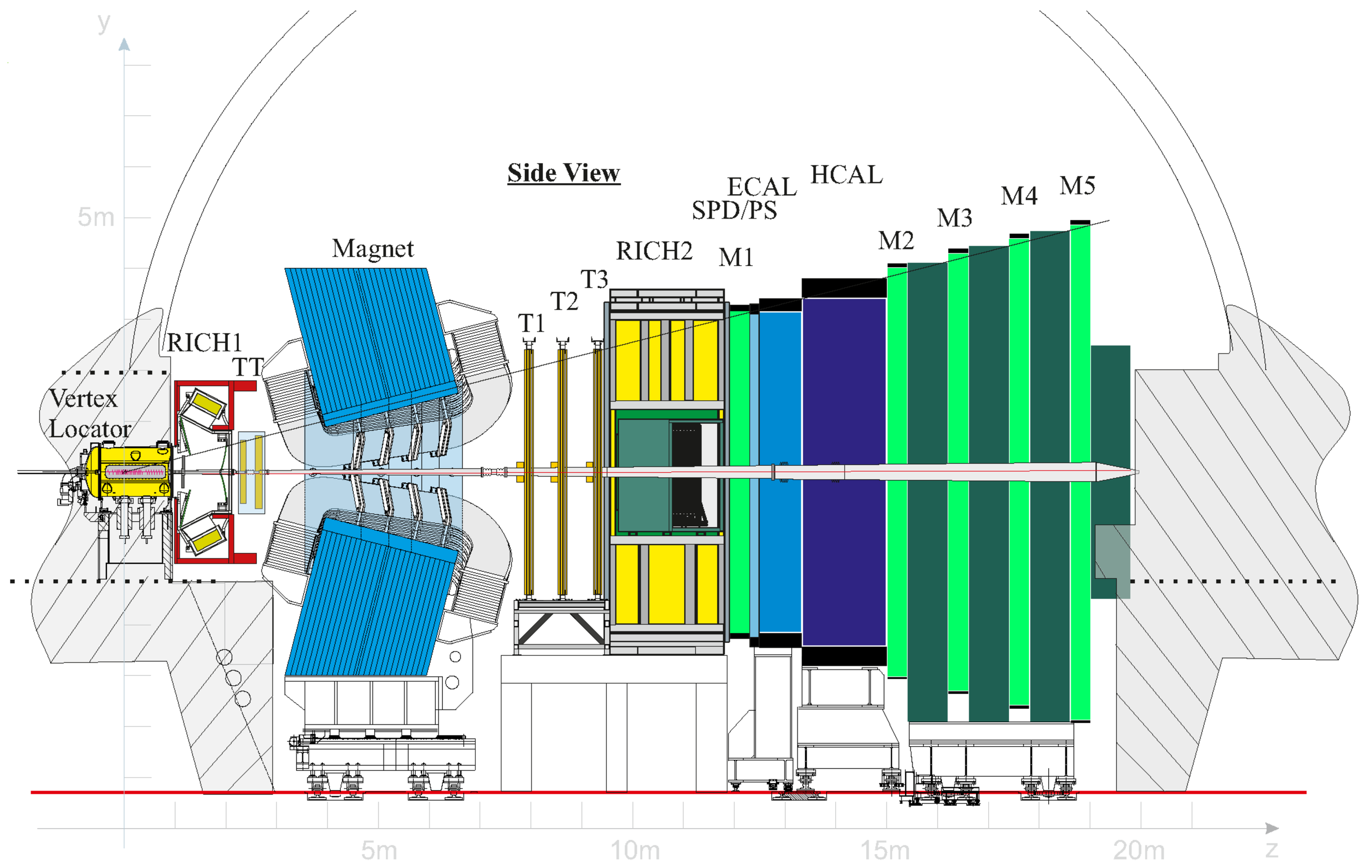

Figure 1: LHCb detector layout. The interaction point is on the left, inside the VELO detector.

Global Particle Identification

Particle identification plays a crucial role in high-energy physics analysis. Global PID at LHCb identifies the charged particle type associated with a given track. There are five particle types: electron, muon, pion, kaon, proton, and ghost track. Ghost tracks are charged tracks that do not correspond to a real particle which passed through the detector. Different particle types have different responses in the LHCb systems:

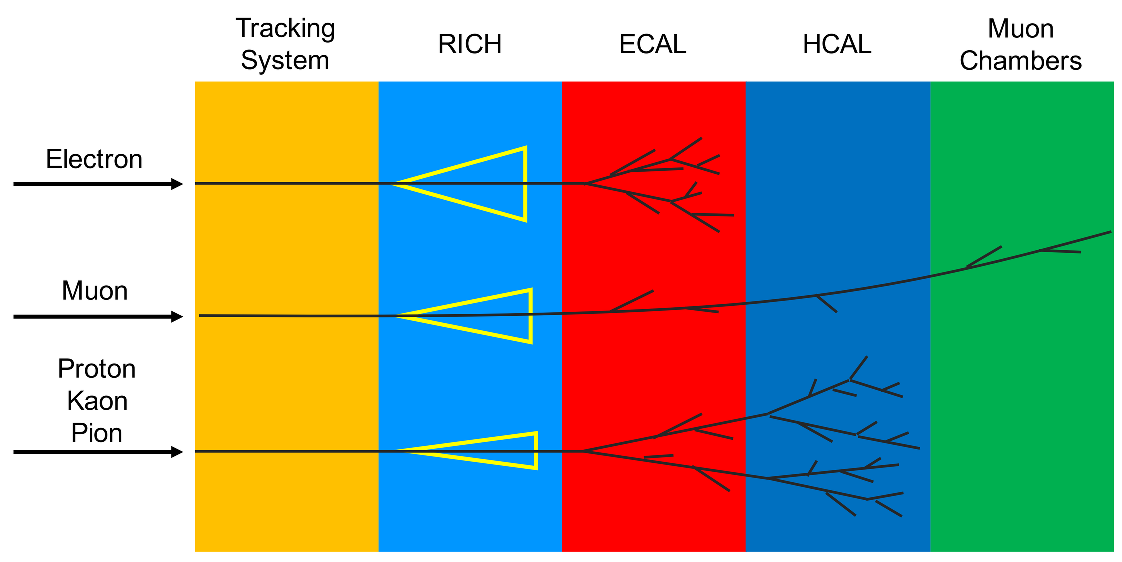

Figure 2: Illustration of different particle type responses in the LHCb systems.

PID is a multiclassification problem in machine learning. Information from the LHCb tracking system, RICHs, calorimeters and muon chambers are used as inputs for the following classifiers to estimate a track type: **ProbNN** [1] (baseline) is an one hidden layer neural network of TMVA library; **Deep NN** is a deep neural network of Keras library; **CatBoost** [2] is a gradient boosting over oblivious decision trees classifier.

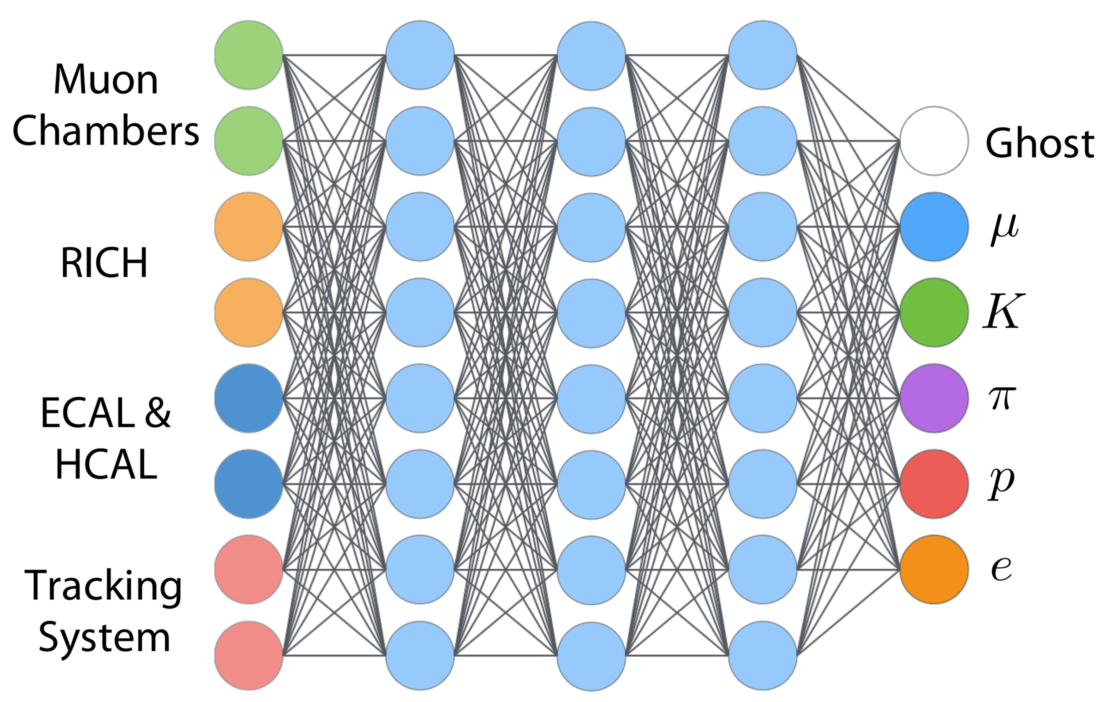

Figure 3: Example of deep neural network used for the particle identification.

The classifiers are trained on a MC sample containing all of the different charged particle types. Calibration samples, containing particles that can be identified purely based on only kinematic properties, are used to estimate the classifier performance on real data. The samples contain decays that allow particle types to be identified based only on kinematic properties. The PID performance of each classifier is shown in Fig 4.

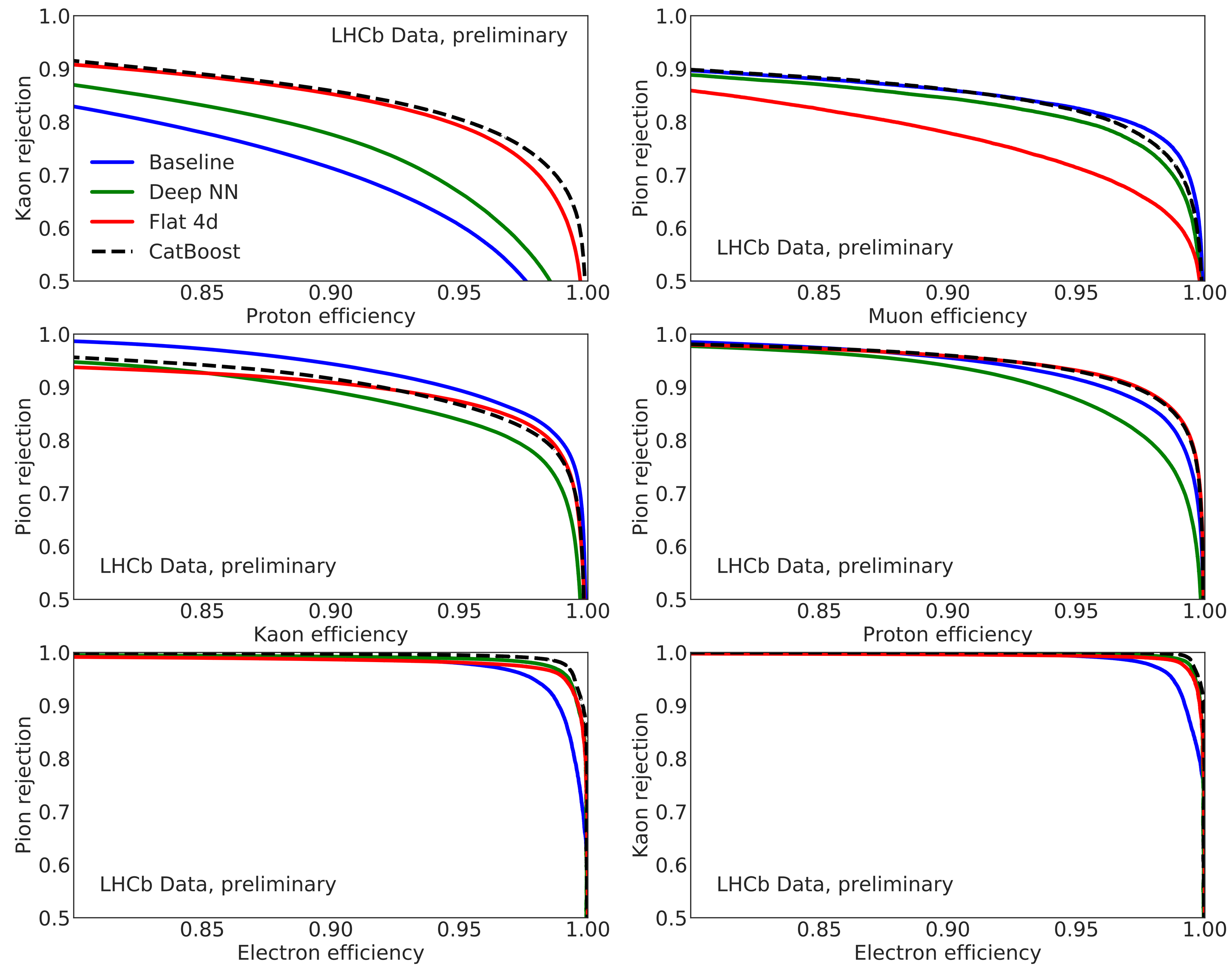

Figure 4: Dependences between background rejection and signal efficiency for six particle pairs.

Flat PID Model

The PID information strongly depends on the kinematic variables. This relationship leads to strong dependency between PID efficiency and kinematic variables as shown in Fig 5. Relative to the baseline model, the **Flat 4d** model, which is a boosted decision trees classifier, has a flatter PID efficiency as a function of particle p, p_T, η and $nTracks$ (event multiplicity) observables. The classifier achieves this flatness using a modified loss function [3].

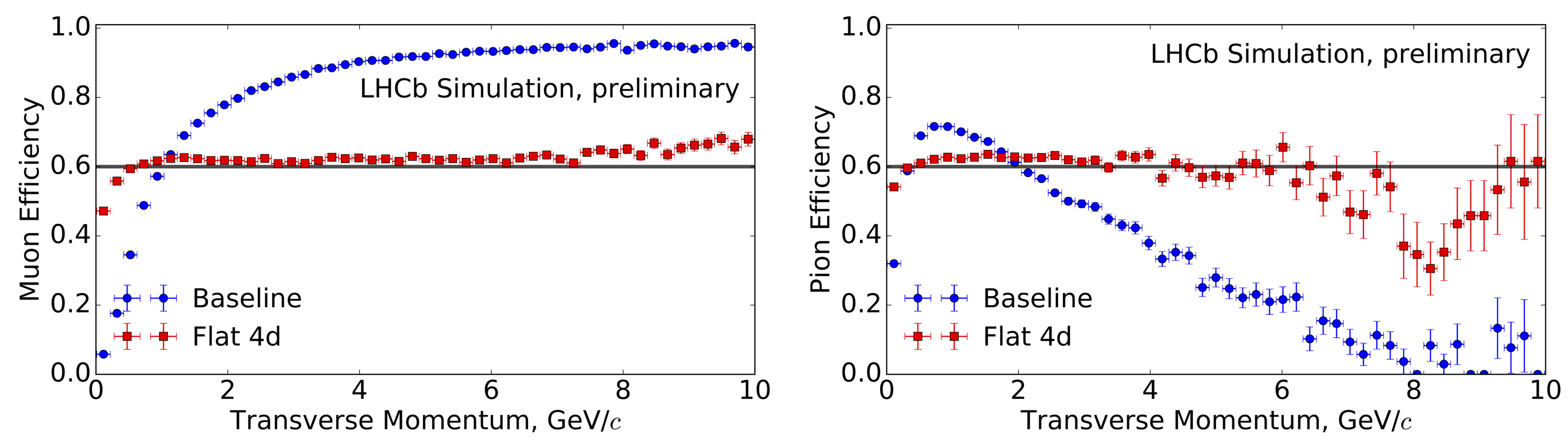

Figure 5: Dependence between Flat 4d model efficiencies and particle transverse momentum for each particle type. The curves correspond to the same global signal efficiency of 60%.

Conclusions

Combining information from the LHCb tracking system, ring-imaging Cherenkov detectors, electromagnetic and hadron calorimeters, and muon chambers using advanced machine learning techniques allows to achieve high quality of global charged particle identification.

Figure 6: Infinity Glove with Infinity Stones. Marvel Entertainment, LLC©.

References

[1] LHCb collaboration. LHCb detector performance. *Int. J. Mod. Phys.*, A(30):1530022, 2015.

[2] A. Veronika Dorogush, A. Gulin, G. Gusev, N. Kazeev, L. Ostroumova Prokhorenkova, and A. Vorobev. Fighting biases with dynamic boosting. *ArXiv e-prints 1706.09516*, June 2017.

[3] Alex Rogozhnikov, Aleksandar Bukva, Vladimir Gligorov, Andrey Ustyuzhanin, and Mike Williams. New approaches for boosting to uniformity. *Journal of Instrumentation*, 10(03):T03002, 2015.