Machine Learning for 40 MHz scouting at CMS

Summer Student Lightning Talk

Gianvito Losapio

06/09/2021

Introduction

- The **Level 1 (L1) trigger** at CMS uses coarse-grained information to search for signatures of interesting physics

40 MHz (LHC collision rate) → 100 KHz (CMS L1 trigger selection rate)

- **L1 scouting** is a new paradigm for data collection at CMS which could help in the early identification of promising potential signals, independently of any trigger selection bias

- **Machine Learning** models can be implemented on Field-Programmable Gate Array devices (FPGAs) for close-to real-time analysis (40 MHz scouting)

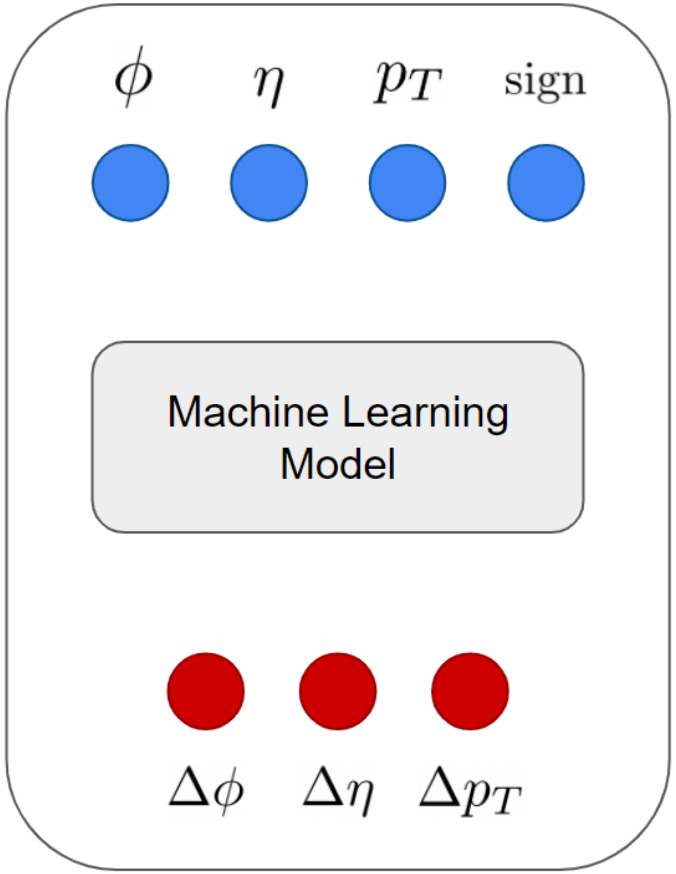
Objective of the project
Investigate efficient machine learning algorithms
with fast inference time for the L1 scouting system

Muon re-calibration

- Given the muon track parameters measured by the L1 trigger train a ML model to predict the correct re-calibration for each of them (multivariate regression)

- Four muon track parameters from the ZeroBias L1 trigger dataset (LHC Run-2, 2017-2018) as inputs:
  - azimuthal angle ϕ
  - pseudorapidity η
  - transverse momentum p_T
  - particle charge sign

- Difference between L1 values and reconstructed values (from offline analysis) as targets: $\Delta\phi$, $\Delta\eta$, Δp_T

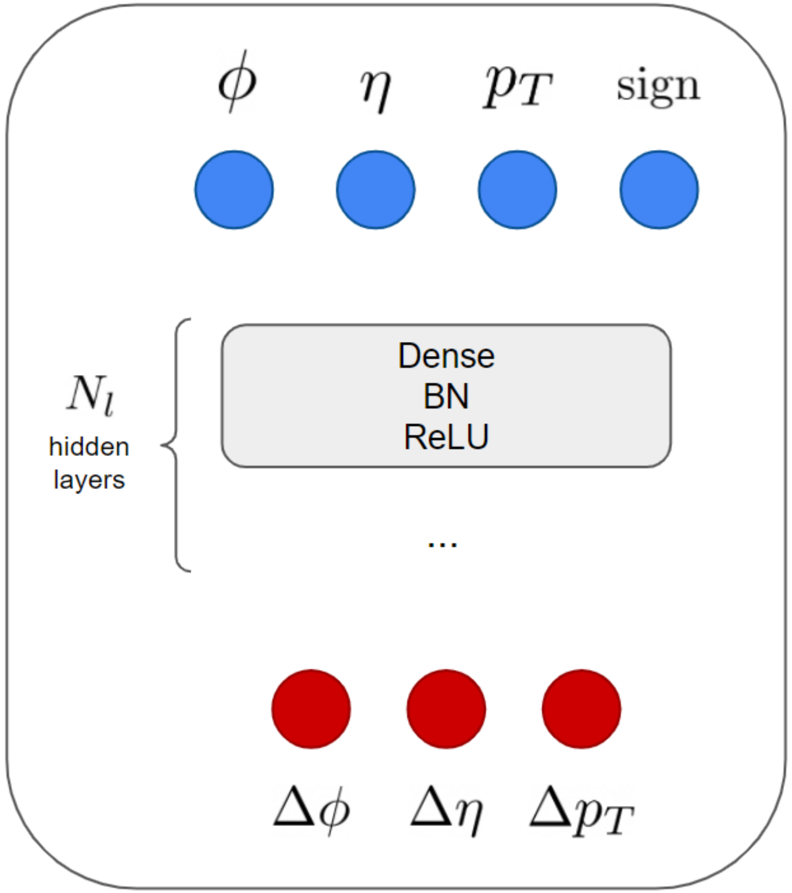


Neural network

- Deep feedforward network (Tensorflow):
  - N_l hidden layers with H_n nodes each
  - Batch normalization
  - ReLU activation function
  - Square/Logcosh loss
  - No/L2 regularization

- Optimization:
  - Adam with default params
  - Different batch sizes
  - Early stopping

- Trained two different models:
  - 4 hidden layers with 128 neurons
  - 3 hidden layers with 32 neurons



Falkon

Fast, large-scale kernel method developed at MaLGa Center, Genoa (*)

- Kernel function (e.g., Gaussian, linear, polynomial kernel)

$$k(x, x_i) = \exp\left(\frac{\|x - x_i\|^2}{2\sigma^2}\right) \qquad k(x, x_i) = \beta + \frac{1}{\sigma^2} x^\top x_i$$

$$k(x, x_i) = (\alpha\, x^\top x_i + \beta)^{\deg}$$

- Nystrom centers selection (e.g., random)

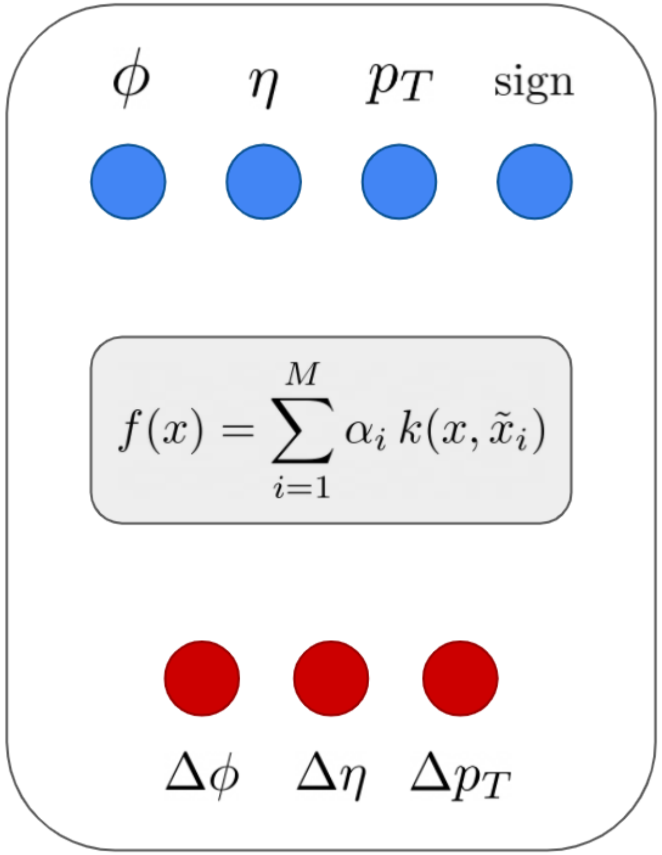
$$\{\tilde{x}_1, \ldots, \tilde{x}_M\} \subset \{x_1, \ldots, x_N\}$$

- 2 relevant hyperparameters + kernel parameters (to be tuned)

M number of Nystrom centers

λ regularization parameter

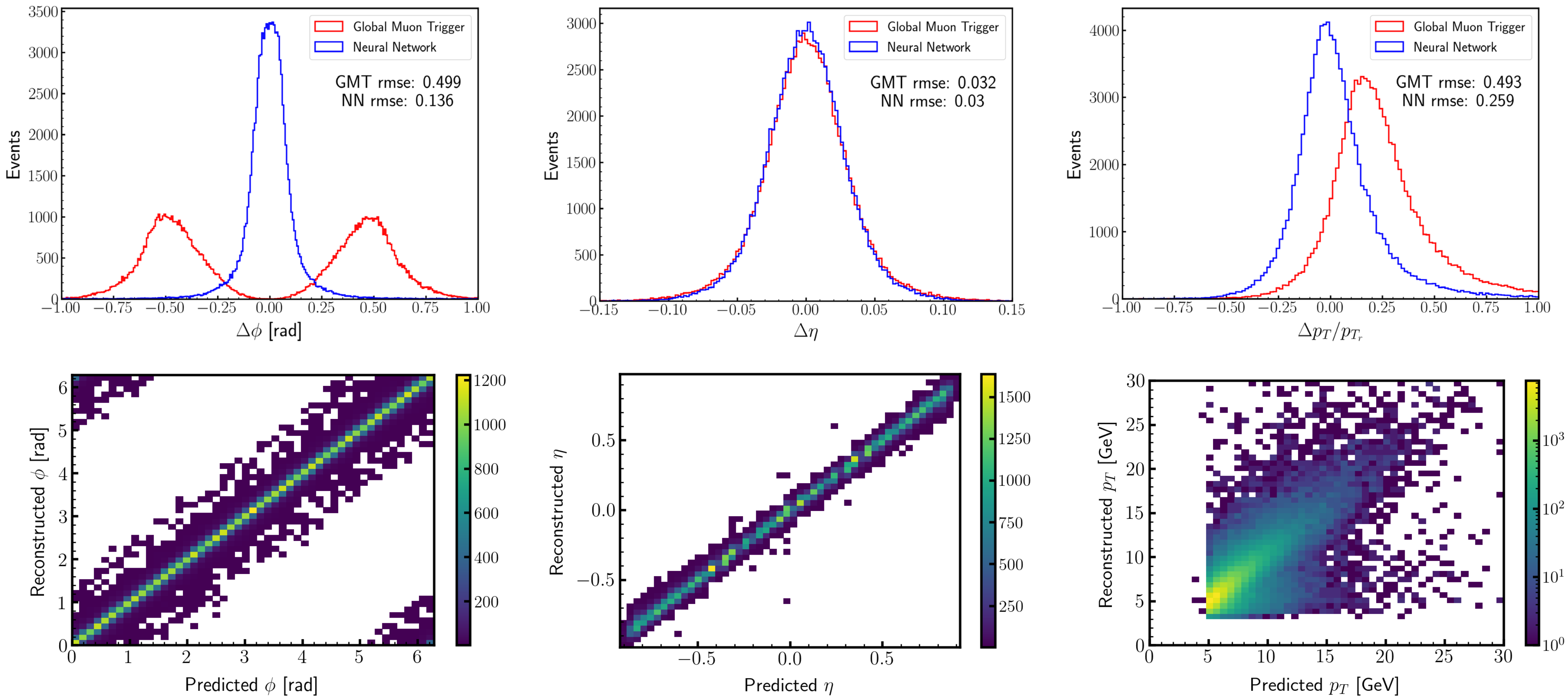
- Trained with different kernels, parameters



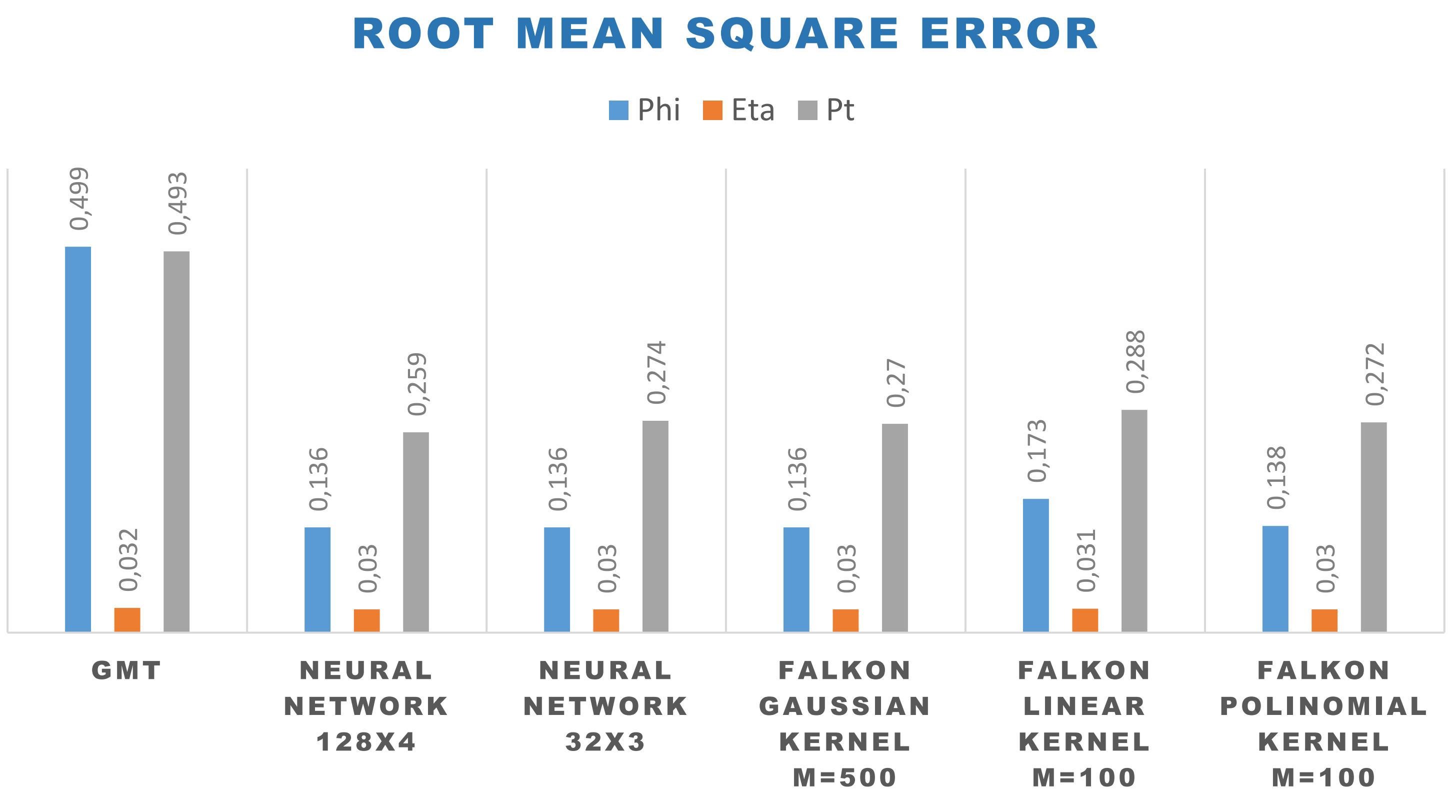
(*) https://falkonml.github.io/falkon/

Results (1/2)

GMT = Global Muon Trigger NN = Neural Network



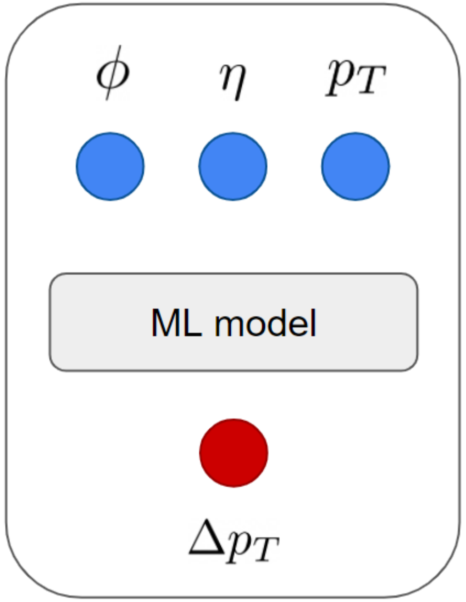
Results (2/2)



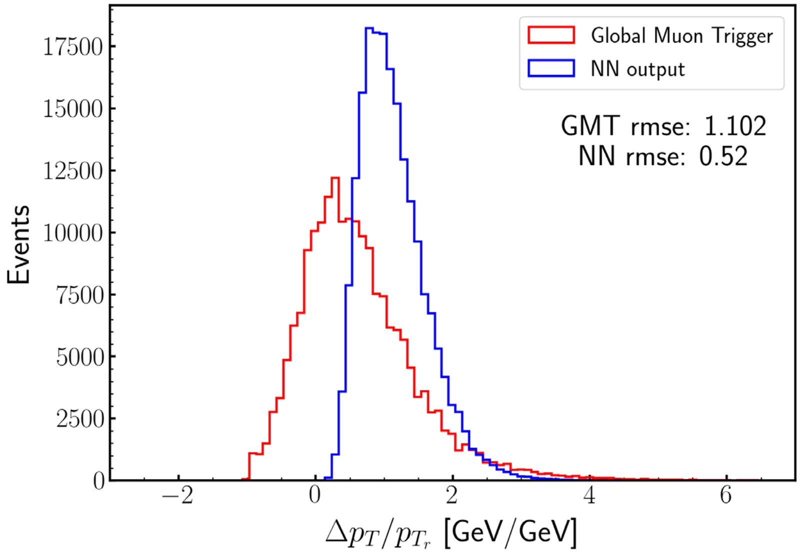
	Approx. Flops (per inference)
Neural Network 128x4	66 432
Neural Network 32x3	3296
Falkon Linear kernel	**304**

Calorimeter data studies

- Learn re-calibration of jet transverse momentum, given L1 measurements
- Match L1 measurements to reconstructed values based on the condition: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$
- Preliminary tests with both ML models show potential for improvement in jet p_T measurements



Preliminary: work in progress



Conclusions and future work

- Neural networks and Falkon are two different ML models providing good results for the muon re-calibration problem

- With proper parameters, Falkon can preserve accuracy and provide advantages in terms of inference time (crucial to meet strict latency requirements on the L1 scouting system)

- Results can be easily extended to calorimeter data. Preliminary experiments show encouraging results with both ML models

- For both neural networks and Falkon, more refined strategies in the training procedure can be used to give more importance to rare particles (crucially important areas of phase space for physics studies)

QUESTIONS?

gianvito97losapio@gmail.com