Towards an Interpretable Data-driven Trigger System for High-throughput Physics Facilities

Kristin Dona, David W. Miller, Benjamin Rosser, Cecilia Tosciri (University of Chicago)

Motivations

- The volume of data available at contemporary experiments for high-energy physics is enormous and complex
- Sophisticated trigger systems for selecting relevant physics processes
- A trigger menu defines the selection criteria
- Trigger design relies heavily on prior knowledge of the feature-space being probed and modularized approaches to optimization are used
- Redundant labeling schemes and cost-ineffective algorithm execution

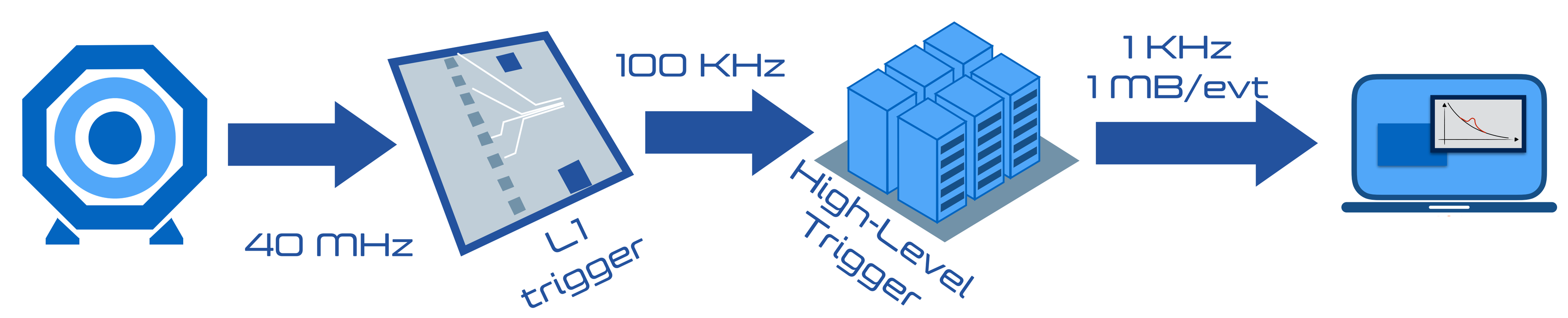
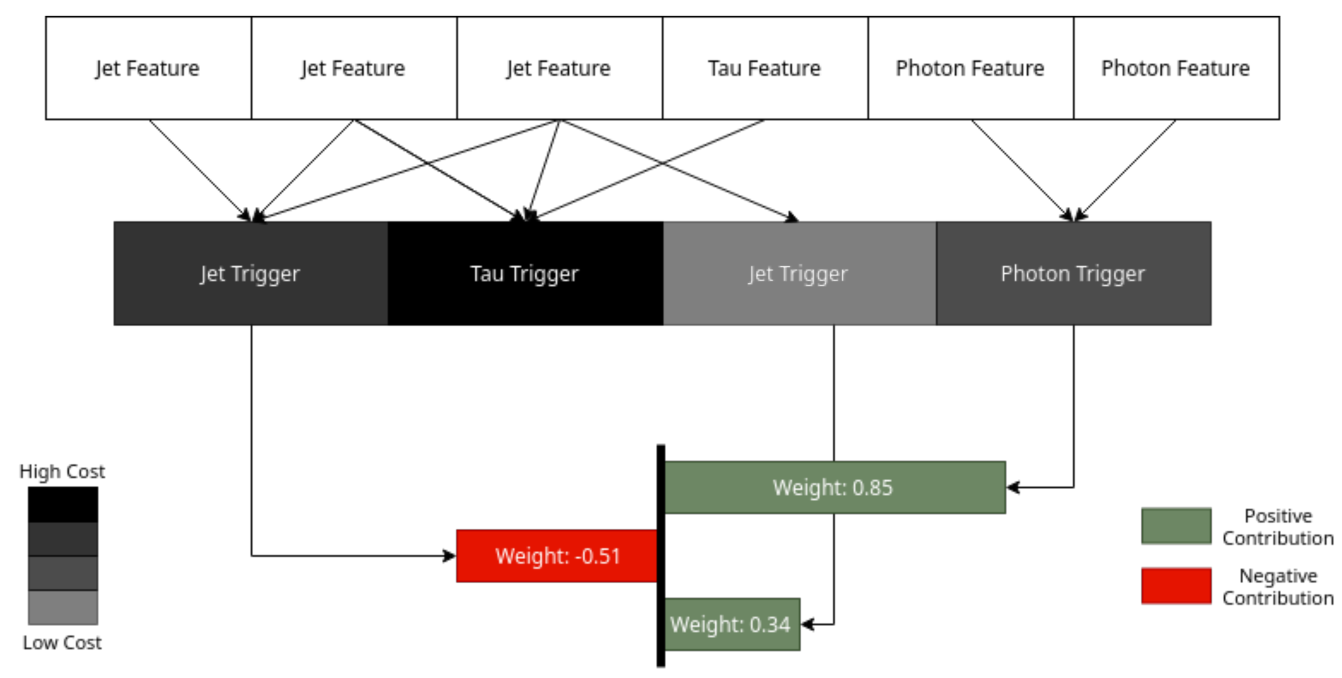


Figure illustrates the complex trigger system used at LHC experiments, as an example

Data-driven Interpretable Trigger

- Replace the hand-designed trigger menu with an optimized data-driven trigger system with **minimal run-time cost**, without compromising **physics coverage**
- **Interpretable** predictive model: for any given event, provides explanation for the trigger decision; allows for pruning costly ineffective individual algorithm labels



Dataset

Simulated **MC dataset** (CMS Collaboration [1]) of top quark pair events generated in a p-p collision with a centre-of-mass energy of 8 TeV. Includes features of the event and the different trigger selections (defined by specific trigger labels) that each event may or may not have passed.

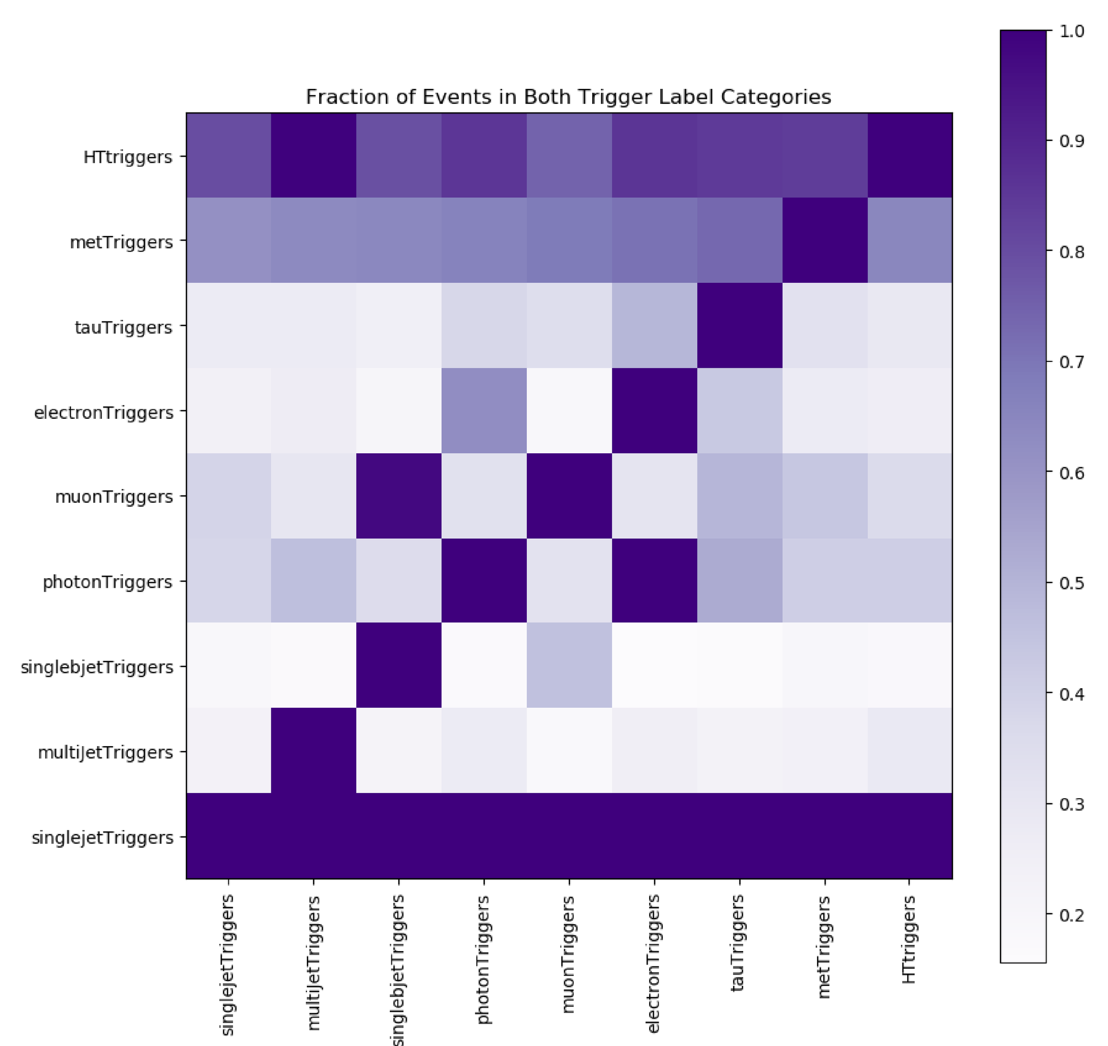


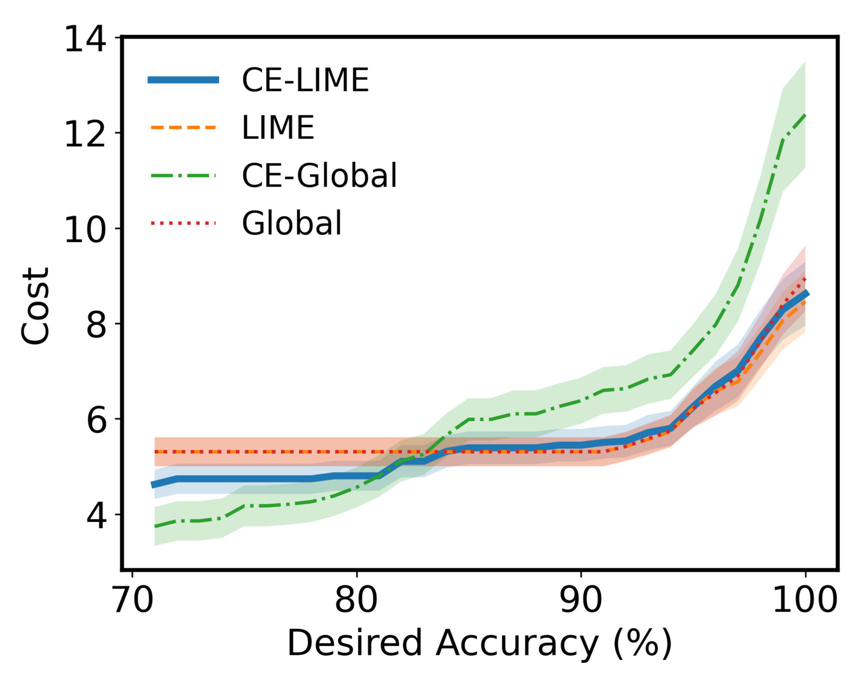
Figure shows the correlation analysis on the dataset: the trigger menu are heavily overlapped. This suggests the potential for these trigger algorithms to be optimized.

Cost-Effective Interpretable Model

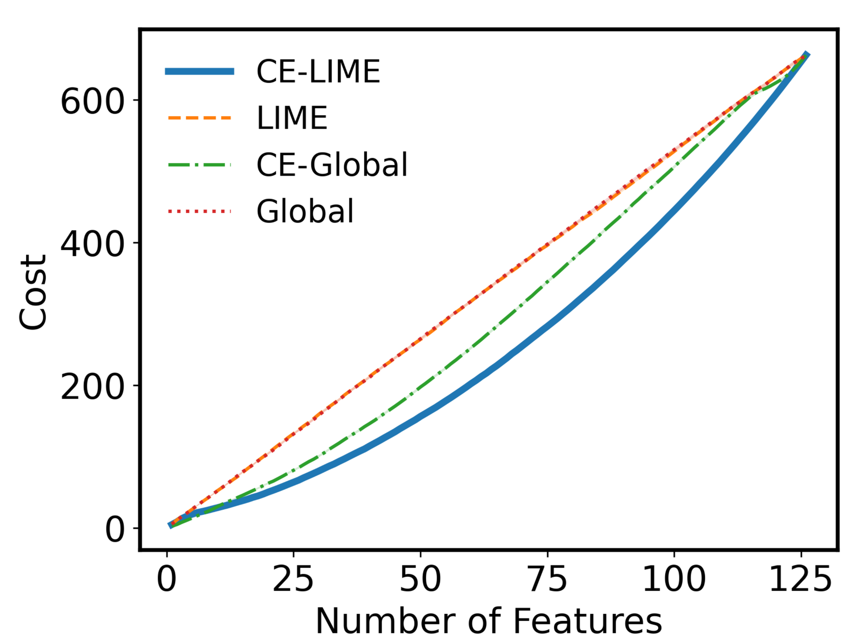
Given a set of candidate trigger labels from the existing trigger menu and a latency cost for each trigger label, we seek an **optimal subset of labels** that make the correct filtering decision with the minimal latency cost. The solution is used as the explanation of our predictive model, which is used to optimize the latency of the existing trigger system, by pruning the costly labels.

A novel **cost-effective elastic net** is used to construct **local interpretable model-agnostic explanations**: **CE-LIME**. The algorithm returns a weight vector β describing the importance of each feature f_i, accounting for the cost $c(f_i)$

$$\hat{\beta} = \arg\min_{\beta} \left(|y - X\beta|^2 + (1-\alpha)\lambda \sum_{i=1}^{p} |\beta_i| \cdot c(f_i) + \alpha\lambda \sum_{i=1}^{p} |\beta_i|^2 \cdot c(f_i) \right)$$



(a) Cost vs Performance, CMS Open Data



(b) Cost vs # Used Features, CMS Open Data

Figure a: the required cost to achieve any accuracy level is the lowest for CE-LIME for most cases

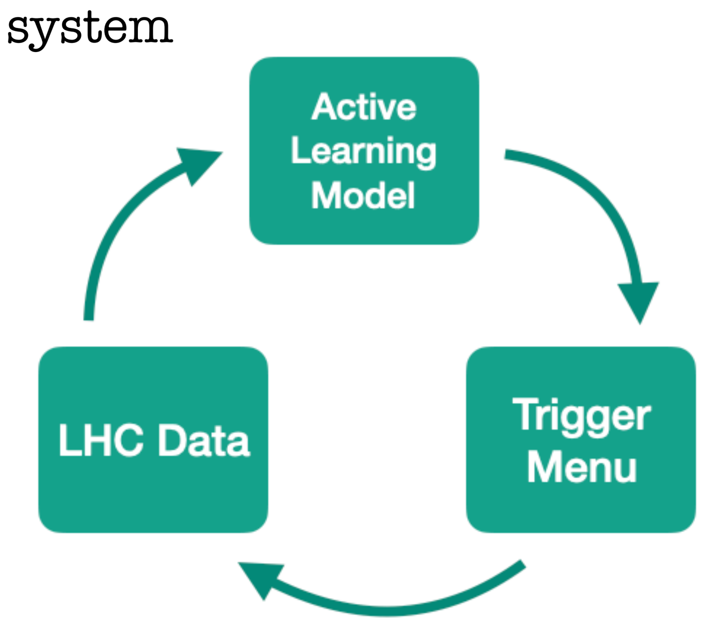
Figure b: the cost of using a certain number of features in the order given by CE-LIME is competitive

Automated Trigger Menu Refinement via Active Learning

- Using the cost-effective interpretable model, we construct an **active learning model** that **continuously updates itself** with incoming data, and **provides explanations** for those updates
- Properly quantify the **uncertainty of the decisions** of the data-driven trigger system

We seek the optimal active learning model that learns from data and decides which trigger label to query in a sequential manner. Only through interpretation of what is learned, we can understand why and how the **trigger menu** should be updated. We construct an **autoencoder** to capture the density estimate for each event.

Biggest challenge: simulating the selection of events via this mechanism, and correctly modeling and understanding the efficiencies of this evolving self-driving trigger system.



Ongoing Work

- We are exploring novel algorithms that account for the objective (anomaly detection) and the budget constraint on the label cost
- Demonstration of the approach is being prototyped using a Xilinx Versal Adaptive Compute Acceleration Platform (ACAP) board [3]
- Xilinx Versal ACAPs contain programmable AI engines, dedicated processors for high-speed, real-time data processing that support dynamic reconfiguration

References

[1] CMS collaboration (2017).DOI:10.7483/OPENDATA.CMS.XH95.JNSE
[2] Y. Chen et al., Self-driving data trigger, filtering, and acquisition, Snowmass2021 LOI (2020)
[3] Xilinx AI Engine Technology

THE UNIVERSITY OF CHICAGO