Using TMine for the Fermi-LAT Event Analysis

Alex Drlica-Wagner (Stanford/SLAC) and Eric Charles (SLAC)
on behalf of the Fermi Large Area Telescope Collaboration

Summary: TMine is a new tool for implementing multivariate classification algorithms in the Fermi-LAT event analysis. In this poster, we present the TMine tool and discuss its application to the creation of high-level science variables (e.g. event energy, incident direction, and particle type).

Abstract

The Fermi-LAT event analysis is the final stage in the creation of high-level science variables (e.g. event energy, incident direction, particle type, etc.). We discuss the development of TMine, a powerful new tool for designing and implementing event classification analyses (i.e. distinguishing photons from charged particles). TMine is structured on ROOT, a data analysis framework that is the de-facto standard for the current high-energy physics experiments, and thus fits naturally into the ROOT-based data processing pipeline of the LAT. TMine utilizes advanced multivariate classification algorithms implemented in ROOT, and allows for the visual development of the LAT event analysis. We discuss the implementation of TMine in the current analyses and improvements that TMine can offer for developing the next iteration of the event analysis (Pass 8).

Fermi-LAT Event Analysis

Thousands of particles trigger the Fermi-LAT every second, many of which undergo full event reconstruction after on-board filter selection. After reconstruction, the LAT event analysis combines information from each detector subsystem (the anticoincidence detector, tracker, and calorimeter) to create a picture of the event as a whole. From this picture, high-level science variables (i.e. energy, direction, and particle type) are assigned to each event [1].

Multivariate Classification:

- Assigning high-level science variables to each event is a difficult task.
- The LAT accepts particles over a wide range in parameter space (both in energy and incident angle) and event topology (e.g. close to detector edges and gaps).
- The LAT event analysis uses multivariate classification algorithms, since classic cut based analyses do not have sufficient signal efficiency and background rejection [1].
- TMine is a new tool for implementing multivariate classification algorithms with the goal of improving the LAT instrument response (i.e. the effective area, energy resolution, and point spread function).

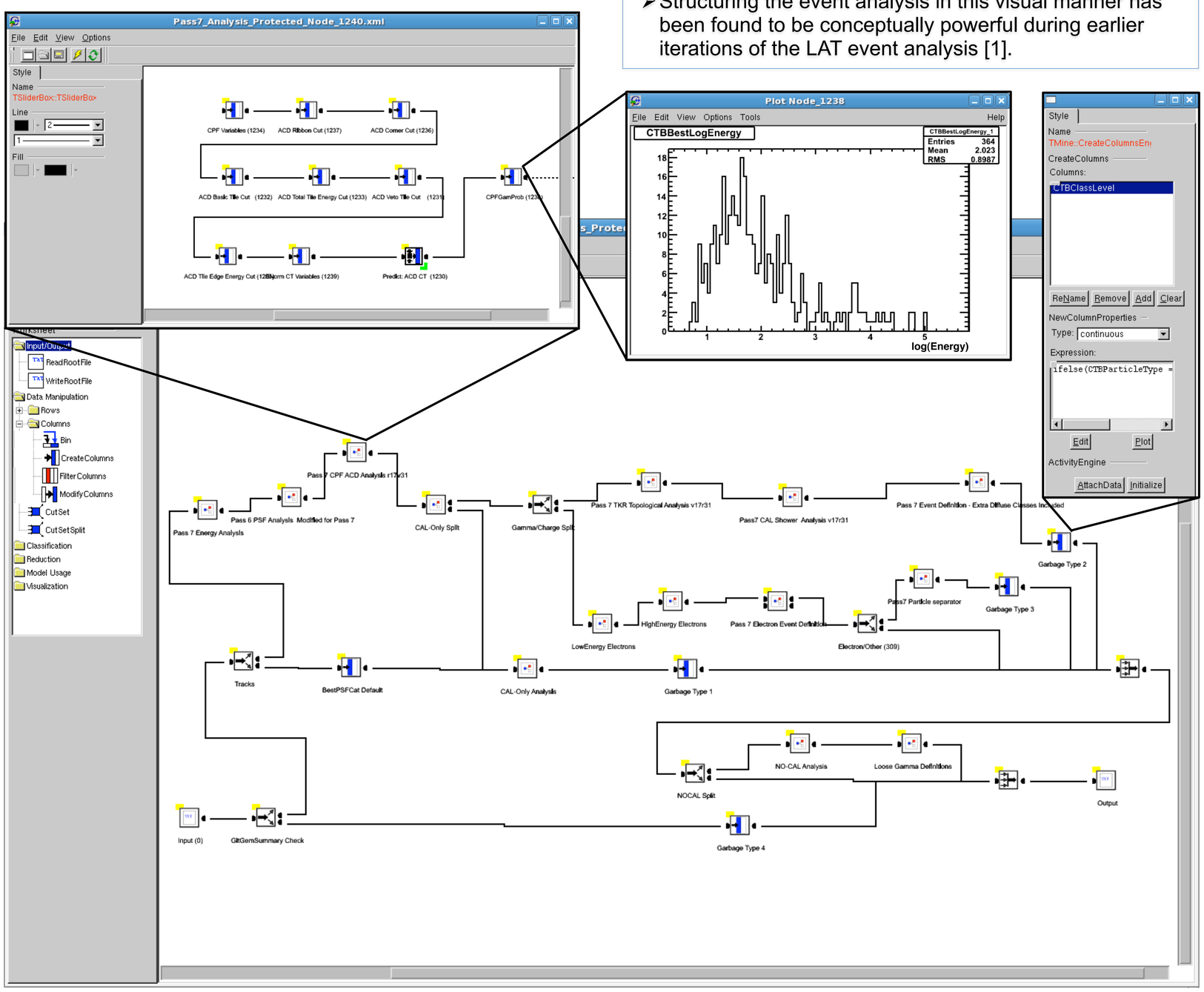

Figure 1 - The current iteration of the Fermi-LAT event analysis as viewed by TMine. Many nodes contain sub-analyses themselves (**inset left**), while the functionality of each active node can be plotted (**inset middle**) and edited through a GUI editor (**inset right**).

TMine Analysis Tool

What is TMine?

- TMine is a tool for designing and implementing complex event classification analyses using ROOT [2], the de-facto data analysis framework for current high-energy physics experiments.
  - Deals with large data sets in a quick and memory efficient manner.
  - Utilizes the ROOT Toolkit for Multivariate Analysis (TMVA) [3] for the processing and parallel evaluation of sophisticated multivariate classification algorithms
  - Preserves command line functionality and provides the user with a visual working environment
- TMine need not be restricted to high energy physics applications.

What is a TMine Analysis?

- A TMine analysis consists of a network of directionally linked nodes to control work flow and operation (Fig. 1).
- Nodes both alter event characteristics (i.e. variable definition and assignment) and direct events through the network.
- Specialized nodes are used for training, testing, and implementing TMVA classification algorithms.
- Structuring the event analysis in this visual manner has been found to be conceptually powerful during earlier iterations of the LAT event analysis [1].

Applications of TMine

TMine and the Development of Pass 8:

- Pass 8 is a complete reworking of the LAT simulation and reconstruction, benefitting from analysis of flight data (which was unavailable before launch).
- TMine will improve the interface between event reconstruction and event classification.
- It also provides improvements to the structure and validation of the Pass 8 event analysis.
- TMine has built in functionality for comparing real and simulated data, an important step towards applying multivariate classification algorithms [4].

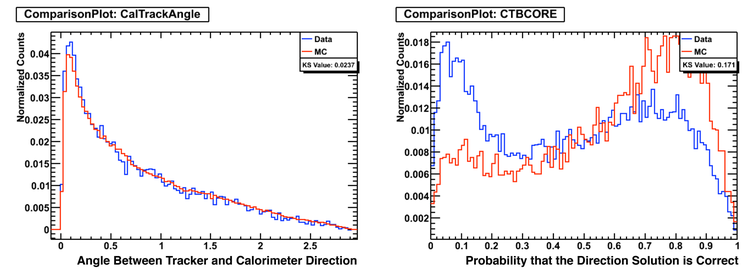

Figure 2 - Comparisons between simulated photons (red) and a statistical sample of photons from flight data (blue). Only variables with good agreement should be used for classification.

Other Applications of TMine in Fermi-LAT Analyses:

- The LAT electron analysis [5] uses TMine for data reprocessing.
- Ongoing analyses to measure the proton spectrum use TMine for event classification (Fig. 2).
- TMine has also been used in on-going attempts to classify unassociated sources in the first LAT catalog [6].

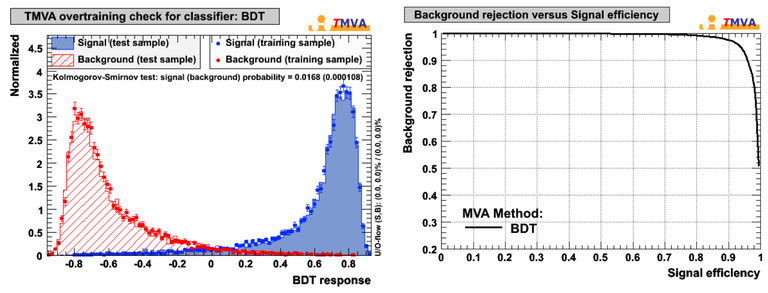

Figure 3 - TMine implementation of the TMVA boosted decision trees to distinguish protons from electrons and positrons.

Acknowledgements

ADW acknowledges support from from the Department of Energy Office of Science Graduate Fellowship Program (DOE SCGF), made possible in part by the American Recovery and Reinvestment Act of 2009.

References

[1] W.B. Atwood, et al. Ap.J. 697 (2009) 1071-1102
[2] R. Brun and F. Rademakers, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86.
[3] A. Hoecker, et al. PoS ACAT 040 (2007)
[4] J. Bregeon. Poster at this Symposium.
[5] A. Abdo, et al. Phys. Rev. Lett. 102 (2009) 181101
[6] M.E. Monzani, et al. Poster at this Symposium.