Use of Machine Learning for gamma/hadron separation with HAWC

T. Capistrán, K. L. Fan, J. T. Linnemann, I. Torres, P. M. Saz Parkinson, Philip L.H. Yu, and for the HAWC Collaboration

We acknowledge support by:

Abstract

Background showers triggered by hadrons represent over 99.9% of all particles arriving at ground-based gamma-ray observatories. An important stage in the data analysis of these observatories, therefore, is the removal of hadron-triggered showers. Currently, the High-Altitude Water Cherenkov (HAWC) gamma-ray observatory employs an algorithm based on a single cut in two variables, unlike other ground-based gamma-ray observatories (e.g. H.E.S.S., VERITAS), which employ a large number of variables to separate the primary particles. In this work, we explore machine learning techniques (Boosted Decision Trees and Neural Networks) to identify the primary particles detected by HAWC. Our new gamma/hadron separation techniques were tested on data from the Crab nebula, the standard reference in Very High Energy astronomy, showing an improvement compared to the standard HAWC background rejection method.

HAWC

- ✴ Ground-base gamma-ray detector,
- ✴ Detect high energy sources,
- ✴ Located in Mexico.

Data/simulation comparison

- ✴ Figure 1 shows the background data near Crab versus the eight nuclei particles simulated for LIC variable.
- ✴ Figure 2 shows the comparison of the gamma ray simulation versus Crab Nebula data.

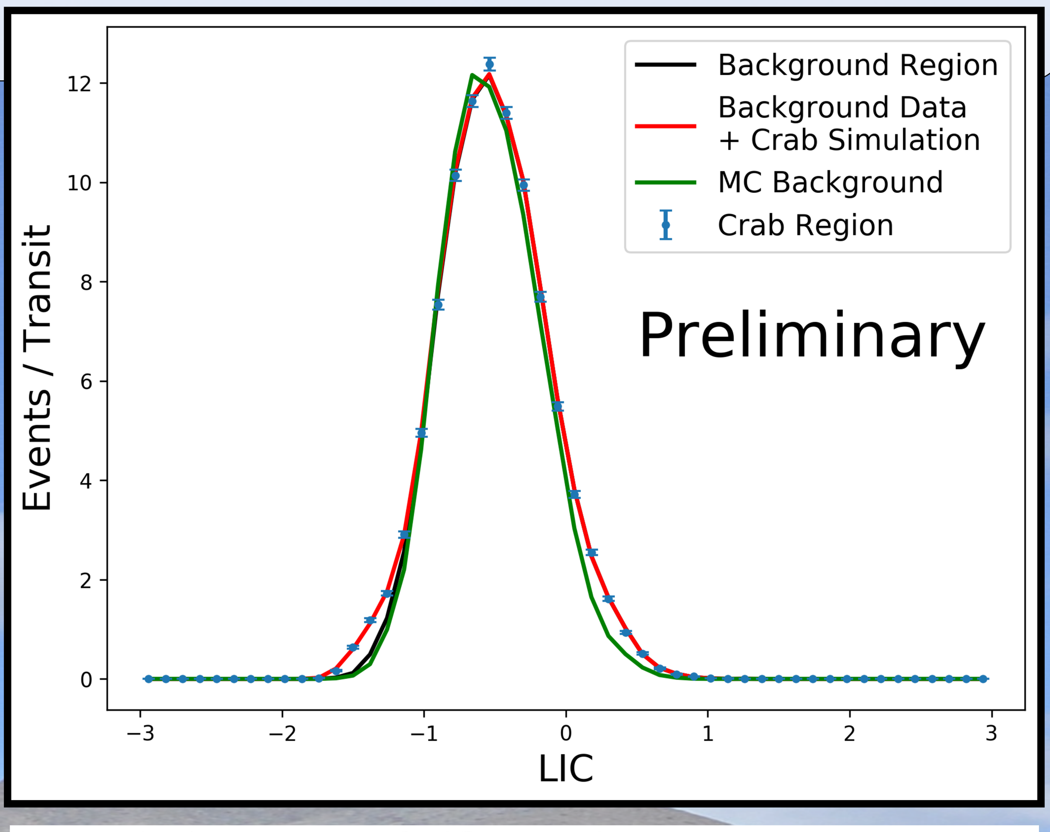

Figure 1: Distribution of LIC variable. First is the background comparison between simulation data (MC Background) and off-Crab (Background region) additional show the events on the Crab Nebula and a prediction of this distribution using simulation data.

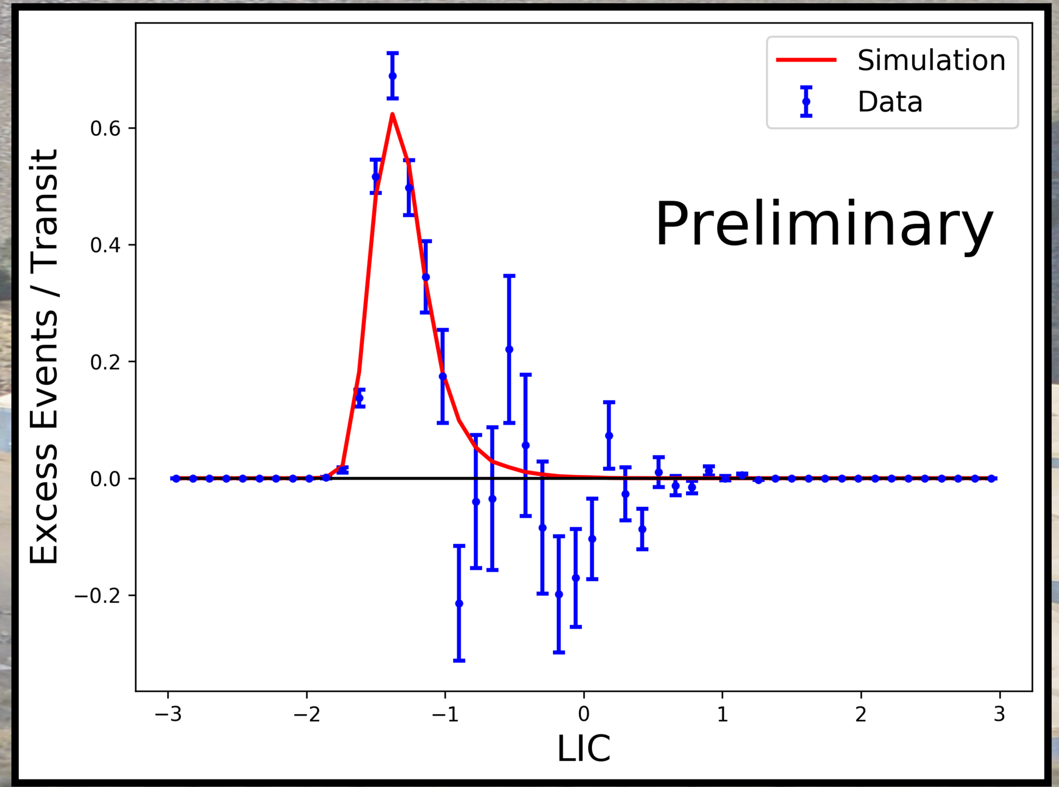

Figure 2: Distribution of LIC variable using gamma-ray simulation (red histogram) and the signal subtraction on the Crab (blue histogram).

G/H separation using MLT

- ✓ Two machine learning technique (MLT) were trained Neural Network (NNModel) and Boosted Decision Tree (BDT).
- ✓ Each model were fed with 7 input variable.
- ✓ The models were training and testing with simulation.

Results

- ✤ MLT are better than SC for identifying the gamma events (Figure 3),
- ✤ SC is the best model to recognize the hadron events (Figure 4),
- ✤ BDT has reported be the best model using the Crab data. (Figure 5 & 6).

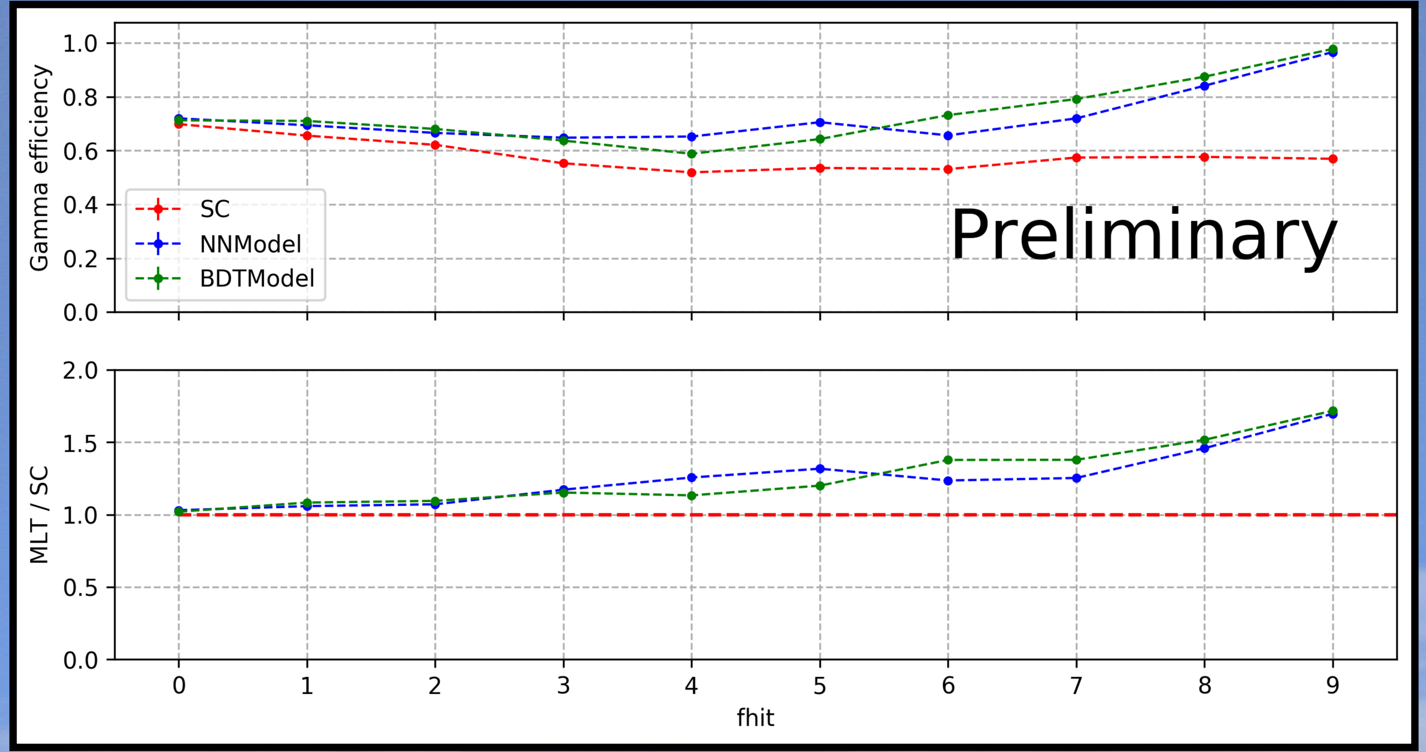

Figure 3: Top panel shows the gamma efficiency for three G/H separation method. The bottom shows the comparison between MLT and SC.

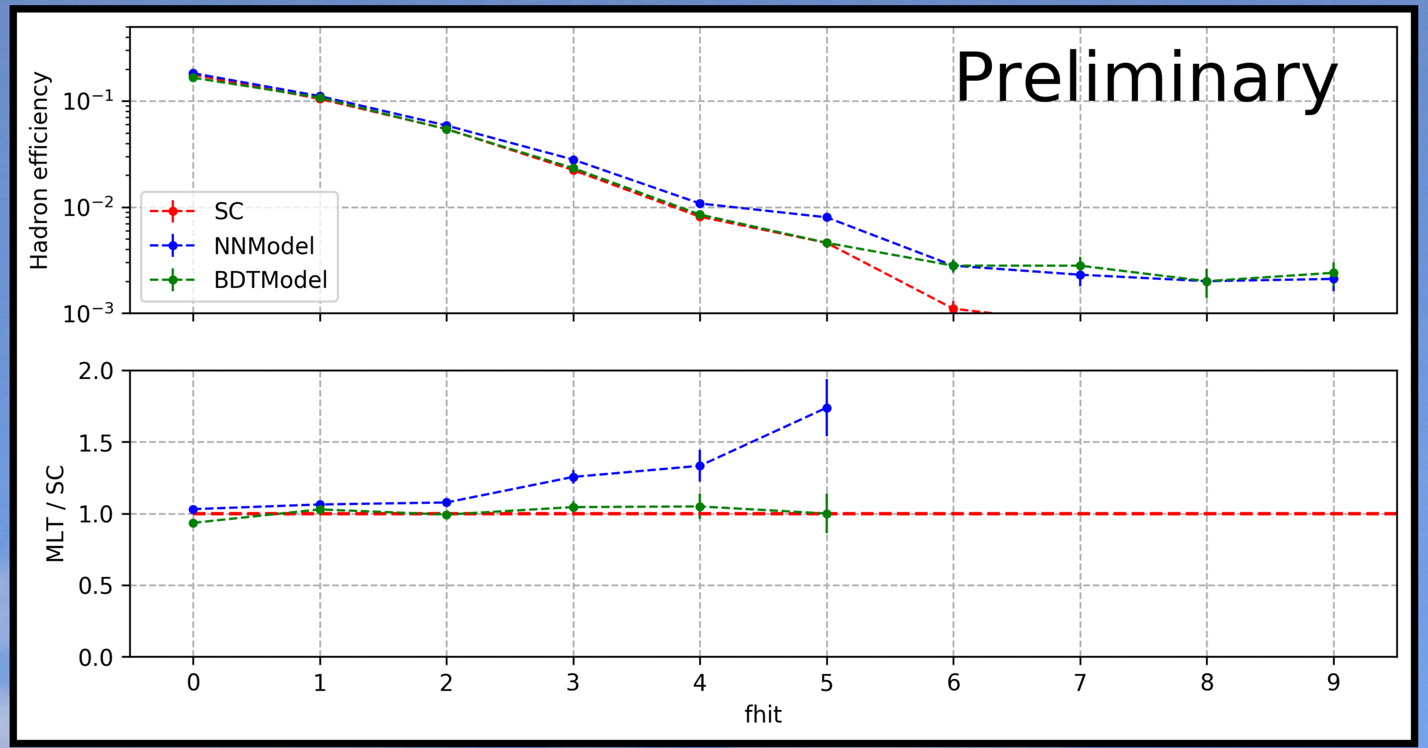

Figure 4: Same as to Figure 3 but here shows the hadron efficiency.

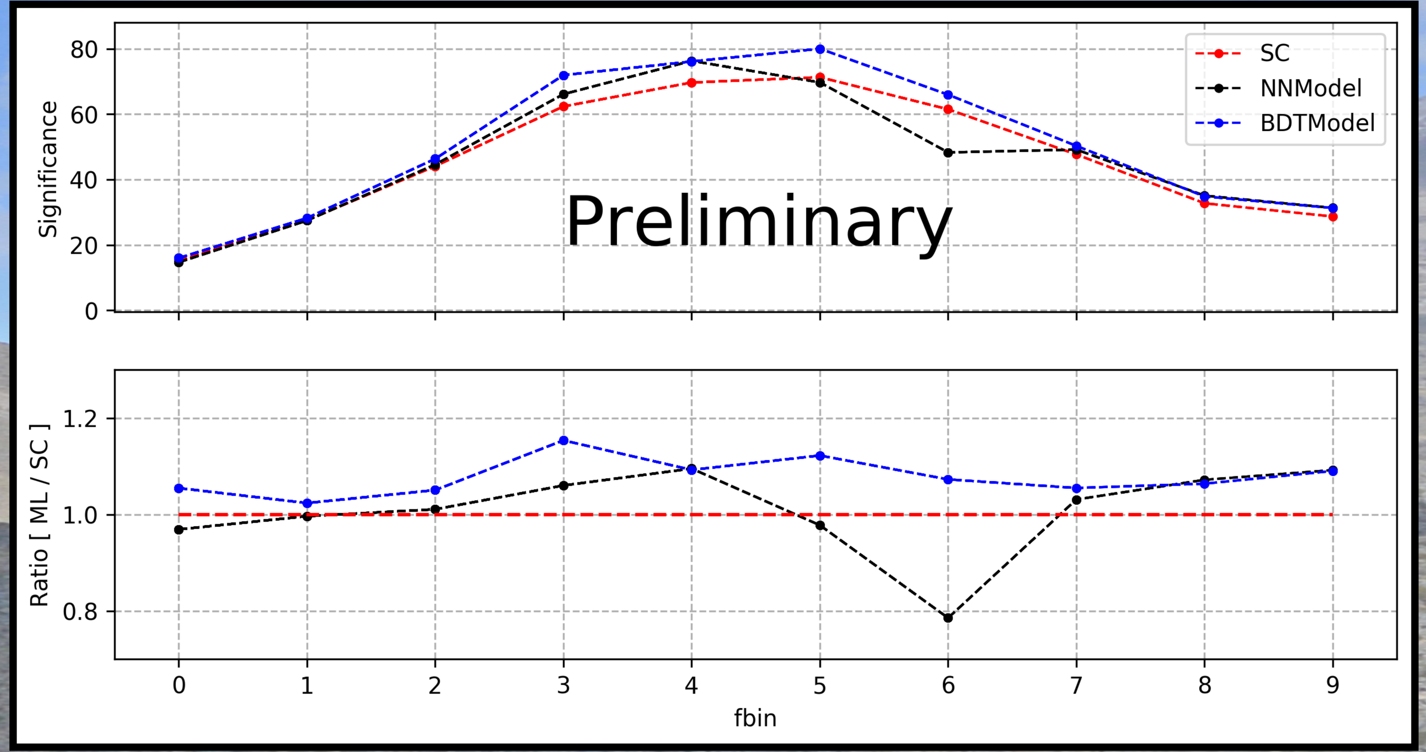

Figure 5: Top panel show the significance on the Crab using the G/H separation model as a function of fbin

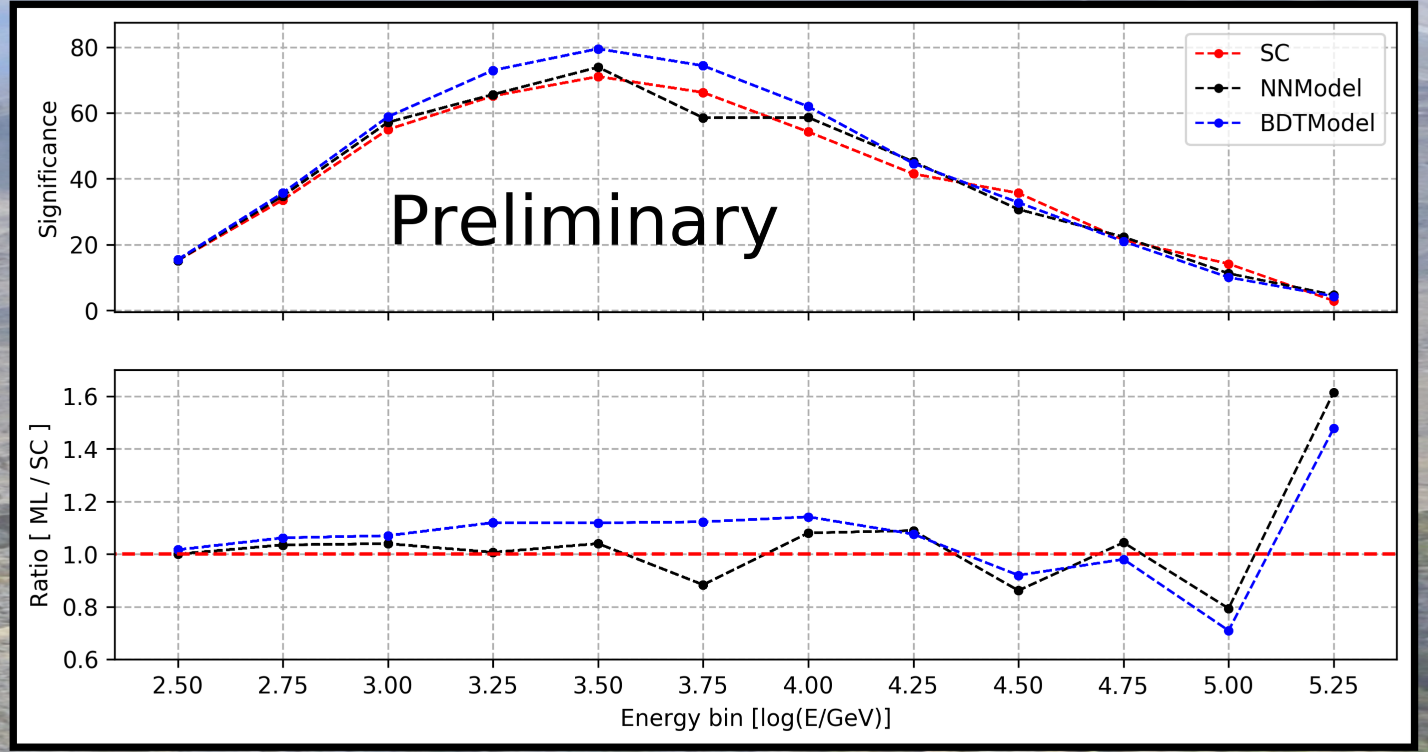

Figure 6: Same as to Figure 5 but here report for each energy bin.

Conclusion and discussion

- ➡ Two MLTs were trained using seven input parameters.
- ➡ Both tested with simulation and Crab Nebula data.
- ➡ Using simulation, both MLTs report higher performance than SC. However, using real data, only the BDT shows that improvement.

Acknowledgments

We acknowledge the support from: the US National Science Foundation (NSF); the US Department of Energy Office of High-Energy Physics; the Laboratory Directed Research and Development (LDRD) program of Los Alamos National Laboratory; Consejo Nacional de Ciencia y Tecnolog\'ia (CONACyT), M\'exico, grants 271051, 232656, 260378, 179588, 254964, 258865, 243290, 132197, A1-S-46288, A1-S-22784, c\'atedras 873, 1563, 341, 323, Red HAWC, M\'exico; DGAPA-UNAM grants IG101320, IN111716-3, IN111419, IA102019, IN110621, IN110521; VIEP-BUAP; PIFI 2012, 2013, PROFOCIE 2014, 2015; the University of Wisconsin Alumni Research Foundation; the Institute of Geophysics, Planetary Physics, and Signatures at Los Alamos National Laboratory; Polish Science Centre grant, DEC-2017/27/B/ST9/02272; Coordinaci\'on de la Investigaci\'on Cient\'ifica de la Universidad Michoacana; Royal Society - Newton Advanced Fellowship 180385; Generalitat Valenciana, grant CIDEGENT/2018/034; Chulalongkorn University's CUniverse (CUAASC) grant; Coordinaci\'on General Acad\'emica e Innovaci\'on (CGAI-UdeG), PRODEP-SEP UDG-CA-499; Institute of Cosmic Ray Research (ICRR), University of Tokyo, H.F. acknowledges support by NASA under award number 80GSFC21M0002. We also acknowledge the significant contributions over many years of Stefan Westerhoff, Gaurang Yodh and Arnulfo Zepeda Dominguez, all deceased members of the HAWC collaboration. Thanks to Scott Delay, Luciano D\'iaz and Eduardo Murrieta for technical support..

References

1. T.C. Weekes et al., *Observation of TeV gamma rays from the Crab nebula using the atmospheric Cerenkov imaging technique*, The Astrophysical Journal **342** (1989) 379.
2. D.J. Fegan, *Gamma/hadron separation at TeV energies*, Journal of Physics G: Nuclear and Particle Physics **23** (1997) 1013.
3. M. Krause et al., *Improved γ/hadron separation for the detection of faint γ-ray sources using boosted decision trees*, Astroparticle Physics **89** (2017) 1 [1701.06928].
4. S. Ohm et al., *γ/hadron separation in very-high-energy γ-ray astronomy using a multivariate analysis method*, Astroparticle Physics **31** (2009) 383 [0904.1136].