Contribution ID: **54** Type: **Talk (12+3min)**

Machine Learning for Classification and Data Analysis in XENON1T

Wednesday 18 July 2018 13:15 (15 minutes)

An important step in the XENON1T analysis pipeline is distinguishing electronic recoil events (ERs) from nuclear recoil events (NRs) inside the detector. ERs come from background, while NRs are evidence of a possible signal. This is currently done by hand using a profile likelihood method. The goal is to maximize the acceptance of correctly classified NRs while rejecting as many ERs as possible; this step has a significant impact on the amount of background that can be discarded and therefore the sensitivity of the detector. This is a binary classification problem, a well-understood domain for machine learning and one of its major strengths. Various supervised learning algorithms were tested on calibration data. The best-performing algorithm, a support vector machine (SVM), approximately matched but did not exceed the performance of the profile likelihood method. Various ways that the quality of the dataset affects the performance of machine learning algorithms are discussed. It is shown that machine learning has a comparative advantage over manual statistical analysis only when dealing with large, high-dimensional, poorly-understood datasets. Although many different types of data are collected on each recoil event, most of it can be used to distinguish calibration events from run events, invalidating its use for machine learning. A more promising approach is to use exploratory machine learning techniques to augment human analysis.

Author: BERNSTEIN, Matthew (Rensselaer Polytechnic Institute)

Presenter: BERNSTEIN, Matthew (Rensselaer Polytechnic Institute)