

Machine Learning in XENON1T Analysis

Dillon A. Davis, Rafael Lang, Darrel Masson
Department of Physics and Astronomy, Purdue University, West Lafayette, IN, USA

ABSTRACT

In process of analyzing large amounts of quantitative data, it can be quite time consuming and challenging to uncover populations of interest contained amongst the background data. Therefore, the ability to partially automate the process while gaining additional insight into the interdependencies of key parameters via machine learning seems quite appealing. As of now, the primary means of reviewing the data is by manually plotting data in different parameter spaces to recognize key features, which is slow and error prone. In this experiment, many well-known machine learning algorithms were applied to a dataset to attempt to semi-automatically identify known populations, and potentially identify other features of interest such as detector artefacts. Additionally, using the results of the machine learning process it became possible to cross-check the results of the XENON1T selection cuts. Clustering algorithms were used to segment the dataset into populations, which then recursively split those into additional subpopulations. Upon capturing a subpopulation, a classifier was trained and used to predict if other data could potentially belong to the same population. From this process, it was observed that there were two clustering algorithms that were capable of identifying the electronic recoil band accurately. It was also seen that a few XENON1T selection cuts may need relaxed. These algorithms may be able to be used to tweak the cuts, or continue in search of artefacts. The process of automating the analysis stage by means of machine learning could be further extended by automating the recognition of waveforms using neural networks.

KEYWORDS

Dark matter, XENON1T, machine learning, data mining, LXe detector, data clustering.