Contribution ID: **14**   Type: **not specified**

Young Scientist Forum : Weakly supervised classifiers in High Energy Physics

Thursday 9 March 2017 10:45 (15 minutes)

As machine learning algorithms become increasingly sophisticated to exploit subtle features of the data, they often become more dependent on simulations. This paper presents a new approach to classification called weak supervision in which class proportions are the only input into the machine learning algorithm. A simple and general regularization technique is used to solve this non-convex problem. Using one of the most important binary classification tasks in high energy physics - quark versus gluon tagging - we show that weak supervision can match the performance of fully supervised algorithms. Furthermore, by design, the new algorithm is insensitive to any mis-modeling of discriminating features in the data by the simulation. Weak supervision is a general procedure that could be applied to a variety of learning problems and as such could add robustness to a wide range of learning problems.

Primary authors: SCHWARTZMAN, Ariel Gustavo (SLAC National Accelerator Laboratory (US)); NACH-MAN, Ben (Lawrence Berkeley National Lab. (US)); RUBBO, Francesco (SLAC National Accelerator Laboratory (US)); DERY, Lucio (Stanford University)

Presenter: RUBBO, Francesco (SLAC National Accelerator Laboratory (US))

Track Classification: 11 : Using tracks