

Abstract Submitted
for the APR20 Meeting of
The American Physical Society

Relegation classifier: a machine-learning approach for optimizing analysis significance in the physical sciences KRIPA GEORGE, MICHAEL MCCRACKEN, Washington Jefferson College, WASHINGTON JEFFERSON COLLEGE TEAM — Machine learning models are used to classify signal and background, a process crucial to many analyses in the physical sciences. Such models are often trained by maximizing classification accuracy. We must also maximize the statistical significance of the signal sample, which, in the physical sciences, is a major factor in determining the merit of an analysis. For datasets where the attributes of the signal and background are largely overlapping or for datasets with imbalanced signal/background populations, making accurate classification while keeping significance high is difficult with standard methods. We present the relegation classifier as a new way to optimize the statistical significance in signal identification, where the model has freedom to ignore some areas of input space and the training loss function combines accuracy and significance. We compare the relegation classifier's performance to that of the logistic regression classifier for toy-model datasets (standard moons benchmark with added mass feature), as well as high-energy physics particle production/decay datasets. We compare the results of these classifiers for various signal-to-background ratios for the same datasets, and compare the results of these classifiers for varying degrees of signal/background overlap.

Kripa George
Washington
Jefferson College

Date submitted: 07 Jan 2020

Electronic form version 1.4