A Comparison of Machine Learning Algorithms for Higgs Boson Identification

Glezakou-Elbert, Ourania-Maria (School: Hanford High School)

Since its detection in the Large Hadron Collider (LHC) in 2012, the Higgs boson has remained a key element in confirming the standard model (SM) of particle physics. Current research to find Higgs boson decay channels predicted by the SM requires being able to accurately identify Higgs bosons or signal events from a noisy background. Through use of machine learning, signal events can be accurately classified in offline reconstruction; however, there exist many classification models which might suit this purpose. This study compares three supervised classification models in their ability to accurately and efficiently identify signal events: boosted decision trees, support vector machine, and neural networks. With an accuracy of 83.88%, F1-score of 81.72% and a training time of 8.4 seconds using the simulated dataset from the 2014 ATLAS Higgs boson Machine Learning Challenge, a histogram gradient boosted decision tree was determined to be the most effective classification model for identifying Higgs boson events. The worst performing algorithm was the support vector machine with the lowest accuracy at 80.35%, F1-score of 77.12%, and the second lowest training time of 50 minutes.

Awards Won: