



Contribution ID: 33

Type: Oral Presentation

An Investigation of overtraining within Semi-Supervised Machine Learning Models in the search for heavy resonances at the LHC

Friday, 30 July 2021 12:45 (15 minutes)

When utilizing semi-supervised techniques in training machine learning models in the search for bosons at ATLAS, the overtraining of the model must be investigated. In particle physics internal fluctuations of the phase space and bias in training can cause semi-supervised models to label false signals within the phase space due to overfitting. The issue of false signal generation in semi-supervised models has not been fully analyzed and therefore utilizing a toy Monte Carlo model, the probability of such situations occurring can be quantified. This investigation of Z resonances is performed using a pure background Monte Carlo sample. Through unique pure background samples extracted to mimic ATLAS data in a background-plus-signal region, multiple runs enable the probability of these fake signals occurring due to overtraining to be thoroughly investigated.

Apply to be considered for a student ; award (Yes / No)?

Yes

Level for award;(Hons, MSc, PhD, N/A)?

PhD

Primary author: LIEBERMAN, Benjamin (University of Witwatersrand)

Co-authors: MELLADO, Bruce (University of the Witwatersrand); RUAN, XIFENG (University of the witwatersrand); LEBESE, Thabang (WITS University); Mr CHOMA, Joshua (University of the Witwatersrand); Dr DAHBI, Salah-Eddine (University of the Witwatersrand)

Presenter: LIEBERMAN, Benjamin (University of Witwatersrand)

Session Classification: Nuclear, Particle and Radiation Physics

Track Classification: Track B - Nuclear, Particle and Radiation Physics