

Abstract Submitted
for the APR21 Meeting of
The American Physical Society

Variable Importance: A Neural Network-Based Variable Ranking Framework in the CMS Search for Four-Top Production DANIEL LI, EMANUELE USAI, MEENAKSHI NARAIN, ULRICH HEINTZ, Brown University, CMS COLLABORATION — A variable ranking system utilizing dense neural networks trained for binary categorization is presented in the context of the search for four-top production in the single-lepton final state. The ranking framework, Variable Importance, defines the importance metric as the difference in the Area Under the Curve scores between two trained networks where one network has N inputs and the other has $N-1$ inputs. The quantity used for ranking is the significance, defined as the ratio between the mean and RMS of the importance metric distributions. Variable Importance is characterized by its large-scale computing approach and incorporates a method for accounting for correlation between variables. In addition to determining a ranking order, the Variable Importance framework includes steps for hyper parameter optimization and k -fold cross validation to obtain a final trained model. Tests on the predictive ability of the significance for model performance are shown as well as the resulting four-top, single-lepton final state simulation-based significance and limits derived from the discriminator of the final trained model.

Daniel Li
Brown University

Date submitted: 05 Jan 2021

Electronic form version 1.4