CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)

Milan Ganai

Project Number

S1807

Project Title

Achieving Improved Accuracy Model for Jet Energy Measurements in the Large Hadron Collider (LHC) Using Machine Learning

Abstract

Objectives/Goals

At LHC, the largest particle accelerator, the discovery of Higgs boson and confirmation of Higgs fields rattled the world in 2013, sealing a major unverified gap in SM. Energy measurements of jets produced by hadronization of quarks and gluons help to search for important rare physics processes beyond the Standard Model (SM). However, the pileup interactions cause inaccuracies in the measurements of jet mass and transverse momentum. I use machine learning (ML) to achieve improved accuracy in jet energy measurements by investigating more variables in the regression. My hypothesis is that ML will provide more accurate, robust, scalable and faster pileup mitigation than previous widely-used approaches at LHC.

Methods/Materials

The most used method at LHC today is area-based pileup mitigation, where a term proportional to jet area is subtracted from the measured jet transverse momentum and mass. Although it is an effective approach, it is inadequate in achieving the needed accuracy for next generation colliders due to its simplistic modeling. I propose a novel pileup mitigation methodology for jet energy measurements using TensorFlow-based machine learning. I investigated 11 most influencing parameters having high correlation with jet measurements and included them in regression training. The training set comprises ~350K full jets (with pileup) and hard jets (without pileup), generated using PYTHIA+FastJet software with varying center-of-mass energies and pileup contaminations.

Results

The error in transverse momentum using my approach was much smaller (mean = 0.53 and variance = 4.88) compared to the area-based approach (mean = -1.42, variance = 29.44). Similar results were obtained for jet mass measurements. Further, my approach is very robust, and performs better over a wider range of transverse momentum, ranging from 7GeV - 300 GeV. The proposed flow is also orders-of-magnitude faster than the area-based.

Conclusions/Discussion

I created and used a novel machine learning methodology to build a robust, scalable solution with better accuracy in predicting jet transverse and mass, thereby confirming and validating my hypothesis. Superiority of Deep Neural Network over linear regressors also shows that linear model used in previous area-based approach is inadequate and has limitation in achieving accuracy. Overall, my approach shows a great potential for exploring in next gen SLHC.

Summary Statement

I create and use a novel pileup mitigation methodology to improve accuracy in jet energy measurements by studying and using many variables in machine learning and achieve more accurate and faster solution than widely-used approaches at LHC

Help Received

I came up with the idea for the project, and developed the methodology and software flow, and analyzed the results. I presented it to my high school science teacher Mr. Leung, who encouraged me to pursue the project.