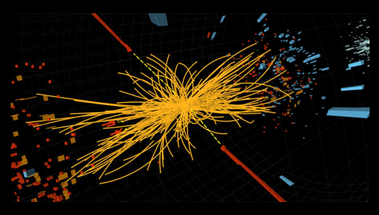
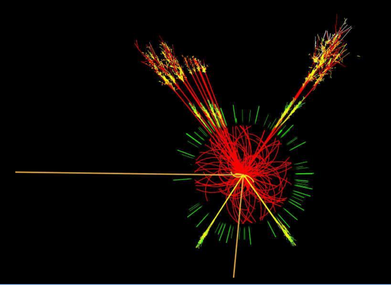
FERMILAB-POSTER-23-245-STUDENT

Higgs event classification using Machine Learning



Mysha Jahin Khan[1],
Supervisor: Dr. Pushpa Bhat[2]
1: Mount Holyoke College 2: Fermi National Accelerator Laboratory



Introduction

The *Higgs* boson is a fundamental particle in the Standard Model of particle physics, theorized in the 1960s and later discovered in 2012 at the Large Hadron Collider (LHC). It plays a pivotal role in the mechanism of mass generation for other Standard Model particles. This poster focuses on the application of Machine Learning algorithms to efficiently classify Higgs boson events.

Purpose

The goal of this project is to utilize machine learning techniques to effectively separate signal and background and distinguish different Higgs production signals in high-energy physics events.

Higgs Production and Decay

Higgs has many production modes such as Vector Boson Fusion (VBF), gluon-gluon fusion (ggF) etc. and different decay channels such as H→ZZ→4l, H→WW→2l2v, H→γγ etc. Where H→ZZ→4l has the largest signal to background ratio, where we can reconstruct the final state leptons.

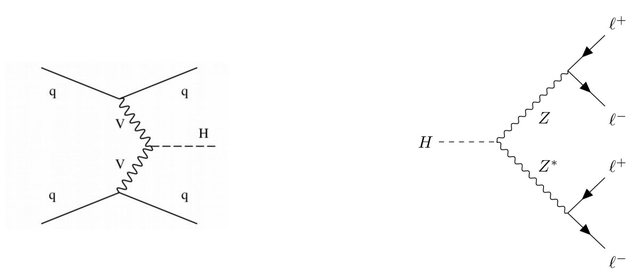


Figure: VBF Higgs production (left) H→ZZ→4l decay (right)

Machine Learning & Neural Networks

- Machine learning, a branch of artificial intelligence, enables computers to algorithmically learn and improve from data, excelling humans at tasks like pattern recognition and decision-making.
- Using interconnected layers, neural networks, one of the popular approaches in machine learning, extract complex patterns from high-dimensional data, enhancing signal-background separation and accuracy of measurements in high-energy physics.
- This approach provides valuable insights for studying the properties of the Higgs boson.

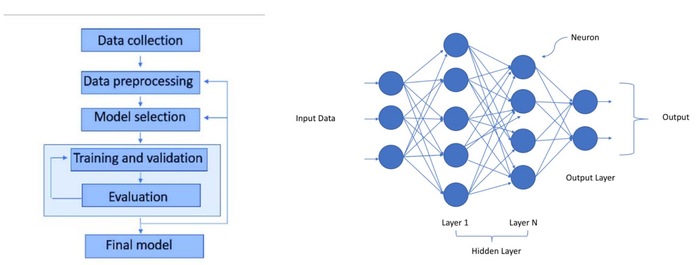


Figure: Machine learning analysis workflow (left); A neural network (right)

Data Splitting and Variables

In our Higgs event analysis using TMVA classification, we divided the data into two sets: training and testing samples. The training set was utilized to train and optimize the model's parameters, while the testing set assessed the learning accuracy in distinguishing signal/background, Higgs decay modes, ensuring reliable and insightful event classification.

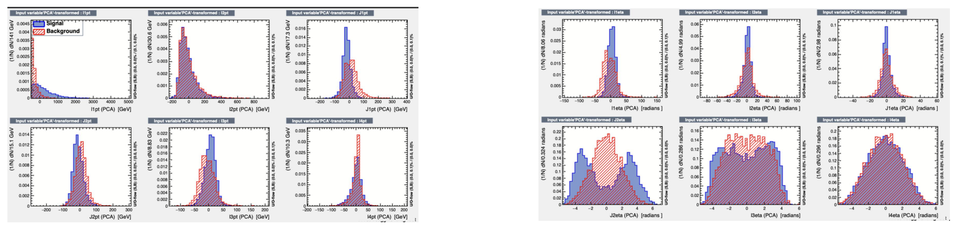


Figure: signal and background distribution of pt and eta for low level variables

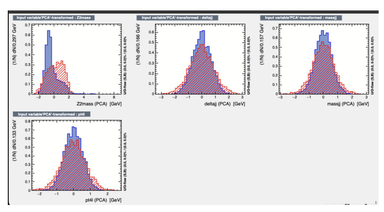


Figure: signal and background distributions low level variables

Training and Testing

Using Neural Networks with TMVA, our Higgs event classification study achieved enhanced accuracy by combining directly measured (low-level) and derived (high-level) variables This approach allowed us to distinguish the Higgs signal and background and different Higgs production (VBF from ggF) modes. The accuracy of distinction could also improve as we add more hidden layers, assuming large samples for training.

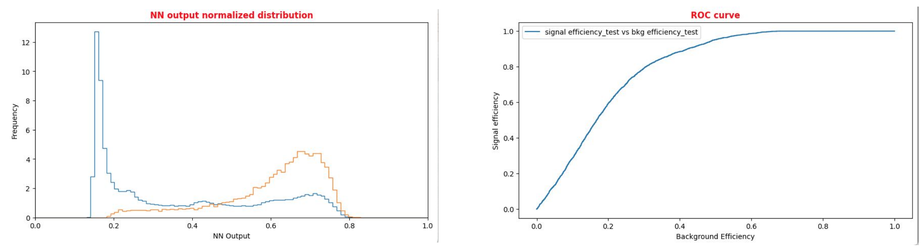


Figure: Sample Signal and Background (left) discrimination and ROC curve (right).

Conclusion

In conclusion, our study demonstrated the effectiveness of Machine Learning and Neural Networks with TMVA for efficiently classifying Higgs boson events. This data-driven approach led to improved accuracy and deeper insights into the properties of the Higgs boson.

Future Directions

In continuation of this project, we can also apply the knowledge gained to discriminate between quark-gluon jets to suppress backgrounds. Additionally, we can use the same approach to distinguish between two different Higgs signals, namely VBF and ggf.