Imperial College London

Using machine learning methods to measure Higgs boson properties

Keqi Zhang Supervisor: Dr Edward J T Scott High Energy Physics Group

INTRODUCTION

- Our current understanding of the fundamental particles and forces is encapsulated in the **Standard Model of Particle Physics**
- On 4th July 2012, the **Higgs boson** was observed by the ATLAS and CMS Collaborations at CERN's Large Hadron Collider [1]
- This project focuses on the precise measurements of the Higgs boson properties, notably the **vector boson fusion** (VBF) production mode and the Higgs **diphoton decay** channel

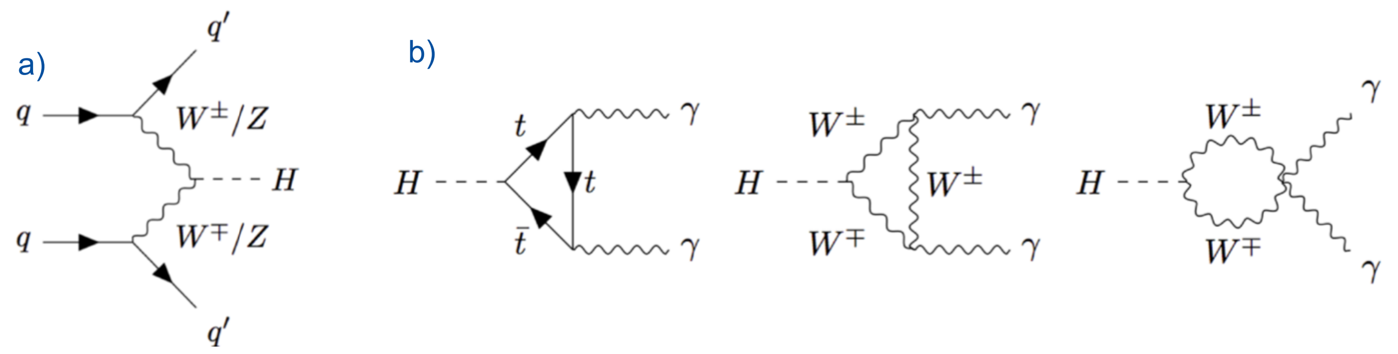


Figure 1. a) Feynman diagram of the vector boson fusion Higgs production mode b) Feynman diagrams of the $H \rightarrow \gamma\gamma$ decay channel.

This project aims to :

- Apply **Boosted Decision Tree** model and **Deep Neural Networks** models to distinguish the signal from the background events
- **Optimise** the classifiers and compare their performances
- Employ **statistical methods** to measure Higgs boson production in data selected using the classifier output score

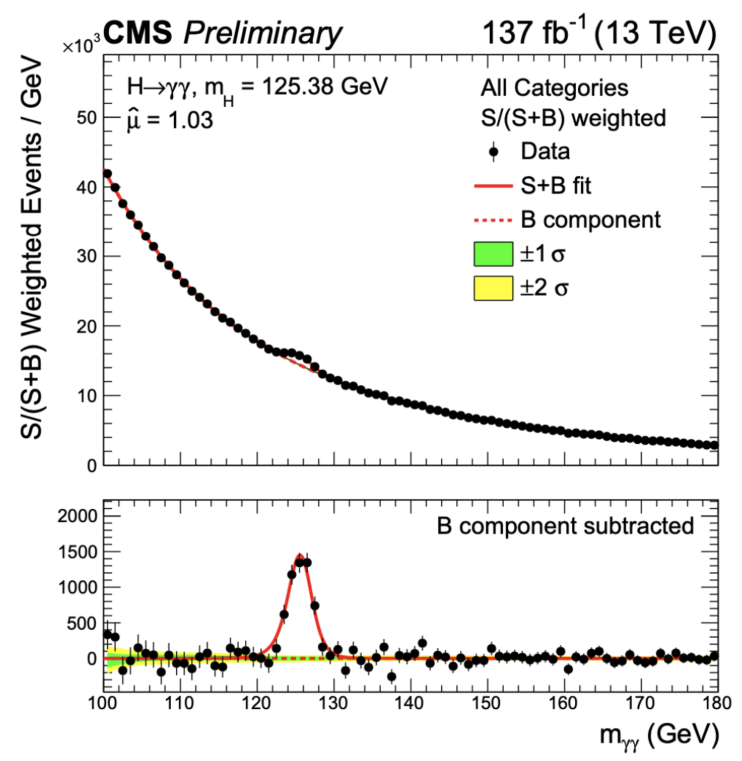


Figure 2. The diphoton invariant mass distribution for the selected data events (black points).

APPROACH

- The simulated **2016-2018 datasets** from the CMS are preprocessed and split into training and testing sets
- Machine learning models are trained using the input data and features
- The model is optimised by changing the **input features** and tuning the **hyperparameters**
- The output scores of the classifier are used to construct **analysis categories** which are used to perform a fit to the diphoton mass

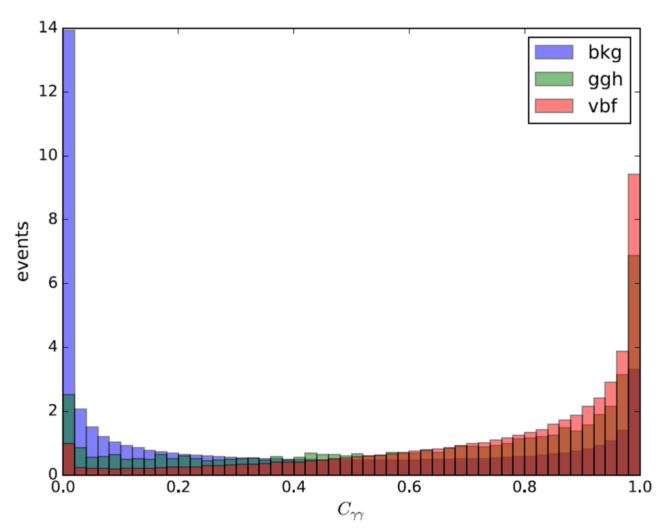
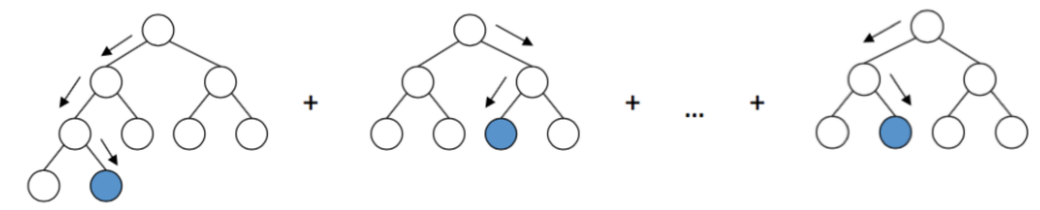


Figure 3. Histograms of the centrality variable for vector boson fusion, gluon fusion and background

MODELS

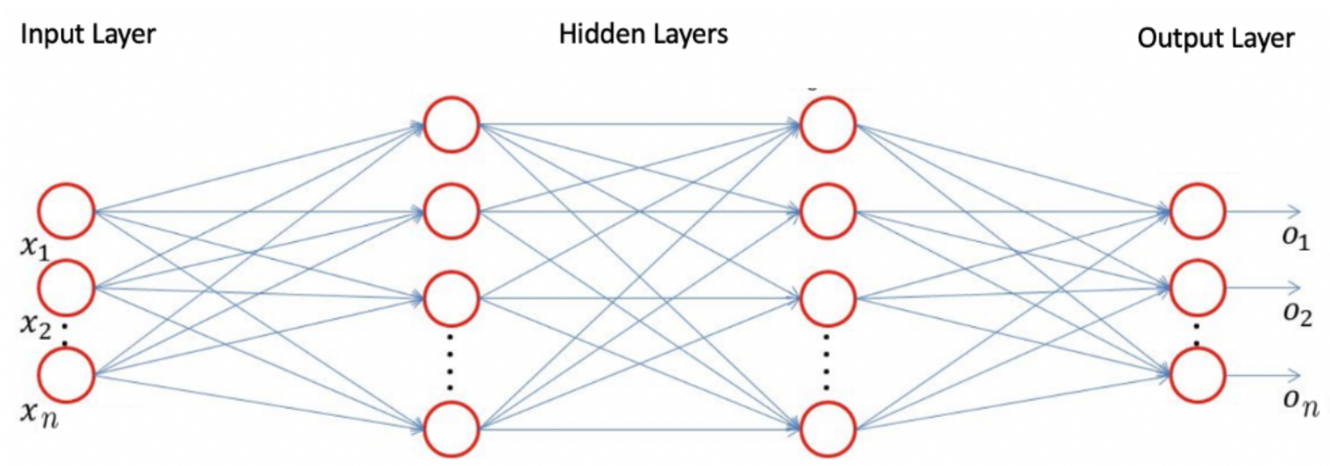
39 input features are used including individual jets variables and diphoton variables. The key hyperparameters are max depth, learning rate and subsample for the BDT and number of layers, nodes per hidden layers and dropout rate for the MLP

1. Boosted Decision Trees (BDT)



BDT combines individual **decision trees**, which takes a set of input features and splits the input data recursively, into a strong classifier with **gradient boosting** [2]. The goal is to minimise the objective function which measures how well the model fits the training data.

2. Multilayer Perceptron Neural Networks (MLP)



MLP trains by updating the weights and the biases of the neurons, using **backpropagation** method, to minimise the objective function.

PRELIMINARY RESULTS

Model	Train Acc.	Test Acc.
BDT baseline	0.915	0.911
BDT with jets vars	0.923	0.917
BDT with jets vars and optimal config.	0.923	0.920
MLP baseline	0.910	0.909
MLP with jets vars	0.926	0.923
MLP with jets vars and optimal config.	0.928	0.927

Table 1. Different models with the train and test accuracy.

The Receiver Operating Characteristic curves of the best BDT and MLP models are:

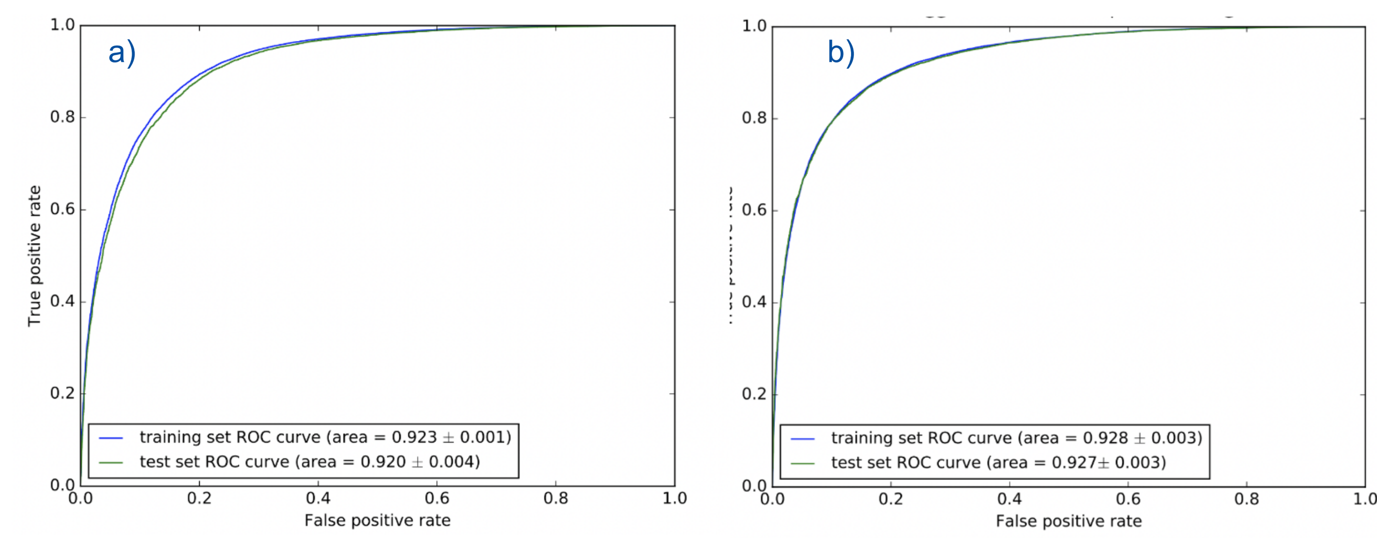


Figure 4. a) ROC curve for BDT with jets variables and optimal configurations of hyperparameters b) ROC curve for MLP with jets variables and optimal configurations

Discussion:

- Including the **jets variables** gives better accuracy for both methods
- BDT test accuracy increases by 0.87 % from baseline, which corresponds to a **17% reduction in background** events for a true positive rate of 70%
- MLP test accuracy increases by 1.98 % from baseline with the jets variables and optimal configuration, which means a **7% reduction in background** events for a true positive rate of 70%
- Using the MLP model gives more than **10%** reduction in background than BDT for the chosen true positive rate

CONCLUSION

- Two machine learning methods are trained and optimised to distinguish between signal and background events
- **Deep learning** increases the sensitivity to background events by more than 10% compared to BDT, showing motivation of using DNN in the Run 3 analysis at the CMS
- Next step: implement **LSTM** neural networks and use **statistical methods** to extract the final diphoton mass distribution

REFERENCE

[1] CMS Collaborations, “Measurements of higgs boson properties in the diphoton decay channel in proton-proton collisions at√s=13 tev,” Journal of High Energy Physics,vol. 2018, no. 11, Nov. 2018,issn:1029-8479.doi:10.1007/jhep11(2018)185.

[2] T. Hastie, R. Tibshirani, and J. Friedman. The Elements of Statistical Learning. Springer, 2009. isbn: 9780387848587.