Machine learning-based optimisation of Higgs coupling measurements in the $H \rightarrow 4l$ decay channel with ATLAS Run 3 data

Luca Spitzauer

The Standard Model of Particle Physics

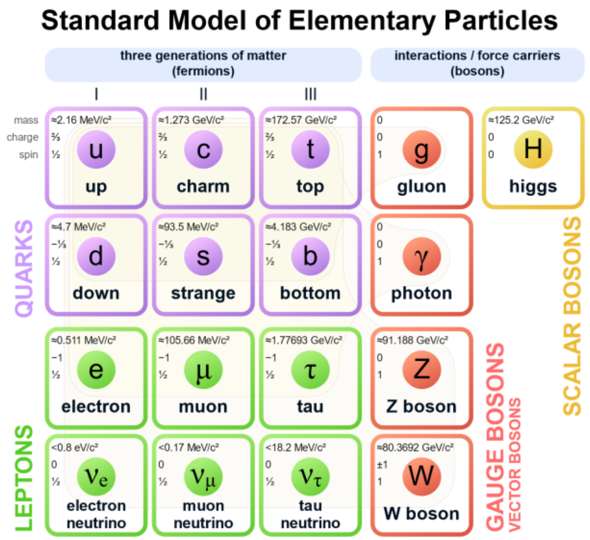
Last discovered elementary particle (July 2012 at the Large Hadron Collider):

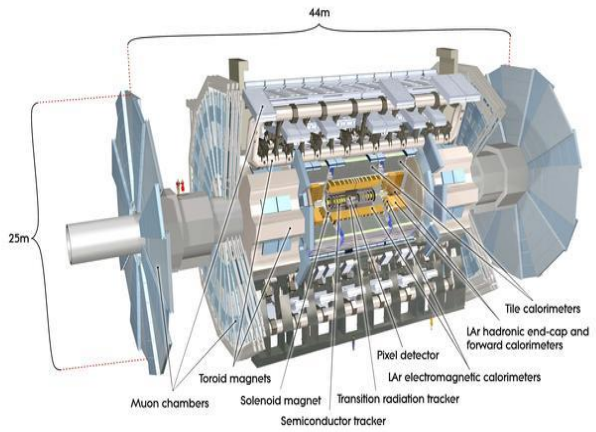
Higgs Boson – the only spin-0 particle

Is the Standard Model enough?
Several observations (dark matter, baryon asymmetry,...)

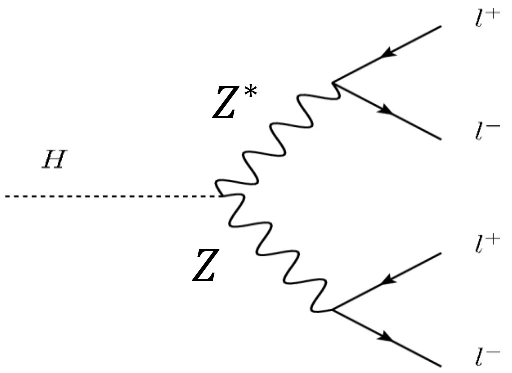
Measurements of Higgs couplings are an important SM test and potential window into BSM physics

Higgs precision measurements at the LHC
with the multi-purpose detectors ATLAS and CMS

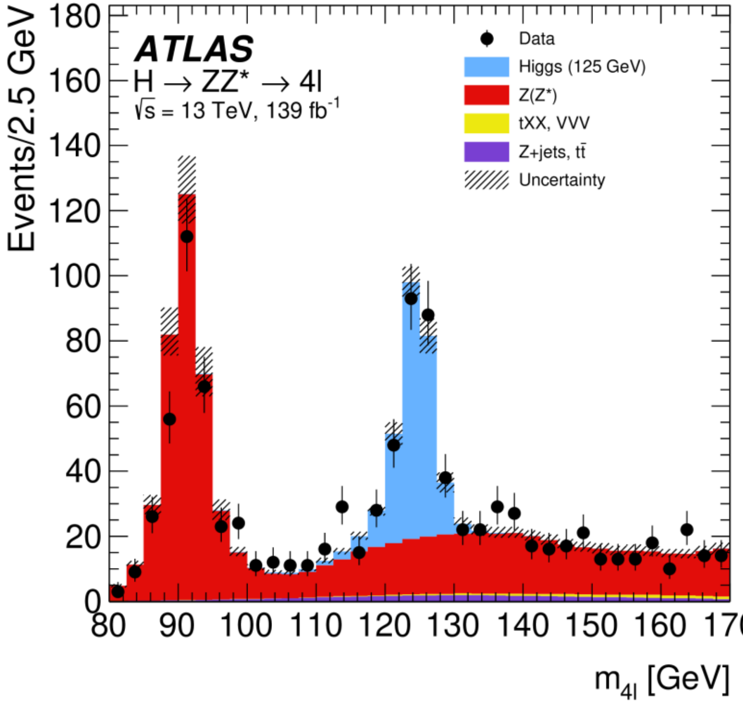




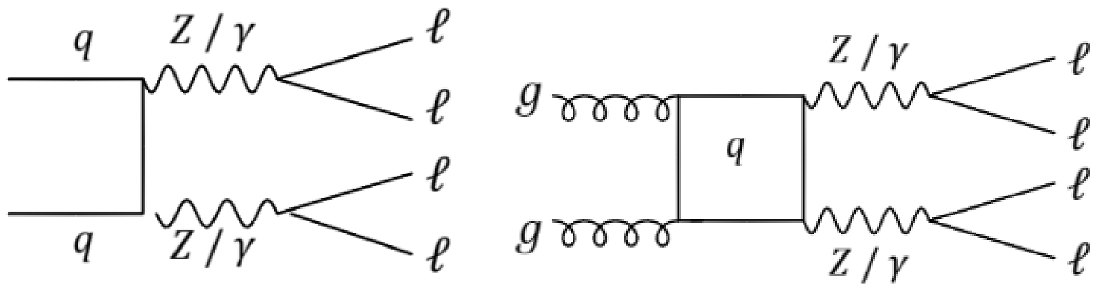
H → ZZ* → 4l The Golden Channel



- Higgs Boson decays into an on-sell and off-shell Z Boson pair
- Leptonic Z decay into electrons and muons of interest
- Final states: 2e2μ, 2μ2e, 4e, 4μ

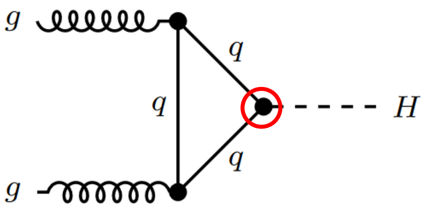


Dominant ZZ Background

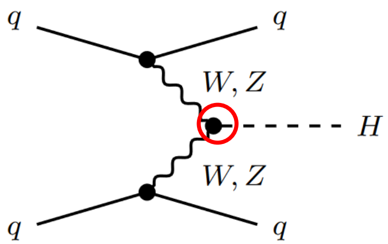


H → ZZ* → 4l Production

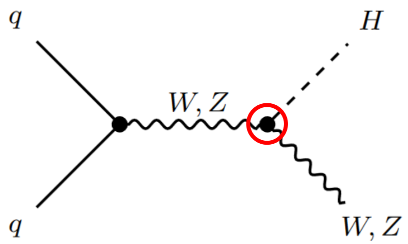
- **Gluon-Gluon-Fusion (ggF) (87%)**



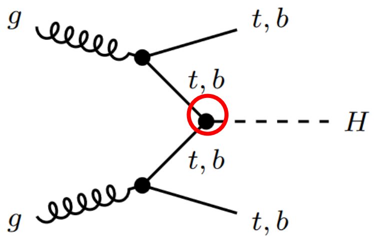
- **Vector-Boson-Fusion (VBF) (7%)**



- **Associated Vector Boson (VH) (4%)**



- **Associated Heavy Quark (ttH/bbH) (2%)**



➢ Cross Section measurements of different Higgs Production Modes important to determine Higgs coupling

H → ZZ* → 4l Event Categorization

- Production processes can be further separated in exclusive phase space regions to improve BSM sensitivity

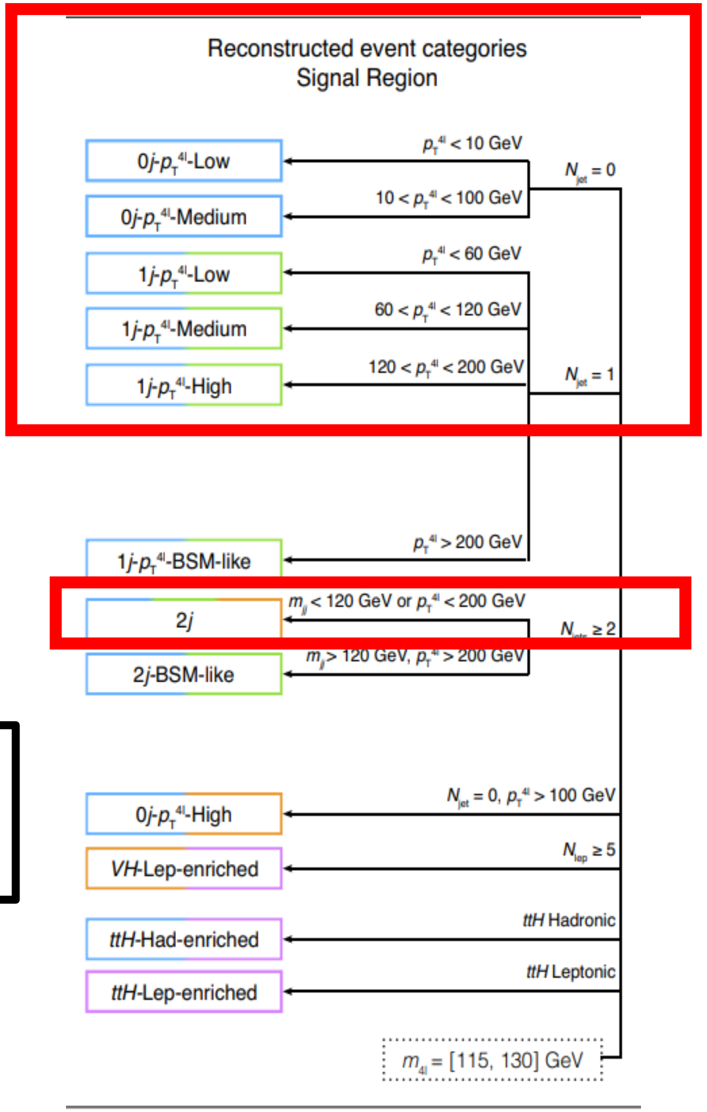
- Reconstruced candidate events classified into event categories based on $p_T^{4l}, m_{jj}, n_{jets}$

Challenge:

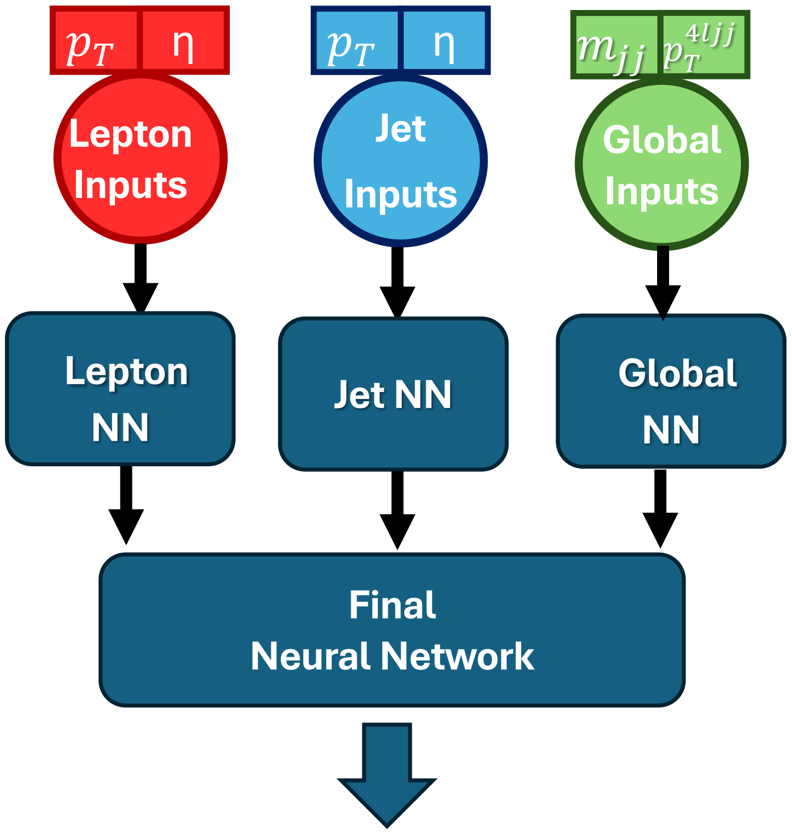
Each event category contains events from different Higgs production modes and ZZ* background

Higgs Production modes:
ggF, VBF, VH, ttH

Separate a given targeted Higgs production mode from other contaminating processes using Neural Networks



Neural Network model for H → 4l analysis



Input variables are separated into 3 groups:

- Lepton variables (p_T^{lep}, η^{lep})
- Jet variables (p_T^{jet}, η^{jet})
- Global event variables (e.g. m_{jj}, p_T^{4l})

➢Different NNs trained separately for each input group

Jet and Lepton Neural Network architectures

NN architectures are chosen that „loop" over each individual lepton or jet iteratively

NN architecture previously used for jets and leptons:

- **Recurrent Neural Network (RNN)**

➢ Data treated as ordered sets

➢ Information from one „loop" passed to next „loop" as input

Alternative architecture investigated in the thesis:

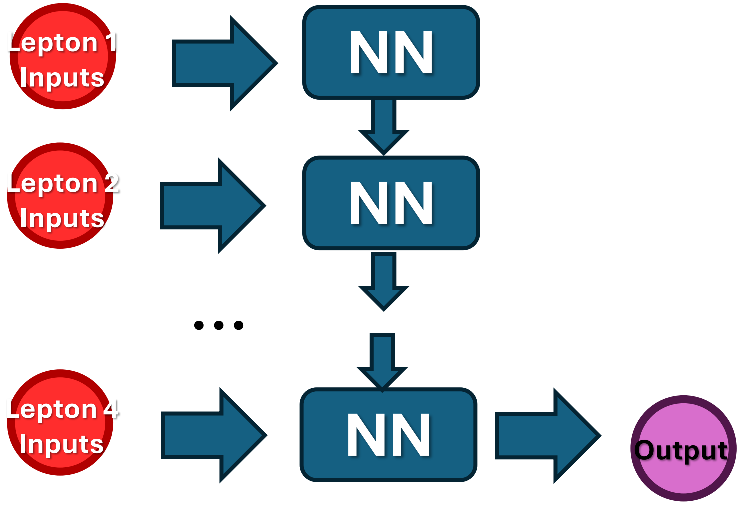
- **Deep Set Neural Network**

➢ Data treated as unordered sets

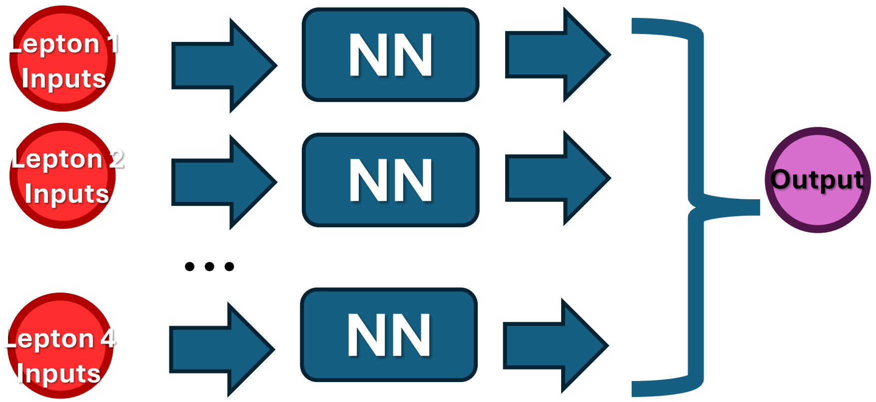
➢ Separate outputs from all „loops" pooled into one outout

Architecure schematics for the lepton information:

RNN for one Higgs event candidate:



Deep Set for one Higgs event candidate:



Overview of Deep Set input variables

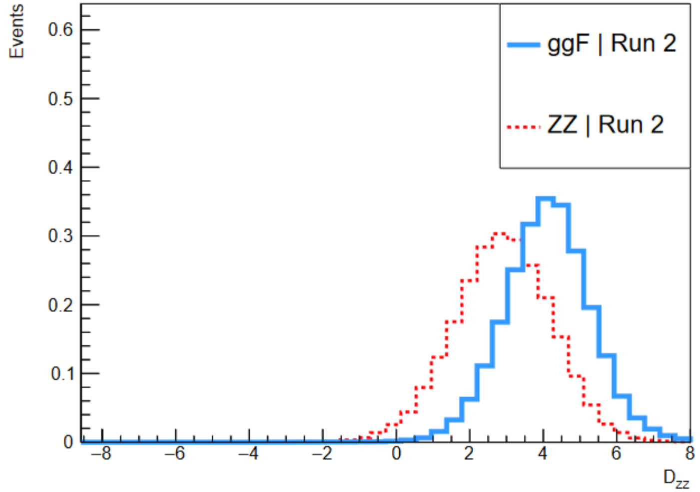
Event Category	Contributing Processes (Targeted Process in red)	Global Variables	Lepton Variables (Variables in brackets not used in RNN training)	Jet Variables
0 Jet 2e2μ η_{jets} = 0	**ggF**, ZZ	$p_T^{4l}, m_{12}, m_{34}, D_{ZZ}, \cos\theta^*, \cos\theta_1, \varphi_{ZZ}$	p_T, η, (Z association)	-
0 Jet 4l η_{jets} = 0	**ggF,** ZZ	$p_T^{4l}, m_{12}, m_{34}, D_{ZZ}, \cos\theta^*, \cos\theta_1, \varphi_{ZZ}$	p_T, η, (Z association)	-
1 Jet Low η_{jets} = 1, $p_T^{4l} <$ 60 GeV	**VBF,ggF,** ZZ	$p_T^{4l}, p_T^j, \eta_j, \Delta R_{4lj}, D_{ZZ}$	p_T, η, (Z association)	-
1 Jet Medium η_{jets} = 1, p_T^{4l} > 60 GeV	**VBF,ggF,** ZZ	$p_T^{4l}, p_T^j, \eta_j, \Delta R_{4lj}, D_{ZZ}, E_T^{Miss}, \eta_{4l},$	p_T, η, (Z association)	-
2 Jet Low η_{jets} = 2, $m_{jj} <$ 120 GeV or $p_T^{4l} <$ 200 GeV	**VH**,ggF, VBF	m_{jj}, p_T^{4ljj}	p_T, η, (Z association)	p_T, η
2 Jet High $m_{jj} >$ 120 GeV and $p_T^{4l} >$ 200 GeV	**VBF**, ggF	$\eta_{ZZ}^{Zepp}, p_T^{4ljj}$	p_T, η, (Z association)	p_T, η

Overview of Deep Set input variables

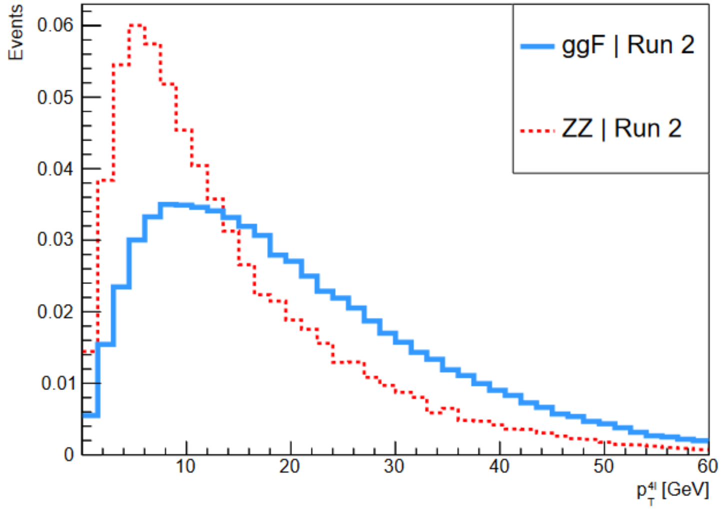
Event Category	Contributing Processes (Targeted Process in red)	Global Variables	Lepton Variables (Variables in brackets not used in RNN training)	Jet Variables
0 Jet 2e2μ η_{jets} = 0	**ggF**, ZZ	$p_T^{4l}, m_{12}, m_{34}, D_{ZZ}, \cos\theta^*, \cos\theta_1, \varphi_{ZZ}$	p_T, η, (Z association)	-
0 Jet 4l η_{jets} = 0	**ggF,** ZZ	$p_T^{4l}, m_{12}, m_{34}, D_{ZZ}, \cos\theta^*, \cos\theta_1, \varphi_{ZZ}$	p_T, η, (Z association)	-
1 Jet Low $\eta_{jets} = 1, p_T^{4l} < 60$ GeV	**VBF,ggF,** ZZ	$p_T^{4l}, p_T^{j}, \eta_j, \Delta R_{4lj}, D_{ZZ}$	p_T, η, (Z association)	-
1 Jet Medium $\eta_{jets} = 1, p_T^{4l} > 60$ GeV	**VBF,ggF,** ZZ	$p_T^{4l}, p_T^{j}, \eta_j, \Delta R_{4lj}, D_{ZZ}, E_T^{Miss}, \eta_{4l},$	p_T, η, (Z association)	-
2 Jet Low $\eta_{jets} = 2,$ $m_{jj} < 120$ GeV or $p_T^{4l} < 200$ GeV	**VH**,ggF, VBF	m_{jj}, p_T^{4ljj}	p_T, η, (Z association)	p_T, η
2 Jet High $m_{jj} > 120$ GeV and $p_T^{4l} > 200$ GeV	**VBF**, ggF	$\eta_{ZZ}^{Zepp}, p_T^{4ljj}$	p_T, η, (Z association)	p_T, η

Important Kinematic Variables (0 –Jet category)

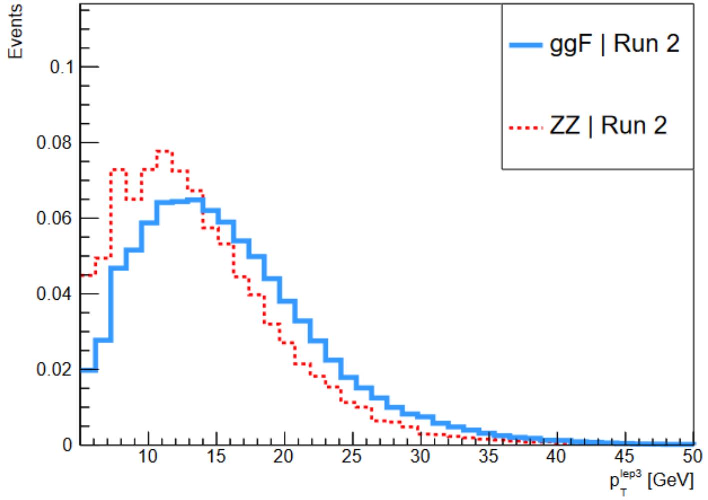
Matrix-element based variable D_{ZZ}



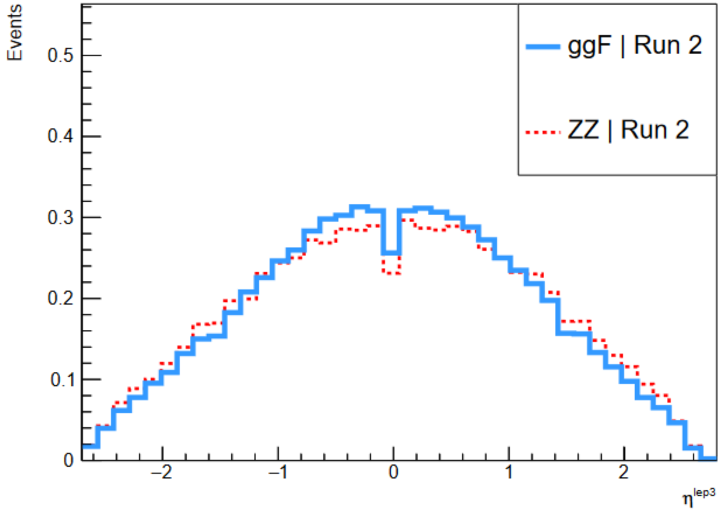
p_T of 4-lepton system



p_T of leptons from off-shell Z boson

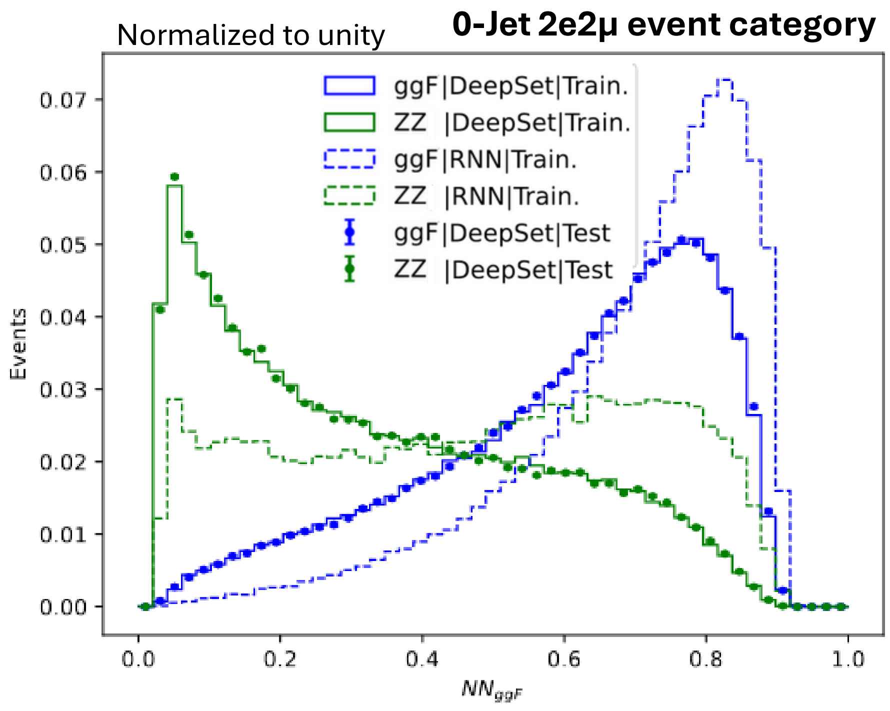


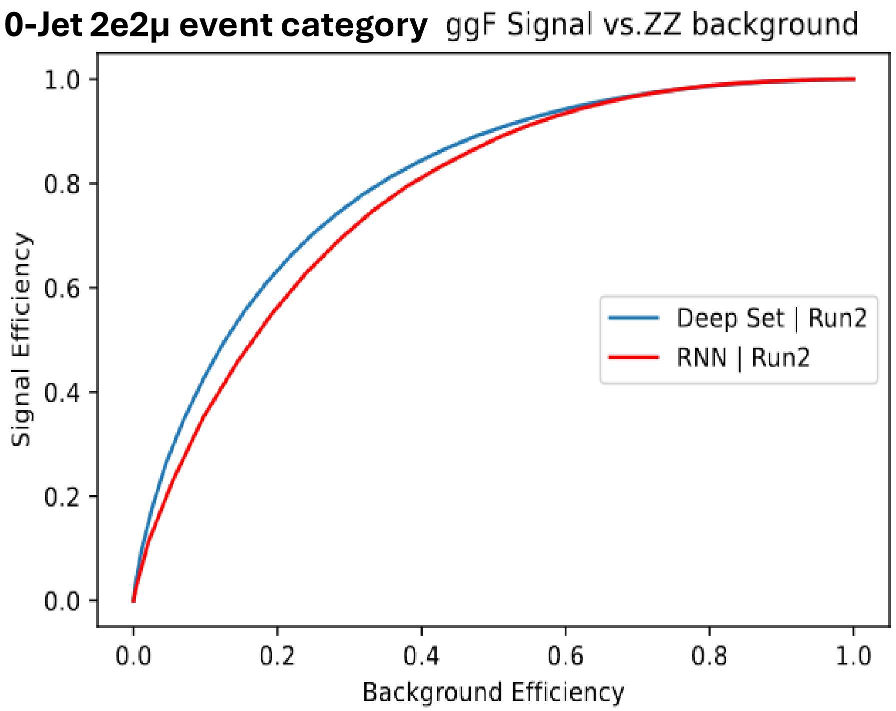
η of leptons from off-shell Z boson



- Clear differences between ggF signal and ZZ background in several kinematic distributions
- Provide strong discrimination power to the Neural Network

Comparison of Deep Sets and RNN Outputs





- Deep Set model generally provides a better indentification of events from the dominant production mode compared to RNN
- The output distributions from the test data-sample agree with those from the training sample, ensuring that there is no overtraining

Overview of Deep Set Performance

- Deep Set / RNN Comparison

Signal Significance:

$$\frac{N_{signal}}{\sqrt{N_{background}}}$$

Event Category	Contributing Processes	Targeted Process	Deep Set/RNN Signal Significance Ratio (at 60% Signal Efficiency)
$0-Jet\ 2e2\mu$	ggF, ZZ	ggF	1.11
$0-Jet\ 4l$	ggF, ZZ	ggF	1.09
$1-Jet$ low-p_T^{4l}	VBF, ggF, ZZ	VBF	1.01
		ggF	1.10
$1-Jet$ medium-p_T^{4l}	VBF, ggF, ZZ	VBF	1.01
		ggF	1.15
$2-Jet$ low-p_T^{4l}	VH, VBF, ggF	VH	1.07
$2-Jet$ high-p_T^{4l}	VBF, ggF	VBF	1.01

- Improved signal significance up to 15% is observed for the Deep Set Neural Network compared to the RNN in all event categories

Summary

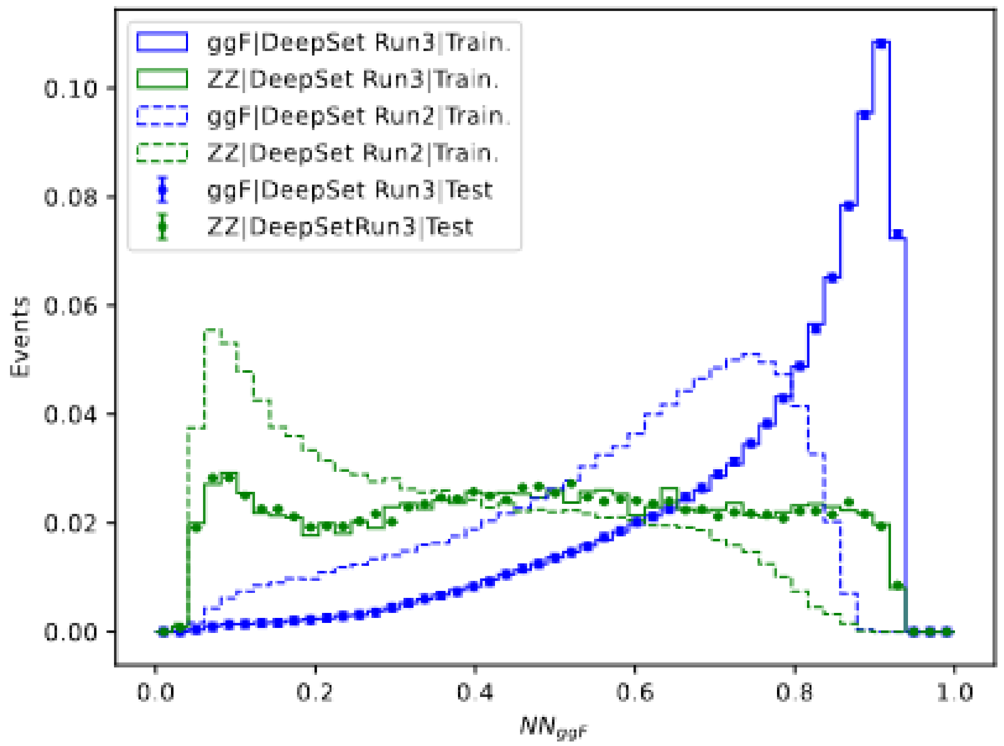
- Deep Set Neural Network architecture introduced and optimized as an alternative to the previously employed RNN model for the classification of Higgs production modes
- Improved performance compared to RNN baseline achieved with the Deep Set Neural Network, allowing for more precise cross section measurement results

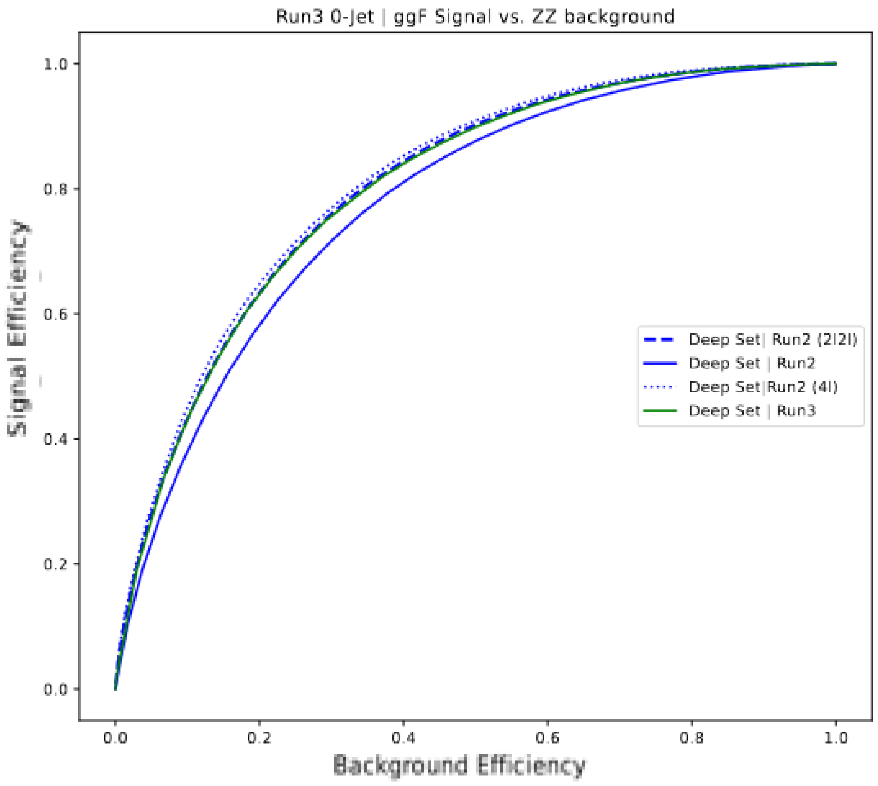
Possible Future Developments:

- Implemention of Deep Set Neural Network models in the official ATLAS $H \rightarrow 4l$ analysis framework
- Optimization with additional extensions to the Deep Set architecture

Backup

Comparison of Deep Set Run 2/Run 3 Outputs





- Comparison of Deep Set output of models trained on Run 2 and Run 3 data
- Deep Set models trained on Run 3 data generally show more confident indentification of events from dominant production mode in all event categories

Overview of Deep Set Run2 vs. Run 3 Performance

- Deep Set Run2/Run3 Comparison

Event Category	Contributing Processes	Targeted Process	Run 3/Run 2 Signal Significance Ratio (at 60% Signal Efficiency)
$0-Jet$	ggF, ZZ	ggF	1.09
$1-Jet$	VBF, ggF, ZZ	VBF ggF	1.00 1.12
$2-Jet$ low-p_T^{4l}	VH, VBF, ggF	VH	1.18
$2-Jet$ high-p_T^{4l}	VBF, ggF	VBF	1.04

Signal Significance:

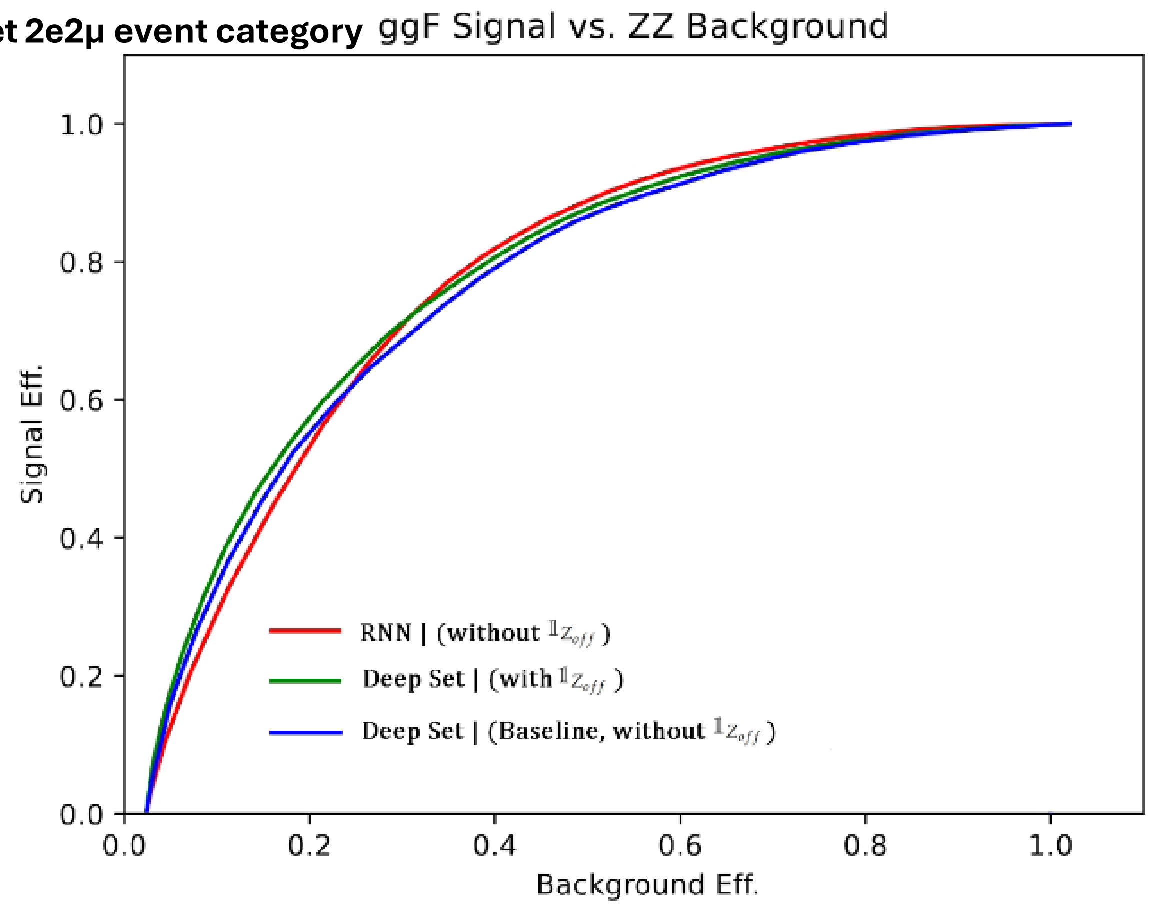
$$\frac{N_{signal}}{\sqrt{N_{background}}}$$

- Equal or improved up to 12% signal significancies are observed for the Deep Set Neural Networks trained on Run 3 data compared to Run 2 data

Impact of Z Association as additional input variable

Inclusion of the Z Association as input variable of the Lepton Deep Set yielded improved signal-to-background separation

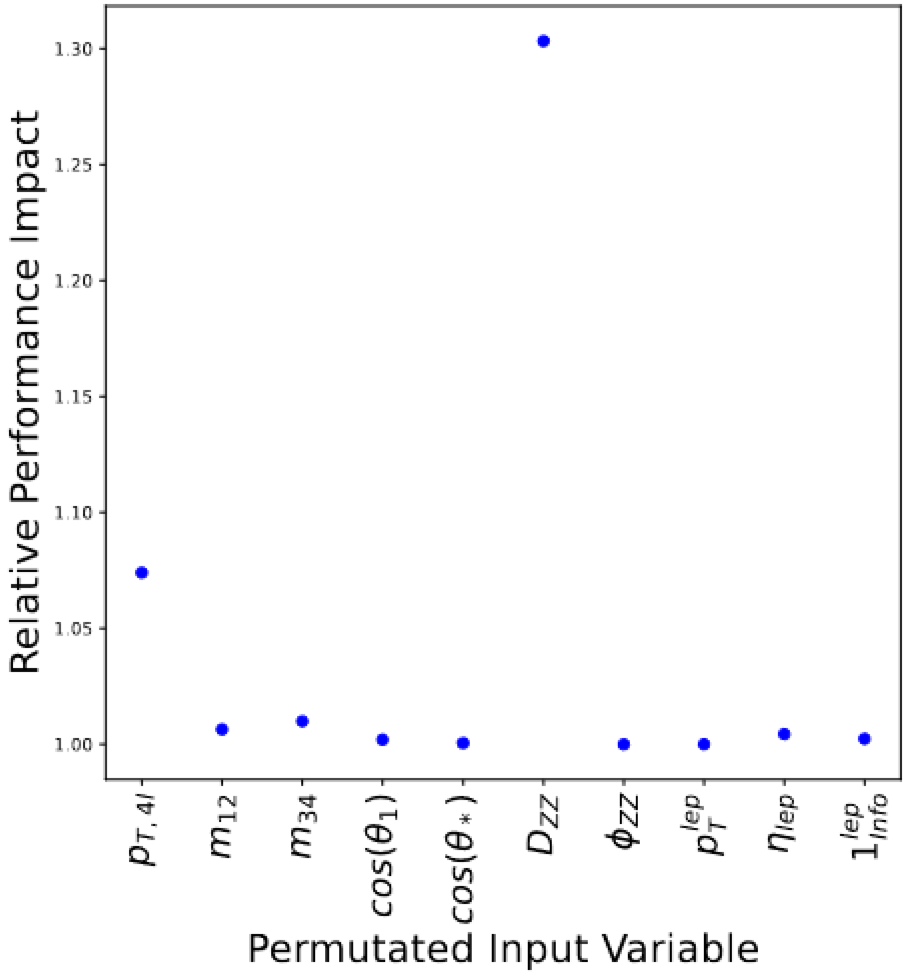
0-Jet 2e2μ event category



Importance Ranking of input variables (0-Jet category)

Impact of input variables on the Deep Set output was quantified by comparing the loss before and after shuffling the input variable data

Impact is dominated by the matrix-element based D_{ZZ} variable



Full Event Categorization Scheme