Master project description

William Hirst

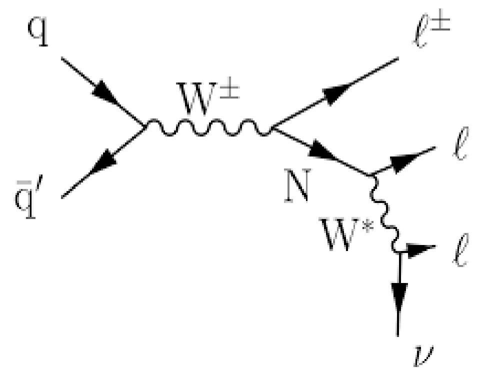
TITLE: Expanding the standard model through the search of a heavy neutrino using supervised machine learning on data from the ATLAS detector at the LHC

Motivation

Every year technology for generating and measuring particle collisions is improved. As a consequence, the amount of data increases drastically. The ATLAS experiment is one of the largest particle detector experiments currently operating at the CERN laboratory near Geneva. ATLAS alone generates approximately 1 petabyte of raw data every second from proton proton collisions at the Large Hadron Collider (LHC). With amounts of data this large, manually analysing the data ourselves would take more time than anyone would like to spend. So-called triggers reduce the data to be analysed somewhat, although physicists would like to take advantage of as much information as possible. However, the amount of data and its complexity will increase drastically in the following years, in particular the High Luminosity LHC towards 2030. Therefore, taking advantage of numerical methods like machine learning is pivotal if scientific development is to keep up with technological development.

This thesis will aim to search for heavy neutral neutrinos in data, produced by the LHC and collected by the ATLAS detector, using supervised machine learning. The physics motivation is the search for hypothetical heavy neutrinos introduced in an attempt to explain the tininess of the observed neutrinos compared to the charged leptons and quarks. The heavy neutrinos could stem from the decay of a hypothetical right-handed W'_R gauge boson, and could also be discovered and measured. In addition to searching for new physics, the thesis will aim to study numerical methods of analysing large sets of data and comparing how different degrees of bias can affect the result of the analysis.

Physics



The Standard Model, the current theory of particle physics, so far explains nearly all experimental measurements. The electroweak part is based on the Glashow Salam Weinberg model based on the symmetry group $U(1)_Y xSU(2)_L$ which unifies the electromagnetic and weak interactions at high energies. The electromagnetic interactions are mediated by the massless photon and the weak interaction by the W^+, W^- and Z gauge bosons. Two features of the SM are that the neutrino masses are assumed to be zero, and the $W^\pm$ boson only couples to left-handed particles. The right-handed neutrino is thus sterile and has no interactions. Another mystery related to massive neutrinos is that they can be their own antiparticle, hence of type Majorana, or neutrino and antineutrino are different particles, hence of type Dirac. We however know from the neutrino oscillation experiments that neutrinos oscillate and hence have tiny but non-zero masses.

The Feynman diagram above sketches how a hypothetical heavy neutrino, *N*, is produced, in proton-proton collisions, through a real $W^{\pm}$ and how *N* further decays to two leptons and a neutrino via a virtual $W^{\pm}$ boson, see reference [1]. Through the search for the heavy neutral neutrino we will also be looking for a new boson, in which case the W and W* in the diagram should be replaced by $W^{\pm}_R$ and $W^{\pm *}_R$. The W'_R is introduced in an attempt to restore parity symmetry in the electroweak interactions, thus extending the SM electroweak symmetry group $U(1)_Y xSU(2)_L$ with a new symmetry group $SU(2)_R$. The decay of the $W^{\pm}_R$ would lead to a right-handed heavy neutrino N_R. It will also be possible to find out whether the heavy neutrino is of type Majorana or Dirac by studying the signs of the leptons to be measured.

As this thesis will mainly focus on supervised learning, a specific physical model must be chosen to train the network. The models will focus on final states with three leptons and missing transverse energy E_T. The feynman diagram shown in the figure above hides several cases depending on the mass of the heavy neutrino and the charges of the leptons measured. The figure shows a quark-antiquark pair annihilation into a W-boson, which decays into a charged lepton and a heavy neutrino. The heavy neutrino decays into charged lepton and a W boson which finally decays into a charged lepton and a neutrino.

There are several cases to study:

- If the mass of the heavy neutrino m_N is lower than the mass of the W it stems from, that W is real and the heavy neutrino will decay through a virtual W*. if m_N is larger than m_W, the first W is virtual and the second can be real. This is valid whether we consider the standard model $W^{\pm}$ or the new $W^{\pm}_R$. the momenta of the charged leptons and missing E_T measured will be affected, as well as the invariant masses involved.
- It is interesting to study whether the heavy neutrino is of type Majorana or Dirac. We need to find out whether the signs of lepton from the first W and the lepton from the *N* decay are the same or opposite, and whether the lepton flavours are the same or different. The first two leptons are always of opposite sign if N is of type Dirac. If N is of type Majorana, both charge configurations can occur.

Among the features (measured variables) to consider we can list the momenta, flavours and charges of the three leptons, the missing transverse energy E_T, the invariant mass of the 3 leptons, the invariant mass M(lll+neutrino), the neutrino being represented by the missing E_T, the so-call transverse mass, which is used instead of the invariant mass in cases where one particle is not reconstructed (such as the neutrino). More features will be evaluated and studied.

Supervised learning methods

For the past couple of decades, machine learning has made its way into many if not all branches of science. Its ability to analyse large sets of data and use it to make real life predictions has led to many great scientific strides. In my thesis I will use a branch of machine learning called supervised learning and apply it to particle collisions in search of new physics, going beyond the standard model. The aspect which makes the learning supervised, is that the methods are dependent on so called labeled data. The most popular methods for supervised learning on the data produced in ATLAS are Gradient Boosted Decision Trees and Deep Neural Networks.

Decision trees focus mainly on creating efficient ways of splitting up the data into categories and then creating simple tests of placing the data in those categories. Gradient boosting deploys an ensemble of such decision trees to create a much stronger statistical learner. Neural networks are not as simple. They are networks of so called layers and nodes, which propagate the data backwards and forward through the network using weights. The network is trained by finely tuning each weight to fit the data it is given. The neural network can be altered by varying many hyperparameters in the network such as the learning rate, the regularisation constant or the momentum. Searching for the ideal value of these parameters will also be a large part of the thesis.

To begin with I will focus on one or both of the methods described above, but given sufficient time alternative methods could also be studied. The methods will be implemented using the software Tensor Flow (alternatively keras). In practice the analysis and most of the methods will be developed on simulated Monte Carlo simulations, where we can directly study theories that go beyond the standard model. The hope is to extend the methods to analyse ATLAS Run2 data and possibly the new data being produced in 2022.

As mentioned in the physics section we will be focusing on the three lepton final state event. Choosing only a subset of the entirety of the ATLAS data such as this implies that we believe that this is the area most beneficial to the search. Given that such a choice will insert our own bias into the system it might weaken the analysis. As this thesis will be focusing on comparing different machine learning algorithms, we could be interested in seeing how analysing larger sets of data could affect any or all findings.

The final part of the thesis is to compare any and all results produced by the supervised learning methods to those produced by Sakarias Frette, another master student at CCSE, who will focus on unsupervised learning. The hope is that comparison could highlight strengths and weaknesses in each method. It will be especially interesting to study how unsupervised learning fares, as this is the method which requires the least amount of bias and preconceived ideas on the subjects we are studying.

Thesis plan

- 2021/H
  - Theory pensum: IN5270, FYS-STK4155, FYS4170
- 2022/V
  - Theory pensum: IN4200, FYS5555, special pensum
  - Begin reading on supervised learning detection used in particle physics
  - Study neutrino physics, and read about heavy neutrino production at the LHC, as well as the competing SM processes
- 2022/H
  - prepare the MC and real data files containing 3 leptons and convert them into numpy arrays or similar
  - define the variables / features and prepare the supervised learning
  - study and tune hyperparameters
  - start training the network and compare Gradient Boosted Decision Trees and Deep Neural Networks
  - Start writing part of thesis
- 2023/V
  - apply on SM MC and on Run 2 ATLAS data and possibly on a part of Run 3
  - Interpretation of the results
  - compare results with unsupervised method of Sakarias
  - Present results locally and in ATLAS group meeting
  - Finish writing thesis, deliver by 15.5, defend thesis mid June

References

[1] Search for heavy Dirac and Majorana neutrinos in tri-leptons at the LHC , C.O. Dib, C.S. Kim, K. Wang, Phys. Rev. D **95**, 115020, DOI: 10.1103/PhysRevD.95.115020
[2] P. Baldi, P. Sadowski & D. Whiteson, DOI: 10.1038/ncomms5308, Searching for exotic particles in high energy collisions with deep learning

Supervisor(s) signature: Prof. Farid Ould-Saada
30.11.2021