Applications of Kolmogorov-Arnold Network classifiers in Particle Physics for Exotic Particle Detection

Ali Al-Shahristani [1] Dominic De Jonge [1] Jakub Skop [1] Kshitij Tyagi [1]

[1]Watford Grammar School for Boys

Abstract

At the Large Hadron Collider (LHC), protons collide at high energies, initiating complex interactions between their constituent particles—quarks and gluons. These interactions result in the release of energy, which can be used in the creation of particles, including electrons. These collisions recreate conditions in the universe about a billionth of a second after the universe came into being. One of the detectors at the LHC is the ATLAS(A Toroidal LHC Apparatus). The ATLAS detector is a versatile particle detector characterized by a symmetric cylindrical design that covers nearly the entire 4π solid angle. This detector comprises an inner tracking system enclosed by a slender superconducting solenoid generating a 2T axial magnetic field. It is designed to take 40 million pictures per second of proton-proton collisions. The LHC incorporates both electromagnetic and hadronic calorimeters, as well as a muon spectrometer. It is crucial for detecting and measuring electrons generated in proton-proton (pp) collisions. It employs components like the Inner Detector to track charged particles, the Electromagnetic Calorimeter to measure their energy, a Trigger System for event selection, and a Magnetic Field to determine their momentum. However, one of the problems remains that particle interactions occur at a very high frequency and in great volume, so data analysis of these interactions to find interesting properties has to be done by computers. In this avenue, there are two main methods, the first of which is applying cuts - removing data points which are obviously not what we are looking for - with approximate values. However, these cuts tend to neglect deeper connections which exist and so the alternative is using AI/ML classifiers as they maintain a high repeatability while being able to obtain a relationship not apparent to humans. In the present paper, we investigate the applications of a new type of classifier called KANs which have the added advantage of being interpretable and so the deeper relationship could be understood by humans.

Kolmogorov-Arnold Networks

Kolmogorov-Arnold Networks (KANs) are a new alternative to Multi-Layer Perceptron (MLP) Neural Networks. While MLPs have fixed activation functions on nodes ("neurons"), KANs have learnable activation functions on edges ("weights"). In fact KANs have no linear weights as all weight parameters are replaced by univariate functions parameterized as splines. A spline is just a function defined piecewise by polynomials and serve as the "next level up" from simply just linear functions.

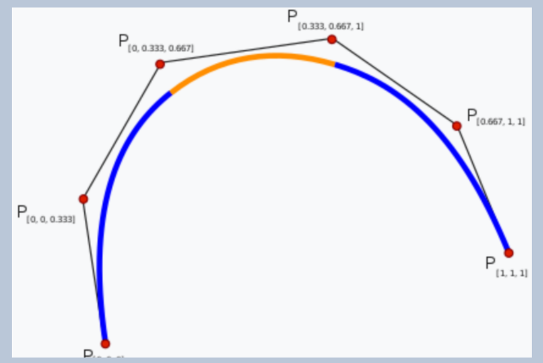

Figure 1. An Example of a spline

Higgs Boson

Below temperatures of $\sim 159\ GeV/k_B$, the electroweak interaction undergoes a spontaneous symmetry breaking, splitting into the electromagnetic and weak interactions. While essential for atoms to form and nuclear reactions to happen, the cause of this was not understood until the 1960s when it was proposed that like all interactions (excluding gravity) this symmetry breaking would also have an associated quantum field: the Higgs field. However, the corresponding quantum, the Higgs Boson, wasn't discovered until 2012. To test KANs, we first set out by trying discover the Higgs Boson in the data.

Discovering Higgs Boson with KANs

In this test we first use simulated data to train the model and then use the model to test it on actual LHC data and prove the existence of the Higgs Boson. To test the success of the model, we use the following metrics:

1. Loss - This is crucial in training any Machine Learning model as it is the metric on which gradient descent is performed to try obtain the best model possible.
2. Accuracy - This is to test how often the model incorrectly classifies a data point in the dataset. When testing, it serves as a key metric of checking if the model is over-fit to the data.
3. Significance Level - When detecting new particles, the Significance Level is the most important factor in dictating whether the experiment provides enough evidence to accept new theory. In certain cases, this may not be true as computational methods are employed however we aim to prove that our classifier has minimal inaccuracies through the other two measures.

To detect the Higgs Boson, we use the $H^0 \to \tau^+\tau^- \to (l^+\nu_l\nu_\tau)(\text{hadrons} + \nu_\tau)$ signal and the background for that will be something that mimics $H \to \tau^+\tau^-$ with the higgs boson so the process $ZZ \to \tau^+\tau^-$.

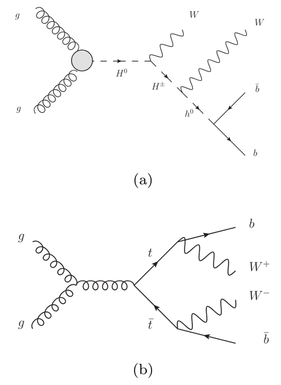

Figure 2. (a) Diagram describing the signal process involving new exotic Higgs bosons H^0 and $H^\pm$. (b) Diagram describing the background process involving top-quarks (t). In both cases, the resulting particles are two W bosons and two b-quarks.

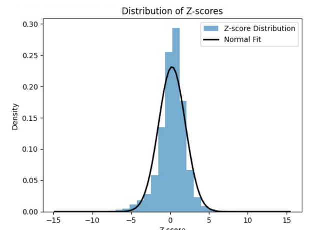

Figure 3. Diagram showing the z-scores evaluated for each neural network output and comapred to a standard normal distribution to evaluation the significance level.

From our plots, we identify a significance of 2.75σ and hence giving great evidence to suggest the existence of the Higgs Boson. In addition, our model achieved a 73% accuracy and a loss of 0.29 which shows that our model was very accurate in its classifications potentially even rivalling that of applying cuts while being able to establish a deeper relationship.

SuperSymmetry

Even after the success of SM as our best model of physics to date, open questions remain. From not being able to account for dark matter and dark energy (which together account for 95% of the universe) as well to the inability to incorporate gravity, new theory was called for. One of the most beautiful ideas that arrived in this intellectual quest was SuperSymmetry (or SUSY for short): the idea that all matter carrying particles (fermions) would have a supersymmetric force carrying particle (boson) and vice versa. In our next example, we wish to try find evidence of this.

Detecting SUSY particles with KANs

Detecting SUSY particles can be very troublesome due to the fact that many of the SUSY particles lack charge or mass effectively making them invisible to the LHC ATLAS detector. As such we wish to investigate a process leading to a final state in which some particles are detectable and others are invisible to the experimental apparatus, and a background process with the same detectable particles but fewer invisible particles and distinct kinematic features.

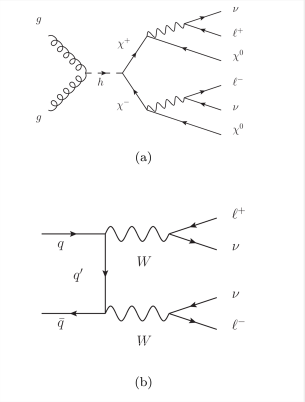

Figure 4. Example diagrams describing the signal process involving hypothetical SUSY particles $\chi^\pm$ and χ^0 along with charged leptons $l^\pm$ and neutrinos ν (a) and the background process involving W bosons (b). In both cases, the resulting observed particles are two charged leptons, as neutrinos and χ^0 escape undetected.

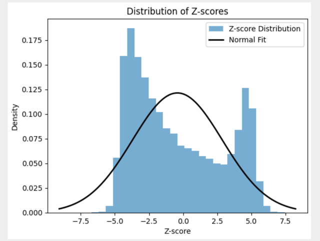

Figure 5. Diagram showing the z-scores evaluated for each neural network output and comapred to a standard normal distribution to evaluation the significance level.

From our plots again, we identify a significance of 1.65 σ which while lower than that of the Higgs Boson makes sense as the Higgs Boson is a confirmed particle while SUSY particles remain hypothetical. This significance is still higher than that of traditional MLPs or by applying Cuts. In addition, our model achieved a higher accuracy of 80% and a loss of 0.373 which shows that our model was very accurate in its classifications potentially even rivalling that of applying cuts while being able to establish a deeper relationship.

Conclusion

Based on the following results in the table below, we believe we have given good reason for the usage of KANs in the hunt for exotic particles to further the physics knowledge as KANs allow for greater interpretability and serve as a great aide in learning deeper laws of the universe.

References

[1] CecileGermain Balazs Kegl.
Higgs boson machine learning challenge, 2014.

[2] Pierre Baldi, Peter Sadowski, and Daniel Whiteson.
Searching for exotic particles in high-energy physics with deep learning.
Nature communications, 5(1), 7 2014.

[3] Aysu Ismayilova and Vugar Ismailov.
On the Kolmogorov neural networks.
arXiv (Cornell University), 1 2023.