INSTITUTE OF EARTH SCIENCES
ACADEMIA SINICA

SUMMER INTERNSHIP PROGRAM

Deep Learning Based Seismic Wave Identification

Author :

Gauss CHANG

Department of Physics, NTU

Lecturer :

Dr. Rih-Teng WU

August 24, 2023

I Abstract

The subject is to classify particle data, e.g. a collection of high momentum Higgs boson [1] events versus high momentum light quarks, as measured in the detector, using neural networks [2].

II Introduction

I. When a "high-momentum" quark is produced, it will hadronize and produce a lot of particles afterwards. These particles would form an object which is called a "Jet", which is the product we can see in the experiments.

II. A typical description of the jet is a collection of particles distributed in a cone. The particles are mostly hadrons, such as kaons, pions, can be either charged or neutral.

III. When a high-momentum Higgs boson decaying into two bottom quarks, both quarks will hadronize into jets, too.

IV. However, since the initial momentum of the Higgs boson is highly "boosted", the open angle of the two jets is small, and would merge into single jet in the end, causing difficulties in identifying processes.

V. That is, the Higgs jet should preserve many physics properties of the original particle, for example, the invariant mass!

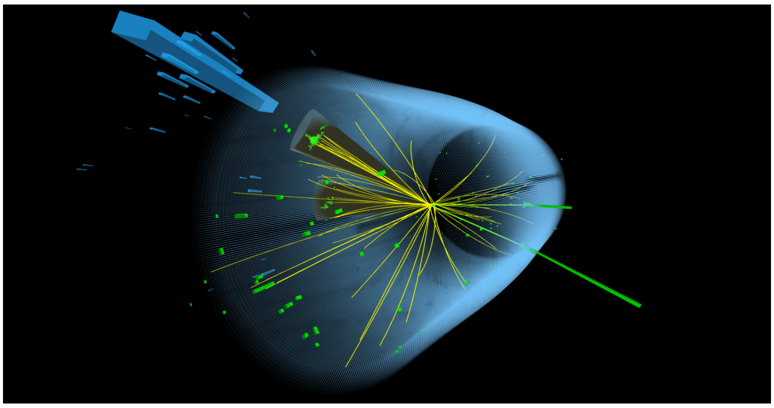

Figure 1: LHC Computer Simulation

In Figure 4 shows how the LHC experiment look like.

II.I Importance of this Work

In a paper published by CERN [2], they proposed a technique to use deep learning to identify the particles produced in the proton-proton collision experiment. Finding the Higgs boson is one major work. Since the higgs predicted a undiscovered boson in the 1964 [1], finding and proving the existence of this boson becomes the first priority of the experimental physicists. In the high energy physics field, every experiment is surprisingly expensive. Thousands of millions of data is being collected in every collision. The data scientists work on the data for years just to discover a new boson, or "Subatomic particle". But since the requirement of accuracy in this field is extremely high, so every new discovery must pass through the "5σ" test. that means, the chance of a false discovery must be at least 5σ out. And since every actual new discovery is a Nobel prize, timing is everything. A well-trained prediction model can tell the researchers at a early stage weather what they are speculating is true.

II.II Possible Model Structure

Following is a model structure they [2] have proposed.

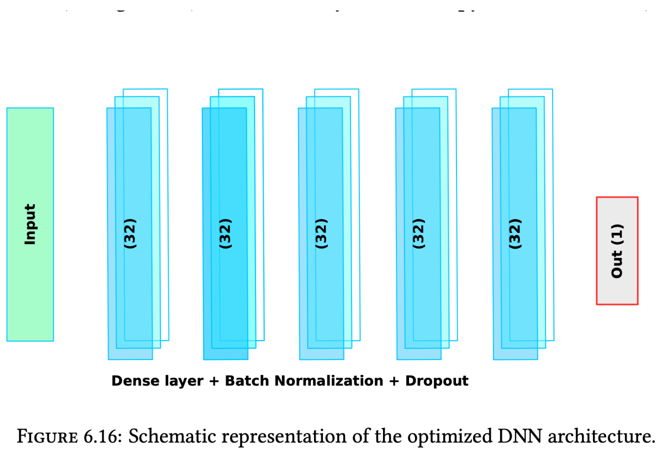

Figure 2: Model Structure 1 [2]

III Training Data Structure

The training data structure is as shown in Table 1

Table 1: Training Data Structural

ID	Code	P_x	P_y	P_z
0	0	-4.0379	1.2086	-3.0469
0	5	-5.6522	1.6572	-4.8261
0	6	-4.6699	1.4765	-4.5763
0	6	-13.3308	2.4805	-13.3598
		⋮		

and the particles is represented by the code shown in Table 2

Table 2: The Particle Code

Code	0	1	2	3	4	5	6	7
Particle name	Photon	Electron	Position	Muon	Anti-muon	Negatively charged hadron	Positively charged hadron	Neutral hadron
Charge	0	-1	+1	-1	+1	-1	+1	0

- P_x is the particle momentum in x direction in unit GeV ${c_0}^{-1}$.
- P_y is the particle momentum in y direction in unit GeV ${c_0}^{-1}$.
- P_z is the particle momentum in z direction in unit GeV ${c_0}^{-1}$.
- ID stands for the event number. The same event share the same ID.
- This data set is obtained from Kai-Feng Chen, Detecting Boosted Higgs Competition 2022 [3].

IV Implementation

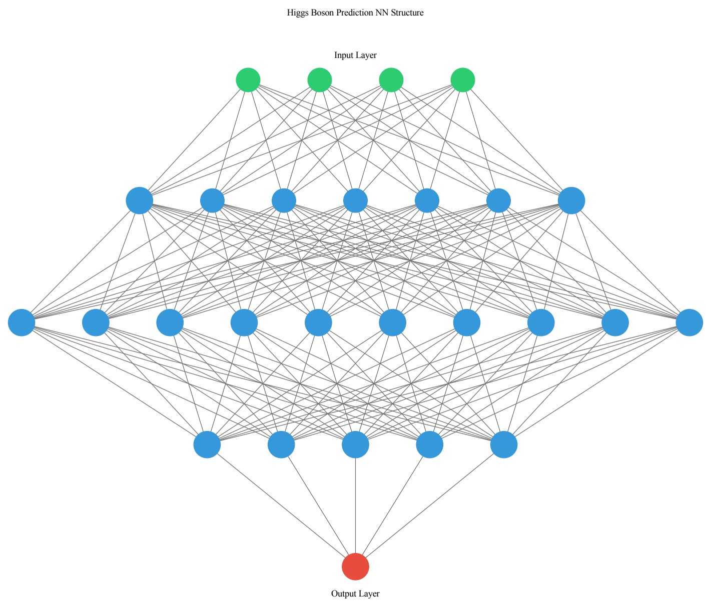

Figure 3: Higgs Boson Prediction NN Structure

In my Neural Network, I use three all-connected hidden layers. And the input features is the total number of particle in one event, the total invariant mass of one event, the number of photon in one event, and the number of "Muon" and "Anti-muon" in one event. The structure is basically a product of try and error method. The reason I use the "number of Photon" and "number of Muons" is because although the Higgs boson may decay into these particles, the chances is very small (0.2% and 0.022%).

V Result

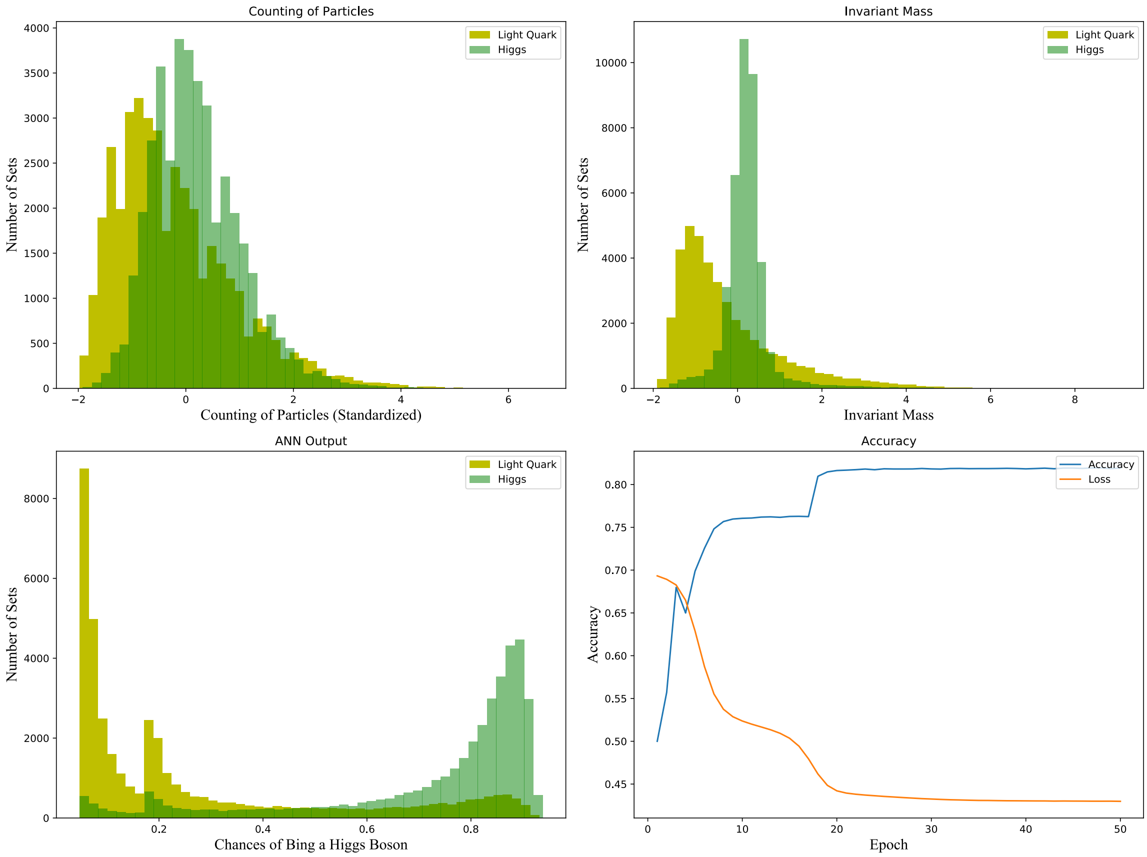

Figure 4: NN Model Prediction Result

As you can see in the Figure 3, after aboue 30 − 40 epoch, the accuracy of my model converged at about 81%.

VI Conclusion

This project is primarily aimed at predicting Higgs boson using deep learning methods. The Higgs boson, a fundamental particle confirmed in 2012, plays a vital role in the formation of the standard model and the deepening understanding of physics. Our objective is to explore how deep learning can be more effectively applied in the field of particle physics, specifically for the prediction of Higgs boson.

Currently, our deep learning model has achieved an accuracy rate of approximately 81%. This result is commensurate with the current average level in the field, indicating that our model has acquired basic predictive capability. However, considering the challenges of Higgs boson prediction and the potential of deep learning, I believe there are still possibilities for further enhancement of the model performance.

In future work, I will focus on further optimization and enhancement of the model. This will involve adjusting the parameters of the deep learning model, adding more features related to Higgs boson, and optimizing the model's training strategy. I believe that through these improvements, T will be able to increase the model's predictive accuracy.

In addition, I will conduct a more in-depth analysis to understand the situations in which the model's predictive capability is weaker, to identify areas needing improvement. This will help us more accurately locate the problem and optimize it in a targeted manner.

I also hope to further deepen the application of deep learning in the field of physics and attempt to apply this technology to more physical problems. I believe that this will not only help promote the development of physics, but will also broaden the application scope of deep learning.

In conclusion, although our model has only barely reached the average level at present, I are still confident about the future possibilities. I will continue to optimize our model and apply deep learning technology.

References

[1] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[2] Leonardo Giannini. Deep learning techniques for the observation of the higgs boson decay to bottom quarks with the cms experiment, 2020. Presented 28 Jul 2020.

[3] Kai-Feng Chen. Detecting boosted higgs competition 2022, 2022.