Organisation: ML4Sci: CMS E2E

Project Title: **Exploring the underlying symmetries in Particle Physics with Equivariant Neural Networks**.

Introduction

About Me

Full Name: Diptarko Choudhury

Time Zone: IST (UTC+5:30)

Email: diptarko.choudhury@niser.ac.in

GitHub ID: dc250601

Matrix Handle: dc2506

Education

University: National Institute of Science Education and Research (NISER), Bhubaneswar

Program: 5-Year Integrated Masters in Physics, with a Minor in Computer Science

Current Year: Third Year (6th semester ongoing)

Personal Background

I am an undergraduate 3rd Year student at the National Institute of Science Education and Research (NISER), pursuing my Integrated Masters in Physical Sciences. Along with my Masters, I am also doing a Minor in Computer Science. My main interest lies in the field of computational physics and Machine Learning. I am proficient in C++, Python and Java and have experience using Machine Learning frameworks such as Tensorflow and PyTorch.

In the past, I have been a Google Summer of Code (2022) student under ML4Sci, where I worked on a project that involved using attention-based vision models on low-level sparse detector images from the CMS detector to classify physics events. I have also worked on other machine learning-based projects, the details of which can be found in my resume and GitHub profile. I have completed several certifications (Coursera) in machine learning, details of which can be found in my resume. I am passionate about continuing my research in Machine Learning.

I am also part of many extracurricular clubs in my college and love participating in events such as debate and extempore. I am also an active member of RTC (Robotech Tech Club) and the Coding Club of NISER.

Motivation for the project

Symmetry is one of the most beautiful and interesting phenomena in physics. Many real-life problems which are otherwise unsolvable can be easily solved by exploiting the underlying symmetries. In physics and nature in general, symmetry has fascinated scientists and mathematicians for ages. One of the human brain's fundamental characteristics is finding symmetries in almost anything and then exploiting the same to solve challenging problems. Conventional neural networks lack the idea of symmetry. Embedding various types of symmetries into neural networks is an active field of research. Neural networks that have a sense of symmetry and transform equivariantly to such symmetry operations are called Equivariant Neural Networks. Equivariant Neural Networks work better where a high degree of symmetry is involved. Moreover, such networks are more robust and generalize better than conventional Neural Networks. These networks have a much greater learning capability, tend to express better, and are more sample efficient. Using Equivariant Neural Networks in places like particle physics where a high degree of symmetry is involved seems promising.

Insights gained from the Evaluation Test

The organization wanted me to complete an evaluation test. I have completed the three evaluation tests and submitted them as per the instruction. Since the evaluation tests were highly correlated with my current project/proposal, I gained quite some insight.

Model	Peak-AUC		Model	Peak-AUC
C4	79.29		C4_lite	79.34
C8	79.52		C8_lite	79.34
D4	79.54		D4_lite	79.41
D8	79.52		D8_lite	79.46
CNN	79.51			

The above table illustrates the performance I got with various architectures. The CNN is the conventional model. The models C_n and D_n belong to the Cyclic Group and Dihedral Groups, respectively. The n signifies the number of elements of the group. The Particle Images, in reality, are SE(2) Equivariant(Can be approximated to); to keep things simple, we used the C_n and D_n groups. Since convolution operation is translational equivariant, we can use C_n and D_n with large values of n to mimic SE(2) equivariance.

Keeping the number of layers and channels fixed, we get better performance on increasing the Equivariance level in our network. D_n has a greater degree of symmetry than C_n; hence, we see better performance with Dn. Moreover, as we increase n, the AUC scores increase, suggesting that Equivariance is indeed helping. The Equivariant networks have slightly greater performance than the Non-Equivariant ones. The margin of improvement should improve with a larger dataset since most Equivariant networks tend to overfit due to higher learning capacity than CNNs.

The Proposal

Abstract:

This project aims to employ Equivariant Neural networks to understand the symmetries in particle physics. The particle collision reactions follow the Lorentz symmetry or are Lorentz invariant. This project's main aim will be to build upon theory and code to achieve near Lorentz invariance inside Neural Networks. We will use data from CMS Open Data, consisting of low-level multi-channel simulated images from the calorimeter itself.

Description:

The idea to use Equivariant Neural Networks for analysis of particle collision data comes from the fact that most of these reactions are indeed Lorentz Invariant. That is, they are not affected by any element of the Lorentz Group. If the idea of Lorentz Invariance can be incorporated into the architecture of the Neural Networks, then much more efficient and expressive networks could be built. In the last couple of years, multiple attempts have been made to analyze such symmetries using CNNs and point cloud based approaches. The Equivariant networks should produce simpler models with fewer learnable parameters, better generalization capability, and sample efficiency than the former methods. To solve this problem, we will build up from the basics in this project and will try to incorporate Soft (Approximate) Lorentz Invariance into our network. The project requires much theoretical groundwork and a sound knowledge of Group and Representation theory.

The project's main aim will be to create a model that will best mimic the symmetries present in the domain of particle physics and, if possible, find new hidden symmetries.

Timeline

Current Progress

I have been working on the Evaluation test of the projects and finished working on them. I gained insight after building the basic architecture and confident enough to build a much more complex and tailor-made architecture to suit my problem the best. I have been doing literature reviews and planning for this project and have also referred to some papers which try to solve this problem.

Community Bonding

I will set up a proper channel and weekly schedule for communications and discussions in consultation with the mentors. I will use this time to further my understanding of the problem statement and clarify doubts (if any) with my mentors.

Week 1-2

During this time, I will work on the underlying mathematics involved in my project. During this short time frame, I will brush up my knowledge of Group Theory and learn more about Representation Theory. These two weeks will be crucial since, for the development of Equivariant Neural Networks, a proper understanding of both is necessary.

Week 3

This week I will do a literature survey and try to identify papers and articles that will help me the most. I will also continue my study of mathematics. This week, I will focus more on understanding Geometric Deep Learning and various types of Equivariant Neural Networks.

Week 4-6

I will implement Steerable Group Convolutions, specifically SE(2) Equivariant networks. This should be straightforward to implement given the availability of standard libraries such as escnn. I will also play around with various hyperparameters of the network and will try to get a better working knowledge of such models. By the end of the 5th week, I will compare them with Regular Group Convolutions and Conventional CNNs. I will also fine-tune the network architectures to get better results.

Week 7

I will again do a short literature survey on Harmonic Networks and Clebsch–Gordan Nets. I will try to determine if they can be applied to this problem.

Week 8-9

I will try to implement Clebsh-Gordon (if possible) and Harmonic Networks. I will train them and fine-tune the Hyper-Parameters for better results. Finally, I will finally compare them to the previously discussed network architectures.

Week 10

This week I will brush up my knowledge on various topics in physics, specifically the Special Theory of Relativity. I will learn more about the Lorentz group and the Lorentz Invariance property of particle physics reactions.

Week 11-12

I will specifically read about Lorentz Invariant Neural Networks in these two weeks. I will refer to various works done in this field to understand how we can implement the same for our project. Given the complexity of the subject, I have allocated two weeks for the same. By the end of the 12th week, I should be comfortable with Lorentz Equivariant Neural Network.

Week 13-16

These 4 four weeks have been allocated explicitly for the design, implementation and training of the Lorentz Equivariant Neural Network. Also, I will try to touch upon the concept of Approximate Lorentz Equivariance and will try to implement the same during these four weeks. By the end of the 16th week, I expect to have a fine-tuned Lorentz Equivariant Neural Network.

Week 17

This week has been kept as a buffer week to deal with unforeseen circumstances.

Week 18-20

These three weeks will be spent comparing the results of all the networks. I will run various types of tests to find the effectiveness and robustness of these networks. We will also compare these networks with Vision Transformers in terms of accuracy and sample efficiency. During these three weeks, I will also try to understand the nature of the particle physics reactions using our network and, if possible, will find a way to find new types of symmetries and anti-symmetries in our data.

Week 21-22

I will merge all my code in these two weeks; add comments and documentation (any remaining) to various parts of my code for it to be easily understandable to the end user. I will also make a pipeline for the end-user to use the network and merge it into the main code.

Future Developments

The main aim of this project is to develop rich Equivariant models which are robust enough to understand the underlying symmetries involved. These networks can be further exploited to understand other degrees of symmetries involved. In future, I would like to extend this project to understand and find new hidden symmetries.