Google Summer of Code

ML4SCI

Google Summer of Code 2024
Project Proposal for ML4Sci

Learning quantum representations of classical high energy physics data with contrastive learning

GSOC Contributor: Sanya Nanda

Email Id: sanya.nanda@gmail.com
Country of Residence: India
Timezone: India/UTC+5:30, Indian Standard Time
Degree: Bachelor of Technology in Computer Engineering, 2023
Primary Languages: English and Hindi
Social Handle: GitHub, LinkedIn

Prerequisite Test: https://github.com/SanyaNanda/ML4Sci-QMLHEP-2024/tree/main

Index of Contents

1. Overview

1.1 Project Synopsis

The ambitious HL-LHC program requires enormous computing resources in the next few decades. New technologies are being sought after to replace the present computing infrastructure. A burning question is whether quantum computers can solve the ever growing demand of computing resources in High Energy Physics (HEP) in general and physics at LHC in particular. Our goal here is to explore and to demonstrate that Quantum Computing can be the new paradigm.

Discovery of new physics requires the identification of rare signals against immense backgrounds. Development of machine learning methods will greatly enhance our ability to achieve this objective. However, with this ever-growing volume of data in the near future, current machine learning algorithms will require large computing resources and excessive computing time to achieve good performance. Quantum Computing, where qubits are used instead of bits in classical computers, has the potential to improve the time complexity of classical algorithms.

With this project we will be implementing Quantum Machine Learning methods for LHC HEP analysis using Pennylane framework. This will enhance the ability of the HEP community to use Quantum Machine Learning methods.

1.2 Benefits to Community

1. Discovery of new physics by identification of rare signals against immense backgrounds
2. Trained quantum model and benchmarking
3. Benchmark of the performance on a HEP dataset against a classical reference model

1.3 Background Research

The High Luminosity Large Hadron Collider (HL-LHC) is an upgrade of the LHC which aims to achieve instantaneous luminosities a factor of 5 to 7.5 larger than the LHC nominal value, thereby enabling the experiments to enlarge their data sample by one order of magnitude during the 12 years of HL-LHC operation compared with the LHC baseline programme.

Following 5 years of design studies and R&D, this challenging project requires about 10 years of developments, prototyping, testing, series production and implementation; hence operation is expected to start at the end of this decade. This timeline is in lieu of the fact that in the coming years, many critical components of the accelerator will reach their End of Life, due to radiation damage and will need replacement. The upgrade phase is therefore crucial for the full

exploitation of the LHC Physics potential, but also to enable operation of the collider beyond the end of the nominal LHC exploitation in 2025.

The HL-LHC will rely on a number of key innovation technologies, including cutting-edge Nb3Sn and Nb-Ti superconducting magnets, compact superconducting crab cavities with ultra-precise phase control for beam rotation, new technology for beam collimation such as bent crystals, and high-power, loss-less MgB2 superconducting links, to name only a few.https://hilumilhc.web.cern.ch/

Let's take the example of quark and gluon classification, instead of using supervised learning where we have the labels and we train the model to understand distinctive features of the 2, we can use contrastive learning. It is a machine learning paradigm that aims to learn representations by contrasting different examples. The core idea is to encourage similar examples to be closer together in a learned embedding space, while dissimilar examples are pushed farther apart. This approach is often used in self-supervised or unsupervised learning settings, where labeled data may be scarce or expensive. Figure 1, shows that embeddings of instances in the same class are closer together compared to other classes.

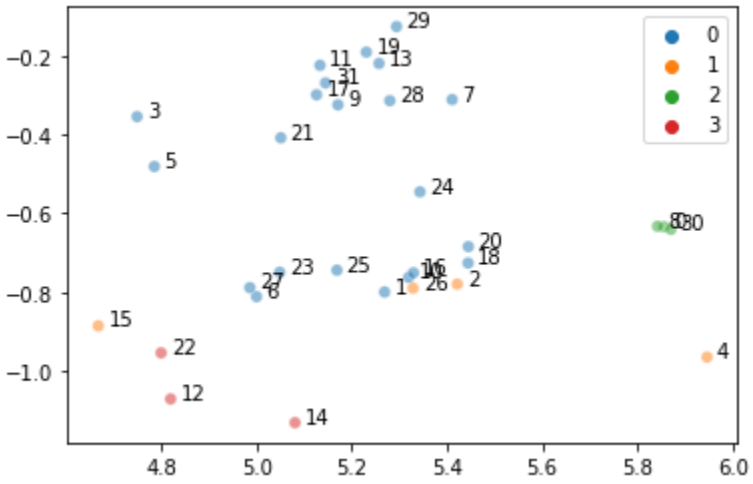

Fig 1: Embeddings of instances

In contrastive learning, two views of the same instance are considered as a positive pair, and the goal is to bring their representations closer together in the embedding space. At the same time, representations of different instances are pushed apart. This is typically achieved through a loss function that encourages the representations of positive pairs to be close and those of negative pairs to be far apart.

It is successful in learning powerful representations from large-scale unlabeled datasets. These learned representations can then be fine-tuned on smaller labeled datasets for downstream tasks, such as classification, object detection, or semantic segmentation, often leading to improved performance compared to models trained from scratch on the labeled data alone.

Further, we can leverage quantum solutions along with contrastive learning which embeds the classical data onto quantum states. B Jaderberg et al authored a paper Quantum self-supervised learning in 2022. Where they took the first steps to understanding whether quantum neural networks (QNNs) could meet the demand for more powerful architectures and test its effectiveness in proof-of-principle hybrid experiments. Interestingly, they observed a numerical advantage for the learning of visual representations using small-scale QNN over equivalently structured classical networks, even when the quantum circuits are sampled with only 100 shots. Furthermore, they applied their best quantum model to classify unseen images on the *ibmq_paris* quantum computer and found that current noisy devices can already achieve equal accuracy to the equivalent classical model on downstream tasks.

Maria Schuld et al, wrote a paper in 2021 on The effect of data encoding on the expressive power of variational quantum machine learning models. Quantum computers can be used for supervised learning by treating parameterized quantum circuits as models that map data inputs to predictions. They investigated how the strategy with which data is encoded into the model influences the expressive power of parameterized quantum circuits as function approximators. They show that there exist quantum models which can realize all possible sets of Fourier coefficients, and therefore, if the accessible frequency spectrum is asymptotically rich enough, such models are universal function approximators.

A. Hammad et al in 2023 published a research article Quantum Metric Learning for New Physics Searches at the LHC. In the NISQ (Noisy intermediate-scale quantum), Quantum computers can be utilized for deep learning by treating variational quantum circuits as neural network models. This can be achieved by first encoding the input data onto quantum computers using nonparametric unitary gates. The separation is achieved with metric loss functions, hence the naming Quantum Metric Learning. With the limited number of qubits in the NISQ area, this approach works naturally as a hybrid classical-quantum computation enabling embedding of high-dimensional feature data into a small number of qubits.

2. Goals and Deliverables

2.1 Deliverables

1. Implement a function encoding classical data on a quantum model with contrastive learning.

2. Experiment with different ideas for embedding functions and contrastive losses for training.

3. Benchmark the trained embedding against a standard encoding on a given QML model.

3. A fully trained embedding function for classical data with e.g. Pennylane framework.

4. Benchmark the performance against a standard encoding.

2.2 Prerequisite Test

I have completed the ML4Sci 2024 Test: (Task I, II, III, VI)
https://github.com/SanyaNanda/ML4Sci-QMLHEP-2024/tree/main

2.3 Outline of Approach

Theory

In Contrastive Learning, embeddings are created in the vector same. Instances of the same class are embedded closer to each other as their similarity or fidelity is higher. Following Fig 2 and 3, are results from Task VI from the Test.

Similarity score: tf.Tensor(**0.9999998895010674**, shape=(), dtype=float64) Label: 1

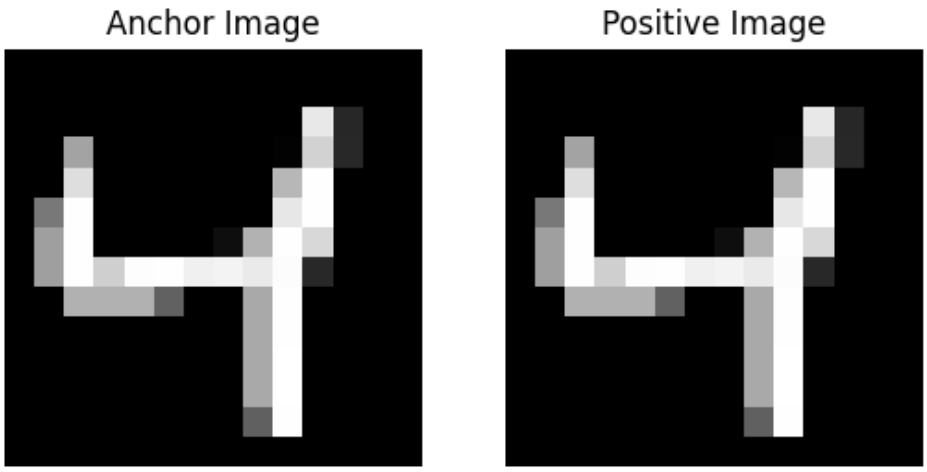

Fig 2: Positive Pair in contrastive Learning

Instances of the different classes are embedded further apart from each other as their similarity or fidelity is lower.

Similarity score: tf.Tensor(**0.087229201481301**, shape=(), dtype=float64) Label: 0

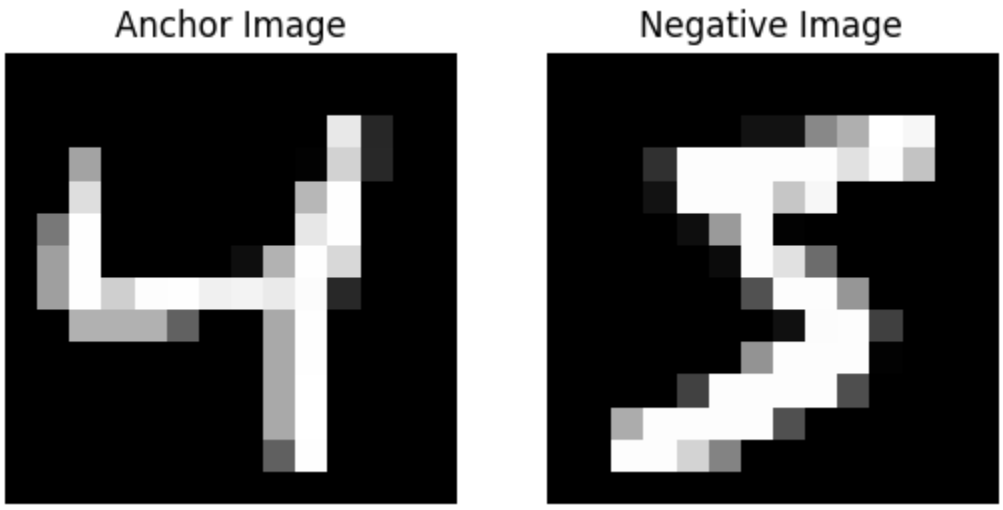

Fig 3: Negative Pair in contrastive Learning

In the field of computer vision, accurately measuring image similarity is a crucial task with a wide range of real-world applications. Quantum networks, coupled with contrastive loss, provide a powerful framework for learning image similarity in a data-driven manner.

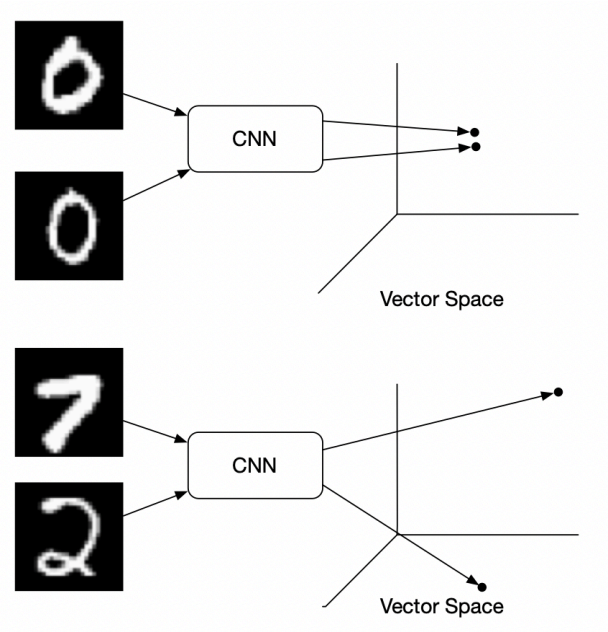

Fig 4: Concept of Contrastive Learning, Reference [2]

Components of Contrastive Learning:

Input Pairs: The contrastive loss function operates on pairs of input samples, where each pair consists of a similar or positive example and a dissimilar or negative example. These pairs are generated during the training process, with positive pairs representing similar instances and negative pairs representing dissimilar instances.

Embeddings: The quantum circuit processes each input sample through a shared network, generating embedding vectors for both samples in the pair. These embeddings are fixed-length representations that capture the essential features of the input samples.

Distance Metric: A distance metric, such as Euclidean distance or cosine similarity, is used to measure the dissimilarity or similarity between the generated embeddings.

Contrastive Loss Calculation: The contrastive loss function computes the loss for each pair of embeddings, encouraging similar pairs to have a smaller distance and dissimilar pairs to have a larger distance. The general formula for contrastive loss is as follows:

$$L = (1 - y) * D^2 + y * \max(0, m - D)^2$$

Where:

- L: Contrastive loss for the pair.
- D: Distance or dissimilarity between the embeddings.
- y: Label indicating whether the pair is similar (0 for similar, 1 for dissimilar).
- m: Margin parameter that defines the threshold for dissimilarity.

The loss term $(1 - y) * D^2$ penalizes similar pairs if their distance exceeds the margin (m), encouraging the network to reduce their distance. The term $y * \max(0, m - D)^2$ penalizes dissimilar pairs if their distance falls below the margin, pushing the network to increase their distance.

Aggregating the Loss: To obtain the overall contrastive loss for the entire batch of input pairs, the individual losses are usually averaged or summed across all the pairs.

By minimizing the contrastive loss through gradient-based optimization methods, such as backpropagation and stochastic gradient descent, the network learns to produce discriminative embeddings that effectively capture the similarity or dissimilarity structure of the input data.

Task VI Submission

We created the quantum states using Amplitude Encoding. In the original MNIST data, the image dimensions are 28 by 28 which is 728 when flattened. Using PCA, we found out that 154 flattened pixels are enough to retain the distinctive attributes of the images. On Applying max pooling on the image data with a kernel of size 2 by 2 we get the 196 flattened pixels which is in range of PCA relevant pixels 154.

In Amplitude Encoding, we use n qubits to represent 2^n features. Therefore, with just 8 qubits we can represent 256 pixels, since we have 196 pixels post data preprocessing, we pad the remaining pixels. We get the following quantum state where the image pixels are represented as features and the weights are optimized upon training of the model.

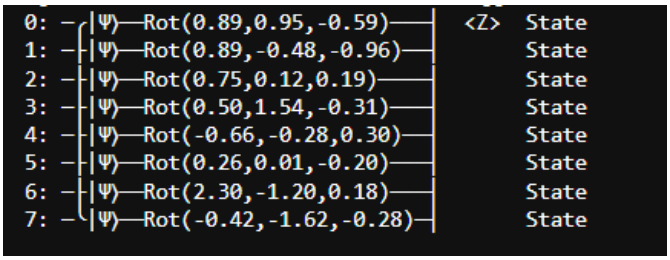

Fig 5: Quantum Circuit representing images

Furthermore, similarity of 2 quantum states can be established via Swap Test which was implemented as shown below.

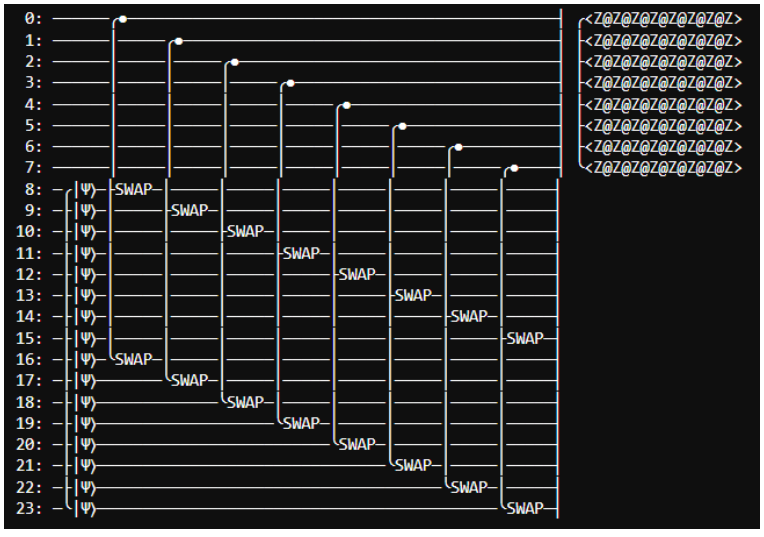

Fig 6: Swap Test to compare for similarities in quantum states

Suggestions for Final Solution in addition to Approach discussed

1. Use Fidelity in Contrastive Loss: Same states showed nearing 1 fidelity and differing states were far apart in their fidelities, quantum state fidelity can be used as a factor in the contrastive loss functions, giving the solution more context on the images.

2. Experiment with different contrastive loss functions: Different Loss Functions can bring forth unique results as show below

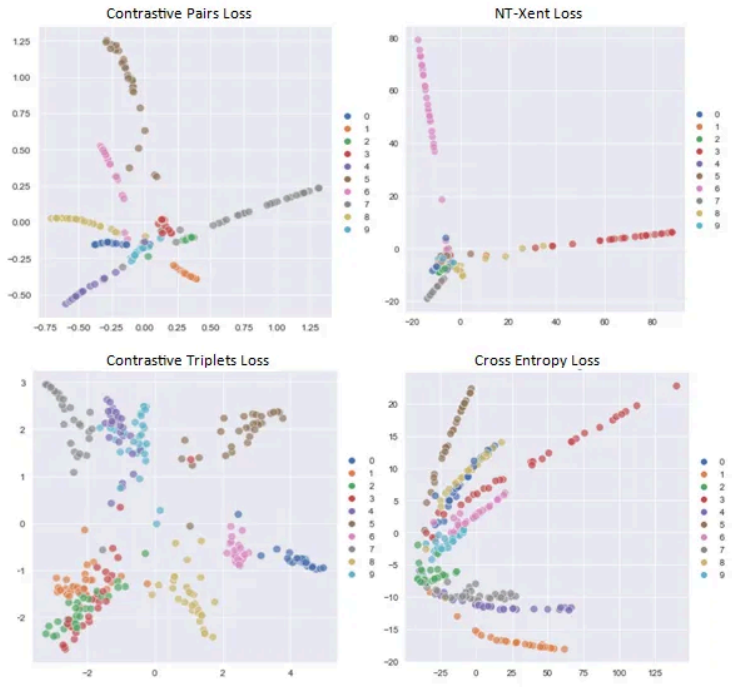

Fig 7: Results from different contrastive loss functions, Ref [8]

3. Use of Quanvolutional Neural Networks: Powering Image Recognition with Quantum Circuits by Maxwell Henderson, et al. *Quanvolutional* layers operate on input data by locally transforming the data using a number of random quantum circuits, in a way that is similar to the transformations performed by random convolution filter layers. Provided these quantum transformations produce meaningful features for classification purposes, then the overall algorithm could be quite useful for near term quantum computing, because it requires small quantum circuits with little to no error correction.

3. Schedule of deliverables or Work Plan

3.1 Application Review Period

This time will be utilized to polish skills relevant to the project which will help in making the solutions better and more accurate. I will contact the mentor to enhance and work on the idea and its implementation. Also, extensive research on implementations and methodologies for the set goals to make the solutions easier will be done.

3.2 Community Bonding Period

This time will be utilized to interact with the mentors and set up communication loops; and continue the process of refining the workflow for the project in consultation with the mentors. Work will be initiated in this period to complete the tasks before the stipulated deadlines. Also, the opportunity of getting involved with other members of the community will be utilized.

3.3 Coding

Before Mid-way Evaluation:

- Quantum Embedding and Contrastive Loss functions will be experimented with and finalized upon
- Preliminary results from Model Training and evaluation
- Experimenting with use of Quanvolutional layers
- Code is modularised, refactored and reusable along with test cases
- Fine Tuning of Hyper-parameters

After the Mid- way Evaluation

- Fine Tuning of Hyper-parameters
- Final Model training and testing
- Benchmarking with classical, hybrid and quantum solutions
- Documenting the work and results

4. Past Experience

4.1 Academic Details

I have completed BTech in computer Engineering at Thapar Institute of Engineering and Technology, Patiala. I have a CGPA of 9.5 on a scale of 10 and am a recipient of a merit-based scholarship. I am proficient in the tech stack relevant to the given project. I have solid

knowledge of machine learning, deep learning, quantum mechanics and quantum computing. I have strong python skills and the ability to work independently and proactively on a research project

4.2 Personal/Open Source Projects related to Quantum

1. **IBM Qiskit Advocate**:
   https://www.credly.com/badges/8e047193-df21-458f-b5e5-c611b8ab7039/public_url
   I am an IBM Qiskit Advocate, as part of it I contributed to QAMP Fall 2022 as a mentee in Qiskit Machine Learning Tutorials. Also, conducted Qiskit Fall Fest in TIET, Patiala in 2022 to spread an understanding of quantum computing among college students through interactive sessions and hands-on challenges.
2. **TCS Quantum Challenge**:
   We worked on Fleet Allocation Model using QAOA on AWS Braket, as part of the challenge I got experiential learning on QAOA and warm start QAOA along with ins and outs of AWS Braket, also enhanced presentation skills by presenting the end results to the jury, results are awaited.
3. **Qiskit Certified:**
   https://www.credly.com/badges/d150279e-68c4-4cf7-b809-727e45bd78c8/public_url
   I am certified in Qiskit which marks my understanding of fundamental quantum computing concepts and use of Qiskit open source SDK. I have experience in using Qiskit in python to create and execute quantum computing programs on IBM Quantum computers and simulators.
4. **Quantum Computing:** https://linktr.ee/gdsctiet_qc
   It covers the basics of quantum computing with notes, reference sheets, code and learning resources. All the concepts are explained on DSC TIET's YouTube channel under the Quantum Computing playlist. I started and led this project and through the process interacted with highly interested individuals that helped improve my understanding of the subject.

4.3 Motivation?

I value this opportunity. This is a problem that I find interesting and I respect and passionately follow the work of the organization. This is great chance for me to harness my skills further and get to interact with like minded people.

5. Availability Schedule and Other Commitments

5.1 Working hours

I can commit the required time to achieve my goals and deliverables.

- **Work Timings for weekdays (4 - 5 hours daily)**
  Early Mornings and Late Evenings IST
- **Work Timings for weekends (7 - 8 hours daily)**
  Mid Morning to Evening IST

Other Commitment: This will be a side project to enhance my skills and pursue my interest in quantum computing. I have commitments as an SDE-1 in my place of work.

5.2 Regular Updates and Meetings with Mentor

1. Written Progress will be shared by the end of every week and code will be regularly committed on Github following the guidelines.
2. I will be available on slack or any other platform suitable for communication.
3. Regular SCRUM meeting to share progress, finalize next steps and acquire guidance on impediments.

5.4 Post GSOC

I believe that open-source contribution is not just restricted to Google Summer of Code. I would love to continue working on the developments and enhancements of the project post-GSoC.

REFERENCES

1. https://arxiv.org/abs/2311.16866
2. https://arxiv.org/abs/2008.08605
3. https://arxiv.org/abs/1904.04767
4. https://iopscience.iop.org/article/10.1088/2058-9565/ac6825
5. https://medium.com/@hayagriva99999/exploring-siamese-networks-for-image-similarity-using-contrastive-loss-f5d5ae5a0cc6
6. https://shubham-shinde.github.io/blogs/contrastive/
7. https://pennylane.ai/qml/demos/tutorial_quanvolution/
8. https://shairozsohail.medium.com/contrastive-representation-learning-a-comprehensive-guide-part-1-foundations-90c1944dbd1e