Calabi-Yau data classification using quantum machine learning algorithms

Mohammadreza Noormandipour

November 2019

With the ever increasing volume of datasets produced after compactification of strings on Calabi-Yau n-folds, the string phenomenologists have tried to find explicit relations between the topological features of these manifolds and physical observables from lower energy theories. Due to the abundance of possible solutions (known as landscape data) after compactification, basic steps towards establishing the landscape data such as finding the Groebner basis is exponentially expensive in computational resources [1706.02714]. Therefore, interests have grown in exploration of the string landscape using machine learning techniques [1706.02714, 1706.03346, 1706.02714, 1812.02893].

As a typical example, Daniel Krefl et al. have used convolutional neural networks to investigate the minimum volume of Sasaki-Einstein base manifolds of non-compact toric Calabi-Yau 3-folds (CY3) [1706.03346]. The bounds on minimum volume are important since using AdS/CFT correspondence, one can have interpretation of these bounds on central charges of class of $4d$ $\mathcal{N}=1$ superconformal field theories. This possibility is important as the neural network is circumventing the expensive calculations behind conventional minimization procedures.

They have characterized the class of $4d$ $\mathcal{N}=1$ supersymmetric gauge theories with convex lattice polygons (known as toric diagrams) [1706.03346]. These theories can be defined on the world volume of a stack of D3-branes probing toric CY3 and are expected to flow towards a superconformal fixed point at low energies [1706.03346]. The topological features of the toric diagrams used in their research comprises of the number of lattice points inside and on the perimeter of the convex lattice polygon. Additionally, they have inserted some other feature of the toric diagram in the form of a square matrix into the neural network as input [1706.03346].

One can simply show that, the above mentioned features can be translated to 2D arrays of numbers. We are mentioning this because we are interested in performing the same tasks of machine learning with a quantum version of that and it is usually the case that the input features should be in the form of 2D arrays for quantum machine learning models. Moreover, Yang-Hui He has used a supercomputer at CERN to find 7890 in-equivalent CY 3-folds, as complete intersections in products of projective spaces and he has shown that they can be represented into 12×15 matrices [1706.02714].

As the dimension of the manifolds increases, the number of in-equivalent CY manifolds increases as well, and it would be costly and inefficient for neural networks to obtain accurate results in classifying them. Recently, a new model of a quantum neural network has been proposed which is able to perform classification tasks for 2D arrays much faster than the classical counterparts [https://www.nature.com/articles/s41534-019-0140-4]. The model can also be implemented on real quantum hardware to perform this task.

The goal of this project is to investigate a possible and efficient mapping of the classification of the CY data to a quantum machine learning problem. This would possibly give the opportunity to probe larger subsections of the string landscape and put more accurate bounds on known physical observables due to the access to a higher dimensional space of models and the speed-up, offered by the qubits used in the quantum neural networks.