2024 International Conference on Compilers, Architecture, and Synthesis for Embedded Systems (CASES)

Primer on Data in Quantum Machine Learning

Aviral Shrivastava
School of Computing and Augmented Intelligence
Arizona State University
Tempe, USA
Aviral.Shrivastava@asu.edu

Vinayak Sharma
School of Computing and Augmented Intelligence
Arizona State University
Tempe, USA
Vinayak.Sharma@asu.edu

***Abstract*—With the increased interest in Quantum Machine Learning (QML), the integration of classical data into quantum systems presents unique challenges and opportunities. The class "Primer on Data in Quantum Machine Learning" delves into the foundational concepts and advanced techniques of embedding classical data into quantum states, a critical process for enhancing the performance of quantum algorithms. By exploring various quantum embedding methods and understanding their strengths and limitations, participants will gain a comprehensive understanding of the impact quantum embeddings can have on machine learning applications. This lesson will cover the following concepts: Fundamental Concepts of Quantum Machine Learning, Limits of NISQ devices and Computing in the NISQ era, Embeddings for QML, and Practical effects of embeddings. The understanding of these topics should provide a better understanding of the importance and effect of embeddings on the overall performance of QML in the NISQ era.**

***Index Terms*—Machine Learning, Quantum Information, Quantum Computing**

I. Introduction

The development of Near Intermediate Scale Quantum (NISQ) devices has empowered a new family of hybrid quantum algorithms that differ from prior approaches by incorporating principles of machine learning. While early approaches in quantum computing focused on exploiting superposition and interference to create algorithms that are faster than their classical counterparts, newer work has focused on optimization problems and trainable networks. This has led to the advent of Quantum Machine Learning (QML), a field that combines quantum computing with machine learning to create more efficient and powerful computational models. A critical consideration of algorithms in the current NISQ era is hardware-aware design. More specifically, NISQ devices have a limited number of noisy qubits (1000 at most). Hybrid algorithms address the issue by only running a small subroutine of the total computation on the quantum processor.

While promising, a major bottleneck in most hybrid algorithms is the embedding of classical data into quantum states. This process is crucial for enhancing the performance of quantum algorithms as qubits have different theoretical and physical properties compared to classical bits. Additionally, due to the limitations of NISQ devices the need for efficient data encoding is even more pronounced. In the rapidly evolving field of Quantum Machine Learning (QML), the integration of classical data into quantum systems presents unique challenges and opportunities. The effect of quantum embeddings on the performance of QML algorithms has been well studied in the literature[7]. We aim to look at the different families of embeddings and study their impact.

II. Quantum Machine Learning

Quantum Machine Learning can be broadly defined as a family of quantum algorithms with trainable parameters. Some of the earliest QML approaches include Quantum principal component analysis and Quantum Support Vector Machines. While these focused on quantum versions of classical ML algorithms, more recent work has focused on creating deep learning-inspired algorithms such as parameterized quantum circuits[2]. The general structure of a hybrid QML algorithm is shown in Figure 1.

While all of these approaches seem promising, there is an open question in the field regarding provable quantum advantage, i.e. the speedup of quantum algorithms over classical algorithms. Recent work [1] has proposed looking at the merits of QML beyond speedups and with a focus on the expressiveness of quantum models due to the non-linear feature space occupied by quantum states.

III. Quantum Embeddings

We can broadly define quantum embeddings as follows —

Definition 1 (Quantum Embedding): A mapping of data or functions between the classical and quantum domains.

This definition allows us to study quantum embeddings in a broader context and in relation to the various critical aspects of quantum computing as a whole.

A. Era based classification

The nature of quantum embeddings has been deeply influenced by the evolution of the era of quantum algorithms. The era of the algorithms informs the requirements and constraints of the embeddings used. We can see this when looking at the pre-Optimization and Optimization algorithm eras —

1) Pre-Optimization Embeddings: The pre-optimization era was characterized by handcrafted algorithms that placed consideration on hardware limitations and as such used functionally infinite qubits without any noise consideration. This was reflected in the embeddings being relatively simple and interpretable. Hence most algorithms utilized Angle encoding, Amplitude encoding, or Basis State encoding.

2643-1726/24/$31.00 ©2024 IEEE
DOI 10.1109/CASES60062.2024.00010

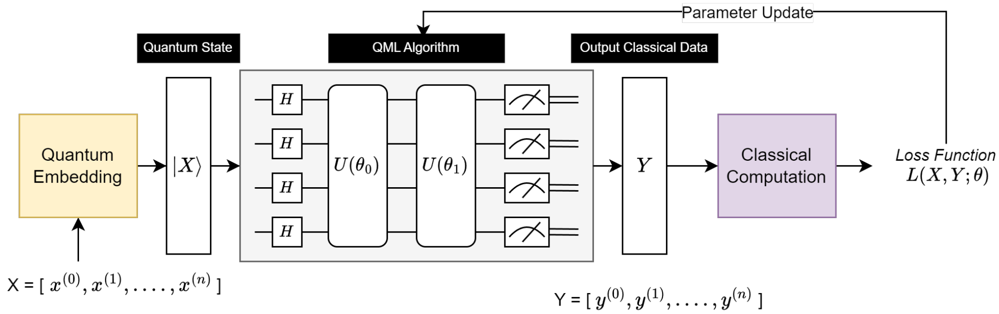

Fig. 1. General Pipeline of a Hybrid QML algorithm. A classical data vector (X) is converted to a quantum state ($|X\rangle$). This state is used as input to a QML algorithm (parameterized by θ), which is run on a quantum computer. The output of the quantum algorithm is then measured and converted back to a classical output(Y). The output is then used for classical computation. Finally, a loss function is used to update the parameters (θ) of the quantum algorithm.

2) Optimization Embeddings: The Optimization era included various hardware considerations and was influenced heavily by differential programming. The embeddings are indicative of this trend and are generally (1). Learned, (2). Applications specific and (3). Resource efficient. These embedding methods were critical parts of the QML algorithms they empowered and were often functions of trainable parameters. Some key examples include Quantum Kernels and Quantum Metric Learning [5].

B. Dataflow classification

Another method to study embeddings is to look at how they relate to the broader context of the data flow in a QML algorithm. Based on this we can classify embeddings into 3 categories —

1) General Vector (GV) Embeddings: These constitute the broadest category of embeddings and are used in a variety of QML algorithms. They are generally used to encode classical data vectors into quantum states. Some examples include Angle Encoding, Amplitude Encoding, Quantum Kernels [3], Metric Learning[5], and so on. A critical consideration in this category is the dimensionality of the data in relation to the number of qubits, and as such are generally paired with dimensionality reduction techniques such as deep neural networks.

2) Sliding Window Embeddings: Sliding window embeddings are similar to GV embeddings in structure but the algorithms they are paired with generally slide over larger classical data. These embeddings are not as limited by dimensionality as the pipelines compensate by patching or windowing the data. A key example of such embeddings is the Quanvolutional Neural Networks[4].

3) Functional Embeddings: Finally, functional embeddings are a specific class of embeddings that encode classical functions or classical data directly into the dynamics of a quantum system. The most well-known example of such an approach is Hamiltonian Embedding[6], both in the data and adiabatic context.

IV. Conclusion

We present here a small part of the study of quantum embeddings in the context of QML. The field is rapidly evolving and the embeddings are a critical part of the algorithms they empower. The understanding of the various embedding methods and their impact on the performance of QML algorithms is a critical part of the study of QML. We hope that this primer provides a good starting point for the study of quantum embeddings and their impact on the field of QML.

References

[1] Amira Abbas et al. “The power of quantum neural networks”. In: *Nature Computational Science* 1.6 (June 2021), pp. 403–409. ISSN: 2662-8457. DOI: 10.1038/s43588-021-00084-1. URL: http://dx.doi.org/10.1038/s43588-021-00084-1.

[2] Edward Farhi and Hartmut Neven. *Classification with Quantum Neural Networks on Near Term Processors*. 2018. arXiv: 1802.06002 [quant-ph]. URL: https://arxiv.org/abs/1802.06002.

[3] Vojtěch Havlíček et al. “Supervised learning with quantum-enhanced feature spaces”. In: *Nature* 567.7747 (Mar. 2019), pp. 209–212. ISSN: 1476-4687. DOI: 10.1038/s41586-019-0980-2. URL: http://dx.doi.org/10.1038/s41586-019-0980-2.

[4] Maxwell Henderson et al. *Quanvolutional Neural Networks: Powering Image Recognition with Quantum Circuits*. 2019. arXiv: 1904.04767 [quant-ph]. URL: https://arxiv.org/abs/1904.04767.

[5] Seth Lloyd et al. *Quantum embeddings for machine learning*. 2020. arXiv: 2001.03622 [quant-ph]. URL: https://arxiv.org/abs/2001.03622.

[6] Maria Schuld and Francesco Petruccione. *Machine learning with quantum computers*. Vol. 676. Springer, 2021, pp. 163–169, 238–243.

[7] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. “Effect of data encoding on the expressive power of variational quantum-machine-learning models”. In: *Physical Review A* 103.3 (Mar. 2021). ISSN: 2469-9934. DOI: 10.1103/physreva.103.032430. URL: http://dx.doi.org/10.1103/PhysRevA.103.032430.