

CSCI 4511W Writing Assignment 5

Brian Cooper
coope824@umn.edu
University of Minnesota

December 1, 2019

*Below is the “Related Work” section that will go into my final project report.
My problem is comparing the performance of quantum machine learning
algorithms to their classical counterparts on bioinformatics classification
problems.*

Related Work

Quantum computing, machine learning, and bioinformatics are established, yet highly plastic fields. Some researchers from the somewhat disparate disciplines have combined their efforts. Quantum machine learning, a synthesis of quantum computing and machine learning, is built upon the foundations of quantum mechanics and artificial intelligence. Such research efforts customarily uses classical machine learning approaches as a baseline for analyzing corresponding quantum counterparts, however they differ in their proposals for analysis and applied solutions. Biamonte, et al. [3] argue that machine learning has a bright future when utilizing quantum processing power, and they mention that significant advancements have been made when it comes to hardware. They clarify that certain devices (such as annealing quantum computers) *have* been built with apparently large qubit counts (2000, for example), although they do not communicate with each other well at such a scale yet, so there is ample room for improvement. Dunjko, et al. [5] primarily discussed complexity improvements for quantum methods. Finally, the paper by Schuld, et al. [10] solely zones in on classification and clustering methods (although in great depth) and emphasizes that we need to be careful in the wishful optimism of quantum computing’s usefulness to machine learning – especially considering its infancy. Despite this caution, the paper expresses hope and considers the field to have a promising future.

When it comes to practicality, it is argued in [3] that the foundation for quantum machine learning algorithms is strong, yet considerably limited by hardware and software challenges. An important point in [10] is that the ability

to simulate the actual, concrete learning process in quantum systems is unsatisfactorily documented, and thus requires more development.

Two dominant machine learning tasks, classification and clustering, are commonly used in modern problems. Although my problem is focused on classification, it is useful to recognize clustering methods within the domain to discover auxiliary information. Each is composed of varying depths of analysis in associated algorithms (such as support vector machines with different kernels for classification and k-means for clustering). Biamonte, et al. [3] provide several runtime complexity difference estimates for various quantum algorithms. Quantum support vector machines, for example, are estimated to perform faster with an $O(\log N)$ speedup (exponential speedup), albeit with certain constraints [2]. In [5], a quadratic improvement is common, and even exponential in relatively short time scales (converges to quadratic as $t \rightarrow \infty$). Schuld, et al. [10] describe the “toolbox of quantum algorithms” to actually be quite well-established and offer promising speedups. The authors detail in great depth various concrete examples of learning problems which are performed on classical computers and consider the same task on a quantum computer. An example they provide is from [9] which proposes that a quantum version of the nearest-centroid algorithm [1] is more efficient than the polynomial runtime required for the same task on a classical computer, even despite the auxiliary quantum operations that need to be performed.

Bioinformatics, in contrast to the analysis of quantum machine learning above, is built upon the foundations of biology, machine learning, and statistical methods. Similar tasks, namely classification and clustering, have been used since the advent of the study of bioinformatics. In [7], the structures of mRNA, tRNA, and rRNA are predicted by testing the energy levels required to form a structure. A structure can be predicted with a certain accuracy by measuring such energy levels. Confidence is proportional to a lower energy level, since it takes less energy to form that structure. It is unlikely for a structure requiring a high energy level to form since it is biologically inefficient, and this data is used to train a machine learning classification model using support vector machines. Other classification problems are introduced and suggested in [4], where support vector machines are demonstrated to perform well on tasks such as identifying molecules that modulate the function of certain protein receptors, compared to other approaches such as neural networks.

Changing perspective, clustering methods are also used in bioinformatics. Higham, et al. [6] cover in great depth the mechanics of spectral clustering, and offer light into its use in bioinformatics. In particular, they address gene expression activity across various tissue samples, and explain how clustering is a natural approach for resolving such gene activity. In [8], the overwhelming amount of modern data (“big data”) is addressed and clustering approaches are suggested for reducing the complexity of bioinformatics problems. A minimum spanning tree representation is utilized within a clustering approach within a

software called CLUMP are introduced, and drastic time complexity improvement is promised with using such software.

Combining bioinformatics with quantum machine learning is relatively new and unexplored space, at least in the public domain. Despite this, I think it is clear that there is much potential in the future of computing.

References

- [1] Rocchio classification. <https://nlp.stanford.edu/IR-book/html/htmledition/rochio-classification-1.html>, April 2009.
- [2] Scott Aaronson. Read the fine print. *Nature*, 2015.
- [3] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.
- [4] E. Byvatov and G. Schneider. Support vector machine applications in bioinformatics. 2003.
- [5] Vedran Dunjko, Jacob M Taylor, and Hans J Briegel. Quantum-enhanced machine learning. *Physical review letters*, 117(13):130501, 2016.
- [6] Kalna Gabriela Higham, Desmond and Milla Kibble. Spectral clustering and its use in bioinformatics. 2007.
- [7] B. Buxton J. Ward, L. McGuffin and D. Jones. Secondary structure prediction with support vector machines. 2003.
- [8] Mao Fenglou Wu Hongwei Olman, Victor and Xu Ying. Parallel clustering algorithm for large data sets with applications in bioinformatics. 2009.
- [9] M. Mohseni S. Lloyd and P. Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. 2013.
- [10] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, 2015.