ISSN XXXX XXXX © 2018 IJESC

Research Article Volume 8 Issue No.8

Text Classification using Quantum Machine Learning: A Review

Aniruddh Goteti
Blekinge Institute of Technology, Sweden

Abstract:
This paper gives a brief overview of the quantum algorithms which are proposed to solve basic machine learning problems such as pattern recognition and classification. Quantum computing and quantum machine learning are briefly introduced. A narrative literature review is done to find out the performance evaluation of quantum machine learning algorithms in compari-son to classical machine learning algorithms.

Keywords: Quantum Computing, Machine Learning, Quantum Machine Learning

I. INTRODUCTION

Machine learning has fundamentally changed the way humans interact with and relate to data. Information is fundamentally governed by the laws of physics. The laws are quantum mechanical at the scales of present day information processing technology, in contrast to the more familiar classical physics at the human scale. The interface of quantum physics and machine learning naturally goes both ways: machine learning algorithms find application in understanding and controlling quantum systems and, on the other hand, quantum computational devices promise enhancement of the performance of machine learning algorithms for problems beyond the reach of classical computing [10].

Recent progress implies that a crossover between machine learning and quantum information processing benefits both fields. Traditional machine learning has dramatically improved the bench marking and control of experimental quantum computing systems, including adaptive quantum phase estimation and designing quantum computing gates. On the other hand, quantum mechanics offers tantalizing prospects to enhance machine learning, ranging from reduced computational complexity to improved generalization performance. The most notable examples include quantum enhanced algorithms for principal component analysis, quantum support vector machines, and quantum Boltzmann machines [10].

Thus the **aim** of this paper is to compare the performance of a quantum algorithm with a classical algorithm when solving a machine learning problem. Thus **objectives** of this paper are identified as follows:

- To briefly describe quantum information processing and quantum machine learning.
- To define a research problem to be solved using quantum machine learning.
- To conduct a narrative literature review to answer the research question.

II. MACHINE LEARNING

According to Arthur Samuel, Machine learning is a sub field of computer science which gives computers the ability to learn without being explicitly programmed. It explores the study and construction of algorithms that can learn from data and make predictions on it. Moreover, machine learning is usually employed in computing tasks where it is difficult to achieve good performance with programming explicit algo-rithms. Some of these tasks include computer vision, email filtering, OCR (optical character recognition) etc.

III. QUANTUM COMPUTATION

Quantum computing is computing using quantum-mechanical phenomena, such as superposition and entanglement. A quantum computer is a device that performs quantum computing. They are different from binary digital electronic computers based on transistors. Whereas common digital computing requires that the data be encoded into binary digits (bits), each of which is always in one of two definite states (0 or 1), quantum computation uses quantum bits, which can be in super-positions of states.

Quantum computation is the field that investigates the computational power and other properties of computers based on quantum-mechanical principles. An important objective is to find quantum algorithms that are significantly faster than any classical algorithm solving the same problem. Three different motivations for studying quantum computers, from practical to more philosophical:

- The process of miniaturization that has made current classical computers so powerful and cheap, has already reached micro-levels where quantum effects occur. Chip-makers tend to go to great lengths to suppress those quantum effects, but instead one might also try to make good use of them [4].

- Making use of quantum effects allows one to speedup certain computations enormously (sometimes exponentially), and even enables some things that are impossible for classical computers [4].

- Finally, one might state the main goal of theoretical computer science as study the power and limitations of the strongest-possible computational devices that nature allows us. Since our current understanding of nature is quantum mechanical, theoretical computer science should be studying the power of quantum computers, not classical ones [4].

IV. QUANTUM MACHINE LEARNING

Quantum machine learning is an emerging interdisciplinary research area at the intersection of quantum physics and machine learning.

Quantum machine learning (QML) is a sub-discipline of quantum information processing research, with the goal of developing quantum algorithms that learn from data in order to improve existing methods in machine learning.

A quantum algorithm is a routine that can be implemented on a quantum computer, a device that exploits the laws of quantum theory in order to process information[10].

A number of quantum algorithms have been proposed for various machine learning models such as neural networks, support vector machines, and graphical models, some of which claim run-times that under certain conditions grow only logarithmic with the size of the input space and/or data-set compared to conventional methods. A crucial point for run-time considerations is to find a procedure that efficiently encodes classical data into the properties of a quantum system.

Besides finding quantum algorithms for pattern recognition and data mining, QML also investigates more fundamental questions about the concept of learning from the perspective of quantum theory. Sometimes the definition of QML is extended by research that applies machine learning to quantum information, such as is frequently done when the full evolution or state of a quantum system has to be reconstructed from limited experimental data.

V. TEXT CLASSIFICATION PROBLEM

Text classification is the process of classifying documents into predefined categories based on their content. It is the automated assignment of natural language texts to predefined categories. Text classification is the primary requirement of text retrieval systems, which retrieve texts in response to a user query, and text understanding systems, which transform text in some way such as producing summaries, answering questions or extracting data. Existing supervised learning algorithms to automatically classify text need sufficient documents to learn accurately [2].

This text classification problem is chosen as to evaluate the performance of quantum algorithms. Thus defining a research question:

RQ: *What is the performance evaluation of quantum algorithms when compared to performance evaluation of classical algorithms when classifying text?*

In order to answer this research question, a narrative literature review is carried out.

VI. LITERATURE REVIEW

A literature review is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic.

The parameters of literature review are as follows:

1) Keywords: Quantum Computers, Machine Learning

2) Inclusion Criteria: Computer Science, Citations having more than 5, All document types which gave 20 papers as a result, out of which 8 relevant papers were chosen for the literature review.

The storytelling or motivation behind the literature review is done in the following way:

1) **Article 1** describes the road-map to address future computing issues, which also describes why quantum computing will be common approach to solve computing issues in the future.

2) **Article 2** describes how quantum computers or annealers can be used to solve complex problems, from the perspective of NASA.

3) **Article 3** introduces basic quantum algorithms and emphasizes on importance of having a quantum computer in the near future.

4) **Article 4** introduces and reviews five quantum versions of machine learning algorithms.

5) **Article 5** surveys the use of quantum algorithms to solve big data problems, and compares it to use of classical algorithms to solve big data problems.

6) **Article 6** introduces quantum version of k-means clustering algorithm and compares it classical version of k-means algorithm to solve a pattern recognition problem.
7) **Article 7 and Article 8** solves a text classification problem using quantum versions of k-nearest neighbour algorithm and k- means clustering algorithm to solve text classification problem and then compares it to the classical versions.

A. Article 1

Article 1[9] proposes a 15 year old road map to a ad-dress escalating complex computing issues such as stalled device-physics advances coupled with big data demands, novel machine-learning problems, and complex software paradigms. Potential solutions range from new transistor technology to quantum computing.

Computing has always been quick to evolve because of changes in semiconductor devices and technologies no more so than at the peak of the high-performance computing (HPC) pyramid. But the current general consensus is that HPC is in its last generation, and a crisis is coming [9] and article proposes solutions to overcome that.

The article describes the challenges which the future computing poses for system software. Problem definition, for example, will require problem-definition languages that are independent of the computation platform. Selecting a computing platform will require making an intelligent decision about which platform will most efficiently execute the problem and then compiling the problem definition into executable code for the chosen platform. Significant challenges remain in creating new OS's and virtualization for these very different computing approaches[9].

The article also expanded the discussion of issues and brought the idea of rebooting computing to more academic, industry, and government research communities.

B. Article 2

There were several instances of quantum algorithms introduced in last few decades which outperformed the performance of classical algorithms, but empirical testing is only possible as quantum computation hardware is build. Article 2[8] introduces several opportunities and challenges in quantum computing, within the perspective of NASA. In order to successfully launch most of the NASAs missions, lots of complex computation problems are needed to be solved.To support NASA's substantial computational needs, NASA Ames Research Center has a world-class super-computing facility with one of the worlds most powerful supercomputers[8].

Quantum Annealers were used as a quantum computation hardware to investigate the quantum annealer performance on small problems from the domains of planning and scheduling, fault diagnosis and machine learning.This article concentrates on the teams quantum annealing work, with only brief mention of research related to capabilities of other near-term quantum computational hardware that will be able to run quantum heuristic algorithms beyond quantum annealing [8].

The power of quantum computation comes from encoding information in a non-classical way, in qubits, that enable computations to take advantage of purely quantum effects, such as quantum tunneling, quantum interference, and quantum entanglement, that are not available classically. The beauty of quantum annealers is that users can program them without needing to know about the underlying quantum mechanical effects. Knowledge of quantum mechanics aids in more ef-fective programming, just as an understanding of compilation procedures can aid classical programming, but it is not neces-sary for a basic understanding [8].

The paper concludes by discussing the future scope of quantum annealers as more sophisticated quantum annealers will be available in the near future thus broadening the areas in which quantum computation has clear applications.

C. Article 3

Article 3[5] discusses two types of quantum algorithms (1) a simple quantum algorithm and (2) a quantum search algorithm. Recent advances in quantum computing and intelligence are reviewed and recent advances and challenges for researchers are also reviewed.

Quantum computing (QC) emerged when scientists were pondering the fundamental limits of computation. They understood that if technology continued to abide by Moores Law, then the continually shrinking size of circuitry packed onto silicon chips would eventually reach a point where individual elements would be no larger than a few atoms [5]. This miniaturization process is now reaching a limit, a quantum threshold below which transistors will cease to function. Present state-of-the-art components possess features only a few hundreds of nanometers across [5]. If these chips were to be miniaturized further to the scale of tens of nanometers then their operation would be disrupted by the emergence of quantum phenomena,such as electrons tunneling through the barriers between wires. Here, a problem arose because at the atomic scale the physical laws that govern the behavior and properties of the circuit are inherently quantum mechanical in nature, not classical [5].This then raised the question of whether a new kind of computer could be devised based on the principles of quantum physics [5].

The article concludes the advancements in quantum information processing which has made numerous promising advancements since its conception, including the building of 2- and 3-qubit quantum computers capable of some simple arithmetic and data sorting. However, a few potentially large obstacles still remain that prevent us from building a quantum computer that can rival todays modern digital computer.

Among these difficulties,error correction, decoherence, and hardware architecture are probably the most formidable [5]. Before any quantum computer will be capable of solving hard problems, research must devise a way to maintain decoherence and other potential sources of error at an acceptable level[5].

D. Article 4

Article 4 This article[7] gives a brief introduction of quantum machine learning and proposes few machine learning methods using quantum computation methods. Differences between classical machine learning algorithms and quantum machine learning algorithms are also described. The quantum version of machine learning algorithms introduced in this article are:

1) Quantum versions of k-nearest neighbour methods
2) Quantum computing for support vector machines
3) Quantum algorithms for clustering
4) A quantum decision tree
5) Quantum state classification with Bayesian methods
6) Hidden quantum Markov models

This introduction into quantum machine learning gave an overview of existing ideas and approaches to quantum machine learning. The focus is this paper was thereby on supervised and unsupervised methods for pattern classification and clustering tasks, and it is, therefore, by no means a complete review [7]. In summary, there are two main approaches to quantum machine learning. Many authors try to find quantum algo-rithms that can take the place of classical machine learning algorithms to solve a problem, and show how an improvement in terms of complexity can be gained [7]. This is dominantly true for nearest neighbour, kernel and clustering methods in which expensive distance calculations are sped up by quantum computation [7].

E. Article 5

Big Data is a term which denotes data that is beyond storage capacity and processing capabilities of classical computer and getting some insight from large amount of data is a very big challenge at hand. Quantum Computing comes to rescue by offering a lot of promises in information processing systems, particularly in Big Data Analytics. In this paper, the available literature is reviewed on Big Data Analytics using Quantum Computing for Machine Learning and its current state of the art. Quantum Walks used to construct Quantum Artificial Neural Networks, which exponentially speed-up the quantum machine learning algorithm is discussed. Quantum Supervised and Unsupervised machine learning and its benefits are com-pared with that of Classical counterpart. The limitations of some of the existing Machine learning techniques and tools are enunciated, and the significance of Quantum computing in Big Data Analytics is incorporated. Being in its infancy as a totally new field, Quantum computing comes up with a lot of open challenges as well. The challenges, promises, future directions and techniques of the Quantum Computing in Machine Learning are also highlighted [3].

In this paper, Review of the available literature on Big Data Analytics using Quantum Computing for Machine Learning and its current state of the art is carried Quantum Supervised and Unsupervised machine learning is also discussed and com-pared its benefits with respect to the Classical Supervised and Unsupervised machine learning techniques. The limitations of some of the existing Machine learning techniques and tools are also enunciated, and the significance of Quantum computing in Big Data Analytics is incorporated [3].

F. Article 6

In this paper quantum algorithms are presented for perform-ing nearest-neighbor learning and k-means clustering.

It is found in this paper that quantum algorithms for machine learning can provide algorithmic improvements over classical machine learning techniques. The algorithms proposed in this paper are a step toward blending fast quantum methods with proven machine learning techniques [1]. Further work will be needed to provide a complete cost assessment in terms of the number of elementary gate operations and logical qubits needed to practically achieve these speedups using a fault tol-erant quantum computer. Possibilities for future work include examining width/depth trade-offs that arise when executing these algorithms on quantum computers and how well the algorithm performs in cases where the oracle represents a database.

G. Article 7

This paper gives a brief introduction into quantum machine learning using the example of pattern classification. A quantum pattern classification algorithm is introduced for measuring the Hamming distance on a quantum computer is introduced and its advantages using handwritten digit recognition as from the MNIST database is discussed [6].

The quantum pattern classification (QPC) algorithm we present here uses the same distance-based classification princi-ple as kNN, only that instead of choosing nearest neighbours, the distance of the entire training vector set is considered. It draws on an algorithm presented in the context of quantum associative memory in [6]. The idea is to create a superposition of the training data set and write the Hamming distance to the input state into the amplitude of each vector in the superposi-tion. Measuring the class-qudit then retrieves the desired class with the highest probability. Even more, if repeated enough times to achieve statistical significance, the algorithm leads to a probability distribution containing information on how close each class members are to the input vector.

It is concluded that approximately 50% of the digits of a test set of 100 examples can be classified correctly by this method (using a training set of 400 examples) using quantum computation algorithm for text classification. Still, the MNIST example helps to demonstrate how quantum computing offers a general advantage in cases of ambiguous inputs [6]. This paper also describes future works that would have to extend the quantum algorithm to allow for continuous inputs, and investigate ways to exploit its advantages using more complex distance measures[6].

H. Article 8

Text classification is one of key problems in pattern recog-nition. The KNN algorithm is a widely used text classification algorithm, because it is simple,valid and non parameters. The main idea of KNN algorithm is to calculate the similarity between the new sample with unknown class label and the training samples, and choosing the class label of the highest k nearest neighbors as the new samples class label. However, the text contains hundreds and thousands of features. The similarity computing in large numbers of vector will cost many time. In fact that, many machine learning algorithms are unable to manipulate and compute large number of vectors in high dimensional space. Quantum Computing algorithms are good at computing high-dimensional vectors in large tensor product spaces. Therefore, this paper introduces a KNN algorithms based on quantum computing, which uses fidelity to compute the similarity between two quantum states[2].

In this paper, the performance of several real world problems are evaluated by simulating the noise incurred by using coher-ent amplitude estimation when solved by Quantum Machine Learning Algorithms and Classical Ma-chine Learning Algo-rithms. As an example, classifying handwritten digits from the MNIST digits database wherein the problem is: given a training set of M handwritten digits and their labels (even or odd), assign a label(even or odd) to an unlabeled test digit. Each digit is represented by a 256-dimensional feature unit vector of pixel values[2].

In this paper, quantum algorithms for performing nearest neighbor classification and k-means clustering are presented that promise significant reductions in their query complexity relative to over their classical counterparts.

The algorithms introduced in this paper enable classification and clustering over data sets with both a high-dimensional feature space as well as a large number of training examples. Computation of distances is extremely common in machine learning algorithms; developing two fast methods for comput-ing the distance between vectors on a quantum computer that can be implemented coherently [2].

It is finally proved that quantum algorithms for machine learning can provide algorithmic improvements over classical machine learning techniques.

VII. RESULTS

The outcome of literature review is as follows:

1) Quantum Computation will be a mandatory approach to solve complex computation problems in the future.
2) Quantum computer provide better computation power than classical computers.
3) For solving a text classification problem, quantum algorithms for machine learning can provide algorithmic

improvements over classical machine learning techniques and accuracy for classifying text is higher for quantum algorithms than classical algorithms.

Thus to answer the research question, the classification performance of the quantum machine learning algorithm is better than the classical machine learning algorithm.

VIII. CONCLUSION

This paper reviews why quantum computation will be necessary in the near future and how quantum computation algorithms provides better algorithmic improvements than classical computation problems for solving machine learning problems.

The computation power of quantum algorithms is computed in quantum annealers or quantum simulation software, thus the results are not completely accurate. But as quantum computers are not available as of now to analyze and evaluate quantum computation algorithms, it is not possible to know the actual computational power of the quantum algorithms, which identifies this as a scope for the future work.

REFERENCES

[1] N. Wiebe, A. Kapoor, and K. Svore, Quantum Algorithms for Nearest-Neighbor Methods for Supervised and Unsuper-vised Learning, arXiv:1401.2142 [quant-ph], Jan. 2014.

[2] S. Shang, M. Shi, W. Shang, and Z. Hong, A text classification algorithm based on quantum information, in 2015 11th International Conference on Natural Computation (ICNC), 2015, pp. 381384.

[3] T. A. Shaikh and R. Ali, Quantum Computing in Big Data Analytics: A Survey, in 2016 IEEE International Conference on Computer and Information Technology (CIT), 2016, pp. 112115.

[4] qcnotes.pdf.

[5] Z. Ezziane, Quantum computing measurement and intel-ligence, Int J Quantum Chem, vol. 110, no. 5, pp. 981992, 2010.

[6] M. Schuld, I. Sinayskiy, and F. Petruccione, Quantum computing for pattern classification, vol. 8862. Springer Verlag, 2014.

[7] M. Schuld, I. Sinayskiy, and F. Petruccione, An introduc-tion to quantum machine learning, Contemp Phys., vol. 56, no. 2, pp. 172185, 2015.

[8] R. Biswas et al., A NASA perspective on quantum computing: Opportunities and challenges, Parallel Comput, vol. 64, pp. 8198, 2017.

[9] T. M. Conte, E. P. DeBenedictis, P. A. Gargini, and E. Track, Rebooting computing: The road ahead, Computer, vol. 50, no. 1, pp. 2029, 2017.

[10] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum Machine Learning, Nature, vol. 549, no. 7671, pp. 195202, Sep. 2017.