Hybrid quantum-classical metaheuristics for automated machine learning applications

Von Dollen, D.J.

Chapter 1

Introduction

1.1 Background

1.1.1 Artificial Intelligence

The modern origin of the branch of computer science known as artificial intelligence (AI) can be traced back to Alan Turing's paper "Computing machinery and intelligence" [216]. This paper defined various objectives for machine intelligence, exploring the potential for computers to learn, play chess, and create music, and introduced the Turing test as a criterion for machine intelligence. In this test, a person (also known as an interrogator) converses with both a machine and a human without knowledge of the identity of either; if the interrogator cannot consistently distinguish the machine from the human, the machine is considered to have passed the test. Although the Turing test has critics [191], the relevance of such tests is increased by the recent prominence of large language models as forms of artificial intelligence agents. At the time of this writing, these models increasingly demonstrate a human-like ability to generate and create responses to queries in natural language and text [55]. Turing's work significantly influenced artificial intelligence and cognitive science, sparking debates which are ongoing today on the nature of machine intelligence, how to measure it, and how to design and build systems with a human-level ability to reason and plan [47].

Within the field of AI, there exist research directions towards what is known as *general artificial intelligence* (AGI) and *narrow artificial intelligence*. AGI refers to a type of intelligence that can reason, learn, plan and apply knowledge across a wide range of tasks, mimicking human cognitive abilities, while narrow artificial intelligence

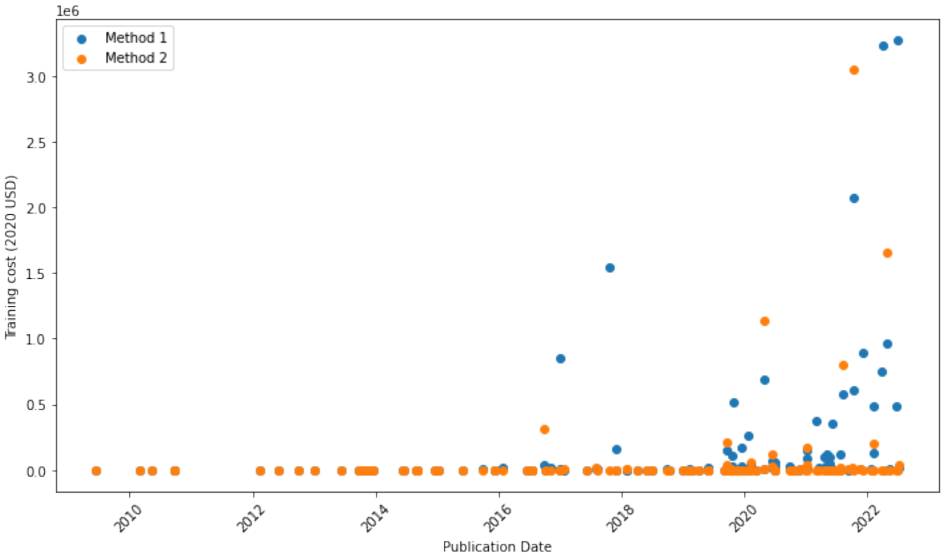

Figure 1.1: Plot of training costs in USD using data collected from academic and industrial publications for machine learning models from [43].

is designed to perform specific tasks within a limited domain or set of activities. AI also encompasses various approaches to learning, such as *knowledge-based artificial intelligence* and *statistical machine learning* [177]. Biologically inspired methods of computing such as *neural networks*, *genetic algorithms*, and *swarm intelligence* are also considered within and adjacent to this field of computer science.

In practical settings, machine learning and other forms of narrow AI have been successfully applied within industry and science [104]. With the advent of increases in the collection and storage of data, algorithms have grown increasingly advanced with respect to prediction tasks such as classification, regression, and clustering in application towards products and processes. As such, as the field strives to solve for AGI, often more powerful forms of narrow AI are produced, opening up new areas of practical applications [78].

1.1.2 Challenges in Machine Learning and Artificial Intelligence

In addition to the pursuit of developing models with higher generalization and representation capabilities, some of the challenges in the field of machine learning include the explainability of model output, data quality, data bias in machine learning [141, 33], and access to computational resources. With the recent development of large language models as text-based artificial intelligence agents, computational cost and considerations of environmental impact also pose a considerable challenge [131]. In recent years, the dollar cost to train machine learning systems has increased by orders of magnitude as ever larger models have been trained by fleets of processing units in cloud data centers on vast amounts of data, as shown in Figure 1.1 [43].

From an environmental point of view, the training of machine learning models is expensive with respect to CO_2 emissions. Work in 2019 [207] showed that training a large transformer model with neural architecture search (NAS) emitted 626,155 pounds of CO_2 into the environment, many times the average household amount in the United States.

The deployment of machine learning systems also presents a challenge for the practitioner. The identification and extraction of correlations and patterns from data sets using machine learning is a core process in data science and engineering applications. Typically, the process follows a flow of *data preprocessing*, *feature selection and engineering*, *model selection*, *hyperparameter optimization*, *model evaluation*, *model deployment*, and *model monitoring*, as shown in Figure 1.2. Each step in this process depends on the skill and prior knowledge of the practitioner and risks introducing bias or error.

Automated machine learning (AutoML) [96] is a subfield of machine learning that automates the design of machine learning systems and process workflows. AutoML aims to speed up and simplify the development and deployment of machine learning models while addressing challenges such as reproducibility, scalability, and efficiency. AutoML enables researchers and machine learning practitioners to effectively handle the task of designing and deploying systems of increased size and complexity.

1.1.3 Quantum Computing

In recent years, quantum computing research and practical applications have accelerated within the computer science and physics research communities and industry [5] [145], and naturally the question arises whether or not access to quantum computation may provide benefits toward artificial intelligence research and applications. Quantum computers have the potential to leverage quantum effects for computation and, therefore, may be more efficient at solving certain classes of problems over classical computers. This is enabled by information processing that is fundamentally different from classical information processing. Classical systems represent information as bits.

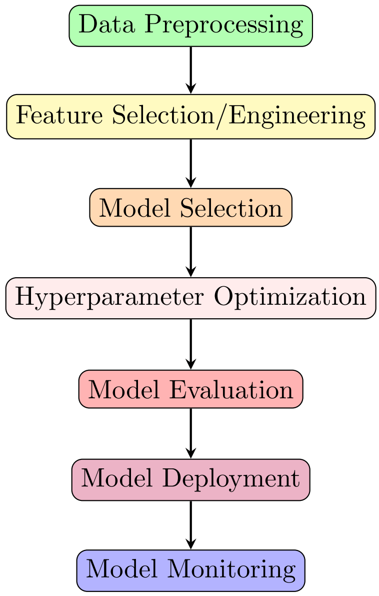

Figure 1.2: Diagram of process flow for machine learning model development for real world applications.

Quantum information is represented by qubits, where n qubits are represented as a vector in a Hilbert space $\mathcal{H}$. We use bra-ket notation (also known as Dirac notation) here to represent the quantum states, popularized by the physicist Paul Dirac in 1939 [54].

To illustrate this difference in information processing capability, let a complex row vector representing a quantum state be notated by $\langle|$ and a complex column vector be notated as $|\rangle$. A qubit, the basic unit of quantum information, can exist in a superposition of states. This superposition is represented as a combination of two possible states: $|0\rangle$ and $|1\rangle$. The likelihood of the qubit being in one state or the other is determined by two values, called probability amplitudes, which must always sum to 1. The squared magnitudes of these amplitudes give the probability of observing a particular basis state for the system upon measurement. The difference in representation power can grow exponentially in the number of qubits in comparison to classical bits, where a quantum system of n qubits in superposition can represent states in a 2^n-dimensional complex vector space in $\mathcal{H}$ versus a classical system represented by n bits.

In theory, certain routines of quantum computation may enable certain calculations that require exponentially less steps than classical computation. An example of this

Algorithm	**Classical**	**Quantum**
Unordered Database Search [82]	$O(n)$	$O(\sqrt{n})$
Integer Factorization [197]	$O\left(\exp\left(c(\log n)^{1/3}(\log\log n)^{2/3}\right)\right)$	$O((\log n)^3)$
Solving Linear Systems [86]	$O(n^3)$	$O(\log(n))$
Simulation of Quantum Systems [66]	$O(\exp(n))$	$O(\text{poly}(n))$

Table 1.1: Table of runtime in the size of the input n computational complexity comparisons between classical and quantum algorithms. Note that a fair comparison between classical and quantum is dependent on details of the problem and assumptions around solutions obtained to the problem. For example, the HHL algorithm requires sparsity in the input matrix as a condition to solve linear systems, and may introduce run-time overhead in producing an efficient encoding for quantum states [86]. For integer factorization, n is the number to be factored, and the classical complexity is for the general number field sieve.

is Shor's algorithm for factoring large numbers [197] as shown in Table 1.1. Quantum computation may also enable the processing of information from data that is quantum generated [94]. Finally, quantum computation for certain calculations may be more energy efficient than classical analogues[40]. We provide an overview of the theoretical speedups for classical versus quantum algorithms and applications in Table 1.1.

At the time of this writing, empirically assessing the efficacy and improvements of artificial intelligence and machine learning augmented with quantum computation remains challenging. Although theoretical achievements have shown increasing evidence for performance gains, success in early-stage quantum computation is limited by ongoing hardware development constraints. This includes quantum processing units (QPUs) or chips with limited numbers of qubits, short coherence times, sensitivity to environmental noise, cost, access, and connectivity. These hardware limitations are representative of the era that we are currently in for quantum computing, which John Preskill termed the "noisy, intermediate-scale, quantum era (NISQ)" [164].

1.2 Motivation for Research

The goal of this thesis is to understand where quantum-inspired computation, combinations of quantum and classical data, and hybrid learning models may offer benefits when applied to real-world optimization and machine learning problems, under the constraints imposed by noisy intermediate-scale quantum (NISQ) devices. We hypothesize that utilization of quantum computation may provide machine learning practitioners with tools that are problem-specific, exploit quantum effects to understand problem structures, and allow for the identification of patterns in quantum data.

As such, the goal of this thesis is to identify and explore problems within the machine learning development pipeline where quantum or quantum-inspired computation may offer benefits. It aims to study and assess any quantitative or qualitative gains access to quantum or quantum inspired computation may offer in the process of developing machine learning and optimization algorithms for real-world applications. This leads us to our central problem statement and research questions.

Problem Statement. *Are there areas of automated machine learning, or areas along the pipeline for developing machine learning models, for which applications of quantum-inspired or quantum computation may offer benefits?*

1.3 Research Questions

In this section we pose the central research questions for this thesis. These questions center around the central problem statement of the study of the intersections of quantum computing, machine learning, and optimization.

RQ1 (Chapter 3 and 6) *Can we observe qualitative or quantitative performance gains in leveraging quantum, quantum inspired, or hybrid-quantum classical algorithms to optimize, identify and search combinations of correlated data?*

RQ2 (Chapter 3) *Can access to quantum or quantum inspired computation generate any qualitative or quantitative performance gains in preprocessing steps such as feature selection and feature engineering in building machine learning models?*

RQ3 (Chapter 4 and 5) *Can access to fully quantum, quantum inspired, or hybrid quantum-classical computation generate any qualitative or quantitative performance gains in optimization metaheuristics?*

RQ4 (Chapter 6) *Can we observe any qualitative performance separation between classical and quantum models for training and inference on classical and quantum data?*

1.4 Outline

We start with some preliminaries and background definitions in chapter 2. In this chapter, we describe and define different areas of machine learning, and briefly discuss the area of deep learning. We define and discuss optimization, and different metaheuristics used in solving optimization problems. We also discuss the area of automated machine learning, an emerging sub-field of artificial intelligence which combines optimization and machine learning to automate areas of the machine learning development pipeline for problems areas such as feature selection and hyperparameter optimization.

In chapter 2, we also include some background on quantum computing. We give a brief introduction and overview of quantum mechanics and quantum mechanical effects used toward quantum computation. We then provide an overview of complexity classes and highlight classes of problems with which access to quantum computation may provide efficient solutions. We then give an overview of quantum logic gates and circuits and discuss hardware for gate-model quantum computing. We discuss adiabatic computing and quantum annealing as a metaheuristic for solving optimization problems. We give an overview of the emerging research area of quantum machine learning, discussing quantum kernel methods and variational quantum algorithms. We discuss the role of quantum data for machine learning and give an overview of combinations of classical-quantum learning algorithms and data. We discuss quantum-inspired methods, presenting an overview of tensor train optimization.

In chapter 3, we address RQ1 and RQ2. We encode the feature selection problem as a QUBO with various distance measures. We examine the effects of various measures when applied to the real-world problem of feature selection for machine learning modeling of vehicle price prediction. We analyze correlations in the features of the *UCI auto* dataset [183], and compare the performance of classical estimators trained on selected data using various methods of selection of quantum and classical features.

In chapters 4 and 5, we address RQ3. We formulate the selection operator for evolutionary algorithms as a QUBO. We benchmark the solvers used for obtaining solutions for QUBO, including the D-Wave quantum annealer, classical solvers, and hybrid quantum-classical solvers. In chapter 4 we use quantum enhanced selection for genetic algorithms for pseudo-Boolean optimization, and in chapter 5 we extend this selection operator to evolution strategies for optimization over continuous domains. We use the selection operator as a mechanism for evolution strategies towards searching for global optima from a variety of multi-modal test objective functions and objective functions with perturbed fitness. We apply quantum-enhanced evolution strategies toward real-world problems such as hyperparameter optimization and neuroevolution of classical neural networks and report on our results.

In chapter 6, we address RQ1 and RQ4. We investigate methods or approximating kernels for sparse Gaussian process regression and compare quantum feature maps to random Fourier features for kernel approximation. We use some of the techniques from previous chapters to formulate QUBO to select subsets of data for sparse estimators and report on the results. We use quantum- and QUBO-enhanced estimators as surrogate models for Bayesian optimization of multi-modal continuous black box functions. We apply these to the real-world problem area of hyperparameter optimization of deep learning models on tabular data. We also explore the effects of learning with quantum data, creating stilted quantum datasets using a projected quantum kernel, and examine the performance characteristics between classical, hybrid quantum-classical, and fully quantum estimators.

The goal of this thesis is to demonstrate where near-term quantum and quantum-inspired computation may show qualitative improvements in the processes of optimization, machine learning, and automated machine learning. This demonstration may illustrate how using quantum and quantum-inspired computing systems can help practitioners extract patterns and correlations in data, potentially improving the search for data, models, and model hyperparameters within the machine learning development pipeline. Crucially, we do not claim that the goal of showing quantum supremacy for all cases of optimization and machine learning is within the scope of this thesis. Rather, within the context of automated machine learning, we explore where near-term quantum devices may yield quality improvements despite current hardware limitations and identify promising directions for future research as these constraints are overcome. We highlight that although near-term quantum chips face strict constraints on their applicability, this thesis may explore practical near-term strategies for hybridizing quantum and classical techniques in machine learning and optimization. These strategies may prove useful to machine learning practitioners as the size, coherence, and robustness to external noise of quantum computing systems improve.

1.5 Contributions of this Thesis

This thesis covers published work completed in pursuit of the research objectives. Subsequently, these works map to chapters in this thesis. Chapter 3 presents work from [228]. The chapters 4 and 5 present the work of [225, 227] in the following. Finally, [229] is presented in Chapter 6.

1. [228] D. Von Dollen, D. Weimer, F. Neukart, and T. Bäck. Predicting vehicle prices via quantum-assisted feature selection. *International Journal of Information Technology*, 15(6):2897–2905, 2023. doi: 10.1007/s41870-023-01370-z. URL `https://doi.org/10.1007/s41870-023-01370-z`.

2. [225] D. Von Dollen, S. Yarkoni, D. Weimer, F. Neukart, and T. Bäck. Quantum-enhanced selection operators for evolutionary algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '22, page 463–466, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392686. doi: 10.1145/3520304.3528915. URL `https://doi.org/10.1145/3520304.3528915`.

3. [227] D. Von Dollen, R. Brasher, F. Neukart, and T. Bäck. Hyperparameter optimization and neuroevolution with binary quadratic meta-heuristics and evolution strategies. In *2023 7th IEEE Congress on Information Science and Technology (CiSt)*, pages 536–540, 2023. doi: 10.1109/CiSt56084.2023.10409984.

4. [229] D. Von Dollen, R. Brasher, H. Wang, F. Neukart, and T. Bäck. Hybrid quantum-classical surrogate models for bayesian optimization of machine learning algorithms. *Quantum Machine Intelligence*, submitted.

1.6 Other Work by the Author

During the course of the study for this thesis, the author completed additional work in collaboration with other coauthors. These works are included in the following.

1. [30] M. Broughton, G. Verdon, T. McCourt, A. J. Martinez, J. H. Yoo, S. V. Isakov, P. Massey, R. Halavati, M. Y. Niu, A. Zlokapa, E. Peters, O. Lockwood, A. Skolik, S. Jerbi, V. Dunjko, M. Leib, M. Streif, D. V. Dollen, H. Chen, S. Cao, R. Wiersema, H.-Y. Huang, J. R. McClean, R. Babbush, S. Boixo, D. Bacon, A. K. Ho, H. Neven, and M. Mohseni. Tensorflow quantum: A software framework for quantum machine learning, 2021. URL `https://arxiv.org/abs/2003.02989`.

2. [179] A. Sagingalieva, M. Kordzanganeh, A. Kurkin, A. Melnikov, D. Kuhmistrov, M. Perelshtein, A. Melnikov, A. Skolik, and D. V. Dollen. Hybrid quantum resnet for car classification and its hyperparameter optimization. *Quantum Machine Intelligence*, 5(2):38, 2023. doi: 10.1007/s42484-023-00123-2. URL `https://doi.org/10.1007/s42484-023-00123-2`.

3. [244] S. Yarkoni, A. Alekseyenko, M. Streif, D. Von Dollen, F. Neukart, and T. Bäck. Multi-car paint shop optimization with quantum annealing. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 35–41, 2021. doi: 10.1109/QCE52317.2021.00019.

4. [146] F. Neukart, D. V. Vreumingen, D. V. Dollen, A.-C. Voigt, M. Hartmann, and C. Othmer. System and method for finite elements-based design optimization with quantum annealing. United States Patent, November 2022. Available at `https://patents.google.com/patent/US11487921B2/en`.