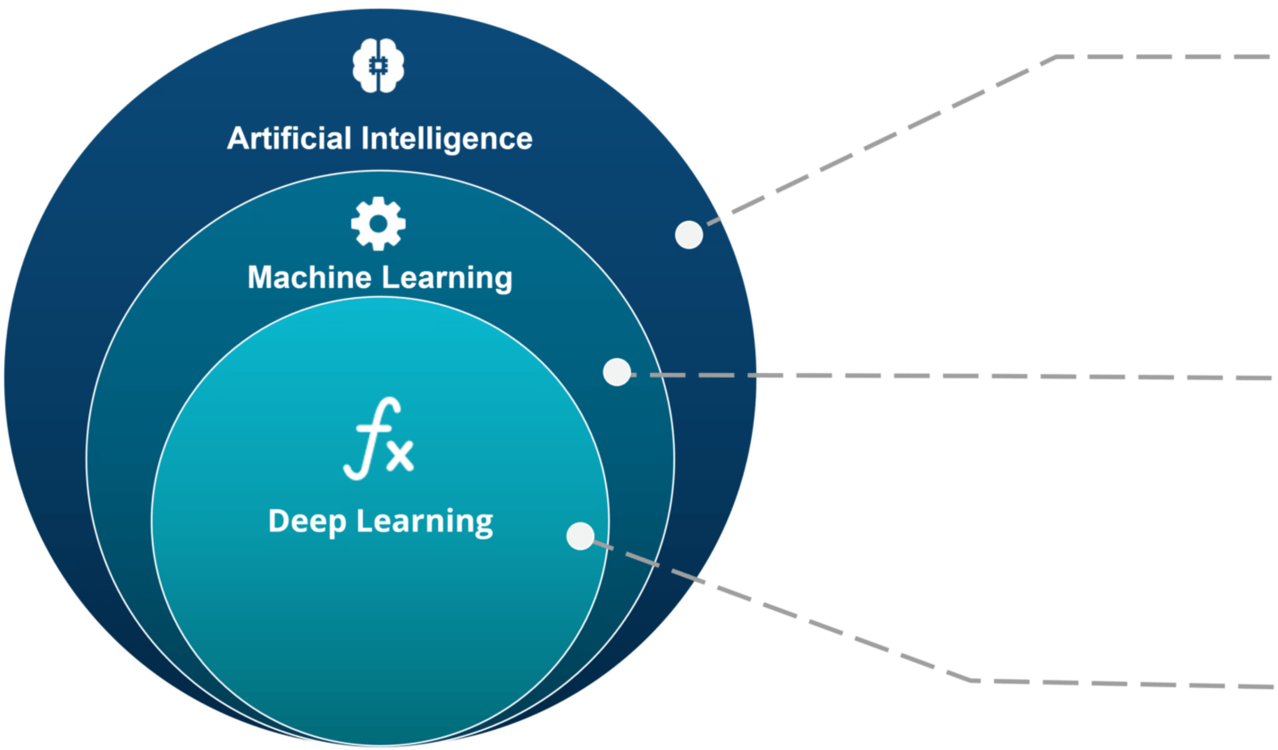
Introduction to quantum machine learning

Goan, Hsi-Sheng
管　希　聖
Department of Physics,
Center for Quantum Science and Engineering,
and Center for Theoretical Physics,
National Taiwan University, Taipei, Taiwan

Artificial intelligence

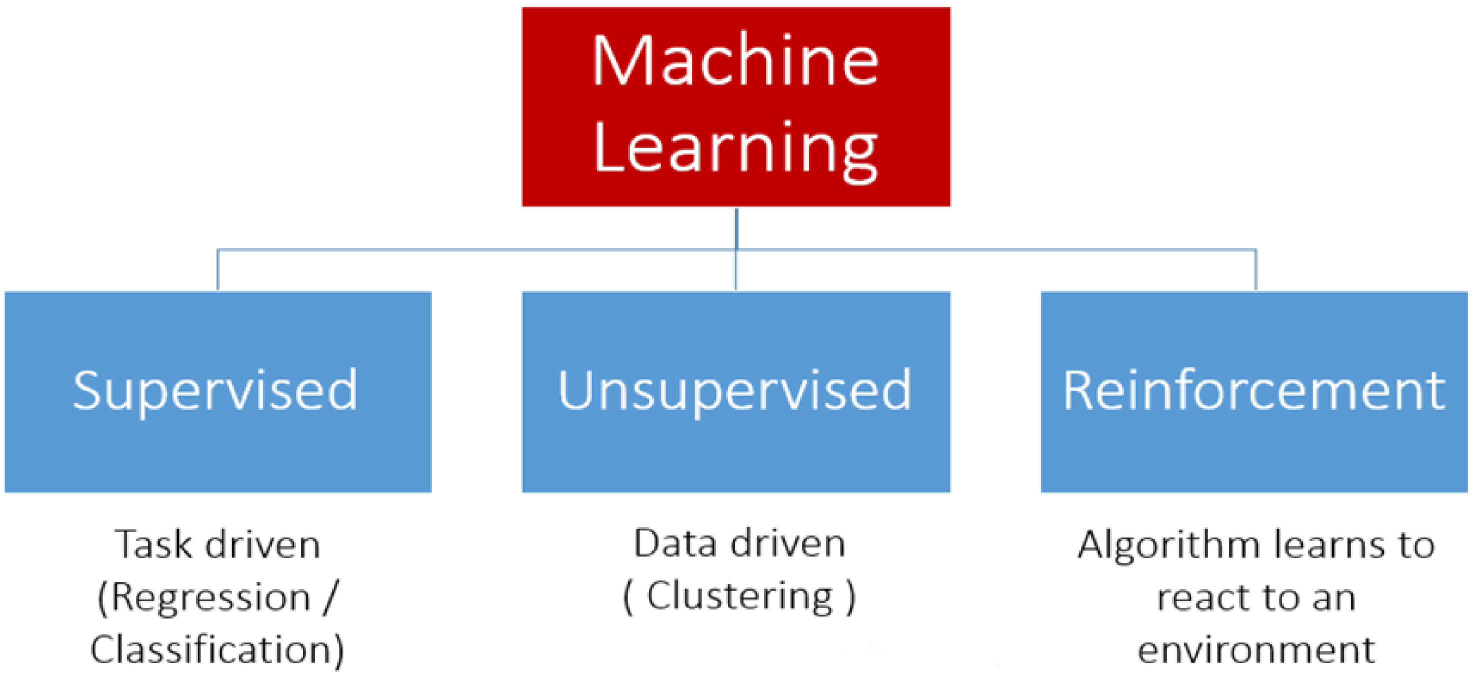


ARTIFICIAL INTELLIGENCE
A technique which enables machines to mimic human behaviour

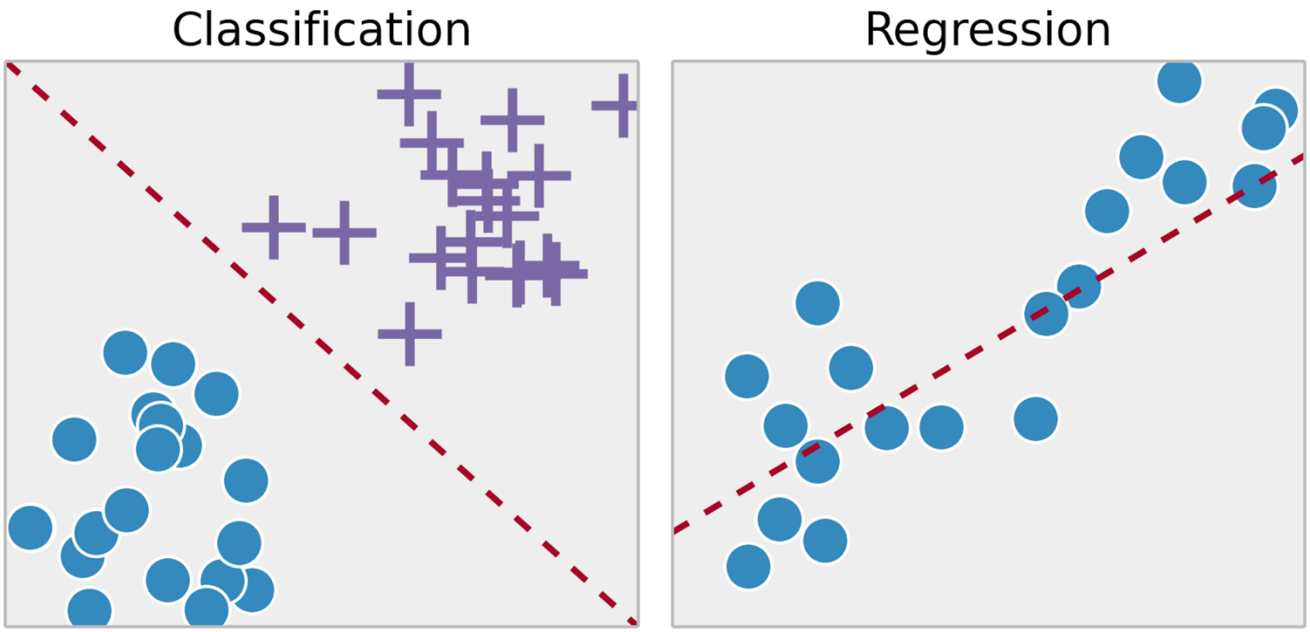
MACHINE LEARNING
Subset of AI technique which use statistical methods to enable machines to improve with experience

DEEP LEARNING
Subset of ML which make the computation of multi-layer neural network feasible

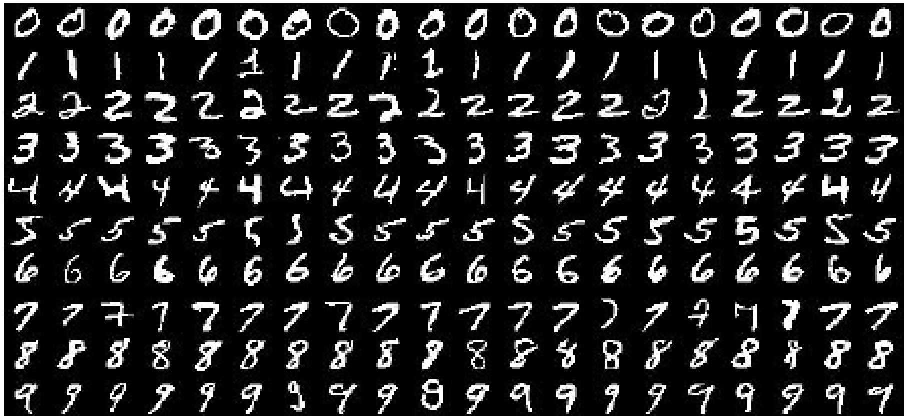
Types of machine learning



Supervised learning



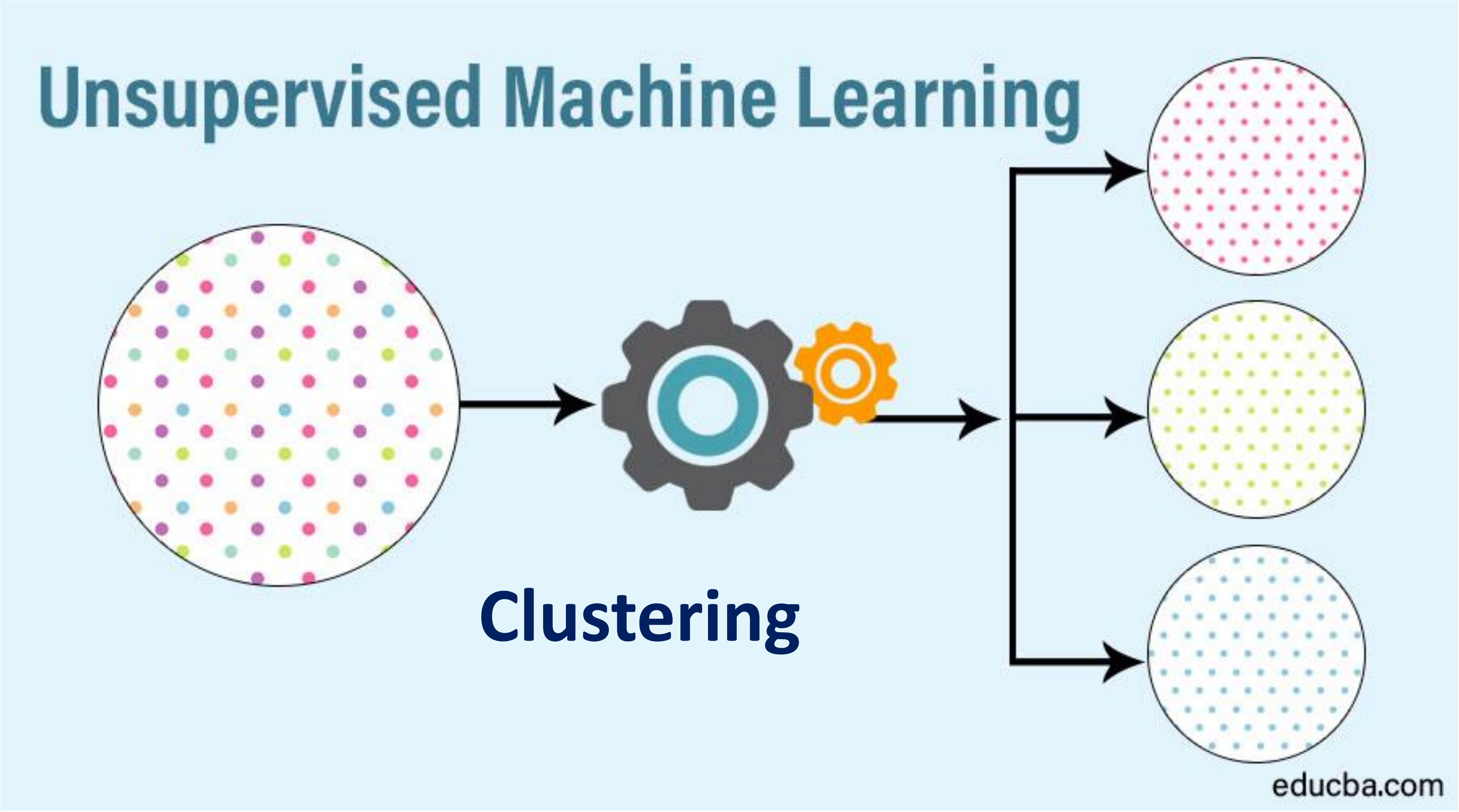
Supervised learning



MNIST

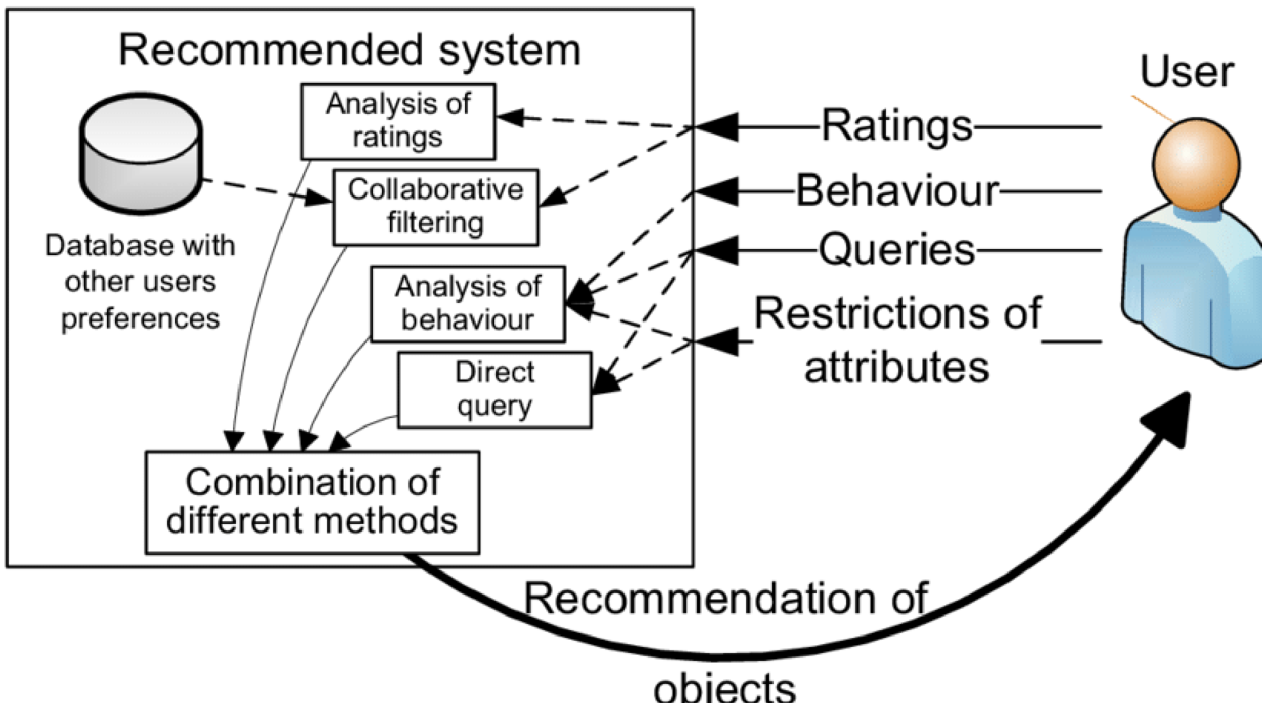
CIFAR-10

Unsupervised Learning

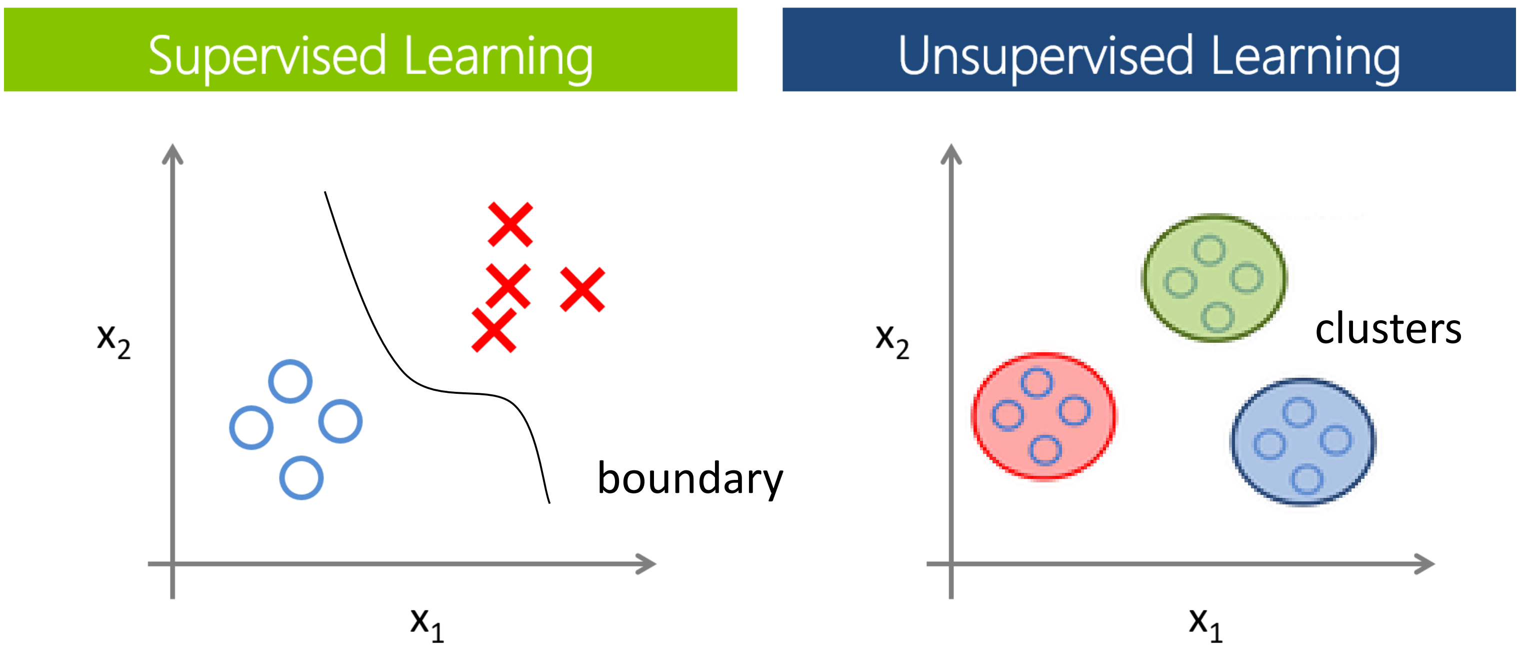


Unsupervised Learning

Recommendation System

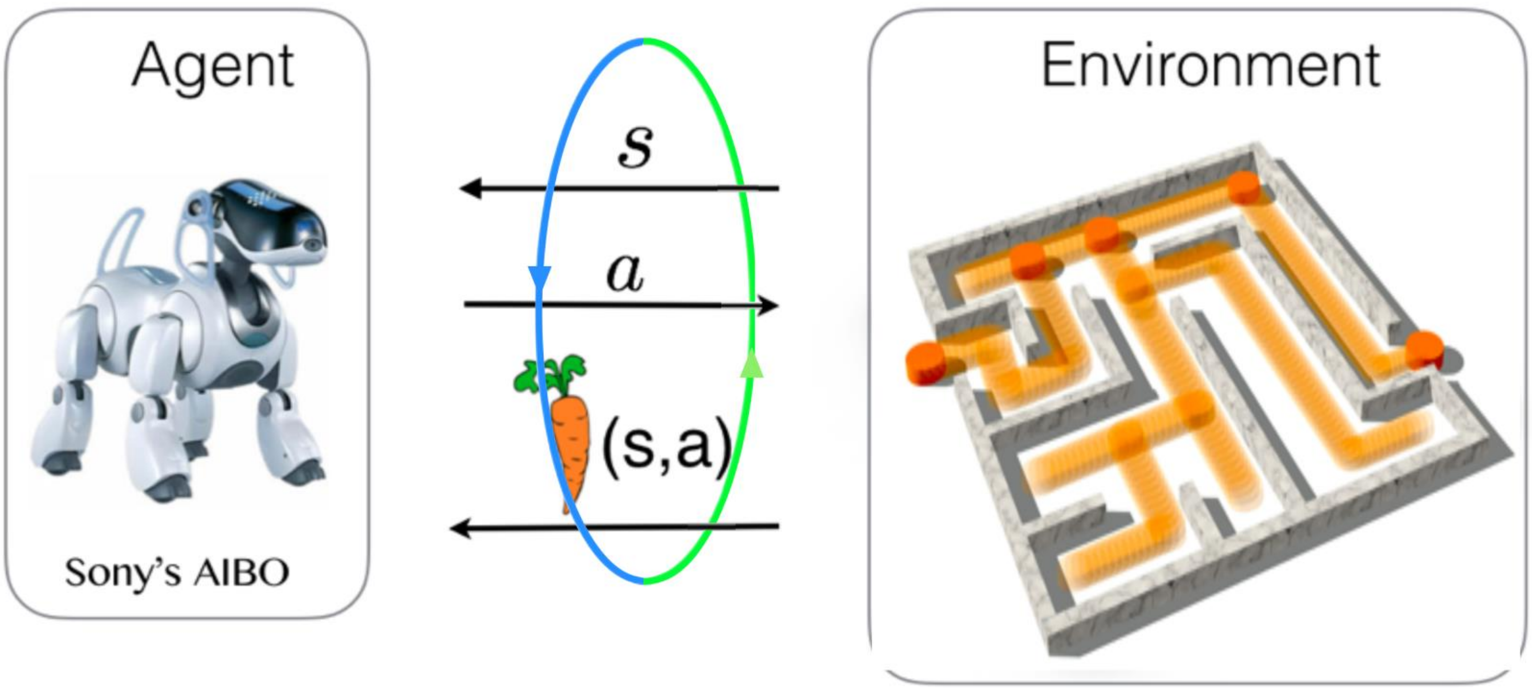


Supervised learning and unsupervised learning



Reinforcement learning

Agent-environment paradigm



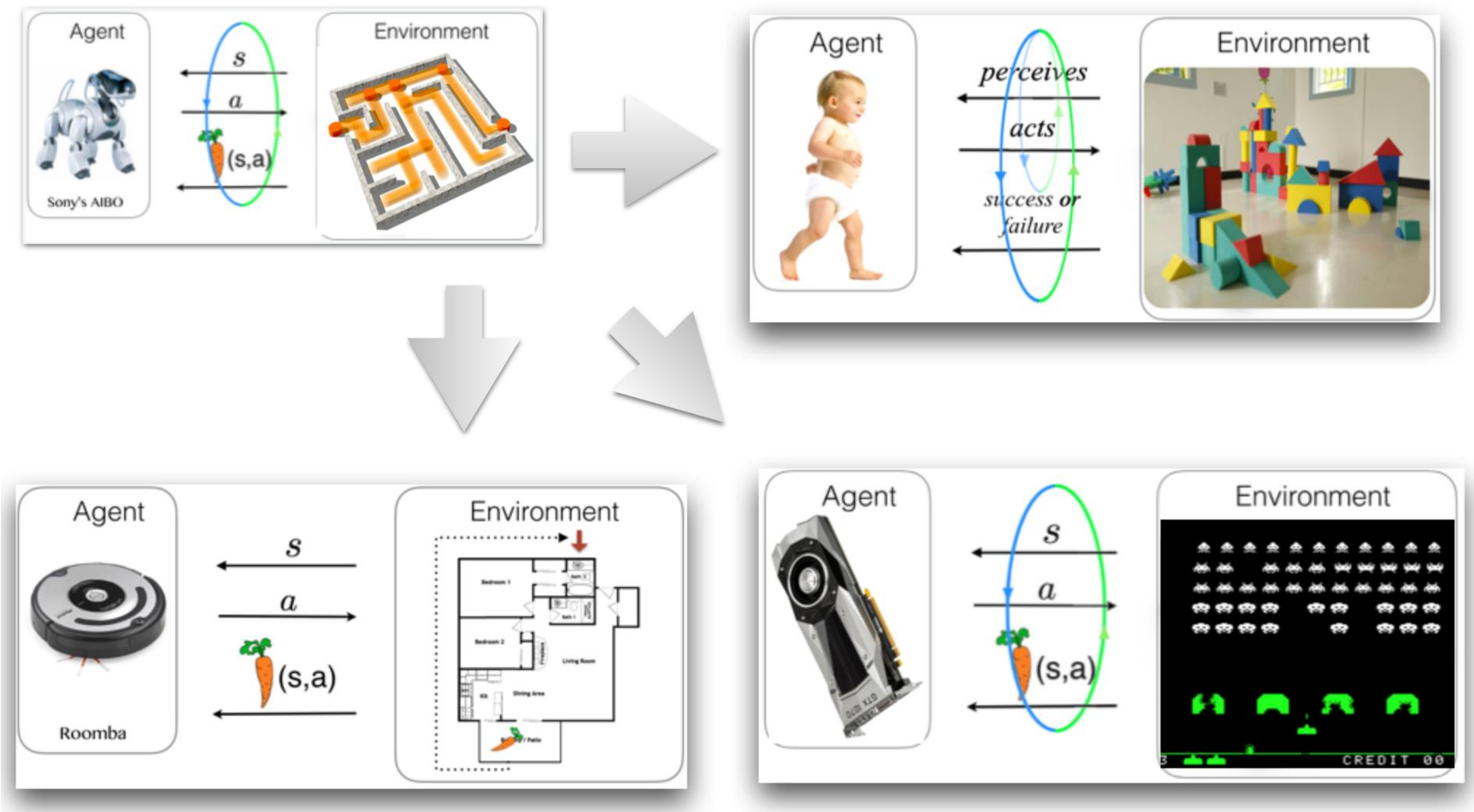
$$\mathcal{S} = \{s_1, s_2, \ldots\}$$
$$\mathcal{A} = \{a_1, a_2, \ldots\}$$

Closer to AI.
There is a body.
Interaction.
Learning.

figures taken from Wikipedia

Reinforcement learning

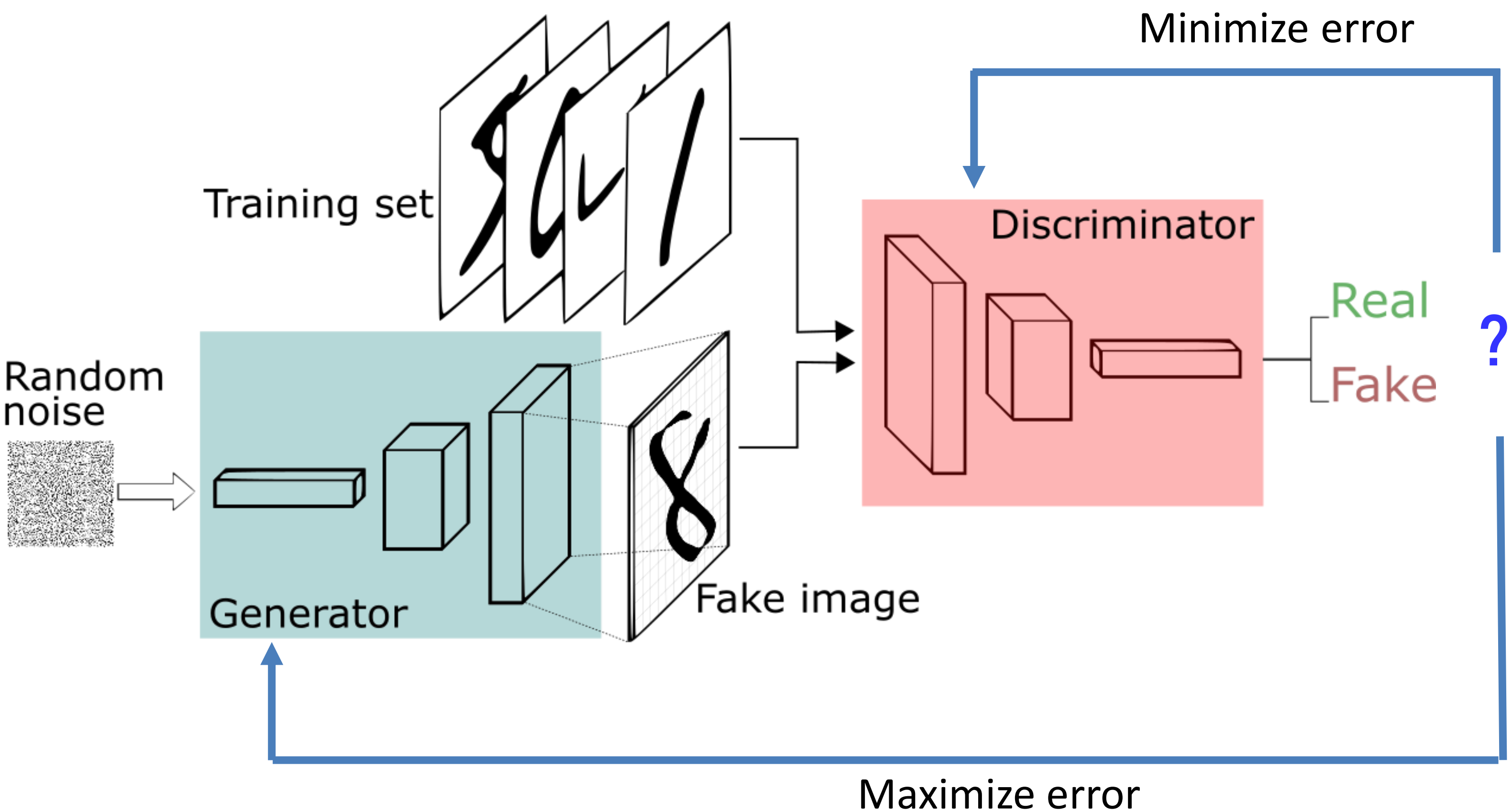
Agent-environment paradigm



figures taken from Wikipedia

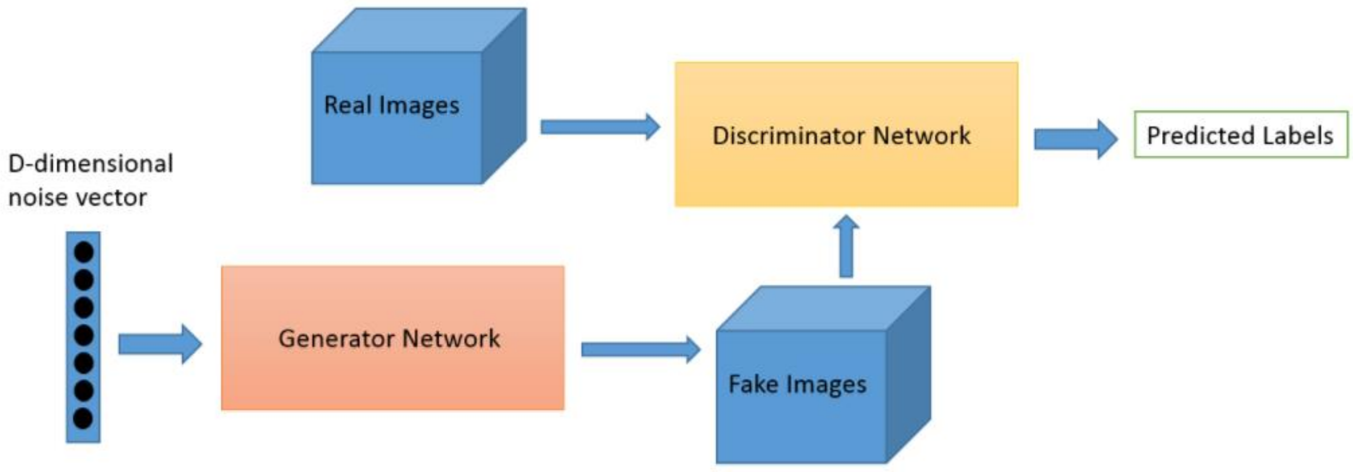
Unsupervised Learning

Generative Adversarial Networks (GAN)



Generative Adversarial Networks (GAN) 生成對抗網路

- 主要概念很簡單，好比一個遊戲有兩個角色，一個是偽造者(counterfeiter)，他不斷製造假鈔，另一個角色是警察，不斷從偽造者那邊拿到假鈔，判斷是真或假，然後，偽造者就根據警察判斷結果的回饋，不斷改良，最後假鈔變成真假難辨，這就是GAN的概念。
- 在GAN架構下，偽造者(counterfeiter)就稱為『生成模型』（generative model），警察稱為『判別模型』（discriminative model）生成模型**利用已知的真鈔加上雜訊(Noise)來製造假鈔**，交給判別模型辨識，它是二元分類模型，只會判斷真或偽，再將結果回饋給生成模型，經過不斷的訓練，就生成越越像的樣本了



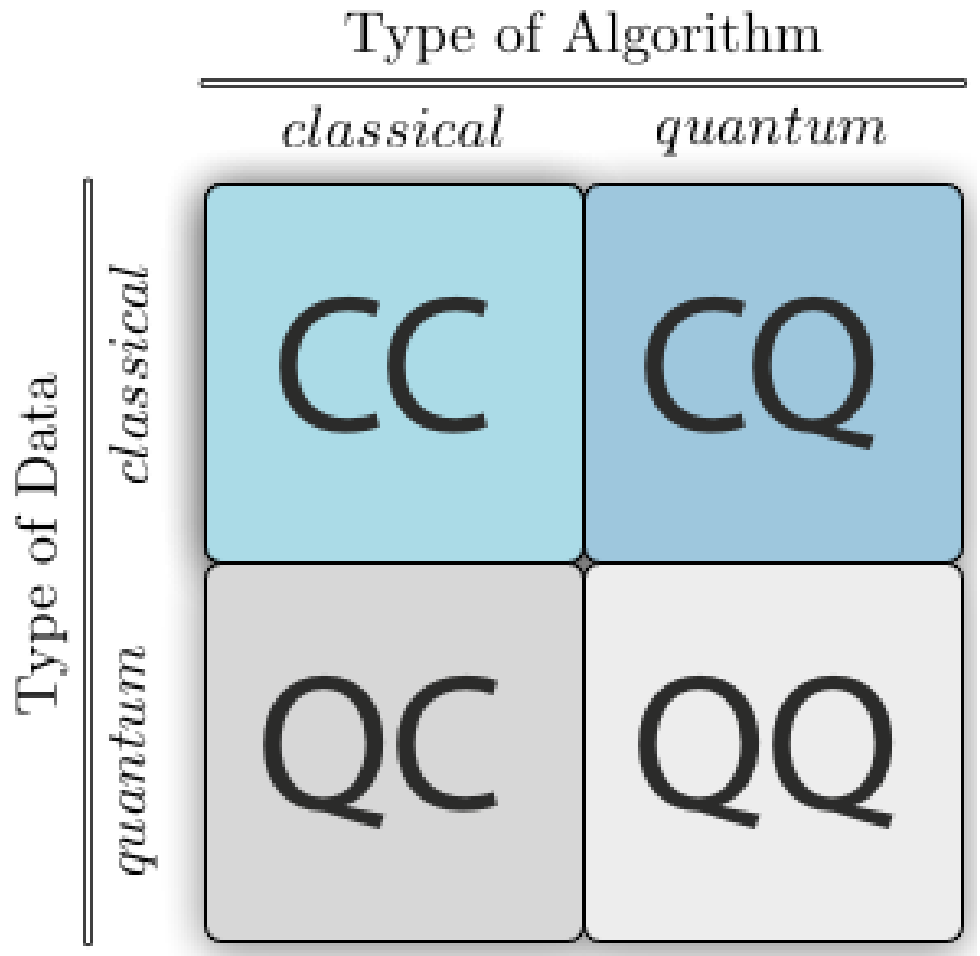
Unsupervised Learning
Generative Adversarial Networks (GAN)

https://thispersondoesnotexist.com

Advantages of GAN

- 『**生成對抗網路』(Generative Adversarial Network，GAN)經由小量真實資料，產生大量的訓練資料，儼然是一個『非監督式』(Unsupervised)的模型**，對照之前的CNN (Convolutional Neural Network)/RNN (Recurrent Neural Network) 都是『監督式』(Supervised)的模型，必須仰賴大量的標註資料，**所以GAN是 Neural Network 的一大進展。**
- **GAN** 透過金庸小說『老頑童周伯通雙手互搏』類似的理念，雙手互相切磋(AlphaGo 好像也是這樣)，**一方面改善模型的準確度，另一方面也可以產生高品質的訓練資料。**

Quantum algorithms and data



Quantum Support Vector Machine for Big Data Classification

Patrick Rebentrost,[1,*] Masoud Mohseni,[2] and Seth Lloyd[1,3,†]
[1]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Google Research, Venice, California 90291, USA*
[3]*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
(Received 12 February 2014; published 25 September 2014)

Supervised machine learning is the classification of new data based on already classified training examples. In this work, we show that the support vector machine, an optimized binary classifier, can be implemented on a quantum computer, with complexity logarithmic in the size of the vectors and the number of training examples. In cases where classical sampling algorithms require polynomial time, an exponential speedup is obtained. At the core of this quantum big data algorithm is a nonsparse matrix exponentiation technique for efficiently performing a matrix inversion of the training data inner-product (kernel) matrix.

- Quantum computers of the future will have the potential to give artificial intelligence (AI) a major boost. The team developed a quantum version of 'machine learning', a type of AI in which programs can learn from previous experience to become progressively better at finding patterns in data. The team's invention would take advantage of quantum computations to speed up machine-learning tasks exponentially.
- **QML takes the results of algebraic manipulations and puts them to good use. Data can be split into groups — a task that is at the core of handwriting- and speech-recognition software — or can be searched for patterns. Massive amounts of information could therefore be manipulated with a relatively small number of qubits.**
- **Such quantum AI techniques could dramatically speed up tasks such as image recognition for comparing photos on the web or for enabling cars to drive themselves** — fields in which companies such as Google have invested considerable resources.

PRL **114,** 140504 (2015) PHYSICAL REVIEW LETTERS week ending 10 APRIL 2015

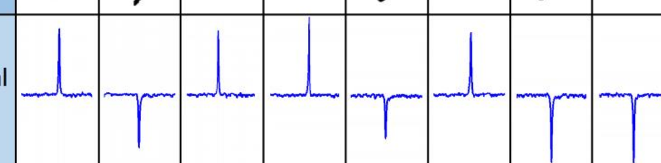
Experimental Realization of a Quantum Support Vector Machine

Zhaokai Li,[1,2] Xiaomei Liu,[1] Nanyang Xu,[1,2,*] and Jiangfeng Du[1,2,†]
[1]*Hefei National Laboratory for Physical Sciences at the Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China*
[2]*Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China*
(Received 1 December 2014; revised manuscript received 17 February 2015; published 8 April 2015)

The fundamental principle of artificial intelligence is the ability of machines to learn from previous experience and do future work accordingly. In the age of big data, classical learning machines often require huge computational resources in many practical cases. Quantum machine learning algorithms, on the other hand, could be exponentially faster than their classical counterparts by utilizing quantum parallelism. Here, we demonstrate a quantum machine learning algorithm to implement handwriting recognition on a four-qubit NMR test bench. The quantum machine learns standard character fonts and then recognizes handwritten characters from a set with two candidates. Because of the wide spread importance of artificial intelligence and its tremendous consumption of computational resources, quantum speedup would be extremely attractive against the challenges of big data.

Training data (printed characters)		label
6	$\vec{x}_1 = (0.987, 0.159)$	$y(\vec{x}_1) = +1$
9	$\vec{x}_2 = (0.354, 0.935)$	$y(\vec{x}_2) = -1$

Handwritten characters			
6	(0.997, -0.072)	9	(0.338, 0.941)
9	(0.147, 0.989)	6	(0.999, 0.025)
6	(0.999, -0.030)	9	(0.439, 0.899)
6	(0.987, -0.161)	9	(0.173, 0.985)

Hand-written characters	6	9	6	6	9	6	9	9
Experimental indicators								
Recognition results	6	9	6	6	9	6	9	9

The feature values are chosen as the vertical ratio and the horizontal ratio, calculated from the pixels in the left (upper) half over the right (lower) half.

Quantum-Enhanced Machine Learning

Vedran Dunjko,[1,*] Jacob M. Taylor,[2,3,†] and Hans J. Briegel[1,‡]
[1]*Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 21a, A-6020 Innsbruck, Austria*
[2]*Joint Quantum Institute, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*
[3]*Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, Maryland 20742, USA*
(Received 15 April 2016; published 20 September 2016)

The emerging field of quantum machine learning has the potential to substantially aid in the problems and scope of artificial intelligence. This is only enhanced by recent successes in the field of classical machine learning. In this work we propose an approach for the systematic treatment of machine learning, from the perspective of quantum information. Our approach is general and covers all three main branches of machine learning: supervised, unsupervised, and reinforcement learning. While quantum improvements in supervised and unsupervised learning have been reported, reinforcement learning has received much less attention. Within our approach, we tackle the problem of quantum enhancements in reinforcement learning as well, and propose a systematic scheme for providing improvements. As an example, we show that quadratic improvements in learning efficiency, and exponential improvements in performance over limited time periods, can be obtained for a broad class of learning problems.

- We are witnessing the emergence of a new field: quantum machine learning (QML), which has a further, profound potential to revolutionize the field of artificial intelligence (AI), much like quantum information processing has influenced its classical counterpart.

REVIEW

14 SEPTEMBER 2017 | VOL 549 | NATURE | 195

doi:10.1038/nature23474

Quantum machine learning

Jacob Biamonte[1,2], Peter Wittek[3], Nicola Pancotti[4], Patrick Rebentrost[5], Nathan Wiebe[6] & Seth Lloyd[7]

Fuelled by increasing computer power and algorithmic advances, machine learning techniques have become powerful tools for finding patterns in data. Quantum systems produce atypical patterns that classical systems are thought not to produce efficiently, so it is reasonable to postulate that quantum computers may outperform classical computers on machine learning tasks. The field of quantum machine learning explores how to devise and implement quantum software that could enable machine learning that is faster than that of classical computers. Recent work has produced quantum algorithms that could act as the building blocks of machine learning programs, but the hardware and software challenges are still considerable.

(1) The input problem. Although quantum algorithms can provide dramatic speedups for processing data, they seldom provide advantages in reading data. This means that the cost of reading in the input can in some cases dominate the cost of quantum algorithms. Understanding this factor is an ongoing challenge.

(2) The output problem. Obtaining the full solution from some quantum algorithms as a string of bits requires learning an exponential number of bits. This makes some applications of quantum machine learning algorithms infeasible. This problem can potentially be sidestepped by learning only summary statistics for the solution state.

(3) The costing problem. Closely related to the input/output problems, at present very little is known about the true number of gates required by quantum machine learning algorithms. Bounds on the complexity suggest that for sufficiently large problems they will offer huge advantages, but it is still unclear when that crossover point occurs.

(4) The benchmarking problem. It is often difficult to assert that a quantum algorithm is ever better than all known classical machine algorithms in practice because this would require extensive benchmarking against modern heuristic methods. Establishing lower bounds for quantum machine learning would partially address this issue.

Machine learning & artificial intelligence in the quantum domain: a review of recent progress

Rep. Prog. Phys. **81** (2018) 074001 (67pp)

Vedran Dunjko[1,2] and **Hans J Briegel**[1,3]

Quantum information technologies, on the one hand, and intelligent learning systems, on the other, are both emergent technologies that are likely to have a transformative impact on our society in the future. The respective underlying fields of basic research—quantum information versus machine learning (ML) and artificial intelligence (AI)—have their own specific questions and challenges, which have hitherto been investigated largely independently. However, in a growing body of recent work, researchers have been probing the question of the extent to which these fields can indeed learn and benefit from each other. Quantum ML explores the interaction between quantum computing and ML, investigating how results and techniques from one field can be used to solve the problems of the other. Recently we have witnessed significant breakthroughs in both directions of influence. For instance, quantum computing is finding a vital application in providing speed-ups for ML problems, critical in our 'big data' world. Conversely, ML already permeates many cutting-edge technologies and may become instrumental in advanced quantum technologies. Aside from quantum speed-up in data analysis, or classical ML optimization used in quantum experiments, quantum enhancements have also been (theoretically) demonstrated for interactive learning tasks, highlighting the potential of quantum-enhanced learning agents. Finally, works exploring the use of AI for the very design of quantum experiments and for performing parts of genuine research autonomously, have reported their first successes. Beyond the topics of mutual enhancement—exploring what ML/AI can do for quantum physics and vice versa—researchers have also broached the fundamental issue of quantum generalizations of learning and AI concepts. This deals with questions of the very meaning of learning and intelligence in a world that is fully described by quantum mechanics. In this review, we describe the main ideas, recent developments and progress in a broad spectrum of research investigating ML and AI in the quantum domain.

PHYSICAL REVIEW LETTERS **120,** 050502 (2018)

Quantum Linear System Algorithm for Dense Matrices

Leonard Wossnig,[1,2] Zhikuan Zhao,[3,4,*] and Anupam Prakash[4]
[1]*Theoretische Physik, ETH Zürich, 8093 Zurich, Switzerland*
[2]*Department of Materials, University of Oxford, Oxford OX1 3PH, United Kingdom*
[3]*Singapore University of Technology and Design, 487372 Singapore*
[4]*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*

(Received 25 May 2017; revised manuscript received 10 November 2017; published 31 January 2018)

Solving linear systems of equations is a frequently encountered problem in machine learning and optimization. Given a matrix A and a vector $\mathbf{b}$ the task is to find the vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$. We describe a quantum algorithm that achieves a sparsity-independent runtime scaling of $\mathcal{O}(\kappa^2 \sqrt{n}\text{polylog}(n)/\epsilon)$ for an $n \times n$ dimensional A with bounded spectral norm, where κ denotes the condition number of A, and ϵ is the desired precision parameter. This amounts to a polynomial improvement over known quantum linear system algorithms when applied to dense matrices, and poses a new state of the art for solving dense linear systems on a quantum computer. Furthermore, an exponential improvement is achievable if the rank of A is polylogarithmic in the matrix dimension. Our algorithm is built upon a singular value estimation subroutine, which makes use of a memory architecture that allows for efficient preparation of quantum states that correspond to the rows of A and the vector of Euclidean norms of the rows of A.

DOI: 10.1103/PhysRevLett.120.050502

- For a sparse and well-conditioned matrix *A* : **A.W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. 103, 150502 (2009).**
- For potentially circumventing the problem of exponential overhead in the initial step of transferring classical data to quantum states : **X. Gao, Z.-Y. Zhang, L.-M. Duan, Sci. Adv. 4, eaat9004 (2018).**

Open questions in quantum-enhanced machine learning

- For quantum-enhanced machine learning, a unified quantum learning theory has not been developed.
- What is the general criterion for determining if a machine-learning task can be significantly expedited by a quantum computer?
- What learning problems can be efficiently solved by a quantum computer but not by a classical one?
- And how can a quantum computer efficiently analyze large quantum data sets that may eventually be available?
- Moreover, a smoking-gun experimental demonstration of quantum speedups in a practical machine-learning task would be an important milestone.

- S. Das Sarma, D.-L. Deng, and L.-M. Duan, PHYSICS TODAY, MARCH 2019, pp. 48-54

Quantum Computing in the NISQ era and beyond

John Preskill

Quantum **2**, 79 (2018)
arXiv: 1801.00862

Institute for Quantum Information and Matter and Walter Burke Institute for Theoretical Physics,
California Institute of Technology, Pasadena CA 91125, USA

30 July 2018

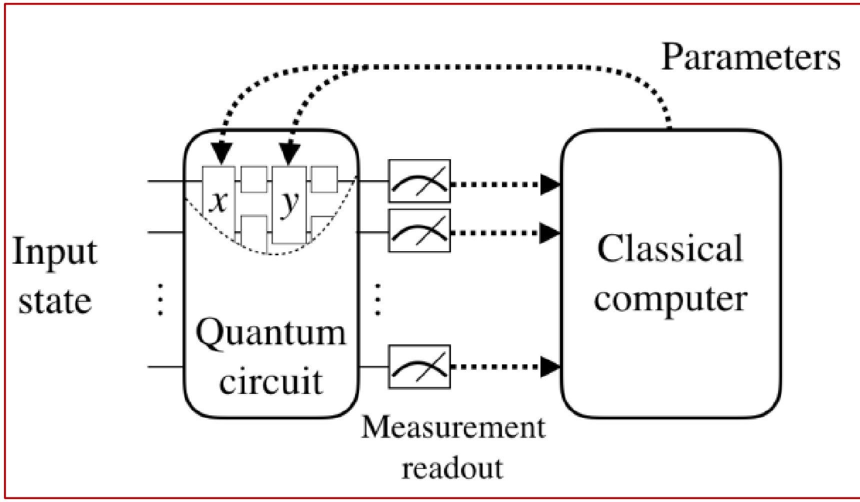
Noisy Intermediate-Scale Quantum (NISQ) technology will be available in the near future. **Quantum computers with 50-100 qubits may be able to perform tasks which surpass the capabilities of today's classical digital computers**, **but noise in quantum gates will limit the size of quantum circuits that can be executed reliably**. NISQ devices will be useful tools for exploring many-body quantum physics, and may have other useful applications, but the 100-qubit quantum computer will not change the world right away — we should regard it as a significant step toward the more powerful quantum technologies of the future. Quantum technologists should continue to strive for more accurate quantum gates and, eventually, fully fault-tolerant quantum computing.

Hybrid Quantum-Classical Algorithm

- **NISQ (Noisy Intermediate Scale Quantum)** Devices:
  - **Imperfect gates, noise measurements and decoherence of the qubits**
    (No error correction)
- **NISQ algorithm solution**: **small number of qubits and low circuit depth**
- **Hybrid Quantum-Classical (HQC) Algorithm**: **leverage strengths of quantum and classical computation**
- **Variational quantum circuit algorithm is the most popular Hybrid Quantum-Classical algorithm**.

Hybrid quantum-classical algorithm

- A lot of optimization methods developed in (classical) machine learning community
- Optimization circuit parameter in classical computer
- Data **encoding** scheme can be tricky, and determine the possible quantum advantages.
  - Amplitude encoding
  - Variational encoding
  - Computational basis encoding



PERSPECTIVE

Quantum Sci. Technol. **3** (2018) 030502

PUBLISHED
19 June 2018

Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers

Alejandro Perdomo-Ortiz[1,2,3,4,5], Marcello Benedetti[1,2,4,5], John Realpe-Gómez[1,6,7] and Rupak Biswas[1,8]

With quantum computing technologies nearing the era of commercialization and quantum supremacy, machine learning (ML) appears as one of the promising 'killer' applications. Despite significant effort, there has been a disconnect between most quantum ML proposals, **the needs of ML practitioners, and the capabilities of near-term quantum devices to demonstrate quantum enhancement in the near future.** In this contribution to the focus collection 'What would you do with 1000 qubits?', we provide concrete examples of intractable ML tasks that could be enhanced with near term devices. We argue that to reach this target, the focus should be on areas where ML researchers are struggling, such as **generative models in unsupervised and semi-supervised learning**, instead of the popular and more tractable supervised learning techniques. We also highlight the case of classical datasets with potential quantum-like statistical correlations where quantum models could be more suitable. We focus on hybrid quantum–classical approaches and illustrate some of the key challenges we foresee for near-term implementations. Finally, we introduce the quantum-assisted Helmholtz machine (QAHM), an attempt to use near-term quantum devices to tackle high-dimensional datasets of continuous variables. Instead of using quantum computers to assist deep learning, as previous approaches do, the QAHM uses deep learning to extract a low-dimensional binary representation of data, suitable for relatively small quantum processors which can assist the training of an unsupervised generative model. Although we illustrate this concept on a quantum annealer, other quantum platforms could benefit as well from this hybrid quantum–classical framework.

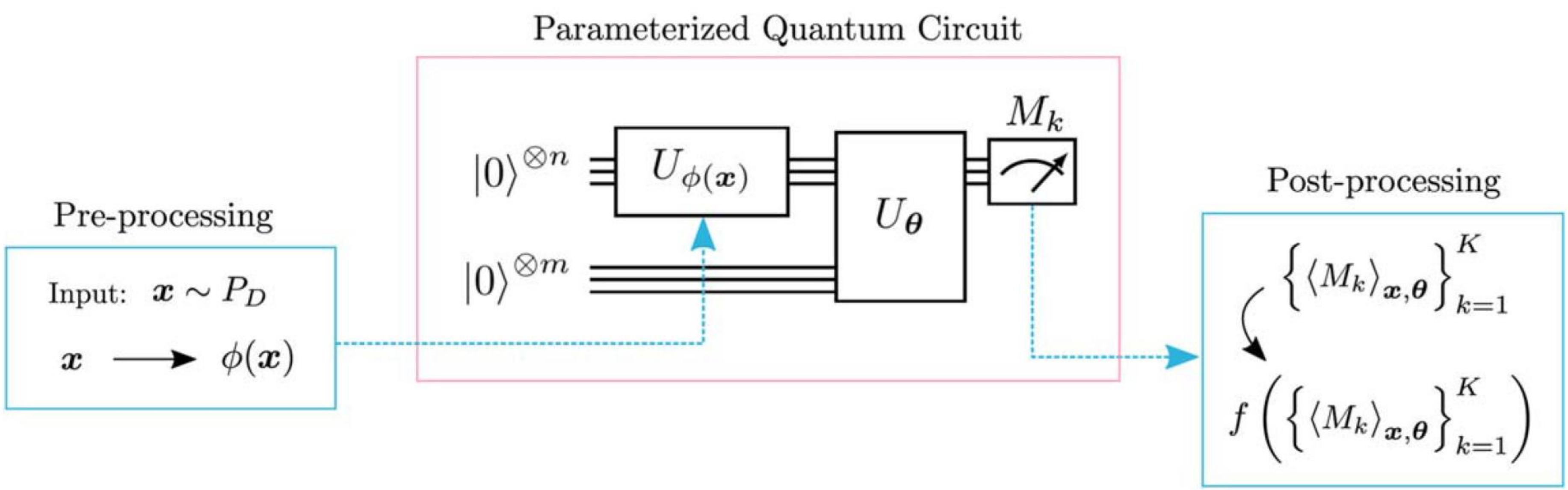
TOPICAL REVIEW

Quantum Sci. Technol. **4** (2019) 043001

Parameterized quantum circuits as machine learning models

Marcello Benedetti[1,2], Erika Lloyd[1], Stefan Sack[1] and Mattia Fiorentini[1]

Hybrid quantum–classical systems make it possible to utilize existing quantum computers to their fullest extent. Within this framework, **parameterized quantum circuits can be regarded as machine learning models with remarkable expressive power.** This Review presents the components of these models and discusses their application to a variety of data-driven tasks, such as supervised learning and generative modeling. With an increasing number of experimental demonstrations carried out on actual quantum hardware and with software being actively developed, this rapidly growing field is poised to have a broad spectrum of real-world applications.

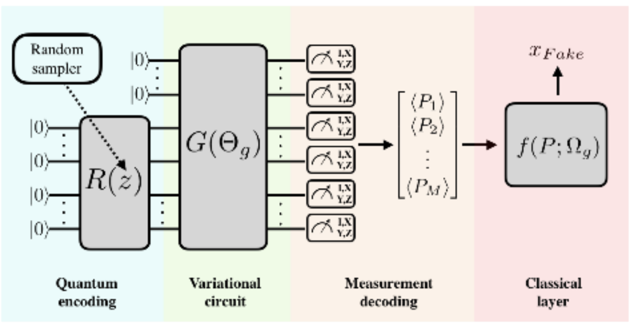


Variational Quantum Circuits

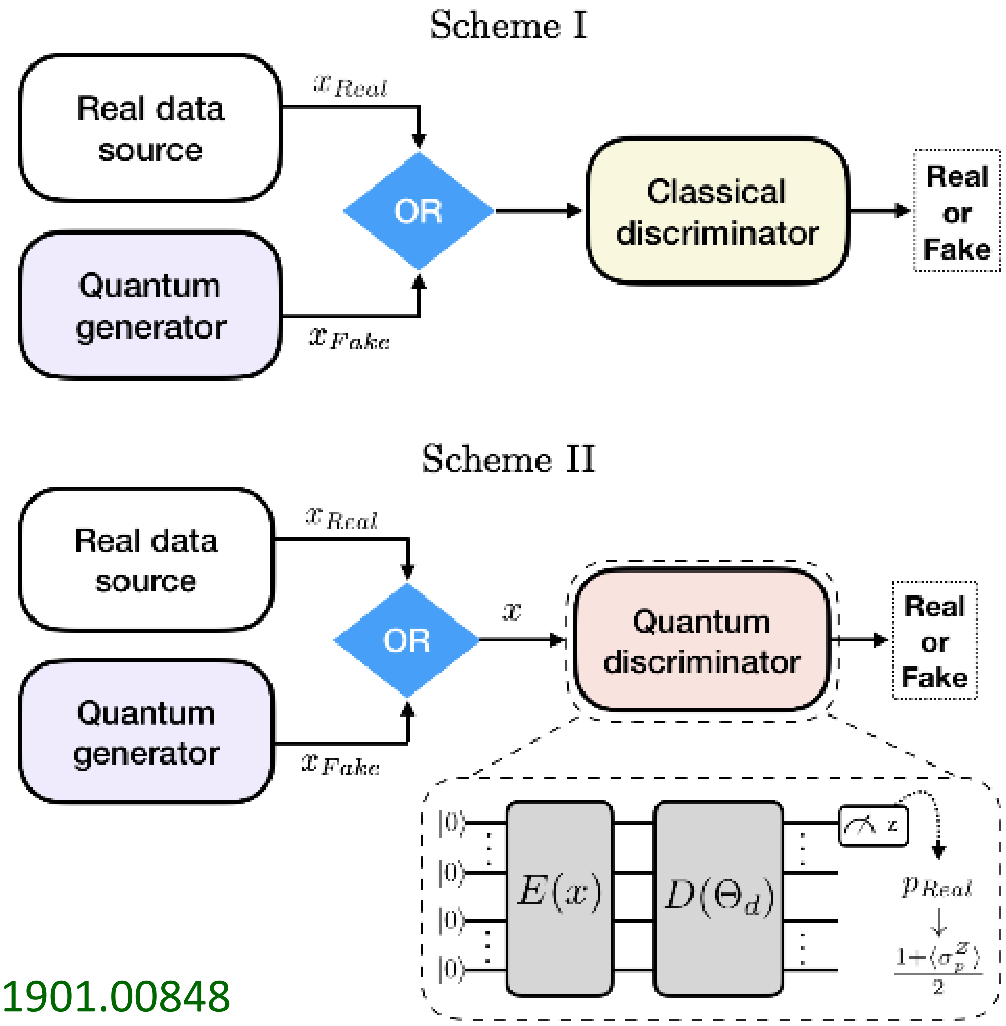
Have already succeeded in:

- Generative Adversarial Networks (GAN)
- Classification
- Function Approximators
- …

Variational quantum circuits - GAN

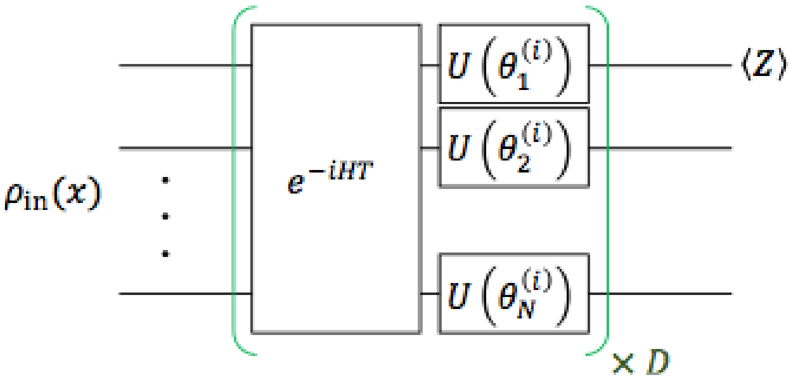


Variational quantum generator

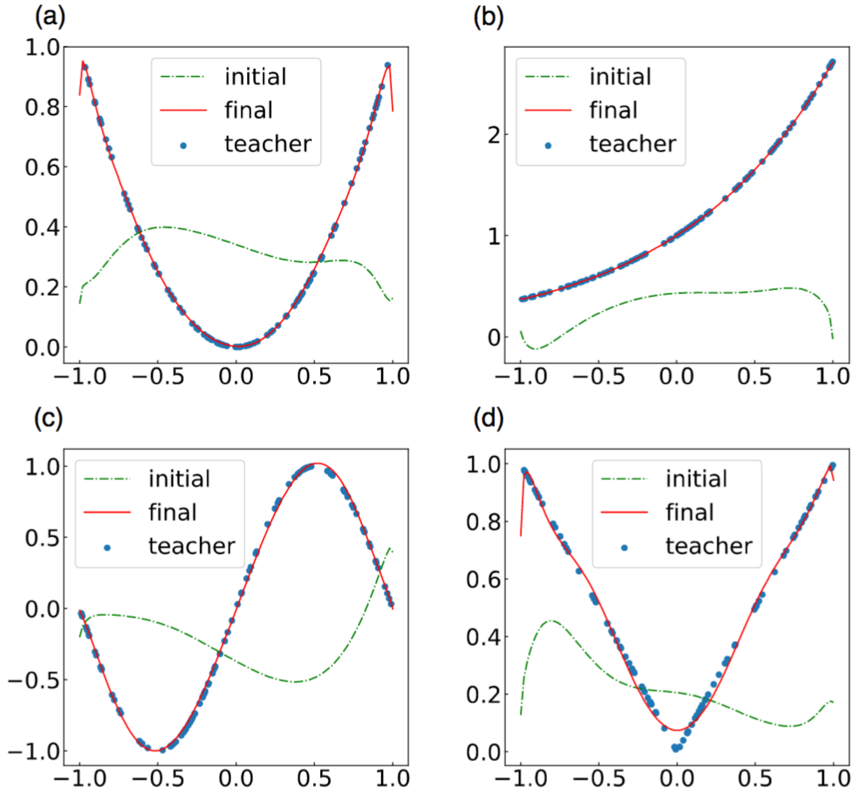


J. Romero, A. Aspuru-Guzik, arXiv:1901.00848

Variational quantum circuits - Function approximator



$$H = \sum_{j=1}^{N} a_j X_j + \sum_{j=1}^{N} \sum_{k=1}^{j-1} J_{jk} Z_j Z_k.$$



K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Phys. Rev. A **98**, 032309 (2018)

Variational quantum circuits - Classification

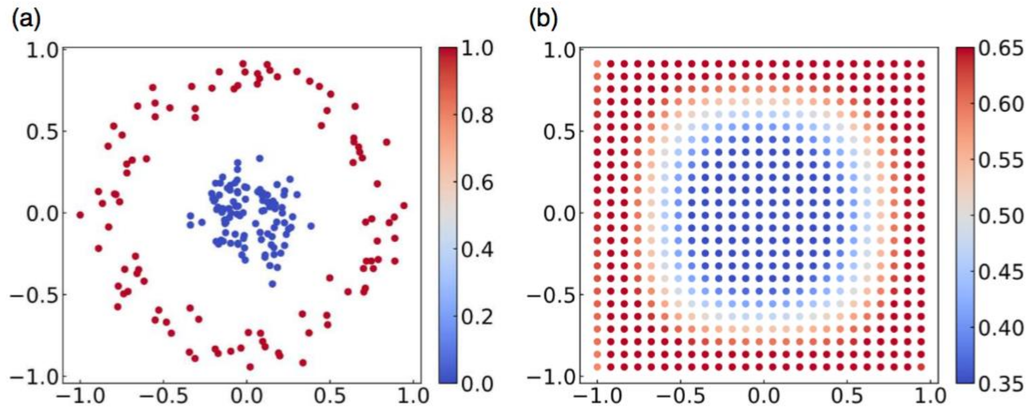
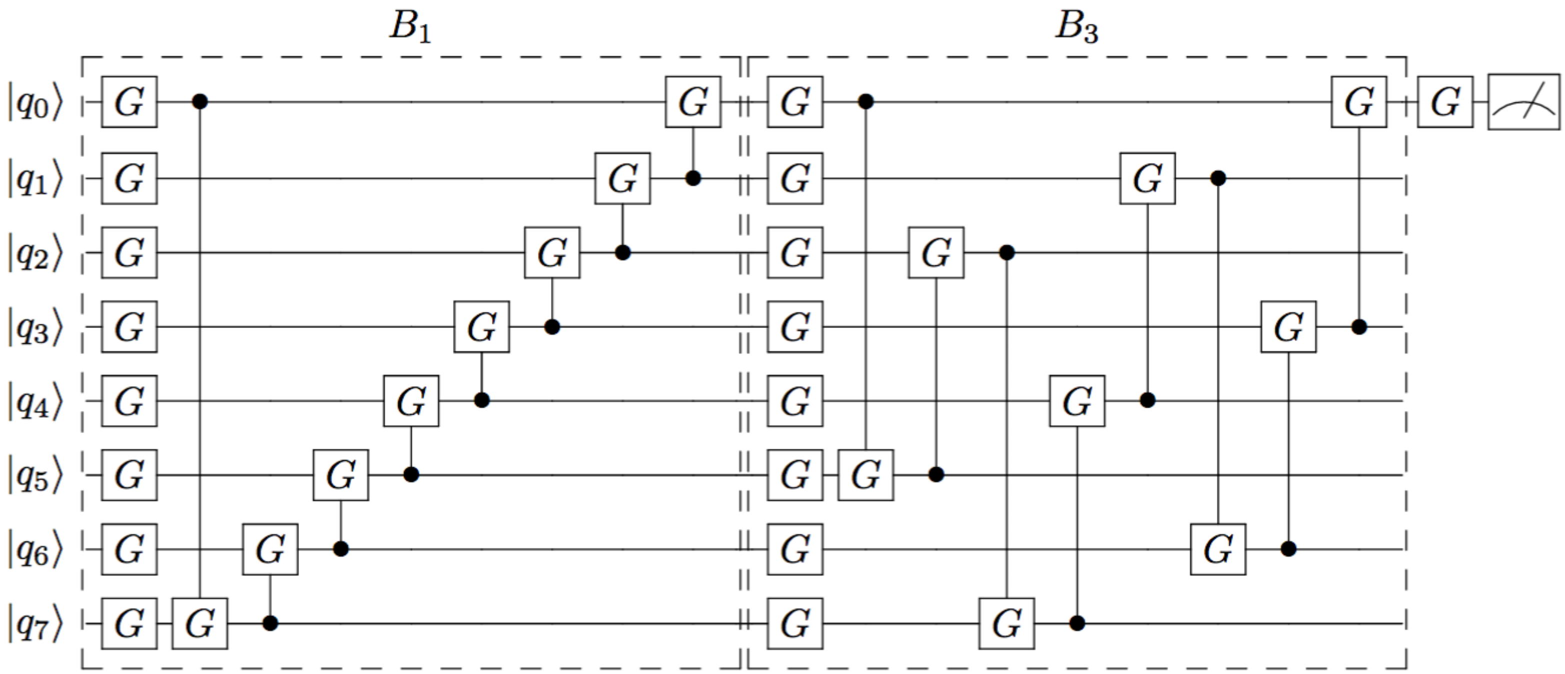


FIG. 4. Demonstration of a simple nonlinear classification task. (a) teacher data. Data points that belong to class 0, 1 is shown as blue and red dot, respectively. (b) Optimized output from first qubit (after softmax transformation). 0.5 is the threshold for classification, less than and greater than 0.5 means that the point is classified as class 0 and 1, respectively.

K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Phys. Rev. A **98**, 032309 (2018)

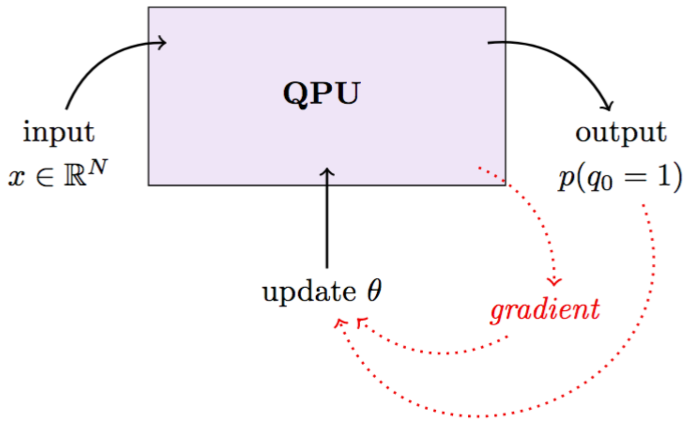
Variational quantum circuits – circuit-centric quantum classifiers



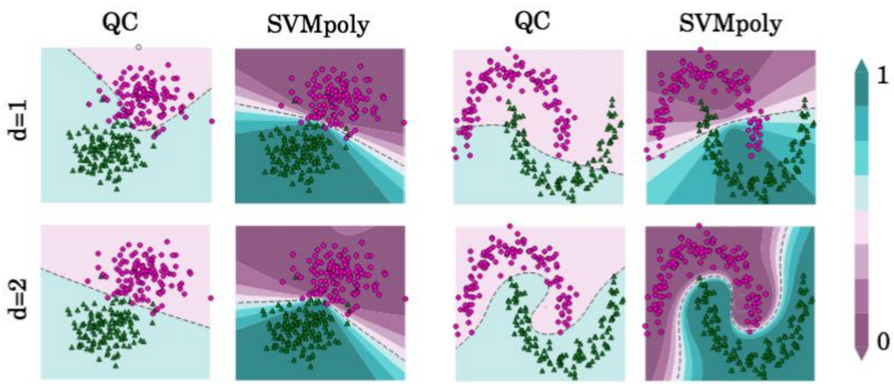
Generic model circuit architecture for 8 qubits.

M. Schuld, A. Bocharov, K. Svore, N. Wiebe, arXiv:1804.00633

Variational quantum circuits – circuit-centric quantum classifiers



Idea of the hybrid training method. The quantum processing unit (QPU) is used to compute outputs and gradients of the model in order to update the parameters for each step of the gradient descent training.

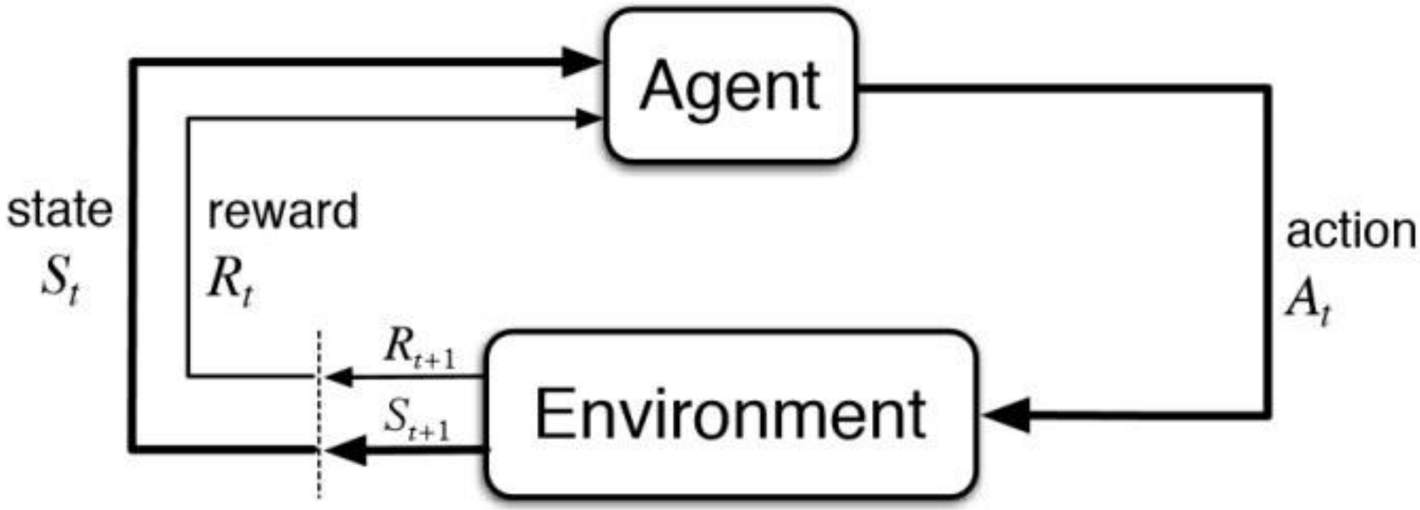


Comparison of the decision boundary for the circuit-centric quantum classifier (QC) and a support vector machine with polynomial kernel (SVMpoly).

M. Schuld, A. Bocharov, K. Svore, N. Wiebe, arXiv:1804.00633

Can we use quantum circuits to do **reinforcement learning** on a **noisy intermediate scale quantum (NISQ) machine**?

Reinforcement learning



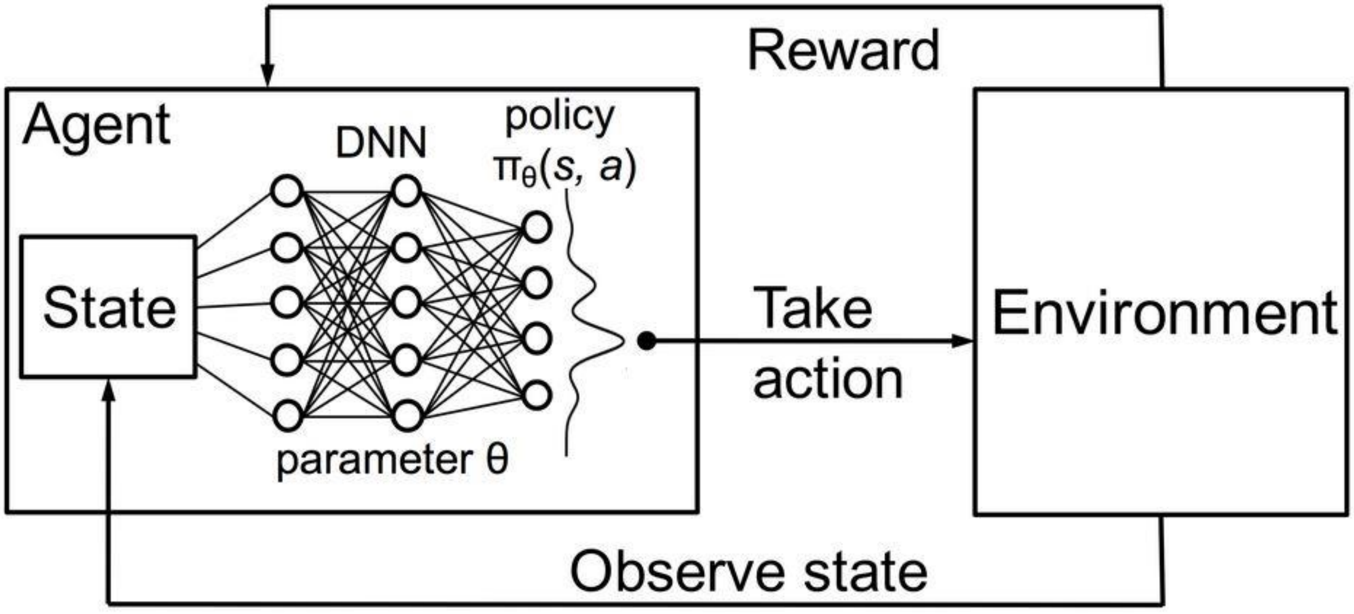
An agent interacts with an environment *E* over a number of discrete time steps. At each time step *t*, the agent receives a state s_t , observes a reward r_t , chooses an action a_t from a set of possible actions $\mathcal{A}$ according to its policy π, and enters a new state s_{t+1} (that may depend on both the previous state and the selected action), and then the action-value function or Q-value function *Q* (expected return) is updated with the learning rate α.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right]$$

γ is the discount factor to control how future rewards are given to the decision making function.

1.Sutton&Barto Reinforcement Learning: An Introduction
2.https://storage.googleapis.com/deepmind-media/dqn/DQNNaturePaper.pdf

Deep reinforcement learning



Q-learning: to directly approximate the optimal action-value function $Q^*(s,a)$ by minimizing the mean square error (MSE) loss function:

$$L(\theta) = \mathbb{E}[(r_t + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s_t, a_t; \theta))^2]$$

target prediction

Using DNN to approximate the value of the loss function
--> suffers from unstable or even divergence

Deep Q-learning (DQL) or deep Q-network (DQN)

Causes of instability:

- the correlations present in the sequence of observations,
- the fact that small updates to Q may significantly change the policy and therefore change the data distribution,
- and the correlations between the action-values (Q) and the target values $r + \gamma \max_{a'} Q(s', a')$, causing $Q(s,a)$ to chase a nonstationary target.

LETTER

doi:10.1038/nature14236

Human-level control through deep reinforcement learning

Volodymyr Mnih[1]*, Koray Kavukcuoglu[1]*, David Silver[1]*, Andrei A. Rusu[1], Joel Veness[1], Marc G. Bellemare[1], Alex Graves[1], Martin Riedmiller[1], Andreas K. Fidjeland[1], Georg Ostrovski[1], Stig Petersen[1], Charles Beattie[1], Amir Sadik[1], Ioannis Antonoglou[1], Helen King[1], Dharshan Kumaran[1], Daan Wierstra[1], Shane Legg[1] & Demis Hassabis[1]

Experience replay and Target network

==> More stable deep RL

Deep Q-learning (DQL) or deep Q-network (DQN)

LETTER

doi:10.1038/nature14236

Human-level control through deep reinforcement learning

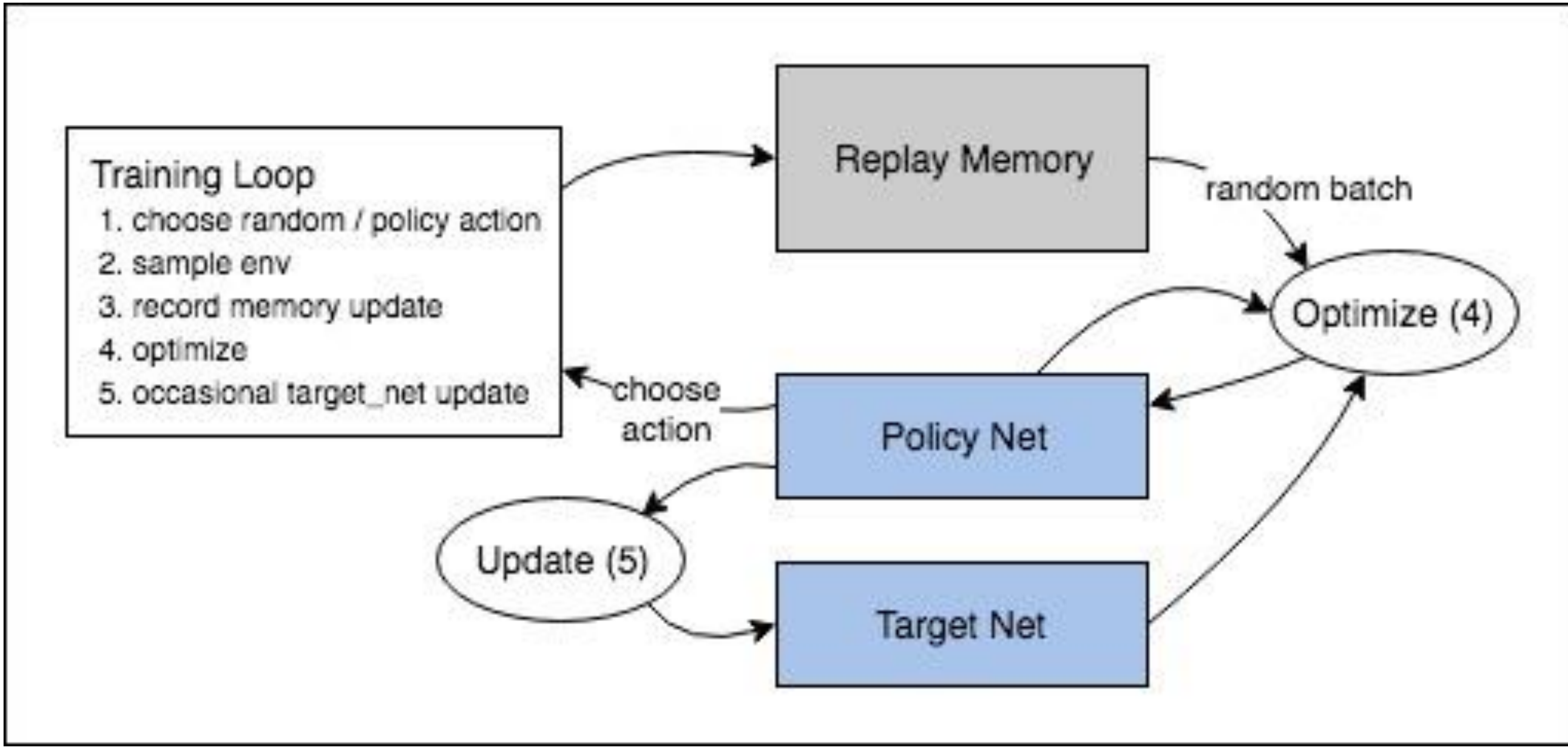
Volodymyr Mnih[1]*, Koray Kavukcuoglu[1]*, David Silver[1]*, Andrei A. Rusu[1], Joel Veness[1], Marc G. Bellemare[1], Alex Graves[1], Martin Riedmiller[1], Andreas K. Fidjeland[1], Georg Ostrovski[1], Stig Petersen[1], Charles Beattie[1], Amir Sadik[1], Ioannis Antonoglou[1], Helen King[1], Dharshan Kumaran[1], Daan Wierstra[1], Shane Legg[1] & Demis Hassabis[1]

Experience replay

- store each transition ($s_t; a_t; r_t; s_{t+1}$) the agent encounters
- update the Q-learning parameters by randomly sampling a batch of experiences from the replay memory
- perform gradient descent with the loss function calculated over the batch sampled from the replay memory
- **lower the correlation of inputs for training the Q-function.**

Target network

- The parameters θ^- of the target are only updated at every finite steps
- This setting helps to **stabilize the Q-value function training since the target is relatively stationary compared to the action-value function**



https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html

Received June 17, 2020, accepted July 6, 2020. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020.3010470

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Variational Quantum Circuits for Deep Reinforcement Learning

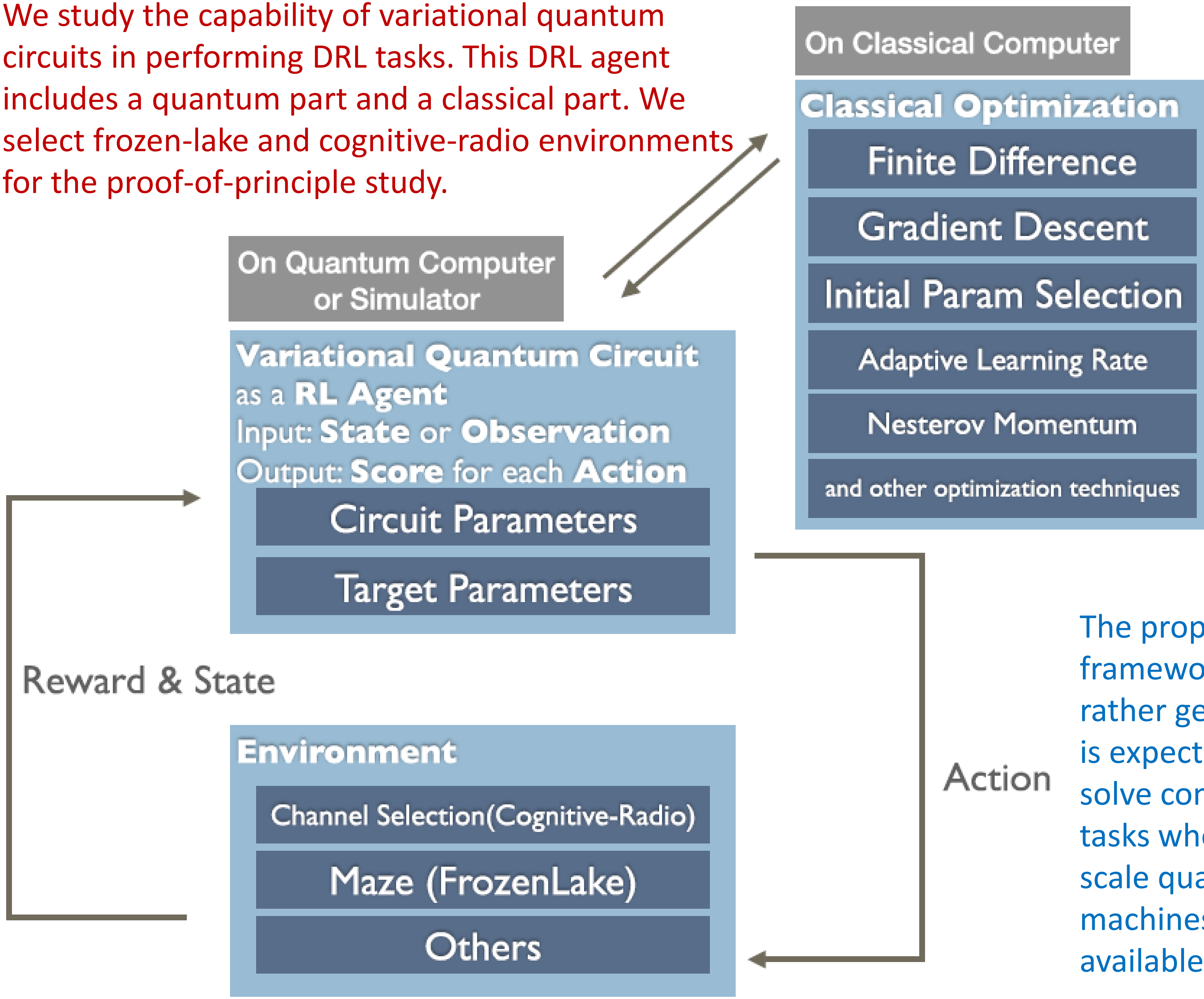
SAMUEL YEN-CHI CHEN[1,2], CHAO-HAN HUCK YANG[3], JUN QI[3], (Member, IEEE), PIN-YU CHEN[4], (Member, IEEE), XIAOLI MA[3], (Fellow, IEEE), AND HSI-SHENG GOAN[1,2,5]

[1]Department of Physics, National Taiwan University, Taipei City 10617, Taiwan
[2]Center for Theoretical Physics, National Taiwan University, Taipei City 10617, Taiwan
[3]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[4]IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA
[5]Center for Quantum Science and Engineering, National Taiwan University, Taipei City 10617, Taiwan

Corresponding authors: Samuel Yen-Chi Chen (ycchen1989@gmail.com) and Hsi-Sheng Goan (goan@phys.ntu.edu.tw)

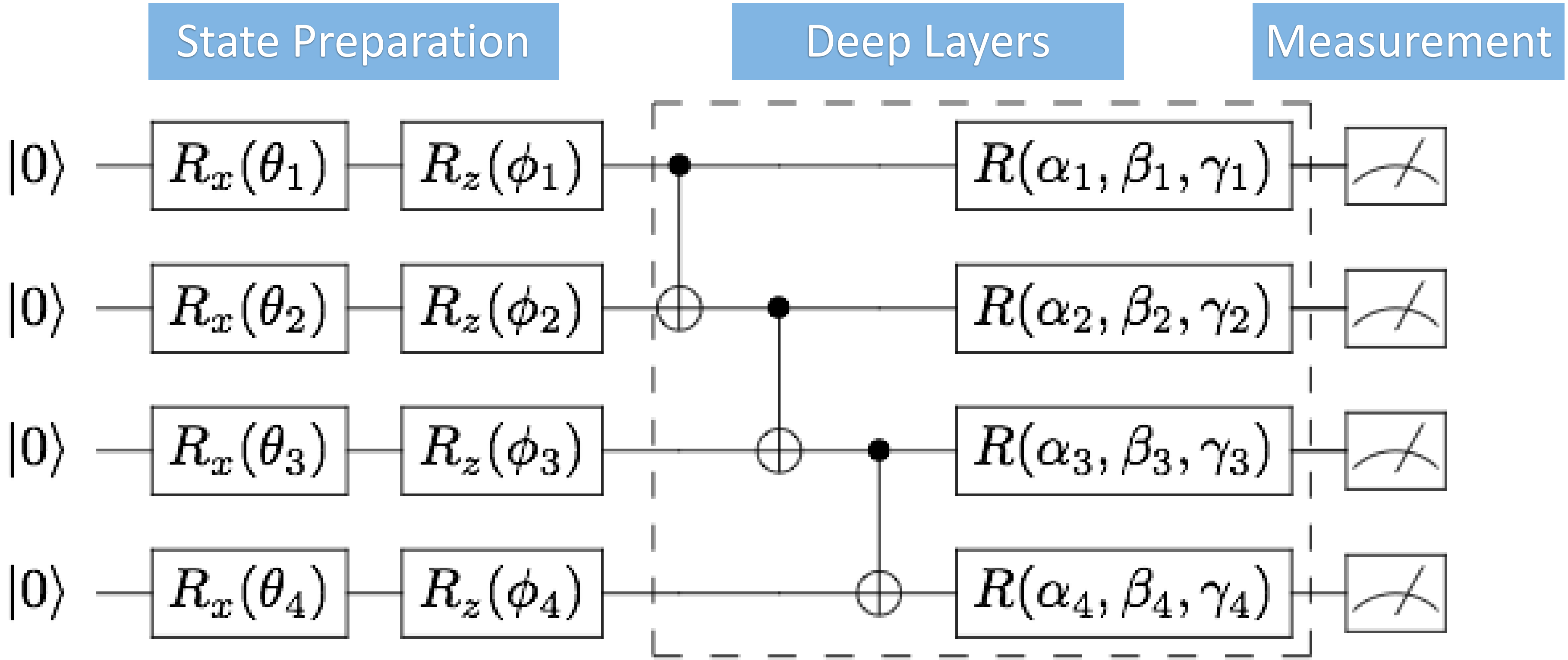
ABSTRACT The state-of-the-art machine learning approaches are based on classical von Neumann computing architectures and have been widely used in many industrial and academic domains. With the recent development of quantum computing, researchers and tech-giants have attempted new quantum circuits for machine learning tasks. However, the existing quantum computing platforms are hard to simulate classical deep learning models or problems because of the intractability of deep quantum circuits. Thus, it is necessary to design feasible quantum algorithms for quantum machine learning for noisy intermediate scale quantum (NISQ) devices. This work explores variational quantum circuits for deep reinforcement learning. Specifically, we reshape classical deep reinforcement learning algorithms like experience replay and target network into a representation of variational quantum circuits. Moreover, we use a quantum information encoding scheme to reduce the number of model parameters compared to classical neural networks. To the best of our knowledge, this work is the first proof-of-principle demonstration of variational quantum circuits to approximate the deep Q-value function for decision-making and policy-selection reinforcement learning with experience replay and target network. Besides, our variational quantum circuits can be deployed in many near-term NISQ machines.

We study the capability of variational quantum circuits in performing DRL tasks. This DRL agent includes a quantum part and a classical part. We select frozen-lake and cognitive-radio environments for the proof-of-principle study.



The proposed framework is rather general and is expected to solve complicated tasks when larger-scale quantum machines are available.

The quantum circuit



Learning Part

$$R(\alpha_i, \beta_i, \gamma_i) = R_z(\alpha_i) R_y(\beta_i) R_z(\gamma_i)$$

The grouped circuit can be repeated *k* times. Increasing the depth *k* enables the circuit to represent more complex structures.

Computational basis encoding and quantum circuit for the frozen lake problem

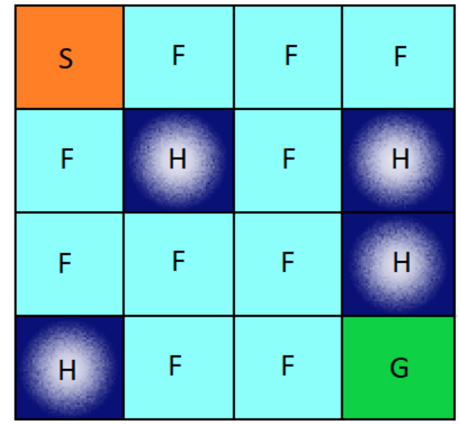
Environment with 16 states.

States numbered as 0~15

b_1, b_2, b_3, b_4

Example: State 12 : 1100 ->1, 1, 0, 0

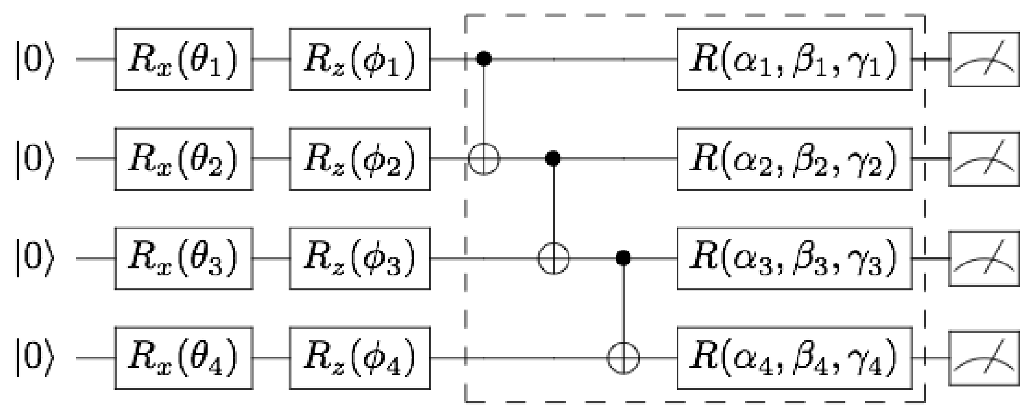
$|b_1\rangle \otimes |b_2\rangle \otimes |b_3\rangle \otimes |b_4\rangle$



Rotations : $\theta_i = \pi \times b_i$

$\phi_i = \pi \times b_i$

Result : $|1\rangle \otimes |1\rangle \otimes |0\rangle \otimes |0\rangle$

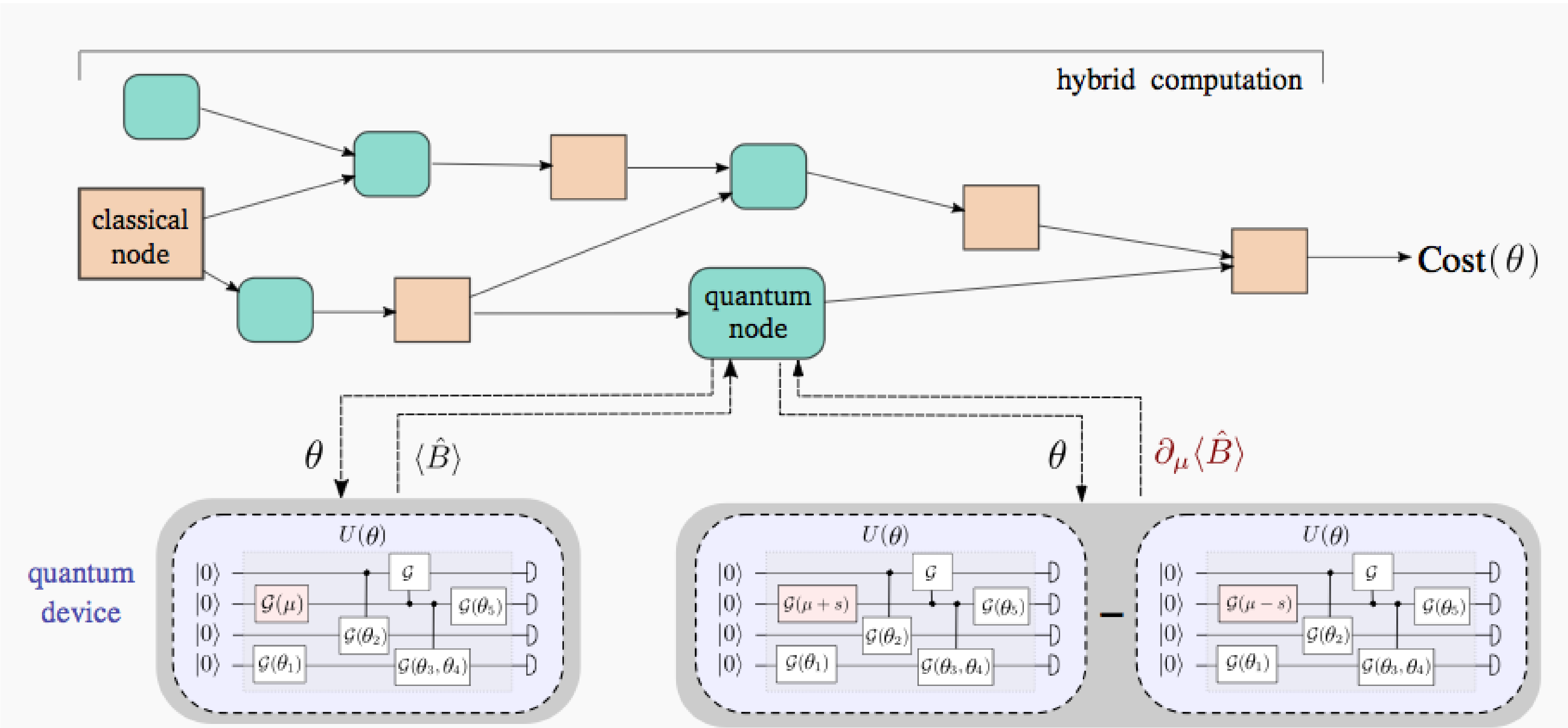


The grouped circuit repeats twice.

$$R_j(\theta) = e^{-i\theta\sigma_j/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_j.$$

$$R(\alpha_i, \beta_i, \gamma_i) = R_z(\alpha_i)R_y(\beta_i)R_z(\gamma_i)$$

Gradients of expectation values: “Parameter shift rule”



In the larger context of hybrid optimization, a quantum node, in which a variational quantum algorithm is executed, can compute derivatives of its outputs with respect to gate parameters by running the original circuit twice, but with a shift in the parameter in question.

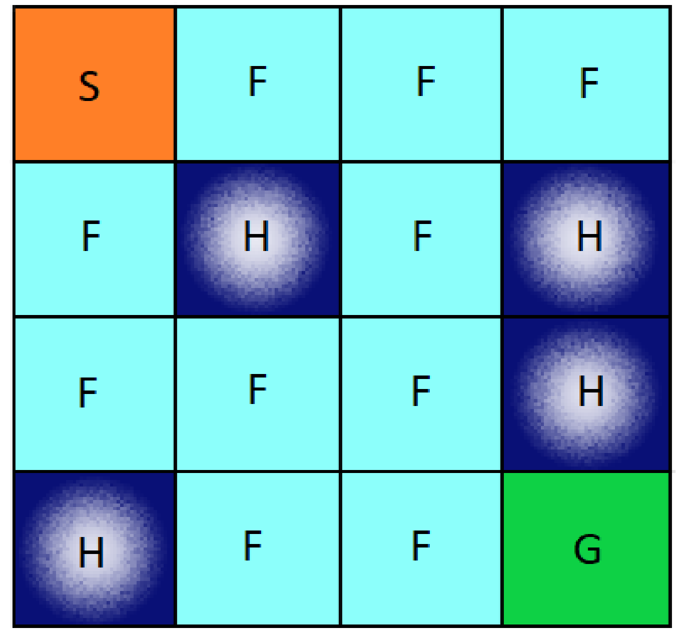
M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, Phys. Rev. A **99**, 032331 (2019)

Testing Env - Frozen-Lake

The agent is expected to go from the start location to the goal location. There are several holes on the way, and the agent should learn to avoid stepping into these hole locations, otherwise the agent will get a large negative reward and the episode will terminate.

LOCATION	REWARD
HOLE	-0.2
GOAL	1.0
OTHER	-0.01

to encourage the agent
to take the shortest path



The target circuit is updated per 20 steps.

- The outputs of the expectation values of qubits 0; 1; 2; 3 from the quantum circuit correspond to the action LEFT, DOWN, RIGHT, UP.
- The qubit *j* with the maximum output expectation value determines the next action.

Source: https://medium.com/swlh/introduction-to-reinforcement-learning-coding-q-learning-part-3-9778366a41c0

Parameters of experience replay and target network

- For **experience replay**, the replay memory is set for the length of 80 to adapt to the frozen-lake testing environment and the length of 1000 for the cognitive-radio testing environment
- The size of training batch is 5 for all of the environments.
- For a **target network**, we construct two sets of circuit parameters with the same circuit architecture.
- The target circuit is updated per 20 steps.

Numerical simulation

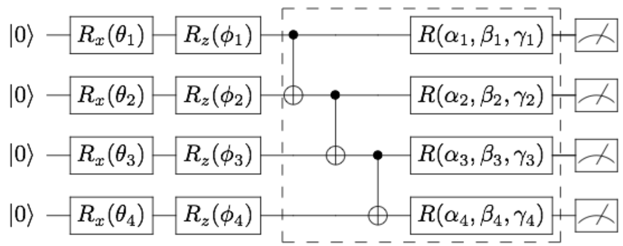
- The grouped box repeats two times regardless of the testing environments.
- The gradient descent optimization method is chosen to be RMSprop [1] in which parameters are set to be learning rate = 0.01, alpha = 0.99 and eps = 10^{-8}, widely used in DRL.
- The ε-greedy strategy used in the frozen-lake environment is the following:

$$\epsilon \leftarrow \frac{\epsilon}{\frac{episode}{100} + 1},$$

- In the cognitive-radio environment, ε is updated in every single step as:

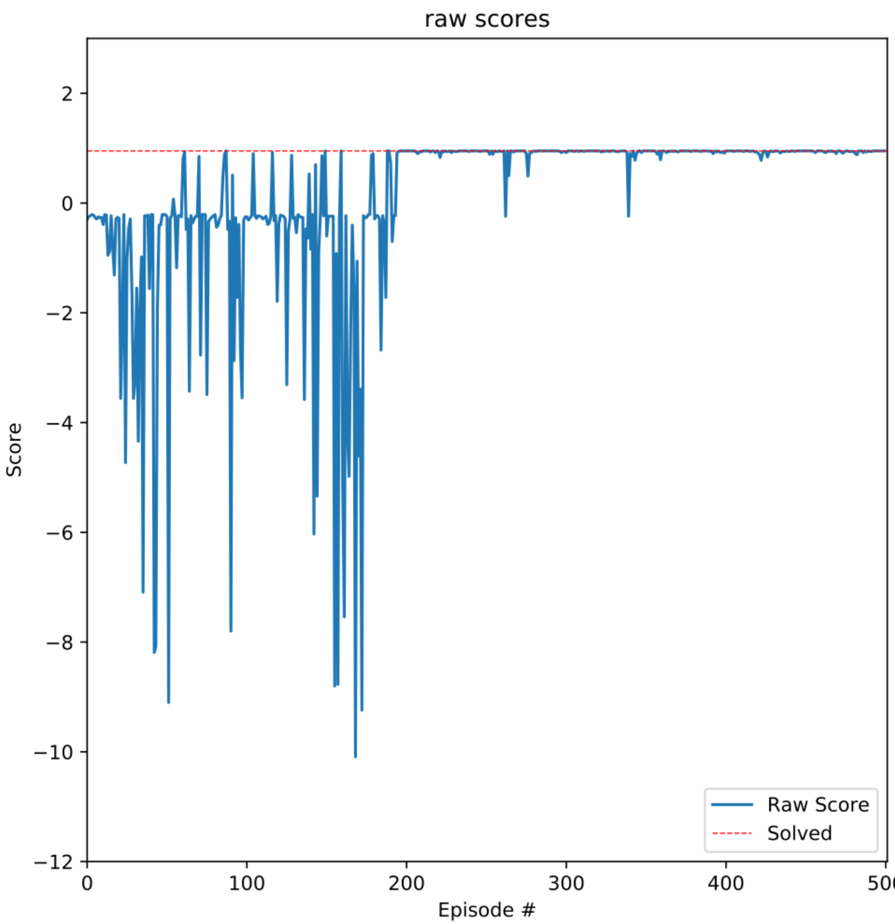
$$\epsilon \leftarrow 0.99\epsilon$$

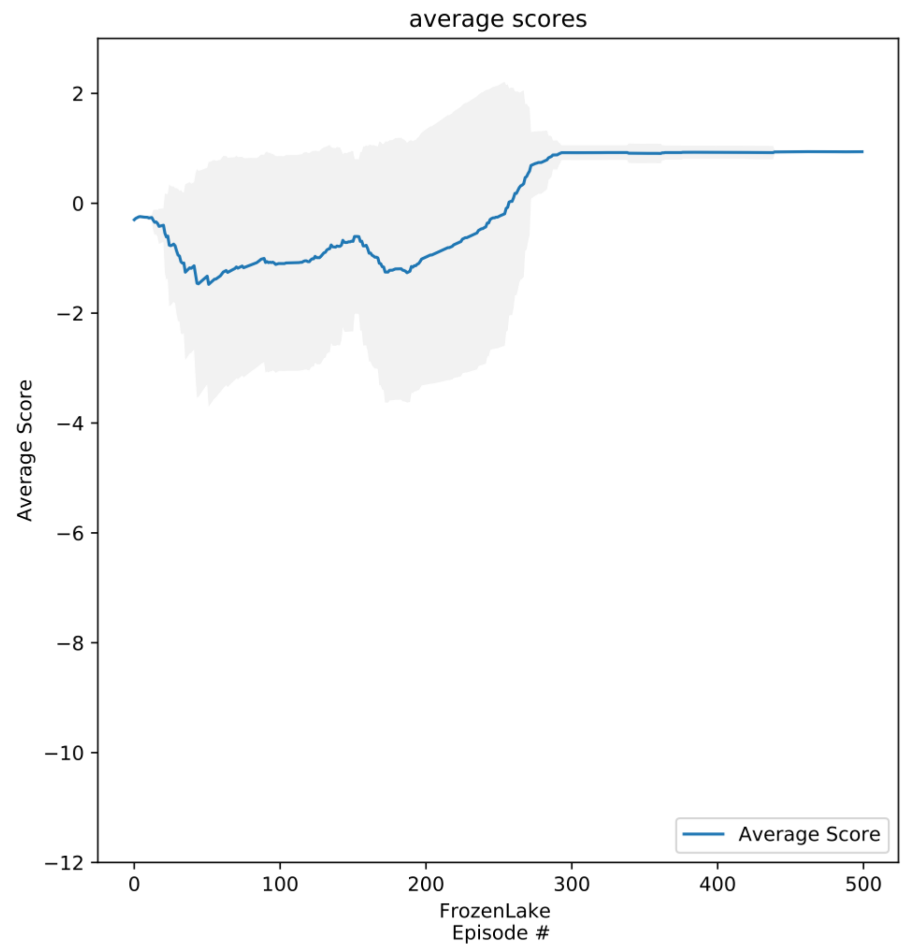
- with initial ε = 1.0 for encouraging more exploration in early episodes and shifting to more exploitation in later episodes.



[1] T. Tieleman and G. Hinton. Lecture 6.5—RmsProp: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for Machine Learning, 2012

Performance of Frozen-Lake testing environment





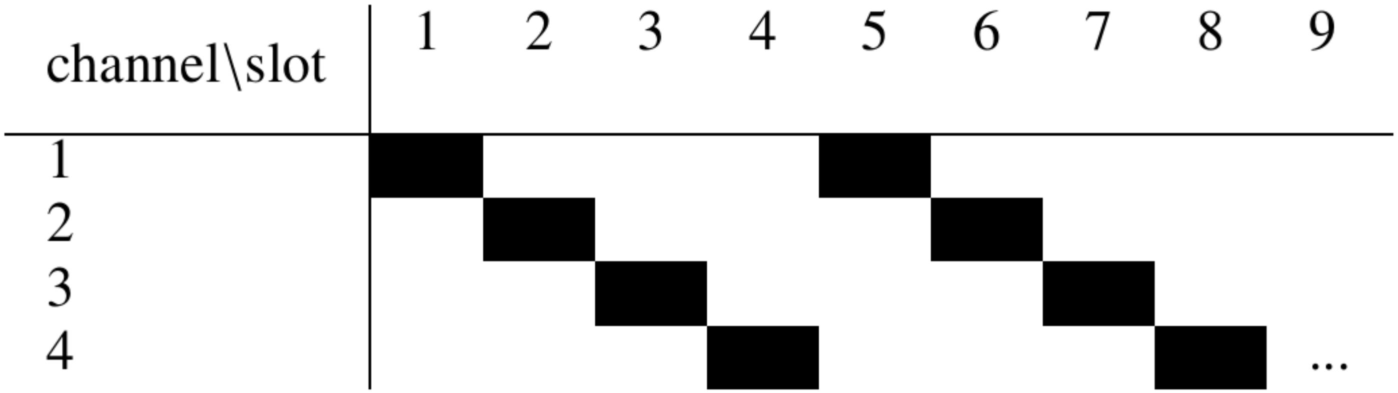
Testing environment - cognitive-radio

- We apply our VQ-DQN to a classical spectrum control problem in cognitive radio with the ns3-gym environment [1].
- We consider the problem of radio channel selection in a wireless multi-channel environment, e.g. 802.11 networks with external interference.
- We first create a scalable reinforcement learning environment sim-radio-spectrum (SRS) with a customized state and an action echo in a real multi-channel spectrum scenario. **The objective of the agent is to select a channel free of interference in the next time slot.**
- We consider a simple illustrative example where the external interference follows a periodic pattern, i.e. sweeping over all channels 1 to *n* in the same order. This environment offers a feasible test-bed for quantum DQL (DQN) with a desirable self-defined environment with a lower action and space complexity working in the NISQ machines.

[1] P. Gawłowicz and A. Zubow. *ns-3 meets OpenAI Gym: The Playground for Machine Learning in Networking Research*. In ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), November 2019.

Testing Env - cognitive-radio

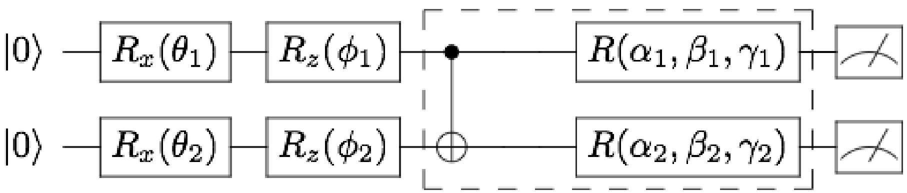
Consider the case that the external interference of channel-changing follows a periodic pattern with *n* time-steps in a full cycle, i.e. sweeping over all channels one to four in the same order.



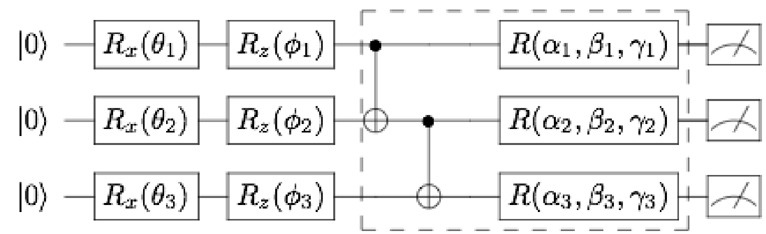
There are *n* possible channels for the agent to select. At each time step, the agent should select one channel that is not occupied or interfered. The episode will terminate if the agent collects three failed selections or the agent plays more than 100 steps.

LOCATION	REWARD
Occupied Channel	-1.0
Available Channel	1.0

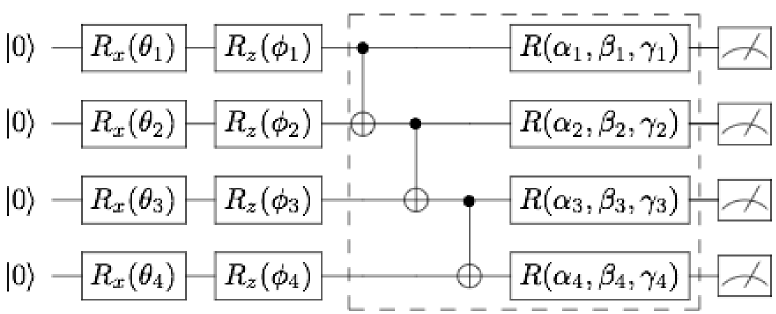
Variational quantum circuits for cognitive-radio testing environment



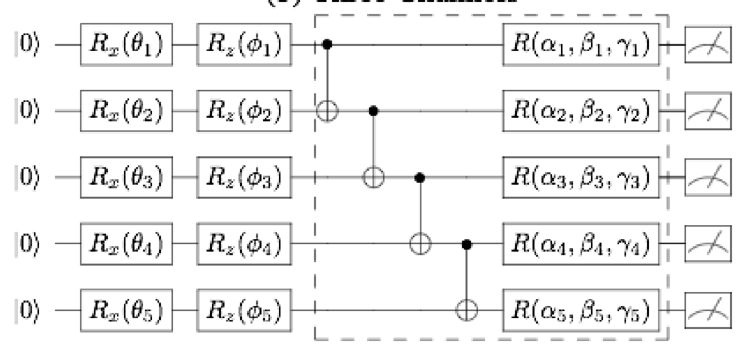
(a) Two Channels



(b) Three Channels



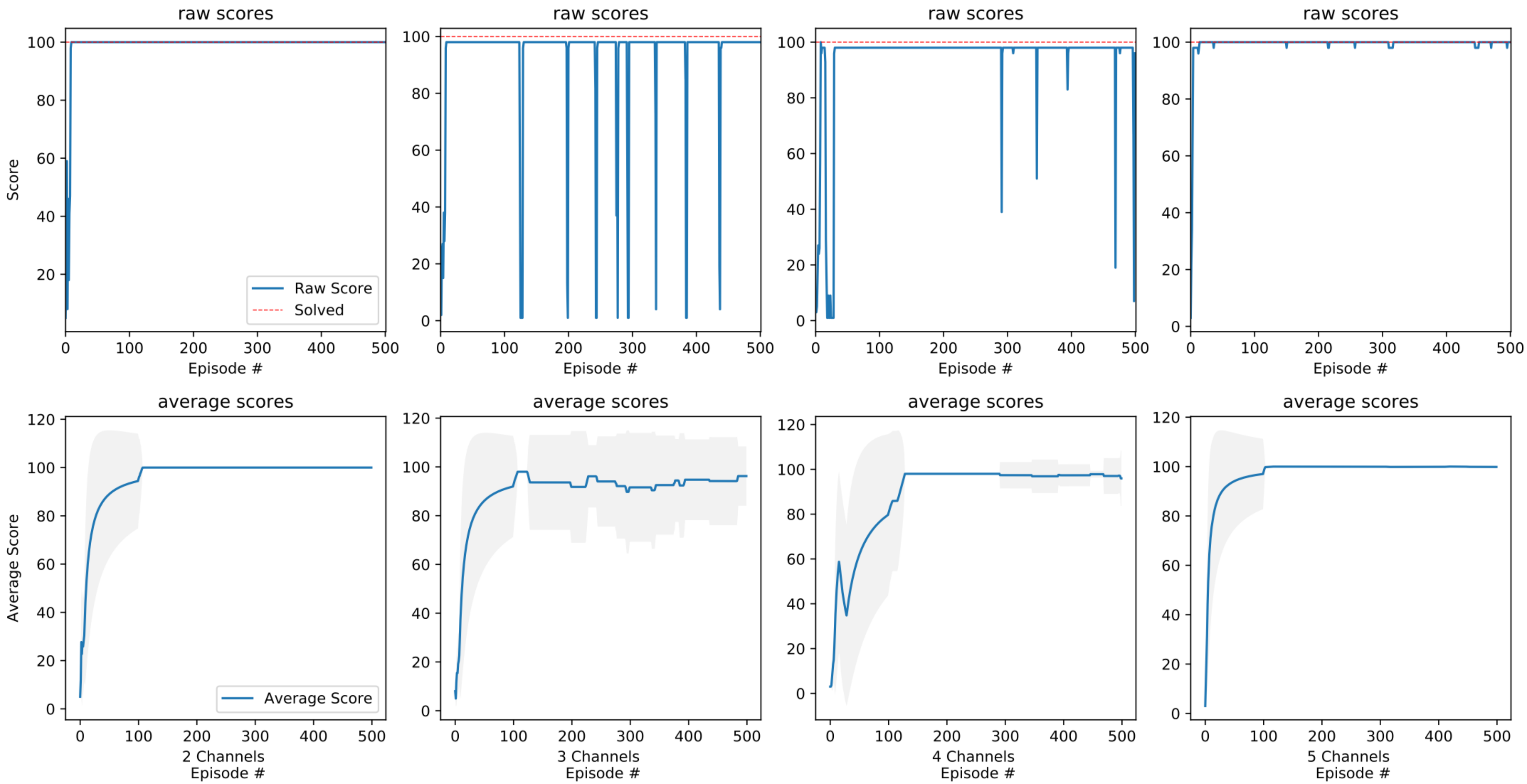
(c) Four Channels



(d) Five Channels

The grouped circuit repeats twice.

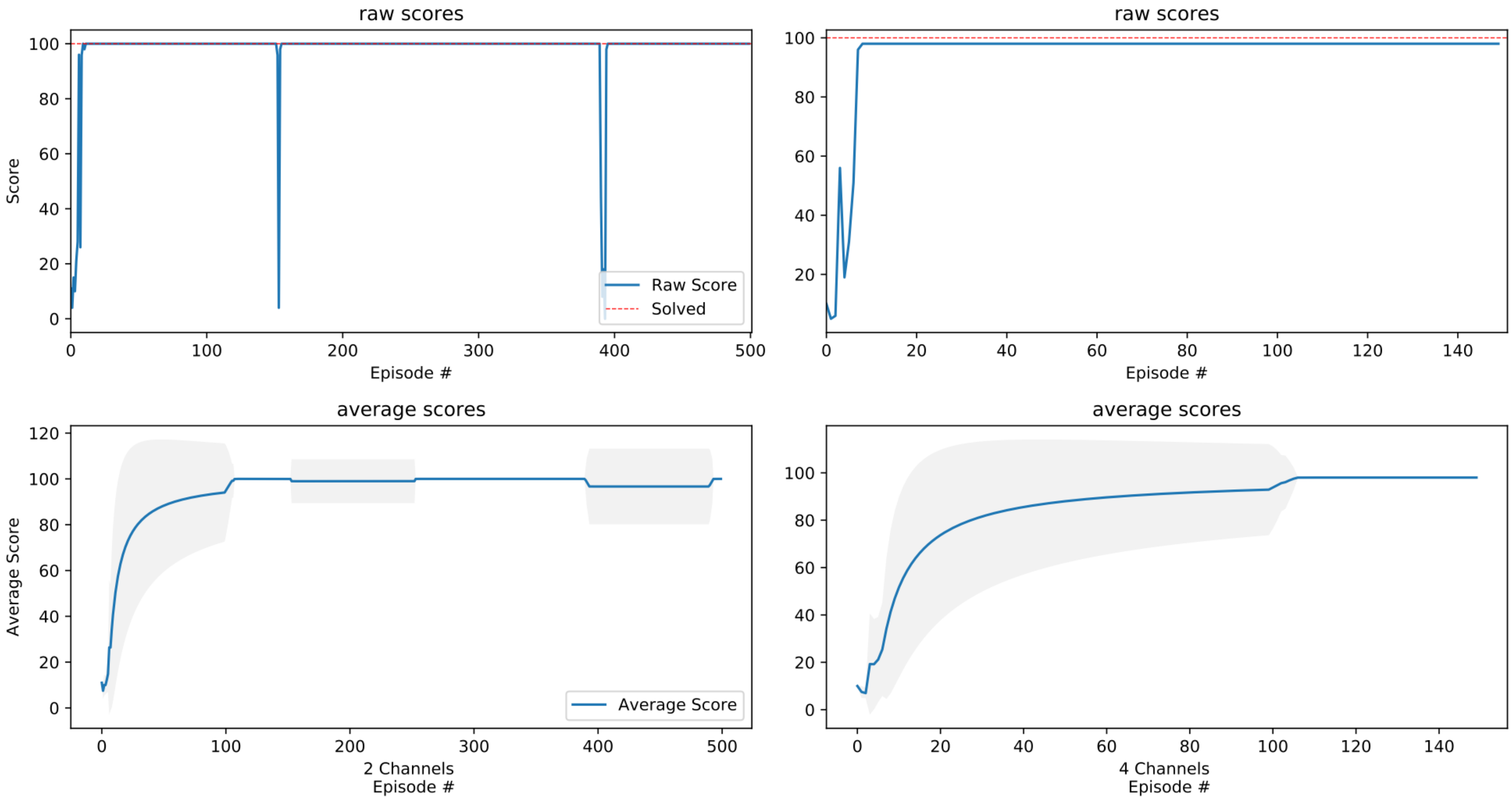
Performance of cognitive-radio testing environment



Simulation with noise model

- To investigate the robustness of our proposed variational quantum circuit-based DRL against the noise from current and possible near-term NISQ devices, we perform additional simulations which include the noises from the real quantum computer.
- We use the IBM **Qiskit-Aer simulator backend**, which has the capability to incorporate the noise model from the IBM quantum computers.
- Noise model from IBM 20-qubit **ibmq_poughkeepsie** machine.

Testing Env - Cognitive-radio with noise



The variational circuits can be relatively robust against noises because the related deviations can be absorbed by the parameters during the iterative optimization process. Therefore, the machine learning algorithms powered by variational quantum circuits can circumvent the complex quantum errors which exist in the available quantum devices.

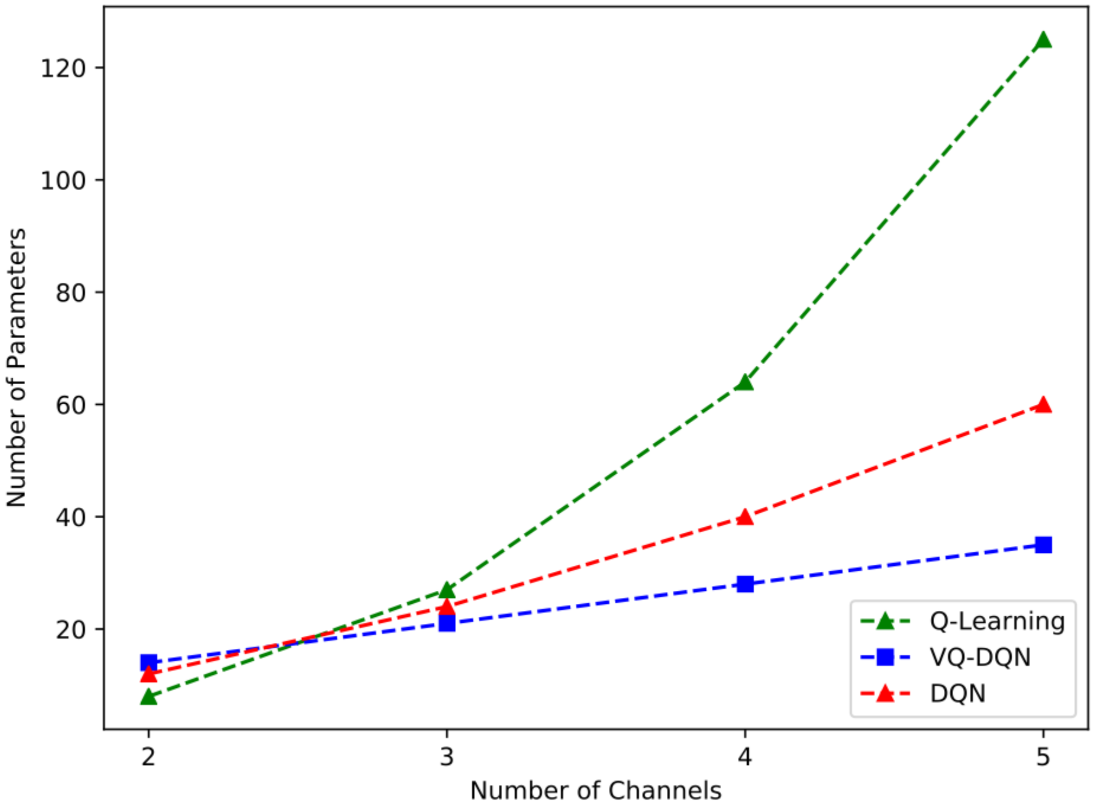
TABLE 3. **Results of the trained VQ-DQL (VQ-DQN) for the cognitive-radio experiment conducted on the IBM Q quantum computer, `ibmq-valencia`. In this experiment, we test the trained quantum DRL model in the cognitive-radio experiment of configuration (c) in the 4-channel case described in Fig. 3. Even if the training is on the simulation software without the quantum noise, the trained model still performs well on the real quantum computer.**

Episodes	1	2	3	4	5
Total Steps	100	100	100	100	100
Total Reward	100	100	100	100	98

TABLE 4. **Results of the trained VQ-DQL (VQ-DQN) for the frozen-lake experiment conducted on the IBM Q quantum computer, `ibmq-valencia`. In this experiment, we test the trained quantum DRL model in the frozen-lake experiment of configuration (c) in Fig. 1. Even if the training is on the simulation software without the quantum noise, the trained model still performs well on the real quantum computer.**

Episode	1	2	3	4	5	6	7
Total Steps	6	6	6	7	7	7	6
Total Reward	0.95	0.95	0.95	0.94	0.94	0.94	0.95

TESTING ENV - COGNITIVE-RADIO - PARAMETERS SAVING



Q-learning: n^3

DQN: $2\times n^2 + 2\times n$

VQ-DQN: $n(3 \times 2+1)$

n is the dimension of input vectors (channel number)

Conclusion

- Demonstration of variational quantum circuits to approximate the deep Q-value function with experience replay and target network.
- Our proposed framework shows the quantum advantage in terms of less memory consumption and the reduction of model parameters:
  - Q-learning: $\mathcal{O}(n^3)$
  - DQN: $\mathcal{O}(n^2)$
  - VQ-DQN: $\mathcal{O}(n)$
- Our *hybrid*-quantum-classical machine learning variational quantum circuits can be deployed in many near-term NISQ machines.
- Current limitations: available quantum devices and simulators
- Choosing a suitable *encoding scheme* is crucial to gain the quantum advantage.