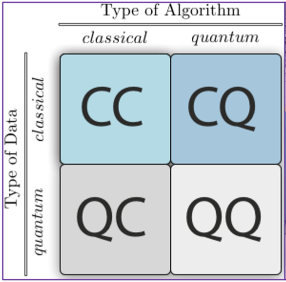
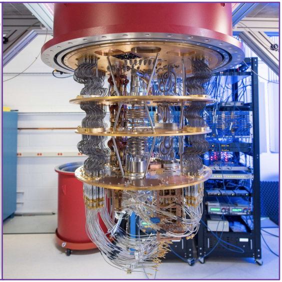
Quantum AI Computing - Beyond the Hype

Data Science Nigeria Bootcamp, 2021





Adewale (Wale) Akinfaderin
{@waleakinfaderin}

November 2, 2021

Talk Outline

1. About Me
2. ML Needs & Hardware Limitations
3. Introduction to Quantum Computing
4. Model of Quantum Computation
5. From Quantum Computing to Quantum Machine Learning

Talk Outline

1. About Me
2. ML Needs & Hardware Limitations
3. Introduction to Quantum Computing
4. Model of Quantum Computation
5. From Quantum Computing to Quantum Machine Learning

About Me

- Scientist at Amazon Web Services working on ML for healthcare
- Google Developer Expert, Machine Learning.
- Former experience developing and productionizing ML algorithms in the retail and renewable energy space.
- Doctoral research in Resonance Spectroscopy (Spin Physics).
- Interested in democratizing ML, ML for Developing World, Responsible AI, and Machine Translation for Low Resource Languages
- Committee and Reviewer: NeurIPS ML4D (2019, 2020), ACL (2021)
- I really like soccer and scrabble.

Talk Outline

Dataset

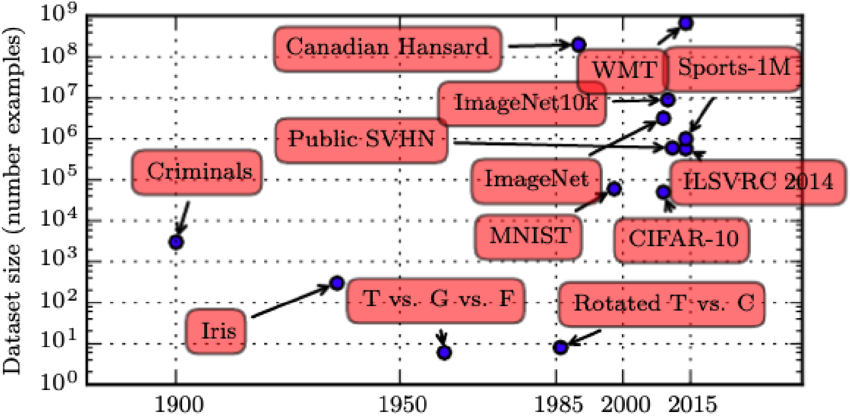


Figure: Increasing Dataset Over Time. *I. Goodfellow, et al. Deep Learning*

Network Size

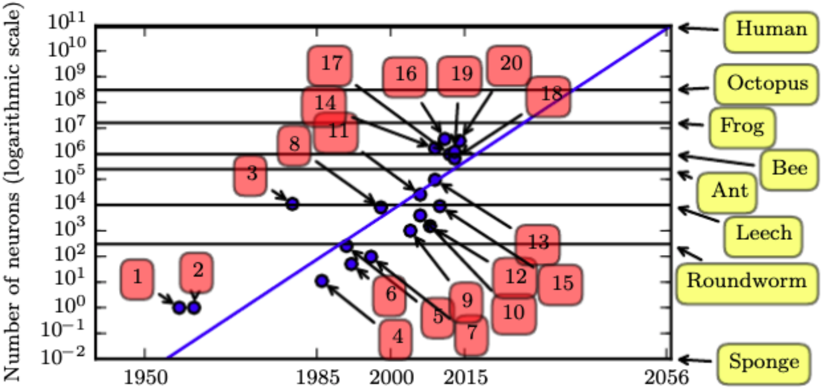


Figure: Increasing NN Size Over Time. *I. Goodfellow, et al. Deep Learning*

Limitations of Conventional Hardware Technology

- Moore's Law
- Moore's Second Law (Economics); Costs of a new technology node is doubling.
- Dennard Scaling is Dead

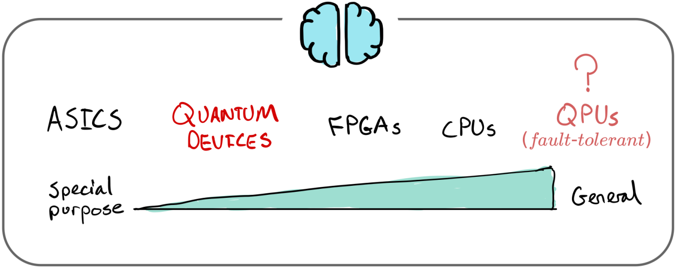
New Architectures for Machine Learning

Same CMOS Technology, New Architectures

- GPU
- ASICs
- FPGA
- Custom Processors for Deep Learning, e.g. TPU (By Google)

Beyond CMOS

- Superconducting Computing
- Memristors
- Stochastic and Chaotic Computing
- **Quantum Computing**



Quantum computers as AI accelerators

The limits of what machines can learn have always been defined by the computer hardware we run our algorithms on—for example, the success of modern-day deep learning with neural networks is enabled by parallel GPU clusters.

Quantum machine learning extends the pool of hardware for machine learning by an entirely new type of computing device—the quantum computer. Information processing with quantum computers relies on substantially different laws of physics known as *quantum theory*.

Talk Outline

Quantum Computing - The Starting Point

International Journal of Theoretical Physics, Vol. 21, Nos. 6/7, 1982

Simulating Physics with Computers

Richard P. Feynman

Department of Physics, California Institute of Technology, Pasadena, California 91107

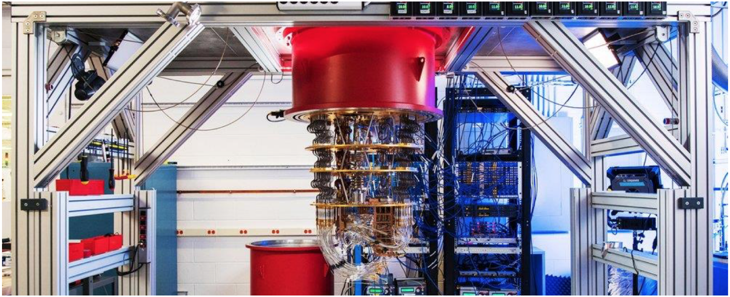
Received May 7, 1981

Elements of Quantum Computing

- Three elements are needed for computations: data, operations, results

For Quantum Computation:

- Data ⇒ **qubit**
- Operation ⇒ **quantum gate**
- Results ⇒ **Measurements**



Bits

A building block of classical computational devices is a two-state system or a classical bit:

$0 \bullet \qquad\qquad\qquad \bullet 1$

Indeed, any system with a finite set of discrete and stable states, with controlled transitions between them, will do:

What is this quantum state?

Richard Feynman (Nobel Prize in Physics, 1965)

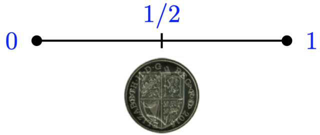
Nobody understands quantum mechanics.

Steven Weinberg (Nobel Prize in Physics, 1979)

My own conclusion is that today there is no interpretation of quantum mechanics that does not have serious flaws. This view is not universally shared. Indeed, many physicists are satisfied with their own interpretation of quantum mechanics. But different physicists are satisfied with different interpretations.

Probabilistic Bits

When you don't know the state of a system exactly but only have partial information, you can use probabilities to describe it:



It is convenient to represent system's state using vectors:

 $= \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  $= \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Then a uniformly random bit is represented by

 $= \frac{1}{2}$  $+ \frac{1}{2}$  $= \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Using probabilities to represent information (or lack of it...) is more useful than you might think!

Quantum Superposition

In nature, the state of an actual physical system is more uncertain than we are used to in our daily lives. . .

© *Charles Addams, The New York Times*

That's why complex amplitudes rather than probabilities are used in quantum mechanics!

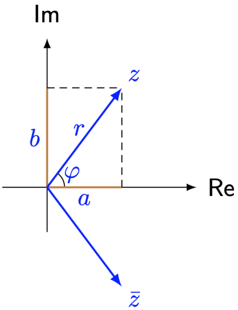
Complex Numbers - A One Slide Review

Representations:

- algebraic: $z = a + ib$
- exponential: $z = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$

Operations:

- addition and subtraction:
  $(a+ib) \pm (c+id) = (a \pm c) + i(b \pm d)$
- multiplication:
  $(a+ib) \cdot (c+id) = (ac - bd) + i(ad + bc)$
  $re^{i\varphi} \cdot r'e^{i\varphi'} = rr'e^{i(\varphi+\varphi')}$
- complex conjugate: $z^* = \bar{z} = a - ib = re^{-i\varphi}$
- absolute value:
  $|z| = \sqrt{a^2 + b^2} = r$, $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$
- absolute value squared: $|z|^2 = a^2 + b^2 = r^2$
  important: $|z|^2 = z\bar{z}$
- inverse: $1/z = \bar{z}/|z|^2$



From Classical to Quantum Bits (Qubits)

Classical

Recall that a random bit can be described by a probability vector:

p 

$+\, q$  $= p \begin{pmatrix} 1 \\ 0 \end{pmatrix} + q \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$

where $p, q \in \mathbb{R}$ such that $p, q \geq 0$ and $p + q = 1$.

Quantum

A quantum bit (or qubit for short) is described by a quantum state:

$$\alpha|0\rangle + \beta|1\rangle = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

where $\alpha, \beta \in \mathbb{C}$ are called amplitudes and satisfy $|\alpha|^2 + |\beta|^2 = 1$. Here $|0\rangle, |1\rangle$ are used as place-holders for the two discernible states of a coin (or any other physical system for that matter).

Any system that can exist in states $|0\rangle$ and $|1\rangle$ can also exist in a superposition $\alpha|0\rangle + \beta|1\rangle$, according to quantum mechanics!

Classical and Quantum Measurements

Classical

Observing a random coin

p  $+\, q$ 

results in heads with probability p and tails with probability q.

Quantum

Measuring the quantum state

$$\alpha|0\rangle + \beta|1\rangle$$

results in $|0\rangle$ with probability $|\alpha|^2$ and $|1\rangle$ with probability $|\beta|^2$.

Important:

- After the measurement, the system is in the measured state, so repeating the measurement will always yield the same value!
- We can only extract one bit of information from a single copy of a random bit or a qubit!

Global and Relative Phases

Phase

If $re^{i\varphi}$ is a complex number, $e^{i\varphi}$ is called phase.

Global phase

The following states differ only by a global phase:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad e^{i\varphi}|\psi\rangle = e^{i\varphi}\alpha|0\rangle + e^{i\varphi}\beta|1\rangle$$

These states are indistinguishable! Why? Because $|\alpha|^2 = |e^{i\varphi}\alpha|^2$ and $|\beta|^2 = |e^{i\varphi}\beta|^2$ so it makes no difference during measurements.

Relative phase

These states differ by a relative phase:

$$|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad |-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

Are they also indistinguishable? No! (Measure in a *different basis*.)

Remember: global phase does not matter, relative phase matters!

Qubit States: the Bloch Sphere

Any qubit state can be written as

$$|\psi\rangle = \underbrace{\cos\tfrac{\theta}{2}}_{\alpha}|0\rangle + \underbrace{e^{i\varphi}\sin\tfrac{\theta}{2}}_{\beta}|1\rangle$$

for some angles $\theta \in [0,\pi]$ and $\varphi \in [0, 2\pi)$.

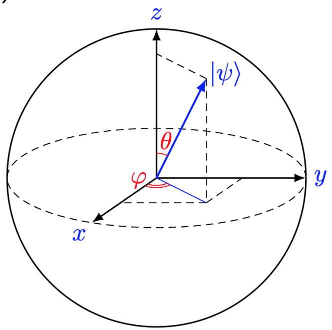
There is a one-to-one correspondence between qubit states and points on a unit sphere (also called Bloch sphere):

Bloch vector of $|\psi\rangle$ in spherical coordinates:

$$\begin{cases} x = \sin\theta\cos\varphi \\ y = \sin\theta\sin\varphi \\ z = \cos\theta \end{cases}$$

Measurement probabilities:

$$|\alpha|^2 = (\cos\tfrac{\theta}{2})^2 = \tfrac{1}{2} + \tfrac{1}{2}\cos\theta$$
$$|\beta|^2 = (\sin\tfrac{\theta}{2})^2 = \tfrac{1}{2} - \tfrac{1}{2}\cos\theta$$

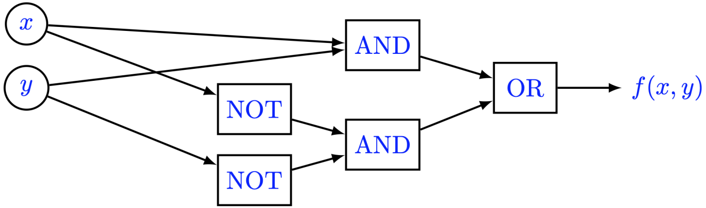


Talk Outline

Boolean Circuits

Logical gates model elementary computational steps in digital electronic circuits. E.g., $\mathrm{XOR}(x, y) = x \oplus y$, $\mathrm{NOT}(x) = x \oplus 1$, $\mathrm{AND}(x, y) = xy$.

Any Boolean function $f : \{0, 1\}^n \to \{0, 1\}$ can be expressed in terms of these. E.g., equality $f(x, y) := (x = y)$ can be expressed as follows:



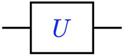
Universal sets of logical gates:

$$\{\mathrm{AND}, \mathrm{OR}, \mathrm{NOT}\} \qquad \{\mathrm{AND}, \mathrm{NOT}\} \qquad \{\mathrm{NAND}\}$$

Single-Qubit Gates

A 1-qubit gate is a 2×2 unitary. It can be written as follows:

$$U = \sum_{i,j \in \{0,1\}} U_{i,j} |i\rangle\langle j|$$



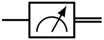
Example: The quantum analogue of logical NOT is the Pauli gate $X = |0\rangle\langle 1| + |1\rangle\langle 0| = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$:

$$X|0\rangle = |1\rangle \qquad\qquad X|1\rangle = |0\rangle$$

Example: The Hadamard gate $H = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ changes basis between the standard basis $\{|0\rangle, |1\rangle\}$ and the Hadamard basis $\{|+\rangle, |-\rangle\}$:

$$H|0\rangle = |+\rangle \qquad\qquad H|+\rangle = |0\rangle$$
$$H|1\rangle = |-\rangle \qquad\qquad H|-\rangle = |1\rangle$$

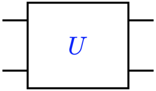
The standard basis measurement of a qubit is given by $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ and is not unitary. It has quantum input and classical output:



Multi-Qubit Quantum Gates

A 2-qubit gate is a 4×4 unitary. It can be written as follows:

$$\begin{aligned} U &= \textstyle\sum_{i,j,k,l \in \{0,1\}} U_{ij,kl} |i,j\rangle\langle k,l| \\ &= \textstyle\sum_{i,j,k,l \in \{0,1\}} U_{ij,kl} |i\rangle\langle k| \otimes |j\rangle\langle l| \end{aligned}$$

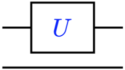


Example: $|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|$ is the 2-qubit identity.

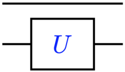
An n-qubit gate is a $2^n \times 2^n$ unitary. It has n input and n output qubits.

If a 1-qubit gate is applied locally on one of two qubits, we can write this as a 2-qubit unitary:

$U \otimes I$



$I \otimes U$

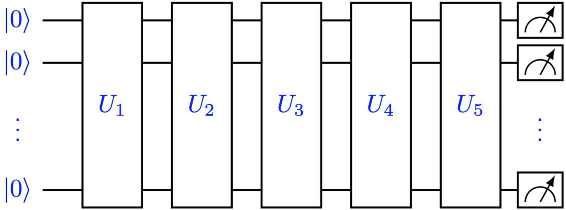


If a 2-qubit gate is applied locally on qubits 1 and 3, we can write this as a 3-qubit unitary:

$$\sum_{i,j,k,l \in \{0,1\}} U_{ij,kl} |i\rangle\langle k| \otimes I \otimes |j\rangle\langle l|$$

Quantum Circuits

An n-qubit quantum circuit is a sequence of n-qubit unitary operations, followed by the measurement in the standard basis:

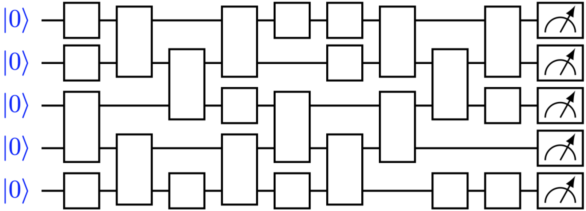


Typically the initial state is $|00\dots0\rangle = |0\rangle^{\otimes n}$ and we measure all qubits in the standard basis at the end.

Fact: While one can imagine more general circuits with intermediate measurements, all measurements can always be deferred to the end and converted into independent standard basis measurements for each qubit.

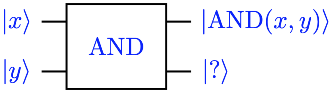
Quantum Algorithm

A quantum algorithm is an infinite family $C_1, C_2, \ldots$ of quantum circuits, where C_n acts on n qubits and consists of a finite sequence of 1-qubit and 2-qubit gates: $C_n = (U_1, U_2, \ldots, U_{L(n)})$ where $L(n)$ denotes the number of gates in the circuit. The map $n \rightarrow C_n$ must be efficiently computable (e.g., in polynomial time on a deterministic Turing machine).



Quantum AND gate

Let's try to find a unitary that computes the logical AND of two bits. Given $x, y \in \{0, 1\}$, it should output $\mathrm{AND}(x, y)$ on the first qubit:



What should it output on the second qubit? Some arbitrary states:

$$
\begin{aligned}
|00\rangle &\mapsto |0\rangle|\psi_1\rangle \\
|01\rangle &\mapsto |0\rangle|\psi_2\rangle \\
|10\rangle &\mapsto |0\rangle|\psi_3\rangle \\
|11\rangle &\mapsto |1\rangle|\psi_4\rangle
\end{aligned}
$$

A unitary gate must preserve orthogonality: since $|00\rangle$, $|01\rangle$, $|10\rangle$ are orthogonal, so must be the three output states $|\psi_1\rangle$, $|\psi_2\rangle$, $|\psi_3\rangle$, but it is not possible to have three mutually orthogonal states in $\mathbb{C}^2$!

Talk Outline

Search Based Algorithms (Grover)

- To measure distances or other metrics between data points.
- Examples: Clustering (K-means, K-medians), K-NN, etc.
- Speedup: Quadratic, Exponential

Quantum Machine Learning

Quantum Basic Linear Algebra Subroutines (qBLAS)

- Exponential speed up in Fourier transforms, finding eigenvectors and eigenvalues, solving linear equations
- Useful in machine learning algorithms such as: Least-square fitting, Gradient dissent, Principal Component Analysis, Single Value Decomposition, Support vector machines

qBLAS Caveats

- The input problem: Quantum algorithms provide dramatic speedups for processing data, they seldom provide advantages in reading data. The cost of reading in the input can dominate the cost of quantum algorithms. This cost can be exponential!
- The output problem. Obtaining the full solution from some quantum algorithms as a string of bits requires learning an exponential number of bits.

Adiabatic Quantum Computing (Quantum Annealing)

- For Solving optimization problems

Grover's Search Problem

One of the two most important algorithms in quantum computing is Grover's search algorithm (invented by Lov Grover in 1996) for searching for a particular value in an unstructured / unsorted search space.

Example: Searching in a sorted vs unsorted database:

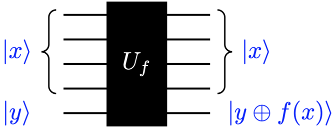
- find a name in a telephone directory
- find a phone number in a telephone directory

Given a black box that for each of N different input strings answers either **yes** or **no**, and there is a unique string with answer **yes**, Grover's algorithm finds this string with $O(\sqrt{N})$ questions (with high probability).

This is a quantum alternative to brute-force search.

Grover's Search Problem

$$U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$$



We refer to U_f as the black box or oracle for computing f.

Suppose there is a unique $a \in \{0,1\}^n$ that yields value 1. Let

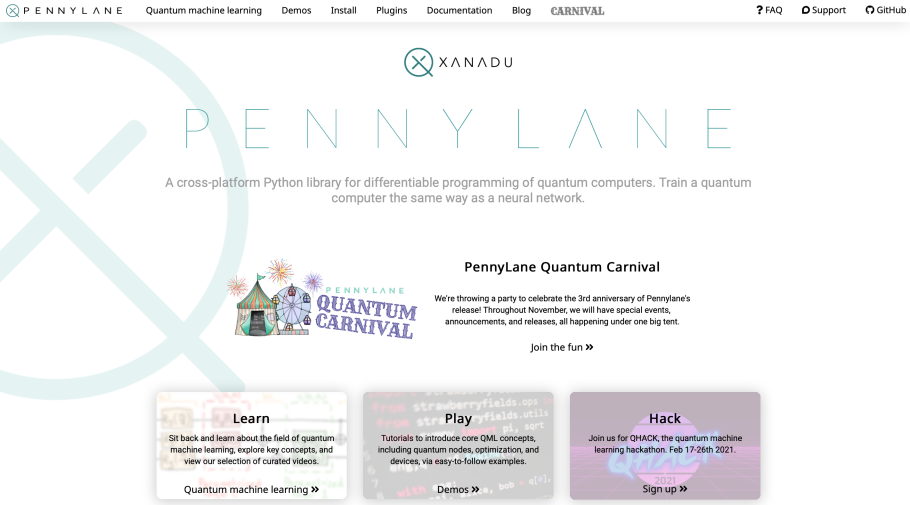
$$f_a(x) = \begin{cases} 1 & \text{if } x = a \\ 0 & \text{otherwise} \end{cases}$$

Grover's algorithm can determine the value of a with $O(\sqrt{N})$ calls to the black box U_{f_a} where $N = |\{0,1\}^n| = 2^n$.

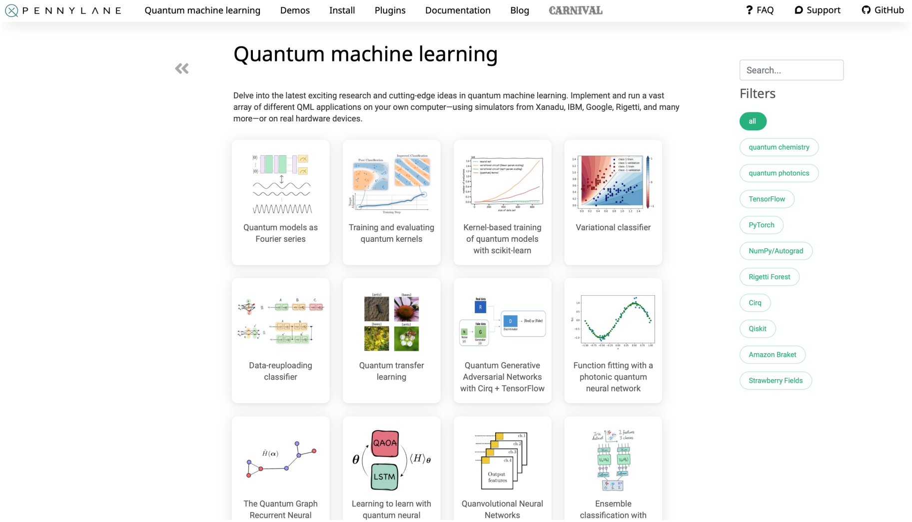
Some Final Notes

- Quantum machine learning is not going to solve all problems of machine learning.
- There are specific limitations that come with quantum machine learning. However, it can be useful in specific tasks.
- There are multiple technical and theoretical challenges in front of quantum machine learning.
- We need multiple breakthrough innovations in order to achieve quantum machine learning. qRAM seems to be the main one.

Quantum ML Python Library: pennylane.ai



PENNYLANE AI

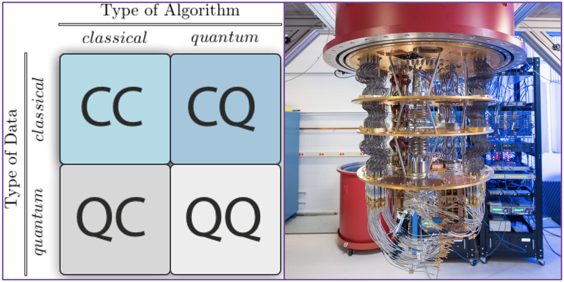


References

- https://pennylane.ai/
- https://qiskit.org/
- Cambridge: https://www.cl.cam.ac.uk/teaching/1617/QuantComp/materials.html

Quantum AI Computing - Beyond the Hype

Data Science Nigeria Bootcamp, 2021



Adewale (Wale) Akinfaderin
{@waleakinfaderin}

November 2, 2021