Using Quantum Computers for Machine Learning

Kaitlin Gili
Stevens Institu
te of Technology

Conference for Undergraduate Women in Physical Sciences
November 8, 2019

$1.2 Billion

$1.5 Billion

Over $10 Billion

$8 Billion...

Eventually

General Outline

- ➢ What is a quantum computer really?
- ➢ How can a quantum computer improve machine learning?
- ➢ How is this going to contribute to the 2027 projection?

What is a quantum computer really?

How Does Your Computer Really Work?

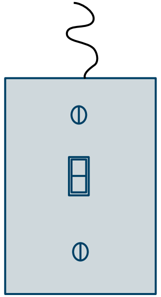
Full Scale Device

Computer Chip

Transistor

001010001110000000001111
000000000000010001000001
1111110010101010101

Bits



OFF

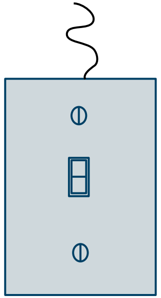
ON

How Does Your Computer Really Work?

Full Scale Device

Computer Chip

Transistor



OFF

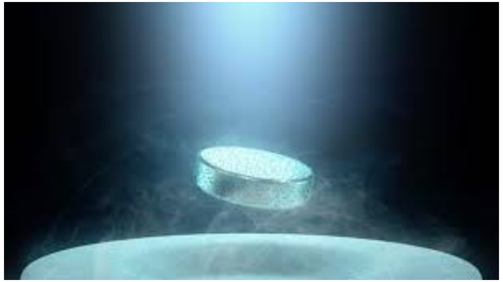
ON

00101000111000000000001111
00000000000000100010000001
1111110010101010101010101

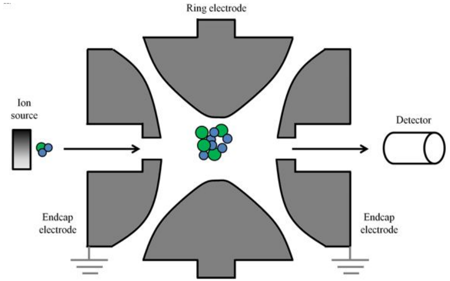
Bits

How Does A Quantum Computer Work?

Photonic



Superconducting

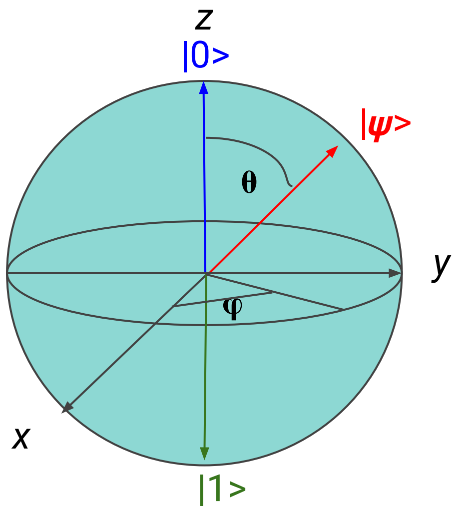


Trapped Ions

0010100011100000000001111 0000000000000100010000001 111111001010101010101

Qubits

What is a Qubit?



Bloch Sphere

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$$

Superposition

A qubit can be both be 0 and 1 at the same time, holding 2^n states per qubit

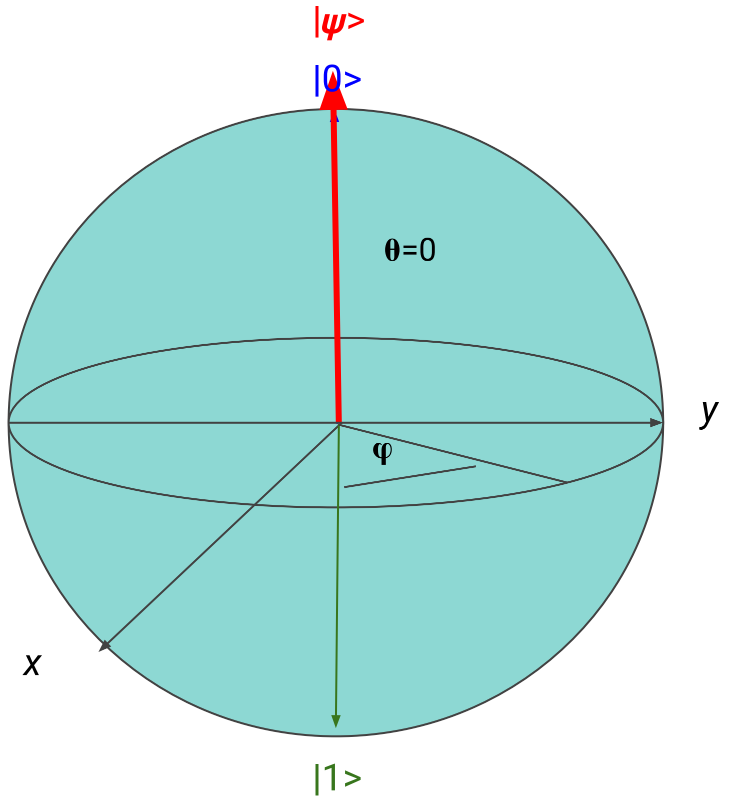
Measurement

A qubit can only be measured classically as a 0 or 1, but it will have some inherent probability of being one or the other

Entanglement

We can connect two qubits so that if we know information about one, then we instantly know some information about the other

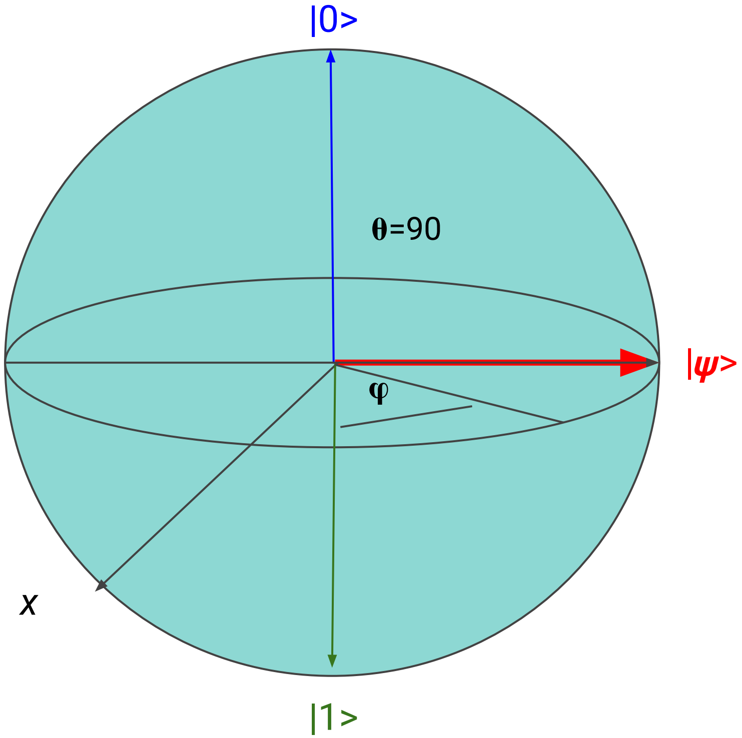
How Can We Control It?



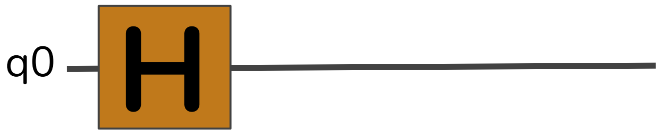
$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$$

q0

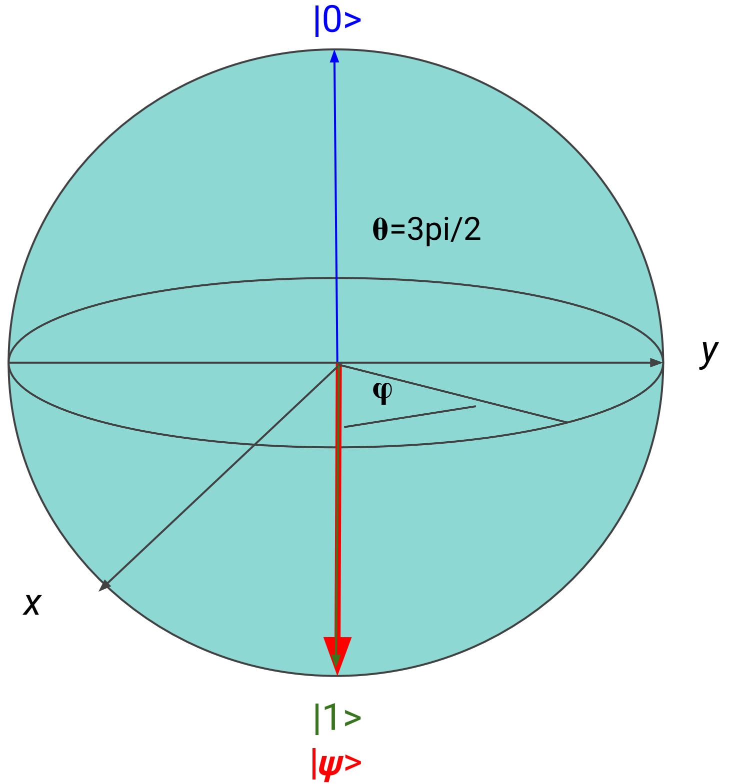
How Can We Control It?



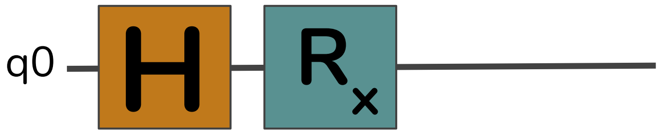
$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$$



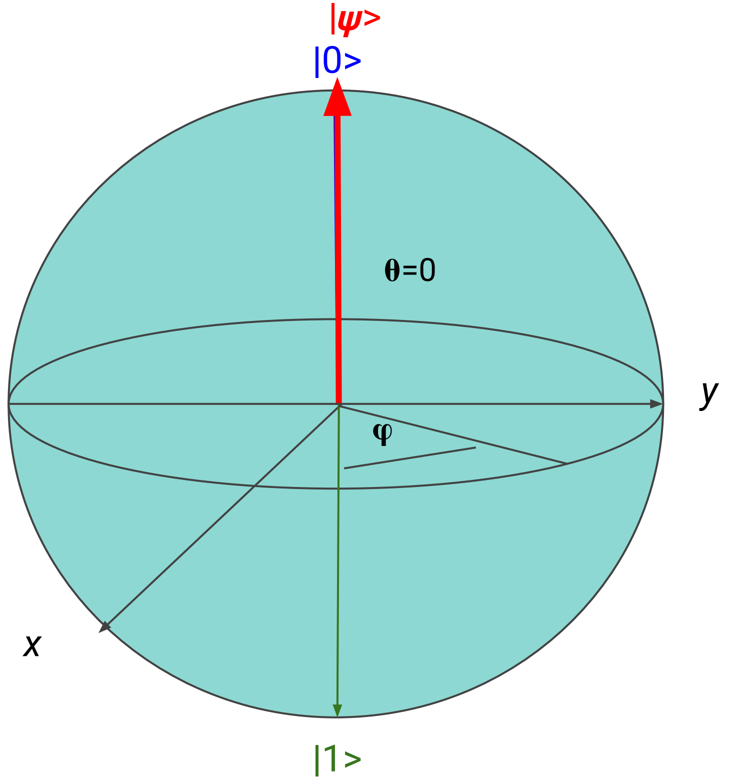
How Can We Control It?



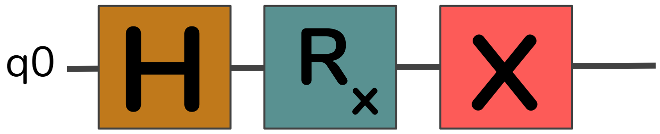
$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$



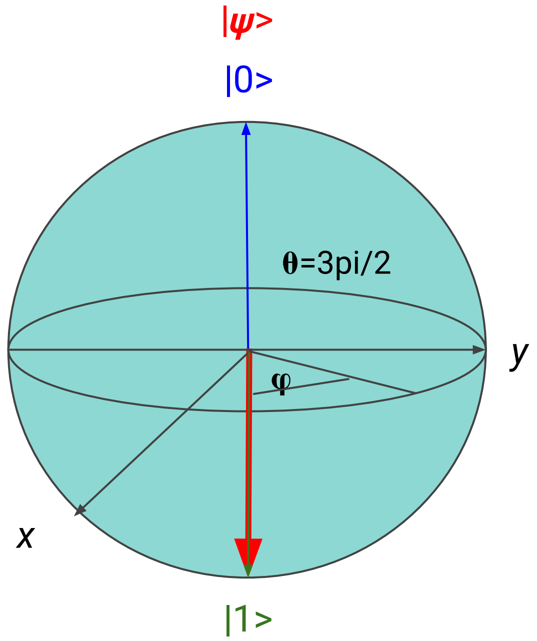
How Can We Control It?



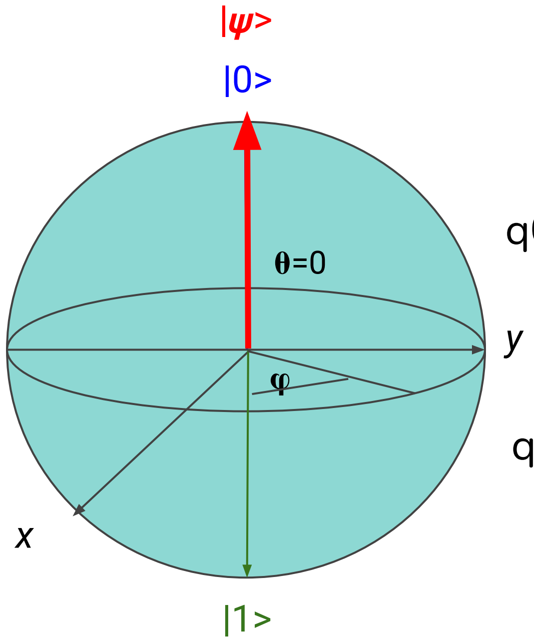
$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$$



How Can We Control It?

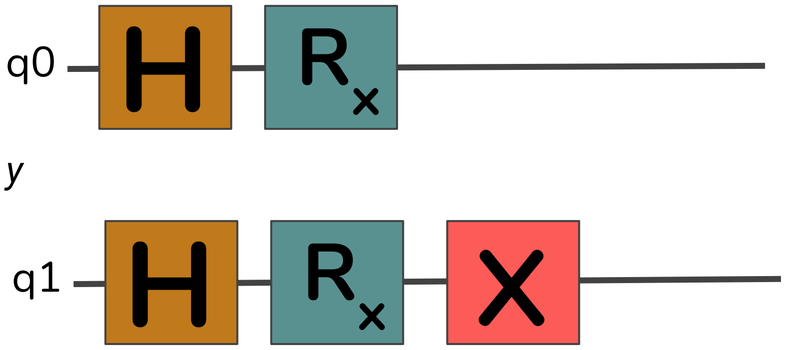


q0

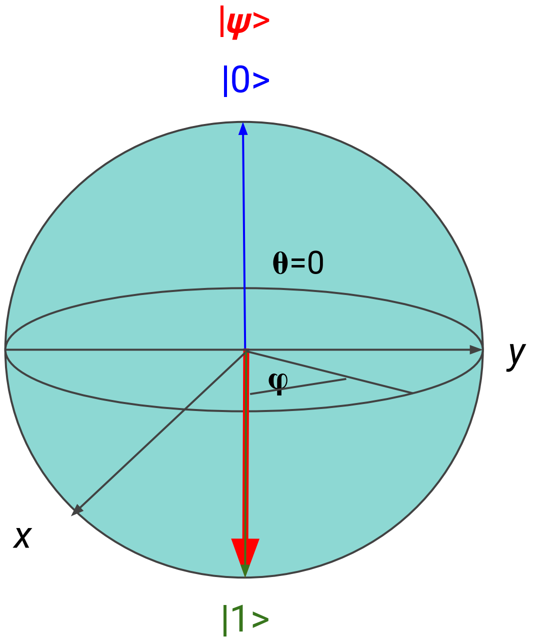


q1

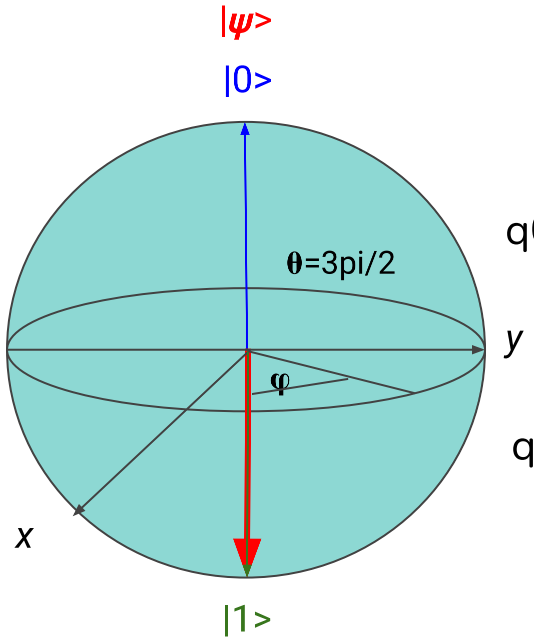
$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$



How Can We Control It?

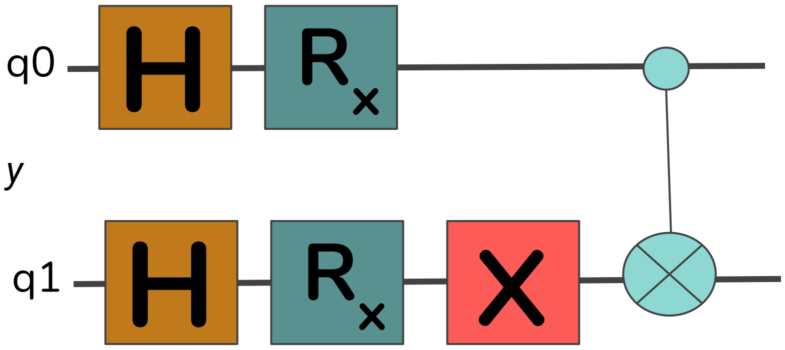


q0

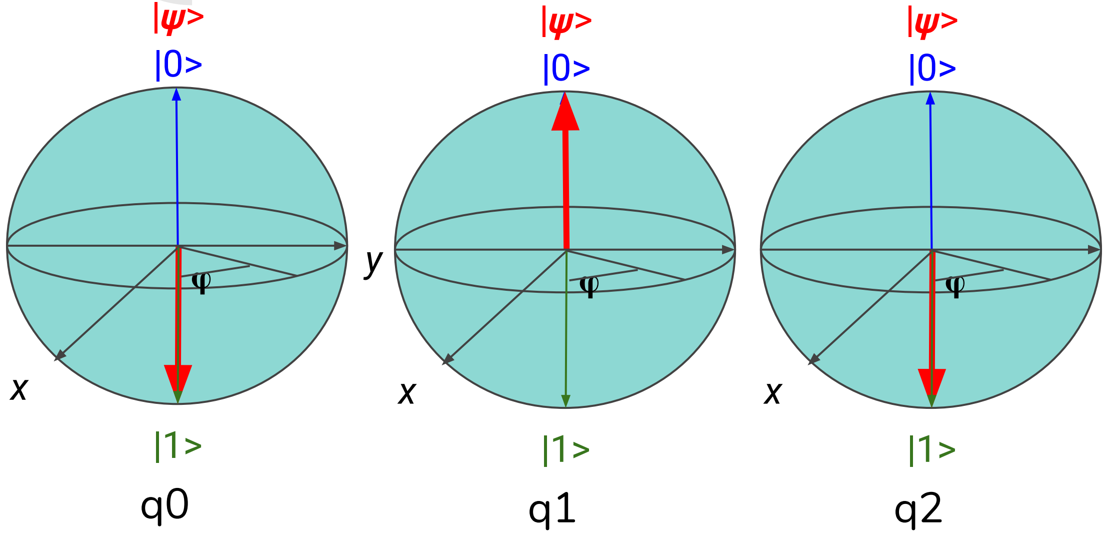


q1

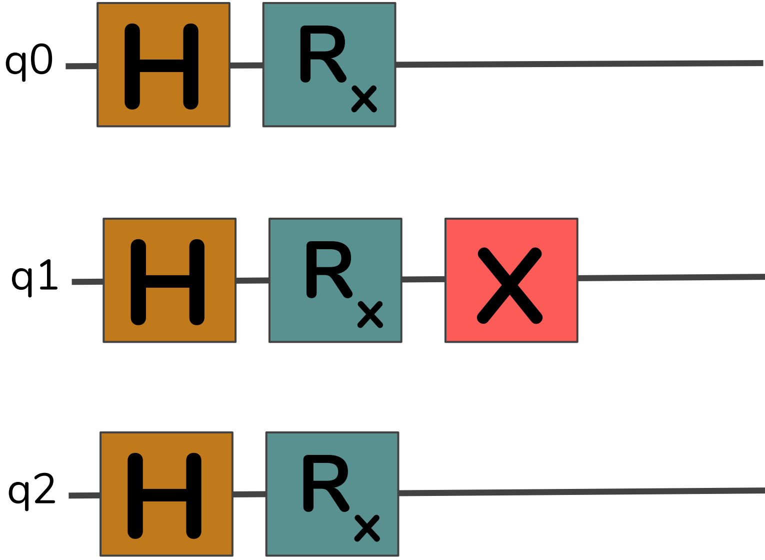
$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle$$



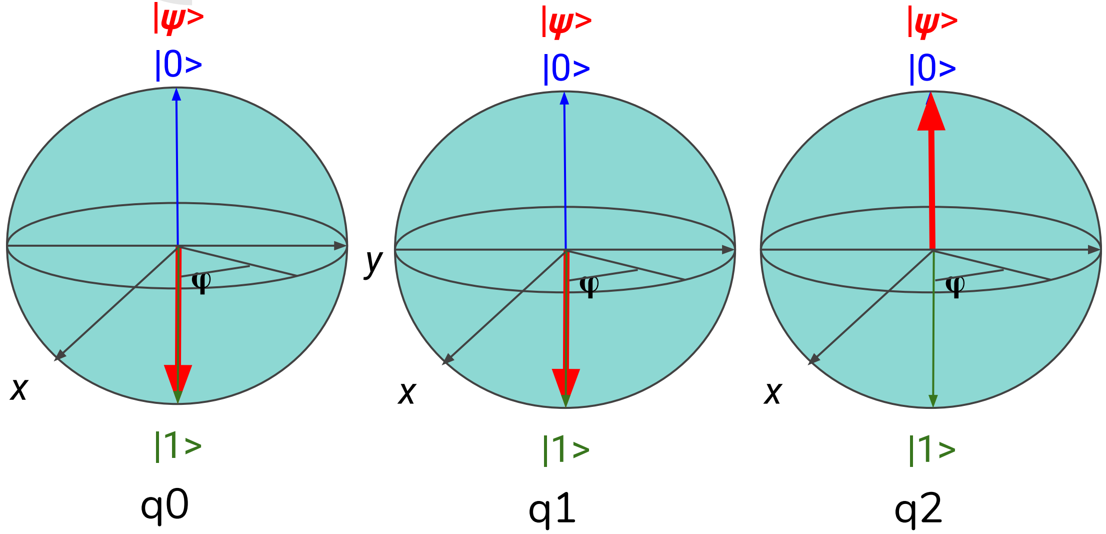
How Can We Control It?



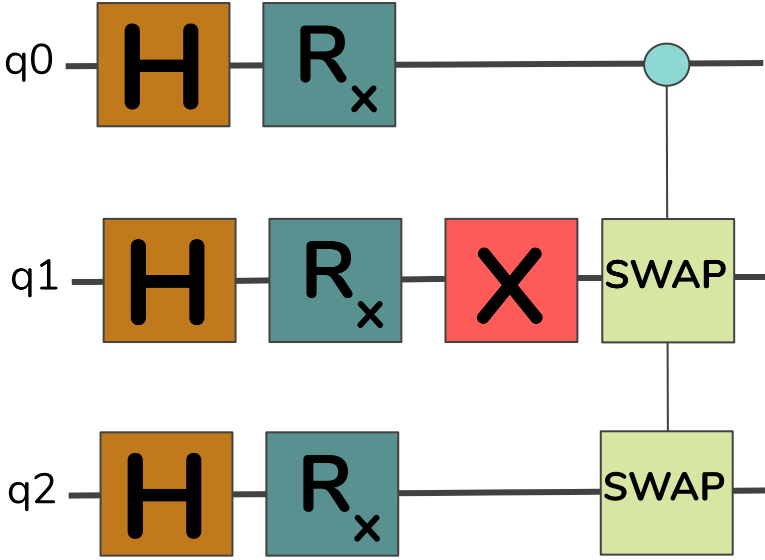
$|\psi> = cos(\theta/2)|0> + e^{i\varphi}sin(\theta/2)|1>$



How Can We Control It?



$$|\psi> = \cos(\theta/2)|0> + e^{i\varphi}\sin(\theta/2)|1>$$



Quantum Computer Requirements

1. Well Defined Qubit System
2. Initialization
3. Universal set of gates
4. Long Coherence Times
5. Measurement Capability

Photonic Quantum Computer

1. Two types of polarized light
   |0> = Vertically polarized
   |1> = Horizontally polarized
2. Polarizer on a coherent laser beam
3. Polarizers, half-quarter wave plates
4. Medium length - milliseconds
5. Photodetectors

How can a quantum computer improve machine learning?

Motivation for Quantum Computing

For every n qubits there are 2^n states at the same time. We can use **fewer qubits** and have **less computationally expensive** algorithms

Shor's Algorithm

- Polynomial Time
- Exponential Speed-up

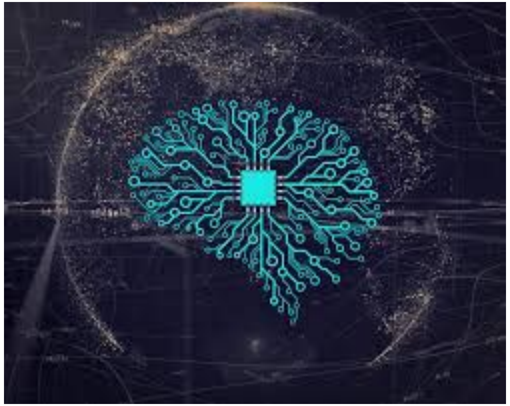
*If performed on a large scale quantum computer, it would break RSA Encryption

Size

Speed

Motivation for Machine Learning

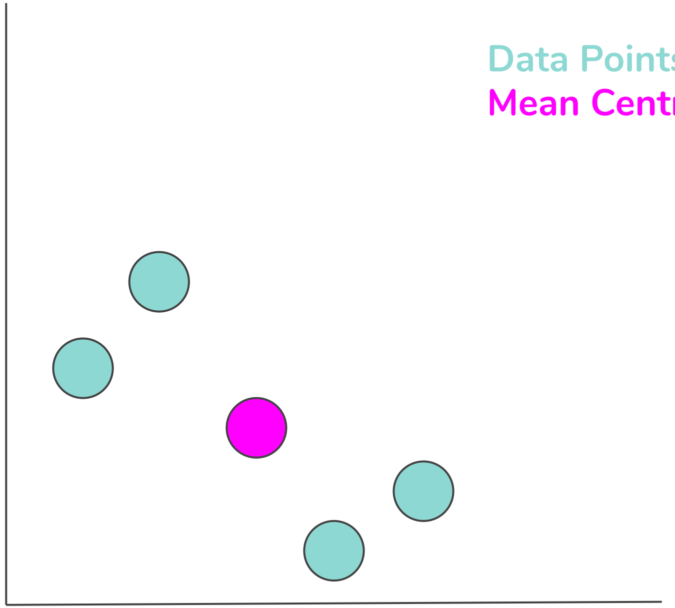
A branch of AI based on the idea that systems can **learn from data** to **update autonomously**



Accuracy

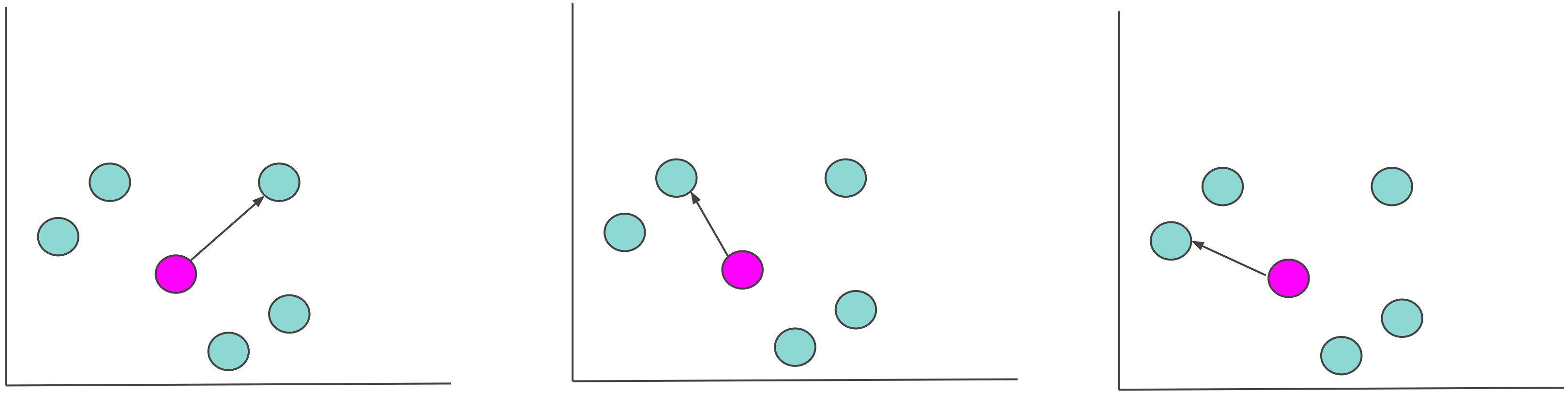
Speed

Updating the Mean Centroid of Data



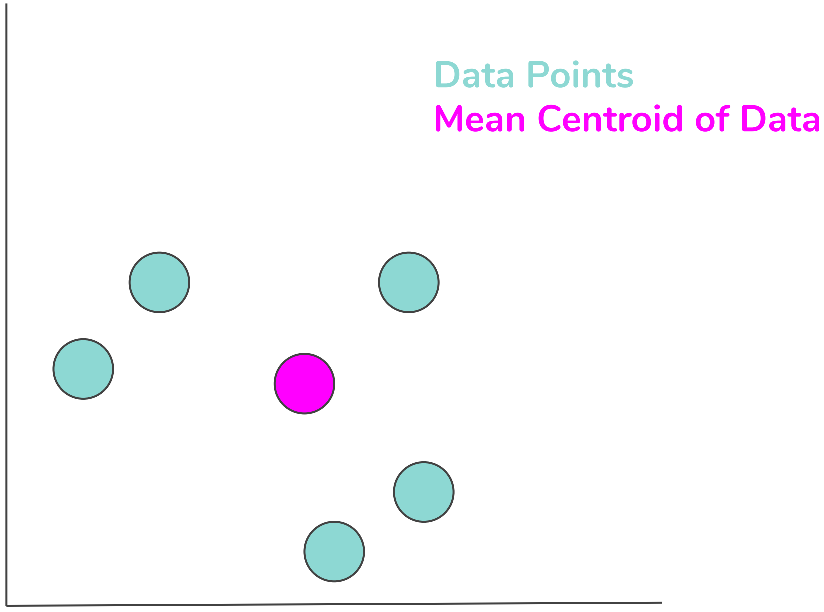
Updating the Mean Centroid of Data

Data Points
Mean Centroid of Data



Euclidean Distance Measurements

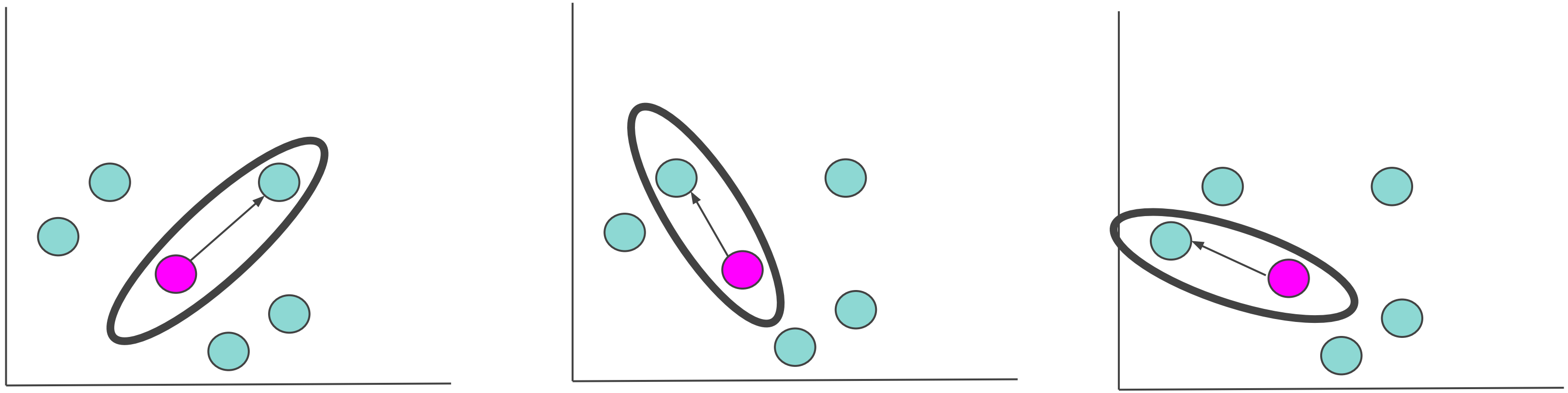
Updating the Mean Centroid of Data



Updated Center

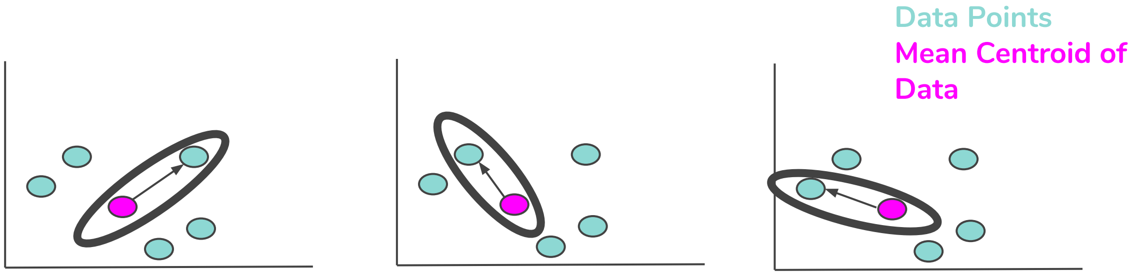
Slipping In the Quantum

Data Points
Mean Centroid of Data



State Overlap Measurements

Slipping In the Quantum



$$|\mathbf{X}^i\rangle = \frac{1}{|\mathbf{X}^i|}\sum_{p=1}^{P} X_p^{(i)}\,|p\rangle$$

We can represent each data point of P features as a quantum state and see how well it overlaps with the state of the centroid.

Operation Steps (P Dimensions)

- 2 Hadamard Gates
- 1 Fredkin Gate

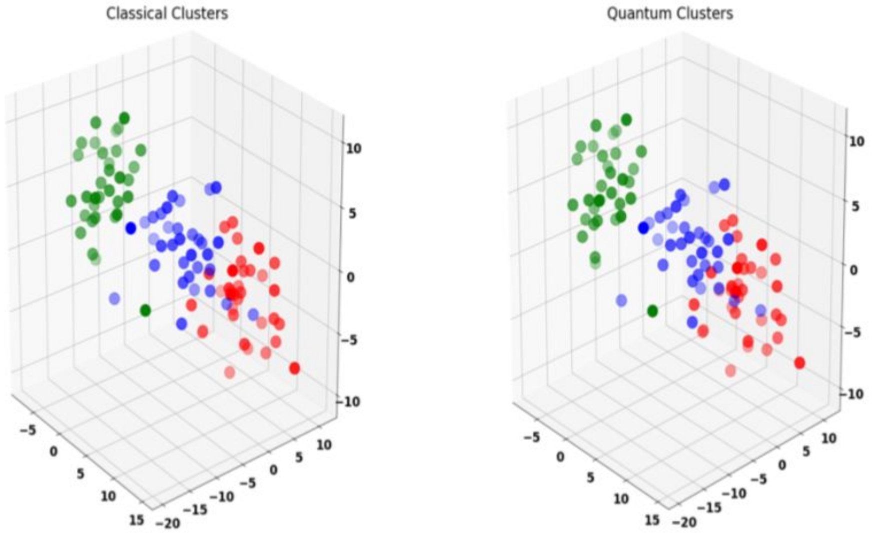
$$\langle \mathbf{X}^i | \mathbf{X}^j \rangle$$

How is this going to contribute to the 2027 projection?

K-Means Clustering on IBM’s Processor

Higher Accuracy

Faster Speed



Can be used for clustering financial datasets for P features

Sarma, A; Chatterjee, R; Gili, K; Yu, T. “Quantum Unsupervised and Supervised Learning on Superconducting Processors, arXiv 1909.04226 [Preprint]. September 10, 2019

K-Means Clustering on IBM's Processor

Higher Accuracy

Faster Speed

Can be used for clustering financial datasets for P features

Algortihm	Accuracy
Classical SVM	96.7%
Quantum SVM by Havlicek et al.	63.3%
Classical K-means	88.7%
Quantum K-means	96.7%

$$O(N^2PK) \rightarrow O(\sqrt{\frac{N}{K}})$$

For Each Iteration

Sarma, A; Chatterjee, R; Gili, K; Yu, T. "Quantum Unsupervised and Supervised Learning on Superconducting Processors, arXiv 1909.04226 [Preprint]. September 10, 2019

Significance of Research Results

Higher Accuracy

Faster Speed

Can be used for clustering financial datasets for P features

1. We can use this method for not only faster, but more accurate machine learning when clustering datasets

2. Practical for industry purposes in the present time

3. Can be extended to datasets of higher order dimensions when larger quantum computers become commercially available

Sarma, A; Chatterjee, R; Gili, K; Yu, T. “Quantum Unsupervised and Supervised Learning on Superconducting Processors, arXiv 1909.04226 [Preprint]. September 10, 2019

Thank You

Supplemental Slides

K-Means Quantum Algorithm

1. Extract each data point and put into the state:

$$|\mathbf{X}^i\rangle = \frac{1}{|\mathbf{X}^i|}\sum_{p=1}^{P} X_p^{(i)}\,|p\rangle$$

2. Set up the following states to perform the gates on

$$|0\rangle \otimes |\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle$$

3. Perform a Hadamard on the 1st qubit and a Fredkin gate on the other two:

$$\begin{aligned}
|\psi\rangle &= (\mathbf{H}\otimes\mathbf{I}\otimes\mathbf{I})\frac{1}{\sqrt{2}}[|0\rangle\otimes|\mathbf{X}^i\rangle\otimes|\mathbf{X}^j\rangle + |1\rangle\otimes|\mathbf{X}^j\rangle\otimes|\mathbf{X}^i\rangle] \\
&= \frac{1}{2}[(|0\rangle+|1\rangle)\otimes|\mathbf{X}^i\rangle\otimes|\mathbf{X}^j\rangle + (|0\rangle-|1\rangle)\otimes|\mathbf{X}^j\rangle\otimes|\mathbf{X}^i\rangle] \\
&= \frac{1}{2}\,|0\rangle\otimes[|\mathbf{X}^i\rangle\otimes|\mathbf{X}^j\rangle + |\mathbf{X}^j\rangle\otimes|\mathbf{X}^i\rangle] + \frac{1}{2}\,|1\rangle\otimes[|\mathbf{X}^i\rangle\otimes|\mathbf{X}^j\rangle - |\mathbf{X}^j\rangle\otimes|\mathbf{X}^i\rangle]
\end{aligned}$$

K-Means Quantum Algorithm

1. Perform another Hadamard gate on the 1st qubit

$$|\psi\rangle = (\mathbf{H} \otimes \mathbf{I} \otimes \mathbf{I}) \frac{1}{\sqrt{2}} [|0\rangle \otimes |\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle + |1\rangle \otimes |\mathbf{X}^j\rangle \otimes |\mathbf{X}^i\rangle]$$
$$= \frac{1}{2} [(|0\rangle + |1\rangle) \otimes |\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle + (|0\rangle - |1\rangle) \otimes |\mathbf{X}^j\rangle \otimes |\mathbf{X}^i\rangle]$$
$$= \frac{1}{2} |0\rangle \otimes [|\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle + |\mathbf{X}^j\rangle \otimes |\mathbf{X}^i\rangle] + \frac{1}{2} |1\rangle \otimes [|\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle - |\mathbf{X}^j\rangle \otimes |\mathbf{X}^i\rangle]$$

2. Measure the auxiliary qubit

$$\langle\psi| \mathbf{M}_0 |\psi\rangle = \frac{1}{4} \{\langle\mathbf{X}^i| \otimes \langle\mathbf{X}^j| + \langle\mathbf{X}^j| \otimes \langle\mathbf{X}^i|\} [|\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle + |\mathbf{X}^j\rangle \otimes |\mathbf{X}^i\rangle]$$
$$= \frac{1}{2} + \frac{1}{4} [(\langle\mathbf{X}^i| \otimes \langle\mathbf{X}^j|)(|\mathbf{X}^j\rangle \otimes |\mathbf{X}^i\rangle) + (\langle\mathbf{X}^j| \otimes \langle\mathbf{X}^i|)(|\mathbf{X}^i\rangle \otimes |\mathbf{X}^j\rangle)]$$
$$= \frac{1}{2} + \frac{1}{4} [\langle\mathbf{X}^i|\mathbf{X}^j\rangle \langle\mathbf{X}^j|\mathbf{X}^i\rangle + \langle\mathbf{X}^j|\mathbf{X}^i\rangle \langle\mathbf{X}^i|\mathbf{X}^j\rangle]$$

3. Use the probability of the auxiliary qubit to retrieve an equation for distance measurement

$$P[|0\rangle] = \langle\psi| \mathbf{M}_0 |\psi\rangle = \frac{1}{2} + \frac{1}{2} |\langle\varphi|\phi\rangle|^2$$
$$= \frac{1}{2} + \frac{1}{4Z} |\mathbf{X}^i - \mathbf{X}^j|^2$$
$$|\mathbf{X}^i - \mathbf{X}^j|^2 = Z\,(4P[|0\rangle] - 2)\,.$$