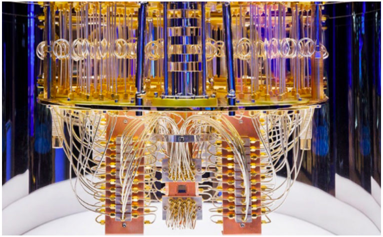
Machine Learning with Quantum Computers

Overview and basic ideas



Davide Pastorello

Department of Information Engineering and Computer Science
University of Trento

Advanced topics in Machine Learning and Optimization

Dec 1, 2022

Premise: The World is quantum...

At microscopic length scales (molecules, atoms, electrons,...):

Newton's laws of dynamics and electromagnetism based on **Maxwell's equations** do not work.

We need **quantum mechanics** to describe physical phenomena.

Quantum things:

- Superposition of states
  (is the electron either "here" or "there"? In both places!).
- Observations alters the state of the observed system.
  (We observe the electron "here" with some probability)
- Entanglement ("spooky action at distance" as Einstein said)

Using physical phenomena to process information is a good and old idea...
What if are we able to use quantum phenomena?

A bit of history...

1981. At the *First Conference of Physics of Computation* (MIT), R. Feynman observed that quantum systems cannot be efficiently simulated by classical computers and proposed a model of quantum computer.

1985. D. Deutsch described the *universal quantum computer* in terms of the *quantum Turing machine*.

1994. P. Shor proposed an algorithm for a quantum computer to factor large integers in polynomial time.

1995. L. Grover presents the *quantum search algorithm* in an unsorted database.

2000s. First prototypes of working quantum computers.

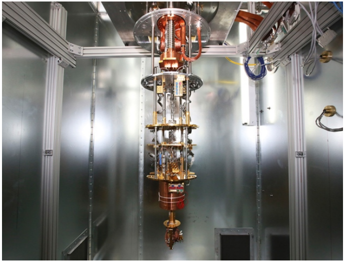
2010s. Quantum computers on the market (D-Wave, IBM Q System 1).

Present days (2022)

- Anyone can run his/her quantum algorithms on available quantum machines (a basic access is usually free).
- Real-life problems are tackled by quantum computers.
- The interest by industry and non academic institutions is dramatically increasing.
- Availability of fault-tolerant, error-corrected, universal quantum computers?
  Not yet, we are in the Noisy Intermediate-Scale Quantum (NISQ) era (few hundred qubits).
- Quantum supremacy? *Not yet.*
- Hot topics: **quantum optimization** and **quantum machine learning**...

Quantum computers exist

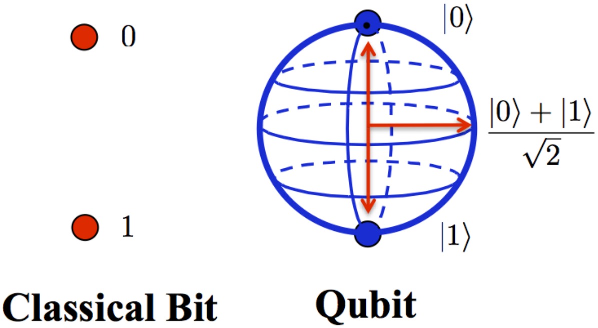
Examples of quantum computers:



D-Wave

IBMQ

Quantum bits

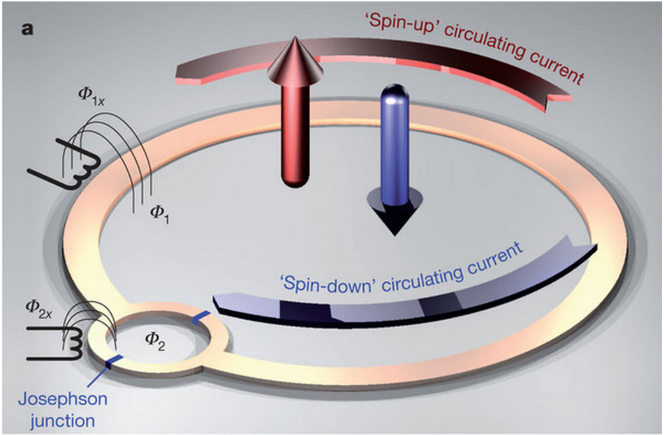


Examples of qubits:

- Particle with spin 1/2;
- Polarized Photons;
- Controlled superconducting circuits.

Quantum bits

Qubit as a superconducting circuit:



Qubits

Quantum state superposition

The state of a qubit is a unit vector in $\mathbb{C}^2$:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \qquad \alpha, \beta \in \mathbb{C}\ , \quad |\alpha|^2 + |\beta|^2 = 1$$

$\{|0\rangle, |1\rangle\}$ computational basis.

A measurement process affects the qubit state

Measurement of a qubit (e.g. we measure the polarization of a photon)

$$\mathbb{P}(0) = |\alpha|^2 \qquad \mathbb{P}(1) = |\beta|^2$$

The qubit state after the measurement is $|i\rangle$ if the outcome is $i = 0, 1$.

Remark:
The vectors $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ with $\theta \in \mathbb{R}$ represent the same *physical state*.

Qubits

Tensor product

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{C}^2 \qquad |\varphi\rangle = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} \in \mathbb{C}^2$$

$$|\psi\rangle \otimes |\varphi\rangle = \begin{pmatrix} \alpha\gamma \\ \alpha\delta \\ \beta\gamma \\ \beta\delta \end{pmatrix} \in \mathbb{C}^4$$

2 qubits (as a composite system) are described in:

$$\mathbb{C}^2 \otimes \mathbb{C}^2 := \mathrm{span}\{|\psi\rangle \otimes |\varphi\rangle \ : \ |\psi\rangle, |\varphi\rangle \in \mathbb{C}^2\} = \mathbb{C}^4.$$

n qubits (as a composite system) are described in:

$$(\mathbb{C}^2)^{\otimes n} = \mathbb{C}^{2^n}.$$

Entangled qubits

Let $|\Psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ be the state of a qubit pair.

$|\Psi\rangle$ is said to be:

- **separable** if it has form $|\Psi\rangle = |\psi\rangle \otimes |\varphi\rangle \equiv |\psi\varphi\rangle$;
- **entangled** otherwise.

Example

Entangled state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Measure the **first qubit**:

- The probability of measuring 0 is $\frac{1}{2}$.
- If the outcome is 0 then the post-measurement state is $|\Psi_0\rangle = |00\rangle$.
- Non-local action on the **second qubit** (quantum correlation).

EPR paradox: inconsistency with QM and Einstein's locality.

A new kind of information

Empirical evidence:
Quantum randomness, state superpositions, entanglement are **physical phenomena** not simply *theoretical interpretations*.

$$\Downarrow$$

Encoding information into qubits (or more general quantum objects) allows completely new kinds of:

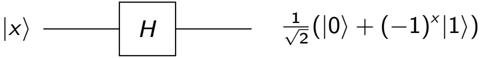
- information storing and processing
- data representation
- telecommunication
- cybersecurity
- ...

In particular we are interested in machine learning!

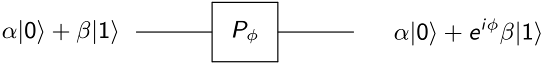
Quantum gates

Let $\{|0\rangle, |1\rangle\}$ be the computational basis

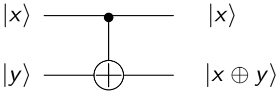
- Hadamard gate:



- Phase gate:



- CNOT gate:



Theorem: $\{H, P_{\pi/4}, CNOT\}$ is a universal set for quantum computation.

Machine learning with quantum computers?

Achievements in Quantum Machine Learning (since 2013)

Algorithm	**Quantum speedup**
K-medians	Quadratic/Exponential
Hierarchical clustering	Quadratic
K-means	Exponential
Principal component analysis	Exponential
Support vector machines	Exponential
Nearest neighbors	Quadratic / Exponential
Neural networks	?

General observation

Quantum architectures enable new paradigms of data representation and information processing.

Current challenge

Devising *quantum learning mechanisms* for available or near-term quantum machines.

A quantum binary classifier

The model

Let $\{(\mathbf{x}_i, y_i)\}_{i=0,\dots,N-1}$ be a training set where:
$x_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$ for any $i = 0, ..., N-1$

$$y(\mathbf{x}) := \operatorname{sgn}\left(\sum_{i=0}^{N-1} y_i \cos(\mathbf{x}_i, \mathbf{x})\right) \qquad \cos(\mathbf{x}, \mathbf{z}) := \frac{\mathbf{x} \cdot \mathbf{z}}{\| \mathbf{x} \| \| \mathbf{z} \|} \qquad \mathbf{x}, \mathbf{z} \in \mathbb{R}^d.$$

Quantum implementation on the IBM ibmq_melbourne

- Quantum superposition of training vectors;
- Test instance in quantum superposition of the two classes;
- Cosine similarities computed by SWAP test.

Classical complexity: $O(Nd)$
Quantum complexity: $O(\log(Nd))$

A quantum binary classifier

Training set stored in a n-qubit register ($n = \log d$).
$\log N$-qubit register, with Hilbert space $\mathcal{H}_{index} \simeq (\mathbb{C}^2)^{\otimes \log N}$, to encode the indexes of training data vectors.

We can construct the state:

$$|X\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle |\mathbf{x}_i\rangle |b_i\rangle \in \mathcal{H}_{index} \otimes \mathcal{H}_n \otimes \mathcal{H}_l,$$

where $\mathcal{H}_l$ is a 1-qubit register encoding the labels with $b_i = \frac{1-y_i}{2} \in \{0, 1\}$.

We can construct also:

$$|\psi_{\mathbf{x}}\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle |\mathbf{x}\rangle |-\rangle \in \mathcal{H}_{index} \otimes \mathcal{H}_n \otimes \mathcal{H}_l,$$

where $|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$.

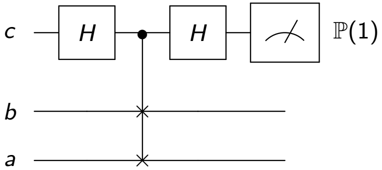
A quantum binary classifier

Add 1 ancillary qubit to the registers and construct:

$$|\Psi\rangle \frac{1}{\sqrt{2}} (|X\rangle|0\rangle + |\psi_{\mathbf{x}}\rangle|1\rangle) \in \mathcal{H}_{index} \otimes \mathcal{H}_n \otimes \mathcal{H}_l \otimes \mathcal{H}_a,$$

that can be retrieved from the QRAM in time $O(log(Nd))$.

Perform the SWAP test:



where qubit b is prepared in $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and qubit c is prepared in $|0\rangle$.
The model:

$$y(\mathbf{x}) := \text{sgn}\left(\sum_{i=0}^{N-1} y_i \cos(\mathbf{x}_i, \mathbf{x})\right),$$

Quantum implementation of the model based on:

$$y(\mathbf{x}) = \text{sgn}\left[1 - 4\,\mathbb{P}(1)\right].$$

A quantum binary classifier

Input: training set $X = \{\mathbf{x}_i, y_i\}_{i=0,\ldots,N-1}$, unclassified instance $\mathbf{x}$.
Result: label y of $\mathbf{x}$.

repeat
    initialize the register $\mathcal{H}_{index} \otimes \mathcal{H}_n \otimes \mathcal{H}_l$ and an ancillary qubit a in the state $|\Psi\rangle$;
    initialize a qubit b in the state $|-\rangle$;
    perform the SWAP test on a and b with control qubit c prepared in $|0\rangle$;
    measure qubit c;
until *desired accuracy on the estimation of* $\mathbb{P}(1)$;
Estimate $\mathbb{P}(1)$ as the relative frequency $\hat{\mathbb{P}}$ of outcome 1;
if $\hat{\mathbb{P}} > 0.25$ **then**
    return $y = -1$
else
    return $y = 1$
end

Overall complexity within an error ϵ in the estimation of $\mathbb{P}(1)$: $O(\epsilon^{-2} \log(Nd))$

Quantum clustering

Qdist is a quantum algorithm based on the SWAP test to calculate Euclidean distance in logarithmic time.

Grover is a quantum search algorithm with quadratic speedup.

Example: K-medians clustering

Input: Data set $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$, number of clusters K
Result: Partition of $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ into K clusters

initialize K centroids $C_1, ..., C_K$ from the elements of the dataset V;
repeat
    foreach $i \leftarrow 1, ..., N$ **do**
        Qdist $(\mathbf{x}_i, C_j)\ \forall j = 1, ..., K$;
        find $\text{argmin}_j \parallel \mathbf{x}_i - C_j \parallel$ with **Grover**;
    end
    construct the cluster $P_j = \{\mathbf{x}_i : C_j \text{ is the nearest centroid}\}$ for all $j = 1, ..., K$;
    foreach $j \leftarrow 1, ..., K$ **do**
        use **Qdist** and **Grover** for centroid calculation;
    end
until *convergence*;
return $P_1, ..., P_K$

Some QML schemes designed for universal quantum computers

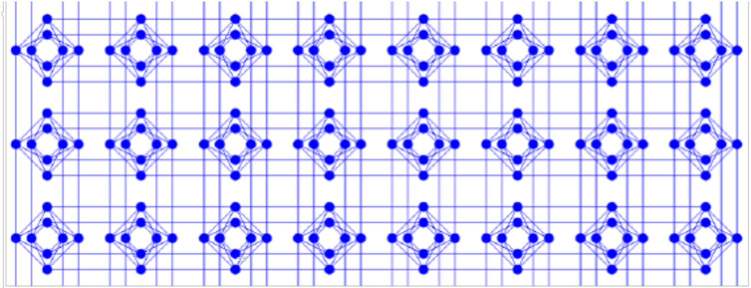
- **Quantum divisive clustering**
  W. Aïmeur et al. *Quantum clustering algorithms* ICML '07: Proceedings of the 24th international conference on Machine learning (2007)
- **Quantum principal component analysis**
  S. Lloyd, et al. *Quantum principal component analysis* Nature Physics 10, 631 (2014)
- **Quantum support vector machine**
  P. Rebentrost et al. *Quantum support vector machine for big data classification* Phys. Rev. Lett. 113, 130503 (2014)
- **Quantum nearest neighbor**
  N. Wiebe et al. *Quantum Algorithms for Nearest-Neighbor Methods for Supervised and Unsupervised Learning* Quantum Information and Computation 15(3,4): 0318- 0358 (2015)
- **Quantum perceptron** M. Schuld *Simulating a perceptron on a quantum computer* Physics Letters A, 379, pp. 660-663 (2015)
- **Quantum meta-learning** M. Wilson et al. *Optimizing quantum heuristic with meta-learning* Quantum Machine Intelligence 3 (2021)

Quantum Annealing

Quantum Annealers

The hardware is a *quantum spin glass*, i.e. a collection of qubits arranged in the vertices of a graph (V, E) where edges represent the interactions between neighbors.

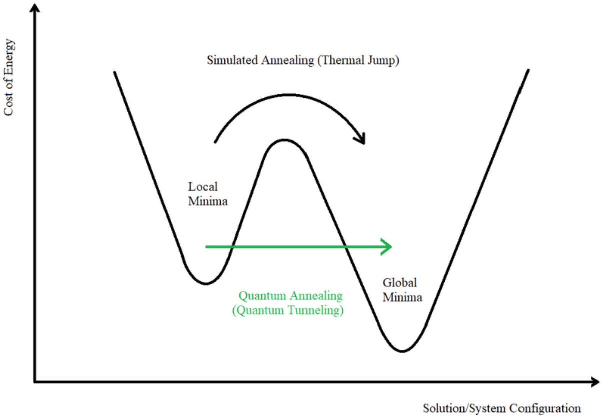
Example: *D-Wave Chimera topology*



Quantum Annealing

Annealing process (annealing time 20μs)

By energy dissipation the quantum system evolves in the **ground state** (the less energetic state) corresponding to the **solution** of a given optimization problem.



Quantum Annealing

Shortcoming

Representing a given problem into the quantum hardware architecture is difficult in general and can destroy the efficiency of the quantum computation.

Proposed solution

Quantum Annealing Learning Search (QALS):

A hybrid quantum-classical algorithm enabling a learning mechanism so that the quantum machine learns the representation of the problem on its own without an expensive classical pre-processing.

References:

-) D.P., E. Blanzieri. **Quantum Annealing Learning Search for solving QUBO problems**. Quantum Information Processing 18: 303 (2019)

-) A. Bonomi, T. De Min, E. Zardini, E. Blanzieri, Valter Cavecchia, D. P. **Quantum annealing learning search implementations** in Quantum Information and Computation, v. 22, n. 3&4 (2022)

ML with quantum annealers

- **Boltzmann machine implementation**
  M.H. Amin *Quantum Boltzmann machine* Phys. Rev. X 8, 021050 (2018)
- **Classification**
  N. T. Nguyen et al. *Image classification using quantum inference on the D-wave 2X*. In: 2018 IEEE International Conference on Rebooting Computing (ICRC), pp. 1-7. IEEE (2018)
- **Clustering**
  V. Kumar et al. *Quantum Annealing for Combinatorial Clustering* Quantum Inf Process (2018) 17: 39
- **Training of a SVM**
  D. Willsch et al. *Support vector machines on the D-wave quantum annealer* Computer Physics Communications 248, 107006 (2020)

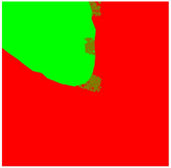
Quantum-inspired ML

Using quantum formalism to devise classical ML algorithms

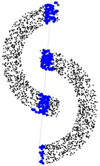
Quantum-inspired classifiers:

- Data encoding into density operators (quantum states);
- Construction of the centroids in the space of quantum states;
- Application of *discrimination of quantum states* to attach a new data instance to the *most similar* centroid.

Example:



Quantum-inspired classifier



SVM with linear kernel

Ref.: R. Leporini and D. P.. *Support vector machines with quantum state discrimination*. Quantum Reports vol. 3, n. 3 (2021)

Some concluding remarks

- Quantum architectures enable new kinds of data processing.
- There is a bottleneck about the efficient data representation into quantum hardware (QRAM are not feasible yet).
- The era of quantum machine learning is just started but it is maybe the most promising road for putting forward the frontiers of quantum computing.
- The quantum-classical hybrid approach is the best way to use the existing quantum machines.
- Providing efficient quantum algorithms is a very difficult task. So we have to devise learning mechanisms of new quantum algorithms towards a "quantum meta-learning"...

In other words, quantum computers must learn how to solve problems on their own...

Thanks for your attention!

Credit: IBM