Sublinear Quantum Algorithms for Training Linear and Kernel-based Classifiers

Tongyang Li, Shouvanik Chakrabarti, Xiaodi Wu

arXiv:1904.02276
ICML 2019
Why Quantum Machine Learning?

- Quantum machine learning is becoming more and more relevant:
  - Theoretical physics has motivated many ML models (Ex. Boltzmann machine, Ising model, Langevin dynamics, etc.)
  - Classical ML techniques can be applied to quantum problems.
  - Quantum computers give speedup for training models.
  - ....
Why Quantum Machine Learning?

- Quantum machine learning is becoming more and more relevant:
  - Theoretical physics has motivated many ML models (Ex. Boltzmann machine, Ising model, Langevin dynamics, etc.)
  - Classical ML techniques can be applied to quantum problems.
  - Quantum computers give speedup for training models.

- Quantum computers are developing fast, having 50-100 qubits now:

  Maryland & IonQ  IBM  Google

Noisy, intermediate-scale quantum computers (NISQ); practical quantum computers to come in 5-10 years.
A promising quantum ML application: **classification**

\[ X^T w \geq \sigma \]

\[ X^T w = 0 \]

\[ X^T w \leq -\sigma \]

Margin = \sigma
Merits of our quantum classifier
Merits of our quantum classifier

- *Near-term implementation*: Highly classical-quantum hybrid with the minimal quantum part; suitable for NISQ computers.
Merits of our quantum classifier

- **Near-term implementation**: Highly classical-quantum hybrid with the minimal quantum part; suitable for NISQ computers.

- **Composability**: Purely classical output, suitable for end-to-end machine learning applications.
Merits of our quantum classifier

- **Near-term implementation**: Highly classical-quantum hybrid with the minimal quantum part; suitable for NISQ computers.

- **Composability**: Purely classical output, suitable for end-to-end machine learning applications.

- **Generality**: The classifier can be kernelized.
Main Results

Given \( n \) data points with dimension \( d \), our quantum algorithms train classifiers for the following problems with complexity \( \tilde{O}(\sqrt{n} + \sqrt{d}) \):

- **Linear classification:** \( X^\top w \)
- **Minimum enclosing ball:** \( \|w - X\|^2 \)
- **\( \ell_2 \)-margin SVM:** \( (X^\top w)^2 \)
- **Kernel-based classification:** \( \langle \Psi(X), w \rangle \), where \( \Psi = \) polynomial kernel or Gaussian kernel.

The optimal classical algorithm runs in \( \tilde{\Theta}(n + d) \) (Clarkson et al. ’12).
Highlights of Our Quantum Algorithm

Standard quantum input: coherently access the coordinates of data, like a Schrödinger's cat:

Speed-up: The classical $\tilde{O}(n + d)$ optimal algorithm by Clarkson et al. uses a primal-dual approach:

$\begin{align*}
\text{Primal:} & \quad O(n) \text{ by multiplicative weight updates.} \\
\text{Dual:} & \quad O(d) \text{ by online gradient descent.}
\end{align*}$

Quantum: quadratic speed-ups for both the primal and dual.

Optimality: We prove quantum lower bounds $\Omega(\sqrt{n} + \sqrt{d})$, meaning that our quantum algorithms are optimal.
Highlights of Our Quantum Algorithm

- **Standard quantum input:** *coherently* access the coordinates of data, like a Schrödinger’s cat: $\frac{1}{\sqrt{2}}|\text{ψ}\rangle + \frac{1}{\sqrt{2}}|\text{φ}\rangle$
Highlights of Our Quantum Algorithm

- **Standard quantum input:** coherently access the coordinates of data, like a Schrödinger’s cat: \[ \frac{1}{\sqrt{2}} |\psi\rangle + \frac{1}{\sqrt{2}} |\phi\rangle \]

- **Speed-up:** The classical \( \tilde{\Theta}(n + d) \) optimal algorithm by Clarkson et al. uses a primal-dual approach:
  - Primal: \( O(n) \) by multiplicative weight updates.
  - Dual: \( O(d) \) by online gradient descent.

Quantum: quadratic speed-ups for both the primal and dual.
Highlights of Our Quantum Algorithm

- **Standard quantum input:** coherently access the coordinates of data, like a Schrödinger’s cat: \[ \frac{1}{\sqrt{2}} |\text{cat}\rangle + \frac{1}{\sqrt{2}} |\text{mouse}\rangle \]

- **Speed-up:** The classical $\tilde{\Theta}(n + d)$ optimal algorithm by Clarkson et al. uses a primal-dual approach:
  - Primal: $O(n)$ by multiplicative weight updates.
  - Dual: $O(d)$ by online gradient descent.

  Quantum: quadratic speed-ups for both the primal and dual.

- **Optimality:** We prove quantum lower bounds $\Omega(\sqrt{n} + \sqrt{d})$, meaning that our quantum algorithms are optimal.
Thank you!

More info: #171 at poster session

arXiv:1904.02276