

量子コンピュータを用いた機械学習ツールの実装と改良

—量子ウォークによるグラフデータ分類—

1. 背景

現在量子コンピューティング技術は大きな注目を集め、世界中で日々研究開発が進められている。化学や金融などの産業応用へ直結するような分野への量子コンピューティング技術の応用があるが、中でも最も注目されている分野の一つに機械学習がある。現在、機械学習は身近な画像認識から化学物質の構造検出まで、様々な場面で活用されており、既存の手法で驚くべき成果が得られているものもあれば、一方で新たなパラダイムを必要とするような問題も存在している。そこで近年様々な研究によって、既存の古典機械学習手法を凌駕する量子的なアプローチの模索が行われており、量子ハードウェアの進歩とともに量子ソフトウェアの進歩によるパラダイムシフトが期待されている。

一方で、近年の機械学習開発ブームは様々なライブラリやパッケージ、プラットフォームが登場し、開発環境が著しく向上したことにも起因している。こうしたツールや学びを提供する潮流は、開発者の参入障壁を下げ、機械学習コミュニティを広げることに貢献した。しかしながら、現在の量子機械学習開発は既存の古典機械学習ツールこそ利用できるものの、一部の研究者に限られている実態がある。

2. 目的

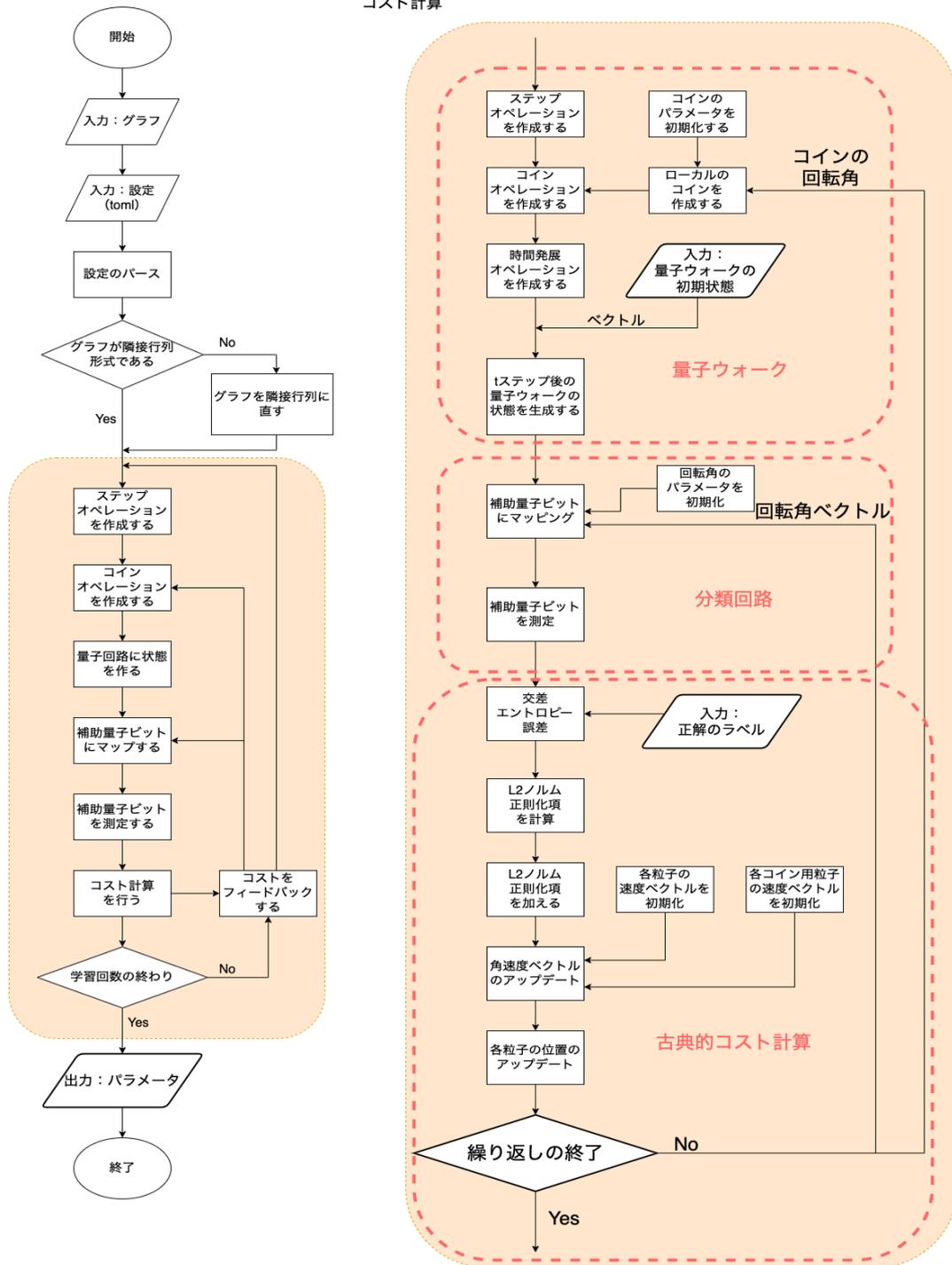
本プロジェクトの目的は、量子機械学習の開発状況の発展を目指すことにある。具体的には技術面と社会面の二つの側面に分けてその実現を目指す。

技術的には新たな量子機械学習のモデルを提案、および開発実装し、実データを用いてそれらの評価実験を行い、その可能性を示す。具体的には IBM の量子計算機システム IBMQ と Qiskit などのモジュールを用いて二つの量子機械学習アルゴリズムの開発を目指す。一つは、既存研究[[SLQ](#)]で提案されている量子変分法によるモデルを改良実装しアヤメや MNIST などの実データセットへの応用すること。二つ目には、量子ウォークという量子アルゴリズムの一種を応用した新しい機械学習のモデルを確立し、ソフトウェアとして実装、評価実験を行い、これに対応する既存の古典アルゴリズムとの比較を行う。

社会的には、こうした新たな手法を開発するだけでなく量子機械学習エンジニアのためのツールや教育環境を提供することを目指す。具体的には、量子コンピューティングの基礎から本プロジェクトで開発予定の先端アルゴリズムまで学ぶことのできるドキュメントや、開発した成果物を何らかの形で体験することのできるツールを提供し、量子機械学習のインパクトや現状の開発状況を正しく理解することのできるWEBプラットフォームの開発を行う。

3. ソフトウェア開発内容

コスト計算



● 構成図

● 解決する課題

このソフトウェアが解決する課題は、グラフのデータというものの処理に、古典コンピュータでは多くのリソースを消費する上、実現している精度に、現状は限界があるということである。

● 動作環境

今回の成果物は基本的にはどのようなプラットフォーム（OS）の上でも動くが、想定している動作環境としては多くのCPUを持ったサーバー環境である。またPythonとPythonのパッケージマネージャーがあれば動作するような環境となっている。

● 機能等

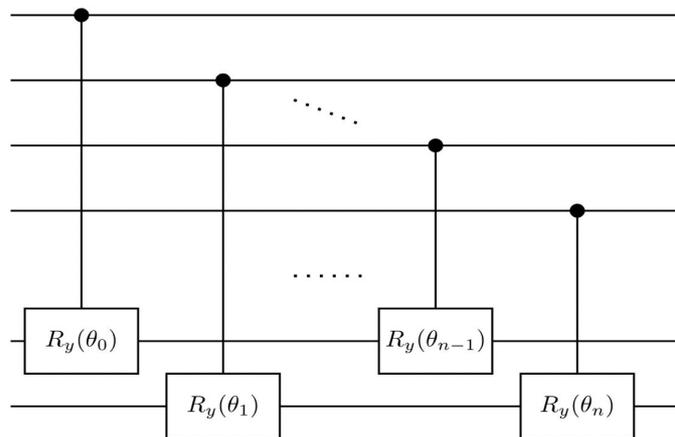
機能としては以下のようなものがある。

- 量子ウォークを用いて、グラフから特徴量を抽出する機能

$$|\psi_t\rangle = (SC)^t |\psi_0\rangle$$

このような式で表される量子ウォークを用いている。具体的には、初期状態にステップオペレータ(S)というものと、コインオペレータ(C)というものの内積を時間発展オペレータを構成し、その時間発展オペレータを時間発展を行うステップ数(t)だけ初期状態に作用させることで、t時間後の量子ウォークの状態というのを見ることが出来る。また、このコインオペレータはどのようなユニタリでも実装可能であることから、パラメトリックなユニタリを利用することで、コストに応じてコインの種類を変えている。

- その量子ウォーク後の状態を用いて、分類を行う機能（量子回路）



上のような回路を用いて、分類を行う。制御側の量子ビット側にt時間後の量子ウォークの状態があり、被制御側に補助量子ビットを挿入し、ある一定の角度を持って制御回転ゲートを作用している。この回転角はランダムに初期化し、以降コストをもとに調整されていく。

- 測定を行い、エラーを計算する機能

制御回転ゲートを作用させたのち、補助量子ビットを測定し、コストの計算を行う。コスト関数としては、

$$E = \frac{1}{N} \sum_{i=0}^N \left(- \sum_k y_{ik} \ln(\hat{y}_{ik}) + \alpha \sum_j |\theta_j|^2 \right)$$

というような関数を用いている。第一項では実験的に得た各クラスに対応した確率分布とワンホットエンコーディングしたラベルとの交差エントロピー誤差をとり、第二項では回転角の二乗和にある小さな定数アルファをかけた正則化項を加えている。

- そのエラーをもとに、コインと分類回路のパラメータを更新する機能
これらのエラーをもとに、コインと分類回路のパラメータを更新する。今回は最適化手法として PSO（粒子群最適化）を使用している。

4. 新規性・優位性

今回開発したソフトウェアの新規性として、コインのパラメータを学習的に生成するという点と、測定を介さずに分類のスキームへ流すことができるという点であると考えられる。現在様々なモデルが提案されているが、今回実験を行ったデータセットの中では、35 個程度のモデルが提案されているが、最大精度では古典最良のモデルと同等、平均精度では 10 個ほどのモデルに有効性を示すことができた。

5. 期待されるユーザー価値と社会へのインパクト

今回の開発したソフトウェアをユーザーが利用することによって得られる価値として、量子ウォークの新たな応用先や、量子コンピュータの新たな応用先という物を探す一つのヒントになりうるのではないかと考えている。グラフデータという物の産業利用は、化学分野やネットワーク分野等で非常に進んできているが、まだまだ発展途上の分野であることは間違いない。そこに対して量子コンピュータを応用する先がまだまだ多くあるのではないかと考えられる。

6. 氏名(所属)

西尾 真(慶應義塾大学 総合政策学部)

佐藤綾祐(慶應義塾大学 環境情報学部)

大倉康寛(慶應義塾大学 総合政策学部)

(参考)関連 URL

我々のプロジェクトページ

<https://qwqmlf.github.io/QuantumFrontier/>

ソースコード

<https://github.com/qwqmlf/qwgc>

<https://github.com/qwqmlf/QC4QWG>