

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ ИМ. М.В.ЛОМОНОСОВА

КУРСОВАЯ РАБОТА СТУДЕНТА 317 ГРУППЫ

ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И
КИБЕРНЕТИКИ
КАФЕДРА МАТЕМАТИЧЕСКИХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ

Методы квантового машинного обучения

Автор:

КОРМАКОВ ГЕОРГИЙ ВЛАДИМИРОВИЧ, СТУДЕНТ 3 КУРСА

Научный руководитель:

ГУРОВ СЕРГЕЙ ИСАЕВИЧ, Д.Ф.-М.Н., ДОЦЕНТ ФАКУЛЬТЕТА ВМК МГУ



May 20, 2019

Contents

1	Введение	3
2	Введение в квантовую теорию	4
3	Использование квантовых вычислений в машинном обучении	5
3.1	Квантовый алгоритм отнесения к ближайшему множеству (возможное применение к KNN)	5
3.2	Адиабатические вычисления	6
3.3	Квантовые нейронные сети	6
3.3.1	Модель перцептрона Алтайского	6
3.3.2	Нейронные сети на клеточных автоматах (CNN)	7
4	Выводы:	8
	Список литературы	8

Abstract

Физические законы, лежащие в основе квантовых вычислений позволяют ускорить, например, процессы вычислений некоторых алгоритмов (алгоритм Шора [11], алгоритм Гровера [4]), и растущее число теоретических результатов показывают, что вычислительное устройство, чьи компоненты напрямую управляются квантовой физикой, могут быть значительно мощнее, чем его классический аналог. Однако построение квантового компьютера на практике сталкивается со многими препятствиями, например, изоляцией системы от внешних возмущений. В связи с этим, пока мы продолжаем лишь рассматривать возможность применения квантовых алгоритмов. Одной из областей применения можно рассматривать машинное обучение. Целью данной работы является рассмотрение возможного применения квантовых эффектов для улучшения механизмов работы методов машинного обучения.

1 Введение

Машинное обучение относится к области информатики, в которой закономерности получены ("изучены") из данных с целью осмысления ранее неизвестных входных данных. Это осмысление обычно включает в себя огромные объемы ранее собранных данных вида ввод-вывод, и алгоритмы машинного обучения должны быть эффективными для того, чтобы иметь дело с так называемыми большими данными. Поскольку объем глобально хранимых данных растет примерно на 20% каждый год (в настоящее время порядка нескольких сотен экзабайт [6]), идёт активный поиск инновационных подходов в машинном обучении.

Перспективная идея, которая в настоящее время исследуется в академических кругах и лабораториях ведущих IT-компаний, использует потенциал квантовых вычислений для оптимизации классических алгоритмов машинного обучения. В последние десятилетия физики уже продемонстрировали впечатляющие мощности квантовых систем обработки информации. В отличие от обычных компьютеров, построенных на физическом представлении '0' и '1', квантовые компьютеры могут использовать кубит – суперпозицию двух квантовых состояний (например, закодированный в двух различных энергетических уровнях атома). Но законы квантовой механики также ограничивают наш доступ к информации, хранящейся в квантовых системах. Технологическая реализация квантовых вычислений появляется [3], и многие считают, что это только вопрос времени, пока многочисленные теоретические результаты смогут быть протестированы на реальных машинах.

В связи с этим, новые исследования в области квантового машинного обучения могут предложить потенциал развития будущих путей интеллектуального анализа данных. В ряде недавних научных работ исследуется идея использования преимуществ квантовой вычислительной техники для улучшения алгоритмов машинного обучения. Например, были приложены определенные усилия в развитие квантовых версий искусственных нейронных сетей (которые широко используются в машинном обучении), но они часто основаны на биологических предположениях и крупного прорыва пока не достигнуто.

Некоторые авторы пытаются разработать целые квантовые алгоритмы, которые решают задачи распознавания образов [13]. Другие предлагают просто запускать подпрограммы классических алгоритмов машинного обучения на квантовом компьютере, надеясь получить ускорение [8]. Интересный подход – адиабатическая квантовая машина обучения, которая кажется особенно подходящей для некоторых классов проблем оптимизации [9]. Стохастический модели, такие как байесовская теория решений или скрытые Марковские модели – элегантная реализация языка открытых квантовых систем [5].

Данная работа даёт обзор развивающейся области квантового машинного обучения. Кратко рассмотрены применения квантовых алгоритмов в классических методах машинного обучения (а именно, в алгоритме класетризации, KNN, K-Means, нейронных сетях и также рассмотрена реализация нейронных сетей с использованием квантового клеточного автомата).

2 Введение в квантовую теорию

Целью данного раздела не является рассмотрение всех постулатов квантовой механики или её законов. Однако считаю необходимым ввести некоторые, свойственные для квантовой механики, обозначения и термины.

Итак, любая изолированная физическая система – это *пространство состояний* (это пространство является Гильбертовым). Оно описывается своим *вектором состояния*. Всюду далее будем обозначать вектор-столбец a как $|a\rangle$, а вектор-строку a – как $\langle a|$. Тогда, например, чтобы получить матрицу, необходимо записать $|a\rangle\langle a|$, а скалярное произведение обозначают $\langle a|a\rangle \equiv \langle a|a\rangle$. Простейшей квантовой системой является *кубит* – суперпозиция двух состояний, обозначаемых $|0\rangle$ и $|1\rangle$. Тогда вектор состояния кубита можно записать как $|\psi\rangle = a|0\rangle + b|1\rangle$, где a и b – комплексные числа. Значения чисел a и b часто называют *амплитудами*, и их модули в квадрате – это вероятности оказаться в состояниях $|0\rangle$ и $|1\rangle$ соответственно. Тогда квантовая динамика влечёт за собой следующее условие: $|a|^2 + |b|^2 = 1$. Таким образом, получаем, что *любое преобразование квантового состояния в другое должно быть унитарным*.

Отсюда, в частности, вытекает, что изменения в квантовой системе описываются уравнением Шрёдингера: $i\hbar |\dot{\psi}(t)\rangle = H(t) |\psi(t)\rangle$, где $\hbar$ – константа Планка, H – Эрмитов оператор, зависящий от времени, и известный как Гамильтониан квантовой системы.

Также необходимо рассмотреть термин *измерения* – множества $\{M_m\}$ операторов измерения, индексы которых относятся к результатам измерений. Если система была в состоянии $|\psi\rangle$ непосредственно до момента измерения, то вероятность наступления результата m равна: $p(m) = \langle \psi | M_m^* M_m | \psi \rangle$, а состояние, в котором окажется система после измерения равно: $\frac{M_m |\psi\rangle}{\sqrt{\langle \psi | M_m^* M_m | \psi \rangle}}$.

Композиция двух систем составляется (описывается) тензорным произведением векторов-состояний систем. К примеру, если система состоит из двух кубитов, то её вектор состояния будет описываться уже как $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$, при этом $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. При этом, уже начиная с суперпозиции двух состояний, существуют так называемые запутанные состояния – состояния, которые не представимы в виде тензорного произведения (например, $|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$).

Итак, для чего же рассматривались все эти определения? Дело в том, что квантовое вычисление строится по следующему порядку: инициализируется квантовое состояние, происходят унитарные преобразования этого состояния на некоторых кубитах и фиксируются измерения интересующего набора кубитов.

Преимущество квантового вычисления вытекает из фактов о композиции состояний. Ведь если система состоит из N кубитов, то вектор состояния порождает 2^N возможных состояний, тогда в результатах измерений можно получить одно из этих состояний с определённой вероятностью.

3 Использование квантовых вычислений в машинном обучении

Сет Ллойд предложил (см. [8]) способ кодирования классической информации в вектор состояния квантовой системы. Это представление основывается на простейшем уравнении: $\langle x|x \rangle = |\vec{x}|^{-1}\vec{x}^2$. Откуда получаем $|x\rangle = |\vec{x}|^{-1/2}\vec{x}$

3.1 Квантовый алгоритм отнесения к ближайшему множеству (возможное применение к KNN)

Перед тем, как рассмотреть квантовый алгоритм кластеризации, рассмотрим решение проблемы об эквивалентности двух квантовых состояний. Итак, задача - выяснить, являются ли два состояния $|\psi\rangle$ и $|\phi\rangle$ эквивалентными. Решение описывается следующими шагами:

1. Добавляем состояние $|0\rangle$
2. Применяем оператор Адамара к этому дополнительному состоянию:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

3. Осуществляем операцию контролируемого обмена (CSWAP)
4. Применяем оператор Адамара
5. Проводим измерение дополнительного состояния. Если $|\psi\rangle = |\phi\rangle$, то состояние $|0\rangle$ возникает вероятностью 1

Схематически этот процесс изобразим в виде следующей схемы:

$$\begin{aligned} |0\rangle |\phi\rangle |\psi\rangle &\xrightarrow{H} \frac{|0\rangle + |1\rangle}{\sqrt{2}} |\phi\rangle |\psi\rangle \xrightarrow{CSWAP} \frac{|0\rangle |\phi\rangle |\psi\rangle + |1\rangle |\psi\rangle |\phi\rangle}{\sqrt{2}} \xrightarrow{H} \\ &\frac{|0\rangle (|\phi\rangle |\psi\rangle + |\psi\rangle |\phi\rangle) + |1\rangle (|\phi\rangle |\psi\rangle - |\psi\rangle |\phi\rangle)}{2} \end{aligned}$$

Схожим образом решается и задача отнесения вектора u к одному из множеств V и W , в каждом из которых имеется по N представителей. Общий метод обределяет, к центру какого класса ближе этот объект. А квантовый алгоритм действует по следующей цепочке:

1. Создаём новую переменную с $N + 1$ состоянием
2. Конструируем, как и в задаче сравнения, состояние

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle |u\rangle) + \frac{1}{\sqrt{N}} \sum_{j=1}^N |j\rangle |v_j\rangle$$

3. Используя задачу сравнения состояний, проверяем, находится ли система в состоянии

$$|\phi\rangle = \frac{1}{\sqrt{Z}}(|u\rangle|0\rangle - \frac{1}{\sqrt{N}} \sum_{j=1}^N |v_j\rangle|j\rangle)$$

, где $Z = |u|^2 + \frac{1}{N} \sum_{j=1}^N |v_j|^2$

4. Тогда вероятность успеха эквивалентности двух состояний (измерения дополнительного состояния) равна расстоянию от объекта до центра, поделённого на Z . Таким образом получают расстояние до центра кластера.

3.2 Адиабатические вычисления

Уравнение Шрёдингера показывает изменения в системе, а Гамильтониан описывает собственные состояния системы своими собственными векторами, а собственные значения – энергии этих состояний. Состояние с наименьшей энергией описывает основное положение. Следуя следующему закону: Физическая система, находящаяся в своём основном положении (с минимальной энергией) стремится сохранить своё положение, можно построить алгоритм минимизации функции на основе адиабатических вычислений:

1. Определяем состояние Гамильтониана H_0 с легко вычисляемыми собственными значениями и устанавливаем систему в это положение
2. Задаём конечное значение Гамильтониана как $H_f = \sum_{z \in \{0,1\}^n} f(z) |z\rangle \langle z|$
3. Преобразуем H_0 к H_f
4. Финальным состоянием системы будет являться минимизатор функции f

Используя этот алгоритм можно получить квантовый алгоритм кластеризации, k -центроид за счёт минимизации функций расстояния, получая состояния минимальности (см. [7])

3.3 Квантовые нейронные сети

За последние 15-20 лет были предложены различные аналоги нейронной сети в квантовом вычислении ([1], [2], [14], [12], [10]).

3.3.1 Модель перцептрона Алтайского

Если рассматривать значения перцептрона как состояние квантовой системы, то выход одного перцептрона может быть получен по следующей формуле: $|y\rangle = \hat{F} \sum_{j=1}^N N \hat{w}_j |x_j\rangle$ с обучающим правилом весов $\hat{w}_j(t+1) = \hat{w}_j(t) + \eta(|d\rangle - |y(t)\rangle) \langle x_j|$

Основываясь на этой модели был получен алгоритм обучения весов и получения выхода перцептрона. Для его понимания потребуется знание дополнительно о псевдообратном преобразовании (Мура-Пенроуза) для вектора:

$$x^\dagger = \begin{cases} 0^T, & \text{если } x \equiv 0; \\ \frac{x^*}{x^* x}, & \text{иначе.} \end{cases}$$

Тогда, взяв $\hat{F} = I$ в представлении Алтайского, можно получить аналитический вид решения для $\hat{w}_j = |y_j\rangle \otimes |x_j^\dagger\rangle$, т.е. тензорное произведение. После этого, просуммировав по всем j : $\hat{w} = \sum_{j=1}^N \hat{w}_j$ получим матрицу весов, однако, она не всегда будет унитарной (как минимум, не всегда квадратной). Если она получилась унитарной, то мы можем вычислить $|y_{output}\rangle$, используя вспомогательную функцию (аналог функции активации в классических нейронных сетях):

$$M(x) = \begin{cases} 1, & \text{если } x > 0.5; \\ 0, & \text{иначе.} \end{cases}$$

Тогда, при квадратной матрице $\hat{w}$, получим $|y_{output}\rangle = M(\hat{w}|x\rangle)$

Если $\hat{w}$ получилась не унитарной, то, при помощи SVD, $U\Sigma V^T = \hat{F}\Sigma_{new}w_{new}$, где Σ_{new} получается заменой сингулярных чисел на 1. И, в этом случае, $|y_{output}\rangle = M(\hat{F}\Sigma_{new}w_{new}|x\rangle)$

3.3.2 Нейронные сети на клеточных автоматах (CNN)

Основная идея клеточных автоматов является представление кубита в виде массива, клетки, которая может находиться в одном из состояний на непрерывном отрезке от -1 до 1 (Figure1):

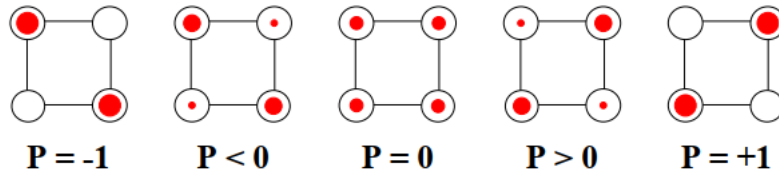


Figure 1: Непрерывное распределение состояния в клетке

Состояние клетки задаётся её поляризацией $P \equiv \frac{(\rho_1 + \rho_3) - (\rho_2 + \rho_4)}{\rho_1 + \rho_3 + \rho_2 + \rho_4}$, где ρ_i – ожидаемые заряды в каждой точке клетки.

Так, на примере выше, будет проще понять идею клеточной нейронной сети. Итак, обычно клеточная нейронная сеть – это, обычно, двумерный или трёхмерный массив аналогичных клеток. Каждая клетка обозначена k -м номером и она имеет динамическую переменную $\hat{x}_k$, внешний вход $\hat{u}_k$ и константную внутреннюю переменную $\hat{z}_k$. Каждая клетка зависит от своих соседей через синаптическую переменную, которая зависит от значений состояний клеток и входов клетки внутри сферы N_k , в центре в клетке k . Синаптический закон описывает влияние других клеток на синаптический вход $I_k^s = \sum_{\lambda \in N_k} (A_k^\lambda \hat{x}_{k+\lambda} + B_k^\lambda f(\hat{x}_k, \hat{x}_{k+\lambda}) + C_k^\lambda \hat{u}_{k+\lambda})$

Динамика CNN описывается системой нелинейных дифференциальных уравнений. Используя простейшую динамику ячеек первого порядка и линейные взаимодействия, уравнение состояния ячейки в положении (i, j) выглядит следующим образом: $\frac{dx_{ij}(t)}{dt} = -x_{ij}(t) + \sum_{(k,l) \in \mathcal{N}(i,j)} A(i,j;k,l) \cdot y_{kl}(t) + \sum_{(k,l) \in \mathcal{N}(i,j)} B(i,j;k,l) \cdot u_{kl} + z(i,j;k,l)$, где u_{ij} , x_{ij} и y_{ij} – вход, состояние и выход ячейки в позиции (i, j) соответственно; индексы k и l обозначают

общую ячейку, принадлежащую окрестности $\mathcal{N}(i, j)$ ячейки в позиции (i, j) . Все переменные являются непрерывными.

Набор матриц и порог A, B, z , который содержит веса нейронной/нелинейной сети, называется *шаблоном клонирования* и определяет операцию, выполняемую сетью. Когда значения шаблона клонирования не зависят от положения ячейки, CNN называется пространственно-инвариантным. В этом случае динамическое поведение сети зависит только от нескольких параметров

Выражение для выхода CNN y_{ij} определяется следующим образом: $y_{ij}(t) = f(x_{ij}(t)) = \frac{1}{2}(|x_{ij}(t) + 1| - |x_{ij}(t) - 1|)$, что соответствует нелинейности классических нейронных сетей.

На границе массива клеток явно меньшая часть попадёт в окрестность клеток, в связи с этим, на граничные клетки накладываются дополнительные условия. Они бывают разных типов:

- Фиксированные (или Дирихле) – если значение граничных ячеек является заданной константой;
- Нулевого потока (или Неймана), если значение граничных ячеек совпадает со значением граничных ячеек;
- Периодические (или тороидальные), если значение граничных ячеек совпадает со значением граничных ячеек на противоположной стороне (например, верхние граничные ячейки имеют значение нижних граничных ячеек).

4 Выводы:

Внедрение квантовой парадигмы влечёт за собой переконструирование общих принципов вычисления. Поэтому важно совершенствовать и рассматривать классические методы в процессах квантового вычисления.

Основным подходом для улучшения работы классических алгоритмов и методов машинного обучения долгое время считалось частичное выполнение вычислений на квантовом компьютере, однако в результате рассмотрения теоретических результатов квантовой физики были получены новые методы прогнозирования. К примеру, квантовые клеточные автоматы нашли применение в задачах, связанных с изображениями и распознаванием отпечатков пальцев.

Несомненно, результаты теоретических исследований в этой области смогут найти применение в ближайшие годы.

References

- [1] *Altaisky M.* Quantum neural network // *arXiv preprint quant-ph.* "— 2001.
- [2] *Fei L., Baoyu Z.* A study of quantum neural networks in neural networks and signal processing // *Proceedings of the 2003 International Conference on.* "— 2003. "— Vol. 1. "— Pp. 539 – 542.

- [3] *Georgescu I. M., Ashhab S., Nori F.* Quantum simulation // *Review of Modern Physics*. "— 2014. "— Pp. 86 – 153.
- [4] *Grover L. K.* A fast quantum mechanical algorithm for database search // *28th Annual ACM Symposium on the Theory of Computation*. "— 1996. "— P. 212.
- [5] Hidden quantum markov models and open quantum systems with instantaneous feedback / L. A. Clark, W. Huang, T. M. Barlow, A. Beige // *arXiv preprint*. "— 2014.
- [6] *Hilbert M., Lopez P.* The world's technological capacity to store, communicate, and compute information // *Science*. "— 2011. "— Vol. 332. "— Pp. 60 – 65.
- [7] *Horn D., Gottlieb A.* Algorithm for data clustering in pattern recognition problems based on quantum mechanics // *Physical Review Letters*. "— 2002. "— Vol. 8, no. 1.
- [8] *Lloyd S., Mohseni M., Rebentrost P.* Quantum algorithms for supervised and unsupervised machine learning // *arXiv preprint*. "— 2013.
- [9] *Pudenz K. L., Lidar D. A.* Quantum adiabatic machine learning // *Quantum Information Processing*. "— 1996. "— Pp. 2027 – 2070.
- [10] *Sagheer A., Zidan M.* Autonomous quantum perceptron neural network // *arXiv preprint*. "— 2013.
- [11] *Shor P.* Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*. "— 1994. "— P. 124.
- [12] *Siomau M.* A quantum model for autonomous learning automata // *Quantum information processing*. "— 2014. "— Vol. 13, no. 5. "— Pp. 1211 — 1221.
- [13] *Trugenberger C. A.* Quantum pattern recognition // *Quantum Information Processing*. "— 2002. "— Pp. 471 – 493.
- [14] *Zhou R., Ding Q.* Quantum mp neural network // *International Journal of Theoretical Physics*. "— 2007. "— Vol. 46, no. 12. "— Pp. 3209 — 3215.