

РАЗРАБОТКА ГИБРИДНЫХ КОНВЕЙЕРОВ АНАЛИЗА ДАННЫХ С ПОМОЩЬЮ ФРЕЙМВОРКА PennyLane

П. В. Зрелов¹, О. В. Иванцова^{1,}, М. С. Катулин¹*

¹ Объединенный институт ядерных исследований, Дубна, Россия

Предложен подход к разработке гибридных квантово-классических конвейеров анализа данных на основе композиции готовых квантовых шаблонов фреймворка PennyLane. Разработан и успешно протестирован гибридный конвейер квантово-классического машинного обучения для задачи классификации, демонстрирующий высокую точность на ряде тестовых наборов данных и возможность масштабирования архитектуры.

An approach to the development of hybrid quantum-classical data analysis pipelines based on the composition of ready-made quantum templates of the PennyLane framework is proposed. A hybrid quantum-classical machine learning pipeline for classification tasks has been developed and successfully tested, demonstrating high accuracy on a number of test datasets and the ability to scale the architecture.

PACS: 89.20.Ff

ВВЕДЕНИЕ

Современный этап развития квантовых вычислений (КВ) характеризуется смещением внимания в сторону прикладных разработок в различных областях [1–3]. Включение квантовых алгоритмов (КА) в конвейеры анализа данных позволяет создавать адаптивные системы, в которых квантовые модули дополняют традиционные методы, усиливая их эффективность. Это способствует развитию гибридных вычислительных платформ и ускоряет внедрение квантовых технологий в практические приложения. Отсутствие как стандартизированных методов интеграции КА в классические конвейеры анализа данных, так и базовых знаний в области квантовой физики у ИТ-специалистов подчеркивает актуальность разработки подходов, упрощающих использование КА для практических задач анализа данных.

Разработка КА на квантовых программных симуляторах (КПС) [4], эмулирующих квантовые вычисления на классических архитектурах, представляет собой важный этап современных исследований. Проектирование КА является сложной научно-технической задачей, требующей

* E-mail: ivancova@jinr.ru

учета фундаментальных квантово-механических эффектов и аппаратных ограничений современных квантовых процессоров. Эмуляция на классических вычислительных системах позволяет изучать поведение КА в контролируемых условиях, а также оптимизировать структуру алгоритмов до их развертывания на физических квантовых устройствах [5]. Использование КПС обеспечивает возможность детального моделирования квантовых систем, что особенно ценно в условиях ограниченной доступности реальных квантовых компьютеров, существенно сокращает временные и финансовые затраты на их разработку. Фреймворк PennyLane [6] представляет собой кроссплатформенное решение для программирования КА и предоставляет библиотеку шаблонов*, которые можно использовать для построения, оценки и обучения вычислительных моделей для решения конкретных практических задач. Шаблоны предназначены для кодирования данных в квантовые состояния, для определения обучаемых слоев квантовых вентилей, для подготовки квантовых состояний в качестве первой операции в схеме, просто как общие подпрограммы, из которых строится схема.

ОПИСАНИЕ

Конвейер анализа данных (рис. 1) представляет собой структурированный подход к обработке и анализу данных. На начальном этапе происходит сбор данных из различных источников. На следующем этапе осуществляется первичная обработка, фильтрация, преобразование и сжатие данных. На этапе анализа производится исследование подготовленных данных с целью выявления скрытых закономерностей, трендов, аномалий и взаимосвязей между признаками. Этап моделирования включает создание прогностических и описательных моделей. На заключительном этапе производится оценка точности моделей и анализ их значимости в контексте конкретных задач. В гибридных приложениях важно определить точки интеграции квантовых модулей в классический конвейер.



Рис. 1. Конвейер анализа данных

На рис. 2 представлены основные варианты встраивания квантовых модулей в классические конвейеры анализа данных.

* Шаблоны PennyLane: <https://docs.pennylane.ai/en/stable/introduction/templates.html>

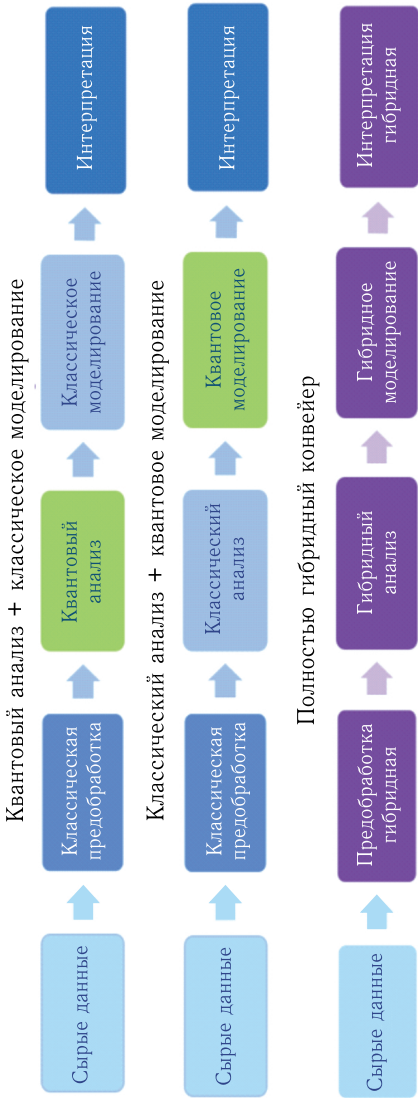


Рис. 2. Возможные варианты встраивания квантовых модулей в конвейеры анализа данных

- «Квантовый анализ и классическое моделирование» предполагает использование квантовых методов для предварительной обработки и преобразования данных (таких как квантовое снижение размерности или вычисление квантовых ядер) с последующим применением классических алгоритмов машинного обучения. Этот вариант особенно эффективен для задач с высокой размерностью признакового пространства.

- «Классический анализ и квантовое моделирование» сохраняет традиционные классические методы анализа данных и использует квантовые алгоритмы на этапе построения предсказательной модели. Такая архитектура часто применяется при работе с параметризованными квантовыми схемами и вариационными квантовыми классификаторами.

- «Полностью гибридный конвейер» является наиболее комплексным решением, комбинирующим квантовые и классические методы на всех этапах обработки и анализа данных. Этот вариант требует особенно тщательного проектирования интерфейсов между компонентами, но позволяет максимально использовать преимущества обеих вычислительных парадигм.

Квантовые методы наиболее органично внедряются в этапы анализа для выявления закономерностей, трендов, других значимых характеристик и моделирования (прежде всего в задачах квантового машинного обучения [7]).

ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ И ИХ АНАЛИЗ

Одной из задач классического машинного обучения является задача классификации. Ниже рассматривается задача по построению квантово-классического конвейера машинного обучения на примере тестовой задачи классификации с использованием стандартного набора данных*, состоящего из 178 образцов, каждый из которых имеет 13 признаков и принадлежит одному из трех классов, в подходе квантовых вариационных схем (VQC) [8].

Архитектура разработанного конвейера имеет следующие компоненты:

- Классические модули предобработки и анализа: нормализация данных и снижение размерности с помощью метода главных компонентов (PCA).

- Параметризованную квантовую схему с использованием шаблонов PennyLane для углового кодирования данных** и реализации квантовых слоев с сильной перепутывающей связью между кубитами***.

* https://scikit-dearn.org/stable/modules/generated/sklearn.datasets.load_wine.html#sklearn.datasets.load_wine

** В PennyLane реализуется с помощью шаблона `AngleEmbedding`.

*** В PennyLane реализуется с помощью шаблона `StronglyEntanglingLayers`.

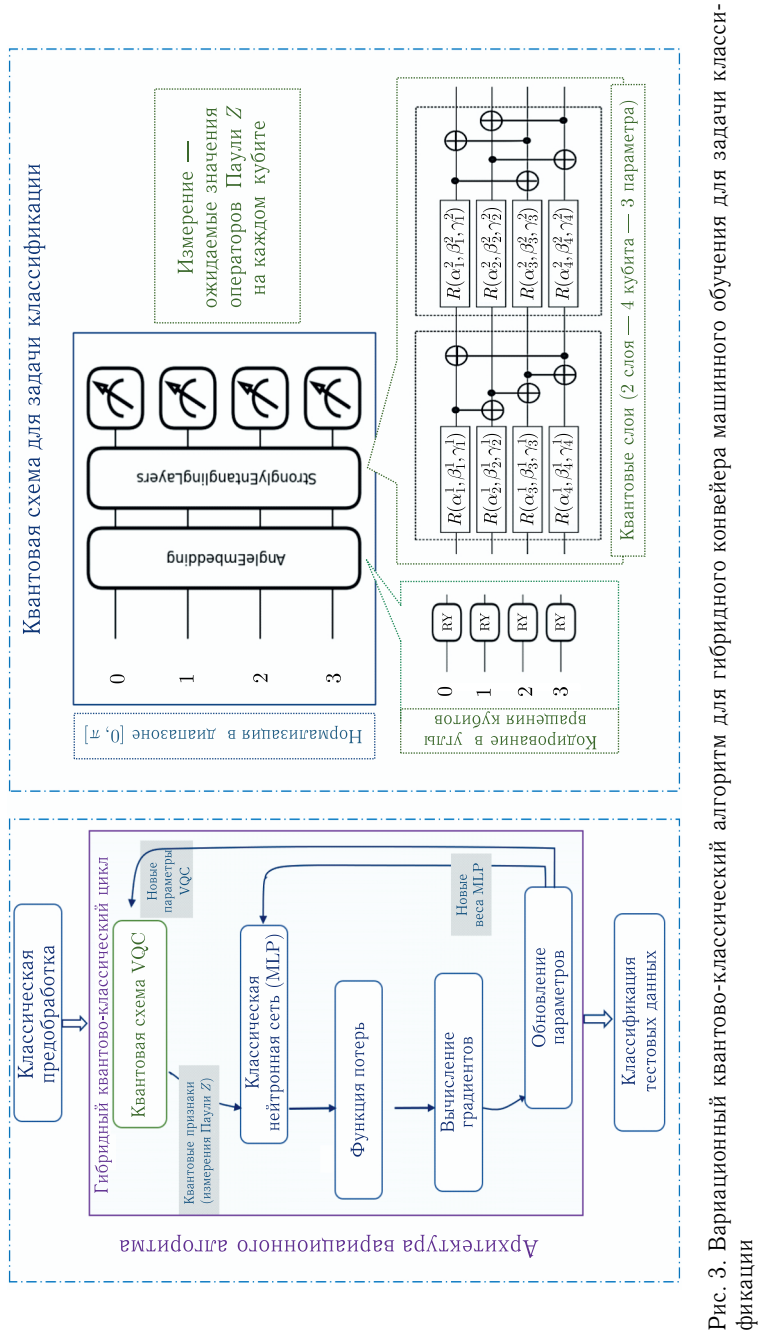


Рис. 3. Вариационный квантово-классический алгоритм для гибридного конвейера машинного обучения для задачи классификации

• Классическую часть модели — классификатор, представляющий собой многослойную нейронную сеть прямого распространения (MLP), на вход которой поступают результаты квантовых измерений. Выходной слой формирует вероятностное распределение событий по заранее заданным классам.

На рис. 3 представлены архитектура вариационного квантово-классического алгоритма (слева) и квантовая схема (справа) для гибридного конвейера машинного обучения для задачи классификации. На этапе предобработки данные нормализуются, и с помощью PCA количество признаков сокращено до 4. Количество признаков преобразованных данных определяет число кубитов квантовой схемы.

Квантовая схема VQC (выполняемая на квантовом симуляторе или квантовом компьютере) кодирует классические данные в квантовое состояние через угловое кодирование, выполняет нелинейные преобразова-

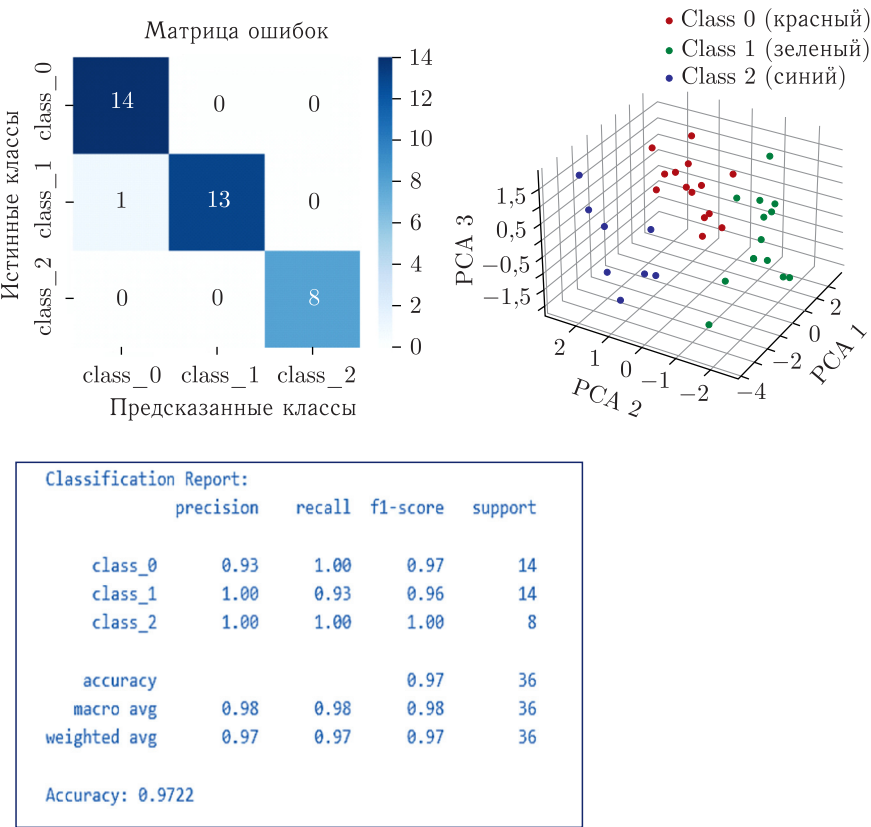


Рис. 4. Результаты классификации

ния, используя вариационные слои с перепутыванием, и извлекает квантовые признаки через измерения операторов Паули. Роль классической нейронной сети состоит в интерпретации квантовых признаков и выполнении финальной классификации. Совместная работа обеспечивается единым циклом оптимизации. В прямом проходе данные последовательно обрабатываются квантовой схемой и классической нейронной сетью. При обратном распространении градиенты вычисляются через всю систему. Оптимизатор одновременно обновляет параметры обоих компонентов: углы вращения в слоях квантовой схемы и веса классической нейронной сети.

Оценка качества модели проводилась на тестовой выборке (20% исходных данных) с вычислением стандартных метрик классификации Precision, Recall и F1 (рис. 4). Модель демонстрирует точность в пределах от 97,22 до 100% на тестовых наборах данных. В качестве базового классического алгоритма для сравнения с гибридным конвейером был выбран алгоритм случайного леса (Random Forest). При работе с небольшими и средними наборами данных этот метод, как правило, демонстрирует высокую эффективность на задачах классификации и регрессии. Алгоритм случайного леса показал на тестовых данных точность, близкую к 100%, что соответствует результатам работы гибридного алгоритма, описанного выше.

ЗАКЛЮЧЕНИЕ

Представлен практический опыт построения гибридных квантово-классических конвейеров с использованием фреймворка PennyLane для реализации параметризованных квантовых схем, библиотеки машинного обучения PyTorch и фреймворка scikit-learn. Основное внимание уделено методологии интеграции параметризованных квантовых схем в традиционные конвейеры анализа данных, где квантовые слои выполняют отображение признаков в пространство состояний кубитов и нелинейные преобразования, а классические нейросетевые компоненты обеспечивают финальную классификацию. Экспериментальная реализация демонстрирует работоспособность подхода на задаче классификации и его конкурентоспособность с классическими методами. Полученные результаты свидетельствуют о перспективности дальнейшего исследования гибридных архитектур для решения прикладных задач анализа данных, в частности, для создания масштабируемых квантово-вдохновленных алгоритмов машинного обучения.

СПИСОК ЛИТЕРАТУРЫ

1. Weder B., Breitenbücher U., Leymann F., Wild K. Integrating Quantum Computing into Workflow Modeling and Execution; doi:10.1109/UCC48980.2020.00046. 2020.
2. Sau Lan Wu et al. Application of Quantum Machine Learning Using the Quantum Variational Classifier Method to High Energy Physics Analysis

- at the LHC on IBM Quantum Computer Simulator and Hardware with 10 Qubits // *Phys. G: Nucl. Part. Phys.* 2021. V.48, No. 12. 125003; doi: 10.1088/1361-6471/ac1391.
3. *Mugel S., Kuchkovsky C., Sanchez E. et al.* Dynamic Portfolio Optimization with Real Datasets Using Quantum Processors and Quantum-Inspired Tensor Networks // *Phys. Rev. Res.* 2022. V.4, No.1. 013006; doi: 10.48550/arXiv.2007.00017.
 4. *Paltenghi M., Pradel M.* A Survey on Testing and Analysis of Quantum Software; doi: 10.48550/arXiv.2410.00650. 2024.
 5. *Zrelov P. V., Ivantsova O. V., Korenkov V. V. et al.* Evaluation of the Capabilities of Classical Computers in the Implementation of Simulators of Quantum Algorithms // *Software Products and Systems.* 2022. V.35, No. 4. P.618–630.
 6. *Bergholm V., Izaac J., Schuld M. et al.* PennyLane: Automatic Differentiation of Hybrid Quantum-Classical Computations. arXiv:1811.04968. 2018.
 7. *Biamonte J. et al.* Quantum Machine Learning // *Nature.* 2017. V.549. P.195–202.
 8. *Ramezani S.B. et al.* Machine Learning Algorithms in Quantum Computing: A Survey. doi: 10.1109/IJCNN48605.2020.9207714. 2020.