

## Документация за решението

# Квантови изчисления в машинното обучение: анализ на ефективността

### 1. Тема

Проектът изследва потенциала на квантовите компютри за усъвършенстване на алгоритмите за машинно обучение, като предоставя на изследователите в областта на изкуствения интелект платформа за оценка на ефикасността на квантовите изчисления в техните специфични приложения. Това включва интегриране на персонализирани датасети и алгоритми за анализ на въздействието на квантовите технологии върху ключови показатели като скорост на обучение и точност на моделиране.

### 2. Автори

**Никита Артемович Черневский**, ЕГН: 0942060063

Гр. Разлог, ул. Гоце делчев, 49. Тел.: +359 87 6401515, +359 87 8181336

Email: [Nikita.Chernevskiy@hotmail.com](mailto:Nikita.Chernevskiy@hotmail.com), [nc78983818@edu.mon.bg](mailto:nc78983818@edu.mon.bg),

9.клас, Професионална Гимназия по Транспорт, гр. Разлог

### 3. Ръководител

**Анета Самуилова Даилова**, Тел.: +359 88 9866 453, [aneta.dailova@pgt-razlog.org](mailto:aneta.dailova@pgt-razlog.org),

Старши учител по ИТ и информатика, Професионална Гимназия по Транспорт, гр. Разлог

### 4. Цел и задачи

Предоставяне на инструмент за анализ на ефективността на квантовите изчисления за техните конкретни датасети и използвани методи на машинно обучение на изследователи, работещи в областта на машинното обучение и изкуствения интелект.

Решаеми задачи:

**1. Разработка на аналитичен инструмент.** Създаване на инструмент, който позволява на изследователи в областта на машинното обучение и изкуствения интелект да вграждат и анализират своите датасети и алгоритми в квантова изчислителна среда.

**2. Изпитване и валидиране на методологията.** Провеждане на обширни експерименти за тестване на различни модели на машинно обучение, използвайки квантови изчисления, за да се оцени тяхната ефективност и точност спрямо стандартните изчислителни методи.

## 5. Основни етапи в реализирането на проекта

1. Обединение на таблиците в една чрез Power BI
2. Преобразуване на получената таблица в CSV формат
3. Допълване на таблицата с рандомни значения използвайки квантова генерация на рандомни числа и шум
4. Превръщане на CSV файлове в Json чрез csvtojson.py
5. Използване на 10те json файлове като Training Data

Използван датасет на **University of Sussex, UK** – специално подготвени данни за машинно обучение:

- ✓ Анализа и класификация на ферми на здравословни и нездравословни за диви и домашни пчели.
- ✓ HalfMoon модел използва тези данни и след обучение класифицира ферми на здравословни (1) или нездравословни (0) във въведените Validation Data.

### Използвани колони

**Farm** – номер на дадена ферма

**Type** – HLS (здравословни) и ELS (нездравословни)

**Year** – Годината на събирането на данните (2013-2014)

**Round** – Период на събирането на данните (3 за година)

**Floral** – Общо количество растения в хиляди

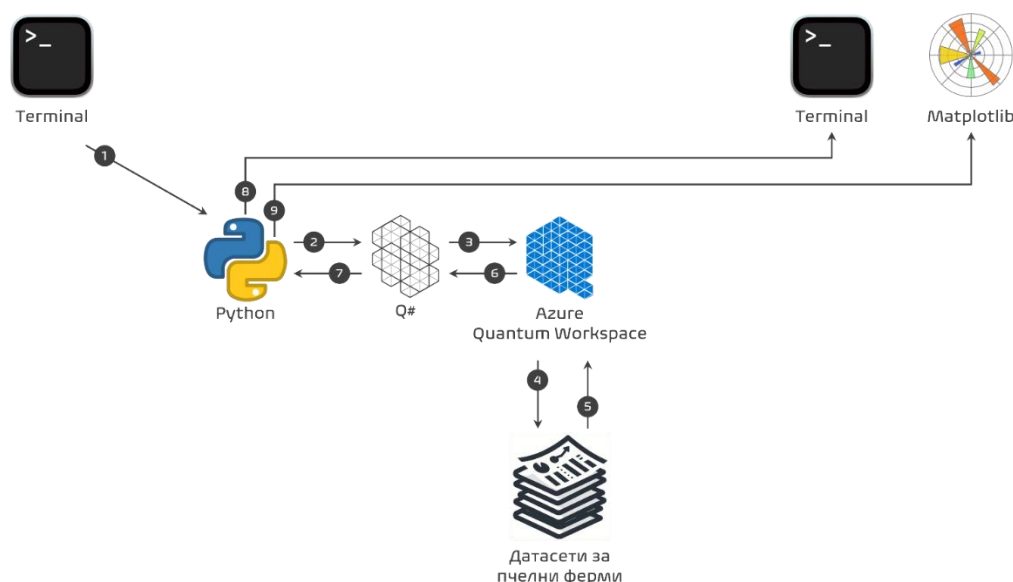
**Bees** – Общо количество пчели в хиляди

## 6. Ниво на сложност на проекта

Проектът представя средно до високо ниво на сложност, поради интегрирането на квантови технологии и специфични алгоритми за машинно обучение. Той изисква задълбочени познания в областта на квантовите изчисления и тяхната практическа приложимост в рамките на изкуствения интелект. Освен техническата компетентност, успешната реализация на проекта също зависи от способността да се адаптират и интегрират нови методологии в съществуващи системи и инфраструктури за машинно обучение.

## 7. Логическо и функционално описание на решението

### 7.1 Архитектура



### 7.2 Предимства на архитектурата

- ✓ **Скалируемост:** Архитектурата е проектирана така, че да може лесно да се адаптира към увеличаващи се изисквания за обработка на данни и сложност на моделите, което осигурява гъвкавост при разширяване на проекта.
- ✓ **Ефективност:** Благодарение на интеграцията с квантови изчислителни ресурси, архитектурата обещава значително подобрене в скоростта на обработка и анализ на данните, което е критично за бързото развитие на модели на машинно обучение.
- ✓ **Сигурност:** Вградените мерки за сигурност и защита на данните гарантират, че всички операции с датасети и модели са защитени от външни заплахи и злоупотреби.
- ✓ **Отворен код и потребителска навигация:** Архитектурата е разработена с акцент на отворен код, което позволява на изследователската общност да допринася и оптимизира системата. Потребителският интерфейс е интуитивен, което улеснява изследователите в процеса на работа с инструмента.

## 8. Реализация

Използването на водещи продукти и езици за програмиране ни позволява да създадем стабилна и функционална система, която да извлече максимални ползи от квантовите изчисления.

Включването на иновативния продукт **Azure Quantum Workspace** от Microsoft позволява директен достъп до възможностите на квантовите изчисления, осигурявайки надеждна платформа за екзекуция на програмни кодове.

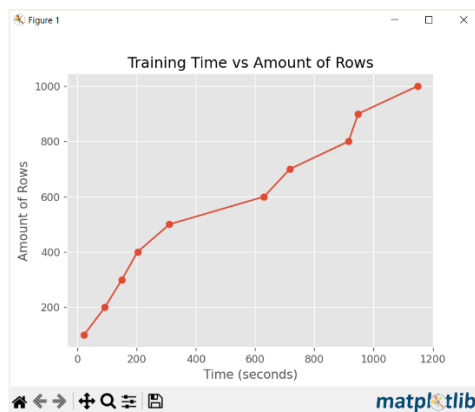
Интеграцията с **Visual Studio Code** като безопасен интегриран развойн среда (IDE) поддържа последователното развитие и подобрене на кода, докато програмният език **Python** се утвърждава като стандарт за разработка и визуализация на квантови изчисления.

Допълнително, използването на **Q#**, специализиран език за квантово програмиране, и адаптирането на **модела HalfMoon** за класификация на данни, демонстрират високата адаптивност и функционалност на системата при работа с двоични класификационни задачи.

Така всеки аспект от технологичната инфраструктура е стратегически подбран и оптимизиран за максимизиране на производителността и ефективността на проекта.

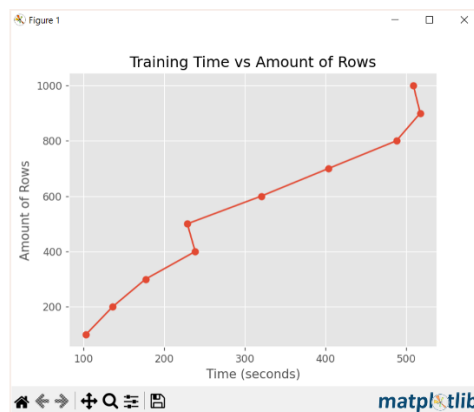
## 9. Описание на проекта

В раздела "Описание на проекта" се излага същността на проведената работа, която включва два ключови експеримента. Първият експеримент е извършен на класически компютър и има за цел да установи базови показатели за времето на обучение и точността на избрани модели за машинно обучение. Вторият експеримент е проведен на квантов компютър, като се използват същите модели, за да се направи сравнение на производителността и да се измери влиянието на квантовите изчисления върху процеса на обучение.



### Класически компютър

Времето за обучение на модела се увеличава експоненциално с нарастването на броя редове в данните. Това показва, че времевите изисквания нарастват пропорционално на размера на обучаващия набор от данни.



### Azure Quantum Workspace

Обучението в Azure Quantum Workspace е значително по-бързо в сравнение с класически компютър, особено при обработката на големи обеми на данни.

По-стабилни изчислителни ресурси.

Резултатите и анализът от извършените експерименти са следните:

### Експеримент на Класически Компютър:

- Моделите за машинно обучение бяха тренирани на стандартен компютърен хардуер.
- Установено беше, че с увеличаването на сложността и обема на данните, времето за обучение нараства значително.

**Експеримент на Квантов Компютър:**

- Същите модели бяха обучени, използвайки квантов компютър.
- Забелязва се значително ускорение във времето за обучение в сравнение с класическия компютър.

С нарастването на значението на големите езикови модели (LLM) и изкуствения интелект (ИИ) като цяло, възможността за оценка на ползата от квантови компютри става все по-важна. С разработката на такива програми, всеки разработчик на ИИ може да изтегли софтуера от GitHub, да добави свои собствени датасети и алгоритми за машинно обучение, и по този начин да оцени ефективността на квантовите изчисления за своя специфичен проект.

Пълните инструкции за установка на решението, както и изходният код, са свободно достъпни за изтегляне от GitHub (<https://github.com/NikitaChernevskiy/EAQCMPLP>). Проектът е лицензиран под Лиценза MIT, което осигурява свободата на използване, копиране, модифициране, сливане, публикуване, разпространение, сублицензиране и/или продажба на копия на софтуера. Това гарантира, че всеки желаещ да работи с проекта, може свободно да го използва за лични и търговски цели, като същевременно споделя усъвършенстванията с общността и допринася за нейното развитие.

Този подход позволява индивидуална оценка и адаптация към конкретните нужди на всеки проект. Разработчиците могат да експериментират с различни видове датасети и алгоритми, за да наблюдават как квантовите изчисления влияят на скоростта и точността на обучение, както и на обработката на данни. Това им дава възможност да идентифицират потенциални области на подобрене и да определят дали квантовите компютри представляват изгодно инвестиране за техния проект.

Освен това, тази платформа може да спомогне за разширяване на разбирането за квантовите технологии в контекста на ИИ. Това включва изследване на нови приложения и използване на квантови алгоритми за решаване на сложни проблеми, които традиционните изчислителни методи не могат ефективно да обработят.

**10. Заключение**

Проектът демонстрира, че понякога използването на квантови компютри в машинното обучение значително подобрява скоростта и ефективността на обработка на големи обеми на данни.

Но е необходимо провеждането на допълнителни изследвания на различни датасети, модели за машинно обучение и квантови компютри от различни поколения, за да се получат точни заключения.