

МЕТОДИ И АЛГОРИТМИ ЗА МАШИННО ОБУЧЕНИЕ ЧРЕЗ КВАНТОВИ СПЛИТАНИЯ

на дисертация за придобиване на образователна и научна степен
"ДОКТОР"

Научна специалност: „Системи с изкуствен интелект“

[illegible]

СОФИЯ, 2020 г.

Дисертационният труд е обсъден и насочен за защита от Катедрения съвет на катедра „Компютърни системи“ към Факултет по компютърни системи и технологии на ТУ-София на редовно заседание, проведено на 16.06.2020 г.

Публичната защита на дисертационния труд ще се състои на 24.09.2020 г. от 13 часа в Конферентната зала на БИЦ на Технически университет – София на открито заседание на научното жури, определено със заповед № ОЖ-.... / дд.мм.гггг г. на Ректора на ТУ-София в състав:

1. Доц. д-р Надежда Ангелова Кучмова
2. Проф. д-р Милена Кирилова Лазарова-Мицева
3. Проф. д-р Сотир Николов Сотиров
4. Проф. д-тн Красимира Петрова Стоилова
5. Проф. д-тн Димитър Грозданов Христозов

Рецензенти:

1. Проф./доц. д-р/д.т.н. Име Фамилия
2. Проф./доц. д-р/д.т.н. Име Фамилия

Материалите по защитата са на разположение на интересуващите се в канцеларията на Факултет по компютърни системи и технологии на ТУ-София, блок № 1, кабинет № 1441.

Дисертантът е задочен докторант към катедра „Компютърни системи“ на факултет по компютърни системи и технологии. Изследванията по дисертационната разработка са направени от автора, като някои от тях са подкрепени от научно-изследователски проекти.

Автор: маг. инж. Десислав Андреев

Заглавие: Методи и алгоритми за машинно обучение чрез квантови сплитания

Тираж: 30 броя

Отпечатано в ИПК на Технически университет – София

I. ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност на проблема

Сложността на системите за изкуствен интелект нараства значително през последното десетилетие, защото те все по-лесно могат да намерят място не само в научните, но и в ежедневните приложения. Увеличената сложност заедно с все по-големия брой входни данни водят до недостиг на изчислителен ресурс при персоналните компютри и затова такъв тип функционалности се прехвърлят постепенно към облачните услуги, осигуряващи достъп до суперкомпютри. Имайки предвид този темп на развитие, се очаква недостатъчно бързодействие дори при подобни постановки, а усложняването на хардуерните системи е възможно да не покрие нуждите на алгоритмите за машинно обучение. От своя страна квантовите компютри са безспорна алтернатива, когато става въпрос за бързодействие. Сложната им природа не позволява лесен способ за работа с тях, а трудната им реализация за различни цели е сериозен недостатък. Някои компании, доставящи изчислителни ресурси чрез облачни услуги, осигуряват и достъп до свои квантови компютри, както и онлайн симулации на такива. Класическите алгоритми за машинно обучение, изискващи голямо време за изпълнение, могат да бъдат преобразувани чрез математически способности и приложени в среда на квантов компютър, като по този начин да се осигури възможност за по-бързото им изчисление. Образуването на комплекс от алгоритми, както и изграждането на комуникационни възли между тях, подобно на невронните мрежи, не би могло да се изпълни без т. нар. квантови сплитания.

Цел на дисертационния труд, основни задачи и методи за изследване

Целта на настоящия дисертационен труд е: „Чрез изследване и анализ на възможностите за използване на комплексни квантови системи за машинно обучение да се предложат алгоритми за машинно обучение в квантово-сплетена среда“. За нейното постигане е необходимо да се решат следните задачи:

1. На базата на литературен преглед на съвременното състояние на квантовите изчисления да се дефинират основните ресурси и средства за създаване на квантова среда и да се анализират предимствата и недостатъците на съществуващите подходи за преобразуване на алгоритми за квантово машинно обучение и основните компоненти за създаване на комплексна квантова система.
2. Да се анализират и сравнят математически модели на алгоритми, използващи квантови сплитания, и да се направи оценка на тяхната възможност за реализация.
3. Да се предложи математически модел за изграждане на съставни квантови алгоритми за машинно обучение в квантово-сплетена среда и да се приложи предложеният математически подход с цел определяне на предпоставките и възможности за прилагане на квантови алгоритми.
4. Да се извърши експериментален анализ относно приложението на предложеният математически модел, на базата на който да се оценят неговите възможности, предимства и недостатъци.

Научна новост

Изграждането на системи за квантово машинно обучение, както и математическият анализ за преминаване от класическите в квантови реализации, е клон в информационните системи, който тепърва започва да се развива, особено с областта на невронните мрежи и клъстериращите алгоритми. Проектирането и имплементацията на такива способности позволява различен, но пълноценен поглед върху системите с изкуствен интелект, изискващи голямо време за изчисление, и е предпоставка за извършване на проведените изследвания в настоящия дисертационен труд.

Практическа приложимост

Практическата приложимост на дисертационния труд е насочена към следните няколко ключови аспекта:

- Анализиране и проектиране на квантови логически схеми и квантови алгоритми за машинно обучение.
- Приложение на квантовите сплитания в алгоритмите за машинно обучение и по-специално – при проектирането на невронни мрежи.
- Прилагане на подход чрез квантови сплитания за скалиране на квантови алгоритми.
- Консолидация на софтуерни компоненти и системи при изграждането на съставни квантови алгоритми за машинно обучение.

Апробация

Апробацията е извършена чрез реализация на предложените квантови системи в симулационни среди и прилагането на представените математически способности. Наблюдаваните им поведения, както и получените резултати, са сравнени с еквивалентни системи в класически компютри.

Публикации

Основни постижения и резултати от дисертационния труд са публикувани в 7 научни статии, от които 1 в "Spring Conference of the Union of Bulgarian Mathematicians" 2018; 1 в Journal of Internet Technology and Secured Transaction, Vol. VI, Issue 2; 1 в "IEEE International Conference for Internet Technology and Secured Transactions" 2018; 1 в "9th Balkan Conference on Informatics" 2019; 1 в "Future Technologies Conference" 2019; 2 в "21st International Conference on Computer Systems and Technologies" 2020. Две от публикациите са самостоятелни.

Структура и обем на дисертационния труд

Дисертационният труд е в обем от 151 страници, като включва увод, 5 глави, списък на основните приноси, списък на публикациите по дисертацията, списък на участия в научно-изследователски проекти и използвана литература. Цитирани са общо 124 литературни източника, като 90 са на латиница и 8 на кирилица, а останалите са URL-връзки. Работата включва общо 43 фигури и 9 таблици. Номерата на фигурите и таблиците в автореферата съответстват на тези в дисертационния труд.

II. СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

ГЛАВА 1. ОСНОВИ ЗА КВАНТОВИ ИЗЧИСЛЕНИЯ В МАШИННОТО ОБУЧЕНИЕ

Когато се говори за оптимизация в машинното обучение или алгоритмите за изкуствен интелект, трябва да се има предвид, че самите техники в тези дисциплини продължават да се развиват благодарение не само на нови подходи в самите алгоритми и техните математически модели, но в хардуера, който се разработва в наши дни. Квантовото машинно обучение е добър пример за двете насоки едновременно: необходима е квантова среда, която да служи за експериментална платформа, и математически апарат, който да позволи превода на алгоритмите от класическия към квантовия свят. До този момент и двете водят до сериозни предизвикателства, основно поради факта, че не съществува универсален системен подход за решението на въпросните проблеми: относно имплементацията на подобен хардуер може да се каже, че такъв се произвежда конкретно за специализираните приложения. Липсва също и системен, общоприложим начин за преобразуване на класически алгоритъм към квантов. Още повече, че не за всеки класически алгоритъм

съществува реална нужда от квантова среда. Всяко приложение на квантовите изчисления се разработва само за себе си и се базира на хардуерна и софтуерна среди, предоставени и подготвени само за специфичния пример. Очевидно е, че това се дължи на самата физическа природа на квантовия свят, който не само все още бива изучаван физически, а и предстои да бъде широко представен и използван като инструмент в компютърните изчисления. Стига се и до невъзможността за описание на цялостната картина, която би могла да се получи, ако всички квантови феномени: *квантов паралелизъм*, *квантова интерференция* и *квантово сплитане*, бяха в сила едновременно. Самото име на квантовото сплитане подсказва за изграждане на мрежи или по-сложни квантови системи. Ако се разгледа хардуерното представяне на феномена, става ясно, че той спомага за решаването на проблема с квантовата оперативна памет: 20 квантови бита могат да съхраняват до 2^{20} на брой стойности. Абстрактно погледнато, подобни топологии могат да се приложат чисто математически за повече битове и съответно да се опишат чрез по-сложни системи. Към това се включват и по-тежките за изпълнение алгоритми за машинно обучение, които за момента набират сила в системите с паралелни изчисления. За да се изпълнят на квантов компютър, те трябва да се преобразуват в подходяща форма. Това предполага и създаване на алгоритми, използващи някакъв вид мрежа, които да бъдат описани и математически оптимизирани чрез квантови сплитания.

Трансформирането на алгоритмите от класическия им вид към техния квантов еквивалент трябва да се поясни чрез базовите понятия, които стоят в основата на квантовите изчислителни машини, буквално компютри. Други важни дефиниции и теореми от областта на квантовата механика, които са важни при описанието и реализацията на квантови алгоритми са *квантово състояние*, *вълнова функция*, *Теорема на Еренфест*, *квантов бит*, *квантови регистри*. Квантовият паралелизъм обединява *принципа на суперпозицията* и *линейността от квантовия свят* с цел изчислението на единствена функция едновременно за произволен брой входни данни. Квантовата интерференция спомага за създаването на логически пътища след определено изчисление, така че те да си влияят: този ефект води до това желаните изходни данни да бъдат подсилени, а нежеланите да отпаднат с времето. Това саморегулиране на системата е очаквано поведение в машинното обучение, особено когато се има предвид изчислението на алгоритми върху големи масиви от данни. Квантовите състояния, са обособени от много частици и не могат да се разгледат като самостоятелни такива с по една частица. Отношението между тях не може да се дефинира чрез средствата на класическата физика и затова се дефинира основният източник за анализ на квантовата информация и също мощно средство за комуникация – квантово сплитане. В текущата глава са разгледани подробно следните понятия: (А) квантов бит; (Б) вълнова функция; (В) измерено състояние; (Г) квантов гейт; (Д) квантова памет; (Е) изпълнение на квантова функция; (Ж) сфера на Блох.

Благодарение на представените дотук градивни елементи, се представя проектирането на разнообразни квантови логически схеми. Както начина на работа на всяка изчислителна машина може да бъде описана чрез логически схеми, така и квантовия компютър бива изграден от такива. Комбинацията от описаните гейтове дава възможността за имплементиране на множество алгоритми и теоретичния дизайн на сложни устройства. Реализацията на тези комбинации е невинаги възможна. По тази причина голяма част от проучванията в сферата на квантовите компютри са основно теоретични. Научните изследвания в дисертационния труд се базират на математически анализ и реализация в симулационни среди. Например, всяка линейна трансформация на комплексно векторно пространство може да се опише чрез матрица. Квантовите логически схеми се представят чрез унитарни матрици, т.е. $MM^\dagger = I$, където I е единична матрица, а $M^\dagger$ е комплексно спрегната и транспонирана матрица на матрицата M . Това прави началната матрица унитарна трансформация, т.е. тя е реверсивна. Видима прилика между квантовите гейтове и класическите логически елементи е работата чрез няколко входни сигнала. Съгласно дефинираното понятие за квантов гейт при k на брой входни квантови бита, е необходима

матрица с размерност $2^k \times 2^k$ на един квантов гейт. Това води до друго основно правило: размерите на входните и изходните матрици трябва да останат равни. За целите на математическото моделиране на квантовата система е необходимо да се дефинира как ще се извършва умножение на унитарната матрица, определена от гейта, с входния вектор – всъщност, квантово състояние на квантовия бит. Именно това са средствата, на които се основават предложените в дисертационния труд решения: (А) логически елементи: гейт на Адамар; гейт на Паули; CNOT-гейт; гейт на Тофоли; swap-гейт; (Б) операции и примерни логически схеми: измерване; близост на две квантови състояния.

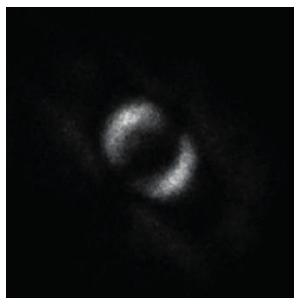
Информацията дотук показва, че е възможно взаимодействие между квантови битове, което се приема за по-сложен казус от изследването на едночастични системи, т.е. гейтове, работещи с един квантов бит. Комбинацията от такива, обаче, води до създаването на контролирано многочастично взаимодействие между самите квантови битове. В такъв случай, това се назовава като *сплетено състояние*, известно още като *състояние на Бел*. Анализирайки го, се достига до следния извод: или измерването на състоянието на кубит А има нелокален ефект върху кубит В, или квантовата механика е непълна теория и тя дава само частична информация за състоянието на система с класически корелации. Този факт води до необходимостта от допълнителни изследвания не само на самия феномен, но и на неговия ефект върху нерелативисткия свят, включително и теорията за квантовите компютри.

Квантовото сплитане е феномен, който има различни изражения във физиката, т.е. пълното му приложение не може да се установи до момента. С разгледаната дотук информация относно квантовите изчисления става ясно, че могат да се изграждат квантови системи с повече от един квантови бита. Също така се налага тези системи да се третират като едно цяло: съществува една вълнова функция, т.е. един вектор на състоянието, описващ дадената система. Подобно на класическите системи, всеки квантов бит може да притежава свои дискретни параметри, като в случая те обозначават *спиновите* състояния. Съществена разлика е наличието само на една времева координата t . Класическа комуникация чрез квантова механика - *квантово сплитане* [1]: В една реална система, макар и експериментално, кубитът може да бъде представен като двете различни поляризации на фотона, проекциите на спина на електрона, както и като две атомни електронни нива. Най-обещаващите кандидати за физична реализация на квантов компютър са йонни уловки, фотони, кондензат на Бозе-Айнщайн в оптична решетка, атоми и квантови точки. Въпросните квантови точки имат пряка връзка с теоретична възможност за изграждане на същински квантови компютри, възползващи се от огромните размери на пространствата от вълнови функции. Това се обяснява със следния пример:

Пример 1: При две частици А и В, всяка от които би могла да се намира на 10 различни места, съществуват 100 различни разположения на въпросната двойка (могат да се намират на едно и също място). Съответно са необходими 100 комплексни числа z_{00-99} , за да се опише вълновата функция на тази система. При нормиране, сумата на квадратите на модулите на тези числа ще бъде 1. Примерен опит може да бъде, където това ще бъде вероятността частицата А да се намира в положение 3, и частицата В – в положение 8, съответно. Аналогично на описанието по-горе, Хилбертово пространство става $\mathcal{H}^{100}$. Следвайки тази логика, то три частици, ще имат вълнова функция, конструирана от 1000 комплексни числа. При три отделни вълнови функции Хилбертово пространство щеше да бъде само $\mathcal{H}^{30}$. Връщайки се към тримерното Евклидово пространство $\mathbb{E}^3$, можем да приемем, че то е свободно избрано гладно поле, т.е. притежава ∞^3 брой точки [74]. Въпросните частици се състоят от два реални компонента, тъй като са комплексни числа, то степените им на свобода ще бъдат $\infty^{2 \times 3}$. Това е истина, когато полето се разгледа в даден момент от време, т.е. t е константа. При времепространство се налага употребата на 4 измерения [5]. Ако се вземат предвид 10-те различни места в примерната Вселена по-горе, то за n на брой частици степените на свобода ще бъдат $\infty^{2 \times 10^n}$. С други думи Хилбертово пространство ще стане

$\mathcal{H}^{10^n}$. Въпросният пример има за цел да представи колко на брой измерения могат да съществуват в един момент от време, дори за малък брой частици. Видимото ускорение е търсен ефект в квантовите компютри и основна причина за тяхното изучаване [21]. Съществуването на квантовото сплитане е доказано експериментално, но не е наблюдавано преди [85]. На Фигура 6 е представено изображение, представляващо наблюдавано сплитане между два фотона:

Въпросният пример има за цел да представи колко на брой измерения могат да съществуват в един момент от време, дори за малък брой частици. Видимото ускорение е търсен ефект в квантовите компютри и основна причина за тяхното изучаване. На Фигура 6 е представено изображение, представляващо наблюдавано сплитане между два фотона:



Фигура 6. Визуализация на квантово сплитане между два фотона.

Трябва да се спомене и още едно събитие, а именно квантовата телепортация. В случая се има предвид размяна на информация със скорост по-висока от скоростта на светлината – проблем, който е несъвместим с релативистката физика. Тук е важно да се уточни, че не става въпрос за удвояване на индивиди, т.е. копиране на неизвестно квантово състояние, а за същински пренос на информация. Произходът на въпросния тунел е неясен, а самият той не може да се изследва, заради принципите на квантовото сплитане, които се разглеждат по-нататък в научния труд. Въпреки това квантовото състояние на даден кубит е всъщност съвкупност от 0 и 1. Например цифрата 2 може да се кодира като 10 чрез двоичен код и така да се предаде като информация. Имайки предвид свойствата на квантовия канал за комуникация, например, че не може да се подслушва, квантовата криптография напредва значително и се използва в реално работещи проекти с висока степен на сигурност. Комбинацията от двете свойства все още не е изучавана в средата на алгоритмите за машинно обучение, основно поради трудността на задачата за създаване на квантов компютър, изпълняващ точно такъв тип задача. Това дава повод за разглеждането именно на алгоритми, изграждащи връзки между елементите си, или такива, изискващи по-бърза комуникация между функционалните си модули. Обобщен концептуален модел за изграждане на квантова информационна система е представен на Фигура 7.



Фигура 7. Системни блокове за изграждане на единна квантова информационна система.

Съществуват три реда в разработването на квантов алгоритъм или квантов компютър – както стана по-ясно по-горе, двата термина могат да се припокриват:

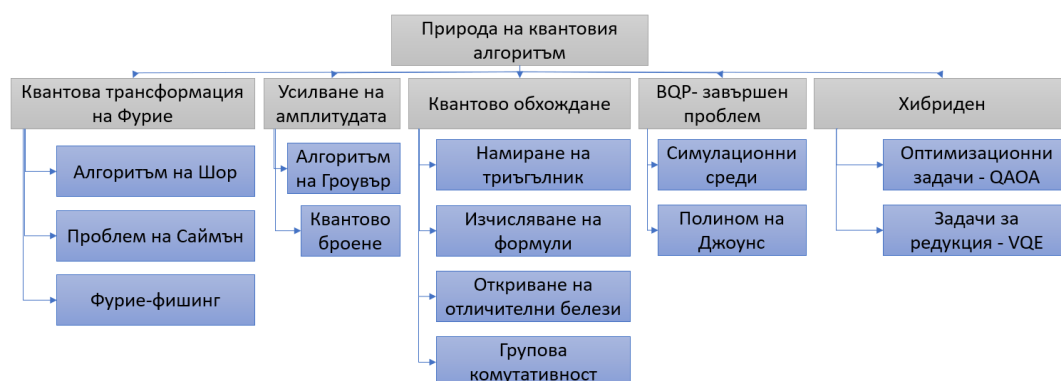
- Ред 1 – квантовите физически модели, изпълняващи различни изчислителни парадигми;

- Ред 2 – математическите модели, описващи алгоритмите, се преработват така, че да могат да се опишат с квантови операции;
- Ред 3 – самият алгоритъм, разработен за квантово изчисление – частично или напълно. Това включва и замяната на стъпки за изпълнение от основен, класически алгоритъм, с квантово-имплементирани такива.

В дисертационния труд не се разглеждат разработки от Ред 1 поради изискването за физическото им съществуване макар и в лабораторни условия. Тъй като целта на дисертационния труд е математическа обосновка на нуждата от квантово машинно обучение, включващо квантови сплитания, то съществуването и употребата на система от Ред 1 няма добавена стойност. От друга страна липсата на източници за постигане на резултати от Ред 3 правят задачата трудна не само за моделиране на такава система, но и за пояснения към бъдещата ѝ имплементация.

ГЛАВА 2. АНАЛИЗ НА ДИЗАЙН И ИМПЛЕМЕНТАЦИЯ НА КВАНТОВИ АЛГОРИТМИ В ИНФОРМАЦИОННИТЕ ТЕХНОЛОГИИ

В текущата глава се разглеждат основополагащите алгоритми в квантовите изчислителни системи. Целта е да се представят основните стъпки при проектирането на по-сложни, наречени още *съставни*, квантови алгоритми. Също така трябва да се провери дали даден алгоритъм може да бъде невъзможен за трансформация от класическа към своята квантова форма, както и квантов алгоритъм да не може да бъде трансформиран към класически. Такъв пример е *алгоритъмът на Гроувър*. Друг такъв е *алгоритъмът на Шор*, който се счита за един от най-добрите примери, тъй като намира решение за дискретен логаритъм или факторизиране на цели числа за много кратко време – притежава полиномна сложност. Интересен пример е *проблемът на Саймън*. Този алгоритъм е първият, показващ експоненциално ускорение спрямо свой аналог в класическите компютри. Освен това включва в себе си т. нар. вероятностни алгоритми с ограничение на грешката при полиномно време. Ключово свойство, което трябва да се отбележи, и което ще се разгледа по-нататък в текущия научен труд, е, че тези алгоритми се базират на квантово сплитане, но по своя характер са коренно различни. Примери за квантови алгоритми са дадени на Фигура 8.



Фигура 8. Примери за квантови алгоритми.

Търсенето в неподреден масив от данни е необходимо да бъде бързо. При класическите компютри се прилагат различни методологии за ускорение като например техниките на паралелното програмиране. При квантовите компютри е възможно квадратично ускорение чрез алгоритъма на Гроувър. Целта е да се дефинира функция за търсения елемент, която да върне резултат *истина* при успех.

Алгоритъмът използва вътрешни извиквания към „оракул“, който определя стойността на функцията. След което трябва да се намери най-малкия възможен брой извиквания на оракула, определящи кои елементи са търсените. Нека съществува база данни от $N = 2^n$ елемента, където $n = \log_2 N$ бита служат за декларацията на всеки такъв. Извършва се

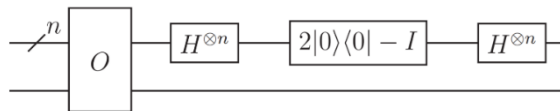
начална инициализация, като се получава равно-претегленото суперпозиционно състояние $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{x=0}^{n-1} |x\rangle$ чрез трансформация на Адамар върху $|0\rangle^{\otimes n}$. Дефинира се оператор на Гроувър G , познат още като дифузен оператор на Гроувър. Той се състои от следните стъпки:

- извикване на оракул $O = (-1)^{f(x)}$, където $f(x) = 1$, ако е решение;
- прилагане на трансформация на Адамар $H^{\otimes n}$;
- прилагане на условно фазово отместване на състоянията, т.е. тези които не са $|0\rangle$ получават фаза -1 : $|x\rangle \mapsto -(-1)^{\delta_{x0}} |x\rangle$;
- повторно прилагане на трансформацията на Адамар.

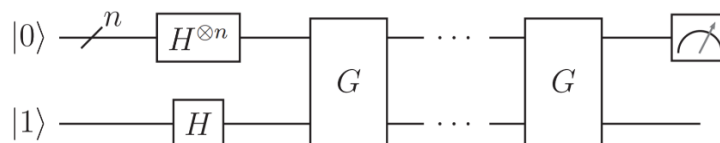
Когато стъпките са приложени заедно, се получава тъждеството:

$$(\mathcal{H}^{\otimes n} (2|0\rangle\langle 0| - I) \mathcal{H}^{\otimes n}) O = (2|\psi\rangle\langle\psi| - I) O \quad (8.1)$$

Графично представяне на тази част от алгоритъма е представено на Фигура 12, както и итерациите при прилагане на оператора $O(\sqrt{N})$ пъти:

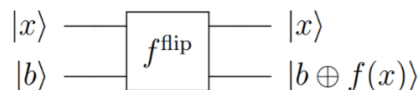


Фигура 12.1. Принципно квантова логическа схема на алгоритъм на Гроувър.



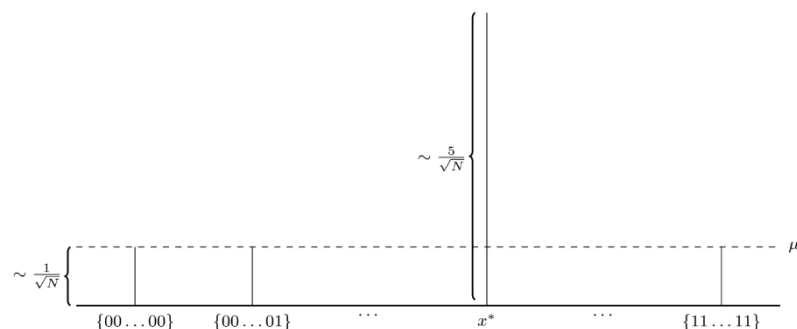
Фигура 12.2. Пълна квантова логическа схема за приложение на алгоритъма на Гроувър.

Наблюдаваното по-горе фазово отместване всъщност възниква от факта, че достъпването на квантова база данни би било еквивалентно на това при класическите компютри, но образуващият се оракул няма да бъде валиден квантов гейт. Причината се основава на факта, че този квантов гейт притежава n -битов входен интерфейс и еднобитов изходен. Освен това гейтът нито е унитарен, нито е реверсивен. Решението, което се използва основно, е третата стъпка от оператора на Гроувър. Предлага се алтернативен подход чрез f^{flip} гейт, където се използва допълнителен $|b\rangle$ бит за изход. Квантова логическа схема на тази операция е показана на Фигура 13.



Фигура 13. Принципно логическа схема на квантов f^{flip} гейт.

След няколко итерации на G може да се наблюдава следният ефект, ако се приеме, че всички амплитуди на входните стойности са предварително изразени чрез бар-графи, както е представено на Фигура 14.4.



Фигура 14.4. Видимо нарастване на амплитудата на търсения елемент и възможно доближаване до край на процеса на итерациите.

2.2.4 Квантови невронни мрежи (QNN)

От самото име на използваната техника става ясно, че това е квантов вариант на невронните мрежи – методология в информационните технологии, описваща поведението на алгоритми чрез аналогични елементи от човешката нервна система. Представено е преобразуването на неврона на Маккълох-Питс $x = \{-1, 1\}$ в т. нар. *кюрон* $|x\rangle$ в двумерното Хилбертово пространство $\mathcal{H}^2$ с базис $\{|0\rangle, |1\rangle\}$. Аналогично, състоянията $|0\rangle$ и $|1\rangle$ представляват пасивното и активното положение на неврона. Състоянието $|\psi\rangle$ на мрежа от N на брой кюрона представлява съответно квантово състояние от 2^N -мерно Хилбертово пространство и може да се представи чрез следния израз:

$$|\psi\rangle = \sum_{i=1}^{2^N} a_i |x_1 x_2 \dots x_N\rangle_i, \quad (11)$$

където a_i , $i = 1, 2^N$ сочи към комплексни амплитуди, присвоени към съответните основни състояния на мрежата. Тъй като при моделирането на една квантова система, реализираща невронна мрежа, се изисква поставянето на граница и ясно ниво на абстракция се дефинират следните минимални изисквания към квантова невронна мрежа:

- Първоначалното състояние на квантовата система съдържа произволен бинарен низ с дължина N . Квантовата невронна мрежа осигурява стабилна изходна конфигурация, съдържаща един от 2^M възможните изходни бинарни низове с дължина M , който е най-близък до входното състояние спрямо някаква метрика за разстояние.
- Квантовата невронна мрежа отразява един или повече от основните механизми за квантови изчисления, например: динамика на привличане, синаптични връзки, произвеждане на сигнал, правила за обучение, структура на невронната мрежа.
- Еволюцията се основава на квантови ефекти като суперпозиция, сплитане и интерференция, като те трябва да спазват законите на квантовата теория.

Чрез описаните по-горе изисквания и познанията за класически и квантови невронни мрежи, могат да се формулират аналозиите, представени в Таблица 1.

Таблица 1. Съпоставка между средствата на квантовата механика и невронните мрежи.

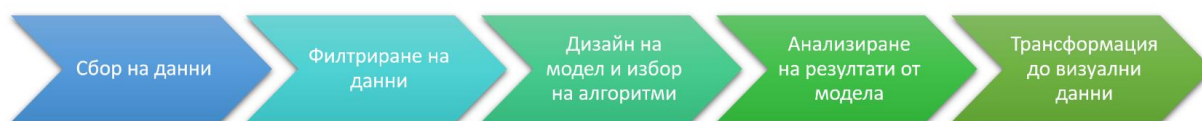
Средства на квантовата механика	Невронни мрежи
Вълнова функция	Неврони
Суперпозиция (кохерентност)	Взаимовръзки (тегловни коефициенти)
Измерване (декохерентност)	Еволюция към привличащия (атрактор)
Сплитане	Обучаващо правило
Унитарна трансформация	Усилваща функция

Първото изискване осигурява възможността на квантовата невронна мрежа да разпознава шаблони, както и да изгражда асоциативна памет. Може да се каже, че то осигурява и някои допълнителни фундаментални способности при обработката на информация. Второто изискване следи за връзка с идеята за невронно изчисление, но остава достатъчно абстрактно при необходимост от по-късна модификация. В случая е важно да се изключи употребата на други квантови алгоритми, симулиращи асоциативна памет, буквално *Quantum Associative Memories*. Третото изискване осигурява класификацията на мрежата като квантова. Имайки предвид уравнения (10), може да се каже, че входният сигнал $\sum_{j=1}^N w_{ji} x_j$ за N неврона x_j , където $j \neq i$, свързани към неврон x_i чрез тегловите коефициенти w_{ij} , бива свързан към изхода на неврон x_i чрез например, стъпкова, сигмоидна, $\tanh$ или ReLU функция, както беше описано по-горе. Важна особеност е, че при подаване на сигналите ще се разпространяват вероятностни стойности, тъй като това е ключова част от квантовата механика, и съответно няма да могат да се дават детерминистични стойности на следващите неврони. Връзката между вероятностите и нелинейните функции противоречи на основния принцип за линейна еволюция, където линейната суперпозиция при решенията на уравнението на Шрьодингер е от изключителна важност, както е описано в **Първа глава**. Ако все пак нарушаването на суперпозицията след измерване на кюронното състояние е допустимо изключение, то тогава то ще бъде спрямо един от базисните вектори $\{|0\rangle, |1\rangle\}$ с вероятност $|\alpha|^2$ или $|\beta|^2$, съответно. Въпреки това, ако бъде изграден квантов персептрон само с цел индуциране на измервания, динамиката на системата ще еволюира спрямо класическите невронни мрежи, а квантовите ефекти ще бъдат унищожени при всяка стъпка за обновяване. В този смисъл разглеждането на квантовата невронна мрежа е ключова предпоставка в сравнение с *тензорната* такава, където за *тензор* се има предвид математическото понятие за многомерен масив.

ГЛАВА 3. АНАЛИЗ И ПРОЕКТИРАНЕ НА КВАНТОВИ АЛГОРИТМИ В ИНФОРМАЦИОННИТЕ ТЕХНОЛОГИИ

Ускорението в изчисленията е неявно и математическите формализми описват единствено възможността на квантовите системи да бъдат скалируеми. Това не е задължително проблем, тъй като скоростта на изчисление в квантовия хардуер не се базира единствено на възможността на квантовите сплитания да спомагат за бързия пренос на данните, но и на факта, че кюбитите просто биват по-бързо обработвани във всичките си състояния. В текущата глава са представени методологии със съществуващи реализации, които могат да бъдат използвани в различни приложения. Важно да се отбележи е, че използваните подходи за всяко от приложенията са иновативни и представляват същински алтернативи не само като квантови способности, но и изобщо като решения с използване на машинно обучение.

Описаните квантови системи са синтезирани за целите на текущия дисертационен труд. Основните етапи на един алгоритъм за машинно обучение, представени на Фигура 20, обобщават генерализирания подход при решаване на всяка задача, използваща машинно обучение.



Фигура 20. Последователност от действия на алгоритъм за машинно обучение.

Прилагането на квантови алгоритми следва да се базира на аналогични стъпки, но не е задължително всички етапи от Фигура 20 да бъдат изпълнявани. Чрез квантовите техники

реално се цели преобразуване на етапите, които са с най-голяма изчислителна сложност. Важно предимство на алгоритмите за машинно обучение е, че в различни етапи от този конвейер могат да се включват други алгоритми например: клъстериращи алгоритми, които се използват за класификация на данни в реално време; *pre-sampling* процедури, които целят намаляване броя на обработваните по-нататък данни; сумиране на данните за получаване на първоначални групи. В последствие те могат да бъдат подадени на невронна мрежа. Единственото предизвикателство, което остава, е сложността и природата на самия квантов компютър. Разгледани са следните новосинтезирани приложни системи: (1) квантова система за препоръки на изображения; (2) квантова система за откриване на фалшиви потребителски профили; (3) квантова конволюционна невронна мрежа.

3.2 Моделиране на алгоритъм при квантова система за препоръки

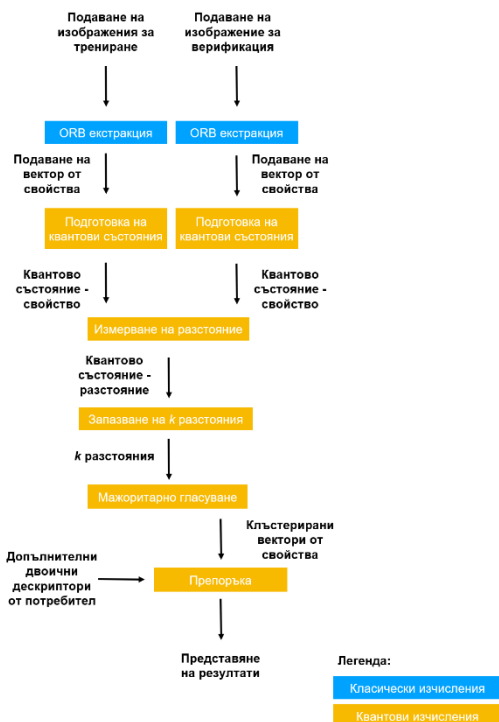
Системата за добавяне на изисквания от клиента с цел ускоряване или подобряване на изходните данни се нарича *система за препоръки*. От гледна точка на квантовата механика този подход няма много алтернативи и съвсем подходящ алгоритъм за решаване на тази задача е алгоритъмът на Гроувър – тук алгоритъмът е насочен към бързо търсене в несортирана база данни. Също така се дава и добро обяснение за природата на самата *препоръка*. Допълнителен интерес представлява и поведението на алгоритъма в контекста на текущия дисертационен труд, тъй като той е базиран на квантови сплитания. Квантовата система за препоръки е представител на *хибридния* клас квантови алгоритми, тъй като съдържа в себе си оптимизационна задача от вида QAOA, техника чрез усилване на амплитудата, подаването на самите входни данни е чрез класически подход, а самият казус може да се приеме за NP-строг. Подзадачите на квантовата система за препоръки са:

- 1) екстракция на ORB дескриптори от множество входни изображения;
- 2) клъстериране на резултатите от 1) чрез квантова версия на *k*-majority алгоритъм, където за измерване на разстояние се използва методът на Жакард-Нийдхам;
- 3) използване на формираните клъстери за откриване на съвпадения в изображения;
- 4) използване на алгоритъма на Гроувър за ускорение чрез препоръка.

Подходът е приложим за различни класификационни задачи, за които са в сила следните особености и се поставят следните изисквания: (1) данните трябва да бъдат представени като вектори, т.е. динамични масиви от двоични стойности; (2) измеренията на данните в случая няма да влияят на скоростта на алгоритъма, защото те реално не съществуват – всяко бинарно число не притежава явни свойства освен стойността си; (3) данните могат да се групират и чрез алгоритми като факторизационна матрица или обучение в дълбочина, но тези алгоритми могат да доведат до по-висока степен на сложност сами по себе си, без тя да е оправдана. Системният модел на квантова система за препоръки чрез свойства от двоични дескриптори е показана на Фигура 21. Не се изисква разглежданата квантова система да подготвя ORB дескрипторите и да ги извлича от изображенията. Приема се, че те ще бъдат подготвени преди това и се предоставят на системата като поток от данни. Тъй като това е вид сурови, необработени данни, няма значение дали са подготвяни от класически компютър или квантов такъв. Системата извършва клъстериране на данните, на базата на което се прави сравнение и класифициране на нови изображения. Стъпката за препоръка или още казано усилване на характеристиките може да доведе до промяна на крайния резултат. Целта е намаляване броя извиквания на клъстериращата техника преди конвергенция, като тук се помага при самото класифициране да се постави новото изображение в правилния клас. Тъй като в основата си това е клъстериращ алгоритъм, се очаква, че новият обект става част от избрания си клъстер и се предизвиква преизчисление, за да се обнови центроида.

За да могат да се извършат квантовите побитови операции са необходими още преобразувания на класическия алгоритъм. Предлага се мажоритарен избор, именно с целите на доближаване до битова репрезентация при измерване на разстоянието или приликата между входните данни. За всеки клъстер *S* се създава т. нар. акумулиращ вектор,

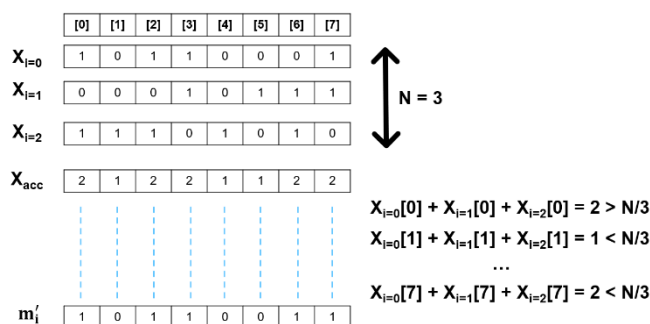
чиято цел е да съхранява броя пъти, където даден бит от всички елементи x на клъстера е бил 1.



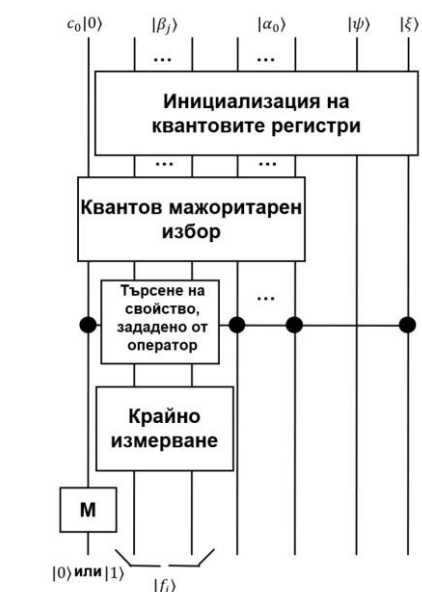
Фигура 21. Системен модел на квантова система за препоръки чрез свойства от двоични дескриптори.

Фигура 24. Примерна квантова логическа схема на предложения реализирания алгоритъм за клъстериране на двоични вектори от дескриптори.

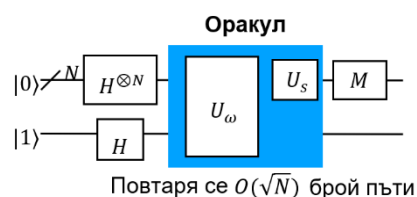
Резултатът от клъстеризацията се използва за изграждането на нов центроид m' , където за всеки елемент x_i в x , ако множеството бинарни вектори „гласуват“, че даден бит е 1, то тогава m'_i става 1. В противен случай m'_i приема стойност 0. С цел онагледяване на този процес може да бъде разгледан и следният пример, показан на Фигура 23. Примерна квантова логическа схема на предложената техника е визуализирана на Фигура 24.



Фигура 23. Примерно изчисление на нов центроид чрез мажоритарен избор.



Фигура 24. Примерна квантова логическа схема на предложения реализирания алгоритъм за клъстериране на двоични вектори от дескриптори.



Фигура 25. Примерна квантова логическа схема за извършване на измерване при препоръка чрез намиране на сходен елемент с алгоритма на Гроувър.

Използваните регистри са следните:

$c_0|0\rangle$ – спомагателен квантов бит, където полученото вероятно разпределение от клъстерираните елементи може да бъде усилено чрез алгоритъма на Гроувър, за да се подсилят вероятностните характеристики, стига да са подадени такива от потребителя. Ако

c_0 измерва вероятност $|1\rangle$, то трябва да се преизчисли стъпката за мажоритарен избор, като опита с друга стойност на k ;
 $|\beta_i\rangle$ – описва състоянията от регистрите на свойствените вектори на всяко тренировъчно изображение;
 $|\alpha_0\rangle$ – описва състоянията от регистрите на свойствените вектори на тестовото изображение;
 $|\psi\rangle$ – състоянието, съдържащо идентификатор на препоръчания ORB дескриптор;
 $|\xi\rangle$ – регистър, представляващ допълнителни данни, зададени от оператор;
 $|f_i\rangle$ – краен резултат - препоръка за присвояване на дадено изображение към определен клъстер.

С цел реализиране на квантова система за препоръки на базата на описаните основни средства и градивни блокове от предложения системен модел в реална система или дори квантова симулационна среда, може да се използва следният псевдо-код (с цел улеснение $\vec{x}$ се приема като променлива, представляваща произволен дескриптор)¹:

1. Екстракция на ORB дескрипторите от тестовото изображение и тренировъчните такива.
2. Генериране на случаен k -брой двоични центроиди C
3. **while** C not changed **do**:

$\vec{x} \in \vec{x}$:
 $c_x \leftarrow \operatorname{argmin}(d(c, x)), c \in C$
 $c \in C$:
 $\vec{x} \in \vec{x}|_{c_x=c}$
 v accumulates x votes
 $c' \leftarrow \operatorname{Majority}(v)$,
 where c' is the new centroid
end while
4. Изпълнение алгоритъма на Гроувър върху получените дотук клъстери, използвайки допълнителни ORB дескриптори, зададени от оператор.
5. Извеждане на резултати.

Блокът *крайно измерване*, където участва и алгоритъмът на Гроувър, е показан на Фигура 25.

3.3.3 Оценка на качеството на получените клъстери

Разглежда се квантов алгоритъм за откриване на шаблони в динамични редове. След като вече е установен методът за клъстериране, може да се премине към следващ етап, а именно представяне на резултатите или качествена оценка на получените клъстери. Втората техника се използва рядко в системите, занимаващи се с динамични редове, защото потокът данни се приема за голям. Това означава, че дори и оптимизирани, алгоритмите за клъстериране остават изчислително сложни и системата не може лесно да се самоверифицира, особено когато времето за реакция е важно. Въпреки това, ако се приеме, че квантовото представяне на тези методологии би довело до значително ускорение, то се предполага и че стъпката за самооценяване, също може да бъде ускорена по подобен начин, като се запази като част от алгоритъма, тъй като реално няма да се оказва мястото, където е възможно да настъпи допълнително забавяне.

¹ Тъй като това е псевдо-код, доближаващ реализацията до истинска имплементация, част от текстовите полета в него са оставени на английски.

Друга важна роля при оценка на качеството на получените клъстери е конвергенция на алгоритъма, т.е. или той приключва до определен брой итерации, или докато не се наблюдават промени в състава на клъстерите. Възможно е нито едно от двете условия да не е приемливо при динамичните редове, затова се налага трети вариант с използване на качествената оценка, а именно чрез т. нар. *методът Силует*. Неговият начин на работа може да се дефинира просто чрез следната система:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i), \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases} \quad (39)$$

където a е средно-статистическото разстояние от наблюдаемата i до точките от клъстера, на който принадлежи. b е минималното средно-статистическо разстояние на i спрямо точките от друг произволен клъстер. С други думи, b е средно-статистическото несъответствие между точка i и точките на най-близкия клъстер, до нейния клъстер. Използвайки метриката за разстояние между два квантови бита, която се нарича *прецизност (fidelity)*, за n -мерност на текущия за итерацията клъстер може да се използва следният израз:

$$a(i) = \frac{1}{n_c} \sum_{j \in c, j=1}^{n_c} \langle F(\rho, \rho_j') \rangle_{\rho_0}, \quad (40.1)$$

$$b(i) = \min_s \frac{1}{s} \sum_{c=1}^s \frac{1}{n_s} \sum_{i \notin S_c, j=1}^{n_s} \langle F(\rho, \rho_j') \rangle_{\rho_0}, \quad (40.2)$$

като близостта на двете квантови състояния се маркира с: $\langle F \rangle = \langle F(\rho, \rho') \rangle_{\rho_0}$, където ρ и ρ' са матриците за плътността на i -тата променлива, както и всички останали, съответно при първоначални състояния ρ_0 . Допълнително, s е броят на клъстерите, а с s се бележи индекса на текущия клъстер. n_s е броят на елементи в клъстера S , за който i не принадлежи на него, т.е. S_c е в контекста на b :

$$s(i) = \frac{b(i)-a(i)}{\max(a(i), b(i))}, \text{ за } \forall i \in \vec{x} \quad (41)$$

3.3.4 Употреба в система за засичане на фалшиви потребителски профили

В случая трябва да се дефинират някои специфични понятия в контекста на предложената технология. Тъй като се има предвид клъстериране на потребителски профили, според които ще се смята вероятността някой да бъде фалшив, трябва да се поясни, че за целите на експеримента в **Петта глава** за *име* се има предвид случайно-генерирано такова, изписано с латински символи. Съкратено представено, след кратък преглед на статистиката за български имена, става ясно, че може да се изгради шаблон, който да отговаря на реалните такива. Параметризацията е следната: уникалните букви в името трябва да са 35-45 %, а имената, изписани с латинска азбука варират до 10 букви като дължина; към това, гласните букви са средно между 45 % и 55 % от думата. Тази информация може да се използва като едно от измеренията при клъстериране. Допълнително към тези вероятности, ще бъдат назначени и брой имена, които се считат за производни на друго или автоматично-засечени, че имат общ корен. Важно уточнение е, че тук не се търси реална връзка между самите имена, а по-скоро съвпадение в шаблоните им на записване, например “George” и “Georgi”. *Разстоянието* между два сравнявани низа s_1 и s_2 се определя с *формулата на Джаро*:

$$\text{sim}_{\text{Jaro}}(s_1, s_2) = \frac{1}{3} \left(\frac{c}{|s_1|} + \frac{c}{|s_2|} + \frac{c-t}{c} \right), \quad (42.1)$$

където sim_{Jaro} е функцията, измерваща близостта на два низа, c е броя съвпадащи символи, а t е минималния брой такива размествания на единични символи от името, които биха довели до по-голяма близост на самите низове. Фигура 26 показва това нагледно.

S1	P	A	U	L
S2	P	U	A	L

Фигура 26. Примерна демонстрация на начина, по който алгоритъмът на Джаро пресмята своите коефициенти.

В примера на Фигура 26 може да се види, че $c = 4$ и $t = \frac{2}{2} = 1$.

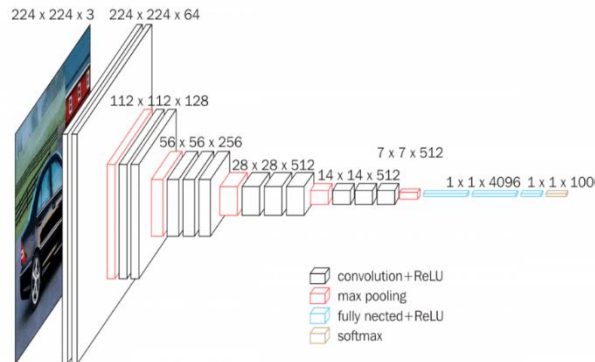
Известно подобрение на този алгоритъм е т. нар. *алгоритъм на Джаро-Винклер*. Той се представя със следната формула:

$$sim_{Wink}(s_1, s_2) = \begin{cases} sim_{Jaro}(s_1, s_2), & \text{когато } sim_{Jaro}(s_1, s_2) < 0.7 \\ sim_{Jaro}(s_1, s_2) + \frac{l}{10} \cdot (1 - sim_{Jaro}(s_1, s_2)) \end{cases}, \quad (42.2)$$

където l е дължината на общия префикс на двата низа, като максималният позволен префикс е с дължина от четири символа. Очевидно е, че формулата на Джаро-Винклер се базира на тази на Джаро, но е важно да се отбележи, че нейният резултат ще се влияе силно, когато думите започват еднакво. Тази метрика за близост до други низове също ще бъде включена като измерение на потребителските имена и ще участва в клъстериращия алгоритъм като мярка за сходство. Следователно крайната форма на системата ще използва за обучение предварително изчислените вероятност и *разстояния* между низовете като входни данни.

3.4.1 Изграждане на квантова конволюционна невронна мрежа

Предлага се конволюционна невронна мрежа за откриване на ретинални заболявания, базирана на VGG16 архитектурата (Фигура 27).



Фигура 27. Примерна архитектура на техниката VGG16.

Нейната цел е да идентифицира следните класове заболявания като използва изображения за входни данни: (1) CNV – хороидална неоваскуларизация; (2) DME – диабетен макулен едем; (3) DRUSEN – друзи; (4) NORMAL – нормална ретина; (5) CSR – централна серозна ретинопатия; (6) MH – макулна дупка. Конволюционните слоеве пресмятат нови пикселови стойности x_{ij}^N чрез линейна комбинация от обкръжаващите ги подобни в предхождащите слоеве: $x_{ij}^N = \sum_{a,b=1}^w w_{a,b} x_{i+a,j+b}^{(N-1)}$, където теглата $w_{a,b}$ образуват матрица с размерност $w \times w$. Обединяващите слоеве редуцират размера на масива от свойства, премахвайки максималните стойности от няколко съседни пиксели, а след тях се изпълнява нелинейната функция *ReLU*. Един от начините за регулиране на теглата и напълно-свързващата функция е чрез употребата на огромни множества от данни за обучение. Тъй като архитектурата на желаната система е установена и по-горе бяха описани квантовите операции, необходими за изграждането ѝ, следва да се подготвят нужните формализми. Нека входът на такава квантова логическа схема е непознато квантово състояние ρ_{in} . Конволюционният слой изпълнява квази-локална унитарност U_i с цел

крайна дълбочина. При обединение се измерват част от набора кубити, а съответните стойности на измерването определят унитарни ротации V_j , приложени върху съседните кубити. Съответно, нелинейността на тази невронна мрежа ще нараства с намаляването на степените на свобода. Последователността от конволюционни и обединяващи слоеве, както се описва по-горе, продължава, докато размерът на анализираната система не намалее достатъчно. След това се прилага напълно-свързан слой F като унитарност на останалите квантови битове. Накрая се измерва определен брой от изходни квантови битове. В последния слой на квантовата невронна мрежа се изпълнява функцията *softmax*, която измерва вероятността дадена частица да бъде открита в определено квантово енергийно състояние, когато, например, атом е част от множество елементи, достигнали температурно равновесие при температура T . Тази функцията се нарича също *разпределение на Гибс* или *разпределение на Болцман*. Както стана ясно във **Втора глава** симулацията на квантова система е особена задача и изпълнението на функцията *softmax* е част от този проблем. Нека съществува система с Хамилтониан $\mathcal{H}$ и нека тя контактува с околната среда при температура T . След някакъв период от време системата ще изпадне в температурно равновесие, което се нарича състояние на Гибс: $\rho_{\text{therm}} = e^{-\mathcal{H}/k_B T} / \mathcal{Z}$, където k_B е константата на Болцман и $\mathcal{Z} = \text{tr} e^{-\mathcal{H}/k_B T}$ – стандартната функция за нормализиране при делене, която гарантира $\text{tr}(\rho) = 1$. Процесът, по който се достига равновесието, все още не е добре изучен. Единствено се задават изисквания за околната среда. Тя трябва да бъде голяма, трябва да има ненулева популация от състояния с енергии, отговарящи на собствените състояния на $\mathcal{H}$, а връзката със системата да бъде слаба. Получаването на ρ_{therm} за случайни $\mathcal{H}$ и T е експоненциално сложен проблем за класически компютър. Все още не е ясно дали квантов компютър би могъл да го реши по-бързо, точно както е описано в предходната глава за VQR-проблемите.

Връщайки се към изграждането на квантовата невронна мрежа, трябва да се има предвид, че класификацията на N -кубитни състояния се описва чрез $O(\log(N))$ параметъра, което е поне двойно намаляване спрямо предишните две предложени системи. Пример за това е пресмятането на целевата функция чрез средноквадратична грешка: $(\vec{w}, \vec{b}) = \frac{1}{2n} \sum_{\vec{x}} |y(\vec{x}) - \vec{a}|^2$. При квантовото представяне при данни за обучение $\{(|\psi_\alpha\rangle, y_\alpha) : \alpha = 1, \dots, M\}$, където $|\psi_\alpha\rangle$ са входните състояния, а $y_\alpha = 0$ или 1 са съответните бинарни резултати от класификацията. Тогава средноквадратичната грешка може да се представи като:

$$C = \frac{1}{2M} \sum_{\alpha=1}^M \left(y_i - f_{\{U_i, V_j, F\}}(|\psi_\alpha\rangle) \right)^2, \quad (43)$$

където $f_{\{U_i, V_j, F\}}(|\psi_\alpha\rangle)$ описва очакваният изходен резултат за вход $|\psi_\alpha\rangle$ на квантовата конволюционна невронна мрежа. Алтернатива би била т. нар. кръстосана ентропия или крос-ентропия:

$$\mathcal{L} = \sum_i \sum_{k \in \{0,1\}} \left(f_{\{U_i, V_j, F\}}(|\psi_\alpha\rangle) \right)_k \log y_{ik}. \quad (44)$$

Този израз всъщност е крос-ентропия на две вероятностни разпределения. Функцията *softmax* предоставя именно такова вероятностно разпределение. Може да се счете, че тя е по-добър вариант от (43), защото от вероятностна гледна точка крос-ентропията расте еднакво с целевата функция.

За целите на имплементацията могат да се използват готови модели на квантови логически схеми. Единият е *анзац на многомащабни ренормализации на сплитанията (AMPC)*, а другият – *квантова корекция на грешката*. Рамката на AMPC предлага ефективен модел на тензорна мрежа с много класове. Може да се обясни като квантово състояние, произведено от последователност на унитарни и изометрични слоеве, и

Diagram illustrating a deep convolutional neural network architecture. The input is a vector p_{in} . The network consists of alternating convolutional layers (blue) and pooling layers (orange). The first convolutional layer has nodes U_1 . The first pooling layer has nodes V_1 . The second convolutional layer has nodes U_2 . The second pooling layer has nodes V_2 . The final convolutional layer has nodes U_2 . The final pooling layer has nodes V_2 . The output of the final pooling layer is connected to a fully connected layer (green) with node F . The final layer is a softmax layer (purple) with node $Gibbs^*$. The output of the softmax layer is connected to a vector of nodes: CN, DM, DR, NO, CS, MR.

The diagram illustrates a quantum communication channel. On the left, an input state $|\psi_I\rangle$ is combined with three ancilla qubits in the $|0\rangle$ state. These enter an encoding unit U_2^{-1} . The output of U_2^{-1} is a 4-qubit state that enters a block labeled "Кодiranje" (Encoding). This block contains three parallel units, each consisting of a unitary U_1^{-1} followed by a measurement (represented by a meter symbol). The outputs of these measurements are classical bits. These classical bits, along with the 4-qubit state, enter a block labeled "Декодiranje" (Decoding). This block contains three parallel units, each consisting of a unitary U_1 followed by a measurement. The outputs of these measurements are classical bits. Finally, the 4-qubit state and these classical bits enter a decoding unit U_2 , which produces the output state ρ .

3.4.2 Обучение на квантовата конволюционна невронна мрежа

$$S_{qb} = Z_{q-1} X_q X_{q+1} \dots X_{b+1} X_b X_{b-1} Z_b, \quad (45)$$

18

вероятност $p = (\langle \psi_{in} | S | \psi_{in} \rangle + 1) / 2$ и с вероятност 1 при $(1 - p)$, тъй като $S^2 = 1$. Важно да се отбележи е, че сложността на извадката на SOP скалира независимо от дължината на низа, докато сложността на извадката на квантовата невронна мрежа постоянно се подобрява, увеличавайки дълбочината.

Демонстрирайки начина на работа на квантовата невронна мрежа, трябва да се посочат и стойностите на хиперпараметрите. Както може да се види от архитектурата на VGG16 (Фигура 28), се създават 13 конволюционни слоя, 5 обединяващи слоя, 3 напълно-свързани слоя и един *softmax*-слой. От съществено значение е да се отбележи, че след втория обединяващ слой конволюционните слоеве биват групирани по три един след друг, а не по два, както е в началото. Тази структура се характеризира също и с числото $n = 2$ (от 224 пиксела стават 112 и т.н.), определящо редукцията на системния размер при всяка

конволюционно-обединяваща конструкция, т.е. $L \rightarrow \frac{L}{n}$. Първият конволюционен слой

изисква $(n + 1)$ -кюбитови унитарни матрици, започващи на всеки n -ти кюбит. Това бива следвано от n слоя на n -кюбитови унитарни матрици, подредени според архитектурната диаграма по-горе. Първият обединяващ слой измерва $(n - 1)$ от всеки близък блок от n кюбита. Всяка от тези операции се асоциира с унитарна матрица V_j , приложена върху оставащия квантов бит, в зависимост резултата от измерването. Тези конструкции се извикват d на брой пъти, където d е дълбочината на мрежата. Напълно-свързаните слоеве съдържат произволни унитарни матрици, приложени към последните останали N / n^d квантови бита, където N са предварително-зададените спинове. Поради невъзможността за реализация на такава система, може да се запази използването на уравнение (23), но при задължителното задоволяване на следните две изисквания [31]:

- Критерий за фиксирана точка – ако входът е силно-сплетено състояние² $|\psi_0\rangle$, изградено от L спина, резултатът от конволюционно-обединяващите слоеве е силно-сплетено състояние $|\psi_0\rangle$ от $L/3$ спина, като всички измервания детерминистично клонят към $|0\rangle$.
- Критерий за корекция на квантовата грешка – ако входните данни не са $|\psi_0\rangle$, но се различават от $|\psi_0\rangle$ с грешка, която може да се замени с глобалната симетрия, резултатът би трябвало все още да бъде силно-сплетено състояние от $L/3$ спина, но поне едно от измерванията ще бъде поставено в $|1\rangle$.

3.4.3 Допълнителни забележки към корекция на грешката

Квантовата конволюционна невронна мрежа, както и нейното обратно изпълнение-Backpropagation, може да се разглежда директно като декодиращ и кодиращ квантов канал между физическите входни квантови битове и логическия изходен квантов бит [89]. Кодиращата схема дефинира нови множества от кюбити в инициализирано състояние $|0\rangle$, докато декодиращата изпълнява измерванията. Примерна квантова логическа схема, представяща право - Feed Forward, и обратно изпълнение на конволюционната невронна мрежа, е представена на Фигура 29.

Нека е даден канал за грешки N и целта на задачата да бъде максимизиране на възстановяващата прецизност, описана чрез:

$$f_q = \sum_{|\psi_l\rangle \in |\pm x, y, z\rangle} \langle \psi_l | \mathcal{M}_q^{-1}(\mathcal{N}(\mathcal{M}_q(|\psi_l\rangle\langle\psi_l|))) | \psi_l \rangle, \quad (46)$$

² Бел. прев. от англ. Cluster state, lattice of qubits – в английската литература се използва въпросния термин в смисъла на силно-сплетено състояние на множество кюбити.

където $\mathcal{M}_q(\mathcal{M}_q^{-1})$ е кодиращата, съответно декодиращата техника, генерирана от невронната мрежа, а $|\pm x, y, z\rangle$ са ± 1 собствените състояния на матриците на Паули. Уравнението (46) се базира на уравненията в система (6) и анализа за прецизност на гейта в 2.2.1. На Фигура 29 е представена невронна мрежа, за чието функциониране са необходими девет кубита, реализиращи всички унитарни операции $U_1^{\pm 1}$, които са осем на брой, и N . Аналогично могат да се пресметнат и необходимите кубити за анализираната тук квантова система. Към всичко това трябва да се има предвид и примерният модел на грешката, който най-общо може да се представи като приложение на квантов канал за деполяризация на всеки от физическите кубити – за примера по-горе това са $i \in \{1, 2, \dots, 9\}$:

$$\mathcal{N}_{1,i}: \rho \mapsto (1 - \sum_{\mu} p_{\mu})\rho + \sum_{\mu} p_{\mu} \sigma_i^{\mu} \rho \sigma_i^{\mu}, \quad (47)$$

където σ_i^{μ} са матриците на Паули, p_{μ} са вероятностите на всеки отделен кубит за $X, Y, Z = \mu$ грешки. След извършен подробен анализ, в [31] са предложени независими едно-кубитови анизотропни грешки с допълнителни дву-кубитови корелирани грешки $X_i X_{i+1}$ с вероятност p_{xx} . Добавяйки това към (47), следва:

$$\mathcal{N}_{2,i}: \rho \mapsto (1 - p_{xx})\rho + p_{xx} X_i X_{i+1} \rho X_i X_{i+1}, \quad (48)$$

за двойки близки кубити с индекси $i \in \{1, 2, 4, 5, 7, 8\}$.

Обобщено, цели се максимизиране на (26), но ако се приложи директно метода на градиентното спускане, базирайки се на f_q за U_1 и U_2 , се рискува достигане до локален оптимум. Вместо това, двете унитарни матрици могат да се оптимизират последователно, както се прави чрез *метода за обратно разпространение на стойността на грешката*³ при класическите конволюционни невронни мрежи. Допълнително, заради дървовидната структура на невронната мрежа, първият слой може да се представи като независим квантов канал:

$$\mathcal{M}_{U_1}: \rho \mapsto \text{tr}_a[U_1 \mathcal{N}(U_1^{\dagger}(|0\rangle \otimes \langle 0| \otimes \rho \otimes |0\rangle \langle 0|) U_1) U_1^{\dagger}], \quad (49)$$

където $\text{tr}_a[\]$ представя трасирането върху спомагателните кубити, които биват измервани в междинните стъпки. От тази гледна точка $\mathcal{M}_{U_1}$ е достатъчен модел за грешка, който може да бъде коригиран от втория слой.

ГЛАВА 4. АНАЛИЗ И РЕАЛИЗАЦИЯ НА КОМПОНЕНТИ ЗА ТЕХНОЛОГИЧНО ПОДПОМАГАНЕ НА ПРОЦЕСА ПО МАШИННО ОБУЧЕНИЕ В КВАНТОВО-СПЛЕТЕНА СРЕДА

В настоящата глава е извършен обзор на различни начини за изграждане на квантови системи за машинно обучение. Направен е и анализ на възможните реализации на различни квантови подходи, като се обръща внимание и на невъзможността някои от тях да бъдат изградени с описаните инструменти, тъй като математическият апарат не обяснява тяхната постановка в квантова среда. Основна част от проблема се състои в това, че все още не съществува понятието *квантов компютър за обща употреба*. Всяка една от частите на една примерна квантова система има собствена реализация и в общия случай тези градивни блокове не могат да се съберат чисто физически в единна система. Има се предвид и консумацията на оперативна памет при класически компютър, имитиращ квантов, защото необходимите клетки памет растат експоненциално с увеличаване на броя кубити в системата. За примери като квантови невронни мрежи това се оказва граница, щом става въпрос за тяхното скалиране. Към това не бива да се пропуска и факта, че реализацията дори на един квантов алгоритъм, участващ в съставен такъв, съответно, може да има висока цена не само като сума, но и като усилие. Конструирането дори на прости алгоритми и

³ Бел. прев. от англ. Backpropagation.

логически гейтове е сложно извън специализирани лабораторни условия. Една от причините за това е дотук изтъкнатата нестабилност на квантовия компютър и това, че може да бъде дестабилизиран от различни фактори – като например незначителна промяна в температурата, което не е проблем за класическия компютър. Също така и достигането на температура близка до 0 K е голямо предизвикателство.

Тези доводи не бива да обезсмислят допълнителния математически анализ, залежал в квантовите алгоритми, както и не бива да спира нови разработки в реализацията на различните техники. От предходните глави се вижда също, че квантовите алгоритми могат да предоставят решения, които не са изпълними на класическите компютри. Подходът за решаване на проблеми в контекста на квантовите изчисления е всъщност възможност за открития в сферата на информационните технологии изобщо, включително и приложения, за които досега не се е и предполагало. Единствено остава затруднението, че извън системите за изкуствен интелект, софтуерните разработки не се интересуват все още толкова много от квантовите изчисления. А от хардуерна гледна точка, може да се направи обзор на съществуващите модели на квантови изчислителни машини, които се дефинират като фундаментални в създаването на квантов компютър. Тъй като те са насочени към решаване на един и същ проблем, чисто формално те могат да се приемат за еквивалентни един на друг, т.е. всеки един от моделите в Таблица 3 може да симулира някой от останалите, с изключение на квантовата машина на Тюринг, която все още не съществува като реализация:

Таблица 3. Съществуващи квантови изчислителни модели.

Квантов модел	Определение
Масив от квантови гейтове	Декомпозиция на изчислението на последователност от няколко кубитови гейтове.
Еднопосочен квантов компютър	Декомпозиция на изчислението до последователност от едно-кубитови измервания, приложени върху силно-сплетено начално състояние.
Адиабатен квантов компютър	Базира се на квантово каляване, описано в Първа глава . В случая представлява декомпозиция на изчислението към бавна, продължителна трансформация на началния Хамилтониан към краен такъв, чиито основни състояния съдържат решението на задачата.
Топологичен квантов компютър	Изчислението се декомпозира, включвайки се в конструкцията от анони ⁴ , изграждаща двумерна мрежа.
Квантова машина на Тюринг	Теоретично необходима, но няма пряка имплементация за момента.

4.1. Физическа реализация на квантови системи

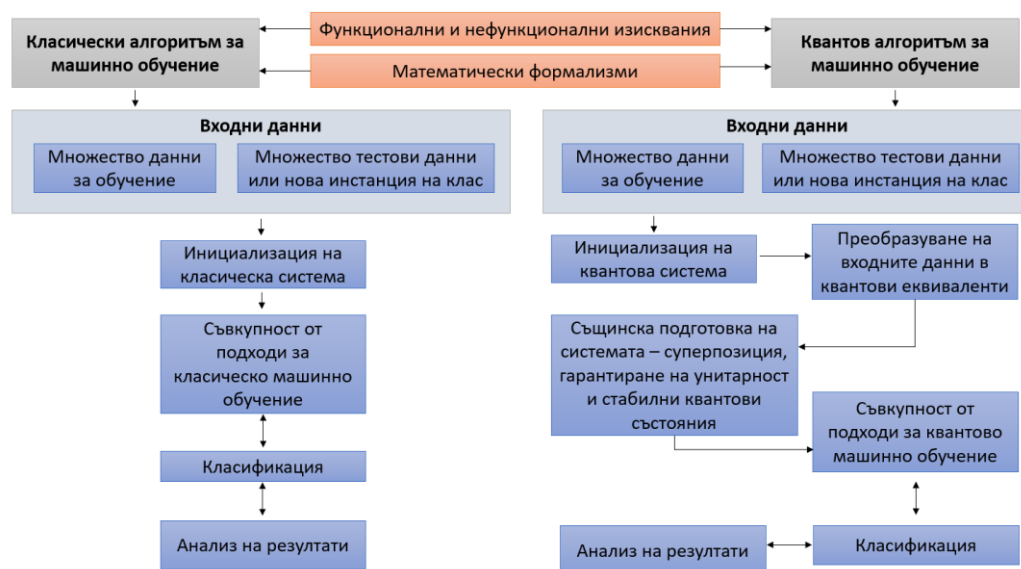
Цифровите квантови изчисления се основават на квантови логически гейтове, които както стана ясно, са основните градивни единици на квантовите логически схеми, и напълно се покриват като смисъл с класическите логически схеми. За двата подхода са необходими квантови битове, наречени още кубитове. Определят се следните функционални изисквания: (1) скалируемост с цел увеличаване броя кубити; (2) кубитите могат да се

⁴ Бел. прев. от англ. Ануон – квази-частица, съществуваща само в двумерните системи. Не бива да се бърка с анион (anion).

инициализират с произволни стойности; (3) квантовите гейтове трябва да са по-бързи от времето за декохеренция – с други думи, пълна изолация от околната среда; (4) универсално множество от гейтове; (5) кубитите трябва да се прочитат лесно.

4.2 Софтуерна реализация на квантови системи и симулационни среди

На база на подробния преглед на съществуващите квантови системи, спомагащи изпълнението на описаните по-горе функционални изисквания, се вижда, че съществува значителен избор измежду тях. Към това се извършва и детайлен обзор върху множеството езици за програмиране, библиотеки, алгоритми, моделиране и симулации, позволяващи имплементацията на разнообразни квантови системи. Тяхната цел е да участват в различните етапи на реализацията на алгоритми за машинно обучение, подобно на класическата такава. Стъпките, които трябва да бъдат изпълнени на ниво системен модел, са представени на Фигура 30.



Фигура 30. Съпоставка на подходите за изграждане на системи за машинно обучение при класически и квантови изчислителни системи.

На базата на информацията от Фигура 30 може да се направи извод, че фундаментално не съществуват особени разлики между двата подхода на изграждане на системи за машинно обучение, но квантовият изисква повече подготовка. Понякога едни и същи математически модели могат да описват стъпките на двата подхода. На ниво системни изисквания се очаква, че и двата ще изпълняват еднакви функционални изисквания, но ще притежават различни нефункционални такива, например когато се дефинира времето до конвергенция или стойностите на входните параметри. След изпълнението на алгоритъма и при двата случая е необходим анализ от оператор на системата. Единствено, но значително, ограничение остава хардуерната реализация. В резултат на направения обзор върху средствата за хардуерна реализация на квантови логически схеми или модели на алгоритми, както и анализа на съществуващите софтуерни системи за квантови изчисления, могат да се обобщят следните изводи:

- Езикът за програмиране Python е основополагащ при дизайна на квантова система, защото дава възможност за употреба на библиотеки за математическо моделиране на квантови системи, притежава разнообразни инструменти за измерване и визуализация, а като най-важна предпоставка може да дефинира неговия интерфейс до квантови облачни среди.
- Използването на реална квантова хардуерна постановка е непреодолимо ограничение към този момент.
- Употребата на продукти на IBM при софтуерните квантови среди е обоснована, след като поддръжката им продължава, и инструментите, които предлагат, биват актуализирани.

Продуктите на D-wave могат да се приемат за скъпо струващи като цена, а тези на Microsoft не се актуализират напълно.

- Останалите квантови симулационни среди не са от полза за текущия дисертационен труд, защото комбинацията от езика Python и продуктите на IBM са напълно достатъчни за реализация на описаните в Трета глава системи.

ГЛАВА 5. ПЛАНИРАНЕ, ПРОВЕЖДАНЕ И АНАЛИЗ НА РЕЗУЛТАТИТЕ ОТ СИМУЛИРАНИ ЕКСПЕРИМЕНТИ

В резултат на проведените експерименти беше установено, че използването на същински квантови логически схеми е непреодолима на този етап трудност, защото изисква скъпи и трудно достъпни лабораторни постановки. Квантовите системи, с изключение на транзисторно-базираните, изискват силно специализиран персонал, особено в сферите на експериментална физика и свръх проводниците. CMOS-базираните квантови изчислителни системи изглеждат като подходящ хардуер за реализация на част от основните квантови алгоритми, но тъй като тяхното прототипно съществуване е от скоро – по-малко от година, достъпът до такива към този етап е невъзможен. Следователно долуизброените експерименти се базират върху работа със симулационни среди. Във връзка с употребата на квантови сплитания трябва да се има предвид и следния факт: Техните конструкции и трансформации са части от фундаменталните статични ресурси в квантовата механика и като такива имат свойства, коренно различни от понятията в класическата теория на информацията, съответно не могат да се дефинират като напълно определени по природа. За тяхното поведение се съди от обобщени наблюдения и доказани експерименти. От тази гледна точка се поставят следните обстоятелства: (1) създаването и употребата на квантовото сплитане е прост динамичен процес в квантовата теория на информацията, т.е. колко квантови бита са необходими на два модула, за да обменят информация, изграждайки сплитане помежду си, при условие, че не са имали такова преди този момент; (2) преобразуването на квантовото сплитане от една форма в друга, е също динамичен процес, който трябва да определи как това ще се реализира без реално двете страни, т.е. въпросните два модула, да си комуникират – ако е през квантов информационен канал, тогава какъв капацитет трябва да има той. Присъствието на квантовите сплитания няма как да се пренебрегне, обаче, защото е основополагащо при изграждането на комуникацията между слоевете на квантовата конволюционна мрежа.

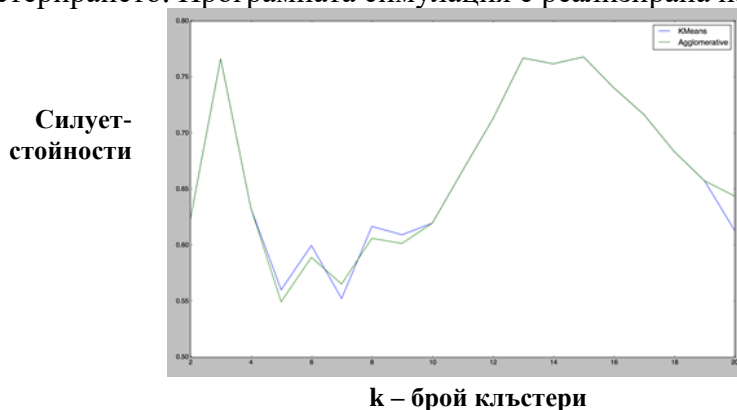
5.1 Експеримент №1 – квантови клъстериращи алгоритми

За да може да се пристъпи към експериментите, трябва да се определи ползата от употребата на квантовите клъстериращи алгоритми. Причината за това е, че те са фундаментални техники за предложените за реализация системи. Основно и значително затруднение може да се отбележи при достигане на стъпката от съставния квантов алгоритъм, изпълняваща именно клъстерирането. Трябва да се уточни, че точно тя е главната причина за това да се търси квантов еквивалент на подобен тип алгоритми, тъй като може да отнеме значително време за изпълнение, което се потвърждава и от резултатите както в този, така и в следващите експерименти. *K-means* в класическата си форма е с асимптотична сложност $O(Ndk)$, което дори за малки стойности на k или d е със значителна тежест, тъй като N е броят на входните данни, а d – броя техните измерения. При агломеративното клъстериране може да се каже, че сложността на алгоритъма по време клони към $O(n^3)$, а по-оптимизирани варианти се стремят да намалят сложността до $O(n^2 \log(n))$. Ако се приеме, че се използва квантова оперативна памет, тогава следва и да се получи т. нар. *квантово паралелно изчисление*, различаващо се по природа от

паралелното изчисление при класическите компютри, но позволяващо ускорение от типа $O(\log(Ndk))$ или $O(\log n^3)$.

Допълнително, когато се погледне самата природа на йерархичния агломеративен алгоритъм, става ясно, че той няма как да гарантира намирането на оптимално решение на задачата. И това може да се отчете като негово ограничение.

За входни данни на този експеримент са използвани множества от консуматори на електроенергия в различни сгради в различни часове на денонощието. Отбелязва се, че и двата алгоритъма могат да се справят с това да открият шаблон в несортиран масив от данни. Текущо избраните принадлежности на входен елемент към даден клъстер означават, че е намерен шаблон в тенденцията на данните, към които този елемент спада с най-голяма вероятност. Пример за това е отчитане на непрекъснато покачваща се консумация. Основни разлики обаче се наблюдават при верността на самите алгоритми, където k -means може да се отчете като по-надежден от агломеративното клъстериране. Друг аргумент срещу агломеративното клъстериране е, че той е по-бавен при текущите входни данни. Данните дотук са според предварителните очаквания. На Фигура 32 е представен сравнителен анализ на качеството на клъстерите при двата подхода за клъстериране. Използвана е силует-оценка за качеството на получените клъстери. По осите Ox и Oy разполагат стойности на различни на брой предварително-зададени клъстери, за агломеративното клъстериране и за k -means клъстерирането. Програмната симулация е реализирана на Python.

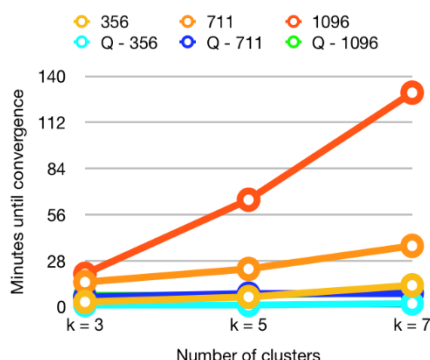


Фигура 32. Графично сравнение на качеството на клъстериране между k -means и агломеративното клъстериране.

5.2 Експеримент №2 – квантова система за препоръки

Разглеждайки основните градивни блокове на системния модел спрямо **Трета глава**, операциите за четене и запис могат да се разглеждат със сложност $O(1)$, както е при класическия асимптотичен анализ. С цел абстракция, такава сложност ще бъде приета за подготовителните стъпки, т.е. поставянето на системата в суперпозиция, както и настройката на регистрите, няма да се включват към основната тежест при изчисленията. Друга стъпка, която може да се пренебрегне за момента, е извличането (екстракцията) на ORB дескриптори, тъй като от гледна точка на проучването в текущия дисертационен труд, приложението на тази операция не води до категоричен положителен ефект, макар че може да бъде интерпретирана и в квантов вариант. Броят на дескрипторите е от значение за броя на изчисленията и тяхната итерация може да се оцени количествено с линейна сложност, т.е. $O(N)$, ако N е броят на входните дескриптори. От значителна важност е стойността на k , заложена при клъстериращия алгоритъм, защото спрямо нея се определя и броят на класовете, по които ще се прогнозира принадлежността на даден входен параметър. В случая показателят d е също толкова съществен като стойност, защото са подадени ORB дескриптори, които имат 256 на брой измерения, т.е. определят се от 256 класически бита. Последната част от системата за препоръки е допълнителната операция за търсене, която класически се изпълнява за линейно време при несортиран масив от данни, или иначе

казано $O(N)$. Алгоритмът на Гроувър, както беше уточнено в предходните глави, ще се изпълни за $O(\sqrt{N})$ време. За целите на експеримента е използвано множеството от данни *Graz-02*. То се състои от 1096 изображения, разделени в три категории: велосипеди, автомобили, хора. Всички изображения са цветни и имат разделителна способност 640x480. Тъй като съществуват три класа, т.е. три разпознаваеми типа обекти, началната стойност на k може да бъде 3. Тази настройка ще се използва за етапа на *обучение* на системата. По-късно, при етапа за *разпознаване на шаблони* ще се въвежда ново изображение, т.е. неговите ORB дескриптори, и спрямо тях ще се извършва класифицирането. За тестовите е използван достъп до облачните услуги на IBM Quantum Experience, които се включват при употребата на *qiskit*, описан в **Четвърта глава**. На Фигура 33 е показана следната зависимост: за брой изображения, използвани при обучението, а именно 356, 711 и 1096 съответно, и брой центроиди k .



Фигура 33. Зависимост между избор на симулирана квантова среда и облачно-базирана квантова услуга. С Q се отбелязват метриките от квантово-базирана облачна среда.

За времето, при което квантовият алгоритъм, изпълнен в облачна среда, може да достигне конвергенция при 356 изображения, симулираният такъв все още не може да раздели центроидите един от друг и съответно няма как да послужи за класификационна техника. След изпълнение на експеримента може да се обобщи следното: Наблюденията показаха, че стъпката за *назначаване* на входящ елемент към центроид е най-времеемка и се оказва като основна трудност при изпълнение на алгоритъма. Стъпката за измерване на разстояние чрез Жакард-Нийдхам, както мажоритарният избор се оказват ефективни не само като бързодействие, което е лесно обяснимо, но и като качество на изчисленията. Друго значително усложнение се оказва броят итерации на клъстериращия алгоритъм до конвергенция, а в случая се достигат 50 на брой. Употребата на QMA и алгоритъм на Гроувър успяват да намалят броя итерации, което е особено полезно при бинарните вектори, защото трудно биват разделени бързо по клъстери. Точността на алгоритъма също е висока, но трябва да се отбележи, че са използвани изображения с ясно-открояващи се обекти в себе си, т.е. няма как да се потвърди дейността на алгоритъма при изображения с влошено качество или трудно различими обекти от познатите класове. Подобно приложение не би било полезно при работа със системи в реално време, но дава повод за допълнителни проучвания.

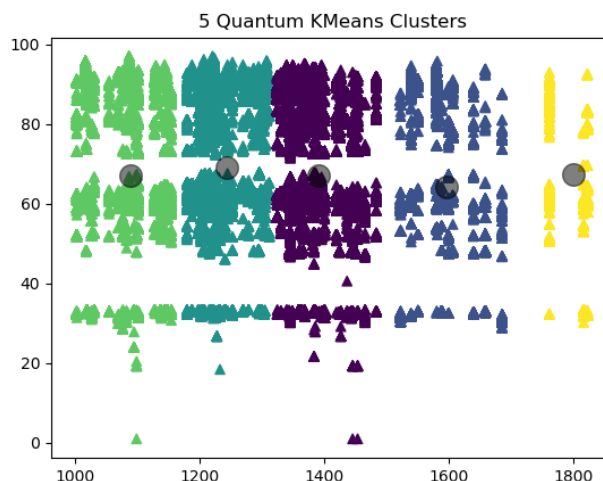
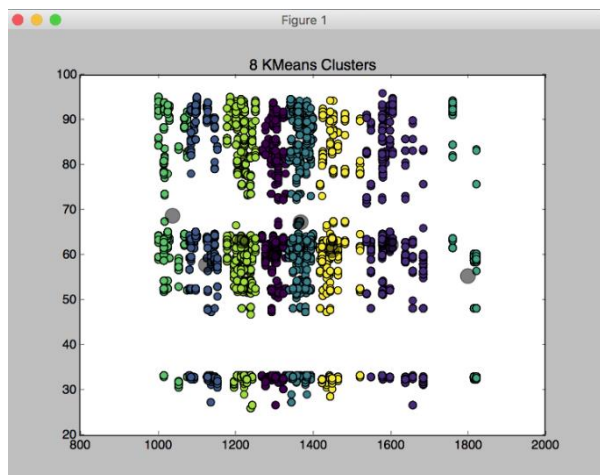
5.3 Експеримент №3 – квантова система за засичане на фалшиви потребителски профили

За целите на експеримента е изготвено изкуствено множество потребителски профили, генерирано чрез хаотичен генератор. Този тип генератори са популярни при подобен тип експерименти, защото могат да генерират голям набор от различаващи се елементи. От друга страна притежават детерминистични свойства като прогнозируемост, т.е. при идентични входни условия дават една и съща изходна последователност от генерирани елементи. Текущият експеримент се нуждае от различни видове имена: такива, които си

приличат; такива, които са анаграми едно на друго; производни имена на вече генерирани такива. За целите на генерирането, ще се създаде хаотична карта, спомагаща създаването на описаните варианти. Стъпките, през които ще се изпълни предложения експеримент, могат да се обобщят по следния начин:

- 1) Генериране на бази данни от комплексни, различни имена чрез псевдо-случаен хаотичен генератор.
- 2) Клъстериране чрез квантова техника за ненадзиравано машинно обучение.
- 3) Оценка на качеството на клъстерите чрез метода *силует*.

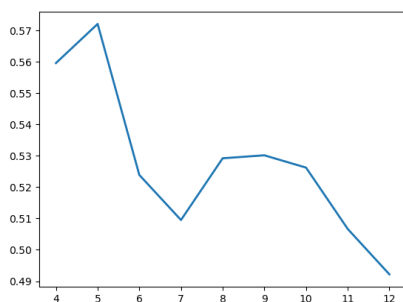
Генерираните потребителски профили се състоят от n на брой символи, където $n > 6$ и също така представлява комбинация от букви и буквено-цифрови символи. Използва се кодиране UTF-8. Началните потребителски профили са 100, като от всеки от тях се генерират поне 1000 варианта. Накрая се генерират 100 000 профила, които се подават на стъпката за клъстериране. За текущото приложение са необходими няколко изисквания на хаотичния генератор: (1) При така поставената задача всяка хаотична карта ще може да се използва. (2) Хаотичният генератор ще се счита за успешен, ако премине NIST-тестовите⁵ за случайност. (3) Ако генерираните стойности са равномерно разпределени, те трябва да бъдат запазвани с еднаква вероятност. Условие (3) е важно, защото за целите на експеримента се предполага, че всички правила на хаотичния генератор от 0 до 9, където съответната цифра е първата за някоя генерирана псевдо-случайна стойност, имат еднаква вероятност да бъдат създадени. За разлика от предходния експеримент тук няма нужда от бинаризиране на входните данни и могат да се използват реални числа и символи. Това е допълнително предимство към текущото приложение на хаотичния генератор, защото не съществува стриктно правило за бинаризиране. И съответно извършвайки този процес, може да се загуби възможността за по-голяма степен на случайност. При генерирана и подадена база данни се генерират клъстерите, представени на Фигура 34.1 и Фигура 34.2, където данните са подредени по брой близки до тях низове (x -оста) и вероятност да са фалшиви имена (y -оста).



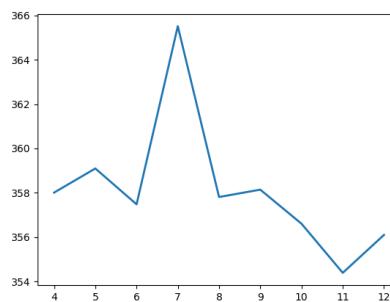
Фигури 34.1. и 34.2. Генерирани клъстери при подаване на потребителски профили на класически k -means алгоритъм.

На Фигурите 34.1 и 34.2 образуваните центроиди са маркирани със сиви окръжности. С цел измерване на качеството на получените клъстери се използва квантовият метод *силует*. При различен брой клъстери k (x -оста) се получава графика, представена на Фигура 35.1. От нея се вижда, че $k=5$ е оптималния избор, когато става въпрос за по-високо качество на клъстерите (y -оста).

⁵ Бел. авт. Има се предвид National Institute of Standards and Technology – U.S. Department of Commerce.

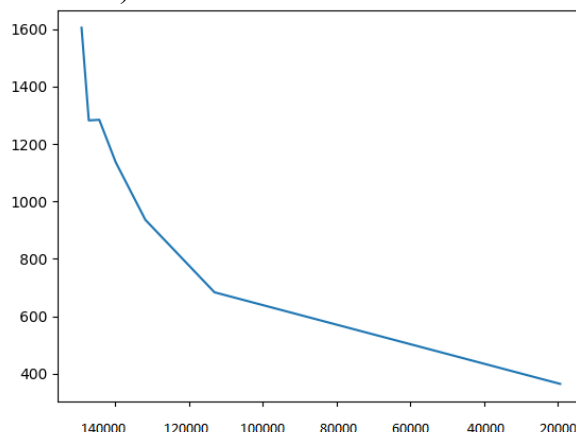


Фигура 35.1. Силует-оценка при различен брой кълстери.



Фигура 35.2. Време за изпълнение на силует-оценка при различен брой кълстери.

При отчитане на времето за изпълнение може да се извършат същите измервания, но този път се засича времето за конвергенция. Резултатите от този анализ са представени на Фигура 35.2, като по оста x отново са нанесени различните стойности за k , а по оста y — времето (в секунди) за изпълнение на силует-оценката. В класическите реализации силует-оценката може да се изчисли със сложност на алгоритъма по време $O(1kN^2)$, т.е. в квантовите се очаква тя да е $O(N \log k)$. Изводът е, че силует-оценката не бива да зависи толкова чувствително от k , но от друга страна стойностите в секунди между 356 и 364 не са много, затова резултатите от Фигура 35.2 могат да се приемат като флуктуации на средата. Това не е неочаквано, заради начина, по който Python изпълнява програмата. На Фигура 36 са представени измерените стойности за време в секунди (по y -оста) и брой елементи в системата (низходящо по x -оста).



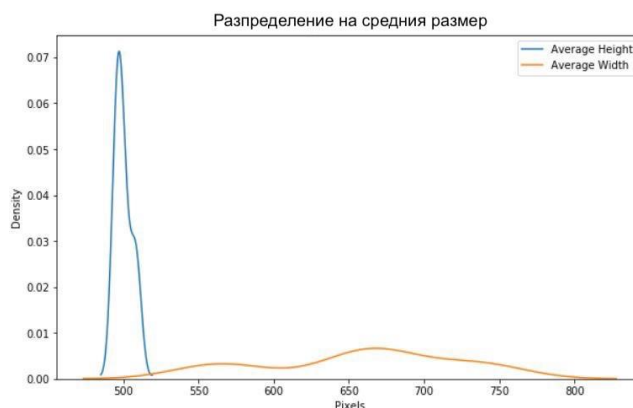
Фигура 36. Време за изпълнение на целия квантов k -means алгоритъм при класификация на потребителски профили.

Резултатите потвърждават теоретично предвиденото чрез средствата на асимптотичния анализ. На тяхна основа е възможно да се формулират следните изводи: Квантовите кълстеризащи алгоритми могат да се използват при реални приложения, но се изисква значително количество памет, където да се разположат елементите им. Използването на квантова симулационна среда показва, че времето за изпълнение на квантовия кълстеризащ алгоритъм е по-малко от това на класическия, като следва логаритмично по-бързо изпълнение. При един и същ алгоритъм за класическа и квантова имплементация се получават сходни кълстери, но центроидите забележимо се различават, въпреки че следват един и същ шаблон на разположение. Силует-оценката може да се интегрира в квантовия алгоритъм чрез своето квантово представяне и да послужи за самооценка на системата в реално време.

5.4 Експеримент №4 – реализиране на конволюционна невронна мрежа

Предварителната обработка на изображения е необходима, за да могат да се използват в контекста на VGG16-модела. Той изисква използване на триканално RGB изображение с размер 224×224 . За да се подготвят изображенията за използвания мрежови модел, трябва

да се скалират до размери 224 x 224 и да се нормализира всеки цветен канал, като се определи средната стойност и се раздели на стандартното отклонение, както е представено на Фигура 38.



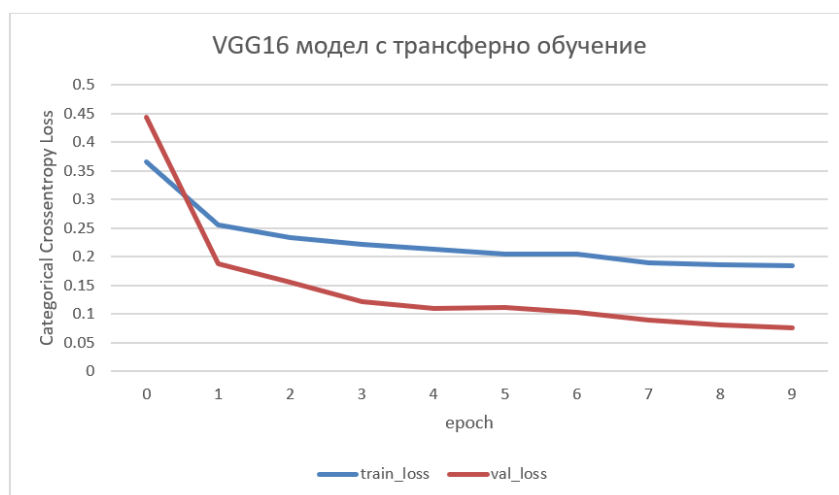
Фигура 38. Обхвата на ширините и височините на изображенията в обучаващото множество изображения.

Използвана е конволюционна невронна мрежа в дълбочина, т.е. със седем конволюционни слоя и два напълно свързани слоя с балансиране на тежестите на класовете, и е използван същият набор от изображения, разделени по начина като в предходния модел. След балансиране на тежестите на класовете спрямо броя налични изображения за всеки клас, се получава грешка от 0.08890 и тестова точност 0.97077. Моделът на конволюционната невронна мрежа със седем конволюционни слоя подобрява производителността при класификация. Стойностите на обучителната и валидационната грешки са с много близки стойности, което означава, че моделът не е нито прекомерно обучен, нито недостатъчно обучен чрез дадените тренировъчни данни. Резултатите са представени в Таблица 9.

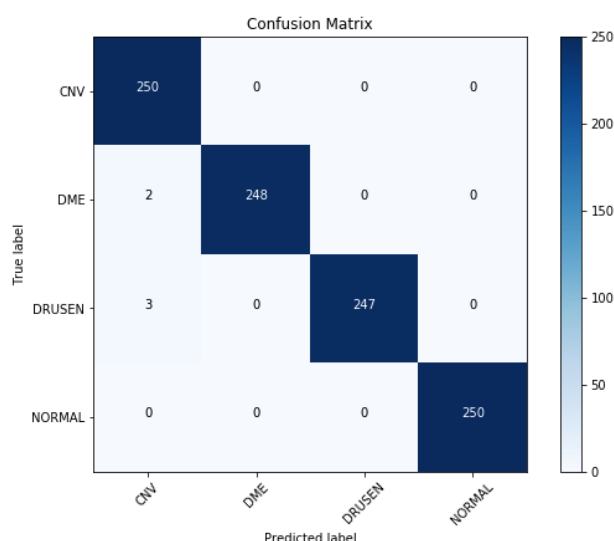
Таблица 9. Обобщение на резултатите от 7-слойния модел с балансиране на тежестите.

	precision	recall	f1-score	support
CNV	0.91	1.00	0.95	250
DME	0.99	0.98	0.98	250
DRUSEN	1.00	0.91	0.95	250
Normal	0.99	1.00	1.00	250
micro avg	0.97	0.97	0.97	1000
macro avg	0.97	0.97	0.97	1000
weighted avg	0.97	0.97	0.97	1000

Тъй като класификаторът може да бъде разширяван с още класове, за които не се разполага с голям набор от изображения, едно от най-добрите решения за това е използването на трансферно обучение. Използва се конволюционната невронна мрежа с модел VGG16, която съдържа тринадесет конволюционни слоя и три напълно свързани слоя. Използват тегловни коефициенти, получени след предварително трениране с ImageNet, тъй като този модел показва добри резултати при извличане (екстракция) на свойства от изображение. За да се приложи този подход за обучение, то се замразява за всички слоеве без тези на класификатора. Отново се използва същият набор от изображения, както в предните модели, разделен по същия начин. Като резултат от използването на готовите тегловни коефициенти и обучение единствено на класификационната част на модела се постига тестова грешка равна на 0.03829 и тестова точност 0.98758. Отново валидационната грешка е със сходни стойности с тези на обучителната, което означава, че моделът се обучава правилно. На база на този модел е изграден и реалният класификатор от шест класа, два от които са много по-ограничени откъм набор от данни. Въпреки това те осигуряват висока точност. Резултатите от изследването са представени на Фигура 43.1 и Фигура 43.2.



Фигура 43.1. Диаграма на тренировъчна и валидационна загуби за всички десет епохи от обучението на VGG16-модел за трансферно обучение.



Фигура 43.2. Матрица на разпределение на VGG16-модел с трансферно обучение с тестовия набор от данни.

На базата на получените резултати могат да се формулират следните изводи: Първо, реализираният модел на конволуционна невронна мрежа е напълно подходящ за представяне на ролята на квантовите сплитания при задачите за квантово машинно обучение, тъй като представеният модел на конволуционна невронна мрежа изпълнява задачата за откриване на класове заболявания на ретината при правилен подход на обучение, а именно чрез VGG16. Второ, предложената невронна мрежа може да се представи чрез квантови математически модели, но имплементацията ѝ остава ограничена към момента, защото не съществува технология, която да конструира мрежата от квантови сплитания, необходима за изграждането ѝ. В контекста на разработената невронна мрежа не може да се направи обстойно и добре обосновано измерване на квантовите сплитания, въпреки че те са фундаментален ресурс, щом става въпрос за комуникация между слоевете и преоценяване на тегловите коефициенти. Реализацията на квантовата конволуционна невронна мрежа е повод за следващи научни изследвания в областта на квантовите сплитания и квантовото машинно обучение.

НАУЧНИ, НАУЧНО-ПРИЛОЖНИ И ПРИЛОЖНИ ПРИНОСИ

В резултат на проведените в дисертационния труд научни изследвания са постигнати следните приноси:

- *Научни приноси:*

- Предложен е подход за математическо моделиране на квантови системи за изкуствен интелект с прилагане принципите на работа на квантови сплитания с цел ускорение в изпълнението на приложения алгоритъм.
- Предложен е подход за синтез на квантови алгоритми за машинно обучение и с използване на квантови канали за пренос на квантова или класическа информации с цел устойчивост на шум при определени параметри на канала.
- Синтезирана е система за добавяне на изисквания от клиента с цел ускоряване или подобряване на изходните данни, наречена система за препоръки, с използване на нов подход за клъстериране на данните, работещ както при квантовите, така и при класическите компютри. В основата на подхода е приложението на алгоритъма на Гроувър, който позволява реализиране на различни по рода си функционалности: бързо откриване на елемент с несортиран масив, подобряване сходимостта при системите за препоръка, бързо намиране на най-малкия елемент, което е съществено за клъстериращите алгоритми.
- Синтезиран е подход за откриване на фалшиви потребителски имена на акаунти с цел намаляване броя на хакерски атаки, като е възможно самоподобряване на клъстериращите алгоритми чрез прилагане на Силует-метод на квантов принцип.
- Проектирана е квантова конволюционна невронна мрежа с приложение за разпознаване на шаблони и изграждане на асоциативна памет. Процесът на обучение цели да определи модел на невронната мрежа, който да класифицира вярно по-голяма част от данните за обучение, с което се осигурява по-малка грешка при класификация с използване на квантовата невронна мрежа.

- *Научно-приложни приноси:*

- Синтезиран е k -majority алгоритъм, модифициран вариант на k -means клъстериране, за намиране на най-предпочитан елемент. Съгласно този алгоритъм се получават предварително k на брой клъстера и е възможно собствените елементи да изпълняват мажоритарен вот. В резултат се позволява определяне на нови центроиди.
- Извършен е обзор на средствата за техническа и програмна реализации на квантови логически схеми или модели на алгоритми за квантови изчисления, и е направен избор за квантови симулационни среди с използване на езика за програмиране Python, съвместно с IBM-продукти за провеждане на експерименти.

- *Приложни приноси:*

- Проведени са четири вида експериментални изследвания въз основа на предложените подходи, доказващи работоспособността им.

СПИСЪК НА ПУБЛИКАЦИИТЕ ПО ДИСЕРТАЦИОННИЯ ТРУД

1. **Andreev D.**, “Analysis of machine learning methods and algorithms in a quantum entanglement-based environment“, *Proceedings of the XLVII Spring Conference of the Union of Bulgarian Mathematicians*, Borovets, April 2–6, 2018, Издателство на БАН, pp.129-137, ISSN 1313-3330.
http://www.math.bas.bg/smb/2018_PK/tom_2018/pdf/129-136.pdf
2. **Andreev D.**, Filipova-Petrakieva S., Taralova I., “A novel approach for protection of accounts’ names against hackers combining cluster analysis and chaotic theory“, *Journal of Internet Technology and Secured Transaction*, Volume VI, Issue 2, Jun 2018, pp. 579–587, ISSN 2046-3723.
https://hal.archives-ouvertes.fr/hal-02060383/file/paper_JITST2018_vol-7_N-1_last-1.pdf
<http://infonomics-society.org/jitst/published-papers/volume-6-2018/>
3. **Andreev D.**, Filipova-Petrakieva S., Taralova I., Qiao Z., “Applying quantum machine learning approach for detecting chaotically generated fake usernames of account: Security systems with heightened safety of information“, *Proceedings of IEEE International Conference for Internet Technology and Secured Transactions (ICITST’2018)*, Dec 2018, Cambridge, UK (статията е отличена с награда ICITST’2018 Best Paper Award).
<https://www.orpems.org/Proceedings/ICITST-WorldCIS-WCST-WCICSS-2018-Proceedings.pdf>
4. **Andreev D.**, Lazarova M., “Applying a quantum amplification over a system for image feature matching and image recommendation“, *Proceedings of the 9th Balkan Conference on Informatics (BCI’19)*, Article No. 28, Sept 2019, Sofia, Bulgaria, ISBN: 978-1-4503-7193-3, DOI>0.1145/3351556.3351564 (индексирана в Scopus)
<https://dl.acm.org/doi/10.1145/3351556.3351564>
5. **Andreev D.**, “Quantum Recommendation System for Image Feature Matching and Pattern Recognition“, In: Arai K., Bhatia R., Kapoor S. (eds) *Proceedings of the Future Technologies Conference (FTC’2019)*, 24-25 Oct. 2019, San Francisco, United States, ISBN: 978-3-030-32519-0, DOI: 10.1007/978-3-030-32520-6_37. *Advances in Intelligent Systems and Computing*, Vol. 1069, Springer Nature, pp.497–511, ISSN: 2194-5357 (индексирана в Scopus, IF 0.57, SJR 0.174)
<https://www.springer.com/gp/book/9783030325190#>
6. Mihalov V., **Andreev D.**, Lazarova M., “Software Platform for Retinal Disease Diagnosis through Deep Convolutional Neural Networks“, *21st International Conference on Computer Systems and Technologies (CompSysTech’20)*, 19-20 June 2020, Ruse, Bulgaria (приета за печат).
7. **Andreev D.**, Lazarova M., “Quantum Entanglement-based Deep Convolutional Neural Networks“, *21st International Conference on Computer Systems and Technologies (CompSysTech’20)*, 19-20 June 2020, Ruse, Bulgaria (приета за печат).

SUMMARY

Methods and algorithms for machine learning in quantum entanglement-based environment

Author: Desislav Andreev

The current work is dedicated to the study of both the constraints and the possibilities in interpreting classical machine learning algorithms to their quantum equivalents. The possible usage of the quantum entanglement and its affect over the routine, the speed, the quality and the scalability of given techniques is also thoroughly explored. Thus, three software applications in various simulation environments have been developed, based on the mathematical approaches, which are derived from this scientific work. The possibilities for implementing quantum entanglement as a technique for machine learning have been studied through four experiments. Problems, that are detected during the research have been identified and noted, mainly considering: the impossibility of interpretation of the classical approach or a part of it, as a simple routine, to its quantum variant; the needless usage of quantum entanglement in some cases – it overcomplicates the mathematical approach and theoretically doesn't bring any added value to the solution of a given problem. The final results and the research's contributions could be used directly in further-provided scientific projects. The mathematical approaches, described by the current work, could be applied by both scientists and engineers in the areas of machine learning and quantum computing. These approaches are the contributions resulting from the following common objectives:

- A bibliographical research with concluded summary of the following subjects shall be conducted: mathematical models of the classical machine learning techniques; the already-studied and implemented quantum circuits and algorithms; the nature of the quantum entanglements and their possibility for integration in quantum systems.
- Conduct a research for using quantum entanglement as networking establishment in quantum machine learning, thus leading to algorithms' scalability.
- Conduct a research and thorough analysis for the possible implementation of complex quantum machine learning algorithm systems, based on quantum entanglement.
- Identify the main resources and environments, including simulated ones, for implementation and execution of quantum machine learning algorithms. Their advantages and disadvantages must be well described and a defined quantum environment shall be concluded.
- A detailed report of the existing tools and software components for the purposes of the current work shall be constructed. The useful ones must be applied with the realization and the execution of the required experiments.

The working hypothesis is the assumption, that the concluded mathematical approaches might lead to advancement in the area of artificial intelligence, specifically in the area of quantum machine learning, e.g. establishing a reliable communication, faster networking and the overall improvement of machine learning's algorithms' evaluation quality and speed.