



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

NICOLAS MELO DE OLIVEIRA

Classificadores Quânticos Baseados em Distância para Dispositivos NISQ

Recife

2024

NICOLAS MELO DE OLIVEIRA

Classificadores Quânticos Baseados em Distância para Dispositivos NISQ

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

Área de Concentração: Inteligência Computacional

Orientador (a): Adenilton José da Silva

Coorientador (a): Teresa Bernarda Ludermir

Recife

2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Oliveira, Nicolas Melo de.

Classificadores quânticos baseados em distância para dispositivos NISQ / Nicolas Melo de Oliveira. - Recife, 2024.
79 f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Informática, Programa de Pós-Graduação em Ciência da Computação, 2024.

Orientação: Adenilton José da Silva.

Coorientação: Teresa Bernarda Ludermir.

Inclui referências.

1. Computação quântica; 2. Aprendizagem de máquina quântica;
3. Classificadores baseados em distância. I. Silva, Adenilton José da. II. Ludermir, Teresa Bernarda. III. Título.

UFPE-Biblioteca Central

AGRADECIMENTOS

Como não poderia deixar de ser, começo meus agradecimentos aos orientadores presentes na minha vida acadêmica. Primeiramente, agradeço ao professor Dr. Wilson de Oliveira que, desde a graduação na UFRPE, me inspirou a seguir a carreira acadêmica e me guiou neste sentido. Agradeço também ao professor Dr. Adenilton Silva e à professora Dra. Teresa Ludermir, essenciais para que fosse possível concluir este doutorado. Deixo aqui um imenso agradecimento aos três por todo o apoio, paciência e conselhos dados durante todo o período que estiveram presentes em minha vida.

Agradeço ao Dr. Israel Araújo e ao Dr. Daniel Park, pela parceria na escrita de um dos artigos que resultaram desta tese. Resultado crucial para que esta tese pudesse ser defendida.

Sendo este o fim de uma jornada e início de um novo ciclo da minha vida profissional e pessoal, aproveito para agradecer a todos os funcionários, professores e colegas da UFPE e UFRPE, com os quais compartilhei momentos fundamentais para a minha formação, desde a graduação, passando pelo mestrado, até culminar neste momento de fechamento.

Agradeço a todos os amigos e pessoas que, de alguma forma, fizeram parte da minha trajetória.

Agradeço à FACEPE por, durante toda minha jornada acadêmica, ter financiado minha pesquisa.

Por fim, durante os últimos momentos de escrita desta tese, eu conheci Sylvia. E, a ela, dedico aqui um agradecimento especial. Apesar do pouco tempo que nos conhecemos, Sylvia me fez acreditar em um projeto de vida e em um potencial ainda não explorado. Me fez acreditar em cenários inimagináveis para mim até o momento que ela entrou na minha vida. Em tão pouco tempo, ela provocou mudanças profundas, motivou o melhor em mim. E, mesmo que ela não acredite, ela é parte significativa deste final de ciclo da minha vida acadêmica. Independente do que o futuro nos reserva, era necessário, para mim, eternizar meu profundo agradecimento pelo nosso encontro.

“Eu prefiro ter perguntas que não podem ser respondidas a ter respostas que não podem ser questionadas”. Richard Feynman.

RESUMO

A aprendizagem de máquina quântica é um campo em rápido crescimento e tem surgido como uma alternativa à aprendizagem de máquina clássica para impulsionar a busca por soluções eficientes para certos problemas. Além disso, o avanço da tecnologia em Computação Quântica trouxe possibilidades para a execução de algoritmos em dispositivos quânticos reais. Aqui, centramos nossa pesquisa no estudo de uma das tarefas mais comuns dentro da aprendizagem de máquina, a *classificação supervisionada*. Nos nossos estudos, limitamos a temática para os classificadores quânticos baseados em distância, visando a execução dos mesmos em dispositivos quânticos ruidosos de escala intermediária (*Noisy Intermediate-Scale Quantum (NISQ)*). Embora algoritmos quânticos anteriores para classificação baseada em distância tenham se mostrado promissores, muitos deles não são adequados com os dispositivos NISQ e não conseguem lidar com problemas envolvendo múltiplas classes. O *Hadamard Classifier (HC)* é um modelo quântico de aprendizagem de máquina baseado em distância para reconhecimento de padrões. A partir dele, desenvolvemos dois classificadores quânticos baseados em distância para execução em dispositivos NISQ: o primeiro deles, *Quantum One-Class Classifier (QOCC)*, é um modelo de classificação elementar. Nosso objetivo com o QOCC é oferecer um classificador que consuma menos recursos computacionais quânticos, estando menos propenso aos erros provenientes de ruídos, ao mesmo tempo em que possui desempenho competitivo quando comparado a análogos clássicos. Os resultados obtidos nos experimentos com o QOCC foram promissores e confirmaram um desempenho equiparável tanto contra classificadores clássicos quanto contra o HC, mesmo usando menos qubits (bits quânticos) e menos operações quânticas. Em seguida, focamos nossos esforços em utilizar o QOCC como base para um classificador multiclasse. Nossa motivação para isso consistiu na carência desse tipo de classificador na literatura da aprendizagem de máquina quântica e na impossibilidade atual de executar, em dispositivos NISQ, algumas rotinas presentes nos classificadores multiclasse existentes. Propomos o *quantum Variational Distance-based Centroid Classifier (qVDCC)*, que faz uso da teoria dos circuitos quânticos variacionais para, em um esquema híbrido de computação quântica-clássica, realizar a classificação de bases de dados multiclasse. Essa abordagem nos permitiu utilizar o maquinário bem-established da aprendizagem de máquina clássica para encontrar e otimizar centroides artificiais para as bases de dados que minimizassem suas distâncias para os demais dados da respectiva classe. Simulamos o funcionamento do qVDCC em dois ambientes distintos: primeiramente em uma simulação ideal (livre de erros); em seguida, em um uma

simulação ruidosa que mimetiza um dispositivo NISQ. Mais uma vez, nossos resultados se mostraram promissores e competitivos em relação aos classificadores clássicos.

Palavras-chaves: Computação Quântica. Aprendizagem de Máquina Quântica. Classificadores Baseados em Distância.

ABSTRACT

Quantum machine learning is a rapidly growing field and has emerged as an alternative to classical machine learning to boost the search for efficient solutions to certain problems. Furthermore, the advancement of technology in quantum computing has brought possibilities for executing algorithms on real quantum devices. Here, we focus our research on studying one of the most common tasks in machine learning, *supervised classification*. In our studies, we narrow the focus to distance-based quantum classifiers, aiming to execute them on Noisy Intermediate-Scale Quantum (NISQ) devices. While previous quantum algorithms for distance-based classification have shown promise, many are not compatible with noisy quantum devices and struggle with problems involving multiple classes. The Hadamard Classifier (HC) is a distance-based quantum machine learning model for pattern recognition. Building upon it, we introduce two distance-based quantum classifiers tailored for execution on NISQ devices: the first, QOCC, is a basic classification model. Our goal with QOCC is to offer a classifier that consumes fewer quantum computational resources, making it less prone to errors from noise, while still maintaining competitive performance compared to classical counterparts. The results obtained in experiments with QOCC were promising and confirmed its competitive performance against classical classifiers and HC, even using fewer qubits and quantum operations. In a second phase, we directed our efforts towards using QOCC as the foundation for a multiclass classifier. Our motivation stemmed from the lack of such classifiers in the quantum machine learning literature and the current inability to execute some routines from existing multiclass classifiers on NISQ devices. We propose the *qVDCC*, which leverages variational quantum circuits theory to, in a quantum-classical hybrid computing scheme, performing the classification of multiclass databases. This approach allowed us to use the well-established machinery of classical machine learning to find and optimize artificial centroids for classes that minimize their distances to other data points within the same class. We ran *qVDCC* in two simulation environments: first in an ideal (error-free) setting; then in a noisy simulation that mimics a NISQ device. Once again, our results were promising and competitive compared to classical classifiers.

Keywords: Quantum Computing. Quantum Machine Learning. Distance-based Classifiers.

LISTA DE FIGURAS

<p>Figura 1 – Representação no circuito quântico de um porta lógica quântica controlada geral, onde $q_0\dots q_i$ são os qubits de controle e $q_{i+1}\dots q_j$ são qubits alvo. O operador U é aplicado apenas aos qubits alvo se todos os qubits de controle tiverem valores verdadeiros (1).</p>	28
<p>Figura 2 – Procedimento para codificar o vetor $0\dots 0\rangle$ a partir de qualquer outro vetor real.</p>	30
<p>Figura 3 – Exemplo de como carregar um padrão nas amplitudes de um estado quântico.</p>	31
<p>Figura 4 – Multiplexador quântico para uma porta R_y multi-controlada. Os CNOTs dentro das linhas tracejadas podem ser cancelados.</p>	31
<p>Figura 5 – Multiplexador quântico para uma porta R_y multi-controlada para incorporar dados clássicos de valores reais com quatro atributos em um estado quântico de dois qubits.</p>	32
<p>Figura 6 – Esquema geral do funcionamento de um algoritmo híbrido quântico-clássico. A execução passo a passo inclui (1) codificar os dados clássicos em registradores quânticos, (2) um modelo quântico, composto por portas lógicas quânticas paramétricas (com ângulos θ a serem otimizados), portas fixas e emaranhamento, (3) medição do circuito quântico, (4) avaliação de uma função de custo e (5) otimização dos parâmetros por meio de um computador clássico. Esse processo é repetido várias vezes para escolher os melhores ângulos θ em cada iteração.</p>	36
<p>Figura 7 – Exemplo do classificador Hadamard. No passo A, os qubits auxiliar (a_0) e de índice (m_0) são postos em uma sobreposição uniforme. No passo B o vetor de teste é carregado e nos passos C e D os vetores de treinamento são incorporados às amplitudes do estado quântico emaranhados com o qubit de índice (i_0). No início do passo E, os qubits de índice (i_0) e de classe (c_0) são trocados de posição devido ao mapeamento físico dos qubits no hardware que foi usado para os experimentos do HC. No passo F, há a interferência sendo realizada entre o vetor de teste e os de treinamento através do operador Hadamard, seguida da medição (condicional) de 2 qubits (auxiliar a_0 e de classe c_0).</p>	46

Figura 8 – Um exemplo artificial em que a probabilidade de pós-seleção tende a zero para um determinado novo padrão de entrada x_0 da classe 0. As estrelas são amostras do conjunto de dados e têm norma unitária. O quadrado preto x_0 (x_1) é uma amostra da classe 0 (1) a ser classificada. A pós-seleção é realizada com base na distância entre a amostra a ser classificada e todas as outras amostras no conjunto de dados. Neste exemplo, a amostra x_0 não seria classificada porque a abordagem em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) requer a saída 0 na pós-seleção para prosseguir com a classificação. A saída 0 na pós-seleção só é alcançada se os dados a serem classificados estiverem próximos aos outros dados no conjunto de dados. 49

Figura 9 – Circuito quântico implementando o classificador quântico de única classe QOCC. No passo A o qubit auxiliar ($|q_0\rangle$) e o qubit de índice ($|q_1\rangle$) são postos em uma sobreposição uniforme através do operador Hadamard. O resultado de sucessivas execuções deste circuito representa uma espécie de grau de semelhança do novo padrão de entrada (passo B) em relação aos padrões já armazenados no classificador (passos C e D). As portas lógicas quânticas R_y são responsáveis por carregar dados (neste exemplo, padrões com 2 atributos) de cada vetor por meio de seus ângulos associados α , β e γ . Dentro das linhas tracejadas está o sub circuito para a preparação do estado. No final da computação (passo E), a porta lógica Hadamard no qubit auxiliar interfere com as cópias do novo padrão de entrada com os padrões carregados e, em seguida, o qubit auxiliar é medido. 51

Figura 10 – Circuito quântico mínimo implementando o classificador quântico de uma classe com uma amostra de treinamento e dois atributos. O passo A coloca o qubit ancilla (índice a) em superposição uniforme. Do passo B ao passo D, são carregados os padrões de teste e treinamento. O passo E realiza a interferência entre o vetor de entrada e o vetor armazenado através da porta Hadamard, e o qubit ancilla (índice a) é medido. 52

- Figura 11 – Circuito quântico representando o esquema para 1 classe (célula) do qVDCC. No *passo A*, o qubit auxiliar (índice 0) é colocado em uma superposição uniforme através do operador Hadamard. O padrão de entrada a ser classificado é carregado no *passo B*. As características dos padrões são codificadas na amplitude dos qubits de dados ($|0\rangle_1, \dots, |0\rangle_m$), onde M é o número de características e $m = \lceil \log(M) \rceil$. O *passo C* inverte o qubit auxiliar e move o padrão de entrada para o estado $|0\rangle$ desse mesmo qubit auxiliar. O *modelo quântico* no *passo D* é o circuito quântico parametrizado responsável por encontrar o centroide. No *passo E*, uma porta $R_y(\theta)$ interfere o centroide com o padrão de entrada. Finalmente, ao final do circuito, temos a medição. A informação dos ângulos das portas lógicas quânticas paramétricas foi omitida por simplicidade. 60
- Figura 12 – O circuito completo do qVDCC para N classes (células) e padrões com M atributos. O detalhamento de cada célula pode ser visto na Figura 11. A informação dos índices dos ângulos θ foi omitida por simplicidade. 61
- Figura 13 – Esquema geral das simulações do qVDCC com um modelo de computação híbrida quântico-clássica para classificação multiclasse. 62
- Figura 14 – Exemplo da motivação para substituir o operador Hadamard por uma porta de rotação parametrizada para realizar a interferência entre as amostras de teste e treinamento no qVDCC. 63
- Figura 15 – Triângulos, pontos e quadrados representam as 3 classes da base de dados *Iris*. Os marcadores 'x' preenchidos representam os centroides virtuais encontrados para as classes, obtidos através da otimização dos parâmetros presentes no circuito quântico variacional do qVDCC. As 2 características mostradas no gráfico de dispersão foram obtidas executando o algoritmo t-SNE (MAATEN; HINTON, 2008) nas 4 características originais do conjunto de dados. 66

LISTA DE TABELAS

Tabela 1 – Características das bases de dados	53
Tabela 2 – Resultados do classificador Hadamard em simulações e experimentos nos dispositivos quânticos reais <code>ibmq_athens</code> and <code>ibmq_santiago</code> (marcado com um *).	55
Tabela 3 – Resultados da execução do QOCC para 2 classes e 2 padrões armazenados em simulação e nos dispositivos quânticos reais <code>ibmq_athens</code> and <code>ibmq_santiago</code> (marcado com um *).	56
Tabela 4 – Acurácia dos classificadores clássicos SVM, DT, KNN e SGD.	56
Tabela 5 – Resultados da execução do QOCC para 2 classes e 1 padrão armazenado em simulação e nos dispositivos quânticos reais <code>ibmq_athens</code> and <code>ibmq_santiago</code> (marcado com um *).	56
Tabela 6 – Características das bases de dados.	64
Tabela 7 – As médias e melhores acurácias do classificador qVDCC (simulação sem ruído) em validação cruzada com 10 partições em cada base de dados.	64
Tabela 8 – As médias e melhores acurácias do classificador qVDCC (simulação ruidosa) em validação cruzada com 10 partições em cada base de dados.	65
Tabela 9 – As médias e melhores acurácias do classificador KNN em validação cruzada com 10 partições em cada base de dados.	65

LISTA DE ABREVIATURAS E SIGLAS

HC	Hadamard Classifier
IA	Inteligência Artificial
IJCNN	International Joint Conference on Neural Networks
KNN	K-Nearest Neighbors
MSE	Mean Squared Error
NISQ	Noisy Intermediate-Scale Quantum
PQM	Probabilistic Quantum Memory
QML	Quantum Machine Learning
QNN	Quantum Neural Network
QOCC	Quantum One-Class Classifier
QUBO	Quadratic Unconstrained Binary Optimization
qVDCC	quantum Variational Distance-based Centroid Classifier
RBF	Radial Basis Function
SVM	Support Vector Machine
VQC	Variational Quantum Circuit

SUMÁRIO

1	INTRODUÇÃO	14
1.1	MOTIVAÇÃO E PROPOSTA	21
1.2	OBJETIVOS	22
1.2.1	Objetivo geral	22
1.2.2	Objetivos Específicos	22
1.3	ORGANIZAÇÃO DESTA TESE	23
2	BASE TEÓRICA	25
2.1	COMPUTAÇÃO QUÂNTICA	25
2.2	CODIFICAÇÃO DE DADOS	28
2.2.1	Codificação de dados clássicos nas amplitudes de registradores quânticos	30
2.3	COMPUTAÇÃO HÍBRIDA E CIRCUITOS QUÂNTICOS PARAMÉTRICOS	33
2.4	ALGORITMOS DE APRENDIZAGEM QUÂNTICA	38
3	CLASSIFICADORES QUÂNTICOS BASEADOS EM DISTÂNCIA	41
3.1	CLASSIFICAÇÃO QUÂNTICA BASEADA EM DISTÂNCIA	41
3.2	CLASSIFICADOR HADAMARD	45
4	CLASSIFICAÇÃO QUÂNTICA E HÍBRIDA BASEADA EM DISTÂNCIA	48
4.1	QOCC - QUANTUM ONE-CLASS CLASSIFIER	48
4.1.1	Experimentos e resultados	52
4.1.2	Discussão	55
4.2	QVDCC - QUANTUM VARIATIONAL DISTANCE-BASED CENTROID CLASSIFIER	58
4.2.1	Classificador Variacional	59
4.2.2	Classificação Multiclasse	59
4.2.3	Interferência Quântica Parametrizada	61
4.2.4	Simulações e Resultados	63
4.2.5	Discussão	67
5	CONCLUSÃO	69
	REFERÊNCIAS	72

1 INTRODUÇÃO

A pressão decorrente dos limites fundamentais na computação clássica (FEDOROV et al., 2022), juntamente com a perspectiva de avanços exponenciais provenientes dos efeitos quânticos, chamou recentemente a atenção para os circuitos quânticos. Preocupações com o tamanho dos transistores, que se aproxima dos limites intrínsecos físicos (MOORE, 2006), além de questões como o desempenho e frequências de *clock*, têm se tornado argumentos para motivar o crescente interesse na computação quântica (FEDOROV et al., 2022). Ainda, mesmo que surjam novas tecnologias que possibilitem o avanço da computação clássica, a computação quântica encontra motivação nas limitações fundamentais da computação clássica em resolver eficientemente certos problemas. Em termos teóricos, os computadores quânticos têm o potencial de superar seus equivalentes clássicos ao resolver determinados problemas discretos (SHOR, 1994; GROVER, 1997).

A computação quântica (FEYNMAN, 2018) é um paradigma computacional que obedece às leis da mecânica quântica e faz uso de seus fenômenos, tais como sobreposição de estados, interferência e emaranhamento (NIELSEN; CHUANG, 2010; SHOR, 1994; GROVER, 1996; ARUTE et al., 2019), para estabelecer um modelo computacional alternativo ao modelo clássico de computar. Assim, propriedades como o paralelismo quântico e os fenômenos já comentados aumentaram o interesse da pesquisa nessa área para problemas sem soluções algorítmicas eficientes conhecidas. Com isso, a computação quântica surge como uma possibilidade mais eficiente para resolver certos problemas que são computacionalmente intratáveis para as máquinas clássicas existentes. Nesse sentido, notadamente, tem-se os algoritmos de fatoração de Shor e o de busca de Grover (SHOR, 1994; GROVER, 1996). Desse modo, uma vez que as pesquisas em computação quântica têm conseguido comprovar sua superioridade em relação ao modelo clássico de computar em alguns cenários, é um resultado natural o crescente interesse em aplicar o paradigma quântico no contexto da aprendizagem de máquina.

Apesar de a computação quântica oferecer o potencial de resolver certos problemas de maneira mais eficiente do que os computadores clássicos, encontrar aplicações comercialmente relevantes para dispositivos NISQ ainda é um desafio (MOHSENI et al., 2017; PRESKILL, 2018). Assim, em virtude do recente progresso observado tanto na evolução do hardware quântico (PRESKILL, 2018) quanto nas plataformas de desenvolvimento (LAROSE, 2019), o interesse na Computação Quântica (NIELSEN; CHUANG, 2010) tem crescido na medida em que se torna

possível testar e comprovar suas vantagens diante da computação clássica (ARUTE et al., 2019).

Posto isso, a *Aprendizagem de Máquina Quântica* (BIAMONTE et al., 2017; SCHULD; SINAYSKIY; PETRUCCIONE, 2015; HUANG et al., 2021; WITTEK, 2014; DUNJKO; TAYLOR; BRIEGEL, 2016; SCHULD; PETRUCCIONE, 2018; DUNJKO; BRIEGEL, 2018; CILIBERTO et al., 2018) (Quantum Machine Learning (QML)) é uma área da computação quântica em rápido crescimento, que tem ganhado especial atenção de pesquisadores e empresas, e explora o uso de dispositivos quânticos para várias tarefas de aprendizado de máquina, combinando técnicas de inteligência artificial com os recursos da computação quântica. Assim, a aproximação entre esses dois campos de estudo busca alternativas para aperfeiçoar soluções existentes para problemas de aprendizado de máquina e, também, objetiva propor novas abordagens para estes. As interpretações quânticas englobam (mas não se limitam a) versões para máquinas de vetores de suporte (Support Vector Machine (SVM)) (REBENTROST; MOHSENI; LLOYD, 2014), redes convolucionais (CONG; CHOI; LUKIN, 2019; HUR; KIM; PARK, 2022), redes neurais (EZHOV; NIFANOVA; VENTURA, 2000; OLIVEIRA et al., 2008; SCHULD; SINAYSKIY; PETRUCCIONE, 2014), algoritmos de agrupamento (AÏMEUR; BRASSARD; GAMBS, 2007) e algoritmos de vizinhos próximos (LI et al., 2022). Porém, uma questão em aberto no QML é como aproveitar efetivamente os dispositivos NISQ para obter vantagens quânticas na abordagem dos atuais problemas de aprendizado de máquina e Inteligência Artificial (IA). Alguns trabalhos que apresentam modelos de aprendizagem de máquina quântica no âmbito da classificação supervisionada e que estão mais relacionados com as propostas desenvolvidas nesta tese são os apresentados nas referências (BLANK et al., 2020; PARK; BLANK; PETRUCCIONE, 2020; KATHURIA et al., 2020; RUAN et al., 2017; SCHULD; FINGERHUTH; PETRUCCIONE, 2017; NEUMANN, 2020).

Cabe pontuar que, apesar de modelos clássicos de classificadores estarem muito bem estabelecidos, fornecendo excelentes resultados em aprendizagem de máquina, ainda existe um gargalo com relação à dimensão dos dados (SALAM et al., 2021). Esse é um dos aspectos em que a computação quântica pode prover soluções competitivas ao passo que tem a capacidade de lidar com dados de alta dimensionalidade de forma eficiente. Com isso, a busca por diferentes soluções e modelos quânticos para a tarefa de classificação é um caminho natural a ser seguido por muitos pesquisadores.

Contudo, mesmo com esse cenário favorável, ainda existem barreiras a serem transpassadas para que a aprendizagem de máquina quântica se torne uma alternativa real e prática para a resolução de problemas dentro do paradigma quântico de computação. Em (SILVER; PATEL; TIWARI, 2022), é argumentado que limitações ainda existentes nos dispositivos quânticos NISQ

dizem respeito à: (1) ausência de um dispositivo quântico com uma quantidade de qubits suficiente para realizar tarefas práticas e do mundo real e (2) falta de tolerância a erros. Com isso, resta aos pesquisadores produzir modelos que sejam aplicáveis a apenas alguns qubits, o que restringe o desenvolvimento de novas aplicações e algoritmos visando a exploração de bases de dados reais nos dispositivos atuais. Dito isto, muitas das abordagens existentes não são adequadas para serem executadas em um dispositivo pequeno e ruidoso. Assim, dois dos obstáculos presentes atualmente no contexto da aprendizagem de máquina quântica estão diretamente relacionados com as abordagens concebidas para esta tese:

1. O número de qubits disponíveis para realizar a computação;
2. O aumento do erro observado ao final da computação, o qual está associado à quantidade de portas lógicas usadas no circuito e à sua profundidade.

Estes entraves caracterizam parte do cenário atual da computação quântica, caracterizado pelos *dispositivos ruidosos de escala intermediária* (NISQ) (PRESKILL, 2018). Tal contexto torna imprescindível que as pesquisas não só foquem em dispositivos ideais (seja em tamanho ou nível de confiabilidade), mas também que esforços sejam direcionados para fazer uso do maquinário disponível atualmente, visando comprovar a efetividade das vantagens quânticas, sendo necessário investigar novas possibilidades de algoritmos que sejam adequados e demonstrem uma capacidade efetiva de classificar corretamente dados de conjuntos de dados reais.

O objetivo desta tese é propor abordagens quânticas para a tarefa de classificação supervisionada baseada em distância. A classificação quântica baseada em distância foi proposta em diversos trabalhos (CHALUMURI; KUNE; MANOJ, 2021; FENG et al., 2023; PIATRENKA; RUSEK, 2022; KOYA; LASKAR; DUTTA, 2023; LI et al., 2022; GAO et al., 2022; SHAO, 2020; SILVER; PATEL; TIWARI, 2022). Ainda, embora os métodos existentes tenham se mostrado promissores, nem todos consideram o cenário atual da tecnologia de dispositivos quânticos ruidosos de escala intermediária. Por exemplo, em (PIATRENKA; RUSEK, 2022) é apresentado um algoritmo com precisão de 71,1% em um dispositivo NISQ. Outros trabalhos utilizam sub-rotinas que não podem ser executadas nos dispositivos NISQ. Por exemplo, (FENG et al., 2023) requer uma estimativa de amplitude, e a versão quântica do KNN proposta por (LI et al., 2022) requer um algoritmo de busca para encontrar a distância mínima entre os dados. O uso de rotinas,

como estimativa de amplitude ou algoritmos de busca, tem seu uso prático em computadores quânticos atuais limitado por várias razões. Algumas delas, listadas a seguir:

- Os computadores quânticos atuais sofrem de erros quânticos significativos devido à decoerência e outros efeitos. Esses erros podem corromper os resultados das medições e das operações quânticas. A estimativa de amplitude é particularmente sensível a ruídos, pois requer uma série de medições e interferências sutis para construir a estimativa correta. Já os algoritmos de busca quântica normalmente requerem várias operações em sequência, aumentando a probabilidade de erro com a profundidade do circuito;
- A estimativa de amplitude muitas vezes requer circuitos quânticos profundos, com várias operações quânticas sequenciais. A profundidade do circuito pode aumentar o tempo de execução e a probabilidade de erro em cada operação. Nos computadores quânticos atuais a execução de circuitos muito profundos pode resultar em uma degradação significativa da qualidade do resultado;
- Embora o algoritmo de Grover ofereça uma aceleração quadrática na busca, a implementação prática desse algoritmo requer circuitos quânticos relativamente complexos. Isso inclui a aplicação de múltiplas portas quânticas, que precisam ser executadas de maneira precisa. A profundidade e complexidade dos circuitos podem levar a erros acumulados, comprometendo os resultados finais;
- Embora a estimativa de amplitude tenha um potencial teórico de aceleração em relação a algoritmos clássicos, a implementação prática desses algoritmos pode ser complexa. O número de medições necessárias para obter uma estimativa precisa da amplitude pode ser muito alto, exigindo um grande número de repetições do experimento quântico. Isso pode levar a uma quantidade de tempo e recursos computacionais que ainda não é viável com a tecnologia atual;
- Os computadores quânticos atuais têm um número limitado de qubits e uma conectividade restrita entre eles. A implementação de algoritmos complexos, como a estimativa de amplitude, pode exigir uma configuração de qubits que não está disponível. Essa limitação de hardware impede a execução eficiente de algoritmos quânticos mais sofisticados;

- A eficiência dos algoritmos de busca quântica geralmente se torna mais evidente em problemas de grande escala. No entanto, os computadores quânticos atuais têm um número limitado de qubits, o que restringe a capacidade de executar algoritmos em conjuntos de dados suficientemente grandes. Para obter o benefício total dos algoritmos quânticos, seria necessário um hardware com mais qubits e melhor conectividade.

Como dito em (SILVER; PATEL; TIWARI, 2022), existe uma lacuna no que se refere a modelos de aprendizagem de máquina quântica que lidem com a tarefa de classificação multiclasse. Com isso, atualmente, as pesquisas focam principalmente em prover e melhorar modelos de classificação binários. A melhor interpretação de como desenvolver classificadores quânticos multiclasse a partir de modelos de classificação binária permanece incerta. Ainda, a escolha de um *ansatz* variacional (um circuito que depende de parâmetros, que serão modificados via otimização clássica) eficiente é um problema em aberto.

Em um primeiro momento, investigamos o Hadamard Classifier (HC) (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) e, a partir da sua concepção, apresentamos um classificador mínimo aprimorado chamado Quantum One-Class Classifier (QOCC), que visa mitigar erros de dispositivos NISQ (PRESKILL, 2018). O QOCC, além de ser baseado no HC, se assemelha à operação de uma *Memória Quântica Probabilística* (Probabilistic Quantum Memory (PQM)) (TRUGENBERGER, 2001). Para o desenvolvimento do QOCC, reduzimos o HC removendo seus registros quânticos de classe e índice, mantendo a precisão. Além disso, o HC requer uma etapa adicional para realizar a classificação, chamada *pós-seleção* (ver Seção 3.2) e precisa de uma medição de 2 qubits para realizar a classificação, enquanto o QOCC requer apenas uma medição de 1 qubit. Essa modificação resulta em uma redução nos erros de medição. Tanto o HC quanto o QOCC são modelos quânticos de aprendizagem de máquina que usam interferência quântica para classificar novos dados de entrada.

Realizamos simulações e experimentos no IBM Quantum Experience (QUANTUM, 2023) para validar o QOCC. O primeiro experimento foi realizado em um ambiente de simulação ideal, sem a presença de erros. Em seguida, conduzimos experimentos em ambientes ruidosos (e, conseqüentemente, com erros), em dois processadores quânticos, especificamente o *ibmq_athens* e o *ibmq_santiago*. Os resultados mostraram que o QOCC é equivalente ao HC em termos de precisão, mesmo com um número menor de qubits e uma única medição, e tem precisão melhor ou equivalente em dispositivos quânticos reais. Assim, apresentamos o QOCC como um classificador mínimo que possui resultados competitivos em comparação com o HC

(mesmo com menos operações e qubits) e classificadores clássicos para os conjuntos de dados utilizados. Além disso, fornecemos uma atualização sobre o desempenho do HC em dispositivos quânticos atuais e explicamos a estratégia usada para a codificação de amplitude, bem como a metodologia usada para os experimentos, permitindo sua replicação adequada. Finalmente, realizamos experimentos em mais conjuntos de dados para sustentar nossos resultados.

Circuitos Quânticos Variacionais (Variational Quantum Circuit (VQC)), também conhecidos como *circuitos quânticos paramétricos*, fazem parte de uma teoria que aborda uma proposta para a realização de tarefas de aprendizado de máquina em dispositivos ruidosos em conjunto com técnicas presentes na aprendizagem de máquina clássica (MCCLEAN et al., 2016; BENEDETTI et al., 2019; CERZO et al., 2021). Os VQCs são compostos por portas lógicas quânticas paramétricas que passam por um procedimento de otimização híbrido quântico-clássico (CROOKS, 2019). Isso permite que o algoritmo tire vantagem do melhor dos recursos disponíveis, tanto do dispositivo quântico quanto do dispositivo clássico. Esse processo é repetido até que um critério de parada seja alcançado.

Assim, após a concepção do QOCC, nosso foco foi desenvolver o qVDCC (*quantum Variational Distance-based Centroid Classifier*). Com essa proposta, além de demonstrar explicitamente como realizar a tarefa de classificação multiclasse utilizando datasets reais, o classificador pôde ser executado em dispositivos NISQ. Ainda, mostramos que o qVDCC possui uma acurácia competitiva frente ao classificador clássico bem estabelecido KNN. Apesar de existirem outras propostas na literatura de classificadores quânticos multiclasse que guardam alguma semelhança com o qVDCC (CHALUMURI; KUNE; MANOJ, 2021; PIATRENKA; RUSEK, 2022), existe uma lacuna no que se refere a demonstrar, experimentalmente em dispositivo quântico ruidoso ou por meio de simulações, que esses classificadores podem competir com abordagens clássicas de classificação.

Portanto, o qVDCC, tem por objetivo demonstrar efetiva e explicitamente a capacidade quântica de lidar eficientemente com bases de dados reais, fornecendo competitividade com os modelos de classificação clássicos. Para o qVDCC, empregamos a teoria dos circuitos quânticos variacionais na construção de um modelo híbrido clássico-quântico de classificação multiclasse que busca encontrar o centroide de cada classe, objetivando realizar uma classificação de novos dados baseada na distância para estes centroides. Ainda, em se tratando de um modelo híbrido, o qVDCC tira vantagem das técnicas de treinamento bem-estabelecidas e eficientes da aprendizagem de máquina clássica. Uma outra vantagem do qVDCC é que ele não necessita de sub-rotinas complexas para funcionar, tais como algoritmos de busca. Essa é uma característica

adicional que permite que ele seja facilmente implementado nos dispositivos NISQ.

Com isso, nesta tese, aprimoramos a viabilidade e aplicabilidade dos classificadores quânticos baseados em distância. Classificadores quânticos baseados em distância, como o Hadamard Classifier (HC) (SCHULD; FINGERHUTH; PETRUCCIONE, 2017), são conhecidos por sua simplicidade, capacidade de alcançar compressão exponencial de dados e potencial para aceleração quântica exponencial quando os dados são eficientemente codificados em estados quânticos. No entanto, o circuito de codificação quântica de dados pode impor demandas significativas aos dispositivos quânticos atuais, e os classificadores quânticos baseados em distância existentes estão limitados à resolução de problemas binários (SCHULD; FINGERHUTH; PETRUCCIONE, 2017; BLANK et al., 2020; PARK; BLANK; PETRUCCIONE, 2020; BLANK et al., 2022). Para lidar com essas limitações, integramos o QOCC, nosso modelo de aprendizado de máquina quântico que requer menos recursos quânticos, com VQCs. Os VQCs identificam um centroide que representa os padrões de dados de treinamento, e o treinamento do classificador de uma classe para cada classe cria naturalmente um classificador multiclasse, expandindo sua variedade de aplicações. Nosso classificador proposto é avaliado por meio de simulações em várias bases de dados de referência, incluindo um ambiente sem ruído e um modelo realista de dispositivo quântico ruidoso. Os resultados dessas simulações demonstram aprimorada viabilidade do classificador quântico.

As principais contribuições específicas desta tese são as seguintes:

- **QOCC** (*Quantum One-Class Classifier*):
 - Formulação de um classificador de única classe baseado no *Hadamard Classifier* (HC) para classificação supervisionada baseada em distância em dispositivos NISQ;
 - Validação da proposta do classificador de única classe através de experimentos em máquinas quânticas ruidosas reais e em ambientes de simulação livres de ruído;
 - Comparação dos resultados do QOCC com o HC e com o KNN clássico.
- **qVDCC** (*quantum Variational Distance-based Centroid Classifier*):
 - Descrição explícita de um classificador multiclasse híbrido clássico-quântico capaz de lidar com conjuntos de dados reais com um número arbitrário de classes;
 - Execução da tarefa de classificação através de simulações com conjuntos de dados reais tanto em ambiente livre de ruído quanto através de simulações ruidosas;

- Modelo de classificação que busca encontrar um centroide artificial para cada classe de forma a minimizar sua distância até novos dados a serem classificados;
- Comparação dos resultados do qVDCC com o classificador clássico KNN.

1.1 MOTIVAÇÃO E PROPOSTA

Existem diversas propostas de classificadores baseados em distância que buscam se beneficiar das vantagens advindas da computação quântica (CHALUMURI; KUNE; MANOJ, 2021; FENG et al., 2023; PIATRENKA; RUSEK, 2022; KOYA; LASKAR; DUTTA, 2023; LI et al., 2022; GAO et al., 2022; SHAO, 2020; SILVER; PATEL; TIWARI, 2022). Porém, uma vez que não existem ainda dispositivos quânticos de larga escala, com grande capacidade de armazenamento e processamento, é fundamental pensar em modelos que consumam menos recursos quânticos e que sejam adequados para os dispositivos atuais. Portanto, modelos que utilizem menos qubits e menos operações quânticas são necessários para a realização de experimentos que visem demonstrar o poder da computação quântica.

A preocupação com relação ao número de operações quânticas necessárias para realizar determinada tarefa (seja ela de preparação de estados ou diretamente relacionada com a classificação em si) faz referência ao ruído presente nos computadores NISQ, que se acumula na medida em que a profundidade dos circuitos quânticos aumenta. Assim, devido ao estado corrente da tecnologia, as máquinas quânticas à disposição para a execução de experimentos acumulam erros que decorrem da dificuldade em manter a coerência quântica, o que decorre da aplicação das operações e do tempo de coerência que, em circuitos mais profundos, é algo que se agrava.

Inicialmente, direcionamos nossa atenção ao desafio de propor um classificador que fosse capaz de determinar o *grau de semelhança* de uma amostra de teste em relação a amostras de treinamento (ver Seção 4.1). Essa medida de similaridade é dada pela distância entre o vetor de teste e vetores de treinamento pertencentes a uma única classe, de forma que é possível recuperar alguma informação sem a imposição de haver um qubit de rótulo. Além disso, nossa proposta tinha como objetivo realizar a tarefa de classificação de forma simples e empregando o mínimo possível de recursos computacionais quânticos, visando diminuir erros provenientes dos ruídos presentes nos dispositivos NISQ e, conseqüentemente, aumentando a eficácia do modelo. O classificador em questão, chamado QOCC, foi construído e teve

seus resultados publicados na conferência International Joint Conference on Neural Networks (IJCNN) 2021 (OLIVEIRA et al., 2021).

Além disso, considerando que a literatura se concentra principalmente em classificadores binários, buscamos estender nossos esforços para o projeto de um classificador quântico baseado em distância que pudesse lidar com bases de dados contendo múltiplas classes. Decidimos adaptar nosso classificador de única classe, o QOCC, como uma célula para a classificação multiclasse (ver Seção 4.2). Esse novo classificador, chamado qVDCC, tem como principal característica o emprego da teoria de circuitos quânticos variacionais (paramétricos) com o uso de computação híbrida quântica-clássica para otimizar os parâmetros livres presentes no circuito responsável pela tarefa de classificação. Essa otimização de parâmetros visa definir centroides para as classes das bases de dados, realizando a classificação baseada na distância a partir desses centroides. Traçamos esse caminho para prover uma maior adaptabilidade ao classificador, permitindo que ele se mantenha simples, adequado para os dispositivos quânticos atuais, e com uma acurácia competitiva em relação aos análogos clássicos. Os resultados obtidos foram publicados no periódico *Neurocomputing* (OLIVEIRA et al., 2024).

1.2 OBJETIVOS

Abaixo, listamos, o objetivo geral desta tese e seus objetivos específicos associados.

1.2.1 Objetivo geral

O objetivo desta tese é investigar os modelos existentes de classificadores quânticos baseados em distância e, em seguida, propor novas abordagens simplificadas que sejam mais adequadas para execução nos atuais dispositivos quânticos ruidosos de escala intermediária (NISQ), alcançando resultados competitivos em comparação com os classificadores clássicos.

1.2.2 Objetivos Específicos

Os objetivos específicos desta tese são os seguintes:

- Investigar se a aprendizagem de máquina quântica tem gerado modelos de classificação que competem com os modelos clássicos existentes;

- Propor um classificador quântico mínimo de 1 classe baseado em distância para execução em dispositivos NISQ;
- Propor um classificador quântico multiclasse baseado em distância que utilize as capacidades dos circuitos quânticos variacionais para encontrar centroides das classes nas bases de dados, visando a execução em dispositivos NISQ;
- Demonstrar a eficiência e viabilidade de modelos de classificação quântica ou híbridos baseados em distância em dispositivos NISQ.

1.3 ORGANIZAÇÃO DESTA TESE

Esta tese está organizada de maneira a fundamentar todos os conceitos que circundam a temática em questão para, então, abordar os trabalhos diretamente relacionados e as soluções propostas. Dessa forma, o conteúdo restante desta tese está organizado nos seguintes quatro capítulos:

O Capítulo 2 fornece uma visão geral de conceitos básicos relacionados com a aprendizagem de máquina quântica (QML), definindo as características que compõem sua base teórica, delineando os conceitos imprescindíveis para o entendimento do trabalho aqui apresentado e servindo de base para o entendimento dos modelos de classificação apresentados, bem como para compreender de modo mais abrangente as motivações apresentadas na Seção 1.1. Mais especificamente, serão abordados neste capítulo:

- Fundamentos básicos da computação quântica;
- Descrição da codificação de dados clássicos nas amplitudes de registradores quânticos (*amplitude encoding*);
- Caracterização da computação híbrida quântica-clássica, bem como dos circuitos quânticos paramétricos;
- Apresentação de alguns dos modelos de aprendizagem de máquina quântica presentes na literatura.

Um levantamento e descrição de algumas abordagens existentes para classificadores quânticos baseados em distância (binários e multiclasse) são apresentados no Capítulo 3;

No Capítulo 4 temos as contribuições principais desta tese, juntamente com os resultados obtidos e uma discussão sobre os mesmos:

- Primeiramente, na Seção 4.1, é exposto o classificador quântico de única classe QOCC;
- Em seguida, na Seção 4.2, temos os detalhes do classificador quântico variacional multiclasse qVDCC.

Os comentários conclusivos e considerações finais serão encontrados no Capítulo 5, seguidos de uma discussão sobre os resultados gerais alcançados e possibilidades de continuidade do trabalho.

Por fim, temos a listagem das referências bibliográficas utilizadas nesta tese.

2 BASE TEÓRICA

Neste capítulo, serão apresentados, de forma abrangente, os principais conceitos e trabalhos que compõem o avanço do estado-da-arte no que tange à aprendizagem de máquina quântica (QML) (BIAMONTE et al., 2017; SCHULD; SINAYSKIY; PETRUCCIONE, 2015; WITTEK, 2014). Aqui, o objetivo está centrado em realizar uma revisão sistemática básica do que de mais relevante circunda a QML, de forma a exibir as técnicas utilizadas e os resultados obtidos de modo geral.

O capítulo está organizado da seguinte maneira:

- Definições elementares sobre a computação quântica serão expostas na Seção 2.1, incluindo *qubits* e portas lógicas quânticas (representação e ação);
- Uma das técnicas utilizadas para codificar dados clássicos nos registradores quânticos é apresentada na Seção 2.2;
- Na Seção 2.3, serão tratados os tópicos referentes à computação híbrida quântica-clássica e à teoria dos circuitos quânticos paramétricos;
- Por fim, na Seção 2.4, serão mencionados alguns algoritmos de aprendizagem de máquina quântica existentes na literatura.

2.1 COMPUTAÇÃO QUÂNTICA

Nesta seção, será apresentada uma breve revisão de conceitos essenciais para uma compreensão básica da computação quântica. Para uma introdução completa, sugerimos o texto da referência (NIELSEN; CHUANG, 2010).

Um computador quântico é uma máquina capaz de realizar cálculos e operações computacionais com base em propriedades inerentemente quânticas e é definido pelo critérios encontrados em (DIVINCENZO, 2000). Analogamente ao bit clássico, a unidade de informação quântica é o bit quântico, ou *qubit*. Valores lógicos “0”, “1” ou qualquer sobreposição destes podem ser atribuídos a um qubit. Essa superposição consiste em uma combinação linear dos estados da base computacional caracterizada por um vetor, conforme descrito na Equação (2.1), onde α e β são amplitudes complexas de probabilidade associadas aos respectivos estados e a propriedade $|\alpha|^2 + |\beta|^2 = 1$ é válida.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (2.1)$$

Uma das principais características da computação quântica quando comparada com a computação clássica é justamente a superposição de estados citada acima. Essa superposição permite que a computação quântica obtenha um alto grau de paralelismo computacional. Para um sistema quântico com n qubits, pode-se criar a superposição descrita na Equação (2.2), onde os α_i são as amplitudes complexas de probabilidade associadas aos estados $|i\rangle$. Assim, n qubits podem representar 2^n combinações de estados clássicos.

$$|\psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle \quad (2.2)$$

As operações sob um estado quântico são realizadas por operadores unitários. Assim, dado um operador $U : \mathbb{V} \rightarrow \mathbb{V}$, onde \mathbb{V} denota um espaço vetorial, U é considerado unitário quando sua inversa é igual ao seu conjugado transposto e \dagger indica o conjugado transposto. Ou seja, $UU^\dagger = U^\dagger U = I$, onde I designa o operador identidade. Desse modo, um operador atuando em n qubits precisa estar em um espaço vetorial complexo de dimensão 2^n . Isso significa que uma única aplicação de U realiza um número exponencial de operações nos estados da base, agindo linearmente sobre os vetores, como descrito na Equação (2.3).

$$U|\psi\rangle = U \left(\sum_{i=0}^{2^n-1} \alpha_i |i\rangle \right) = \sum_{i=0}^{2^n-1} \alpha_i U|i\rangle \quad (2.3)$$

Uma particularidade observada no modelo de computação quântica diz respeito às *medições*. Medições quânticas usam operadores que atuam sobre o estado quântico para determinar o resultado da computação que foi executada e, conseqüentemente, extrair alguma informação deste estado. O resultado de uma medição quântica será sempre colapsado para um único estado dentre os que estavam presentes na superposição quântica. Portanto, dado o estado mostrado na Equação (2.1), podemos usar os operadores de medição $M_0 = |0\rangle\langle 0|$ e $M_1 = |1\rangle\langle 1|$ para obter a probabilidade de que o resultado da medição seja 0 ($|\alpha|^2$) ou 1 ($|\beta|^2$), respectivamente. A probabilidade de obter a saída i é $p_i = |\alpha_i|^2$.

Circuitos quânticos são uma das formas disponíveis para representar e executar a computação quântica. São os circuitos quânticos que determinam em que ordem e quais operadores são aplicados a um ou mais qubits. Exemplos de operadores (ou, ainda, portas lógicas) quân-

ticos e suas ações respectivas incluem: X (2.4), *Hadamard* (2.5), R_y (2.6) e *Controlled-not* ($CNOT$) (2.7).

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{aligned} X|0\rangle &= |1\rangle \\ X|1\rangle &= |0\rangle \end{aligned} \quad (2.4)$$

$$H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{aligned} H|0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ H|1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{aligned} \quad (2.5)$$

$$R_y(\theta) = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{bmatrix} \quad (2.6)$$

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \begin{aligned} CNOT|0\rangle|0\rangle &= |0\rangle|0\rangle & CNOT|1\rangle|0\rangle &= |1\rangle|1\rangle \\ CNOT|0\rangle|1\rangle &= |0\rangle|1\rangle & CNOT|1\rangle|1\rangle &= |1\rangle|0\rangle \end{aligned} \quad (2.7)$$

Além desses operadores, há uma generalização da porta lógica quântica $CNOT$, que pode incluir mais de um qubit desempenhando a função de controle e mais de um qubit como alvo, além de poder aplicar um operador arbitrário U aos qubits alvo. O operador de controle geral é representado na Figura 1, onde $0, \dots, i$ são os índices dos qubits de controle e $i + 1, \dots, j$ são os índices dos qubits alvo.

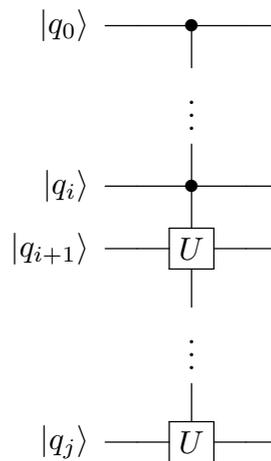


Figura 1 – Representação no circuito quântico de um porta lógica quântica controlada geral, onde $q_0 \dots q_i$ são os qubits de controle e $q_{i+1} \dots q_j$ são qubits alvo. O operador U é aplicado apenas aos qubits alvo se todos os qubits de controle tiverem valores verdadeiros (1).

2.2 CODIFICAÇÃO DE DADOS

Para a concepção de algoritmos de aprendizagem de máquina quântica, é crucial que se pense em como incorporar os dados a serem analisados nos circuitos quânticos (CEREZO et al., 2022). Uma vez que a quantidade de operações quânticas presentes em um circuito quântico é um parâmetro crítico para determinar a eficiência deste em retornar a resposta correta com o máximo de precisão possível, pensar em abordagens que busquem reduzir esse número de operações é uma tarefa primordial na computação quântica. Logo, estratégias eficientes de codificação de dados em registradores quânticos se tornam cada vez mais desejáveis. Trataremos, aqui, do método de codificação de estados clássicos em amplitudes quânticas (SCHULD; PETRUCCIONE, 2021; MOTTONEN et al., 2004; BERGHOLM et al., 2005). Em se tratando da codificação de dados em dispositivos quânticos, tem-se, em geral, circuitos quânticos que desempenham duas tarefas: (i) implementar operações quânticas genéricas e (ii) inicializar registradores quânticos. A segunda tarefa não é trivial porque o espaço de estados de um registrador de n qubits não é finito e contém sobreposições exponenciais de cadeias de bits clássicos.

Nas tecnologias de implementação existentes em dispositivos quânticos, portas lógicas que atuam em três ou mais qubits são difíceis de implementar diretamente. E essa é uma das motivações dos classificadores apresentados neste trabalho. Assim, implementar uma computação quântica como uma sequência de portas lógicas de dois qubits é de crucial importância. Estas,

por sua vez, podem ser decompostas em circuitos contendo portas de um qubit e uma porta de dois qubits padrão, geralmente a porta lógica quântica CNOT (do inglês, *Controlled-NOT*). Em (SHENDE; BULLOCK; MARKOV, 2006), é apresentada uma técnica assintoticamente ótima para a inicialização de um registrador quântico.

A codificação de estados clássicos em registradores quânticos, também chamada de *preparação de estados quânticos* é uma etapa crucial no desenvolvimento de aplicações em QML, uma vez que a eficiência em codificar dados nos qubits pode impactar da eficácia do algoritmo como um todo. Algoritmos de QML realizam análises em dados representados como um estado quântico. Assim, quando os dados são inerentemente quânticos, como os coletados via sensoriamento quântico, técnicas de QML podem ser aplicadas diretamente (CEREZO et al., 2022). Para dados inerentemente clássicos, é necessário primeiro mapeá-los para um estado quântico. Embora os modelos apresentados aqui nesta tese acomodem tanto dados clássicos quanto quânticos, focamos no caso clássico em nossas implementações, experimentos e simulações, de forma que isto requer uma etapa adicional de codificação de dados. Dentre os diversos métodos disponíveis para incorporação de dados em registradores quânticos (HAVLÍČEK et al., 2019; LAROSE; COYLE, 2020; PÉREZ-SALINAS et al., 2020; ARAUJO et al., 2021b; ARAUJO et al., 2023; VERAS; SILVA; SILVA, 2022), utilizamos a codificação dos dados nas amplitudes dos estados quânticos, que proporciona compressão exponencial de dados e tem sido fundamental para muitos algoritmos quânticos de aprendizado de máquina (GIOVANNETTI; LLOYD; MACCONE, 2008; HARROW; HASSIDIM; LLOYD, 2009; REBENTROST; MOHSENI; LLOYD, 2014).

Existem duas abordagens muito utilizadas para a codificação de dados em registradores quânticos. A primeira é aplicada em qubits individuais (estados separados). A segunda é nas amplitudes de um estado composto por vários qubits (estado entrelaçado). Para a primeira classe, existem dois algoritmos comuns: *codificação nos estados da base* e *codificação nos qubits*. Para cada uma dessas classes, há diversos algoritmos. A *codificação nos estados da base* consiste apenas em configurar o qubit como $|0\rangle$ ou $|1\rangle$. Ou seja, deve-se aplicar uma porta lógica X nos qubits que devem ser $|1\rangle$ para codificar um dado binário. Já o *a codificação nos qubits* (também conhecido com *angle encoding* -codificação por ou nos ângulos) codifica informação contínua na amplitude do qubit (um número real ou complexo). Esse procedimento é feito redimensionando os dados para um intervalo entre 0 e π , seguido de uma rotação do ângulo proporcional ao valor do dado.

A segunda classe exige um procedimento mais complicado, mas oferece uma vantagem exponencial sobre o equivalente clássico. Ou seja, a quantidade de informação armazenada

crece exponencialmente com a quantidade de qubits (2^n , para n qubits). Outra vantagem dessa segunda classe de algoritmos para codificação dos dados é que ela permite operações simultâneas nos dados armazenados, por causa da superposição. Veremos como incorporar dados nas amplitudes dos estados quânticos na seção seguinte

2.2.1 Codificação de dados clássicos nas amplitudes de registradores quânticos

Para realizar a codificação dos dados de entrada nas amplitudes do estado quântico, seguimos a estratégia apresentada em (SCHULD; PETRUCCIONE, 2018). Essa estratégia de codificação é adequada ao lidar com dados cujos valores pertençam a \mathbb{R} ou a \mathbb{C} .

A Figura 2 ilustra como partir do vetor desejado até o estado $|0\dots 0\rangle$. Para realizar a operação que transforma o estado inicial $|0\dots 0\rangle$ no estado quântico desejado que codifica os dados do padrão de entrada, realizamos as operações ilustradas na Figura 2 na ordem inversa e com a operação inversa.

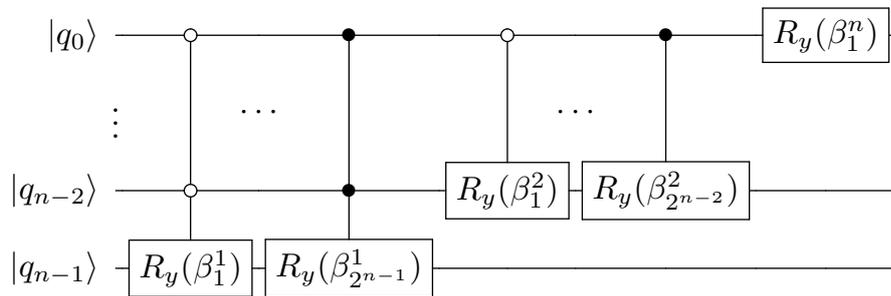


Figura 2 – Procedimento para codificar o vetor $|0\dots 0\rangle$ a partir de qualquer outro vetor real.

Ainda na Figura 2, cada β é um ângulo de rotação a ser realizado pela respectiva porta lógica quântica R_y e é calculado de acordo com a Equação (2.8), onde $s = \{1, \dots, n\}$, n é o número de qubits que armazenam 2^n atributos, $j = \{1, \dots, 2^{i-1}\}$ e $i - 1$ é o índice do qubit no qual a porta de rotação está sendo aplicada.

$$\beta_j^s = 2 \cdot \arcsen \left(\frac{\sqrt{\sum_{l=1}^{2^{s-1}} |\alpha_{(2j-1)2^{s-1}+l}|^2}}{\sqrt{\sum_{l=1}^{2^s} |\alpha_{(j-1)2^s+l}|^2}} \right) \quad (2.8)$$

A codificação de amplitude carrega dados de valores complexos nas amplitudes de probabilidade de um estado quântico. Isso é alcançado mapeando um vetor normalizado em um estado quântico. A codificação de amplitude pode ser implementada em um circuito quântico.

tico por meio de operadores de rotação multi-controladas, com os ângulos de rotação sendo determinados pelo vetor de dados alvo através da Equação (2.8), como mostrado na Figura 2.

A versão controlada dessa mesma porta lógica R_y é mostrada na Equação 2.9.

$$CR_y(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\left(\frac{\theta}{2}\right) & -\text{sen}\left(\frac{\theta}{2}\right) \\ 0 & 0 & \text{sen}\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} \quad (2.9)$$

Um exemplo de como a codificação de amplitude funciona é mostrado na Figura 3, que ilustra o procedimento sendo realizado para carregar um vetor de entrada contendo 4 atributos (amostra 20 da base de dados Iris) nas amplitudes do estado quântico, ou seja, $|\psi\rangle = -0.286|00\rangle + 0.723|01\rangle - 0.464|10\rangle - 0.425|11\rangle$.

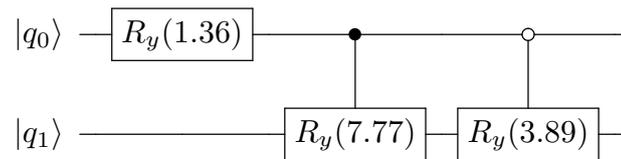


Figura 3 – Exemplo de como carregar um padrão nas amplitudes de um estado quântico.

É importante observar que as portas lógicas quânticas de rotação multi-controladas (SHENDE; BULLOCK; MARKOV, 2006)) representadas na Figuras 2 não são nativamente suportadas pelos dispositivos quânticos ruidosos comerciais atualmente disponíveis, criando uma imposição no seu uso. Portanto, para implementar a codificação de amplitude, o circuito quântico precisa ser decomposto em uma sequência de portas lógicas quânticas elementares de um ou dois qubits, como a porta lógica CNOT, por exemplo. Nas simulações e experimentos mostrados nesta tese, utilizamos essa abordagem para executar uma versão da porta de rotação multi-controlada. Essa decomposição pode ser vista na Figura 4.

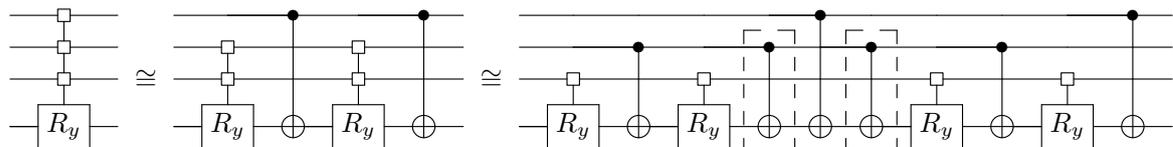


Figura 4 – Multiplexador quântico para uma porta R_y multi-controlada. Os CNOTs dentro das linhas tracejadas podem ser cancelados.

Já a Figura 5 mostra um exemplo utilizando essa abordagem de decomposição, onde incorpora-se quatro atributos de um padrão de dados em dois qubits, usando três ângulos de rotação independentes.

Na computação quântica, a incorporação de dados clássicos nas amplitudes dos estados quânticos é um passo fundamental para a implementação de algoritmos eficientes. Cabe pontuar que a fase (ou sinal) dessas amplitudes desempenha um papel crucial na interferência quântica, principalmente quando do uso de dados quânticos (complexos). Cada amplitude pode ser expressa na forma polar, onde a magnitude representa a probabilidade de encontrar o sistema em um determinado estado, enquanto a fase influencia a relação entre os diferentes estados quânticos. Portanto, ao calcular os ângulos de rotação necessários para codificar dados, é vital considerar como essas fases afetam a interferência quântica.

Quando se fala em rotação de estados quânticos, o uso de operadores quânticos, como os de rotação $R_x(\theta)$ e $R_z(\phi)$, permite manipular os estados de forma a alterar suas amplitudes e fases. O ângulo de rotação escolhido deve levar em conta a fase das amplitudes originais. A fase relativa entre os estados quânticos pode causar interferência construtiva ou destrutiva durante o processamento, o que pode amplificar ou reduzir a probabilidade de medir um resultado desejado. Se a fase não for corretamente considerada, pode ocorrer um cancelamento das probabilidades, levando a resultados imprecisos ou até errôneos na medição final. Além disso, a manipulação adequada das fases permite explorar a totalidade do espaço de estados quânticos, aproveitando o fenômeno da superposição. Essa capacidade de manipular a fase é uma das características distintivas da computação quântica e é essencial para o desempenho de algoritmos como o de Grover e o de Shor, onde a fase desempenha um papel central na eficiência da busca e na fatoração de números.

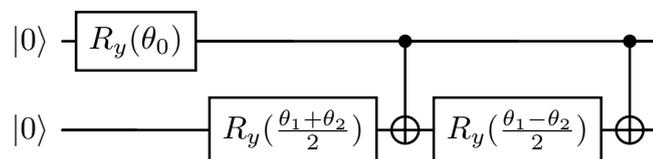


Figura 5 – Multiplexador quântico para uma porta R_y multi-controlada para incorporar dados clássicos de valores reais com quatro atributos em um estado quântico de dois qubits.

2.3 COMPUTAÇÃO HÍBRIDA E CIRCUITOS QUÂNTICOS PARAMÉTRICOS

Nos últimos anos, a aprendizagem de máquina quântica testemunhou o crescimento de abordagens que visam empregar recursos computacionais de dispositivos quânticos para resolver determinadas tarefas de aprendizado de máquina (WITTEK, 2014; SCHULD; PETRUCCIONE, 2018; DUNJKO; BRIEGEL, 2018; CILIBERTO et al., 2018; DUNJKO; TAYLOR; BRIEGEL, 2016). No entanto, uma questão em aberto neste contexto é como usar os atuais dispositivos quânticos ruidosos para demonstrar de maneira prática as vantagens dos métodos quânticos na resolução de certos problemas quando comparados com as propostas clássicas existentes. Uma das principais tarefas em aprendizado de máquina é o reconhecimento de padrões, onde há um algoritmo que deve ser capaz de prever o rótulo de uma nova entrada usando algum método que o relacione a um determinado conjunto de treinamento (BISHOP, 2006).

Tarefas de aprendizagem de máquina podem ser solucionadas de maneira mais eficiente em computadores quânticos devido à sua alta capacidade de paralelismo (TYCHOLA; KALAMPOKAS; PAPAKOSTAS, 2023), permitindo que, por exemplo, o conjunto de treinamento seja inteiramente posto em sobreposição. Essa característica diminui a quantidade de unidades de informação necessárias para se realizar a computação.

É importante saber que a implementação de uma abordagem dentro da aprendizagem de máquina quântica tem, principalmente, as seguintes etapas:

1. Codificação e mapeamento dos dados (geralmente clássicos, mas podendo ser puramente quânticos) nos registradores quânticos (diretamente nos qubits ou nas amplitudes dos estados);
2. O algoritmo quântico em si, através da execução dos circuitos quânticos (ou seja, uma transformação unitária);
3. Medições, análise e tratamento dos resultados obtidos;
4. No caso da computação híbrida, tem-se, ainda, uma etapa adicional de otimização de parâmetros (tipicamente realizada através do maquinário da aprendizagem de máquina clássica).

A geração atual de tecnologias quânticas de computação demanda algoritmos quânticos que requerem um número limitado de qubits e portas lógicas quânticas, além de serem robustos

contra erros. Portanto, devido aos atuais entraves dos dispositivos quânticos com relação ao número de qubits e à profundidade do circuito quântico que podem ser executados de forma confiável, é necessário criar estratégias para reduzir o tamanho de um circuito quântico (quantidade de operações e de qubits). Neste cenário, muitos algoritmos quânticos apresentam requisitos de recursos inviáveis quando comparados com o hardware que está disponível hoje e, com isso, tem ganhado força pesquisas que abordam a aprendizagem de máquina quântica a partir de uma concepção de computação híbrida, a qual beneficia-se das capacidades quânticas na mesma medida em que emprega técnicas clássicas bem estabelecidas para otimização de parâmetros. Além disso, a sobrecarga de alguns algoritmos assintoticamente ótimos é tal que mesmo os primeiros computadores quânticos competitivos com supercomputadores clássicos podem não ser capazes de executá-los (MCCLEAN et al., 2016).

Assim, surgiu a teoria dos *circuitos quânticos variacional* (MCCLEAN et al., 2016; BENEDETTI et al., 2019; CERZO et al., 2021; PERUZZO et al., 2014), uma estratégia flexível e adaptável para tirar proveito do hardware NISQ atual em aplicações do mundo real que consiste em circuitos quânticos com parâmetros livres que podem ser ajustados por meio de algoritmos de otimização clássicos (por exemplo, através de algoritmos baseados em gradientes). Circuitos quânticos variacionais foram primeiramente apresentados em (MCCLEAN et al., 2016) como componentes para a construção de algoritmos cuja execução fosse um híbrido, ora sendo executados no computador quântico, ora no clássico. VQCs surgiram com a filosofia de que mesmo recursos quânticos mínimos poderiam ser úteis quando usados em conjunto com rotinas clássicas. Também em (MCCLEAN et al., 2016), a teoria geral desse tipo algoritmo é estendida e são sugeridas melhorias algorítmicas para implementações práticas.

Uma maneira comum de utilizar um VQC é construir uma função de custo de um problema de otimização (mensurada pelas medições observadas no estado final do circuito) apropriada para um problema de otimização, como os encontrados na aprendizagem de máquina clássica, com base na medição de um circuito quântico de baixa profundidade. Com isso, o treinamento dos parâmetros das portas lógicas quânticas de rotação pode reduzir os erros presentes dos dispositivos NISQ. Portanto, o uso de VQCs é considerado uma abordagem adequada para o hardware quântico atual. A atualização dos parâmetros das portas quânticas em um VQC é realizada, tipicamente, por um computador clássico. Acredita-se que essa abordagem evite as limitações atuais dos dispositivos quânticos ruidosos, pois a baixa profundidade minimiza os efeitos do ruído e o treinamento dos parâmetros das portas lógicas pode compensar erros sistemáticos. Dessa forma, VQCs têm o potencial de serem aplicados de forma eficiente a

aplicações do mundo real em dispositivos quânticos e os circuitos quânticos parametrizados aplicados a certas tarefas de aprendizado de máquina (BENEDETTI et al., 2019) se configuram como uma alternativa tangível para validar as vantagens do aprendizado de máquina quântico em comparação com algoritmos análogos clássicos.

Em um esquema híbrido de computação quântica-clássica (BENEDETTI et al., 2019), existe o esforço em extrair o máximo de desempenho presente nas duas arquiteturas. De forma geral, temos a presença de um circuito quântico variacional (ou paramétrico) que possui portas lógicas quânticas de rotação com parâmetros livres. Em seguida, esse conjunto de parâmetros determina a saída do circuito a ser medida. É nesse momento que o maquinário clássico da aprendizagem de máquina exerce seu papel: uma vez definida uma função de custo, é realizada a otimização dos parâmetros, ajustando-os. Esse processo transforma a informação quântica em informação clássica. Dessa forma, mesmo sem empregar técnicas de correção de erro ou mitigação de ruído, é possível observar um desempenho competitivo dos algoritmos quânticos em comparação com seus análogos clássicos.

A otimização aqui citada consiste em um procedimento quântico-clássico híbrido onde a porção quântica é responsável pela preparação do estado quântico, pelo “modelo quântico” (circuito quântico variacional) e pelas medições. Já a parte clássica se encarrega de realizar o pré-processamento dos dados e receber a saída das medidas quânticas aplicando os resultados obtidos em algum algoritmo para ajuste dos parâmetros θ , calculando uma função de custo e realizando a otimização dos parâmetros. Os parâmetros atualizados são, então, usados para alimentar o circuito quântico novamente. Esse processo continua, em geral, até que algum critério de parada seja atendido. Os algoritmos quânticos variacionais são promissores pois habilitam que os circuitos sejam executados em dispositivos NISQ. Na Figura 6, podemos ver um esquema geral para implementação de um algoritmo de aprendizagem de máquina quântica que faz uso de circuitos quânticos paramétricos. É possível observar que esta é uma abordagem iterativa que se alterna entre os paradigmas quântico e clássico de computação.

A saída de um circuito quântico variacional como, por exemplo, o valor esperado de um observável, pode ser escrita como uma função quântica $f(\theta)$ parametrizada por $\theta = \theta_1, \theta_2, \dots$. Em muitos casos, a derivada parcial de $f(\theta)$ pode ser expressa como uma combinação linear de outras funções quânticas. É importante notar que essas outras funções quânticas, em geral, usam o mesmo circuito, diferindo somente por uma mudança (*shift*) do argumento. Isso significa que as derivadas parciais de um circuito variacional podem ser computadas usando a mesma arquitetura do circuito variacional.

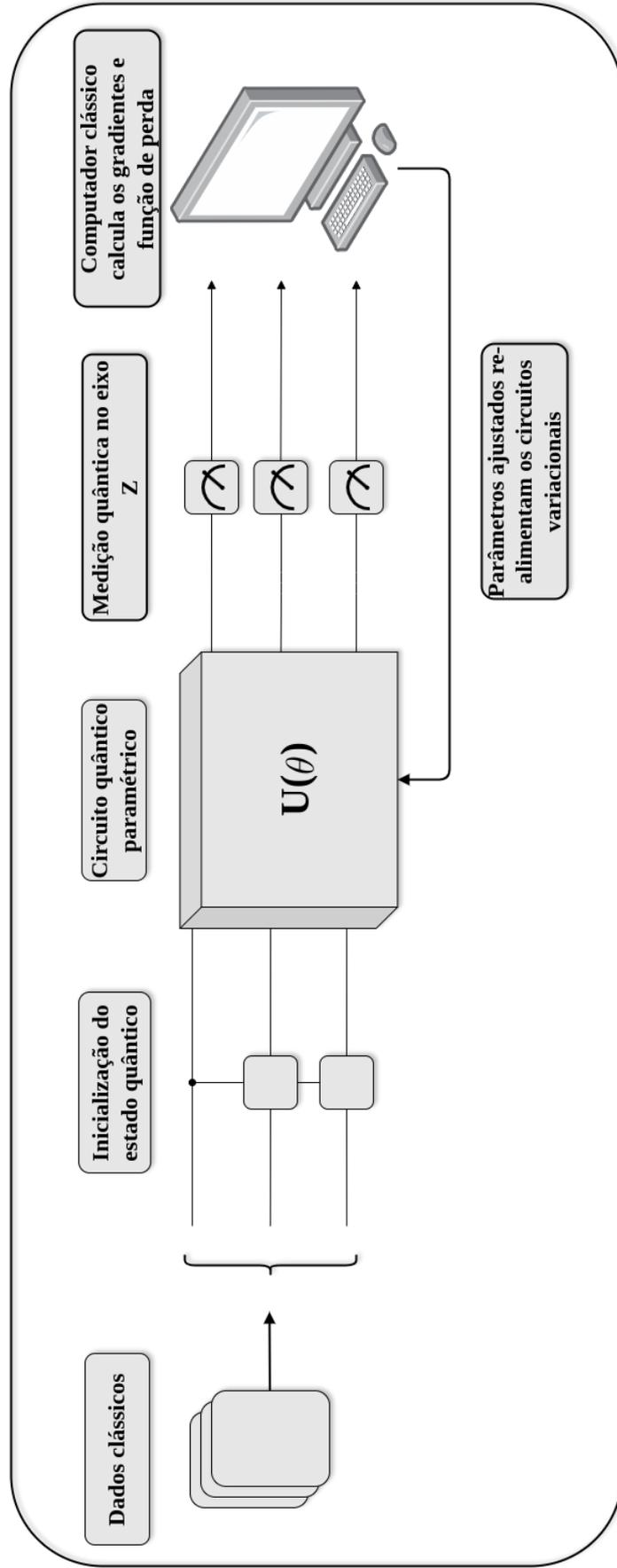


Figura 6 – Esquema geral do funcionamento de um algoritmo híbrido quântico-clássico. A execução passo a passo inclui (1) codificar os dados clássicos em registradores quânticos, (2) um modelo quântico, composto por portas lógicas quânticas paramétricas (com ângulos θ a serem otimizados), portas fixas e emaranhamento, (3) medição do circuito quântico, (4) avaliação de uma função de custo e (5) otimização dos parâmetros por meio de um computador clássico. Esse processo é repetido várias vezes para escolher os melhores ângulos θ em cada iteração.

Maneiras de como obter as derivadas parciais ao avaliar instâncias com mudanças de parâmetros de um circuito variacional são chamadas de regras *parameter-shift*, e foram primeiramente introduzidas em aprendizagem de máquina quântica em (MITARAI et al., 2018) e (SCHULD et al., 2019). Fazendo uma analogia aproximada com funções computáveis classicamente, isto é similar a como a derivada da função $f(x) = \text{sen}(x)$ é idêntica à Equação 2.10. Portanto, o mesmo algoritmo subjacente pode ser reutilizado para computar tanto $\text{sen}(x)$ quanto sua derivada (ao avaliar em $x \pm \frac{\pi}{2}$). Essa intuição é válida para muitas funções quânticas de interesse: *o mesmo circuito pode ser usado para computar ambos, a função quântica e o gradiente da função quântica.*

$$f(x) = \frac{1}{2}\text{sin}\left(x + \frac{\pi}{2}\right) - \frac{1}{2}\text{sin}\left(x - \frac{\pi}{2}\right) \quad (2.10)$$

Paralelamente ao surgimento dos circuitos quânticos variacionais, as técnicas de otimização de parâmetros baseadas em gradientes também foram aperfeiçoadas para se adequar aos algoritmos paramétricos. Graças a esse avanço, os métodos de mudança de parâmetros para computar gradientes, mais especificamente a “regra analítica de mudança de parâmetros” (tradução livre do inglês para *analytic parameter-shift rule*) (MITARAI et al., 2018; SCHULD et al., 2019), tornaram praticáveis e mais eficientes os experimentos envolvendo o treinamento de circuitos quânticos paramétricos no contexto da aprendizagem de máquina. Nesse algoritmo, os circuitos quânticos são construídos a partir de sequências de operadores (portas lógicas) quânticos unitários, conforme mostrado na Equação 2.11, onde x é o conjunto de parâmetros não-livres (também chamado de *dados de entrada*, com parâmetros não-livres) e $\theta = (\theta_1, \theta_2, \dots, \theta_N)$ é o conjunto de parâmetros ajustáveis. Neste trabalho, o gradiente analítico do valor esperado em relação a um parâmetro alvo é calculado justamente pelo algoritmo quântico usando a regra *parameter-shift*.

$$U(x; \theta) = U_N(\theta_N)U_{N-1}(\theta_{N-1}), \dots, U_1(\theta_1)U_0(x) \quad (2.11)$$

Assim, dado um circuito quântico variacional $U(\theta)$ e um observável \hat{B} , o valor esperado mostrado na Equação 2.12 tem sua derivada com relação aos parâmetros $\theta = (\theta_1, \theta_2, \dots, \theta_i)$ dada pela Equação 2.13, que é calculada com valores de parâmetros deslocados ao computar os gradientes do valor esperado do mesmo circuito quântico. Assim, os valores clássicos de θ e x são convertidos em informação quântica através do estado quântico $U(x; \theta)|0\rangle$.

$$\langle \hat{B} \rangle(\theta) = \langle 0|U(\theta)^\dagger \hat{B}U(\theta)|0\rangle \quad (2.12)$$

$$\nabla_{\theta_i} \langle \hat{B} \rangle(\theta) = \frac{1}{2} \left[\langle \hat{B} \rangle \left(\theta + \frac{\pi}{2} \hat{e}_i \right) - \langle \hat{B} \rangle \left(\theta - \frac{\pi}{2} \hat{e}_i \right) \right] \quad (2.13)$$

Ao final do circuito quântico, é necessário medir o valor esperado de um observável (que depende do problema) para obter a saída prevista. Uma etapa de pós-processamento é realizada para calcular a função de custo que será usada para atualizar os parâmetros θ do circuito quântico. O algoritmo da regra *parameter-shift* utiliza o mesmo circuito quântico para obter a previsão atual e seu gradiente. Portanto, são necessárias duas execuções por parâmetro para estimar corretamente a saída esperada do circuito.

Com o maquinário da aprendizagem de máquina clássica fornecida por bibliotecas de programação diferenciável, como TensorFlow (ABADI et al., 2015), PyTorch (PASZKE et al., 2017) e PennyLane (BERGHOLM et al., 2018), podemos treinar circuitos quânticos usando um computador quântico para calcular o valor esperado da saída e um computador clássico para realizar atualizações de parâmetros usando técnicas como o gradiente descendente (FARHI; NEVEN, 2018). Embora qualquer algoritmo clássico de otimização possa ser utilizado, optamos, nesta tese, por técnicas de gradiente descendente devido ao uso da regra *parameter-shift*, que fornece uma estimativa do gradiente.

2.4 ALGORITMOS DE APRENDIZAGEM QUÂNTICA

Existem vários modelos que empregam o conceito de aprendizado de máquina no contexto da computação quântica. No que diz respeito aos modelos de aprendizagem quântica, podemos realizar uma divisão entre dois grandes grupos: (1) modelos puramente quânticos e (2) modelos híbridos quântico-clássicos. Além disso, estes grupos englobam algoritmos pensados exclusivamente para o paradigma quântico e algoritmos que são análogos quânticos para os clássicos já existentes. Nesta seção, apresentaremos uma síntese de algumas das diversas abordagens de QML existentes na literatura.

Em (BLANK et al., 2020), é apresentado um classificador quântico binário baseado em métodos de kernel. Os autores demonstram que a computação quântica pode melhorar tarefas de classificação baseadas em distância na aprendizagem de máquina com base em métodos de kernel. Essa conclusão é obtida pela eficiência com que a computação quântica pode lidar

com espaços de características grandes. Em (PARK; BLANK; PETRUCCIONE, 2020), o qubit de rótulo é removido, reduzindo o classificador original proposto por (SCHULD; FINGERHUTH; PETRUCCIONE, 2017). Eles propõem um classificador quântico baseado em distância que utiliza o teste de SWAP para obter uma métrica de similaridade entre as amostras de treinamento e teste.

Em (RUAN et al., 2017), é proposta uma versão quântica do conhecido algoritmo KNN. Os autores utilizaram a distância de Hamming como medida de similaridade entre as amostras de teste e treinamento. Para realizar a classificação, as características do conjunto de dados são armazenadas em qubits quânticos. Para comprovar a eficácia do modelo, foram realizados experimentos utilizando a base de dados MNIST.

O classificador proposto em (KATHURIA et al., 2020), do mesmo modo que (RUAN et al., 2017), também utiliza a distância de Hamming como métrica de similaridade. Assim como em (BLANK et al., 2020), o teste de SWAP também foi usado para obter a medida de similaridade. Os experimentos foram realizados nos dispositivos quânticos reais *ibmqx2* e *ibmq_16_melbourne* da IBM Quantum Experience. Apesar de demonstrar a aplicabilidade e eficiência do classificador em alguns exemplos, os resultados mostram que a execução ainda não é totalmente viável nos dispositivos quânticos atuais.

Finalmente, durante o período de escrita desta tese, observamos o trabalho relacionado apresentado em (NEUMANN, 2020). Apesar de ser uma abordagem relacionada na qual os autores buscam construir um classificador mínimo, em (NEUMANN, 2020) ainda está presente a medida condicional mostrada no classificador em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017).

Willsch implementa em (WILLSCH et al., 2020) um modelo quântico para o bem conhecido algoritmo clássico de *máquinas de vetores de suporte* (SVM). Tal implementação passa pelo mapeamento do problema em um formato de problema de *otimização binária quadrática* (Quadratic Unconstrained Binary Optimization (QUBO)) e foi pensada para execução em dispositivos de arrefecimento quântico. Outra versão do algoritmo SVM pode ser vista em (REBENTROST; MOHSENI; LLOYD, 2014) e consiste em tirar proveito de técnicas de exponenciação de matrizes não esparsas para classificação de *big data*. Além disso, ainda em (REBENTROST; MOHSENI; LLOYD, 2014), o tempo de execução alcançado para esta versão quântica do algoritmo SVM é da ordem $O(\log(NM))$, para N características e M instâncias de treinamento. Porém, neste caso, é exigido que os dados sejam quânticos, ao invés de clássicos.

Em se tratando de redes neurais quânticas (Quantum Neural Network (QNN)), (SILVA; LU-

DERMIR; OLIVEIRA, 2016) introduziu um modelo quântico para o perceptron cujo aprendizado é realizado através de sobreposição de características. Uma alternativa quântica para o algoritmo clássico de regressão linear foi proposto por (SCHULD; SINAYSKIY; PETRUCCIONE, 2016). Os autores construíram o modelo com base no método de otimização dos mínimos quadrados e obtiveram um tempo de execução logarítmico com relação ao tamanho da entrada.

3 CLASSIFICADORES QUÂNTICOS BASEADOS EM DISTÂNCIA

Como visto no Capítulo 2.4, várias abordagens empregam o conceito de aprendizado de máquina quântica para realizar tarefas de aprendizagem de máquina. Neste capítulo, apresentaremos alguns dos trabalhos mais relacionados à nossa proposta. Ou seja, exibiremos modelos de aprendizagem focados na classificação, binária ou multiclasse, baseada em métricas de distância como parâmetro.

Este Capítulo está organizado da seguinte maneira:

- De forma mais específica, na Seção 3.1, serão expostos classificadores quânticos baseados em distância;
- Na Seção 3.2 veremos o funcionamento do classificador Hadamard.

3.1 CLASSIFICAÇÃO QUÂNTICA BASEADA EM DISTÂNCIA

O classificador proposto em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017), chamado aqui de *classificador Hadamard* foi usado como base para as nossas propostas. Sua definição, características e resultados obtidos serão detalhados na Seção 3.2. Além da proposta apresentada em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017), outros trabalhos já propuseram seus próprios classificadores quânticos baseados em distância (BLANK et al., 2020; PARK; BLANK; PETRUCCIONE, 2020; KATHURIA et al., 2020; RUAN et al., 2017; NEUMANN, 2020; HAVLÍČEK et al., 2019; SCHULD; SINAYSKIY; PETRUCCIONE, 2016).

Em (CHALUMURI; KUNE; MANOJ, 2021) é proposto um classificador quântico variacional chamado QMCC (*Quantum Multi-Class Classifier*). Os resultados obtidos foram demonstrados através de simulações em um ambiente que mimetiza o computador quântico, porém livre de ruído. As bases de dados utilizadas foram o *Iris*, *Banknote Authentication* e *Wireless Indoor Localization*. A acurácia para a base *Iris* foi de 92.10%, para o *Banknote* foi de 89.50% e para o *Wireless Indoor Localization* foi de 91.73%. Apesar dos resultados obtidos, o classificador proposto foi projetado para lidar com datasets de apenas 3 classes. Os autores argumentam ser possível estender o modelo para classificar um número arbitrário de classes, mas isso não foi mostrado explicitamente.

O algoritmo QKNN (*Quantum K-Nearest Neighbor*), proposto em (FENG et al., 2023), foi construído para lidar com o problema de desempenho do KNN clássico quando da presença

de *big data*. O QKNN substitui a distância Euclidiana por uma nova medida de similaridade baseada no *sistema de coordenada Polar*, chamada de *distância Polar*. Foram feitos experimentos com as bases de dados *Iris*, *Wine*, *Liver* e *Overflow Vulnerability*. Os autores compararam os resultados obtidos entre o uso da distância Euclidiana e a distância Polar no KNN e no QKNN. Os resultados médios do QKNN com a distância Polar foram: *Iris* (95.82%), *Wine* (95.86%), *Overflow* (89.19%) e *Liver* (63.42%). Para determinar a saída do classificador, os autores usaram a técnica de *amplificação de amplitude*, que não é viável no hardware quântico atual.

Em (PIATRENKA; RUSEK, 2022), é proposto um classificador quântico variacional multi-classe a ser aplicado na base de dados *Iris*. Os autores realizaram experimentos e simulações para verificar a acurácia obtida e atestar a qualidade do modelo apresentado. Na construção do modelo, foi empregado o uso do circuito **RealAmplitudes**, disponível no Qiskit, como *ansatz* no modelo quântico para a tarefa de classificação. Para analisar os resultados, os autores utilizaram 3 tipos de circuitos quânticos também presentes no Qiskit para realizar o data encoding: (ZFeatureMap, ZZFeatureMap e PauliFeatureMap). Além disso, foram usados os seguintes otimizadores: SPSA (SPALL, 1992), COBYLA (*Constrained Optimization by Linear Approximation*) (POWELL, 1994) e SLSQP (*Sequential Least Squares Programming*) (KRAFT, 1988). Todas as combinações testadas obtiveram uma acurácia maior que 90% em um ambiente de simulação sem ruído. Os experimentos no computador quântico ocorreram somente na fase de testes. Os resultados na máquina quântica real alcançaram, no máximo, uma acurácia de 71.1%. Não foram realizadas comparações com modelos clássicos.

O QKNN (KOYA; LASKAR; DUTTA, 2023) é uma versão quântica modificada do algoritmo clássico KNN para detecção de símbolo em um sistema de comunicação no contexto da tecnologia veicular. Os autores usaram o *SWAP-test* para calcular a distância (também chamada de *fidelidade quântica*) entre os dados de treinamento e de teste. Foram feitas simulações para demonstrar a eficiência do modelo apresentado. O algoritmo proposto faz uso de uma *técnica de atualização de memória baseada em estratégia*, que os autores argumentam ser melhor que outras técnicas, uma vez que elimina a busca quântica. Além disso, foi realizada uma análise com relação à complexidade de consulta, de tempo e de sub-rotinas quânticas, comparando o QKNN proposto com os resultados obtidos em (BASHEER; GOYAL et al., 2020), (DANG et al., 2018) e (LI et al., 2022), tendo exibido vantagens perante estes.

Os trabalhos de (LI et al., 2022) e (GAO et al., 2022) propuseram versões quânticas do algoritmo clássico KNN baseadas, respectivamente, nas distâncias de Hamming e de Mahalanobis.

Na interpretação quântica do KNN proposta por (LI et al., 2022), o modelo está centrado em um algoritmo de busca para encontrar a distância mínima entre os dados. Os autores focaram em demonstrar a eficácia do algoritmo ao analisar que a complexidade de tempo mostrou uma aceleração quadrática com relação ao análogo clássico. Contudo, essa aceleração só é alcançada quando os padrões de entrada estão em um espaço de características de baixa dimensão. Tanto para o cálculo da distância de Hamming quanto para a busca pelo mínimo de uma sequência desordenada, os sub-algoritmos proposto pelos autores necessitam fazer uso de um oráculo, de um algoritmo de busca quântico e de uma QRAM (*Quantum Random Access Memory*) (GIOVANNETTI; LLOYD; MACCONE, 2008). De forma similar, (GAO et al., 2022) aborda a construção de um classificador quântico baseado no KNN, porém usando como métrica a distância Mahalanobis. A proposta dos autores consiste em unir o KNN clássico também com um algoritmo quântico de busca para realizar a tarefa de classificação. O diferencial do classificador proposto em (GAO et al., 2022) é o uso de distância Mahalanobis, que considera as correlações entre atributos individuais nos dados de entrada. Os resultados da análise de complexidade mostraram uma aceleração quadrática em termos de complexidade de tempo em comparação com o KNN clássico. Não foram realizados simulações ou experimentos com bases de dados reais em ambos os trabalhos.

Shao (SHAO, 2020) propôs um modelo quântico de rede Radial Basis Function (RBF) cuja velocidade de treinamento é acelerada em quase um fator quadrático com relação ao análogo clássico. Os resultados alcançados foram demonstrados com experimentos realizados através de um modelo de rede RBF clássico generalizado. O autor propôs dois modelos de redes RBF: (1) uma rede RBF clássica generalizada que é considerada equivalente a uma rede RBF quântica e (2) uma rede RBF quântica para tarefas de classificação binária que usa técnicas algébricas lineares quânticas para acelerar o treinamento. O treinamento da rede quântica RBF foi realizado utilizando um esquema híbrido quântico-clássico. Devido à ausência de um computador quântico em larga escala, o desempenho da rede quântica RBF foi medido resolvendo problemas de classificação binária usando a rede RBF clássica generalizada. Embora o RBF quântico apresentado em (SHAO, 2020) tenha demonstrado alta eficácia em tarefas de classificação, existem algumas questões que precisam ser abordadas em relação aos computadores quânticos atuais (ou seja, dispositivos NISQ). Em primeiro lugar, para calcular a saída quântica de RBF, foi utilizada a técnica de estimativa de amplitude. Em segundo lugar, para implementar os dados de entrada no computador quântico, (SHAO, 2020) aplicou as técnicas algébricas lineares quânticas de codificação de blocos e estados coerentes. Portanto, embora

tais técnicas permitam acelerar o procedimento de treinamento, a rede RBF quântica proposta ainda não é adequada para os dispositivos quânticos atuais.

Em (SILVER; PATEL; TIWARI, 2022) é apresentado um *framework* para classificação multiclasse de imagens chamado Quilt, no qual foi construído um *ensemble* de classificadores quânticos. Os autores demonstraram seu desempenho na base de dados *MNIST*, tendo alcançado 85% de acurácia, além de rodar o *framework* para as bases *Fashion-MNIST* e *Cifar*. O objetivo dos autores foi permitir classificação multiclasse ao mesmo tempo que mantém os classificadores simples provendo, ainda, um aumento na acurácia. O Quilt possui um mecanismo de correção de erro para classificadores binários, o que aumenta a probabilidade de predição correta do *ensemble* de classificadores. O processo de classificação do Quilt requer vários estágios para determinar a saída. Para lidar com bases de dados de alta dimensionalidade, o modelo proposto usa o algoritmo PCA (*Principal Component Analysis*). Os resultados foram obtidos através de simulações com e sem ruído, além de experimentos em uma máquina quântica real.

Um trabalho que enfoca a teoria e o desenvolvimento de classificadores quânticos binários baseados em *métodos de kernel* é apresentado em (BLANK et al., 2020). Os autores mostram que a computação quântica pode melhorar as tarefas de classificação em aprendizado de máquina que são baseadas no conceito do método kernel. Esta conclusão é obtida observando a eficiência com a qual a computação quântica pode lidar com grandes espaços de características. Também empregando métodos de kernel, (PARK; BLANK; PETRUCCIONE, 2020) propõe um classificador quântico baseado em distância que faz uso do teste SWAP para obter uma métrica de similaridade entre as amostras de treinamento e teste. Além disso, os autores mostraram como usar o classificador proposto em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) para realizar a tarefa de classificação sem a necessidade de um qubit de classe.

Uma versão quântica para o conhecido algoritmo KNN é apresentada em (RUAN et al., 2017), que usa a distância de Hamming como uma medida de similaridade entre as amostras de teste e treinamento. Para realizar a classificação, os atributos do conjunto de dados são armazenados em bits quânticos. O classificador proposto em (KATHURIA et al., 2020) também usa a distância de Hamming e o teste SWAP para classificar novas entradas. No entanto, o classificador binário proposto em (KATHURIA et al., 2020) tem um propósito mais específico, visando identificar uma doença a partir de um atributo genômico funcional. Apesar de demonstrar a aplicabilidade e eficiência do classificador em alguns exemplos, os resultados mostram que a execução ainda não é totalmente viável nos dispositivos quânticos atuais.

Em (HAVLÍČEK et al., 2019), dois métodos são propostos para serem implementados em processadores supercondutores: primeiro, classificadores quânticos variacionais são usados para realizar uma classificação análoga ao SVM clássico; o segundo método consiste em um estimador de kernel quântico a ser otimizado para então implementar um SVM clássico. O kernel é estimado através da amplitude de transição para estabelecer vetores de suporte e classificar novos dados.

Uma alternativa quântica para o algoritmo clássico de regressão linear foi proposto por (SCHULD; SINAYSKIY; PETRUCCIONE, 2016). Os autores construíram o modelo com base no método de otimização dos mínimos quadrados e obtiveram um tempo de execução logarítmico com relação ao tamanho da entrada.

Apesar de algumas das propostas apresentadas aqui nesta seção serem promissoras, elas não preenchem todas as lacunas elencadas para esta tese. É possível observar que parte desses modelos de aprendizagem existentes deixa a desejar no que se refere à viabilidade de execução nos dispositivos NISQ, por necessitarem de rotinas adicionais que não são adequadas para serem executadas nos dispositivos quânticos atuais. Ainda, uma outra parcela foca unicamente na tarefa de classificação binária, sem dar uma descrição explícita de como adaptar os modelos para a tarefa de classificação multiclasse. Por fim, existe uma carência de comparações dos resultados obtidos com os modelos quânticos em relação a modelos clássicos.

3.2 CLASSIFICADOR HADAMARD

O *classificador Hadamard* (HC) foi introduzido em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) como um classificador binário minimalista que pode ser construído usando um circuito quântico. A estratégia usada no HC é utilizar a *codificação de amplitude* para codificar as características dos padrões de entrada e realizar interferência quântica para determinar a classificação de um novo vetor de entrada avaliando a distância deste para os dados de treinamento (dois padrões conhecidos, um de cada classe). O HC é um modelo de aprendizado de máquina quântico que pode ser implementado em dispositivos NISQ. Para validar o HC, os autores realizaram experimentos de classificação supervisionada usando o conjunto de dados Iris (FISHER, 1936) (disponível no Scikit-learn (PEDREGOSA et al., 2011)).

O processo de classificação do HC começa com o estado quântico descrito na Equação (3.1), onde o registrador de índice $|m\rangle$ enumera os dados de treinamento, $|\psi_{\mathbf{x}^m}\rangle$ representa o m -ésimo padrão de treinamento, $|\psi_{\mathbf{x}}\rangle$ corresponde à entrada de teste, e $|y^m\rangle \in \{|0\rangle, |1\rangle\}$

é o qubit de rótulo.

$$|\mathcal{D}\rangle = \frac{1}{\sqrt{2M}} \sum_{m=1}^M |m\rangle (|0\rangle|\psi_{\bar{x}}\rangle + |1\rangle|\psi_{\mathbf{x}^m}\rangle) |y^m\rangle \quad (3.1)$$

Observe que a codificação de amplitude (GIOVANNETTI; LLOYD; MACCONE, 2008; ARAUJO et al., 2021b; Veras et al., 2020; ARAUJO et al., 2021a) descrita na Seção 2.2.1 permite que um vetor clássico de dimensão N seja mapeado em $\lceil \log_2(N) \rceil$ qubits. Após a codificação do estado, a porta lógica quântica Hadamard é aplicada ao segundo registrador do estado quântico $|\mathcal{D}\rangle$, resultando no estado transformado mostrado na Equação 3.2.

$$|\mathcal{D}'\rangle = \frac{1}{2\sqrt{M}} \sum_{m=1}^M |m\rangle (|0\rangle (|\psi_{\bar{x}}\rangle + |\psi_{\mathbf{x}^m}\rangle) + |1\rangle (|\psi_{\bar{x}}\rangle - |\psi_{\mathbf{x}^m}\rangle)) |y^m\rangle \quad (3.2)$$

A principal limitação do HC é uma medição condicional (ver passo F da Figura 7), chamada de *pós-seleção*, que depende da probabilidade de medir $|0\rangle$ no *qubit auxiliar*. Uma medição é realizada no *qubit de classe* apenas após o sucesso da pós-seleção. Resultados experimentais mostraram uma precisão de 100% para as classes 1 e 2 da base de dados *Iris*. No entanto, devido à dependência da pós-seleção, são necessárias mais repetições para obter uma estimativa realista do resultado. Além disso, com a presença dessa pós-seleção, é necessário realizar uma medição de 2 qubits, o que pode aumentar a probabilidade de erros nos dispositivos quânticos atuais sujeitos a ruído.

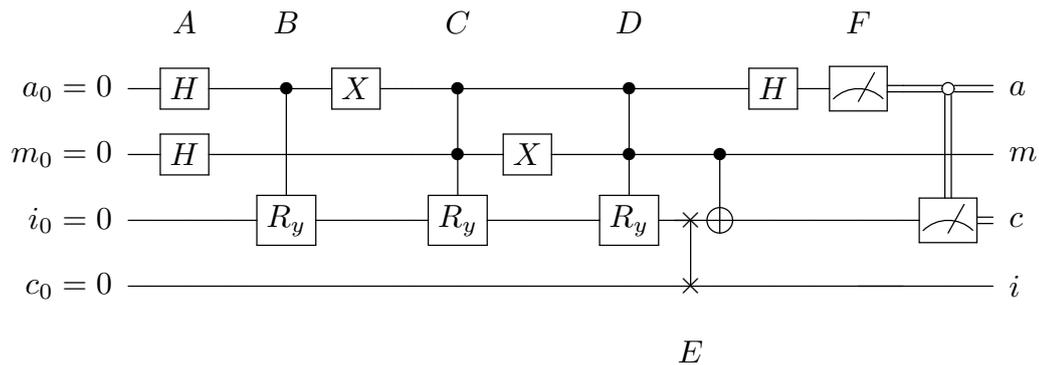


Figura 7 – Exemplo do classificador Hadamard. No passo A, os qubits auxiliar (a_0) e de índice (m_0) são postos em uma sobreposição uniforme. No passo B o vetor de teste é carregado e nos passos C e D os vetores de treinamento são incorporados às amplitudes do estado quântico emaranhados com o qubit de índice (i_0). No início do passo E, os qubits de índice (i_0) e de classe (c_0) são trocados de posição devido ao mapeamento físico dos qubits no hardware que foi usado para os experimentos do HC. No passo F, há a interferência sendo realizada entre o vetor de teste e os de treinamento através do operador Hadamard, seguida da medição (condicional) de 2 qubits (auxiliar a_0 e de classe c_0).

A probabilidade de sucesso dessa pós-seleção, p_{acc} , é dada pela Equação 3.3, onde $\tilde{\mathbf{x}}$ representa os dados de teste e \mathbf{x}^m representa o m -ésimo dado de treinamento. O passo final do algoritmo HC envolve a medição do último registrador da Equação 3.2 (ou seja, o qubit de rótulo), o que resulta na atribuição de um rótulo ao padrão de teste. Portanto, como já dito, o HC requer uma operação de medição de dois qubits para classificação, como observado em (BLANK et al., 2020; PARK; BLANK; PETRUCCIONE, 2020).

$$p_{acc} = \frac{1}{4M} \sum_m |\tilde{\mathbf{x}} + \mathbf{x}^m|^2 \quad (3.3)$$

Como a probabilidade de sucesso da pós-seleção é a probabilidade de medir $|0\rangle$ no qubit auxiliar, para alcançar esse resultado é necessário realizar o traço parcial para um dado estado. Quando lidamos com sistemas compostos (por exemplo $|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$), como dois qubits, a probabilidade de medir um resultado em um qubit depende do estado do sistema como um todo. Para medir o primeiro qubit e determinar a probabilidade de obter o resultado $|0\rangle$ utilizamos $P(0) = \langle\psi|P_0|\psi\rangle$ com o operador projetor correspondente a esse resultado, que é $P_0 = |0\rangle\langle 0|\otimes I$, onde I é o operador identidade que age sobre o segundo qubit. Essa abordagem nos permite focar apenas na parte do estado quântico que está relacionada ao primeiro qubit.

Para calcular essa probabilidade, transformamos o estado quântico em uma matriz densidade ρ , que encapsula todas as informações sobre o sistema. A partir dessa matriz densidade, aplicamos a operação de traço parcial sobre o segundo qubit. O traço parcial nos permite "ignorar" o segundo qubit, resultando em uma matriz densidade reduzida que descreve apenas o primeiro qubit. Essa matriz densidade reduzida contém informações sobre as amplitudes associadas aos estados do primeiro qubit e é fundamental para determinar a probabilidade de resultados em medições.

Finalmente, a probabilidade de medir $|0\rangle$ no primeiro qubit é obtida somando as amplitudes correspondentes na matriz densidade reduzida. Especificamente, essa probabilidade é dada pela soma dos quadrados das amplitudes associadas aos estados $|00\rangle$ e $|01\rangle$. Assim, a regra de medição, aliada ao conceito de traço parcial, fornece uma ferramenta poderosa para analisar e interpretar medições em sistemas quânticos complexos.

4 CLASSIFICAÇÃO QUÂNTICA E HÍBRIDA BASEADA EM DISTÂNCIA

Neste capítulo, serão apresentadas as principais contribuições desta tese. A estrutura do capítulo está organizada da seguinte maneira:

- Na Seção 4.1 será apresentado o *Quantum One-Class Classifier* (QOCC), um modelo de classificação quântico de única classe que utiliza recursos computacionais quânticos mínimos, visando reduzir a probabilidade de erros interferirem na acurácia do classificador;
 - A Seção 4.1.1 contém detalhes dos experimentos realizados e dos resultados obtidos;
 - Uma discussão sobre os resultados alcançados é feita na Seção 4.1.2;
 - Ao final da seção, tem-se o link com os dados e software utilizados no desenvolvimento do QOCC.

- O classificador quântico variacional multiclasse *quantum Variational Distance-based Centroid Classifier* (qVDCC) será exposto na Seção 4.2;
 - A Seção 4.2.1 trata da abordagem paramétrica do circuito quântico do qVDCC;
 - A modelagem multiclasse do classificador é retratada na Seção 4.2.2;
 - Na Seção 4.2.3 é apresentada a motivação por trás da escolha de utilizar uma porta lógica quântica parametrizada ao final do circuito, substituindo a porta lógica quântica *Hadamard*, presente no QOCC;
 - A descrição das simulações realizadas e a exposição dos resultados obtidos estão presentes na Seção 4.2.4;
 - Ao final da seção, tem-se o link com os dados e software utilizados no desenvolvimento do qVDCC.

4.1 QOCC - QUANTUM ONE-CLASS CLASSIFIER

A pós-seleção do classificador Hadamard (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) (ver Seção 3.2) é bem-sucedida com a probabilidade descrita na Equação 4.1. Essa probabilidade depende da distribuição dos dados e pode tender a zero. Além disso, essa etapa de pós-seleção

determina quais dados serão, de fato, classificados dentre as classes existentes na base de dados.

$$p_{acc} = \frac{1}{4M} \sum_m |\tilde{\mathbf{x}} + \mathbf{x}^m|^2 \quad (4.1)$$

A Figura 8 apresenta um conjunto de dados artificial onde a probabilidade de pós-seleção é aproximadamente 0,02 para o padrão \mathbf{x}_0 e 0,98 para o padrão \mathbf{x}_1 . A probabilidade de pós-seleção é uma função da distância Euclidiana do novo padrão a ser classificado em relação aos padrões do conjunto de dados armazenados no classificador e retorna 0 com uma probabilidade maior se o novo padrão estiver próximo aos padrões no conjunto de dados. É importante ressaltar que o problema de pós-seleção encontrado em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) já foi abordado em (BLANK et al., 2020), onde foi mostrado que o resultado da medição de apenas um único qubit é necessário para determinar a classe. Assim, a modificação feita aqui nos permite abstrair a questão da distribuição de dados, que poderia levar a uma probabilidade muito baixa de pós-seleção.

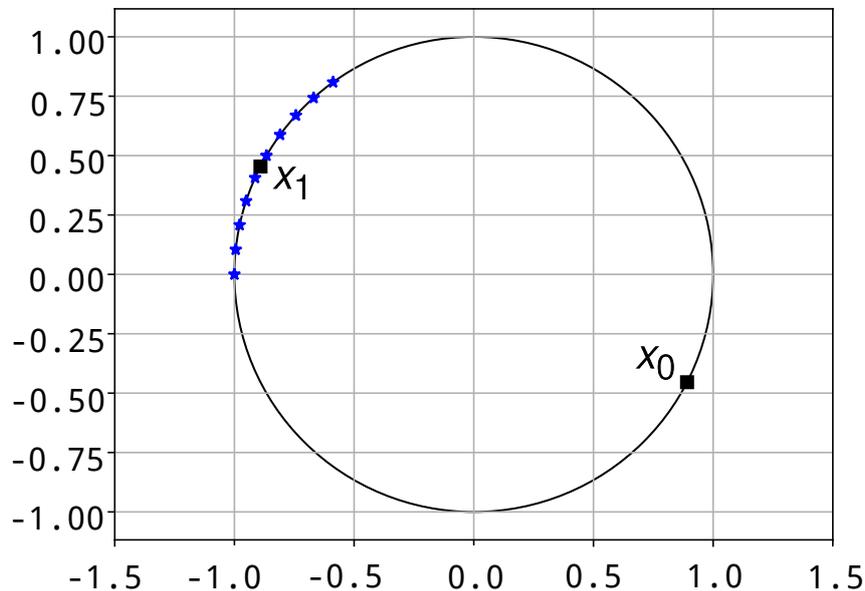


Figura 8 – Um exemplo artificial em que a probabilidade de pós-seleção tende a zero para um determinado novo padrão de entrada \mathbf{x}_0 da classe 0. As estrelas são amostras do conjunto de dados e têm norma unitária. O quadrado preto \mathbf{x}_0 (\mathbf{x}_1) é uma amostra da classe 0 (1) a ser classificada. A pós-seleção é realizada com base na distância entre a amostra a ser classificada e todas as outras amostras no conjunto de dados. Neste exemplo, a amostra \mathbf{x}_0 não seria classificada porque a abordagem em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) requer a saída 0 na pós-seleção para prosseguir com a classificação. A saída 0 na pós-seleção só é alcançada se os dados a serem classificados estiverem próximos aos outros dados no conjunto de dados.

Assim, para construir o classificador QOCC, redefinimos o classificador Hadamard para usar o resultado da pós-seleção como a saída do circuito quântico que determina a classificação. Com isso, removemos o qubit de classe, reduzimos o número de repetições necessárias para estimar a saída do classificador e diminuimos o número de operações necessárias para realizar a classificação.

O QOCC é um classificador binário e a tarefa de classificação supervisionada é realizada na presença dos dados de apenas 1 das classes possíveis, indicando se o novo padrão a ser classificado pertence ou não a esta única classe. Portanto, o QOCC é baseado no HC e amplia sua aplicabilidade para permitir a classificação de um novo padrão de entrada a partir de um classificador quântico de única classe indicando a probabilidade de que o padrão esteja associado ao conjunto de padrões carregados (treinamento) no classificador. Essa característica nos permite executar o classificador com menos repetições para obter o resultado e realizar apenas uma medição de 1 qubit, em vez da medição de 2 qubits presente no HC. Conforme apontado em (KIM; PARK; RHEE, 2020), a taxa de erro presente nos dispositivos quânticos atuais pode ocultar o potencial da computação quântica. Portanto, a redução de uma medição de 2 qubits para uma medição de 1 qubit também representa uma tentativa de mitigar os erros dos atuais dispositivos NISQ sem necessariamente fazer uso de um procedimento específico para isso.

Em nosso modelo, utilizamos a mesma estratégia de pré-processamento de dados utilizada em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017). No entanto, a classificação de um novo padrão de entrada funciona de maneira semelhante a uma memória quântica probabilística (TRUGENBERGER, 2001). Essa conexão com memórias quânticas probabilísticas decorre da semelhança na forma como a saída é definida: um novo padrão de entrada é comparado com as amostras já presentes no classificador, e a saída é dada por uma distribuição de probabilidade. Assim, quanto mais distante (diferente) o novo padrão e as amostras armazenadas estiverem entre si, maior a probabilidade de observarmos $|1\rangle$ ao medir o qubit auxiliar. Por outro lado, se a similaridade entre eles for alta, teremos na saída uma maior probabilidade de observar $|0\rangle$ na medição do qubit auxiliar. Portanto, um novo padrão de entrada é classificado de acordo com um *grau de semelhança* desse novo padrão em relação aos padrões já armazenados no classificador. Esse grau de semelhança é a probabilidade de medir 0 na saída onde anteriormente era feita a pós-seleção. Portanto, ao carregar amostras de treinamento de uma única classe, o classificador proposto mostra-se mais flexível ao simplificar a classificação/associação de um novo padrão de entrada em qualquer classe.

A Figura 9 mostra o circuito do QOCC recebendo entradas com duas características. Na etapa E, a saída calculada depois de sucessivas medições representa o grau de semelhança do novo padrão de entrada com relação aos padrões da classe c representada pelo classificador. Portanto, um grau de semelhança maior que 0,5 significa que o novo padrão de entrada foi classificado como classe c . Ainda na Figura 9, as portas quânticas de rotação R_y nos passos B, C e D carregam os padrões de entrada clássicos nas amplitudes do estado quântico. No entanto, este não é um procedimento direto e requer uma estratégia auxiliar para realizar tal incorporação, como foi explicado na Seção 2.2.

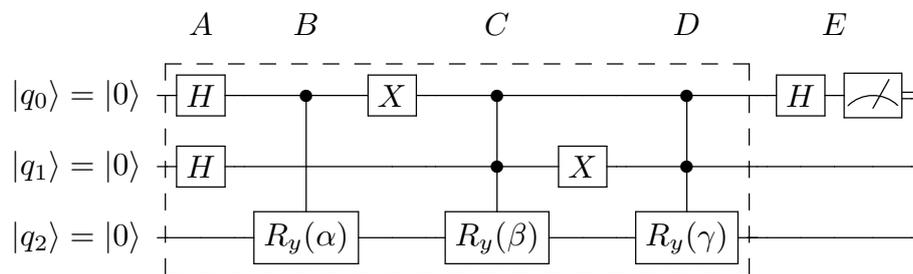


Figura 9 – Circuito quântico implementando o classificador quântico de única classe QOCC. No passo A o qubit auxiliar ($|q_0\rangle$) e o qubit de índice ($|q_1\rangle$) são postos em uma sobreposição uniforme através do operador Hadamard. O resultado de sucessivas execuções deste circuito representa uma espécie de grau de semelhança do novo padrão de entrada (passo B) em relação aos padrões já armazenados no classificador (passos C e D). As portas lógicas quânticas R_y são responsáveis por carregar dados (neste exemplo, padrões com 2 atributos) de cada vetor por meio de seus ângulos associados α , β e γ . Dentro das linhas tracejadas está o sub circuito para a preparação do estado. No final da computação (passo E), a porta lógica Hadamard no qubit auxiliar interfere com as cópias do novo padrão de entrada com os padrões carregados e, em seguida, o qubit auxiliar é medido.

O procedimento para realizar o QOCC com 2 amostras armazenadas é mostrado no Algoritmo 1.

Por fim, ao realizar os experimentos, notamos que, com apenas 1 padrão de treinamento armazenado no classificador, foi possível obter uma precisão equivalente ao caso em que são utilizados 2 padrões de treinamento. Essa equivalência de desempenho entre o uso de 1 ou 2 amostras armazenadas justifica-se pois essa única amostra pode estar próxima da representação do centroide da classe. Dessa forma, o QOCC ficou ainda mais minimalista, eliminando o qubit de índice $|q_1\rangle$ presente na Figura 9. Da mesma forma, em relação ao Algoritmo 1, para realizar a classificação com 1 amostra armazenada, não é necessário ter o qubit de índice. Assim, o QOCC, como um classificador mínimo, pode ser visto na Figura 10. Os resultados desses experimentos podem ser vistos na Seção 4.1.1.

A Equação (4.2) descreve o estado após cada etapa do circuito mostrado na Figura 10.

Algoritmo 1: Quantum One-class Classifier (QOCC)

Input: *teste, treinamento*

- 1 Inicializar registradores quânticos $auxiliar = |a\rangle$, $indice = |m\rangle$, $dados = |i\rangle$
 - 2 Realizar $H|a\rangle$ e $H|m\rangle$
 - 3 Fazer $C-R_y(teste)|a\rangle|i\rangle$ para carregar o padrão a ser classificado
 - 4 Aplicar $X|auxiliar\rangle$ para emaranhar o padrão de teste com o estado $|0\rangle$ do qubit auxiliar
 - 5 Fazer $CC-R_y(treinamento[0])|a\rangle|m\rangle|i\rangle$ para carregar o primeiro padrão de treinamento
 - 6 Aplicar $X|m\rangle$ para emaranhar o primeiro padrão de treinamento com o estado $|0\rangle$ do qubit de índice e o estado $|1\rangle$ do qubit auxiliar
 - 7 Fazer $CC-R_y(treinamento[1])|a\rangle|m\rangle|i\rangle$ para carregar o segundo padrão de treinamento
 - 8 Aplicar $H|a\rangle$ para interferir o padrão de teste com os padrões de treinamento
 - 9 Medir $|a\rangle$ para obter a probabilidade da saída ser $|0\rangle$ Retorne o grau de semelhança de acordo com o resultado da medição
-

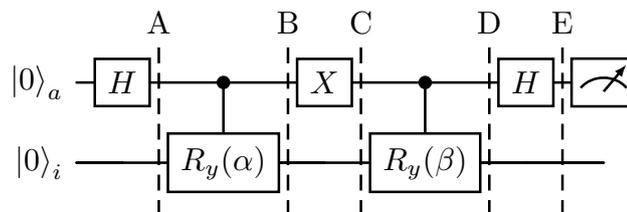


Figura 10 – Circuito quântico mínimo implementando o classificador quântico de uma classe com uma amostra de treinamento e dois atributos. O passo A coloca o qubit ancilla (índice a) em superposição uniforme. Do passo B ao passo D, são carregados os padrões de teste e treinamento. O passo E realiza a interferência entre o vetor de entrada e o vetor armazenado através da porta Hadamard, e o qubit ancilla (índice a) é medido.

$$\begin{aligned}
 A &: \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) |0\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |10\rangle) \\
 B &: \frac{1}{\sqrt{2}} (|00\rangle + |1\rangle \otimes R_y(\alpha) |0\rangle) \\
 C &: \frac{1}{\sqrt{2}} (|10\rangle + |0\rangle \otimes R_y(\alpha) |0\rangle) \\
 D &: \frac{1}{\sqrt{2}} (|1\rangle \otimes R_y(\beta) |0\rangle + |0\rangle \otimes R_y(\alpha) |0\rangle) \\
 E &: \frac{1}{2} [|0\rangle \otimes (R_y(\alpha) + R_y(\beta)) |0\rangle + |1\rangle \otimes (R_y(\alpha) - R_y(\beta)) |0\rangle]
 \end{aligned} \tag{4.2}$$

4.1.1 Experimentos e resultados

Nesta seção, mostraremos as simulações realizadas usando um simulador quântico presente no Qiskit e experimentos utilizando dispositivos quânticos NISQ reais. Assim como

em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017), utilizamos o conjunto de dados Iris (FISHER, 1936) em nossos experimentos e simulações. Além disso, realizamos experimentos (simulações) com as bases de dados *Haberman's Survival* (HABERMAN, 1976) e *Skin Segmentation* (BHATT et al., 2009). As características dos conjuntos de dados podem ser vistas na Tabela 1.

Tabela 1 – Características das bases de dados

Base de dados	Classes	Atributos	Instâncias
Iris	3	4	150
Haberman's Survival	2	3	306
Skin Segmentation	2	3	245057

Devido aos dados desbalanceados nos conjuntos de dados *Haberman's* (HABERMAN, 1976) e *Skin* (BHATT et al., 2009), utilizamos o procedimento de re-amostragem SMOTE (CHAWLA et al., 2002) para aumentar o número de amostras da classe minoritária. Para os experimentos e simulações com a base de dados *Iris*, utilizamos as duas primeiras classes da mesma. Para usar o menor número possível de qubits, alcançando o classificador quântico mínimo, escolhemos 2 atributos dos conjuntos de dados. No conjunto de dados de *Haberman's*, excluimos o atributo que armazena o ano da operação do paciente. Os atributos escolhidos no conjunto de dados *Iris* são largura da sépala e comprimento da pétala. Os atributos do conjunto de dados *Skin* são as cores RGB do pixel, onde escolhemos arbitrariamente as cores R (vermelho) e B (azul). Para estimar o desempenho com um número igual de amostras em todos os conjuntos de dados, separamos o conjunto de dados de *Haberman's* e *Skin* em 4 lotes (*batches*) com o mesmo número de instâncias do conjunto de dados *Iris*, considerando as duas primeiras classes (100 amostras).

Para poder usar a codificação de amplitude para incorporar os padrões nos registradores no circuito quântico a amostra deve ser normalizada. Portanto, padronizamos os atributos dos dados para ter média zero e variância unitária, e normalizamos os padrões. Assim como no HC, a escolha adequada de vetores de treinamento influencia fortemente a probabilidade de sucesso na classificação. Idealmente, deveríamos carregar todo o conjunto de treinamento mas, devido às limitações e erros do hardware quântico disponível, decidimos, inicialmente, manter a proposta original do HC e carregar apenas 2 vetores de treinamento com 2 atributos para o classificador - essa escolha visa manter o classificador o mais simples possível, também configurando-o como uma *prova de conceito*. Aqui, chamamos de *etapa de treinamento* o procedimento que escolhe o par de amostras que alcança melhor precisão em um conjunto

de treinamento. Esta etapa é realizada exclusivamente no simulador quântico presente no Qiskit (QISKIT, 2023) em um conjunto de treinamento que corresponde a 70% do conjunto de dados total. Para não ter que executar todas as combinações presentes no conjunto de treinamento, decidimos pegar (dentre esses 70%) 30 pares (ou amostras, no caso da abordagem com um único padrão armazenado) para realizar a etapa de treinamento. Os 30% restantes são usados como conjunto de validação. Para aplicar uma comparação estatística justa, usamos as mesmas partições (*folds*) de amostras no HC e no QOCC.

Na etapa de validação, realizamos as inferências no conjunto de validação usando o par fixo de padrões escolhido na etapa de treinamento. Essa escolha tem uma abordagem diferente dependendo do classificador quântico: se o experimento for no QOCC, esse par tem a mesma classe; se for no HC, o par é composto por classes distintas. Além da simulação usando o simulador presente no Qiskit, a classificação o QOCC e o HC são executados nos dispositivos quânticos IBM (QUANTUM, 2023) (*ibmq_athens* e *ibmq_santiago*), seguindo o padrão de 1024 execuções para cada experimento. Devido ao ruído presente nos computadores quânticos reais atuais, obtemos a precisão média ao longo de 5 execuções dos circuitos de validação resultantes do Algoritmo 1, com o par de treinamento e uma amostra do conjunto de validação como entrada. O procedimento completo do experimento está presente no Algoritmo 2.

Para validar nossa abordagem, replicamos o HC, atualizando os resultados originais, e comparamos com o QOCC (Tabela 2 e Tabela 3, respectivamente). Para fins de comparação, realizamos as mesmas tarefas de classificação usando classificadores clássicos conhecidos. Os resultados obtidos são mostrados na Tabela 4. Como já mencionado, o QOCC carrega apenas os padrões de entrada de uma única classe. Portanto, em nossos resultados, onde é mostrado QOCC C_k , k representa a classe do vetor de entrada. A precisão leva em consideração tanto a classificação correta das amostras da classe k quanto as amostras da outra classe, que devem ser identificadas pelo classificador com um baixo grau de associação.

Como mencionado, além de remover o qubit de classe, também foi possível obter resultados significativos com precisão competitiva ao executar o QOCC com apenas 1 padrão armazenado. Com isso, removemos o qubit de índice e aprimoramos o QOCC de forma a ser um modelo mínimo de classificação baseado em distância. Os resultados para esses experimentos (simulações) podem ser vistos na Tabela 5. Em relação aos Algoritmos 1 e 2, ao usar apenas 1 padrão armazenado, ao invés de carregar pares de amostras, carregamos apenas 1 amostra de treinamento.

Algoritmo 2: Experimentos para a classificação com o QOCC

Entrada: Conjunto de treinamento \mathcal{T} e conjunto de validação \mathcal{V} contendo padrões de uma classe

- 1 Definir variável de acurácia $a_{melhor} = 0$
- 2 Definir o índice das melhores amostras k_{melhor}
- 3 Definir vetor de inferências $I_{inf}^{treinamento}$
- 4 **for** $(\mathbf{x}^k, \mathbf{x}^{k+1}) \in \mathcal{T} \times \mathcal{T}$ onde $k \in \mathbb{N}$ **do**
- 5 Calcular os ângulos \mathcal{P}_{angles} dos pares de treinamento para as portas R_y
- 6 $\mathcal{X} = \mathcal{T} - \{\mathbf{x}^k, \mathbf{x}^{k+1}\}$
- 7 **for** $x \in \mathcal{X}$ **do**
- 8 Calcular os ângulos para o padrão de teste x_{angle} para as portas R_y
- 9 $y = \text{QOCC}_{simulação}(x_{ângulos}, \mathcal{P}_{ângulos})$
- 10 $I_{inf}^{treinamento} \leftarrow (x, y)$
- 11 **end**
- 12 $a_{atual} = \text{acurácia}(\mathcal{X}, I_{inf}^{treinamento})$
- 13 **if** $a_{atual} \geq a_{melhor}$ **then**
- 14 $a_{melhor} \leftarrow a_{atual}$
- 15 $k_{melhor} \leftarrow k$
- 16 **end**
- 17 **end**
- 18 Calcular os ângulos $\mathcal{P}_{ângulos}^{melhor}$ dos pares de treinamento com a melhor acurácia
- 19 **for** $x \in \mathcal{V}$ **do**
- 20 Calcular os ângulos para o padrão de teste $x_{ângulos}$ para as portas R_y
- 21 $y = \text{QOCC}_{dispositivo_real}(x_{ângulos}, \mathcal{P}_{ângulos}^{melhor})$
- 22 $I_{inf}^{val} \leftarrow (x, y)$
- 23 **end**
- 24 Retorne $a_{val} = \text{acurácia}(\mathcal{V}, I_{inf}^{val})$

Tabela 2 – Resultados do classificador Hadamard em simulações e experimentos nos dispositivos quânticos reais `ibmq_athens` and `ibmq_santiago` (marcado com um *).

Base de dados	Simul./real	HC
Iris	simul.	97.78%
Iris	real	92,44% / 92%*
Haberman's	simul.	64.72%
Haberman's	real	62.50% / 62.72%*
Skin	simul.	95.56%
Skin	real	87.56% / 87.56%*

4.1.2 Discussão

A precisão apresentada em (SCHULD; FINGERHUTH; PETRUCCIONE, 2017) é alcançada apenas se a pós-seleção for bem-sucedida (ou seja, quando o qubit auxiliar está no estado $|0\rangle$).

Tabela 3 – Resultados da execução do QOCC para 2 classes e 2 padrões armazenados em simulação e nos dispositivos quânticos reais *ibmq_athens* and *ibmq_santiago* (marcado com um *).

Base de dados	Simul./real	QOCC C ₁	QOCC C ₂
Iris	simul.	98.89%	100%
Iris	real	98.89% / 91.11%*	98.67% / 97.33%*
Haberman's	simul.	65%	64.17%
Haberman's	real	64.61% / 63.28%*	63.11% / 63.50%*
Skin	simul.	94.17%	95.56%
Skin	real	89.17% / 83.44%*	90.44% / 88.44%*

Tabela 4 – Acurácia dos classificadores clássicos SVM, DT, KNN e SGD.

Base de dados	SVM	DT	KNN	SGD
Iris	100%	100%	100%	100%
Haberman's	65%	65.83%	73.33%	58.33%
Skin	95.56%	95.83%	95.56%	86.11%

Tabela 5 – Resultados da execução do QOCC para 2 classes e 1 padrão armazenado em simulação e nos dispositivos quânticos reais *ibmq_athens* and *ibmq_santiago* (marcado com um *).

Base de dados	Simul./real	QOCC C ₁	QOCC C ₂
Iris	simul.	98.89%	98.89%
Iris	real	97.78% / 98.89%*	98.89% / 98.89%*
Haberman's	simul.	62.78%	64.17%
Haberman's	real	63.33% / 61.94%*	64.44% / 61.11%*
Skin	simul.	95.56%	96.67%
Skin	real	95% / 95.83%*	96.11% / 96.39%*

Nos experimentos realizados, a pós-seleção teve uma probabilidade de sucesso maior que 50% em apenas 47,66% das execuções no computador quântico real. Isso mostra que o HC requer mais execuções para fornecer uma resposta mais precisa.

No modelo apresentado por (SCHULD; FINGERHUTH; PETRUCCIONE, 2017), não foi mostrado como proceder na presença de mais atributos para o vetor de entrada. Isso ocorre porque o exemplo utilizado no HC foi preparado para lidar com um único qubit de dados armazenando as amplitudes correspondentes às características das amostras, não deixando explícita a generalização das operações envolvendo o qubit que armazena a classe. Para o nosso classificador, aplicamos diretamente o procedimento mostrado em (SCHULD; PETRUCCIONE, 2018) (ver Seção 2.2).

Em relação à impossibilidade de classificar vetores da classe 2, apontada em (SCHULD;

FINGERHUTH; PETRUCCIONE, 2017), esse obstáculo foi baseado na rápida decoerência do qubit de classe ao armazenar o estado $|1\rangle$ nos dispositivos utilizados nos experimentos com o HC. Em nossos experimentos, não foi observado esse impedimento na classificação de padrões da classe 2. Isso reflete os novos dispositivos quânticos disponibilizados pela IBM Quantum Experience, que são modelos aprimorados e com menos erros, em comparação com aqueles utilizados nos experimentos apresentados originalmente.

Como nosso objetivo foi construir um classificador mínimo baseado em distância, realizamos nossos experimentos (simulações) finais com apenas 1 amostra armazenada. Os resultados desses experimentos foram expostos na Tabela 5 e mostraram que a precisão permaneceu consistente com os experimentos que continham 2 amostras armazenadas (Tabela 3). Também podemos observar que, para o conjunto de dados *Skin*, o QOCC com 1 amostra armazenada teve um desempenho melhor do que o com 2 amostras armazenadas. A vantagem do QOCC sobre o HC neste ponto pode não parecer significativa, uma vez que estamos lidando com apenas 1 classe por vez. No entanto, essa redução torna-se mais perceptível em uma abordagem multiclasse com vários classificadores QOCC sendo executados em paralelo. Assim, a execução de um classificador multiclasse baseado no QOCC (ver Seção 4.2) em dispositivos NISQ se tornaria mais viável em relação ao número de qubits necessários. Outra possibilidade para flexibilizar o QOCC e torná-lo ainda mais adequado para os dispositivos quânticos ruidosos atuais é parametrizar o circuito quântico para avaliar seu desempenho na classificação, inserindo portas quânticas de rotação paramétricas com base na teoria de circuitos variacionais apresentada em (MCCLEAN et al., 2016).

Assim como no HC, os exemplos com o QOCC foram feitos apenas com a parte real dos estados quânticos. Porém, ambos os modelos são adequados para lidar com dados complexos. Uma abordagem para a classificação quântica que também leva em consideração a parte imaginária dos estados quânticos é aquela que utiliza o teste SWAP, como em (BLANK et al., 2020).

Finalmente, é necessário observar que em conjuntos de dados com apenas duas classes, o QOCC funciona como um classificador binário quântico completo. Portanto, é possível classificar todo o conjunto de dados apenas em relação às amostras de uma única classe armazenada no classificador, de modo que o alto (baixo) grau de semelhança das amostras de teste em relação aos padrões já armazenados nos dá a classificação entre as duas classes.

Disponibilidade dos dados

O site <<https://github.com/lucasponteslpa/QOCCClassifier>> contém todos os dados e o software usados para gerar os resultados apresentados na Seção 4.1.1.

4.2 QVDCC - QUANTUM VARIATIONAL DISTANCE-BASED CENTROID CLASSIFIER

O classificador aqui proposto, chamado qVDCC (*quantum Variational Distance-based Centroid Classifier*) (OLIVEIRA et al., 2024), é uma extensão do QOCC apresentado na Seção 4.1. O qVDCC é definido como um modelo de classificação multiclasse baseado em distância, obtido transformando o QOCC em um circuito quântico variacional e realizando otimização paramétrica sobre ele. No QOCC, um padrão de entrada tem sua distância euclidiana medida em relação aos padrões de treinamento por meio de interferência causada por um operador Hadamard. No entanto, visando uma maior flexibilidade, o qVDCC possui uma etapa de treinamento que busca por um centroide para cada classe da base de dados que tenta minimizar a distância destes para o padrão de teste. Além disso, o qVDCC possui uma porta lógica quântica de rotação R_y parametrizada substituindo o operador Hadamard.

Como já destacado, o objetivo principal desta tese é fornecer modelos de classificação baseados em distância adequados para os dispositivos quânticos ruidosos atuais e que sejam competitivos quando comparados a modelos de classificação clássicos. Para alcançar isso, utilizamos a teoria de circuitos quânticos variacionais na construção de um modelo de classificação híbrido clássico-quântico multiclasse, que busca encontrar o centroide de cada classe, com o objetivo de realizar uma classificação de novos dados com base na distância destes para os centroides.

Portanto, propomos um esquema de treinamento quântico-clássico no qual os gradientes analíticos do modelo podem ser estimados executando várias versões ligeiramente adaptadas do circuito variacional a ter seus parâmetros otimizados. Com isso, mostramos, por meio de simulações, que o classificador quântico qVDCC tem um desempenho satisfatório em bases de dados clássicos de referência, exigindo menos parâmetros do que outros métodos.

4.2.1 Classificador Variacional

Para estender os resultados do QOCC, precisamos modificar as portas quânticas de rotação responsáveis pelo o que seria o circuito de codificação dos padrões armazenados no circuito do QOCC visto na Figura 9, tornando-as parametrizadas. Isso foi feito ao substituir os ângulos α , β e γ (para o caso onde tem-se 2 padrões de treinamento), da Figura 9, por parâmetros livres θ_i a serem otimizados através de um algoritmo híbrido quântico-clássico.

A parametrização do nosso classificador tem como objetivos:

1. Otimizar os parâmetros do centroide virtual das classes, tornando-o representativo para a respectiva classe desejada;
2. Permitir a parametrização da porta lógica quântica de interferência entre os dados codificados nos registradores quânticos, tornando o modelo mais flexível e geral.

A ação conjunta do centroide e da interferência, ambas etapas parametrizadas, resulta na minimização da distância entre um dado de entrada a ser classificado e o padrão armazenado. Portanto, o qVDCC utiliza computação híbrida quântico-clássica para realizar uma busca por centroides, aprendendo e otimizando os parâmetros de rotação que codificam o padrão artificial e proporcionando uma melhor precisão. O importante a se observar aqui é que os centroides não são padrões reais presentes no conjunto de dados, mas sim um padrão artificial com parâmetros otimizados que visa minimizar a distância para os dados a serem classificados.

A Figura 11 mostra o ponto de partida para o esquema de classificação do qVDCC, onde *Entrada* representa uma codificação de amplitude usual de um padrão de entrada e *Modelo Quântico* é uma versão parametrizada dessa mesma codificação de amplitude, que, por sua vez, desempenha o papel do centroide.

4.2.2 Classificação Multiclasse

A Figura 12 retrata nossa abordagem proposta para classificação multiclasse, que envolve a execução de N células individuais em paralelo, onde N denota o número de classes presentes no conjunto de dados de treinamento. Nesse esquema, cada célula do qVDCC é designada a uma única classe e, coletivamente, o conjunto de todas as células é responsável por realizar a

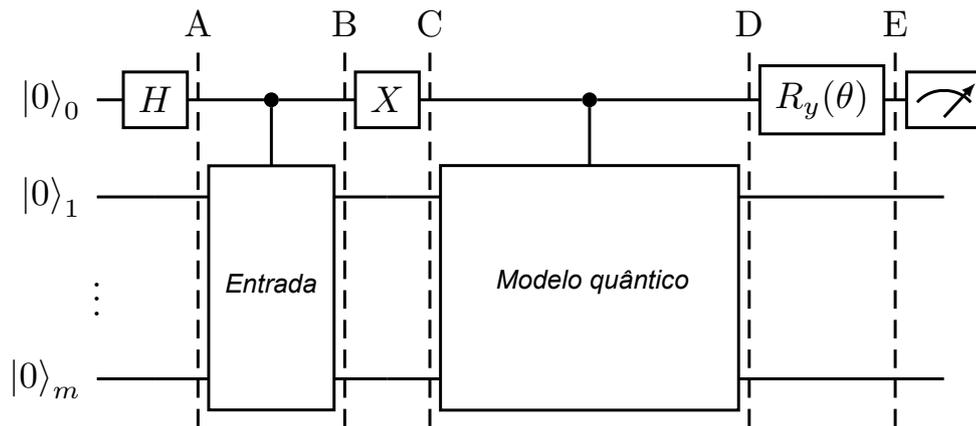


Figura 11 – Circuito quântico representando o esquema para 1 classe (célula) do qVDCC. No *passo A*, o qubit auxiliar (índice 0) é colocado em uma superposição uniforme através do operador Hadamard. O padrão de entrada a ser classificado é carregado no *passo B*. As características dos padrões são codificadas na amplitude dos qubits de dados ($|0\rangle_1, \dots, |0\rangle_m$), onde M é o número de características e $m = \lceil \log(M) \rceil$. O *passo C* inverte o qubit auxiliar e move o padrão de entrada para o estado $|0\rangle$ desse mesmo qubit auxiliar. O *modelo quântico* no *passo D* é o circuito quântico parametrizado responsável por encontrar o centroide. No *passo E*, uma porta $R_y(\theta)$ interfere o centroide com o padrão de entrada. Finalmente, ao final do circuito, temos a medição. A informação dos ângulos das portas lógicas quânticas paramétricas foi omitida por simplicidade.

tarefa de classificação. Para um conjunto de dados composto por N classes, o qVDCC requer N qubits auxiliares e N operações de medição, ou seja, uma para cada classe (célula).

Após o treinamento do modelo, o resultado da medição de cada célula indica o grau de semelhança entre o padrão de entrada e o centroide codificado nos registradores quânticos daquela célula específica. Uma probabilidade mais alta de observar o estado $|0\rangle$ implica em uma maior semelhança entre os padrões. Posteriormente, um vetor de N posições é gerado com as probabilidades de se medir o estado $|0\rangle$ em cada uma das células do classificador ao final da computação. A classificação é realizada com base na posição da célula que retorna o maior grau de semelhança. Ou seja, a posição do vetor que retorna a medição mais provável no estado $|0\rangle$ indica a classe prevista para o padrão de entrada.

Na Figura 13, é ilustrado o processo de otimização dos parâmetros do circuito quântico variacional e de classificação geral do qVDCC. Inicialmente, os dados clássicos dos padrões a serem classificados são codificados nos registradores quânticos, que servirão como entrada para o qVDCC. Na primeira iteração, o classificador é iniciado com parâmetros aleatórios, e o qubit auxiliar de cada célula é medido. Essas medições são, então, transformadas em um vetor, onde, em cada posição, é indicada a probabilidade de o estado de entrada estar associado à respectiva classe (probabilidade de observação/medição do estado $|0\rangle$). Essa classificação a partir das medições é usada para calcular a função de custo do processo de treinamento e

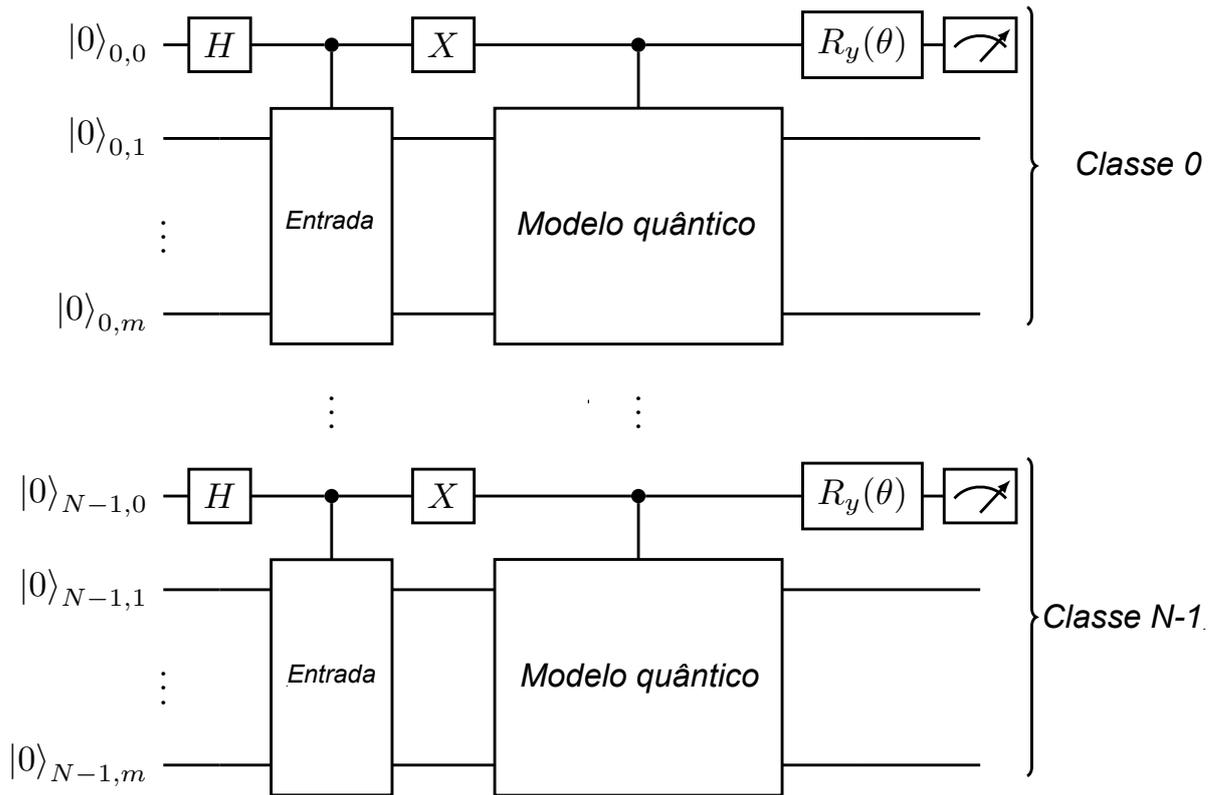


Figura 12 – O circuito completo do qVDCC para N classes (células) e padrões com M atributos. O detalhamento de cada célula pode ser visto na Figura 11. A informação dos índices dos ângulos θ foi omitida por simplicidade.

os parâmetros atuais são usados como entrada em um algoritmo clássico de otimização. Os parâmetros atualizados são retroalimentados no qVDCC, e esse processo continua até que um critério específico de parada seja alcançado.

4.2.3 Interferência Quântica Parametrizada

Conforme descrito anteriormente, a aplicação da porta lógica quântica Hadamard presente no final do circuito do QOCC, responsável pela interferência no estado quântico do classificador, foi substituída por uma porta quântica de rotação paramétrica $R_y(\theta)$. A motivação para essa parametrização reside na flexibilidade para classificar padrões de entrada que teriam um grau de semelhança muito baixo quando medidos em relação a qualquer uma das classes do qVDCC, ou poderiam ser classificados de forma inadequada se os graus de semelhança a duas (ou mais) células fossem muito próximos. Em essência, a porta de rotação age como um viés que pode ajustar as probabilidades de saída para se adequar aos dados de treinamento.

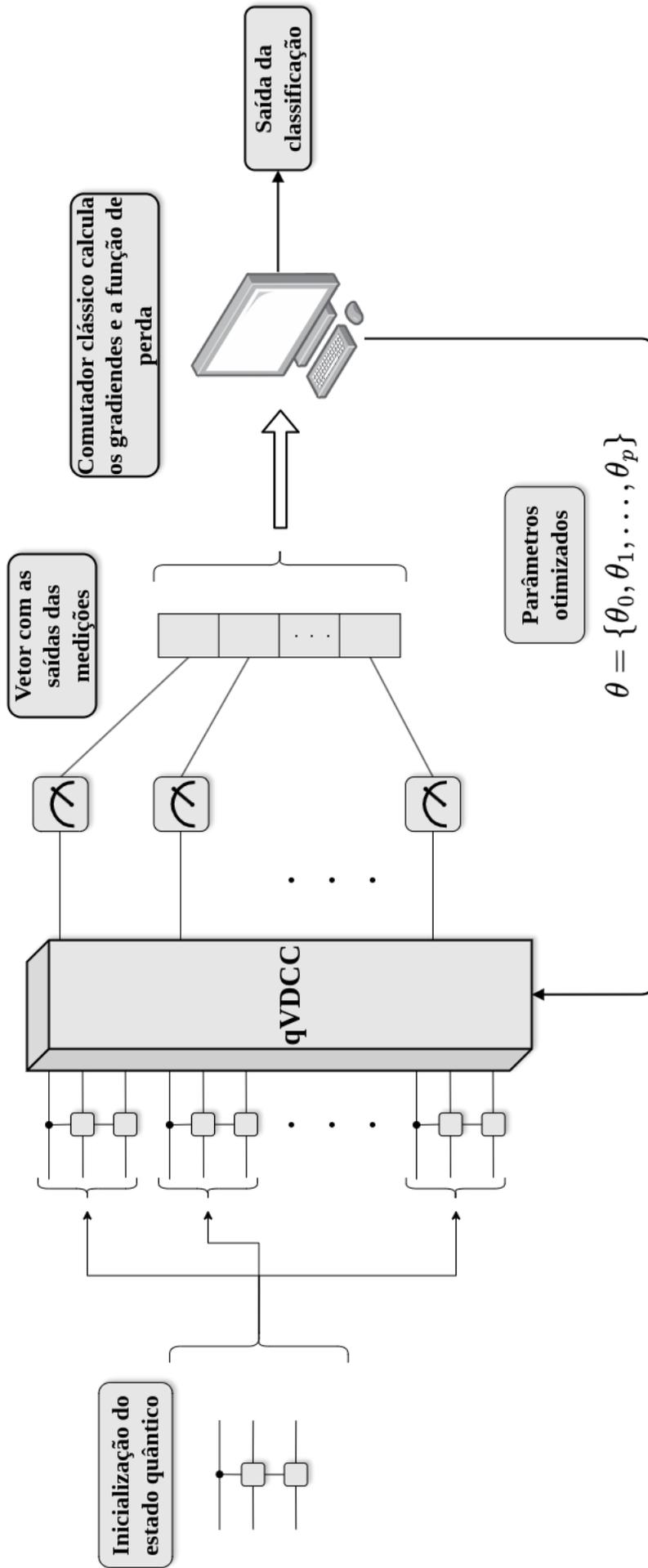


Figura 13 – Esquema geral das simulações do qVDCC com um modelo de computação híbrida quântico-clássica para classificação multiclasse.

Observe o exemplo da Figura 14. Suponha que temos 2 células no classificador qVDCC, cada uma armazenando um padrão de sua respectiva classe (0 ou 1). Nesse exemplo, a célula do qVDCC da classe 0 armazena o padrão $c_0 = \{-0.5, 0.866\}$ e a célula da classe 1 armazena o padrão $c_1 = \{0.946, -0.326\}$, com a amostra de teste sendo $t = \{-0.707, -0.707\}$. Usando a porta Hadamard para a interferência, teríamos probabilidades muito próximas (e baixas) de o padrão de teste ser classificado como classe 0 (0.362) e também de ser classificado como sendo da classe 1 (0.274), gerando uma incerteza. No entanto, com o uso da porta $R_y(\theta)$, é possível definir diferentes ângulos para a interferência. Assim, por exemplo, se a amostra de teste na Figura 14 for da classe 0, é possível definir um valor para o $R_y(\theta)$ da célula da classe 0 que aumenta a probabilidade de uma classificação correta. Por outro lado, a porta para a classe 1 é ajustado para minimizar o grau de semelhança para com o padrão de teste. Portanto, a parametrização na interferência permite uma maior flexibilidade, o que não é alcançada usando o operador Hadamard, que não é ajustável.

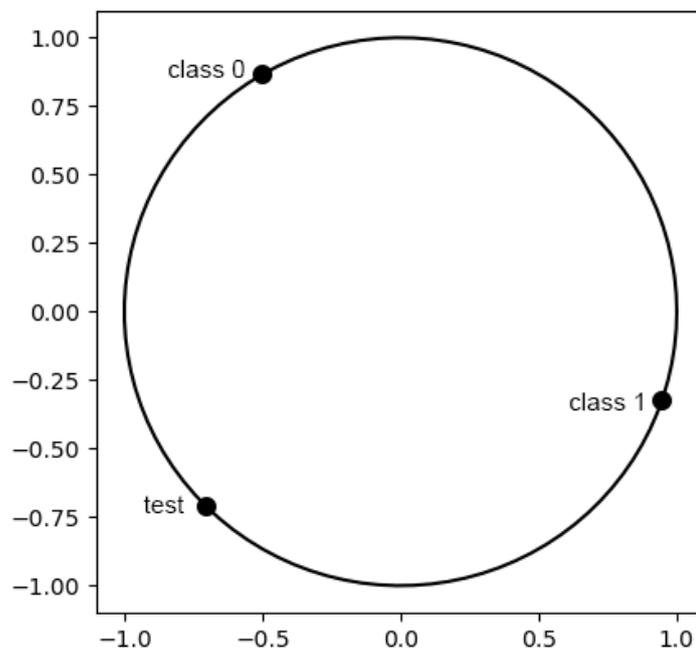


Figura 14 – Exemplo da motivação para substituir o operador Hadamard por uma porta de rotação parametrizada para realizar a interferência entre as amostras de teste e treinamento no qVDCC.

4.2.4 Simulações e Resultados

Realizamos simulações utilizando sete bases de dados: *Balance Scale*, *Banknote Authentication*, *Breast Cancer*, *Iris*, *Mammographic*, *Voting Records* e *Wine*. A Tabela 6 fornece deta-

lhes sobre as características desses conjuntos de dados. Avaliamos o desempenho do qVDCC em comparação com o algoritmo *K-vizinhos mais próximos* (K-Nearest Neighbors (KNN)) usando validação cruzada com 10 partições. As médias e as melhores acurácias alcançadas são relatadas nas Tabelas 7 (simulação do qVDCC em um ambiente sem ruído), 8 (simulação do qVDCC em um ambiente ruidoso que mimetiza uma máquina quântica real) e 9 (execução do KNN para fins de comparação). Para demonstrar a eficácia do qVDCC em dispositivos quânticos ruidosos de escala intermediária (NISQ), realizamos as simulações também em um ambiente de simulação ruidosa, presente no IBM Quantum Experience. O modelo ruidoso e o mapa de acoplamento foram importados do dispositivo *ibm_vigo*.

Tabela 6 – Características das bases de dados.

Base de dados	Classes	Atributos	Instâncias
Iris	3	4	150
Wine	3	13 (PCA=4)	178
Breast cancer	2	30 (PCA=4)	569
Balance scale	3	4	625
Mammographic	2	6 (PCA=4)	961
Banknote	2	4	1372
Voting Records	2	16 (PCA=4)	435

Tabela 7 – As médias e melhores acurácias do classificador qVDCC (simulação sem ruído) em validação cruzada com 10 partições em cada base de dados.

Base de dados	Melhor	Média	Desvio padrão
Iris	0.9333	0.8333	0.0956
Wine	0.9444	0.9042	0.0540
Breast cancer	0.9846	0.9486	0.0266
Balance scale	0.9206	0.8839	0.0256
Mammographic	0.8916	0.8215	0.0397
Banknote	0.9927	0.9707	0.0134
Voting records	1.0	0.9045	0.0643

A entrada para o qVDCC é o padrão a ser classificado, e a saída esperada é um vetor com N posições, onde N é o número de classes. Cada posição corresponde ao valor esperado de observar o estado $|0\rangle$, estimado por meio de sucessivas medições no qubit auxiliar da respectiva classe. Na etapa de treinamento/otimização dos parâmetros, para calcular a função de custo, os rótulos associados aos padrões de entrada de uma determinada classe n têm a n -ésima

Tabela 8 – As médias e melhores acurácias do classificador qVDCC (simulação ruidosa) em validação cruzada com 10 partições em cada base de dados.

Base de dados	Melhor	Média	Desvio padrão
Iris	0.9333	0.8066	0.1063
Wine	1.0	0.9268	0.0661
Breast cancer	0.9688	0.9330	0.0222
Balance scale	0.9365	0.8813	0.0334
Mammographic	0.8795	0.8154	0.0301
Banknote	0.9927	0.9751	0.0116
Voting records	1.0	0.9087	0.0662

Tabela 9 – As médias e melhores acurácias do classificador KNN em validação cruzada com 10 partições em cada base de dados.

Base de dados	Melhor	Média	Desvio padrão
Iris	1.0	0.88	0.0757
Wine	0.9444	0.9098	0.0476
Breast cancer	0.9846	0.961	0.0248
Balance scale	0.8571	0.822	0.0261
Mammographic	0.8554	0.8032	0.0510
Banknote	1.0	0.9927	0.0069
Voting records	1.0	0.9007	0.0703

posição marcada com 1, o que corresponde à medição do estado $|0\rangle$ — as outras posições são marcadas com -1, correspondendo à medição do estado $|1\rangle$.

Na Figura 15, utilizamos o operador Hadamard ao final do circuito do qVDCC no lugar da porta R_y (passo E na Figura 11). Conforme explicado na Seção 4.2.3, substituir a porta Hadamard por uma versão parametrizada aumenta a precisão do modelo em alguns casos. No entanto, para observar os centroides das 3 classes da base de dados Iris com apenas dois atributos, a aplicação da porta Hadamard é suficiente para fins de visualização.

Para treinar os parâmetros do qVDCC, utilizamos o otimizador clássico Adam (KINGMA; BA, 2017). Os parâmetros resultantes já treinados foram carregados nos circuitos quânticos de teste para avaliação de desempenho. Para nossas simulações, empregamos a implementação padrão do algoritmo KNN da biblioteca Scikit-learn em Python (PEDREGOSA et al., 2011). Durante o treinamento do classificador variacional, atualizamos os parâmetros para minimizar a perda de *erro quadrático médio* (*Mean Squared Error (MSE)*). Nossas simulações para o qVDCC foram conduzidas de acordo com o Algoritmo 3.

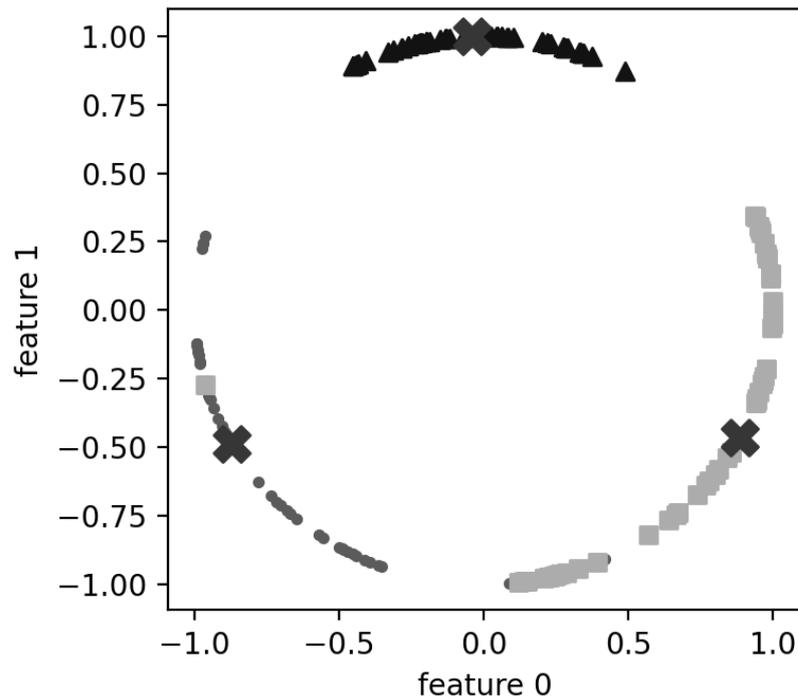


Figura 15 – Triângulos, pontos e quadrados representam as 3 classes da base de dados *Iris*. Os marcadores 'x' preenchidos representam os centroides virtuais encontrados para as classes, obtidos através da otimização dos parâmetros presentes no circuito quântico variacional do qVDCC. As 2 características mostradas no gráfico de dispersão foram obtidas executando o algoritmo t-SNE (MAATEN; HINTON, 2008) nas 4 características originais do conjunto de dados.

Algoritmo 3: Simulações do qVDCC

- 1 Carregar a base de dados \mathcal{D} (classe = $\{0, \dots, c\}$, atributos = $\{0, \dots, f\}$)
 - 2 Padronizar e normalizar \mathcal{D}
 - 3 Dividir \mathcal{D} em subconjuntos de treinamento e teste
 - 4 Treinar o modelo
 - 5 Obter os parâmetros ajustados
 - 6 **for** $x = 1$ até o tamanho do conjunto de teste **do**
 - 7 Definir x como entrada para cada célula do qVDCC
 - 8 Executar qVDCC com os parâmetros otimizados
 - 9 Medir o valor esperado para cada **qubit auxiliar**
 - 10 Representar o conjunto de valores esperados como um vetor de N posições
 - 11 **end**
 - 12 Retornar o número de amostras que foram associadas corretamente à sua classe/célula
-

Disponibilidade dos dados

O site <<https://github.com/nicolas-melo/qVDCC>> contém todos os dados e o software usados para gerar os resultados apresentados na Seção 4.2.4.

4.2.5 Discussão

Uma possibilidade não explorada nas simulações foi a busca puramente clássica pelos centroides da classe para, então, carregar esses dados como estados quânticos. Isso não foi feito devido ao objetivo da tese de empregar a computação híbrida com os algoritmos de otimização visando nas características da computação quântica. Cabe pontuar que usar centroides otimizados classicamente pode levar a uma inicialização mais estável para o classificador quântico. Se os centroides forem suficientemente representativos, o classificador pode funcionar melhor em termos de desempenho e robustez em relação ao ruído e outras imperfeições típicas de estados quânticos.

Ainda, uma possível modificação seria realizar o treinamento totalmente quântico. Essa abordagem puramente quântica poderia explorar as propriedades quânticas, como superposição e emaranhamento, que podem levar a melhorias significativas de desempenho em cenários específicos, especialmente quando se trata de espaços de parâmetros complexos ou de operações de classificação em dados de alta dimensão. No entanto, isso depende muito da eficácia do algoritmo quântico implementado. Contudo, a otimização clássica dos centroides pode ser mais eficiente em termos de tempo, atualmente, especialmente em conjuntos de dados grandes, onde os algoritmos clássicos podem convergir rapidamente. Com isso, a otimização quântica, embora potencialmente mais poderosa em alguns casos, ainda está em desenvolvimento e pode não ser a solução mais rápida para o problema específico de encontrar centroides, além de exigir mais pesquisa e desenvolvimento. Posto isso, a abordagem híbrida se configurou como uma possibilidade mais direta para a implementação do algoritmo proposto, já que muitas técnicas clássicas de otimização estão bem estabelecidas e disponíveis em bibliotecas de aprendizagem de máquina.

Outro aspecto a ser mencionado, tanto com relação ao QOCC quanto ao qVDCC, diz respeito ao uso de dados reais para as simulações. Essa escolha se deu em virtude do uso trivial de bases de dados de treinamento clássicas para validação de algumas abordagens de aprendizagem de máquina em computação quântica. Porém, é importante frisar que utilizar amplitudes complexas em vez de reais na codificação de dados em algoritmos quânticos traz várias vantagens importantes:

- As amplitudes complexas permitem a interferência construtiva e destrutiva, o que é fundamental em muitos algoritmos quânticos. Isso permite que estados quânticos que

têm fases diferentes interajam de maneiras que não seriam possíveis se estivéssemos usando apenas amplitudes reais;

- As amplitudes complexas oferecem uma representação mais rica dos estados quânticos, pois cada amplitude é definida por uma magnitude e uma fase. Isso permite explorar um espaço de soluções mais amplo, potencialmente levando a resultados mais eficientes;
- Muitas técnicas de otimização em computação quântica, como o algoritmo de Grover e o algoritmo de Shor, dependem da manipulação de fases para maximizar a probabilidade de medir os resultados desejados. Usar amplitudes complexas é essencial para essas manipulações;
- Sistemas quânticos que utilizam amplitudes complexas podem ser mais robustos a ruídos e erros, já que a informação está codificada tanto na magnitude quanto na fase das amplitudes;
- A própria formalização da mecânica quântica é baseada em números complexos, com a função de onda sendo uma função complexa. Utilizar amplitudes complexas na computação quântica alinha-se com os fundamentos teóricos da física quântica.

5 CONCLUSÃO

Nesta tese, inicialmente, foram investigados modelos de aprendizagem de máquina quântica. Especificamente, aqueles voltados para classificação baseada em distância. A intenção por trás disso foi identificar lacunas e possibilidades existentes a fim de propor novos modelos que fossem compatíveis com os dispositivos quânticos ruidosos atuais e que tivessem desempenho competitivo com os análogos clássicos. Tais lacunas foram referentes a: poucos modelos realmente capazes de serem executados em dispositivos NISQ, mesmo para bases de dados pequenas com poucos atributos, escassez de abordagens que focassem em classificação multiclasse e comparações insuficientes com modelos clássicos.

Assim, foi proposto o *Quantum One-Class Classifier* (QOCC), um classificador quântico mínimo de única classe baseado na investigação do comportamento do classificador Hadamard (HC) e no conceito de memória quântica probabilística. Em comparação com o HC, o QOCC demonstrou uma precisão equivalente ou aprimorada em dispositivos quânticos reais, além de utilizar menos recursos quânticos. Além disso, o QOCC se mostrou competitivo quando comparado a classificadores clássicos. Com isso, demos os primeiros passos em direção ao que poderia se tornar uma generalização de um classificador baseado em distância.

O QOCC apresenta uma vantagem em relação ao HC no que diz respeito ao seu tamanho e à redução da medição quântica ao final da computação - saindo de uma medição de 2 qubits para uma medição única de 1 qubit. Em conjuntos de dados com duas classes, o QOCC se comporta como um classificador binário quântico completo, indicando, no qubit auxiliar, o grau de semelhança entre o padrão de entrada e a classe da qual pertencem os dados armazenados no classificador. Dessa forma, o QOCC necessita, assim, carregar apenas dados de uma das classes. Também foi fornecida uma atualização do desempenho do HC em dispositivos quânticos atuais e tornou-se explícita a metodologia usada para a codificação de dados e para os experimentos.

Possíveis melhorias identificadas neste ponto incluíam investigar como usar o QOCC para classificar conjuntos de dados multiclasse e explorar como inserir parâmetros ajustáveis no classificador para aprimorar sua precisão. Portanto, nesta tese, também foi apresentado o qVDCC (*quantum Variational Distance-based Centroid Classifier*), um algoritmo híbrido clássico-quântico que combina técnicas de aprendizagem de máquina clássicas e quânticas para classificação multiclasse.

Embora os dispositivos NISQ prometam avanços na computação quântica, seu número limitado de qubits e suscetibilidade a erros quando na presença da aplicação de muitos operadores quânticos são obstáculos significativos para problemas do mundo real com conjuntos de dados grandes. Por isto, torna-se desafiador utilizar esses dispositivos para lidar com um grande número de amostras de treinamento no classificador sem incorrer em erros significativos. O qVDCC utiliza o classificador quântico de uma classe QOCC e a teoria dos circuitos quânticos variacionais para superar as limitações do hardware quântico atual, tornando-o adequado para aplicações de aprendizagem de máquina quântica em curto prazo. O qVDCC é um classificador multiclasse simples e viável para dispositivos NISQ que requer apenas a busca por um centroide por classe, evitando a necessidade de carregar todos os dados de treinamento em um estado quântico. Essa abordagem permite o treinamento híbrido e a execução do qVDCC em dispositivos quânticos de escala intermediária sem exigir recursos computacionais excessivos.

Assim, ao empregar um paradigma híbrido, o qVDCC forneceu resultados imediatos em hardware quântico real, possibilitando uma busca mais direta por vantagens quânticas sobre dispositivos clássicos. Nossos resultados de simulação demonstraram a eficácia do qVDCC em tarefas de classificação, mesmo em ambientes quânticos ruidosos, o que o estabelece como uma base para o desenvolvimento de classificadores quânticos multiclasse baseados em distância mais complexos e eficientes. Os resultados em simuladores de ambientes quânticos ruidosos foram consistentes com aqueles obtidos em simulações sem erros, demonstrando que o qVDCC é um classificador competitivo, apesar das limitações dos dispositivos NISQ atuais. Também comparamos o desempenho do qVDCC com o KNN, e nossos resultados mostram que o qVDCC foi competitivo em termos de precisão de classificação. Notavelmente, o qVDCC oferece compressão exponencial de dados em comparação com o bem-estabelecido algoritmo clássico KNN, tornando-o uma opção promissora para aprendizado de máquina quântica.

Ao aplicar os classificadores propostos aqui a conjuntos de dados reais, pudemos observar, como evidenciado pelos resultados apresentados, que tanto o QOCC quanto o qVDCC alcançaram os objetivos pretendidos por esta tese: fornecer modelos de classificação quântica simples, adequados para os dispositivos quânticos ruidosos atuais e competitivos (em termos de acurácia e de compressão de dados) para com os análogos clássicos. Essas abordagens apresentadas têm a vantagem adicional de não exigir nenhuma sub-rotina complementar, como algoritmos de busca ou amplificação de amplitude. Tais rotinas não são adequadas para os dispositivos quânticos ruidosos atuais.

Trabalhos futuros podem focar em pesquisas adicionais para otimizar o circuito visando

uma execução eficiente em dispositivos quânticos reais, explorar medidas de distância além da Euclidiana (Manhattan, Hamming, Mahalanobis, entre outras), realizar simulações com bases de dados quânticas e estender o qVDCC para suportar, por exemplo, aprendizado por agrupamento (*ensemble learning*). No caso do aprendizado por agrupamento, uma possibilidade seria agrupar diversos classificadores, cada um deles utilizando uma medida de distância específica, combinando suas previsões para uma decisão final. Outro trabalho futuro com potencial de melhorar nossa abordagem consiste em avaliar o impacto do ruído na busca por centroides na medida em que a profundidade dos circuitos quânticos aumenta, ou seja, na presença de mais padrões de treinamento no classificador. Ainda, experimentos adicionais podem ser feitos para observar nos resultados de classificação o impacto de utilizar centroides otimizados em simulações ideais (sem ruído) em simulações ruidosas.

REFERÊNCIAS

- ABADI, M.; AGARWAL, A.; BARHAM, P.; BREVDO, E.; CHEN, Z.; CITRO, C.; CORRADO, G. S.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; GOODFELLOW, I.; HARP, A.; IRVING, G.; ISARD, M.; JIA, Y.; JOZEFOWICZ, R.; KAISER, L.; KUDLUR, M.; LEVENBERG, J.; MANÉ, D.; MONGA, R.; MOORE, S.; MURRAY, D.; OLAH, C.; SCHUSTER, M.; SHLENS, J.; STEINER, B.; SUTSKEVER, I.; TALWAR, K.; TUCKER, P.; VANHOUCHE, V.; VASUDEVAN, V.; VIÉGAS, F.; VINYALS, O.; WARDEN, P.; WATTENBERG, M.; WICKE, M.; YU, Y.; ZHENG, X. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.
- AÏMEUR, E.; BRASSARD, G.; GAMBS, S. Quantum clustering algorithms. In: *Proceedings of the 24th international conference on machine learning*. [S.l.: s.n.], 2007. p. 1–8.
- ARAUJO, I. F.; BLANK, C.; ARAÚJO, I. C.; SILVA, A. J. da. Low-rank quantum state preparation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, IEEE, 2023.
- ARAUJO, I. F.; PARK, D. K.; LUDERMIR, T. B.; OLIVEIRA, W. R.; PETRUCCIONE, F.; SILVA, A. J. da. Configurable sublinear circuits for quantum state preparation. *arXiv preprint arXiv:2108.10182*, 2021.
- ARAUJO, I. F.; PARK, D. K.; PETRUCCIONE, F.; SILVA, A. J. da. A divide-and-conquer algorithm for quantum state preparation. *Scientific Reports*, Nature Publishing Group, v. 11, n. 1, p. 1–12, 2021.
- ARUTE, F.; ARYA, K.; BABBUSH, R.; BACON, D.; BARDIN, J. C.; BARENDS, R.; BISWAS, R.; BOIXO, S.; BRANDAO, F. G.; BUELL, D. A. et al. Quantum supremacy using a programmable superconducting processor. *Nature*, Nature Publishing Group, v. 574, n. 7779, p. 505–510, 2019.
- BASHEER, A.; GOYAL, S. K. et al. Quantum k-nearest neighbor machine learning algorithm. *arXiv preprint arXiv:2003.09187*, 2020.
- BENEDETTI, M.; LLOYD, E.; SACK, S.; FIORENTINI, M. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, IOP Publishing, v. 4, n. 4, p. 043001, 2019.
- BERGHOLM, V.; IZAAC, J.; SCHULD, M.; GOGOLIN, C.; ALAM, M. S.; AHMED, S.; ARRAZOLA, J. M.; BLANK, C.; DELGADO, A.; JAHANGIRI, S. et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.
- BERGHOLM, V.; VARTIAINEN, J. J.; MÖTTÖNEN, M.; SALOMAA, M. M. Quantum circuits with uniformly controlled one-qubit gates. *Physical Review A—Atomic, Molecular, and Optical Physics*, APS, v. 71, n. 5, p. 052330, 2005.
- BHATT, R. B.; SHARMA, G.; DHALL, A.; CHAUDHURY, S. Efficient skin region segmentation using low complexity fuzzy decision tree model. In: IEEE. *2009 Annual IEEE India Conference*. [S.l.], 2009. p. 1–4.

- BIAMONTE, J.; WITTEK, P.; PANCOTTI, N.; REBENTROST, P.; WIEBE, N.; LLOYD, S. Quantum machine learning. *Nature*, Nature Publishing Group, v. 549, n. 7671, p. 195–202, 2017.
- BISHOP, C. M. Pattern recognition and machine learning. *Springer google schola*, v. 2, p. 1122–1128, 2006.
- BLANK, C.; PARK, D. K.; RHEE, J.-K. K.; PETRUCCIONE, F. Quantum classifier with tailored quantum kernel. *npj Quantum Information*, Nature Publishing Group, v. 6, n. 1, p. 1–7, 2020.
- BLANK, C.; SILVA, A. J. da; ALBUQUERQUE, L. P. de; PETRUCCIONE, F.; PARK, D. K. Compact quantum kernel-based binary classifier. *Quantum Science and Technology*, IOP Publishing, v. 7, n. 4, p. 045007, jul 2022. Disponível em: <<https://dx.doi.org/10.1088/2058-9565/ac7ba3>>.
- CEREZO, M.; ARRASMITH, A.; BABBUSH, R.; BENJAMIN, S. C.; ENDO, S.; FUJII, K.; McClean, J. R.; MITARAI, K.; YUAN, X.; CINCIO, L.; COLES, P. J. Variational quantum algorithms. *Nature Reviews Physics*, v. 3, n. 9, p. 625–644, 2021. ISSN 2522-5820. Disponível em: <<https://doi.org/10.1038/s42254-021-00348-9>>.
- CEREZO, M.; VERDON, G.; HUANG, H.-Y.; CINCIO, L.; COLES, P. J. Challenges and opportunities in quantum machine learning. *Nature Computational Science*, Nature Publishing Group US New York, v. 2, n. 9, p. 567–576, 2022.
- CHALUMURI, A.; KUNE, R.; MANOJ, B. A hybrid classical-quantum approach for multi-class classification. *Quantum Information Processing*, Springer, v. 20, n. 3, p. 119, 2021.
- CHAWLA, N. V.; BOWYER, K. W.; HALL, L. O.; KEGELMEYER, W. P. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, v. 16, p. 321–357, 2002.
- CILIBERTO, C.; HERBSTER, M.; IALONGO, A. D.; PONTIL, M.; ROCCHETTO, A.; SEVERINI, S.; WOSSNIG, L. Quantum machine learning: a classical perspective. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, The Royal Society Publishing, v. 474, n. 2209, p. 20170551, 2018.
- CONG, I.; CHOI, S.; LUKIN, M. D. Quantum convolutional neural networks. *Nature Physics*, Nature Publishing Group UK London, v. 15, n. 12, p. 1273–1278, 2019.
- CROOKS, G. E. Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition. *arXiv preprint arXiv:1905.13311*, 2019.
- DANG, Y.; JIANG, N.; HU, H.; JI, Z.; ZHANG, W. Image classification based on quantum k-nearest-neighbor algorithm. *Quantum Information Processing*, Springer, v. 17, p. 1–18, 2018.
- DIVINCENZO, D. P. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, Wiley Online Library, v. 48, n. 9-11, p. 771–783, 2000.
- DUNJKO, V.; BRIEGEL, H. J. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Reports on Progress in Physics*, IOP Publishing, v. 81, n. 7, p. 074001, 2018.

- DUNJKO, V.; TAYLOR, J. M.; BRIEGEL, H. J. Quantum-enhanced machine learning. *Physical review letters*, APS, v. 117, n. 13, p. 130501, 2016.
- EZHOF, A.; NIFANOVA, A.; VENTURA, D. Quantum associative memory with distributed queries. *Information Sciences*, Elsevier, v. 128, n. 3-4, p. 271–293, 2000.
- FARHI, E.; NEVEN, H. Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*, 2018.
- FEDOROV, A. K.; GISIN, N.; BELOUSSOV, S. M.; LVOVSKY, A. I. Quantum computing at the quantum advantage threshold: a down-to-business review. *arXiv preprint arXiv:2203.17181*, 2022.
- FENG, C.; ZHAO, B.; ZHOU, X.; DING, X.; SHAN, Z. An enhanced quantum k-nearest neighbor classification algorithm based on polar distance. *Entropy*, MDPI, v. 25, n. 1, p. 127, 2023.
- FEYNMAN, R. P. Simulating physics with computers. In: *Feynman and computation*. [S.l.]: CRC Press, 2018. p. 133–153.
- FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, Wiley Online Library, v. 7, n. 2, p. 179–188, 1936.
- GAO, L.-Z.; LU, C.-Y.; GUO, G.-D.; ZHANG, X.; LIN, S. Quantum k-nearest neighbors classification algorithm based on mahalanobis distance. *Frontiers in Physics*, Frontiers, v. 10, p. 1047466, 2022.
- GIOVANNETTI, V.; LLOYD, S.; MACCONE, L. Quantum random access memory. *Physical review letters*, APS, v. 100, n. 16, p. 160501, 2008.
- GROVER, L. K. A fast quantum mechanical algorithm for database search. In: *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*. New York, NY, USA: ACM, 1996. (STOC '96), p. 212–219. ISBN 0-89791-785-5. Disponível em: <<http://doi.acm.org/10.1145/237814.237866>>.
- GROVER, L. K. Quantum mechanics helps in searching for a needle in a haystack. *Physical review letters*, APS, v. 79, n. 2, p. 325, 1997.
- HABERMAN, S. J. Generalized residuals for log-linear models. In: *Proceedings of the 9th international biometrics conference*. [S.l.: s.n.], 1976. p. 104–122.
- HARROW, A. W.; HASSIDIM, A.; LLOYD, S. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, American Physical Society, v. 103, p. 150502, Oct 2009. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.103.150502>>.
- HAVLÍČEK, V.; CÓRCOLES, A. D.; TEMME, K.; HARROW, A. W.; KANDALA, A.; CHOW, J. M.; GAMBETTA, J. M. Supervised learning with quantum-enhanced feature spaces. *Nature*, Nature Publishing Group, v. 567, n. 7747, p. 209–212, 2019.
- HUANG, H.-Y.; BROUGHTON, M.; MOHSENI, M.; BABBUSH, R.; BOIXO, S.; NEVEN, H.; MCCLEAN, J. R. Power of data in quantum machine learning. *Nature communications*, Nature Publishing Group, v. 12, n. 1, p. 1–9, 2021.

- HUR, T.; KIM, L.; PARK, D. K. Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence*, Springer, v. 4, n. 1, p. 3, 2022.
- KATHURIA, K.; RATAN, A.; MCCONNELL, M.; BEKIRANOV, S. Implementation of a hamming distance-like genomic quantum classifier using inner products on ibmqx2 and ibmq_16_melbourne. *Quantum Machine Intelligence*, Springer, v. 2, n. 1, p. 1–26, 2020.
- KIM, C.; PARK, K. D.; RHEE, J.-K. Quantum error mitigation with artificial neural network. *IEEE Access*, IEEE, v. 8, p. 188853–188860, 2020.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2017.
- KOYA, S.; LASKAR, M. R.; DUTTA, A. K. A proposed quantum classification algorithm for symbol detection with noisy observation. In: IEEE. *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*. [S.l.], 2023. p. 1–6.
- KRAFT, D. A software package for sequential quadratic programming. *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt*, 1988.
- LAROSE, R. Overview and comparison of gate level quantum software platforms. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 3, p. 130, 2019.
- LAROSE, R.; COYLE, B. Robust data encodings for quantum classifiers. *Phys. Rev. A*, American Physical Society, v. 102, p. 032420, Sep 2020. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.102.032420>>.
- LI, J.; LIN, S.; YU, K.; GUO, G. Quantum k-nearest neighbor classification algorithm based on hamming distance. *Quantum Information Processing*, Springer, v. 21, n. 1, p. 18, 2022.
- MAATEN, L. Van der; HINTON, G. Visualizing data using t-sne. *Journal of machine learning research*, v. 9, n. 11, 2008.
- MCCLEAN, J. R.; ROMERO, J.; BABBUSH, R.; ASPURU-GUZIK, A. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, IOP Publishing, v. 18, n. 2, p. 023023, 2016.
- MITARAI, K.; NEGORO, M.; KITAGAWA, M.; FUJII, K. Quantum circuit learning. *Physical Review A*, APS, v. 98, n. 3, p. 032309, 2018.
- MOHSENI, M.; READ, P.; NEVEN, H.; BOIXO, S.; DENCHEV, V.; BABBUSH, R.; FOWLER, A.; SMELYANSKIY, V.; MARTINIS, J. Commercialize quantum technologies in five years. *Nature News*, v. 543, n. 7644, p. 171, 2017.
- MOORE, G. E. Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp. 114 ff. *IEEE solid-state circuits society newsletter*, IEEE, v. 11, n. 3, p. 33–35, 2006.
- MOTTONEN, M.; VARTIAINEN, J. J.; BERGHOLM, V.; SALOMAA, M. M. Transformation of quantum states using uniformly controlled rotations. *arXiv preprint quant-ph/0407010*, 2004.

- NEUMANN, N. M. Classification using a two-qubit quantum chip. *arXiv preprint arXiv:2004.10426*, 2020.
- NIELSEN, M. A.; CHUANG, I. L. *Quantum Computation and Quantum Information*. [S.l.]: Cambridge University Press, 2010.
- OLIVEIRA, N. M. D.; ALBUQUERQUE, L. P. D.; OLIVEIRA, W. R. D.; LUDERMIR, T. B.; SILVA, A. J. D. Quantum one-class classification with a distance-based classifier. In: *2021 International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2021. p. 1–7.
- OLIVEIRA, N. M. de; PARK, D. K.; ARAUJO, I. F.; SILVA, A. J. da. Quantum variational distance-based centroid classifier. *Neurocomputing*, Elsevier, p. 127356, 2024.
- OLIVEIRA, W. R. de; SILVA, A. J.; LUDERMIR, T. B.; LEONEL, A.; GALINDO, W. R.; PEREIRA, J. C. Quantum logical neural networks. In: IEEE. *2008 10th Brazilian Symposium on Neural Networks*. [S.l.], 2008. p. 147–152.
- PARK, D. K.; BLANK, C.; PETRUCCIONE, F. The theory of the quantum kernel-based binary classifier. *Physics Letters A*, Elsevier, p. 126422, 2020.
- PASZKE, A.; GROSS, S.; CHINTALA, S.; CHANAN, G.; YANG, E.; DEVITO, Z.; LIN, Z.; DESMAISON, A.; ANTIGA, L.; LERER, A. Automatic differentiation in pytorch. 2017.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011. Disponível em: <<http://jmlr.org/papers/v12/pedregosa11a.html>>.
- PÉREZ-SALINAS, A.; CERVERA-LIERTA, A.; GIL-FUSTER, E.; LATORRE, J. I. Data re-uploading for a universal quantum classifier. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 4, p. 226, fev. 2020. ISSN 2521-327X. Disponível em: <<https://doi.org/10.22331/q-2020-02-06-226>>.
- PERUZZO, A.; MCCLEAN, J.; SHADBOLT, P.; YUNG, M.-H.; ZHOU, X.-Q.; LOVE, P. J.; ASPURU-GUZZIK, A.; O'BRIEN, J. L. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, Nature Publishing Group UK London, v. 5, n. 1, p. 4213, 2014.
- PIATRENKA, I.; RUSEK, M. Quantum variational multi-class classifier for the iris data set. In: SPRINGER. *International Conference on Computational Science*. [S.l.], 2022. p. 247–260.
- POWELL, M. J. *A direct search optimization method that models the objective and constraint functions by linear interpolation*. [S.l.]: Springer, 1994.
- PRESKILL, J. Quantum Computing in the NISQ era and beyond. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 2, p. 79, ago. 2018. ISSN 2521-327X.
- QISKIT. 2023. <https://qiskit.org/>. Last accessed: 2019-11-14.
- QUANTUM, I. 2023. <https://quantum.ibm.com/>. Last accessed: 2023-10-14.

- REBENTROST, P.; MOHSENI, M.; LLOYD, S. Quantum support vector machine for big data classification. *Physical review letters*, APS, v. 113, n. 13, p. 130503, 2014.
- RUAN, Y.; XUE, X.; LIU, H.; TAN, J.; LI, X. Quantum algorithm for k-nearest neighbors classification based on the metric of hamming distance. *International Journal of Theoretical Physics*, Springer, v. 56, n. 11, p. 3496–3507, 2017.
- SALAM, M. A.; AZAR, A. T.; ELGENDY, M. S.; FOUAD, K. M. The effect of different dimensionality reduction techniques on machine learning overfitting problem. *Int. J. Adv. Comput. Sci. Appl*, v. 12, n. 4, p. 641–655, 2021.
- SCHULD, M.; BERGHOLM, V.; GOGOLIN, C.; IZAAC, J.; KILLORAN, N. Evaluating analytic gradients on quantum hardware. *Physical Review A*, APS, v. 99, n. 3, p. 032331, 2019.
- SCHULD, M.; FINGERHUTH, M.; PETRUCCIONE, F. Implementing a distance-based classifier with a quantum interference circuit. *EPL (Europhysics Letters)*, IOP Publishing, v. 119, n. 6, p. 60002, 2017.
- SCHULD, M.; PETRUCCIONE, F. *Supervised learning with quantum computers*. [S.l.]: Springer, 2018. v. 17.
- SCHULD, M.; PETRUCCIONE, F. Representing data on a quantum computer. In: _____. *Machine Learning with Quantum Computers*. Cham: Springer International Publishing, 2021. p. 147–176. ISBN 978-3-030-83098-4. Disponível em: <https://doi.org/10.1007/978-3-030-83098-4_4>.
- SCHULD, M.; SINAYSKIY, I.; PETRUCCIONE, F. The quest for a quantum neural network. *Quantum Information Processing*, Springer, v. 13, n. 11, p. 2567–2586, 2014.
- SCHULD, M.; SINAYSKIY, I.; PETRUCCIONE, F. An introduction to quantum machine learning. *Contemporary Physics*, Taylor & Francis, v. 56, n. 2, p. 172–185, 2015.
- SCHULD, M.; SINAYSKIY, I.; PETRUCCIONE, F. Prediction by linear regression on a quantum computer. *Physical Review A*, APS, v. 94, n. 2, p. 022342, 2016.
- SHAO, C. Data classification by quantum radial-basis-function networks. *Physical Review A*, APS, v. 102, n. 4, p. 042418, 2020.
- SHENDE, V. V.; BULLOCK, S. S.; MARKOV, I. L. Synthesis of quantum-logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, IEEE, v. 25, n. 6, p. 1000–1010, 2006.
- SHOR, P. W. Algorithms for quantum computation: Discrete logarithms and factoring. In: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*. Washington, DC, USA: IEEE Computer Society, 1994. (SFCS '94), p. 124–134. ISBN 0-8186-6580-7. Disponível em: <<http://dx.doi.org/10.1109/SFCS.1994.365700>>.
- SILVA, A. J. da; LUDERMIR, T. B.; OLIVEIRA, W. R. de. Quantum perceptron over a field and neural network architecture selection in a quantum computer. *Neural Networks*, Elsevier, v. 76, p. 55–64, 2016.

-
- SILVER, D.; PATEL, T.; TIWARI, D. Quilt: Effective multi-class classification on quantum computers using an ensemble of diverse quantum classifiers. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2022. v. 36, n. 8, p. 8324–8332.
- SPALL, J. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, v. 37, n. 3, p. 332–341, 1992.
- TRUGENBERGER, C. A. Probabilistic quantum memories. *Physical Review Letters*, APS, v. 87, n. 6, p. 067901, 2001.
- TYCHOLA, K. A.; KALAMPOKAS, T.; PAPAKOSTAS, G. A. Quantum machine learning—an overview. *Electronics*, MDPI, v. 12, n. 11, p. 2379, 2023.
- VERAS, T. M. de; SILVA, L. D. da; SILVA, A. J. da. Double sparse quantum state preparation. *Quantum Information Processing*, Springer, v. 21, n. 6, p. 204, 2022.
- Veras, T. M. L.; De Araujo, I. C. S.; Park, K. D.; da Silva, A. J. Circuit-based quantum memory for classical data with continuous amplitudes. *IEEE Transactions on Computers*, p. 1–1, 2020.
- WILLSCH, D.; WILLSCH, M.; RAEDT, H. D.; MICHIELSEN, K. Support vector machines on the d-wave quantum annealer. *Computer physics communications*, Elsevier, v. 248, p. 107006, 2020.
- WITTEK, P. *Quantum machine learning: what quantum computing means to data mining*. [S.l.]: Academic Press, 2014.