

Priscila Gabriele Marques dos Santos

**UM ESTUDO SOBRE MODELOS DE REDES NEURAIIS QUÂNTICAS**

Trabalho de Graduação



Universidade Federal Rural de Pernambuco  
secretaria@preg.ufrpe.br  
<http://www.ufrpe.br/br/graduacao>

RECIFE  
2016



Universidade Federal Rural de Pernambuco  
Departamento de Estatística e Informática  
Bacharelado em Ciência da Computação

Priscila Gabriele Marques dos Santos

## **UM ESTUDO SOBRE MODELOS DE REDES NEURAIIS QUÂNTICAS**

*Monografia apresentada ao Curso de Bacharelado em Ciência da Computação da Universidade Federal Rural de Pernambuco como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.*

Orientador: *Adenilton José da Silva*

RECIFE  
2016



MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO  
UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO (UFRPE)  
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

<http://www.bcc.ufrpe.br>

**FICHA DE APROVAÇÃO DO TRABALHO DE CONCLUSÃO DE CURSO**

Trabalho defendido por Priscila Gabriele Marques dos Santos como requisito para conclusão do curso de Bacharelado em Ciência da Computação da Universidade Federal Rural de Pernambuco, intitulado **Um Estudo sobre Modelos de Redes Neurais Quânticas**, orientado por Adenilton José da Silva e aprovado pela seguinte banca examinadora:

*Adenilton José da Silva*

Adenilton José da Silva  
DEINFO/UFRPE

*Wilson Rosa de Oliveira*

Wilson Rosa de Oliveira  
DEINFO/UFRPE

*Fernando de Paula Neto*

Fernando de Paula Neto  
CIn/UFPE

# Agradecimentos

Aos meus familiares. Ao meu pai, Waldecy Filho, o ensinamento de que o conhecimento é a herança mais importante que poderia me oferecer. À minha mãe, Gilvanete Silva, a dedicação e suporte que me foi oferecido. Ao meu irmão Leonardo Santos, o silêncio que me foi proporcionado nos momentos em que mais precisei de concentração. Em especial, agradeço ao meu avô Waldecy Marques, que veio a falecer durante a escrita desse trabalho. Agradeço, querido avô, os ensinamentos de vida e todo o amor.

Aos meus amigos. Às minhas grandes amigas Lara Cardoso Figueirêdo e Izabella Gonçalves de Moraes, que são como uma família para mim, o apoio em todos os momentos. Da mesma forma, ao meu namorado Rodrigo Sousa, que esteve sempre me motivando e me proporcionando forças para continuar. Agradeço também a Vinícius de Moraes e Maria Luísa dos Santos, a amizade e união.

Ao meu orientador. Agradeço ao meu orientador, o professor Adenilton Silva, a dedicação e disponibilidade, assim como todos os ensinamentos e orientações. Agradeço também ao professor Wilson Rosa que muito contribuiu para minha formação acadêmica.

A todos os professores do curso. Dos professores com quem tive a oportunidade de estudar, em particular agradeço a Pablo Sampaio, Rodrigo de Souza, Abner Barros e Suzana Sampaio. Por todos os projetos que tive a oportunidade de participar nos meus primeiros anos de curso, agradeço ao professor Gilberto Cysneiros que me proporcionou diversas oportunidades. Também, agradeço ao professor George Valença com quem aprendi sob sua orientação em um projeto de extensão.

Aos colegas de curso. Por todo companheirismo, agradeço a Allan Costa, Allan Monteiro, Carlos Ferraz, Cleberson Barbosa, Gleidson Campos, João Otávio e Márdeni Ferreira. Em especial gostaria de agradecer também a Thays Silva, com quem muito aprendi, a todas as madrugadas de estudo, conversas e planos.

*To confine our attention to terrestrial matters would be to limit the human  
spirit.*

—STEPHEN HAWKING

# Resumo

Uma Rede Neural Quântica é um modelo que combina a Computação Quântica com propriedades das Redes Neurais Artificiais. Redes Neurais Artificiais são estruturas matemáticas que surgiram por inspiração no funcionamento do cérebro humano, sendo capazes de adquirir conhecimento através de um processo de aprendizagem e de armazenar conhecimento em pesos sinápticos. Por sua vez, a Computação Quântica é um modelo de computação que usa efeitos quânticos de nível microscópico para executar tarefas computacionais, tendo produzido resultados que em alguns casos são mais rápidos do que suas contrapartes clássicas conhecidas.

Por unirem as vantagens da superposição quântica das informações ao processamento paralelo das redes neurais, as Redes Neurais Quânticas possuem grande potencial como modelos computacionais. Dessa forma, diversos modelos que se propõe a ser de Redes Neurais Quânticas foram propostos. Porém, ainda não existe na literatura uma definição formal do que se constitui como uma Rede Neural Quântica. Isso se deve ao fato de que as Redes Neurais Artificiais e a Computação Quântica são campos que possuem dinâmicas distintas.

Nesse contexto, o presente trabalho tem seu foco nos modelos de Redes Neurais Quânticas, contribuindo com a área a partir do estudo de propostas diversas. Busca-se, dessa forma, revisar o modo como as propostas associam a dinâmica não-linear das redes neurais com a dinâmica linear e unitária da computação quântica.

**Palavras-chave:** computação quântica, redes neurais artificiais, redes neurais, redes neurais quânticas, redes quânticas sem peso, redes quânticas de Hopfield, modelos de redes quânticas.

# Abstract

A Quantum Neural Network is a model that combines Quantum Computing with Artificial Neural Network's properties. Artificial Neural Networks are mathematical structures that were created with an inspiration in human brain functioning, being able to gain knowledge through a learning process and to storage knowledge in synaptic weights. On the other hand, Quantum Computing is a computing model that uses microscopic level quantum effects to execute computing tasks, generating results that, in some cases, are faster than in its know classical counterparts.

By uniting the quantum superposition of information advantages to the neural networks parallel processing, Quantum Neural Networks have great potential as computing models. Therefore, several Quantum Neural Network models have been proposed. However, there is not yet in literature a formal definition of what is the constitution of a Quantum Neural Network. That is due to the fact that Artificial Neural Networks and Quantum Computing are fields that have fundamentally distinct dynamics.

In this context, this paper focuses on Quantum Neural Networks, contributing with its area through studying different propositions. We seek, therefore, to review how those propositions associate the nonlinear dynamic from neural networks with the linear and unitary dynamic from quantum computing.

**Keywords:** quantum computing, artificial neural networks, quantum neural networks.

# Lista de Figuras

3.1	Modelo de Neurônio Artificial . . . . .	25
-----	---	----



# Lista de Quadros

4.1	Resumo da avaliação das propostas . . . . .	42
-----	---	----

# Lista de Acrônimos

<b>CQ</b>	Computação Quântica
<b>IA</b>	Inteligência Artificial
<b>RNA</b>	Rede Neural Artificial
<b>RNQ</b>	Rede Neural Quântica
<b>PMC</b>	Perceptron Multicamadas
<b>NNQA</b>	Rede Neural com Arquitetura Quântica
<b>QAM</b>	Memória Associativa Quântica

# Sumário

<b>1</b>	<b>Introdução</b>	<b>13</b>
1.1	Motivação . . . . .	13
1.2	Problema de Pesquisa . . . . .	13
1.3	Objetivos . . . . .	14
1.3.1	Geral . . . . .	14
1.3.2	Específicos . . . . .	14
1.4	Organização do trabalho . . . . .	14
<b>2</b>	<b>Computação Quântica</b>	<b>15</b>
2.1	Da Física Clássica à Física Quântica . . . . .	15
2.2	Da Física Quântica à Computação Quântica . . . . .	16
2.3	Bits Quânticos . . . . .	17
2.4	Múltiplos Qbits . . . . .	18
2.5	Portas Quânticas . . . . .	19
2.5.1	Porta NOT . . . . .	20
2.5.2	Inversão de Fase . . . . .	20
2.5.3	Hadamard . . . . .	20
2.5.4	NOT Controlado . . . . .	21
2.5.5	Porta Toffoli . . . . .	21
2.6	Circuitos Quânticos . . . . .	21
2.6.1	Portas Quânticas em Série . . . . .	22
2.6.2	Portas Quânticas em Paralelo . . . . .	22
2.7	Algoritmo de Grover . . . . .	23
<b>3</b>	<b>Redes Neurais</b>	<b>24</b>
3.1	Redes Neurais Artificiais . . . . .	24
3.2	Neurônio Artificiais . . . . .	24
3.3	Características das Redes Neurais . . . . .	26
3.4	Topologias . . . . .	27
3.4.1	Redes de Alimentação Direta . . . . .	27
3.4.2	Redes Recorrentes . . . . .	27
3.4.2.1	Redes de Hopfield . . . . .	27
3.5	Aprendizagem . . . . .	28
3.5.1	Aprendizagem Supervisionada . . . . .	28
3.5.2	Aprendizagem Não-Supervisionada . . . . .	28

---

3.6	Perceptron . . . . .	28
3.7	Perceptron Multicamadas . . . . .	29
<b>4</b>	<b>Modelos</b>	<b>32</b>
4.1	Redes Neurais com Arquitetura Quântica . . . . .	32
4.2	Perceptron Quântico sobre um Corpo . . . . .	35
4.3	A busca por uma Rede Neural Artificial . . . . .	38
4.4	Treinando um Rede Neural Quântica . . . . .	42
<b>5</b>	<b>Conclusão</b>	<b>45</b>
5.1	Trabalhos Futuros . . . . .	46
	<b>Referências</b>	<b>47</b>

# 1

## Introdução

### 1.1 Motivação

A Computação Quântica (CQ) teve início através dos trabalhos de Paul Benioff, Yuri Manin, Richard Feynman e David Deutsch. Richard Feynman propôs que para simular sistemas de mecânica quântica de forma eficaz, computadores que se utilizassem de fenômenos quânticos para realizar cálculos poderiam ser necessários. A área da CQ começou a ganhar destaque através dos trabalhos de David Deutsch, Peter Shor e Lov Grover, que mostraram existir possibilidade de construção de algoritmos quânticos com possibilidade de aplicações mais rápidas que os computadores clássicos. O campo da computação quântica investiga o poder de sistemas quânticos usados como máquinas computacionais.

Por sua vez, uma Rede Neural Artificial (RNA) é um modelo computacional paralelo de Inteligência Artificial (IA) que surgiu por inspiração nas capacidades cognitivas das redes neurais biológicas presentes no cérebro humano, em tentativa de mapear seu funcionamento para uma estrutura computacional. O poder computacional das redes neurais clássicas está no seu processamento de informações distribuído massivamente paralelo e na não-linearidade das transformações realizadas pelas unidades de processamento da rede.

As Rede Neural Quântica (RNQ) são modelos que unem os campos da Computação Quântica com as Redes Neurais Artificiais. Dessa forma, as propostas de RNQ buscam obter vantagens da CQ para produzir um modelo de rede neural com melhor tempo de execução que o dos modelos clássicos, por exemplo (EZHOV; VENTURA, 2000). Assim, as RNQ compõem uma área de estudo que desperta interesse a partir de seu grande potencial como um modelo computacional, sendo resultante da combinação de campos de estudo consagrados e com grande possibilidade de pesquisa.

### 1.2 Problema de Pesquisa

Apesar de ser uma área que vem ganhando destaque nos últimos anos, ainda não existe na literatura uma definição formal do que se constitui como uma Rede Neural Quântica. O que

se têm é uma grande quantidade de abordagens diferentes, aonde muitos pesquisadores acabam por usar suas próprias analogias ao estabelecer uma conexão entre a teoria quântica e redes neurais. A razão para isso está no fato de que a dinâmica não-linear dissipativa do neurônio computacional é fundamentalmente diferente da dinâmica linear unitária da computação quântica (SCHULD; SINAYSKIY; PETRUCCIONE, 2014). Nesse contexto, o presente trabalho aborda os modelos de RNQ, investigando-os quanto proposta quântica de rede neural.

## 1.3 Objetivos

### 1.3.1 Geral

Averiguar propostas de redes neurais quânticas, através de uma revisão do estado da arte.

### 1.3.2 Específicos

- Apresentar conceitos fundamentais a compreensão do campo;
- Expor propostas de correspondências entre o modelo clássico de rede neural e a teoria quântica;
- Verificar consistência dos modelos com a teoria quântica.

## 1.4 Organização do trabalho

Por se tratar de um trabalho sobre modelos de Redes Neurais Quânticas, os campos de redes neurais artificiais e computação quântica estão diretamente envolvidos. Essa pesquisa está organizada de forma a prover os conceitos básicos necessários ao entendimento das propostas de modelos a serem estudados. Dessa forma, o capítulo 2 tem seu foco na Computação Quântica, contemplando uma visão geral sobre a área e sua distinção da computação clássica. São também apresentados no capítulo 2 os conceitos fundamentais que servem de base para o estudo da quântica. Por sua vez, o capítulo 3 aborda Redes Neurais Artificiais. Nele, é apresentado o conceito de neurônio artificial e sua inspiração biológica no cérebro humano. Paradigmas de aprendizagem, arquiteturas e características gerais das redes neurais também são contempladas no capítulo 3. O capítulo 4 expõe as propostas de modelo de redes neurais artificiais aqui estudadas. Por fim, a conclusão encontra-se no capítulo 5.

# 2

## Computação Quântica

Neste capítulo serão apresentados os conceitos que permeiam a computação quântica, sendo necessários ao entendimento do conteúdo a ser exposto. Tendo seu foco nos aspectos lógicos da computação quântica, não se aprofundando nos aspectos físicos da construção dos computadores quânticos.

### 2.1 Da Física Clássica à Física Quântica

Até o início do século XIX, as descrições físicas na época se ajustavam de tal forma que alguns físicos acreditavam que não havia mais o que ser descoberto, através da mecânica de Newton, do eletromagnetismo de Maxwell, da termodinâmica e da mecânica estatística de Boltzmann e outros. As teorias físicas da época, hoje chamada de física clássica, pareciam conseguir descrever com suas leis grande parte dos fenômenos naturais. No entanto, no começo do século XX, através de observações da radiação emitida por corpos quentes, as teorias da física clássica estavam prevendo fenômenos absurdos tais como a existência do que foi denominado de Catástrofe do Ultravioleta.

Na física, o modelo teórico de um objeto capaz de absorver toda radiação eletromagnética que nele incidir, é denominado de corpo negro, devido ao fato de que um corpo com essa propriedade, em princípio, não poderia ser visto. A primeira menção a esse modelo deve-se a Kirchhoff, que também demonstrou que a intensidade de emissão de um corpo negro em função de sua frequência depende exclusivamente da temperatura do corpo, aumentando-se a temperatura, a radiação também aumenta.

A partir disso, Max Planck introduziu a ideia de osciladores elétricos no interior de uma cavidade de um corpo negro. Esses osciladores vibravam para frente e para trás através da agitação térmica. De acordo com a teoria clássica, a energia emitida deveria tender ao infinito à medida que o comprimento de onda da radiação fosse diminuído. Porém, os dados experimentais não confirmavam essa previsão quando os comprimentos de onda estavam próximos do ultravioleta, que tem comprimentos de onda curtos.

A Catástrofe do ultravioleta foi o nome dado a tal discordância entre o teórico e o

experimental, que culminou para o início da mecânica quântica. Ao tentar solucionar a questão em 1900, o físico alemão Max Planck deu o início ao desenvolvimento da teoria quântica. Planck propôs que a energia não era contínua, mas fragmentada em pequenos pacotes, que denominou de quantum, que vem do latim e significa “quantidade”, a partir da noção de unidade mínima, indivisível; já que o quantum seria uma unidade definida de energia proporcional à frequência da radiação. A partir de então surgiu a mecânica quântica.

Este foi o início da trajetória da Mecânica Quântica, estudando os eventos que transcorrem nas camadas atômicas e sub-atômicas, entre as moléculas, átomos, elétrons, prótons, pósitrons, e outras partículas. A mecânica quântica se tornou uma parte indispensável da ciência, sendo aplicada com sucesso em fenômenos que vão desde o interior do sol, incluindo a estrutura do átomo, fusão nuclear em estrelas, supercondutores, a estrutura do DNA, dentre outros. As regras da mecânica quântica são simples mas até especialistas apontam que podem ser contra intuitivas.

Os primeiros antecedentes da computação quântica podem ser encontrados no desejo de longa data dos físicos em melhor entender a mecânica quântica. Esforços de construir um sistema de processamento de informação quântico resultaram em um sucesso modesto até o momento. Pequenos computadores quânticos, capazes de realizar dezenas de operações em alguns qbits representam, o estado da arte da computação quântica. Computação quântica é o estudo das tarefas de processamento de informação que podem ser realizados utilizando sistemas que possuem propriedades inerentes a computação quântica.

## 2.2 Da Física Quântica à Computação Quântica

A área de computação quântica é interdisciplinar (NIELSEN; CHUANG, 2010), envolvendo física, no que diz respeito a conceitos e princípios da mecânica quântica; matemática, principalmente álgebra linear e espaços de Hilbert; estatística, probabilidade e processos estocásticos. Além disso, ela pode abranger diversas áreas da computação: lógica, teoria da computação, algoritmos e estruturas de dados, inteligência artificial e redes neurais.

Na década de 1980, Richard Feynman, físico ganhador do prêmio Nobel, mostrou em (FEYNMAN, 1982) que pareciam haver dificuldades essenciais de custo computacional ao se utilizar os computadores clássicos para simular sistemas quânticos, apontando que a construção de computadores baseados nos princípios da mecânica quântica possibilitariam que tais dificuldades pudessem ser contornadas.

Além das pesquisas sobre a possibilidade de implementação física de um computador que se utiliza da estrutura quântica da matéria, outro aspecto que despertou o interesse de físicos, matemáticos e cientistas da computação foi quanto ao ganho que computadores quânticos poderiam oferecer em relação aos computadores clássicos. Deutsch (DEUTSCH, 1985) levantou a questão sobre a possibilidade de um computador quântico ser capaz de resolver, de forma eficiente, problemas computacionais que não têm solução em um computador clássico ou mesmo



em uma máquina de Turing probabilística.

Deutsch em (DEUTSCH, 1985), além de levantar o questionamento sobre a capacidade dos computadores quânticos, constrói um exemplo simples demonstrando que os mesmos podem ter um poder computacional muito superior ao dos computadores atuais. A teoria vigente de computação quântica foi proposta por Deutsch e utiliza Máquinas de Turing Quânticas (DEUTSCH, 1985) e circuitos quânticos (DEUTSCH, 1989), constituindo uma abordagem que alavancou o crescimento da área. Demonstrou-se em (DEUTSCH, 1989) que os dois modelos propostos são equivalentes, possibilitando a criação de algoritmos quânticos tendo em vista o modelo formalizado.

A computação quântica ganhou grande notoriedade em 1994, quando Shor (SHOR, 1994) conseguiu resultados inéditos em algoritmos quânticos, demonstrando que dois problemas de extrema relevância (o problema de achar fatores primos de um número inteiro e o então chamado problema do ‘logaritmo discreto’), poderiam ser resolvidos eficientemente em um computador quântico. Esses problemas são importantes porque sua intratabilidade computacional, no que se refere a computadores clássicos, é o padrão computacional atualmente assumido por alguns dos métodos criptográficos mais utilizados para garantir a segurança dos dados.

Outra evidência da capacidade dos computadores quânticos veio com o algoritmo de busca proposto por Grover. Realizar uma busca por um elemento em uma base de dados não estruturada de  $n$  elementos requer tempo  $O(n)$ , tanto classicamente quanto probabilisticamente. Grover desenvolveu em (GROVER, 1996) um algoritmo quântico com uma técnica que realiza tal busca em tempo  $O(\sqrt{n})$ , obtendo um ganho polinomial. Apesar do algoritmo de Grover não possuir ganho exponencial como o de Shor, a aplicabilidade do resultado é um grande avanço, pois com ele é possível obter uma aceleração quadrática na resolução de problemas da classe NP.

## 2.3 Bits Quânticos

Na computação clássica, o bit é a estrutura básica de informação. A computação quântica se utiliza de um conceito análogo: o bit quântico ou qbit. O bit clássico assume apenas um estado em um determinado instante, esse estado é 0 ou 1. Um qbit, da mesma forma, assume um estado. Dois possíveis estados para um qbit são  $|0\rangle$  e  $|1\rangle$ , correspondendo aos dois estados do bit clássico. Porém, diferente do bit clássico, o qbit não assume apenas um estado fixo  $|0\rangle$  ou  $|1\rangle$  em determinado instante. O qbit está em uma sobreposição de estados ao mesmo tempo. É possível formar uma sobreposição, combinação de estados, na forma

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

onde,  $\alpha$  e  $\beta$  são números complexos. Os estados  $|0\rangle$  e  $|1\rangle$  são chamados de base computacional e formam uma base ortonormal para esse espaço vetorial. Não se pode examinar um qbit para determinar o seu estado, ou seja, os valores de  $\alpha$  e  $\beta$ . A mecânica quântica possibilita a obtenção

de informações mais restritas. Pode-se medir um qbit e obter 0 com probabilidade  $|\alpha|^2$  ou obter 1 com probabilidade  $|\beta|^2$ . Nota-se que  $|\alpha|^2 + |\beta|^2 = 1$ , as probabilidades devem somar 1. Geometricamente, pode-se interpretar isso como a condição para que o estado do qbit esteja normalizado com tamanho 1. Assim, o estado de um qbit é um vector unitário em um espaço vectorial complexo bidimensional. Enquanto não for medido, um qbit existe em um estado contínuo entre 0 e 1. Após a medição, ele colapsa probabilisticamente para um dos estados  $|0\rangle$  ou  $|1\rangle$ .

## 2.4 Múltiplos Qbits

Tratando de múltiplos qbits no mesmo sistema quântico, a mecânica quântica reserva propriedades adicionais. Um sistema composto por dois qbits possui quatro estados base, denotados por  $|00\rangle$ ,  $|01\rangle$ ,  $|10\rangle$  e  $|11\rangle$ . Um par de qbits pode existir em superposição desses estados, então o estado dos qbits implica associar um coeficiente complexo, chamada amplitude, com cada estado da base computacional. Assim, o estado descrevendo os qbits é

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle.$$

O resultado da medição de cada estado terá probabilidade  $|\alpha_x|^2$ , para o estado dos qbits após a medição  $|x\rangle$ , onde  $x$  é 00, 01, 10 ou 11. A condição de que as probabilidades devem somar 1 é expressada pela condição de normalização

$$\sum_{x \in \{0,1\}^2} |\alpha_x|^2 = 1$$

onde a notação  $\{0,1\}^2$  representa o conjunto de strings de tamanho dois nos quais cada símbolo é 0 ou 1. É possível medir apenas um subconjunto dos qbits. Ao se medir o primeiro qbit, resultará 0 com probabilidade  $|\alpha_{00}|^2 + |\alpha_{01}|^2$ , deixando o estado após a medição

$$|\psi'\rangle = \frac{\alpha_{00}|00\rangle + \alpha_{01}|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}.$$

O estado após a medição está novamente normalizado pelo fator  $|\alpha_{00}|^2 + |\alpha_{01}|^2$ , de modo que satisfaz a condição de normalização, como se espera de um estado quântico.

Tendo dois sistemas, configurados separadamente nos estados  $|\psi_1\rangle$  e  $|\psi_2\rangle$ , posteriormente unidos. Se não ocorreu interação entre os sistemas,  $|\psi_1\rangle$  e  $|\psi_2\rangle$  estão em espaços de Hilbert distintos. O estado  $|\psi\rangle$  do sistema global constituindo através dos dois sistemas pode ser representado como

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \in H_1 \otimes H_2$$

onde  $\otimes$  representa o produto tensorial. Os estados da computação clássica costumam ser descritos

através de produto cartesiano. Da mesma forma, os estados da computação quântica em geral são descritos através do produto tensorial. Sejam  $\{v_1, v_2\}$  e  $\{w_1, w_2\}$  as bases de dois espaços vetoriais complexos bidimensionais  $V$  e  $W$ . O produto tensorial entre  $V$  e  $W$  possui como base  $\{v_1 \otimes w_1, v_1 \otimes w_2, v_2 \otimes w_1, v_2 \otimes w_2\}$

Um estado importante é o estado de Bell ou par EPR

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

Esse estado não pode ser descrito como produto tensorial de cada um de seus qbits. Ou seja, não existem qbits tais que

$$(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) = |00\rangle + |11\rangle$$

pois,

$$(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle) = \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle$$

e  $\alpha_1\beta_2 = 0$  implica  $\alpha_1\alpha_2 = 0$  ou  $\beta_1\beta_2 = 0$ . Estados como esse, que não podem ser decompostos em produtos tensoriais, são denominados de estados emaranhados.

## 2.5 Portas Quânticas

Análogo ao modo como computadores clássicos são formados por circuitos elétricos contendo fios e portas lógicas, um computador quântico é formado por um circuito quântico contendo fios e portas quânticas para transportar e manipular a informação quântica. O princípio fundamental na mudança de estados em um sistema quântico são as transformações unitárias. Uma transformação unitária é uma transformação linear que é reversível e cuja inversa é igual à sua conjugada transposta. Ou seja, a matriz  $U$  é unitária se

$$UU^\dagger = U^\dagger U = I,$$

onde  $U^\dagger$  representa a conjugada transposta da matriz  $U$  e  $I$  é a matriz identidade. As transformações unitárias são inerentes a qualquer sistema quântico que sofre mudanças em seus estados. Vetores de estados são transformados por matrizes unitárias. Uma transformação unitária preserva o mesmo comprimento do vetor e a mesma quantidade de informação do sistema. A preservação do comprimento do vetor garante que a probabilidade total de um conjunto de estados sempre permaneça igual a 1. Portas quânticas em um único qbit podem ser descritas por matrizes  $2 \times 2$ . Qualquer matriz unitária específica uma porta quântica válida.

### 2.5.1 Porta NOT

A porta NOT quântica, atua linearmente, ou seja, leva do estado  $\alpha|0\rangle + \beta|1\rangle$  para  $\alpha|1\rangle + \beta|0\rangle$ . A porta quântica NOT pode ser expressa na forma de matriz:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Se o estado quântico  $\alpha|0\rangle + \beta|1\rangle$  é escrito na notação de vetor

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix},$$

com a entrada na primeira linha correspondendo a amplitude de  $|0\rangle$  e a segunda linha a amplitude de  $|1\rangle$ , a saída após a aplicação da porta NOT é

$$X \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}.$$

### 2.5.2 Inversão de Fase

A porta de inversão de fase não realiza alterações do estado  $|0\rangle$  e troca o sinal do estado  $|1\rangle$ , sendo descrita pela matriz

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

### 2.5.3 Hadamard

Uma das mais importantes portas da computação quântica, a porta de Hadamard é descrita pela matriz

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

Hadamard tem por função

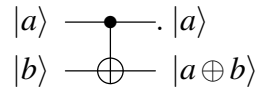
$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$|1\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle),$$

sendo representada em um circuito pelo símbolo  $\boxed{H}$

### 2.5.4 NOT Controlado

A porta NOT Controlado, ou CNOT, é o análogo quântico da porta clássica XOR reversível. A CNOT possui o funcionamento descrito por

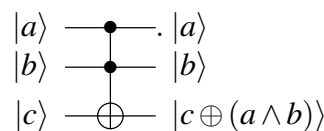


O estado  $|a\rangle$  é fornecido como entrada, resultando em  $|a\rangle$ . Também, é fornecido como entrada o estado  $|b\rangle$ , resultando  $|a \oplus b\rangle$ , onde  $\oplus$  é a operação de ou-exclusivo (XOR). Pode-se interpretar o qbit  $|a\rangle$  como um sinal de controle indicando se o qbit  $|b\rangle$  deve ser negado. Caso o qbit  $|a\rangle$  esteja ativo,  $|b\rangle$  é negado. Caso contrário,  $|b\rangle$  não sofre modificação. Essa transformação é descrita pela matriz

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

### 2.5.5 Porta Toffoli

A porta Toffoli é similar a porta CNOT, mas possui dois bits de controle. A porta Toffoli nega o último qbit apenas quando ambos os qbits acima estiverem no estado  $|1\rangle$ . A operação pode ser escrita como levando o estado  $|a, b, c\rangle$  para  $|a, b, c \oplus (a \wedge b)\rangle$



A porta Toffoli é universal. De modo que, com duas portas Toffoli, se pode construir qualquer porta lógica da computação clássica.

## 2.6 Circuitos Quânticos

Um circuito quântico lida com os princípios operacionais da computação quântica com múltiplos qbits. Em geral, uma porta quântica é representada por uma matriz unitária. O diagrama de um circuito quântico ilustra a combinação de portas quânticas que o compõe. Diferente da versão clássica, é possível realizar uma combinação linear dos qbits. A entrada dos circuitos são estados quânticos, cuja representação matricial é expressa por

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

para um único qbit. Da mesma forma, para dois qbits

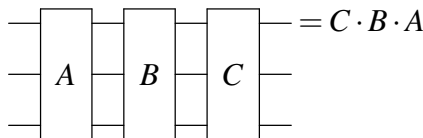
$$|0\rangle \otimes |0\rangle = |00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; |0\rangle \otimes |1\rangle = |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$|1\rangle \otimes |0\rangle = |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}; |1\rangle \otimes |1\rangle = |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Essas matrizes que representam os estados são transformadas de acordo com as matrizes das portas quânticas do circuito.

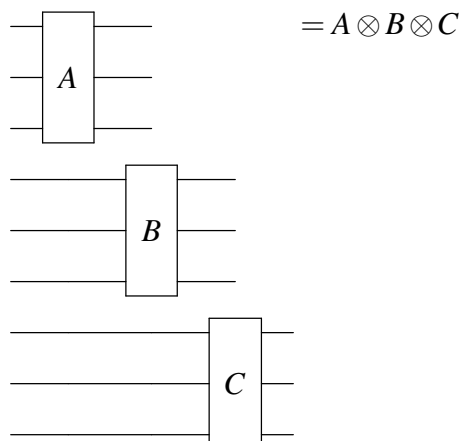
### 2.6.1 Portas Quânticas em Série

Quando portas quânticas estão conectadas em série, utiliza-se o produto escalar das matrizes. Considere várias portas  $A$ ,  $B$  e  $C$  agindo sobre o mesmo subconjunto de qbits, essas portas devem ser aplicadas em série e sua ação pode ser obtida a partir do produto escalar das matrizes em sua ordem inversa.



### 2.6.2 Portas Quânticas em Paralelo

Quando as portas quânticas estão dispostas em paralelo, a ação das portas nos qbits pode ser obtida através do produto tensorial das representações matriciais dos qbits.



## 2.7 Algoritmo de Grover

O algoritmo de Grover é um algoritmo probabilístico capaz de proporcionar uma melhora quadrática no desempenho em buscas não estruturadas, sendo sua complexidade  $O(\sqrt{N})$ . Assumindo uma função  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  e que existe exatamente uma cadeia binária

$$f(x) = \begin{cases} 1 & \text{se } x = x_o \\ 0 & \text{se } x \neq x_o, \end{cases}$$

é necessário encontrar  $x_o$ . Dessa forma, supõe-se um sistema com  $N = 2^n$  estados  $S_1, S_2, \dots, S_n$  representados por cadeias de  $n$  bits, onde um único estado  $S_m$  satisfaz um C que pode ser executado em uma unidade de tempo:  $C(S_m) = 1$ . E, para os demais estados  $C(S) = 0$ . O algoritmo de Grover pode ser descrito a partir dos passos:

1. Inicia com um estado  $|0\rangle$
2. Aplica  $H^{\otimes n}$
3. Por  $\sqrt{2^n}$  vezes:
  - (a) Aplica o operação de inversão de fase:  $U_f(I \otimes H)$
  - (b) Aplica a operação de inversão sobre a média:  $-I + 2A$
4. Realiza uma medição nos qbits

No passo 2, se obtém uma superposição de todos os estados. Com a operação de inversão de fase, estando o sistema em um estado  $S$ , caso  $C(S) = 1$ , sua fase será rotacionada em  $\pi$  radianos. Caso contrário, o sistema permanece inalterado. Dessa forma, apenas o estado desejado terá o sinal da amplitude invertido. Com a operação de inversão sobre a média, a amplitude de um estado será aumentada quando o estado estiver abaixo da média e será diminuída quando estiver acima da média antes da operação. Assim, a amplitude do estado desejado será aumentada, enquanto que a dos demais será diminuída. Após  $\sqrt{2^n}$  repetições, é realizada a medição do registrador que resultará no rótulo de  $n$  bits do estado marcado  $C(S_m) = 1$  com probabilidade mínima 0,5 (GROVER, 1996).

# 3

## Redes Neurais

Neste capítulo serão apresentados conceitos fundamentais ao estudo de Redes Neurais Artificiais como o de unidade de processamento, o neurônio, as operações realizadas pelo neurônio, que em geral, soma as entradas e mapeia não-linearmente o resultado em um valor de saída. Além disso, o capítulo contempla a estrutura de interconexão entre os neurônios, as dinâmicas das redes, paradigmas e regras de aprendizagem, que regulam a modificação das forças de interligação.

### 3.1 Redes Neurais Artificiais

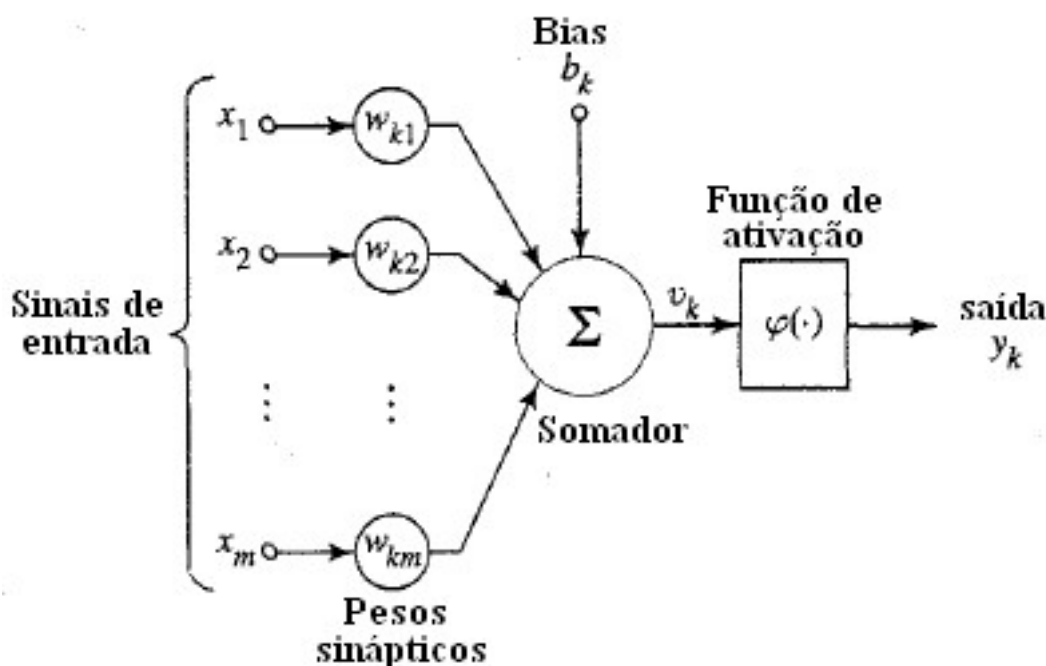
As Redes Neurais Artificiais podem ser definidas como estruturas matemáticas que surgiram por inspiração no funcionamento do cérebro humano. Em (FIESLER; BEALE, 1996) uma Rede neural Artificial é descrita como um processador distribuído massivamente paralelo composto por unidades de processamento mais simples. Esse processador possui uma inclinação natural para armazenar conhecimento experimental e torná-lo propenso para uso, se assemelhando ao cérebro humano por adquirir conhecimento através de um processo de aprendizagem e por armazenar o conhecimento adquirido em pesos sinápticos.

O cérebro é não linear, possui arquitetura altamente complexa e realiza processamento paralelo de informação a partir das unidades funcionais densamente interconectadas que o compõem (HAYKIN; NETWORK, 2004). Essas unidades do cérebro são denominados de neurônios e proporcionam ao cérebro a capacidade de realizar algumas tarefas, como reconhecimento de padrões, de maneiras mais rápida que os computadores atuais. Dessa forma, diversos modelos matemáticos de neurônios artificiais foram desenvolvidos tendo inspiração no conhecimento biológico sobre o funcionamento do cérebro humano (TAGLIARINI; CHRIST; PAGE, 1991).

### 3.2 Neurônio Artificiais

Um neurônio artificial é um modelo computacional inspirado no comportamento do neurônio biológico. Um neurônio artificial tem em sua estrutura entradas, os pesos atribuídos a





**Figura 3.1:** Modelo de Neurônio Artificial

Fonte:([HAYKIN; NETWORK, 2004](#))

cada conexão, um somador para os valores dos sinais de entrada multiplicados por seus respectivos pesos e uma função de ativação para limitar a saída do neurônio. A sua representação pode ser vista na Figura 3.1. Os estímulos são captados pelas entradas e ponderados. Especificamente, um sinal  $x_j$  na entrada da sinapse  $j$  conectada ao neurônio  $k$  é multiplicada por seu peso sináptico  $w_{kj}$ . Esses valores são então somados pelo somador, ou função somatório:

$$u_k = \sum_{j=1}^m w_{kj}x_j,$$

onde  $x_1, x_2, \dots, x_m$ , são os sinais de entrada e  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos do neurônio  $k$ . Assim,  $u_k$  é o valor que representa a ativação global dos sinais de entrada. Por sua vez, a função de ativação  $\phi$  calcula a saída do neurônio

$$y_k = \phi(u_k + b_k),$$

onde  $u_k$  é a combinação linear das entradas ponderadas e  $b_k$  é o bias. O bias é utilizado para aumentar a flexibilidade do modelo para que se adeque melhor aos dados, possibilitando o deslocamento da linha de separação. A função de ativação define as propriedades do neurônio artificial, podendo ser qualquer função matemática ([KRENKER; KOS; BEŠTER, 2011](#)). Deve-se utilizar a função de ativação que mais adequada ao problema que a rede neural se propõe a resolver.

O primeiro modelo de unidade de processamento de uma rede neural foi proposto em

([MCCULLOCH; PITTS, 1943](#)), artigo pioneiro no tema, contendo um estudo sobre o comportamento do neurônio biológico objetivando a criação de um modelo matemático correspondente. Eles mostraram como excitação, inibição e threshold podem ser usados para modelar neurônios. No modelo matemático proposto, cada sinapse possui um peso correspondente e cada neurônio possui um limiar  $\theta$ . Sendo  $w_i$  o peso da  $i$ -ésima sinapse para um neurônio, essa sinapse é dita ser excitatória se  $w_i > 0$  e inibitória quando  $w_i < 0$ . O neurônio estará ativo, indicando que produz um valor de saída, apenas quando a soma ponderada das entradas em um tempo  $t$  for pelo menos  $\theta$ .

### 3.3 Características das Redes Neurais

As redes neurais artificiais derivam seu poder computacional de sua estrutura distribuída massivamente paralela e de sua capacidade de aprender e, conseqüentemente, generalizar ([HAYKIN; NETWORK, 2004](#)). Onde, nesse contexto, generalização diz respeito a habilidade de se obter saídas adequadas para entradas previamente desconhecidas. Por conta disso, as Redes Neurais Artificiais costumam ser empregadas principalmente em campos relacionados com o processamento de informações, como na resolução de problemas de otimização, reconhecimento e classificação de padrões e previsão de séries temporais, por apresentarem características como capacidade de aprendizado, tomada de decisões e generalização ([BITTENCOURT, 2001](#)).

Em ([HAYKIN; NETWORK, 2004](#)) são apresentadas importantes propriedades das redes neurais artificiais, descritas a seguir:

- Não linearidade: Uma RNA que se configura como uma interconexão de neurônios não-lineares, também se faz não-linear.
- Mapeamento entrada-saída: Refere-se ao paradigma de aprendizado supervisionado, no qual os pesos são ajustados a partir de uma amostra de exemplos de treinamento devidamente classificada. Assim, cada exemplo da amostra consiste em um sinal de entrada com uma correspondente resposta desejada. A partir da exposição dos exemplos da amostra, os pesos são atualizados para minimizar a diferença entre a resposta desejada e a obtida. Esse processo de treinamento é realizado até que a rede não apresente mudanças significativas nos pesos. Os exemplos de treinamento previamente apresentados a rede podem ser novamente expostos em uma ordem diferente. Dessa forma, a rede aprende a partir dos exemplos através da construção de um mapeamento entrada-saída para um dado problema.
- Adaptabilidade: As RNAs possuem a habilidade embutida de adaptar seus pesos a mudanças. Uma rede neural que opera em um ambiente que sofre alterações com o tempo, pode ser projetada para atualizar seus pesos em tempo real.

- Resposta evidencial: Além da capacidade de prover informação a respeito de que padrão selecionar, uma rede neural também pode ser projetada para informar sobre a confiança no padrão apresentado.
- Informação contextual: Uma RNA é capaz de lidar naturalmente com informação contextual na medida em que cada neurônio na rede é potencialmente afetado pela atividade global dos demais neurônios na rede.

## 3.4 Topologias

O modo como os neurônios artificiais estão interconectados constitui a topologia, arquitetura ou grafo de uma rede neural (KRENKER; KOS; BEŠTER, 2011). A topologia é um fator importante a ser considerado na escolha do algoritmo de aprendizagem a ser utilizado. Os neurônios podem ser interconectados de diversas maneiras, mas em geral existe duas classes fundamentalmente distintas topologia de alimentação direta e topologia recorrente.

### 3.4.1 Redes de Alimentação Direta

Nas Redes de Alimentação Direta, a informação é passada das entradas até a saída em apenas uma direção, sem que existam ciclos. Não existe restrições quanto a quantidade de camadas, tipo da função de ativação usada dos neurônios ou número de sinapses entre os neurônios. São ditas serem estáticas por produzir apenas um conjunto de valores de saída, ou seja, a resposta de entrada independe do estado prévio da rede.

### 3.4.2 Redes Recorrentes

Nas Redes Recorrentes, parte da informação é passada em mais de uma direção. Assim, diferenciam-se das redes de alimentação direta por necessariamente conterem ao menos um ciclo onde ocorre a realimentação de informação. A presença da realimentação possibilita a criação de representações internas, proporciona um aumento na capacidade de aprendizado e performance da rede (HAYKIN; NETWORK, 2004). Dessa forma, as redes recorrentes são ditas dinâmicas.

#### 3.4.2.1 Redes de Hopfield

Ao se associar a função de energia com uma rede, onde apenas um neurônio muda de estado por vez, uma rede com conexões simétricas é estabelecida para um mínimo local. A partir disso, muitos problemas de otimização poderiam ser mapeados para funções de energia para redes neurais simétricas. A ideia básica é, dado um critério  $J$  a ser minimizado, achar uma rede de Hopfield cuja função de energia  $E$ , se aproxima de  $J$  e então deixar a rede estabelecer um equilíbrio e ler as soluções do estado da rede.

Uma rede de Hopfield possui pesos simétricos, indicando que  $w_{ij} = w_{ji}$ , e não ocorre auto-alimentação, ou realimentação de um neurônio para si mesmo,  $w_{ii} = 0$ . Além disso, a atualização é realizada de forma assíncrona. O estado de um neurônio  $i$ , denotado por  $S_i$ , pode assumir os valores 1, indicando que está ativo ou 0 quando não está ativo. A cada passo de tempo, um neurônio é selecionado aleatoriamente. Caso o neurônio escolhido seja  $i$ ,  $S_i$  é setado para 1 se e somente se  $\sum w_{ij}S_j \geq \theta_i$ , caso contrário,  $S_i$  é setado para 0.

## 3.5 Aprendizagem

Uma rede neural precisa ter a capacidade de aprender e melhorar sua performance através do aprendizado. Para tanto, deve-se saber que informações serão submetidas a rede. Serão apresentados abaixo dois grandes paradigmas de aprendizagem: aprendizagem supervisionada e não-supervisionada.

### 3.5.1 Aprendizagem Supervisionada

O processo de aprendizagem consiste no ajuste dos pesos da rede, quando necessário, de modo a minimizar o erro. Os pesos deverão ser ajustados, a partir do critério escolhido, quando a saída obtida pela rede for diferente da resposta desejada para aquele padrão. O conjunto de treinamento para o aprendizado supervisionado deve ser constituído de pares de entrada e saída desejada, tradicionalmente representados como vetores de dados.

### 3.5.2 Aprendizagem Não-Supervisionada

Também conhecido como aprendizagem auto-supervisionada, possui processo de aprendizagem no qual busca-se determinar como os dados são organizados. A rede processa as entradas, buscando categorizá-las de acordo com algum critério de semelhança. O conjunto de treinamento para aprendizado não supervisionado consiste apenas em exemplos que não estão rotulados.

## 3.6 Perceptron

O Perceptron, proposto em (ROSENBLATT, 1958) foi o primeiro modelo de rede neural implementado, consistindo em um modelo cognitivo de entradas conectadas a uma única camada de neurônios, que se comportam tal como o neurônio proposto em (MCCULLOCH; PITTS, 1943), porém com o acréscimo de sinapses ajustáveis. As redes Perceptron são capazes de treinar para classificar padrões apenas de classes linearmente separáveis, o que constitui uma grande limitação (GUYON, 1991).

Assim, o perceptron é a forma mais simples de rede neural usada para classificação de padrões linearmente separáveis. Geometricamente, quando treinado com padrões de duas classes

linearmente separáveis, o perceptron convergirá para uma superfície de decisão formada por um hiperplano entre as duas classes. Para tanto, o perceptron ajusta os pesos de cada componente do vetor de entrada, para que os exemplos de entrada, que pertencem a uma das classes, sejam corretamente classificados. Após o treinamento, outros exemplos de entrada de classificação desconhecida podem ser classificados de acordo com o lado do hiperplano em que estarão. Cada um dos componentes do vetor de entrada é multiplicado por seu peso correspondente, esses produtos são então somados e a função degrau é aplicada para obter a ativação, que será 1 ou -1

$$f(u_k) = \begin{cases} 1 & \text{se } u_k \geq 0 \\ -1 & \text{se } u_k < 0, \end{cases}$$

onde

$$u_k = \sum_{j=1}^m w_{kj}x_j + bias.$$

A regra de decisão é classificar o ponto representado pelas entradas  $x_1, x_2, \dots, x_m$  como classe 1 se a saída for 1 e como classe 2 caso seja  $-1$ .

O aprendizado é realizado de modo supervisionado e os pesos sinápticos do perceptron podem ser ajustados a cada iteração a partir do teorema de convergência do perceptron, que atualiza os pesos de acordo com a fórmula

$$w_i(n+1) = w_i(n) + \eta \varepsilon(n)x_i(n),$$

onde  $n$  é o contador de iterações,  $\eta > 0$  é a taxa de aprendizagem e  $\varepsilon(n)$  é o erro produzido pelo vetor de entrada na iteração  $n$ , dado por

$$\varepsilon(n) = d(n) - y(n),$$

onde  $d(n)$  é a saída desejada e  $y(n)$  é a saída do perceptron, ambos no tempo  $n$ . O valor da taxa de aprendizagem  $\eta$  afeta a velocidade do processo de aprendizado mas independente dela o algoritmo eventualmente convergirá.

## 3.7 Perceptron Multicamadas

O Perceptron Multicamadas (PMC) é o tipo de rede neural mais conhecido e mais utilizado (FIESLER; BEALE, 1996). As redes PMC são constituídas por uma camada de entrada, uma ou mais camadas intermediárias, também denominadas de camadas escondidas, e uma camada de saída. As múltiplas camadas do PMC são formadas por unidades básicas de processamento do tipo Perceptron (GUYON, 1991). Cada camada pode conter inúmeros neurônios e os neurônios de uma camada estão conectados a todos os neurônios da camada seguinte. As redes PMC possuem alimentação direta, indicando que as interconexões não formam

ciclos. Assim, as saídas dos neurônios de uma camada compõe as entradas dos neurônios da camada subsequente.

O treinamento das redes PMC é realizado de forma supervisionada através do algoritmo de retro-propagação do erro. O processo de aprendizagem consiste em duas diferentes etapas: a informação é propagada da camada de entrada até a de saída e, na segunda, o erro é retro-propagado da camada saída até a de entrada. Cada neurônio da rede possui uma função de ativação não-linear. A função sigmoide é comumente utilizada, sendo definida por:

$$y_j = \frac{1}{1 + e^{(-v_j)}},$$

onde  $y_j$  é a saída do neurônio  $j$  e  $v_j$  é o campo local induzido, ou seja, a soma ponderada das entradas mais o bias. Sem a presença da não linearidade, a relação entre entrada e saída poderia ser reduzida a de um perceptron de uma camada.

Na primeira etapa do treinamento o vetor de entrada de dados é processado, passando pelo número de camadas intermediárias que compõe a rede, camada a camada, até atingir a saída. Cada neurônio realiza a soma ponderada das entradas associadas a ele mais o bias e produz uma saída aplicando o valor a função de ativação. Esse processo ocorre até se atingir a camada de saída, que fornecerá a saída que representa a resposta da rede. Durante essa etapa os pesos sinápticos não sofrem alterações. Durante a segunda etapa, os pesos da rede são ajustados de acordo com uma regra de correção de erros. A resposta da rede é subtraída da resposta desejada para produzir um sinal de erro, que será propagado no sentido inverso das conexões sinápticas. Esse sinal é definido por

$$e_j(n) = d_j(n) - y_j(n),$$

onde  $j$  é o neurônio da camada de saída. A medida que o erro é retro-propagado, os pesos são ajustados visando minimizar o erro a cada iteração. Essas etapas são repetidas para os pares do conjunto de treinamento até que seja obtido um erro de saída mínimo desejável.

Pode-se tomar a norma quadrada do vetor de erro,  $E(n) = ||e(n)||^2$ , como a medida escalar do desvio da rede de seu comportamento ideal para aquela entrada. Dessa forma, pode-se medir o desvio da rede de seu ideal para todo o conjunto de treinamento

$$E = \sum_{n=1}^m E(n),$$

onde  $m$  é o número de amostras no conjunto de treinamento. Se o conjunto de treinamento e a arquitetura da rede são fixos,  $E$  é apenas uma função dos pesos da rede:  $E = E(w)$ .

Assim, treinar uma rede PMC corresponde a procurar pesos que minimizem  $E$ . Para tanto, pode-se utilizar o método gradiente, que consiste em iterativamente alterar o espaço de pesos proporcionalmente ao gradiente negativo da função a ser minimizada. Dessa forma, os

pesos são iterativamente atualizados de acordo com

$$w^{n+1} = w^n - \eta \nabla E,$$

onde  $\nabla E$  representa o gradiente de  $E$  em relação a  $w$  e  $\eta$  é a taxa de aprendizagem.

# 4

## Modelos

As redes neurais são muito dependentes de abordagens não-lineares para realizar seu processamento de dados. Por sua vez, a quântica é uma teoria linear. Dessa forma, criar um modelo de rede neural quântica, unindo esses dois campos distintos, não é algo trivial. Para superar essa dificuldade, modelos muito distintos foram propostos, em termos de diferentes correspondências entre o modelo clássico de rede neural e a teoria quântica. Neste capítulo são apresentados algumas dessas propostas de modelo de redes neurais quânticas.

### 4.1 Redes Neurais com Arquitetura Quântica

Em (PANELLA; MARTINELLI, 2011) foi definido a Rede Neural com Arquitetura Quântica (NNQA), modelo de rede neural que se utiliza do paralelismo quântico, aplicando uma busca exaustiva pela rede ótima através do uso de operadores não-lineares. Um operador é um mapeamento de um espaço vetorial para outro. Um operador não linear é um operador que não satisfaz as condições de linearidade, que são:  $A(x + y) = Ax + Ay$  e  $A(\alpha x) = \alpha Ax$ ,  $\forall x, y \in V$  e  $\forall \alpha \in K$ , onde  $V$  é um vetor e  $K$  um corpo. Ou seja, para quaisquer espaços vetoriais  $U$  e  $V$  sobre um corpo  $K$ , o operador  $A : U \rightarrow V$  é não linear se  $A(\alpha x + \beta y) \neq \alpha Ax + \beta Ay$  para todo  $x, y \in U$  e para todo  $\alpha, \beta \in K$ .

No modelo NNQA, as entradas e os pesos são representados por qbits, e seu procedimento requer a obtenção de uma sequência de qbits para representar em superposição a rede neural artificial e sua performance. Dessa forma, espera-se obter uma superposição de estados, os quais representam todas as possíveis redes neurais artificiais obtidas com uma arquitetura escolhida.

As entradas da rede neural artificial e dos neurônios, assim como as sinapses e bias da rede neural quântica, são considerados como ângulos normalizados no intervalo  $-\frac{\pi}{2}$  e  $\frac{\pi}{2}$ . Os neurônios são conduzidos por números complexos de magnitude 1 e argumentos dados pelos blocos que realizam as sinapses de entrada através da função booleana  $F_0$ . Cada um dos componentes da entrada para a rede neural quântica é acoplado com sua sinapse através de  $F_0$ . Dessa forma, a função  $F_0 : B^{2N_q} \rightarrow B^{N_q}$  realiza a soma de dois números complexos de magnitude



1 e possui como saída o argumento da soma

$$z = \arctg \frac{\text{sen}(y) + \text{sen}(\theta)}{\text{cos}(y) + \text{cos}(\theta)},$$

onde  $y$  é a entrada para um neurônio e  $\theta$  é o parâmetro da sinapse. Os valores  $y$ ,  $\theta$  e  $z$  são quantizados na forma

$$\beta = \pi(-0.5 + \frac{k}{2^{N_q} - 1}), k = 0, 1, \dots, (2^{N_q} - 1),$$

sendo representados pelos  $N_q$  bits utilizados para o inteiro  $k$ .

A função de ativação do neurônio sobre as  $N_n$  entradas é realizada pela função booleana não linear  $F_1 : B^{N_n N_q} \rightarrow B^{N_q}$  que calcula

$$\alpha = \arctg\left(\frac{X}{R}\right),$$

onde  $R$  representa a parte positiva e  $X$  a parte imaginária do resultado da soma dos números complexos realizada em  $F_0$ . O valor de saída  $y$  é obtido através da adição de  $\alpha$  ao bias  $\zeta$ . Caso  $y$  não esteja no intervalo  $-\frac{\pi}{2}$  e  $\frac{\pi}{2}$ , o valor é reajustado para o intervalo. Assim, a saída do neurônio é obtida a partir do cálculo da soma dos argumentos de sua entrada (realizado por  $F_0$ ), seguido da adição desse resultado ao bias  $\zeta$  (realizada por  $F_1$ ).

Dessa forma, a quantidade de parâmetros na camada é  $N_x(N_0 + 1)$ , onde  $N_x$  é a quantidade de neurônios na camada  $x$  e  $N_0$  é o número de componentes que chegam no neurônio (correspondente, na primeira camada, ao número de componentes de  $X^{(p)}$  do conjunto de treinamento e, para as demais, ao número de neurônios da camada anterior a ela). A rede neural quântica é representada por uma cadeia de qbits. Sendo  $n$  o número de qbits, pode-se representar todas as possíveis redes neurais quânticas pelos  $N = 2^n$  estados da superposição

$$|\Psi\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^N |\Psi_1^{(k)}\rangle \Psi_2^{(k)} \dots \Psi_n^{(k)}\rangle,$$

onde  $|\Psi^{(k)}\rangle$  corresponde a  $k$ -ésima rede neural quântica,  $k = 1, \dots, N$ , e  $|\Psi_j^{(k)}\rangle$ ,  $j = 1, \dots, n$ , ao  $j$ -ésimo qbit da representação.

O desempenho da rede neural quântica na  $p$ -ésima amostra é medido pela magnitude  $\Delta^{(p)}$  do erro absoluto entre a saída desejada  $t^{(p)}$  e a saída da rede  $u^{(p)}$ . A função booleana  $F_2 : B^{2N_q} \rightarrow B^{N_q}$  é responsável por determinar o valor absoluto do erro da rede neural quântica em relação a uma amostra do conjunto de treinamento. Os valores de  $t^p$  e  $u^p$  são definidos no intervalo  $-\frac{\pi}{2}$  e  $\frac{\pi}{2}$ , e  $\Delta^{(p)}$  está entre 0 e  $\pi$ . A performance geral  $\Delta$  é calculada através da soma de todos os  $\Delta^p$ ,

$$\Delta = \sum_{p=1}^{N_p} \Delta^p = \sum_{p=1}^{N_p} |t^p - u^p|.$$

A função  $F_2$  pode ser considerada uma função booleana  $F_3$  da entrada  $X^p$ , ou seja,  $\Delta^{(p)} = F_3(X^{(p)})$ , já que  $u^{(p)}$  é uma função booleana de  $X^{(p)}$ . Dessa forma,  $F_3$  pode ser implementada como uma porta quântica linear  $U_f$ , tendo como entrada a superposição  $|\Psi\rangle$  de todas as redes neurais quânticas e  $|0N_q\rangle$ , onde  $|0N_q\rangle$  é uma cadeia de  $N_q$  qbits de  $|0\rangle$ . A saída de  $U_f$  é a superposição  $|\Psi, \Delta^{(p)}\rangle$ , onde cada rede neural quântica está emaranhada com seu desempenho.

Na superposição  $|\Psi, \Delta^{(p)}\rangle$ , o desempenho em todo o conjunto de treinamento é medido pela adição dos qbits  $|\Delta^{(p)}\rangle$  relacionados aos erros para cada  $p = 1, 2, \dots, N_p$ , até que se tenha  $|\Psi, \Delta\rangle$ , onde cada rede neural quântica está emaranhada com seu desempenho sobre todas as amostras do conjunto de treinamento. Para reduzir o número de qbits a serem manipulados, pode ser considerado um valor  $\delta$  representando aproximadamente a média dos desempenhos  $\Delta^{(p)}$ . Assim, a superposição será

$$|\Phi\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^N |\Psi^{(k)}, \delta^{(k)}\rangle,$$

onde  $\delta^{(k)}$  é o desempenho médio para a  $k$ -ésima rede neural quântica.

A rede neural quântica ótima é sinalizada por um rótulo adequado, o que indica a condição de que o seu valor  $\delta$  é menor ou igual a um determinado limiar. A rede neural quântica ótima é sinalizada através de uma função booleana de  $\delta g(\cdot)$  que sinaliza a rede neural quântica ótima. Essa rotulação é feita pela porta  $U_g$  que realiza a função booleana  $c = g(\delta)$  para determinar a verificação. A superposição  $\Phi$  é rotulada com um qbit  $|c\rangle$ , que assume valor  $|1\rangle$  quando a rede neural quântica correspondente na superposição possui um erro médio menor ou igual a um limiar e, do contrário,  $|0\rangle$ . Cada qbit  $|c\rangle$  é emaranhado com a rede neural quântica correspondente em  $|\Phi\rangle$ .

$$U_g : |\Phi, 0\rangle = |\Omega\rangle$$

$$|\Omega\rangle = |\Phi, c\rangle = \frac{1}{\sqrt{N}} \sum_{c^{(k)}=1} |\Psi^{(k)}, \delta^{(k)}, c^{(k)}\rangle + \frac{1}{\sqrt{N}} \sum_{c^{(k)}=0} |\Psi^{(k)}, \delta^{(k)}, c^{(k)}\rangle$$

É proposto um circuito quântico que realiza uma busca exaustiva por uma rede neural quântica que atenda ao requisito, caso exista. O circuito realiza  $n$  iterações, sendo  $n$  o número de qbits necessários para representar a rede neural quântica. Na  $m$ -ésima iteração, sua entrada é constituída por  $|\eta_m, \epsilon_{n-m}\rangle$ , onde  $|\eta_m\rangle$  é uma sequência de qbits determinada nas iterações anteriores e  $|\epsilon_{n-m}\rangle$  é uma superposição de todos os estados possíveis obtidos pelos últimos  $(n - m)$  qbits de  $|\Psi\rangle$ . A iteração  $m$  determina o valor do  $(m + 1)$ -ésimo qbit da solução ótima. Para esta finalidade, a entrada é dividida em duas superposições com base no valor do  $(m + 1)$ -ésimo qbit. Estas duas superposições são marcadas pela porta  $U_g$  introduzindo  $|c\rangle$ . As duas superposições resultantes são processadas pelo bloco AL. Sua saída é a superposição  $|\eta_{m+1}, \epsilon_{n-m-1}\rangle$ , onde o número de qbits determinados presentes em  $|\eta_{m+1}\rangle$  é aumentado em 1 em relação à entrada. No caso da iteração final,  $m = n - 1$ , a saída do circuito é a solução ótima, ou seja,  $|\eta_m\rangle$ . Esta saída é uma das soluções ótimas do problema, caso exista mais de uma. Na primeira iteração,  $m = 0$ , resulta  $|\eta_0, \epsilon_n\rangle = |\Psi\rangle$ , onde  $|\eta_0\rangle$  é nulo e  $|\epsilon_n\rangle$  coincide com  $|\Psi\rangle$ . Além

disso, o bloco AL verifica a existência de soluções ótimas. Se não existir, interrompe a operação do circuito.

## 4.2 Perceptron Quântico sobre um Corpo

Em (SILVA; LUDERMIR; OLIVEIRA, 2016) é proposto o Perceptron Quântico sobre um corpo, propondo uma generalização direta do Perceptron Clássico. Também é proposto um algoritmo de aprendizagem de arquitetura baseado em superposição, que otimiza pesos e arquiteturas. O algoritmo busca a melhor arquitetura em um conjunto finito de arquiteturas em tempo linear sobre o número de amostras no conjunto de treinamento. Se utiliza, para isso, de paralelismo quântico e de um operador quântico não-linear.

O Perceptron quântico proposto pode ser treinado com algoritmo quântico ou clássico, colocando todas as redes neurais para uma determinada arquitetura em superposição. Se os pesos estão na base computacional, o Perceptron quântico se comporta como o clássico. É definido um operador unitário para realizar soma  $\oplus$  e produto  $\odot$  de qbits baseado nas operações de um corpo, pois o conjunto de qbits n-dimensionais, operador de soma e produto tensorial não formam um corpo. Pode-se associar  $a \in F$  a qbits  $|a\rangle$  em uma base de um espaço vetorial  $\mathbb{V}$ , onde  $F(\oplus, \odot)$  é um corpo finito. Se F possui n elementos, o espaço vetorial será n dimensional. Dessa forma, se usa a operação de produto  $\odot$  em F para definir um novo produto entre vetores em  $\mathbb{V}$ . Foi definido  $|a\rangle \odot |b\rangle = |a \odot b\rangle$ , tal que  $|a \odot b\rangle$  está relacionado ao escalar  $a \odot b$ , onde  $|a\rangle$  e  $|b\rangle$  são qbits associados a escalares a e b. Assim, se definiu o operador linear  $P : \mathbb{V}^3 \rightarrow \mathbb{V}^3$

$$P|a\rangle|b\rangle|c\rangle = |a\rangle|b\rangle|c \oplus (a \odot b)\rangle.$$

É mostrado que P é um operador unitário por ser injetivo e enviar vetores na base  $B^3$  para vetores na base  $B^3$  considerando que um operador sobre um espaço vetorial é unitário se e somente se o operador envia uma base ortornormal para uma base ortornormal (HOFFMAN; KUNZE, 1990). Para a base computacional  $B = |a_1\rangle, |a_2\rangle, \dots, |a_n\rangle$  de um espaço vetorial  $\mathbb{V}$ , onde  $|a_i\rangle$  está associado ao elemento  $a_i$  do corpo finito  $F(\oplus, \odot)$ , o conjunto  $B^3 = |a_i\rangle|a_j\rangle|a_k\rangle$ , com  $1 \leq i, j, k \leq n$ , é uma base computacional de  $\mathbb{V}^3$ . Têm-se  $P|a\rangle|b\rangle|c\rangle = |a\rangle|b\rangle|c \oplus (a \odot b)\rangle$ , para a, b, c em B. Então,  $|c \oplus (a \odot b)\rangle \in B$  e  $|a\rangle|b\rangle|c \oplus (a \odot b)\rangle \in B^3$ . Por fim, é mostrado que P é injetivo definindo-se  $a, b, c, a_1, b_1 \in F$  de tal forma que  $P|a\rangle|b\rangle|c\rangle = P|a_1\rangle|b_1\rangle|c_1\rangle$ . Pela definição de P,  $a = a_1$  e  $b = b_1$ , aplicando o operador:

$$|a\rangle|b\rangle|c \oplus (a \odot b)\rangle = |a\rangle|b\rangle|c_1 \oplus (a \odot b)\rangle.$$

Assim,  $c_1 = c$ , já que  $|c_1 \oplus (a \odot b)\rangle = |c \oplus (a \odot b)\rangle$ .

Uma construção similar é utilizada para definir um outro operador que usa a operação  $\oplus$  de soma em F, o operador unitário de soma  $S : \mathbb{V}^3 \rightarrow \mathbb{V}^3$ . Para qbits  $|a\rangle$  e  $|b\rangle$  associados aos

escalares  $a$  e  $b$ , é definido  $|a\rangle \oplus |b\rangle = |a \oplus b\rangle$ , tal que  $|a \oplus b\rangle$  está relacionado ao escalar  $a \oplus b$ .  $S$  é definido por

$$S|a\rangle|b\rangle|c\rangle = |a\rangle|b\rangle|c \oplus (a \oplus b)\rangle.$$

Dessa forma, foi definido o produto de vetores sobre  $F$  por  $|a\rangle \odot |b\rangle = |a \odot b\rangle$  referente a  $P|a\rangle|b\rangle|0\rangle = |a\rangle|b\rangle|a \odot b\rangle$  e a soma  $|a\rangle \oplus |b\rangle = |a \oplus b\rangle$  referente a  $S|a\rangle|b\rangle|0\rangle = |a\rangle|b\rangle|a \oplus b\rangle$ .

No modelo proposto, as entradas  $|x_i\rangle$ , os pesos  $|w_i\rangle$  e a saída  $|y\rangle$  serão vetores unitários em  $\mathbb{V}$  representando escalares em um corpo  $F$ . É descrito na forma

$$|y\rangle = \bigoplus_{i=1}^n |x_i\rangle \odot |w_i\rangle.$$

A descrição acima não apresenta os qbits auxiliares, de forma que é apresentado uma descrição da configuração completa do Perceptron quântico proposto

$$|\Psi\rangle = |x_1, \dots, x_n, w_1, \dots, w_n, p_1, \dots, p_n, s_2, \dots, s_{n-1}, y\rangle,$$

onde  $|x\rangle = |x_1, \dots, x_n\rangle$  é o registrador quântico das entradas,  $|w\rangle = |w_1, \dots, w_n\rangle$  é o registrador quântico de pesos,  $|p\rangle = |p_1, \dots, p_n\rangle$  é o registrador auxiliar usado para armazenar os produtos  $|x_i \odot w_i\rangle$ ,  $|s\rangle = |s_2, \dots, s_{n-1}\rangle$  é o registrador auxiliar usado para armazenar as somas e, por fim,  $|y\rangle$  é o registrador de saída. Para realizar a superposição de todas as possíveis redes neurais quânticas, pode-se colocar o registrador quântico de pesos em superposição. Dessa forma, um único neurônio quântico pode estar em uma superposição de múltiplos neurônios simultaneamente.

Um perceptron quântico sobre um corpo  $d$ -dimensional finito e com  $n$  entradas necessita  $(4n - 1) \cdot d$  qbits para executar sua computação. Existem  $n$  registradores quânticos para armazenar as entradas  $x_i$ ,  $n$  registradores quânticos para armazenar os pesos  $w_i$ ,  $n$  registradores quânticos  $p_i$  para armazenar os produtos  $w_i \odot x_i$ ,  $n - 2$  registradores quânticos para armazenar  $|s_i\rangle$  somas  $\sum_{k=1}^i p_k$  e um registrador quântico para armazenar a saída  $|y\rangle = \sum_{k=1}^n p_k$ .

Pesos, entradas e saída em uma rede neural clássica possuem valores reais. No modelo proposto supõe-se memória finita e se utiliza elementos de um corpo finito  $(F, \oplus, \odot)$  para representar os parâmetros da rede neural. Sugere-se o uso de um operador quântico que realiza multiplicação matricial.  $M_{n \times n, n \times 1}$  multiplica uma matriz  $n \times n$ , matriz que descreve os pesos de uma camada, pela matriz  $n \times 1$  dos valores de entrada. É exemplificado o caso de uma rede neural  $N$  com duas camadas  $L_1$  e  $L_2$ , com 3 e 2 neurônios, respectivamente. A saída de  $N$  pode ser calculada por  $y = L_2 \cdot L_1 \cdot x$ , usando três matrizes  $L_1, L_2$  e  $x$ , onde

$$L_1 = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} L_2 = \begin{bmatrix} w_{14} & w_{24} & w_{34} \\ w_{15} & w_{25} & w_{35} \end{bmatrix} X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Para  $L_1 \cdot x = [o_1 o_2 o_3]^t$ , a ação de  $M$  para o exemplo pode ser definida como

$$M_{3 \times 3, 3 \times 1} |w_1, w_2, w_3, x_1, x_2, x_3, 0, 0, 0\rangle = |w_1, w_2, w_3, x_1, x_2, x_3, o_1, o_2, o_3\rangle,$$

onde  $w_i = w_{i1}, w_{i2}, w_{i3}$ . Cada camada do Perceptron Quântico sobre um corpo pode ser representada por uma matriz arbitrária

$$M_{2 \times 3, 3 \times 1} M_{3 \times 3, 3 \times 1} |L_2\rangle |L_1\rangle |x\rangle |000\rangle |00\rangle,$$

onde  $M_{3 \times 3, 3 \times 1}$  age sobre  $|L_1\rangle, |x\rangle$  com saída inicializada por  $|000\rangle$ .  $M_{2 \times 3, 3 \times 1}$  age sobre  $|L_2\rangle$ , a saída da primeira operação e o último registrador quântico. Dessa forma foi definida a abordagem matricial que pode ser usada para representar qualquer modelo de perceptron quântico multicamada de alimentação direta.

Supõe-se que o conjunto de treinamento e a saída desejada são compostos de dados clássicos e os dados são de alimentação direta. A superposição da saída clássica desejada possibilitará sobrepor as configurações de rede neural com seu desempenho. O algoritmo proposto usa superposição quântica para treinar um perceptron quântico sobre um corpo. No algoritmo de aprendizado com arquitetura baseada em superposição, assim como no modelo proposto em (PANELLA; MARTINELLI, 2011), a superposição das redes neurais armazena seu desempenho emaranhado com sua configuração. É utilizado um operador quântico não-linear para recuperar a arquitetura e os pesos da configuração da rede neural com melhor desempenho.

Para algumas arquiteturas de rede neural, todas as amostras do conjunto de treinamento  $P = p_1, p_2, \dots, p_k$  são apresentadas para cada uma das configurações de rede neural. Calcula-se a saída e então se pode buscar pelos melhores parâmetros. Sejam  $N_0, N_1, \dots, N_{m-1}$   $m$  operadores quânticos representando redes neurais com arquiteturas diferentes. Um circuito quântico com registradores quânticos pra selecionar arquitetura  $|a\rangle$  com  $\lceil \log_2(m) \rceil$  qbits, entrada  $|x\rangle$ , pesos  $|w\rangle$  e saída  $|o\rangle$  pode ser criado, onde o operador  $N_i$  é aplicado a  $|x, w, o\rangle$  se e somente se  $|a\rangle = |i\rangle$ . Se os qbits nos registradores quânticos  $|a\rangle$  e  $|w\rangle$  forem inicializados com o estado quântico  $H|0\rangle$ , o circuito estará em uma superposição representando todas as configurações de pesos possíveis para cada arquitetura. Inicializando o registrador quântico  $|x\rangle$  com uma amostra  $p$ , é possível apresentar a amostra  $p$  em todas as configurações de redes neurais na superposição simultaneamente. O algoritmo de aprendizado com arquitetura baseada em superposição é um algoritmo de aprendizagem quântica para qualquer modelo de rede neural quântica em que a entrada  $|p\rangle$ , saída  $|o\rangle$ , pesos  $|w\rangle$ , seletores de arquitetura  $|a\rangle$  e saída desejada  $|d\rangle$  são representados em registradores quânticos separados. A ideia principal do algoritmo é criar uma superposição de todas as possíveis configurações de redes neurais em um conjunto finito de arquiteturas e aplicar um operador quântico não-linear para recuperar a arquitetura e os pesos de uma configuração de rede neural com um desempenho desejado.

## 4.3 A busca por uma Rede Neural Artificial

Em (SCHULD; SINAYSKIY; PETRUCCIONE, 2014) é levantada a discussão de que não existe uma definição exata para redes neurais quânticas. O artigo, a partir do estudo dos modelos de redes neurais do tipo Hopfield, propõe três requisitos necessários para que um modelo possa se configurar como uma rede neural quântica, sendo eles:

1. O estado inicial do sistema quântico codifica qualquer cadeia binária de comprimento  $N$ . A rede neural quântica produz uma configuração de saída estável codificando o estado das  $2^M$  cadeias binárias de saída possíveis de comprimento  $M$ , estando mais próximo da entrada para alguma distância.
2. A rede neural quântica reflete um ou mais mecanismos básicos de computação neural
3. A evolução é baseada em efeitos quânticos, sendo totalmente consistente com a teoria quântica.

Por se tratar de um estudo voltado para redes neurais do tipo Hopfield com propriedade de memória associativa, o primeiro requisito busca assegurar que a rede neural quântica apresentará a característica de memória associativa, reconhecimento de padrões e outras propriedades centrais do processamento de informação neural em redes neurais de Hopfield. O segundo requisito proposto procura garantir uma relação com a computação neural, ainda que básica, pois exigir apenas que se tenha ao menos uma das características. É afirmado que o segundo requisito foi proposto com essa abrangência para que se possa atender à variedade de abordagens existentes e futuras, se fazendo necessária por excluir as chamadas Memórias Associativas Quânticas, algoritmos de computação quântica que simulam memória associativa sem que para isso tenham relação com redes neurais. Ainda assim, realiza-se uma análise das Memórias Associativas Quânticas como modelo de rede neural quântica. Por fim, o último requisito procura assegurar que a rede de fato se qualifica como uma rede neural quântica.

Os pontos levantados pelos autores parecem se constituir como bons requisitos para auxiliar na busca por modelos que de fato possam ser classificados como redes neurais quânticas, por solicitar que tenham relação direta e coerente com ambos os campos, a partir dos requisitos 2 e 3. Porém, o segundo requisito, que estabelece a relação com a computação neural, se faz pouco restritivo. A tentativa de abranger mais propostas a partir de uma exigência tão ampla no segundo requisito se apresenta talvez como uma das fraquezas da proposta.

Tento em vista os requisitos levantados, é realizada uma releitura da literatura existente para analisar se os modelos de fato se configuram como redes neurais quânticas, chegando a conclusão que nenhuma das propostas avaliadas atende a todos os requisitos. Para auxiliar na avaliação, as abordagens de redes neurais quânticas são agrupadas em quatro categorias que compartilham características semelhantes.

Uma das categorias avalia modelos para interpretar a função degrau como uma medição quântica. Das propostas que buscaram conseguir a dinâmica das redes neurais através de medições quânticas, são analisadas (KAK, 1995), (PERUŠ, 2000), (MENNEER; NARAYANAN, 1995) e (ZAK; WILLIAMS, 1998). Nenhuma delas consegue atender ao primeiro requisito proposto, por não possuírem a propriedade de memória associativa. Porém, na análise de (SCHULD; SINAYSKIY; PETRUCCIONE, 2014), a noção de usar medições quânticas para simular a convergência não-linear de memórias associativas mostra-se como a solução mais madura para o problema de incompatibilidade dinâmica encontrada na busca de um modelo de rede neural quântica.

Em (KAK, 1995) foi introduzida a noção do que seria uma computação neural quântica, através da interpretação da condição necessária para um estado estável  $x^0 = (x_1^0, \dots, x_N^0)$  em uma rede que quase se configura como uma rede de Hopfield definida pela matriz de pesos  $w$  com entradas  $w_{ij}$ ,  $sgm(wx^0) = x_0$ , como uma equação do autovalor de um sistema quântico,  $\hat{w}|x_0\rangle = \lambda|x_0\rangle$ . Dessa forma, a matriz de peso  $w$  opera como um operador com autovetor  $|x_0\rangle$  e autovalor  $\lambda = 1$ . Com isso, atualizar uma rede equivale a realizar uma medição quântica que seleciona os autoestados de um sistema. É observado em (KAK, 1995) que a função sigmoide é uma não-linearidade que não tem equivalente no formalismo da mecânica quântica.

Em (MENNEER; NARAYANAN, 1995) é proposta a chamada Rede Neural com Inspiração Quântica, concebida a partir da interpretação de muitos universos da mecânica quântica. A proposta é a de que, em vez de se ter uma rede armazenando vários padrões, pode-se modelar uma superposição de redes, cada uma armazenando um padrão. Com isso, têm-se o estado quântico da rede dado pelo seu vetor de pesos. Nela, para recuperar um padrão colapsa-se a superposição de vetores de peso para selecionar uma rede com o resultado desejado. Busca-se, em (ZAK; WILLIAMS, 1998), um formalismo quântico que capture as duas principais propriedades das redes de Hopfield: dissipação e não-linearidade. Para tanto, substitui-se a função degrau ou função de ativação sigmoide por uma medição quântica.

A segunda categoria busca analisar as propostas que envolvem a interação de pontos quânticos. Em (BEHRMAN et al., 2000) é apresentada uma proposta de rede neural que se utiliza da semelhança entre a função de atualização e a evolução de um estado quântico, evidenciados em (PERUŠ, 2000). Em (FABER; GIRALDI, 2002) é discutida a abordagem de (BEHRMAN et al., 2000) no contexto da solução do problema de incompatibilidade entre a computação neural e a computação quântica. Questiona-se como a dinâmica não linear da rede neural pode ser simulada por um sistema quântico. Como observado anteriormente por Behrman, a não linearidade surge naturalmente na evolução temporal de um sistema quântico através da função exponencial e da energia cinética não linear que alimenta o potencial  $V(\phi, t)$ . No entanto, Behrman afirma que para cálculos maiores do que simples portas, incluindo propriedades mais avançadas de redes Hopfield como reconhecimento de padrões (e consequentemente memória associativa), um modelo de matriz espacial é necessário.

A proposta de (BEHRMAN et al., 2000) mostra que a evolução natural de um sistema

de pontos quânticos interativos pode servir como um computador neural quântico, na medida em que o mapeamento desejado de uma entrada para uma saída pode ser projetado sob certas condições. Desta forma, qualquer sistema quântico com uma evolução dependente do estado inicial e com parâmetros apropriados é algum tipo de dispositivo de computação analógico para um determinado problema, assim como qualquer sistema físico dependente de entrada pode ser visto como um computador analógico. Tendo isso em vista, (SCHULD; SINAYSKIY; PETRUCCIONE, 2014) aponta ser discutível se o segundo requisito é suficientemente satisfeito. Embora se caracterize como um candidato natural para uma rede neural quântica, as interações entre sistemas quânticos ainda agem muito diferente de perceptrons neurais, e encontrar os parâmetros específicos que levariam à dinâmica correspondente a uma rede neural quântica totalmente funcional não se constitui uma tarefa trivial.

Em outra categoria são analisadas as propostas que buscam encontrar uma rede neural quântica a partir da perspectiva da computação quântica, onde apresentam-se construções de circuitos quânticos inspiradas pela dinâmica de redes neurais. Em (GUPTA; ZIA, 2001) propõe-se circuitos quânticos para qbits nos quais cada operação de computação quântica executada por operadores unitários  $\hat{U}$  é seguida por um operador dissipativo  $\hat{D}$ . Estas portas dissipativas mapeiam as amplitudes do estado quântico em  $c \in \mathbb{C}$  ou 0, dependendo se excede ou não um limiar  $\delta$ . O operador  $\hat{D}$ , conseqüentemente, imita a função de ativação do perceptron. É apontado por (SCHULD; SINAYSKIY; PETRUCCIONE, 2014) que os autores não dão um exemplo de um sistema quântico que poderia servir como uma porta. Em (FABER; GIRALDI, 2002) é sugerido a realização de (BEHRMAN et al., 2000) de uma rede neural quântica para servir como o operador dissipador não linear  $\hat{D}$ .

Em (PANELLA; MARTINELLI, 2011) são sugeridos operadores quânticos não-lineares gerais para construir uma rede de alimentação direta na qual os qbits são sucessivamente emaranhados uns com os outros. Em (SILVA; OLIVEIRA; LUDERMIR, 2016) desenvolveu-se um modelo de rede neural quântica lógica que se baseia em redes neurais clássicas sem peso nas quais funções de atualização neural são armazenadas em uma tabela semelhante a uma Memória de Acesso Aleatório em um computador. Uma cadeia de entrada binária endereçaria um dos registradores e conduziria a uma determinada saída. Este modelo contém a capacidade de RNAs para ser treinado por um algoritmo de aprendizagem, porém não obtém a dinâmica de convergência não-linear das redes neurais ordinárias.

Em outra categoria, é realizada uma análise das propostas de Memória Associativa Quântica (QAM), algoritmos de computação quântica que simulam a propriedade da memória associativa sem a intenção de usar características de redes neurais. Após a inicialização da QAM com um padrão de entrada, um circuito quântico seleciona o padrão de memória mais próximo em termos de distância de Hamming. Percebe-se por definição que as QAM não compõem modelos de redes neurais quânticas segundo a classificação proposta, uma vez que não cumprem o segundo requisito, considerando que não se baseiam em modelos de redes neurais.

A maioria das contribuições nesta categoria baseiam-se na proposta de Ventura e Martinez



em (VENTURA; MARTINEZ, 2000). A idéia básica é executar um algoritmo quântico em uma superposição de todos os estados memorizados  $|M\rangle$  que após a medição final recupera o estado de saída desejado com uma alta probabilidade. Seja  $X^P = x_1^{(1)}, \dots, x_N^{(1)}, \dots, x_1^{(P)}, \dots, x_N^{(P)}$ ,  $P$  padrões a serem armazenados na QAM, a superposição de memória lê

$$|M\rangle = \frac{1}{\sqrt{P}} \sum_{p=1}^P |x_1^{(p)}, \dots, x_m^{(p)}\rangle.$$

Um algoritmo para criar  $|M\rangle$  proposto em (BOYER et al., 1996) usa o algoritmo de busca de Grover para criar a superposição de memória. Como o resultado é probabilístico e o estado destruído após a medição, é necessário construir um número de duplicatas de  $|M\rangle$ . A recuperação de um estado memorizado, proposto inicialmente em (VENTURA; MARTINEZ, 2000), é baseada no algoritmo de Grover. É realizada uma modificação no algoritmo para trabalhar no estado inicial  $|M\rangle$ , com isso, o algoritmo se torna capaz de recuperar todos os padrões que contêm uma determinada sequência de entrada.

Em outra categoria são analisadas duas propostas que abordam o problema de descrever os perceptrons com um formalismo quântico. Em (ALTAISKY, 2001) o perceptron é modelado pela função de atualização quântica

$$|y(t)\rangle = \hat{F} \sum_{i=1}^m \hat{w}_{iy}(t|x_i\rangle),$$

onde  $\hat{F}$  é uma porta quântica arbitrária e os operadores  $\hat{w}_{ij}$  representam os pesos sinápticos sobre os qbits de entrada  $m$ . O perceptron pode ser treinado pelo equivalente quântico de uma regra de aprendizagem:  $w_{jy}(t+1) = w_{jy}(t) + \eta(|d\rangle - |y(t)\rangle) \langle x_j|$ , onde  $|d\rangle$  é o estado desejado,  $|y(t)\rangle$  é o estado do neurônio  $y$  no passo de tempo discreto  $t$  e  $\eta \in [0, 1]$  é a taxa de aprendizagem. Altaisky observa que esta regra de aprendizagem não é unitária em relação às entradas da matriz de peso  $w$ , de modo que a atualização não conseguiria preservar a probabilidade total do sistema. O Quadro 4.1 sumariza as avaliações realizadas, indicando o critério de reprovação dos modelos tal qual proposto.

**Quadro 4.1:** Resumo da avaliação das propostas

Proposta	Reprovado no requisito
(ALTAISKY, 2001)	3
(BEHRMAN et al., 2000)	2
(FABER; GIRALDI, 2002)	3
(GUPTA; ZIA, 2001)	2
(KAK, 1995)	1
(MENNEER; NARAYANAN, 1995)	1
(PANELLA; MARTINELLI, 2011)	3
(PERUŠ, 2000)	1
(SILVA; OLIVEIRA; LUDERMIR, 2016)	3
(VENTURA; MARTINEZ, 2000)	3
(ZAK; WILLIAMS, 1998)	1

## 4.4 Treinando um Rede Neural Quântica

Em (RICKS; VENTURA, 2003) é abordada a questão do treinamento de uma rede neural quântica. Para tanto, é proposta uma rede neural quântica simples e um método de treinamento para ela. A rede neural quântica proposta possui peso e usa algoritmos e portas quânticas. Além disso, funciona em simulações clássicas de tamanho razoável, sendo capaz de transferir informações para sistemas clássicos.

A proposta se assemelha a uma rede neural clássica composta de várias camadas de perceptrons, tendo uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída, onde cada camada está totalmente conectada à camada anterior. Cada camada intermediária calcula a soma de cada saída pelo peso das saídas da camada anterior e, se esta soma é maior que um limiar, o valor do nó é setado para 1, caso contrário, permanece 0. Por sua vez, camada de saída também realiza esses cálculos, porém também verifica sua precisão em relação à saída desejada da rede. A rede como um todo verifica qual bit de saída está setado em 1, mas não há nenhuma verificação para certificar-se que exatamente uma saída está em 1.

Como exemplo dessa proposta é apresentada uma rede capaz de computar a função XOR, composta por duas camadas escondidas e dois registradores de entrada. Os nós de entrada  $i$  são representados por registradores  $|\alpha\rangle_i$  e os nós escondidos calculam a soma das entradas pelos pesos,  $|\psi\rangle_{i1}$  e  $|\psi\rangle_{i2}$ , e comparam a soma com um peso limiar  $|\psi\rangle_{i0}$ , mudando a saída do nó para 1. Tem-se  $|\beta\rangle_k$  representando esses cálculos internos que ocorrem em cada nó. A camada de saída funciona da mesma forma, calculando a soma das saídas pelos pesos dos nós intermediários e verificando o limiar. Por fim, a neural rede quântica verifica cada saída calculada, comparando-a com a saída de desejada,  $|\Omega\rangle_j$ , setando  $|\phi\rangle_j$  para 1 quando são equivalentes. O

desempenho da rede é denotado por  $|\rho\rangle$ , que é o número de saídas calculadas equivalentes à sua saída alvo correspondente.

Ainda considerando o exemplo proposto, os  $n$  exemplos de treinamento terão dois registradores de entrada,  $|\alpha\rangle_{11}$  até  $|\alpha\rangle_{n2}$ , considerando que a rede possui dois parâmetros de entrada. As saídas desejadas são mantidas nos registradores  $|\Omega\rangle_{11}$  até  $|\Omega\rangle_{n2}$ . Os vetores de pesos de cada camada intermediária ou camada de saída são representados por  $|\psi\rangle_i$ , onde cada vetor contém pesos para cada um de seus inputs. Após a classificação de um exemplo de treinamento, os registradores  $|\phi\rangle_1$  e  $|\phi\rangle_2$  refletem a capacidade da rede neural de classificar o exemplo. Incrementa-se  $|\rho\rangle$  pela de todos os  $|\phi\rangle_i$ , de forma que, quando todos os exemplos de treinamento forem classificados,  $|\rho\rangle$  será a soma dos nós de saída que têm a resposta correta ao longo do conjunto de treinamento, variando entre zero e o número de exemplos de treinamento pelo número de nós de saída.

A proposta apresentada de treinamento para essa rede foi pesquisar através dos possíveis vetores de peso,  $|\psi\rangle_i$ , até encontrar um que seja consistente com os dados de treinamento. A ideia é encontrar uma solução que classifique todos os exemplos de treinamento corretamente, dessa forma, deseja-se achar  $|\rho\rangle = n \times m$ , onde  $n$  é o número de exemplos de treinamento e  $m$  é o número de nós de saída. Como não se sabe quantos vetores de peso serão necessários, utilizou-se uma generalização do algoritmo de Grover (BOYER et al., 1996), destinado a problemas onde o número de soluções é desconhecido. Assim, coloca-se os  $|\psi\rangle$  em uma superposição de todos os vetores de peso possíveis, buscando por um que classifica todos os exemplos de treinamento corretamente.

O primeiro passo é criar  $|\psi\rangle$  como uma superposição de todos os vetores de peso possíveis. Os demais registradores,  $|\beta\rangle, |\phi\rangle e |\rho\rangle$ , além das entradas e saídas de destino, são inicializados no estado  $|0\rangle$ . Após isso, são classificados os exemplos de treinamento, atualizando o registro de desempenho,  $|\rho\rangle$ . Através da superposição, simultaneamente são classificados os exemplos de treinamento em relação a cada possível vetor de peso. Com isso, cada vetor de pesos estará emaranhado com  $|\rho\rangle$  de tal forma que ele corresponde a quão bem cada vetor de peso classifica todos os dados de treinamento.

Todos os  $|\psi\rangle$  são desemaranhados dos outros registradores após inverter a fase dos pesos que correspondem aos critérios de busca, com base em  $|\rho\rangle$ . Para fazer isso a rede inteira precisará ser descomputada, o que desemaranhará todos os registradores e configurá-los de volta para seus valores iniciais. Isso significa que a rede precisará ser recomputada toda vez que for feita uma chamada ao operador oráculo e após cada medição.

Nem todos os dados de treinamento terão uma rede que classifique corretamente todas as instâncias de treinamento. Nesses casos, nada será marcado pelo oráculo de busca, de modo que cada um dos vetores de peso terá uma chance igual de ser medido. Também é possível que, mesmo quando existe uma solução, ela não é desejável por se ajustar demais aos dados de treinamento. A quantidade de tempo necessário para encontrar um vetor que classifica corretamente o conjunto de treinamento é  $\mathcal{O}(\sqrt{2^b/t})$ , que tem complexidade exponencial em

relação ao número de bits no vetor de peso.

Por ser um algoritmo exponencial no número total de pesos na rede e bits por peso, é proposto um algoritmo de treinamento randomizado que pesquisa independentemente cada vetor de peso de cada nó. Nele, a rede também é inicializada com exemplos de treinamento em  $|\alpha\rangle$ , as saídas desejadas em  $|\Omega\rangle$  e com zeros nos demais registradores. Um nó é selecionado aleatoriamente e seu vetor de peso,  $|\psi\rangle_i$ , é posto em superposição. Por sua vez, os demais vetores de peso começam com pesos iniciais clássicos aleatórios. Em seguida, busca-se um vetor de peso para este nó que faz com que toda a rede classifique certa porcentagem,  $p$ , dos exemplos de treinamento corretamente. Repete-se este passo diminuindo iterativamente  $p$ , até que um novo vetor de peso seja encontrado. Então, este peso é fixo classicamente e o processo é repetido aleatoriamente para os outros nós.

Procurar o vetor de pesos de cada nó separadamente é uma busca aleatória através do espaço de peso, onde são selecionados vetores de peso que dão um bom nível de desempenho para cada nó. Cada nó assume vetores de peso que tendem a aumentar o desempenho em uma quantidade de aleatoriedade que ajuda a mantê-lo fora dos mínimos locais. Esta pesquisa pode ser terminada quando um nível aceitável de desempenho foi alcançado.

São propostas também algumas melhorias no design básico que ajudam a acelerar a convergência. Primeiro, para garantir que os nós escondidos encontrem vetores de peso que calculam algo útil, uma pequena penalidade de desempenho é adicionada aos vetores de peso que fazem com que um nó oculto produza o mesmo valor para todos os exemplos de treinamento. Isso ajuda a selecionar vetores de peso que contêm informações úteis para os nós de saída.

# 5

## Conclusão

O presente trabalho buscou produzir uma revisão compreensiva sobre o campo das redes neurais quânticas. Para tanto, foram expostos aqui modelos de redes neurais quânticas, fornecendo também material de fundamentação teórica suficiente para sua compreensão. As propostas foram descritas de forma imparcial, sendo as propriedades dos modelos detalhadas.

Por não ser objetivo do trabalho, não foram apresentadas considerações sobre a implementação física dos modelos estudados, assim como não foram apresentadas propostas voltadas para a implementação física de redes neurais quânticas. Da mesma forma, não são contemplados modelos de redes neurais que possuem inspiração na computação quântica, mas que não se utilizam dos princípios da teoria quântica.

As propostas de redes neurais quânticas são muito diversas, o que pode ser notado em (SCHULD; SINAYSKIY; PETRUCCIONE, 2014) que, apesar de propor uma divisão de propostas em categorias segundo suas características comuns, apresenta propostas muito distintas na mesma categoria. Têm-se, por exemplo, as propostas de (PERUŠ, 2000) e (MENNEER; NARAYANAN, 1995) na mesma categoria, onde as correspondências de ambas para os conceitos de redes neurais diferentes diferem.

Buscou-se aqui, focar modelos com propostas voltadas a perspectiva da computação quântica. É possível notar que, apesar disso, mesmo fazendo uso da teoria quântica, principalmente utilizando-se da superposição, esses modelos ainda assim fazem uso de operadores não-lineares. Quanto as topologias das redes neurais quânticas propostas, podem ser vistos também uma diversidade de modelos como Perceptron, MLP e redes de Hopfield.

Por fim, com este trabalho pretendeu-se constituir uma base para um estudo mais aprofundado sobre as redes neurais quânticas. Assim, espera-se que a forma com que os conceitos e modelos foram apresentados, possibilite boa compreensão, servindo de base para consulta ou para inspiração na proposição de novos modelos.

## 5.1 Trabalhos Futuros

Um ponto a ser futuramente trabalhado versa sobre um estudo de classificação de propostas de redes neurais quânticas sem peso. Para tanto, podem vir a ser levantados requisitos específicos de classificação para as redes sem peso, tal qual realizado em (SCHULD; SINAYSKIY; PETRUCCIONE, 2014) para redes neurais do tipo Hopfield com propriedade de memória associativa. Também, os resultados obtidos podem ser comparados entre os dois tipos de rede no sentido de verificar uma possível direção a ser seguida no que tange as propostas de redes neurais quânticas.

# Referências

- ALTAISKY, M. Quantum neural network. **arXiv preprint quant-ph/0107012**, [S.l.], 2001.
- BEHRMAN, E. C. et al. Simulations of quantum neural networks. **Information Sciences**, [S.l.], v.128, n.3, p.257–269, 2000.
- BITTENCOURT, G. Inteligência Artificial Ferramentas e Teorias, Editora da UFSC. **Campus Universitário, Trindade-Florianópolis, 2a. edição**, [S.l.], 2001.
- BOYER, M. et al. Tight bounds on quantum searching. **arXiv preprint quant-ph/9605034**, [S.l.], 1996.
- DEUTSCH, D. Quantum theory, the Church-Turing principle and the universal quantum computer. In: ROYAL SOCIETY OF LONDON A: MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCES. **Proceedings...** [S.l.: s.n.], 1985. v.400, n.1818, p.97–117.
- DEUTSCH, D. Quantum computational networks. In: ROYAL SOCIETY OF LONDON A: MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCES. **Proceedings...** [S.l.: s.n.], 1989. v.425, n.1868, p.73–90.
- EZHOV, A. A.; VENTURA, D. Quantum neural networks. **Future directions for intelligent systems and information sciences**, [S.l.], v.45, p.213–235, 2000.
- FABER, J.; GIRALDI, G. A. Quantum models for artificial neural networks. **Electronically available: <http://arquivosweb.incc.br/pdfs/QNN-Review.pdf>**, [S.l.], v.5, n.7.2, p.5–7, 2002.
- FEYNMAN, R. P. Simulating physics with computers. **International journal of theoretical physics**, [S.l.], v.21, n.6, p.467–488, 1982.
- FIESLER, E.; BEALE, R. **Handbook of neural computation**. [S.l.]: Oxford University Press, 1996.
- GROVER, L. K. A fast quantum mechanical algorithm for database search. In: ACM SYMPOSIUM ON THEORY OF COMPUTING. **Proceedings...** [S.l.: s.n.], 1996. p.212–219.
- GUPTA, S.; ZIA, R. Quantum neural networks. **Journal of Computer and System Sciences**, [S.l.], v.63, n.3, p.355–383, 2001.
- GUYON, I. Neural networks and applications tutorial. **Physics Reports**, [S.l.], v.207, n.3-5, p.215–259, 1991.
- HAYKIN, S.; NETWORK, N. A comprehensive foundation. **Neural Networks**, [S.l.], v.2, n.2004, p.41, 2004.
- HOFFMAN, K.; KUNZE, R. **Linear algebra, 2nd**. [S.l.]: Prentice Hall of India, New Delhi, 1990.
- KAK, S. C. Quantum neural computing. **Advances in Imaging and Electron Physics**, [S.l.], v.94, p.259–313, 1995.

- KRENKER, A.; KOS, A.; BEŠTER, J. **Introduction to the artificial neural networks**. [S.l.]: INTECH Open Access Publisher, 2011.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, [S.l.], v.5, n.4, p.115–133, 1943.
- MENNEER, T.; NARAYANAN, A. Quantum-inspired neural networks. **Department of Computer Science, University of Exeter, Exeter, United Kingdom, Technical Report**, [S.l.], v.329, p.1995, 1995.
- NIELSEN, M. A.; CHUANG, I. L. **Quantum computation and quantum information**. [S.l.]: Cambridge university press, 2010.
- PANELLA, M.; MARTINELLI, G. Neural networks with quantum architecture and quantum learning. **International Journal of Circuit Theory and Applications**, [S.l.], v.39, n.1, p.61–77, 2011.
- PERUŠ, M. Neural networks as a basis for quantum associative networks. **Neural Netw. World**, [S.l.], v.10, n.6, p.1001–1013, 2000.
- RICKS, B.; VENTURA, D. Training a quantum neural network. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS. Anais...** [S.l.: s.n.], 2003. p.None.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, [S.l.], v.65, n.6, p.386, 1958.
- SCHULD, M.; SINAYSKIY, I.; PETRUCCIONE, F. The quest for a quantum neural network. **Quantum Information Processing**, [S.l.], v.13, n.11, p.2567–2586, 2014.
- SHOR, P. W. Algorithms for quantum computation: discrete logarithms and factoring. In: **FOUNDATIONS OF COMPUTER SCIENCE, 1994 PROCEEDINGS., 35TH ANNUAL SYMPOSIUM ON. Anais...** [S.l.: s.n.], 1994. p.124–134.
- 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**, [S.l.], v.76, p.55–64, 2016.
- SILVA, A. J. da; OLIVEIRA, W. R. de; LUDERMIR, T. B. Weightless neural network parameters and architecture selection in a quantum computer. **Neurocomputing**, [S.l.], v.183, p.13–22, 2016.
- TAGLIARINI, G. A.; CHRIST, J. F.; PAGE, E. W. Optimization using neural networks. **IEEE transactions on computers**, [S.l.], v.40, n.12, p.1347–1358, 1991.
- VENTURA, D.; MARTINEZ, T. Quantum associative memory. **Information Sciences**, [S.l.], v.124, n.1, p.273–296, 2000.
- ZAK, M.; WILLIAMS, C. P. Quantum neural nets. **International journal of theoretical physics**, [S.l.], v.37, n.2, p.651–684, 1998.